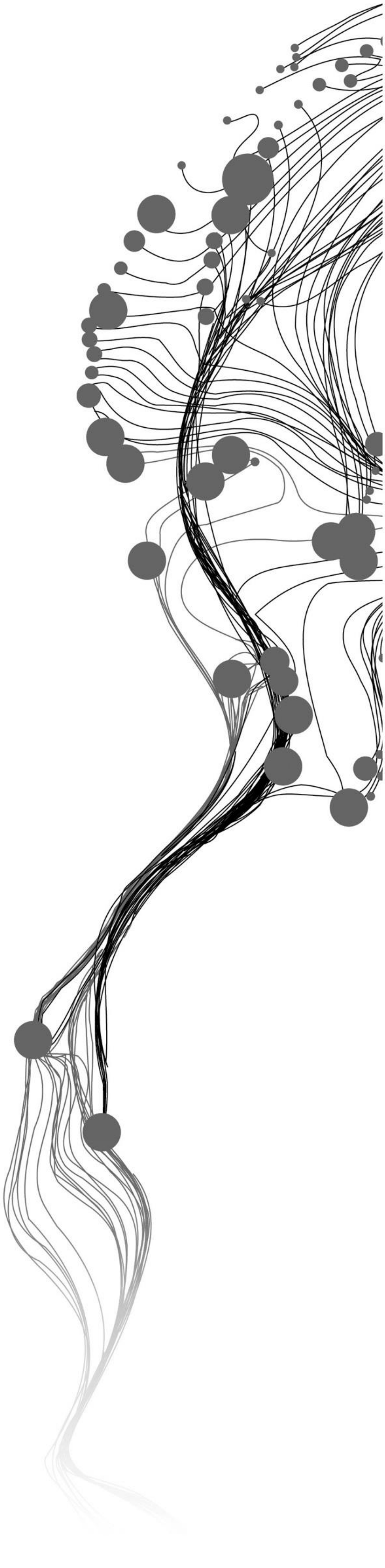
Application of Hyper-Temporal NDVI Data to Improve Homogeneity of the currently used Agricultural Area Frames

MAURICE MUGABOWINDEKWE
February, 2016

SUPERVISORS:
Drs. Eduard Westinga
Dr. Ir. C.A.J.M. (Kees) de Bie



Application of Hyper-Temporal NDVI Data to Improve Homogeneity of the currently used Agricultural Area Frames

MAURICE MUGABOWINDEKWE
Enschede, The Netherlands, February, 2016

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.
Specialization: Geo-information Science and Earth Observation for Natural Resources Management

SUPERVISORS:
Drs. Eduard Westinga (First supervisor)
Dr. Ir. C.A.J.M. (Kees) de Bie (Second supervisor)

THESIS ASSESSMENT BOARD:
Dr. Y.A. Hussin (Chair)
Dr. Ir. B.G.J.S. (Ben) Sonneveld (External Examiner, VU Amsterdam)
Drs. Eduard Westinga (First supervisor)
Dr. Ir. C.A.J.M. (Kees) de Bie (Second supervisor)

ABSTRACT

World's population continuously increases, resulting in an increase of food basic needs from limited land resources. This leads the countries to strive for better policies and plans for food security. To reach the goal, agricultural information/statistics have served indispensably. To collect the data, various methods including census, registers and administrative data, and frame sampling (area frame sampling, list frame sampling or multiple frame sampling) are applied. Because of saving time, labor and money, area frame sampling has been the most popular method. Most importantly, the method is claimed to give accurate statistics, which is said to result from consideration of areas' homogeneity while sampling. Nevertheless, looking at area frames construction process and sampling methodology, the homogeneity claimed by the method might be dubious as the frames are delineated using visual interpretation at the initial stage, and sometimes using outdated data. In Rwanda, one of the countries where the area frame sampling method is applied to stratify land for seasonal agricultural surveys, homogeneous area frames have been designed for the year 2012, 2013, 2014 and 2015 using aerial photographs from 2008. This indicates that the spatial-temporal heterogeneity existing in nature might not have been effectively considered. It is in this regard that the present study used a technique which considers the spatial-temporal heterogeneity in order to stratify the land.

The present study applied MODIS hyper-temporal NDVI data to stratify land in Rwanda. NDVI, which captures a behaviour of an area based on photosynthetic activity over the area, was considered appropriate by the present study for agricultural land stratification. To detect areas' variabilities over time, a temporal window of 10 years from 2004 to 2014 was taken. ISODATA clustering technique was applied for the classification, and 95 best separable classes were identified for the country. Through intersection with the recent (2010) Rwanda land use data, 24 NDVI classes dominated by agriculture were identified, from which 4 sample classes were selected given remarkable differences in temporal behaviour. From the 4 sample NDVI classes, 48 sample sites were selected using random-representative clustered sampling technique. The land covers present in the sample sites and their areas were collected from the field.

To investigate significant differences between and within NDVI classes and Rwanda strata, ANOVA and Fisher's LSD statistical methods were applied. The analysis revealed that NDVI classes were significantly different in terms of Rwanda season A main crops: banana, maize, beans and cassava, at 95% confidence level. Further, the analysis revealed that there was 15% of significant differences within the NDVI classes, and 85% of no significant differences.

On the side of Rwanda strata, statistical analysis was possible to be performed only within Rwanda stratum 1 "intensive cropland and houses", because of insufficient data to represent other strata. The analysis showed that there was 31% of significant differences between clusters within Rwanda stratum 1, and other 69% had no significant differences in terms of the Rwanda season A main crops' area at 95% confidence level. Through the use of NDVI classes, Rwanda stratum 1 was corrected according to the ground truth. The correction resulted in decreasing the significant differences from 31% to 26% within the stratum. This showed that, not only can the NDVI classes be used as independent strata for agricultural surveys to maximize the areas' homogeneity, but they can also be used to improve the homogeneity within the currently used agricultural area frames in Rwanda.

Keywords: Hyper-temporal NDVI data, agricultural statistics, spatial-temporal heterogeneity, homogeneity, area frame sampling

ACKNOWLEDGEMENTS

First and foremost, immeasurable and deepest gratitude goes to The Almighty God for His grace, protection and guidance in life, most importantly through this 18 months journey of MSc studies at ITC, University of Twente, The Netherlands.

I would like to acknowledge with profound gratitude my mom, "Without you, I couldn't have reached whatever I can have today". I can't forget your most precious inheritance: "through obedience and genuine fear of God, you will have all the wealth of the earth and heaven". Thank you very much my lovely and caring mum, I will always keep this in mind.

I am hugely indebted to the Dutch government through the Netherlands Fellowship Programme, and the Centre of GIS and Remote Sensing of the University of Rwanda, most especially the director Dr. Gaspard Rwanyiziri, for granting me the opportunity to pursue my MSc studies at ITC, University of Twente.

My sincere gratitude goes to my first supervisor Drs. Eduard Westinga, for his prompt inspiration, timely suggestions with kindness and patience, which enabled me to keep the focus on the goal of the study, and motivated me throughout the whole process of the thesis.

I profusely owe a deep sense of gratitude to my second supervisor Dr. Ir. C.A.J.M. de Bie (Kees) for his conscientious guidance, kind advice and immense support, most importantly while data processing, field work preparation and throughout the course of the whole thesis.

I would like to thank NRS department at ITC, through the course coordinator Drs. R.G. Nijmeijer (Raymond) who tirelessly helped us to get whatever we needed, including priceless advice in time of need, to comfortably carry on the studies.

I am grateful to the CGIS-UR staff, most particularly Mr. Alban Uwacu Singirankabo and Mr. Ernest Uwayezu, who helped me very much in getting the data I was not able to get for the success of the study. Further, I am very thankful to Mr. Elias Nyandwi, for his kind advice that comforted and motivated me throughout my MSc studies at ITC.

I gratefully would like to thank Mr. Alfred Gasore, and Mr. Desire Uwitonze, for the support they did not cease to provide, whenever I was asking for the data I could not get because I was not in Rwanda.

I would appreciatively like to thank the government of Rwanda through the National Institute of Statistics, and staff of the districts: Nyabihu, Nyagatare, Bugesera, Nyamagabe, Nyaruguru, Musanze and Burera district, for their support and collaboration in data collection and guidance on field work.

My expression of deep and sincere gratitude extends to my family members, brothers and sisters, who have supported me morally, economically, and most importantly through prayers. Because of you, I stayed motivated in the course of my studies.

My special thanks extend to Mr. Prosper Ngabonziza and his wife Claudine. I appreciated very much all the support you provided for me while my MSc studies, especially through spiritual development by prayers and faith strengthening discussions, together with the entire bible study group.

I am extremely thankful for the fellow classmates in NRM/GEM class of 2014-2016 at ITC, for the very much appreciated friendship, moral support, encouragements and companionship throughout the studies, especially Aline, Steve and Edossa.

Finally, I would sincerely and deeply wish to acknowledge and express my gratitude to everyone who made a contribution in my life.

In the absence of you, today could not be my day!

May God plentifully bless you all!

Maurice Mugabowindekwe

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATION

AEZ:	Agro-Ecological Zone
AFS:	Area Frame Sampling
ANOVA:	Analysis of Variance
AVHRR:	Advanced Very High Resolution Radiometer
CLGF:	Commonwealth Local Government Forum
DN:	Digital Number
EACSOF:	East African Civil Society Forum
FAO:	Food and Agricultural Organization of the United Nations
Fcrit:	F Critical
GPS:	Global Positioning System
IDB:	Index Database
IPAQ:	International Physical Activity Questionnaire
ISODATA:	Iterative Self-Organizing Data Analysis Technique
LSD:	Least Significant Differences
MBA:	Master of Business Administration
MERIS:	Medium Resolution Imaging Spectrometer
MINAGRI:	Ministry of Agriculture and Animal Resources
MINECOFIN:	Ministry of Finance and Economic Planning
MININFRA:	Ministry of Infrastructure
MINIRENA:	Ministry of Natural Resources
MOD13Q1:	Product name for "MODIS Vegetation Indices 16-day L3 Global 250 m"
MODIS:	Moderate-Resolution Imaging Spectroradiometer
MRTWeb:	Modis Reprojection Tool Web Interface
NAS:	National Agricultural Survey
NASA:	National Aeronautics and Space Administration
NASS:	National Agricultural Statistics Service
NDVI:	Normalized Difference Vegetation Index
NIR:	Near-Infrared band
NISR:	National Institute of Statistics of Rwanda
NOAA:	National Oceanic and Atmospheric Administration
PPS:	Probability Proportional to Size
PSU:	Primary Sampling Unit
SAS:	Seasonal Agricultural Survey
SSU:	Secondary Sampling Unit
SUPARCO:	Space and Upper Atmosphere Research Commission
TIFF:	Tagged Image File Format
USDA:	United States Department of Agriculture

1. INTRODUCTION

1.1. Background and justification

The population growth continuously increases worldwide (Worldometers, 2015). Hence, the countries need to produce more food from limited resources to ensure that "all people at all times have physical, social and economic access to sufficient, safe and nutritious food which meets their dietary needs and food preferences for an active and healthy life" (FAO, 1996b). To achieve the responsibility, the countries crucially need agricultural information (or statistics) (Boyko & Hill, 2009). Agricultural information includes all information that serves as a basis to contribute or intervene in agriculture development and food security improvement (FAO, 2015b; Vidanapathirana, 2012). This introduction chapter gives a wide introduction on agricultural information and/or statistics, their relevance to countries' food security, different methods currently used for their collection and generation with emphasis on area frame sampling method, impact of nature's heterogeneity on the currently used area frames and on quality of the generated statistics, and how to decrease the nature's heterogeneity in the currently area frames by stratifying land using Normalized Difference Vegetation Index (NDVI) hyper-temporal data as studied and applied by the present study.

- **Agricultural information/statistics**

Agricultural information is a wide terminology incorporating all information contributing to agricultural development and food security in general; including social, environmental and economic aspects in relation with agriculture (SUPARCO, 2015). Talking of agricultural statistics, they refer to agricultural information provided at administrative level with statistical progressions, and they normally include agricultural land area, land productivity, labor and capital in agriculture, available farming technology, farming practices, and demographic and social characteristics of the people involved in agriculture (FAO, 2015b; Vidanapathirana, 2012).

Agricultural statistics are very important for a country in general, and for individuals in particular. On one hand, they are required mainly to underpin the planning processes of a country; provide information for public policy analysis, debate and advice; observe the agricultural sector performance in the country; monitor and evaluate the impact of policies and programmes; and enlighten the decision-making processes (Kiregyera, Megill, José, & Eding, 2007). Additionally, the information help in monitoring, evaluation, improvement and strengthening of early warning systems in regard of country's food security (FAO, 2016). On the other hand, agricultural statistics serve ideas development and decision making of local people, private sectors, and importantly the farmers about their agricultural businesses. Maningas, Villagonzalo, and Macaraig (2004) mentioned that agricultural information empowers farmers through control over their resources and decision-making processes. Nevertheless, the agricultural statistics to be useful with good quality, they have to be accurate: conforming exactly or almost exactly to the ground truth; reliable: showing all details about the reality they represent and errors they contain; and timely: available on time of need (Cotter, Davies, Nealon, & Roberts, 2010; Wigton & Bormann, 1978). To generate agricultural data with good quality, various methods have been developed and applied in different countries worldwide.

- **Current methods for agricultural statistics generation**

Methods including census, registers and administrative data, and sampling frames (area frame sampling, list frame sampling and multiple frame sampling) (Benedetti, Bee, Espa, & Piersimoni, 2010; Eurostat, 2015; Kennel, 2008; Väisänen, 2009) have been in use for collection of the agricultural data by various organizations worldwide. However, these methods are different and apply different principles. First, agricultural census collects information on all agricultural holdings of a country (Eurostat, 2015). But, due, mainly, to their time consumption, the censuses normally take place every five to ten years in different countries. Taking Rwanda as an example, the census takes place exactly every 10 years (MINECOFIN & NISR, 2014). Much time and labor consumption due to the very large population to survey, are also some of the main factors making the census method not very frequently applied. Nevertheless, the method is credited to result in highly accurate and most reliable outputs (Farooq, 2013).

Further, with registers and administrative data method; registers available in a country, such as population registers (births, deaths, marriages,…), business registers, registers of completed education and degrees, tax registers, and farm registers are used for extraction of agricultural statistics (Väisänen, 2009). The registers are used and considered in accordance with their administrative boundaries, as they are primarily to serve administration purposes. However, they are rarely used. For the example of Rwanda, there is no published agricultural survey where the method was applied, though it is claimed to have been used in some agricultural surveys (NISR, 2012b). This might be due to complexity in the method implementation; through the integration of different registers from different institutions with different primary goals as explained by Väisänen (2009). In addition, the statistics from registers and administrative data are said to be incomplete, inaccurate, with loss of information due to a broad level of aggregation, and there is inaccessibility of some data sources (The Republic of Pakistan, 2008).

Finally, with sampling frame methods, frames: sets of all agricultural holdings in an area of interest based on their homogeneity, are constructed using visual interpretation method (Wigton & Bormann, 1978). The authors further explain that, next, the segments: homogeneous land units enumerable in one day, are drawn from the frames. Then, representative sample segments are randomly selected from all the segments (Kennel, 2008). The frame can be an area (area frame) or a list (list frame) or combined (multiple frames) (FAO, 1996a). In contrast to the previous methods, the sampling frame methods are popular in agricultural surveys because they are cheaper, save time, reliable, and most importantly suitable in case of large population (MBA Official, 2015; Menza, Caldow, Jeffrey, & Monaco, 2008). To exemplify, in Rwanda, multiple frame sampling method is applied to design area frames for seasonal agricultural surveys (three times a year) (NISR, 2013, 2015a, 2015b, 2015c).

Moreover, renowned organizations such as Food and Agriculture Organization of the United Nations (FAO), National Agricultural Statistics Service (NASS) of the United States Department of Agriculture (USDA) (Cotter et al., 2010), and many countries worldwide adopted the area frame sampling method as a basic and most suitable method for collecting and generating agricultural statistics for quick and comprehensive agricultural information systems (Wigton & Bormann, 1978). With the method, agricultural statistics are collected from sample segments and aggregated to their respective strata, then to the entire area of interest (Cotter et al., 2010). Wigton and Bormann (1978) indicated that the land stratification for this method does not follow country's administrative units, but homogeneity in nature in order to produce representative and reliable agricultural estimates. Nevertheless, though the area frame sampling method is claimed to be accurate in agricultural surveys with the ability to select representative sample segments, it may be an ideal situation different from reality, given the degree of heterogeneity in nature (Turner & Gardner, 1994).

- **Homogeneity and heterogeneity in nature**

Heterogeneity is a characteristic of a landscape (Kiss, Tokody, Deák, & Moskát, 2016). It determines, characterizes patterns, and benefits the landscape's biodiversity (Hiron et al., 2015; Kiss et al., 2016). There exist both spatial and temporal heterogeneity (Turner & Gardner, 1994) and inter-annual heterogeneity (de Bie, Nguyen, Ali, Scarrott, & Skidmore, 2012). This heterogeneity is due to natural processes in a landscape on one hand, and anthropogenic activities on another hand. For instance, in Rwanda, the inter-annual variability is mainly due to having three different agricultural seasons per year (NISR, 2015b) with different characteristics of each season. In addition, other differences in the country are due to natural factors including topography, hydrography and geomorphology (Twagiramungu, 2006). Because of these differences, the country counts 12 different agro-ecological zones (Clay & Dejaegher, 1987).

The degree of heterogeneity makes it complex to find homogeneous areas in nature unless some level of aggregation is adopted, and this becomes even more complex in agricultural areas which change very fast especially in developing countries (Roser, 2015). So, in order to capture the heterogeneity in landscape and group relatively homogeneous areas, methods with regular spatial data and available for a long time can be effective. Contrastingly, in Rwanda, since 2012 till recently in 2015, seasonal agricultural surveys have been designed and being conducted based on the use of single time aerial photographs, from 2008, to design homogeneous strata and sample units (segments) (NISR, 2013, 2015a, 2015b, 2015c). The rationale was that the aerial photographs were the most recent spatial data available with high spatial resolution (25 cm). Given the five to eight years of difference (from photographs acquisition date: 2008 to agricultural surveys dates, 2012, 2013, 2014, 2015), there might have been many changes in the country's landscape components, especially in agricultural land use. Subsequently, the generated statistics might be subject to errors in terms of accuracy, representativeness and completeness. Therefore, to determine better homogeneous strata for agricultural surveys, there is a need to apply a technique that detects changes in the landscape over time based on vegetation content in an area.

- **Hyper-temporal MODIS NDVI data**

Normalized Difference Vegetation Index (NDVI) (Rouse, 1974) is one of the radiometric measures of the amount of greenness on a land by determining the photosynthetic activity present on that land (NASA, 2015; Tucker & Sellers, 1986). Given that vegetation reflects very low in red satellite band and reflects very high in near-infrared satellite band, the NDVI combines and makes a ratio between these two bands in order to quantify the greenness of the smallest land area (image pixel size) as follows:

$$\boxed{NDVI = \frac{(NIR - RED)}{(NIR + RED)}}$$

(Gillan, 2013; Matsushita, Yang, Chen, Onda, & Qiu, 2007)

Equation 1: NDVI calculation

The NDVI data recorded with high and regular temporal frequency with long temporal sequence, also known as hyper-temporal NDVI data (Ali, de Bie, Scarrott, Ha, & Skidmore, 2012; de Bie, Khan, Toxopeus, Venus, & Skidmore, 2008) have been proven to be strong in studying dynamic aspects such as cropping systems and crops phenology (Ali et al., 2012). In addition, availability of the hyper-temporal NDVI data for a long time and for free, have been also the main strengths of the method. The hyper-temporal data have become a solution to many issues that were challenging to study before, such as agro-ecosystems mapping due to their high variability and rapid changes (de Bie et al., 2008). The authors point out that the temporal changes are more frequent than the spatial ones in agro-ecosystems.

NDVI data from MODIS have been used by various scientists in phenological studies and other crops related studies (Bolton & Friedl, 2013; Funk & Budde, 2009). The data for an area are recorded every two days, and the records are from the year 1999 till date (January 11, 2016) (NASA, 2002; University of Wisconsin-Madison, 2015). MODIS NDVI data allowed monitoring of the Earth's terrestrial photosynthetic vegetation activity, and hence, possibility to consistently compare spatial and temporal vegetation changes (Huete, Didan, & Van Leeuwen, 1999).

Furthermore, various studies used hyper-temporal NDVI technique to study various landscape and ecological phenomena (Ali et al., 2012; Funk & Budde, 2009; Hamad, 2010; Lunetta, Knight, Ediriwickrema, Lyon, & Worthy, 2006). Others used the approach for agricultural ecosystems studies (Benedetti & Rossini, 1993; Sisilana, 2008; Walker & Mallawaarachchi, 1998). Nevertheless, no study yet has used the hyper-temporal NDVI techniques in integration with currently used area frame sampling in order to improve the homogeneity of the designed area frames for agricultural surveys, which would result in better agricultural population representativeness, thus increasing the accuracy of the estimated agricultural statistics.

In this regard, the current study has investigated the possibility to improve homogeneity of the currently used area frames/strata for agricultural surveys, by integration of MODIS hyper-temporal NDVI data. In the present study, the MODIS hyper-temporal NDVI stratification resulted in better homogeneous strata (NDVI classes), and also succeeded in improving homogeneity of the currently used strata in Rwanda.

1.2. Research problem

Agricultural information is crucial for countries' development and decision-making about improvement and interventions for people's food security. So, they should be collected and estimated with care, in order to be of good quality and reliable, with effective representativeness of the entire agricultural aspects of interest. Nevertheless, the currently used and popular methods for the data collection: area frame sampling, seem not to effectively consider the main aspect contributing to good quality of aggregated agricultural statistics. This main aspect is the homogeneity and heterogeneity existing in agricultural ecosystems.

In fact, heterogeneity is a characteristic of a landscape and the main factor which determine the distribution of various important features of the landscape such as biodiversity patterns among others (Kiss et al., 2016; Tuanmu & Jetz, 2015). Turner and Gardner (1994) clarified that there exist both spatial and temporal heterogeneity in the landscape. In addition, there exist inter-annual heterogeneity which occurs mainly in agricultural areas (de Bie et al., 2012). Given the fact that agro-ecosystems are parts of the landscape with fast-changing land use (Roser, 2015), the temporal heterogeneity is more prevailing than spatial one in agricultural land. The degree of heterogeneity makes it complex to find homogeneous areas in nature, especially in agricultural areas with their rapid changes, unless a level of aggregation is adopted.

The area frame sampling method, the most popular in agricultural surveys, is accredited to base upon areas' homogeneity to define the area frames (Cotter et al., 2010; Wigton & Bormann, 1978). Nevertheless, the support technique applied by the method to define the homogeneous area frames "image visual interpretation" (Wigton & Bormann, 1978), depends on the professional experience of the interpreter, data available and provided guidelines (Baks, Janssen, Schetselaar, & Tolpekin, 2013). This is the same technique applied in Rwandan agricultural surveys using aerial photographs from 2008 (NISR, 2015a, 2015b, 2015c). Given the factors that the visual interpretation technique depends on, the homogeneity of the current area frames may be judged as not naturally unique and unstable. Hence, the determined homogeneous area frames/strata may not be characterized as most effective for agricultural surveys, especially in a landscape with as large degree of heterogeneity as Rwanda; a country different from an area to another with differences in topography, hydrography, geomorphology and many other natural factors (Twagiramungu, 2006) resulting in 12 different agro-ecological zones in the country (Clay & Dejaegher, 1987).

The current study aimed at applying a method which effectively considers both spatial and temporal heterogeneity in the landscape, to define better homogenous area frames and improve the homogeneity of the currently used ones in Rwanda agricultural surveys. The used method applied hyper-temporal NDVI data and ISODATA clustering technique to determine various NDVI land classes based on similarities in patches of greenness existing in an area over time (Memarsadeghi, Mount, Netanyahu, & Le Moigne, 2007; NASA, 2015).

The used NDVI data are from Moderate-Resolution Imaging Spectroradiometer (MODIS) (NASA, 2002), which acquires the NDVI images every two days and produces 16 days composites. Ten years (from 2004 to 2014) data were used to ensure enough temporal period for similar behaviours of various areas to be detected. So, the heterogeneity of the same area in every two days for ten years was encountered while distinguishing and producing different NDVI classes. In addition, the data were incorporated with recent Rwanda land use data (RNRA, 2010) in order to raise the richness in spatial dimension as well.

1.3. Research objectives

The aim of the present study was to integrate hyper-temporal NDVI data with the currently used area frame sampling method for agricultural surveys, to produce land stratification with better homogeneous strata (NDVI classes) with improved spatial representativeness of agricultural areas for agricultural surveys, compared to the current Rwanda strata.

In order to reach the aim, the following specific objectives were attained:

1. To identify statistical differences in crops area coverage between NDVI classes;

2. To identify statistical differences in crops area coverage from sample sites within the same NDVI class;

3. To identify statistical differences in crops area coverage from sample sites within Rwanda strata;

4. To analyse and compare statistical differences between crops area coverage in sample sites within NDVI classes and crops area coverage in sample sites within Rwanda strata.

1.4. Research questions

The following are research questions that were answered and led to successful accomplishment of the study's objectives:

1. Are there significant differences in crops area coverage data between NDVI classes?

2. What is the significance level of differences in crops area coverage data from sample sites within the same NDVI class?

3. What is the significance level of differences in crops area coverage data from sample sites within the same Rwanda stratum?

4. How is the significance level of differences in crops area coverage data within NDVI classes compared to the significance level of differences in crops area coverage within Rwanda strata?

1.5. Research Hypotheses

A. **H_0a**: There are significant differences in crops area coverage data between NDVI classes at 95% confidence level.

H_1a: There are no significant differences in crops area coverage data between NDVI classes at 95% confidence level.

B. **H_0b**: There are significant differences in crops area coverage data from sample sites within the same NDVI class (heterogeneous NDVI class) at 95% confidence level.

H_1b: There are no significant differences in crops area coverage data from sample sites within NDVI class (homogeneous NDVI class) at 95% confidence level.

C. **H_0c**: There are significant differences in crops area coverage data from sample sites within the same Rwanda stratum (heterogeneous Rwanda stratum) at 95% confidence level.

H_1c: There are no significant differences in crops area coverage data from sample sites within the same Rwanda stratum (homogeneous Rwanda stratum) at 95% confidence level.

D. **H_0d**: Significant differences in crops area coverage data, are as the same for sample sites within NDVI classes as for sample sites within Rwanda strata (equal homogeneity between NDVI classes and Rwanda strata) at 95% confidence level.

H_1d: Significant differences in crops area coverage data, are smaller for sample sites within NDVI classes than for sample sites within Rwanda strata (NDVI classes are more homogeneous than Rwanda strata) at 95% confidence level.

The following figure 1 is a conceptual diagram presented in style of systems and subsystems, illustrating the issue of agricultural statistics collection (sample site level) and generation (aggregation to national level), according to the currently used area frame sampling method, and how the situation may improve if the area frames are improved in their homogeneity as studied by the present study.

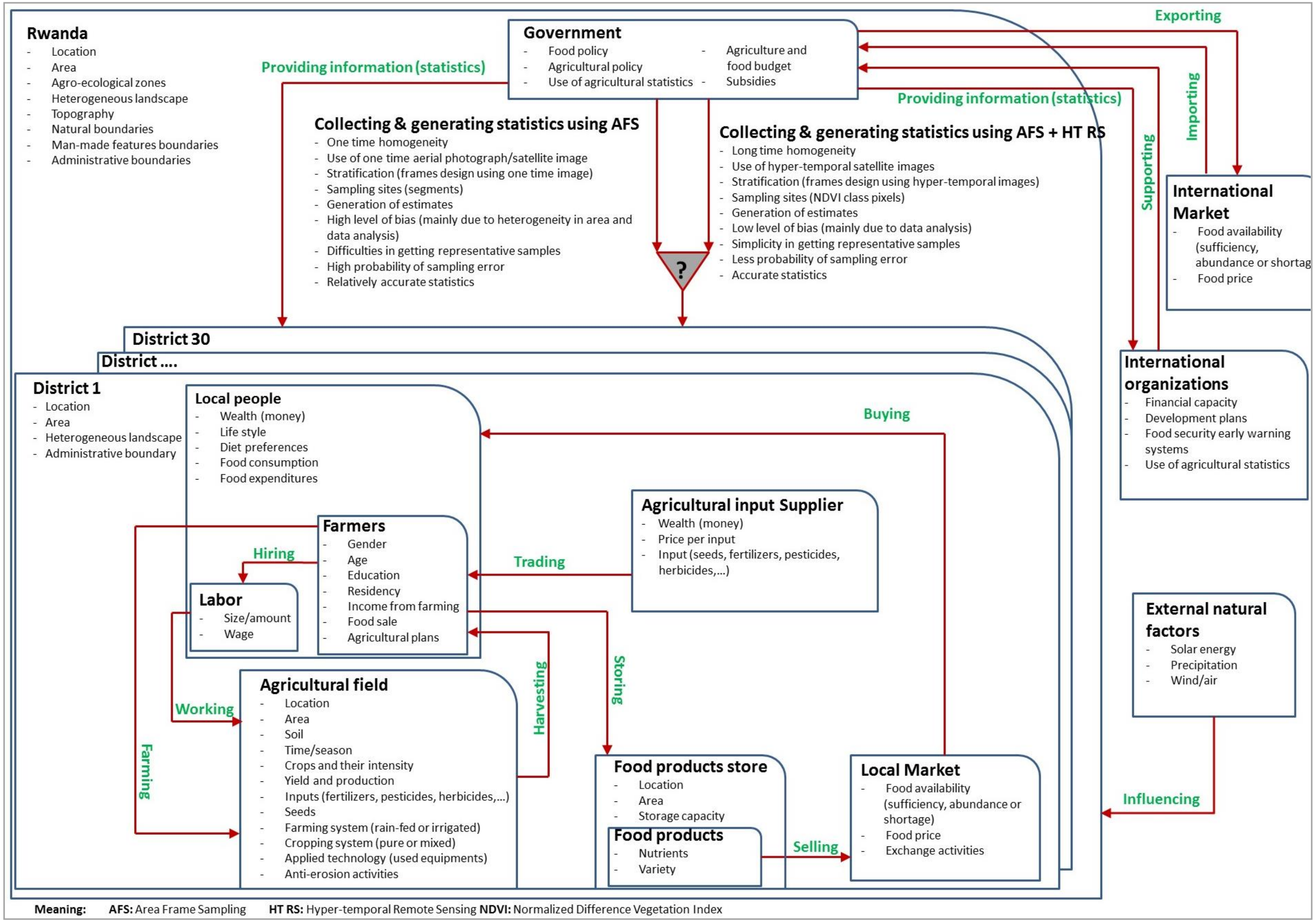


Figure 1: Conceptual diagram illustrating the flow of agricultural statistics in Rwanda

2. MATERIALS AND METHOD

2.1. Study area

The study area was not based on country's administrative boundaries, but on NDVI classes' spatial distribution which are results of hyper-temporal MODIS-NDVI data classification (section 2.4). The study area was composed of four NDVI Classes: NDVI class 24, NDVI class 54, NDVI class 70 and NDVI class 82. Looking at administrative boundaries in Rwanda; with 30 districts making up the whole country, the four sample NDVI classes intersect with almost all the districts of the country. However, the sample classes' biggest entities are located in seven districts namely: Bugesera district in Eastern province for class 24, Nyaruguru and Nyamagabe districts in Southern Province for class 54, Nyabihu in Western province and Musanze district in Northern province for class 70, and Musanze and Burera districts in Northern Province and Nyagatare district in Eastern province for class 82. The sample NDVI classes spatial distribution in Rwanda is presented by map in figure 2 below.

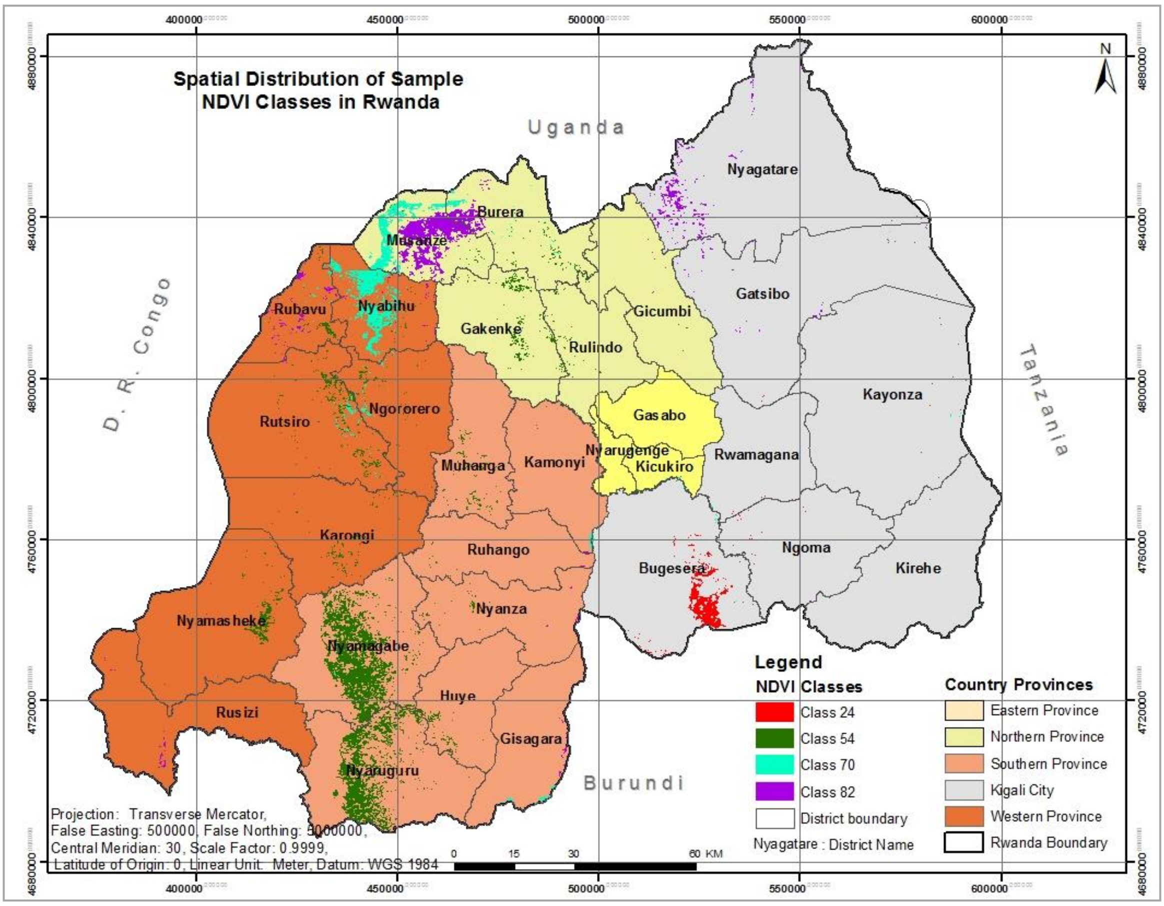


Figure 2: Study area map

Regarding the size of the sample NDVI classes; class 24 covers an area of 6,485 ha, class 54 covers an area of 57,031 ha, class 70 covers an area of 19,583 ha, and class 82 covers an area of 20,769 ha. The whole study area made by the four classes covers an area of 103,868 ha in total, which is equal to 9.3% of entire cropland area in Rwanda.

The spread of the sample NDVI classes in different parts of the country from West to East was one of the good factors making the study area to better represent the variability of the country's landscape. This is supported by the fact that Rwanda is composed of different 12 Agro-Ecological Zones (AEZ), with the main factor of diversity being topography, where in West is high altitude, medium altitude in the Centre and low altitude in East (Clay & Dejaegher, 1987). All these parts of the country were covered by part of the sample NDVI classes.

2.2. NDVI data

Many various satellite products and vegetation indices exist, with advantages and drawbacks. There exist more than 250 remote sensing indices till August 24, 2015 (The IDB Project, 2015). Possibility to allow separability of different land cover classes based on their differences in agro-ecological aspects and photosynthetic activities of vegetation, were the main criteria to select NDVI as the most appropriate remote sensing index for the present study.

There are a variety of differences in Rwanda, especially in agro-ecosystems. The differences are due to factors including applied agricultural practices, soil characteristics, phenological behaviours of the vegetation, water availability, and other various agro-ecological characteristics (Clay & Dejaegher, 1987; Twagiramungu, 2006). Mixture and variability making the nature a complex system, makes it also challenging to find a remote sensing product or index that can effectively approach the analysis and separability of areas with such differences. Nevertheless, various studies indicated that there is a strong correlation between NDVI and relevant agro-ecological factors of vegetation behaviours; including rainfall (water availability), topography, differences in soil type and management, biomass, vegetation cover density, and vegetation greenness (Benedetti & Rossini, 1993; Gillan, 2013; Matsushita et al., 2007; Nicholson, Davenport, & Malo, 1990). Most importantly, the researchers found a strong correlation between NDVI and plant photosynthetic activity, which actually depend mainly on the amount of sunlight reaching vegetation, and also water and nutrients available. However, NDVI is not a direct measure of any of the agro-ecological factors, being one of its limitations (Gillan, 2013).

Various studies used NDVI products from sensors such as AVHRR, MODIS, MERIS (Barbosa, Huete, & Baethgen, 2006; Benedetti & Rossini, 1993; Bolton & Friedl, 2013; Funk & Budde, 2009), either for natural phenomena studies or agro-ecological studies. For the current study, NDVI data from MODIS sensor onboard terra satellite, with 250 m spatial resolution (short product name: MOD13Q1) was used. The rationale was the high revisit time of 2 days, a reasonable spatial resolution than AVHRR and MERIS sensors given the small size of farms in the study area, and wide swath covering a large area. The MOD13Q1 data are produced as 16 days maximum value composites to reduce effects of clouds, directional reflectance and off-nadir viewing effects, minimize sun angle and shadow effects, and aerosol and water vapour effects (Holben, 1986; King, Closs, Wharton, & Myers, 2003). The data used in the present study are for 10 years, from 2004 till 2014, to ensure continuous and long-term homogeneity identification in different places within the study area.

2.3. NDVI data pre-processing and processing

The MOD13Q1 metadata documentation was the first step in data pre-processing. It provided information on the quality of the data, the level of pre-processing performed before the data release, and the remaining pre-processing tasks to be done for the data to be used for the research objective. 221 images from 2004 to 2014 were downloaded from MODIS data pool using MODIS Reprojection Tool Web interface (MRTWeb). The tool did not only facilitate data download but also provided different relevant options such as geographical subsetting and mosaicking. However, though the tool is called "reprojection tool", the data were downloaded with original characteristics, in TIFF format; in order to prevent possible big pixel shifts in case of early reprojection, as explained by Neteler (2011).

The next step was to rescale the NDVI data values. Original values were ranging from -3000 to 10 000 and stored under 16-bit signed integer (LP DAAC, 2014). They were rescaled to the normal pixel value range (0 to 255 for pixels stored under 8 bits). The rescale applied the following equation:

$$\boxed{\text{Rescaled NDVI} = \text{Original NDVI} * 0.02125 + 42.5 + 0.5} \quad \text{(de Bie, 2015)}$$

Equation 2: NDVI rescaling

To enable further analysis, all the 221 images were stacked into one layer. Then, the layer stack was cleaned and smoothed by application of Savitzky-Golay smoothing and differentiation filter to remove possible noise in the data. The Savitzky-Golay smoothing and differentiation filter fits the data points optimally to a polynomial in the least square logic, resulting in elimination of noise in the data, and fit with the optimal data (Luo, Ying, & Bai, 2005). The filter is integrated into ENVI software which was used for this task.

2.4. Integration of MODIS NDVI data with area frame sampling

In line with the objectives of the present study, this section details the improved methodology of integrating the hyper-temporal NDVI data with the area frame sampling method, which led to the design of better homogeneous area frames (NDVI classes) and sample sites (with the size of MODIS NDVI pixel size).

- **Land stratification**

The land stratification was based on similarity in vegetation cover behavior in an area over time. To determine the similar classes, unsupervised classification based on NDVI values was performed. This classification technique is the mostly used technique (Yuan, Lv, & Lu, 2015), because of its ability to obtain relatively precise performance through exploration of reflectance values over large electromagnetic spectrum domain, and get all information about the changes in a specific area in a satellite image (Bovolo & Bruzzone, 2007).

To achieve the optimal classification by the present study, the unsupervised classification applied Iterative Self-Organizing Data Analysis Technique (ISODATA) (Ball & Hall, 1965) as explained by Memarsadeghi, Mount, Netanyahu and Le Moigne (2007), through ERDAS software. The technique repeatedly performs entire classification, recalculating statistics as well, in an iterative process till an optimal threshold is reached (INTERGRAPH, 2013). To assign a class to a particular pixel, the technique uses minimum spectral distance. This gave confidence that the technique was appropriate to distinguish various homogeneous NDVI classes for the present study.

To perform the NDVI data classification, 10 was taken as a minimum number of classes, and 100 as a maximum number of classes, with a maximum number of iterations 50 and convergence threshold 1.0 (100%); as adapted from Ali et al. (2012). The authors observed that after 50 iterations there are no big changes in the classification results. For 100% convergence threshold, it is to make the classification stop when it is no longer possible for any pixel in the image to be assigned to a new class between iterations (INTERGRAPH, 2013).

To find optimal number of NDVI classes, separability analysis using best average and best minimum separability values, as detailed by Swain and Davis (1978) cited in Ali et al. (2012) was applied, and the values were plotted in excel application software in order to observe the peak in the values indicating the optimal classes number. The optimal NDVI classes were found to be 95.

Further, the classification results were overlaid with the recent Rwanda land use data, in order to find the amount of land cover content per NDVI class, and then be able to focus on the classes dominated with agriculture (cropland NDVI classes). The cropland NDVI classes were identified as those containing more than 50% of agriculture and were found to be 24. From the cropland NDVI classes, 4 sample NDVI classes were selected, considering their remarkable differences in temporal behaviours to represent various categories of the cropland NDVI classes. The temporal behaviour was observed from median NDVI values profiles. Medians were preferred over means, as mean tends to be affected by outliers in case of natural systems data which are usually skewed (Lund Research Ltd, 2013). Sample NDVI classes were limited to 4 due to, mainly, time available for the research.

❖ **Random-representative clustered sampling**

The area size of a sample site was equal to the pixel size of the used MODIS NDVI data; which was originally 231.66 * 231.66 (53 666 m2 ~ 5.37 ha) right after the data acquisition. However, after data conversion from raster data into shapefile, and conversion from original projection system "sinusoidal" (Enrique, 2010) to WGS84 (a projection system of other used Rwanda data), the pixel size became 231.92 * 230.37 m ~5.34 ha. With comparison to the shapefiles of Rwanda strata, it was noticed that this geometric transformation did not only result in pixel size change, but also in a small shift. The left upper corner of NDVI data pixel was shifted 1.69 m to the right, and the right lower corner was shifted 2.51 m to the right too. Nevertheless, this error was considered minor, given the size of the pixel and the focus of the study. So, finally, a sample site was 231.92 * 230.37 m (5.34 ha) in size.

The determination of sample size was mainly affected by time available for the research. In addition, it was challenging to determine the optimal sample size, because this can only be effectively done if a researcher is aware of heterogeneity level existing in the population (NIST/SEMATECH, 2013). In this regard, with the possibility to be surveyed in available field work time, 12 well-distributed sample sites were selected and surveyed per sample NDVI class. So, the sample size was 48 sample sites in total for all the four sample NDVI classes.

In order to select representative sample sites, random representative clustered sampling technique was applied. The clustering was due to logistic constraints in order to, mainly, save time. In every NDVI class was three clusters, and every cluster was composed of four sample sites (appendix 1). The following are the constraints applied to the random selection of sample sites:

- Location in inner part of NDVI class, in order to avoid influences from other classes outside;
- Proximity to the road, for accessibility;
- Distribution to be four sample sites (one cluster) in the North or West, other four sample sites (another cluster) in the Centre, and last four sample sites (last cluster) in the South or East of sample NDVI class.

The fact that the NDVI classes were determined based on the homogeneous behavior of an area over time, gave confidence that the 12 sample sites were effectively representative for each NDVI class. In addition, the spread of the sample sites covering all the main parts of NDVI class was important to cover maximum variability that might occur in a class due to its spatial distribution.

Finally, to ease the data collection process with a guide on the field, Rwanda aerial photographs used for delineation of the current Rwanda strata (MINIRENA & RNRA, 2009) were overlaid with the sample sites. This facilitated recognition of the fields' patterns and easily ensure right location within a sample site while on the field, as addition to the GPS navigation system.

2.5. Current area frame sampling in Rwanda

To generate agricultural statistics for the country, National Institute of Statistics of Rwanda (NISR) with other agriculture-related institutions including Ministry of Agriculture (MINAGRI), frequently conduct agricultural surveys all over the country. Since 2012, the frequency of the surveys is three times per year (every agricultural season) (NISR, 2013, 2015a, 2015b, 2015c). To make it possible, the institutions apply the multiple frame sampling method for sampling design and samples selection. The multiple frame sampling is made of two frame sampling methods: area frame sampling and list frame sampling. For the latter, a list of big farmers owning farmland over 3 ha (NISR, 2012a) is acquired, and then a complete survey is carried out on their farms, as they occupy a large area in the agricultural land and are not too many.

According to NISR (2013, 2015a, 2015b, 2015c), in regard to area frame sampling method, area frames (strata) for the surveys are constructed using 2008 aerial photographs (MINIRENA & RNRA, 2009). Through visual interpretation technique, homogeneous strata are delineated based on crops intensity. Later, in the defined homogeneous strata, the segments are delineated by the use of physical features such as roads, paths, rivers, etc. Table 1 presents the designed strata for the entire country during seasonal agricultural survey of 2012-2013. This is the country's agricultural survey that has been of focus by the present study in comparison with NDVI stratification, as it is the one with the data publicly available.

Strata	Description	Land Area (ha)
1	Intensive hillside cropland (50-100% cultivated)	1,535,000
2	Intensive marshland cropland (50-100% cultivated)	55,100
3	Extensive cropland (15-50% cultivated)	192,800
4	Non-cropland (0-15% cultivated)	73,900
5	Cities and towns (0-15% cultivated)	47,700
6	Water	130,200
7	National parks (defined by political boundaries)	219,000
8	Marshlands, riverbeds with potential for rice (0-15% cultivated	79,200
9	Forest	172,200
10	Tea plantation	23,200

Table 1: 2012-2013 Rwanda national agriculture survey strata (NISR, 2015b)

Based on the 10 designed homogeneous strata (table 1), three strata dominated by agriculture were used for further sampling design and survey. The survey strata were; **1**: intensive hillside cropland, **2**: intensive marshland cropland and **3**: extensive cropland, making up an area of 1,782,900 ha (68% of the country size) of agricultural land. Regarding the coverage of the survey strata, stratum 1 covers 86% of the entire national agricultural survey area. Stratum 2 covers 3%, and the stratum 3 covers 11% of the total agricultural survey area.

The map in figure 3 below illustrates spatial distribution of the 2012-2013 Rwanda strata for the national agricultural survey.

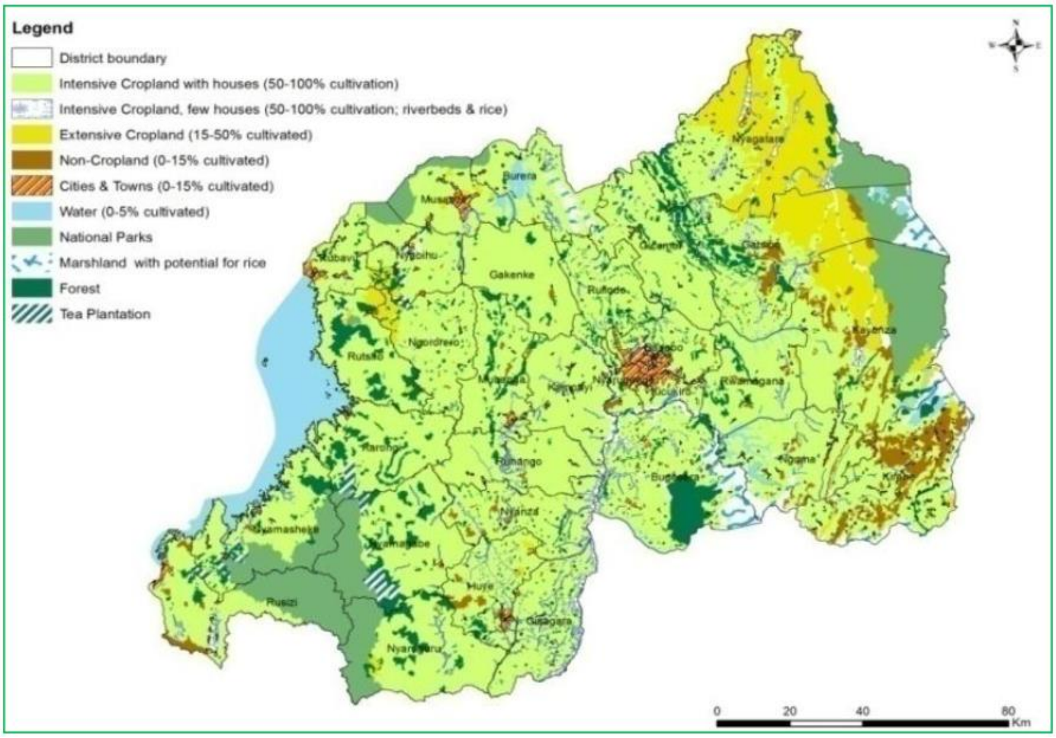


Figure 3: Spatial distribution of 2012-2013 Rwanda strata (NISR, 2015b)

The three survey strata were divided into Primary Sampling Units (PSUs) using physical boundaries (roads, paths, rivers…), from which representative sample PSUs were randomly selected using Probability Proportional to Size (PPS) technique. Per sample PSU, Secondary Sampling Units (SSUs or segments) were constructed, and one segment was randomly selected as a representative for the PSU. The size of one segment was about 20 hectares, except for stratum 3 where a segment was about 50 ha in size.

To exemplify the process, a PSU of 225 ha was divided into 11 segments of approximately 20 ha each, and one segment was randomly selected for the survey representing the 225 ha PSU (NISR, 2013). The sample segments for the whole country were 327, as presented in the following table 2.

Strata	Area in Ha	Number of sample segments
Stratum 1	1,535 000	295
Stratum 2	55,100	14
Stratum 3	192,800	18
Total	**1,782,900**	**327**

Table 2: Number of sample segments per sample Rwanda strata (NISR, 2015b)

Table 2 shows that the number of segments was in respect to the size of a stratum. The bigger the stratum, the higher the number of representative segments. Spatial distribution of the sample segments throughout the country is presented by the map in the next figure 4.

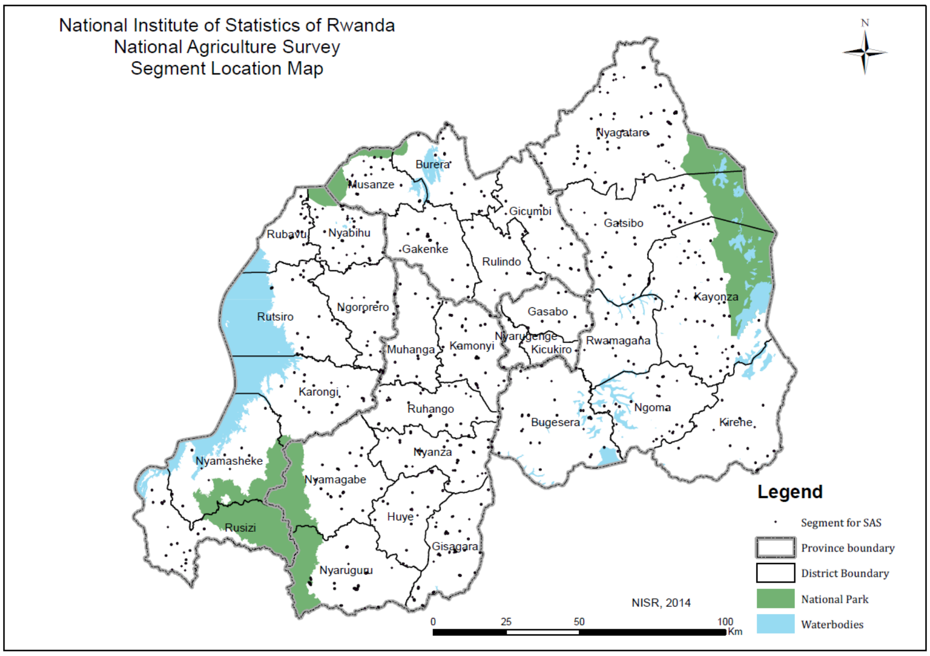


Figure 4: Spatial distribution of 2012-2013 NAS sample segments in Rwanda (NISR, 2015b)

All agricultural households in a sample segment were surveyed using questionnaire and interview. From the survey, agricultural information on crop area, crop type, crop yield, production, agricultural systems and applied farming techniques were collected, along with information on demographic and social characteristics of the respondents (NISR, 2015b). Then, the results from the segments were extrapolated to their respective PSUs, thereafter, to the entire country.

2.6. Spatial relationship between NDVI classes and Rwanda strata

In order to compare and assess spatial differences and similarities between NDVI classes and Rwanda strata, the 48 NDVI sample sites by the present study were overlaid with Rwanda strata (figure 5).

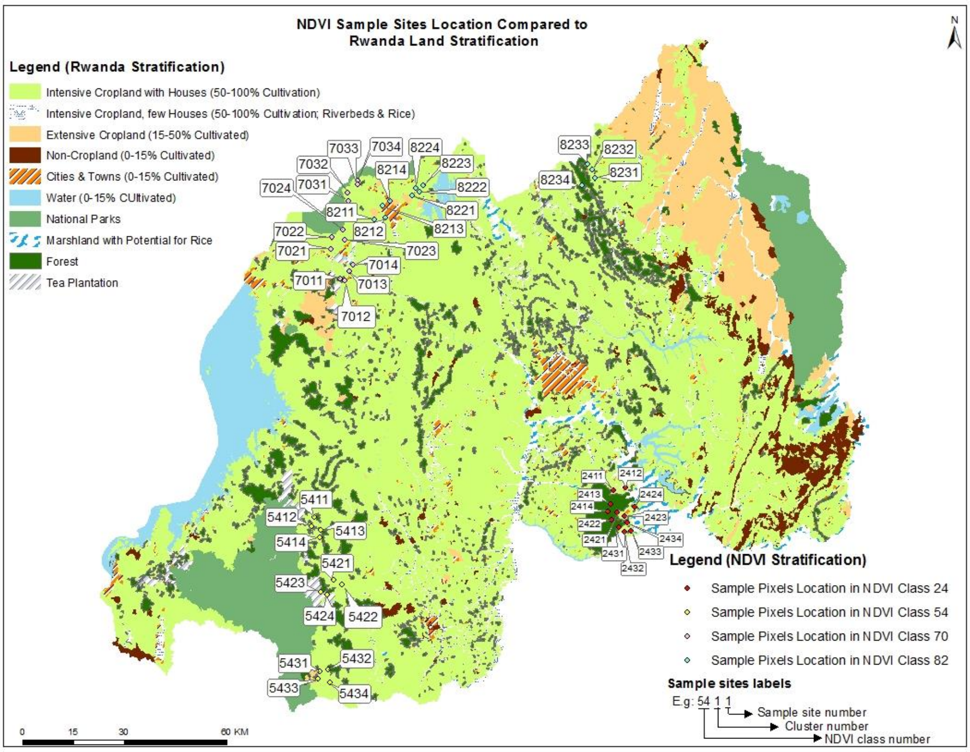


Figure 5: Spatial comparison between NDVI sample sites and Rwanda strata

By evaluating the overlay results in figure 5, it was noticed that majority of the sample sites (40 out of 47) were fully located inside Rwanda stratum 1 "intensive cropland with houses", one sample site (8414) fully in Rwanda stratum 5 "Cities and towns", two sample sites (2414 and 2422) fully in Rwanda stratum 9 "Forest", and two sample sites (5423 and 5424) fully in Rwanda stratum 10 "Tea plantation". The left two sample sites (2431 and 2432) were found mixed, with 18% of 2431 covered by Rwanda stratum 1 and 82% covered by Rwanda stratum 9. For sample site 2432, 50% was covered by Rwanda stratum 1 and other 50% covered by Rwanda stratum 5.

Some of the sample sites were noticed to have been misclassified by Rwanda stratification. For instance, sample site 2431 was classified with 18% of cropland and 82% of forest, while there was found no forest at all from the field. Instead, the sample site was covered by about 98% of agriculture and about 2% of bare soil. Sample site 2432 was classified with 50% of forest and 50% of agriculture, while actually, from the field, it contained about 66% of forest (which is part of its 79% non-agriculture land)) and agriculture occupies 21% (appendix 14).

Regarding geometric accuracy of the data overlay in figure 5, there was a shift between the data due to geometric transformation. First, the transformation of the NDVI data from raster with sinusoidal datum into shapefile data with wgs 84 datum resulted in a shift of 2.10 meters to the right. Second, overlaying the NDVI shapefiles with Rwanda strata, NDVI data shifted 7.55 meters to the left considering the Rwanda strata shapefiles as a reference. So, the overall geometric error was 5.45 meters to the left. Nevertheless, this error could not affect pure sample sites, only the mixed ones could be affected. For example, among all pure sample sites, site 8234 located in Rwanda stratum 1 was the closest sample site to a different Rwanda stratum "9" with a distance of about 20 meters to its left, which is smaller to the geometric error.

2.7. Research design

The following figure 6 contains flowchart summarizing the design of the present study to achieve its objectives, in comparison with the currently used area frame sampling method for agricultural surveys in Rwanda.

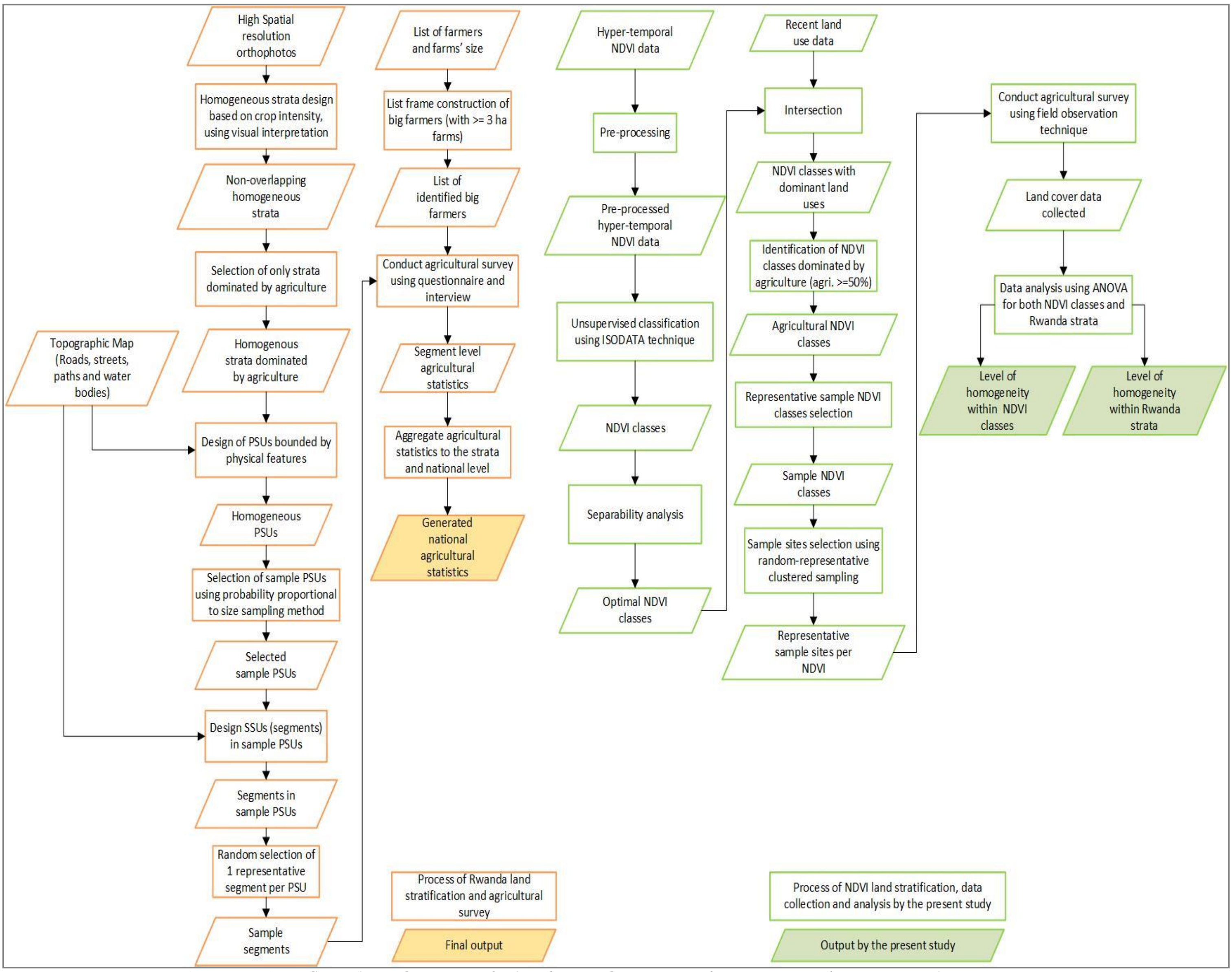


Figure 6: Comparative flowchart for Rwanda land stratification and present study's approach

2.8. Software used

For successful accomplishment of the present study, the following are the software used and their specific applications to the research:

No	Software	Application
1	ArcGIS 10.3	Spatial data analysis and visualization
2	Erdas Imagine 2015	Image processing
3	ENVI classic 5.2	Image processing
4	MRTWeb	Image data download
5	aNimVis	Visualization of temporal behaviour of NDVI
6	MS Office (Excel, Word)	Statistical data analysis, thesis writing
7	MS Visio	Illustrations

Table 3: Software used and their applications to the present study

2.9. Data collection

To reach the goal of the study, land covers and their area coverage data were collected in the field from October 10 till November 6, 2015. The photos in figure 7 present an example of field map used for the data collection on the field (for first sample site of NDVI class 54 as an example). The data were collected using field observation technique. Each land cover present in the sample site was recorded using a pen and later digitized and entered into computer system (figure 8).

Figure 7: Example of fieldwork map before (left) and after (right) data collection

To navigate to, and within the sample sites, IPAQ navigation system integrated with GPS was used. The data collected were later entered into the computer system through digitizing, and their areas were computed as presented by the figure 8 and table 4 below.

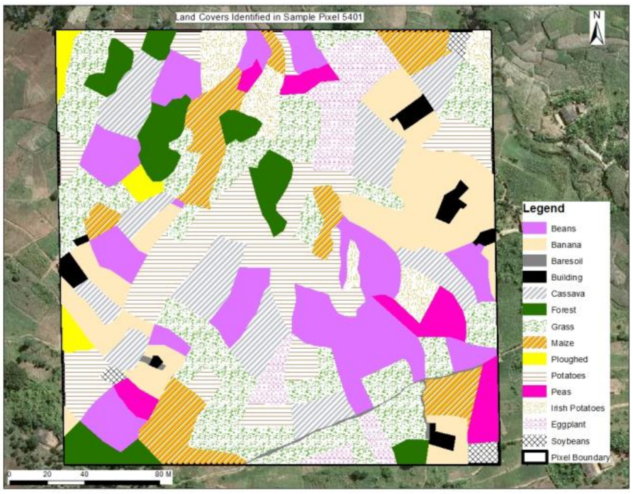


Figure 8: Collected land cover data after digitization

No	Cover	Area (ha)	Area (%)
1	Banana	0.61	11.49
2	Beans	0.77	14.47
3	Bare Soil	0.02	0.37
4	Building	0.08	1.52
5	Cassava	0.59	11.01
6	Eggplant	0.27	5.07
7	Forest	0.32	5.94
8	Grass	1.03	19.35
9	Irish potatoes	0.20	3.69
10	Maize	0.43	8.12
11	Peas	0.19	3.54
12	Sweet potatoes	0.71	13.32
13	Ploughed	0.07	1.36
14	Soybean	0.04	0.75
Total		**5.34**	**100**

Table 4: Land covers and their area

The next figure 9 summarises the frequency of identified land cover in all sample sites.

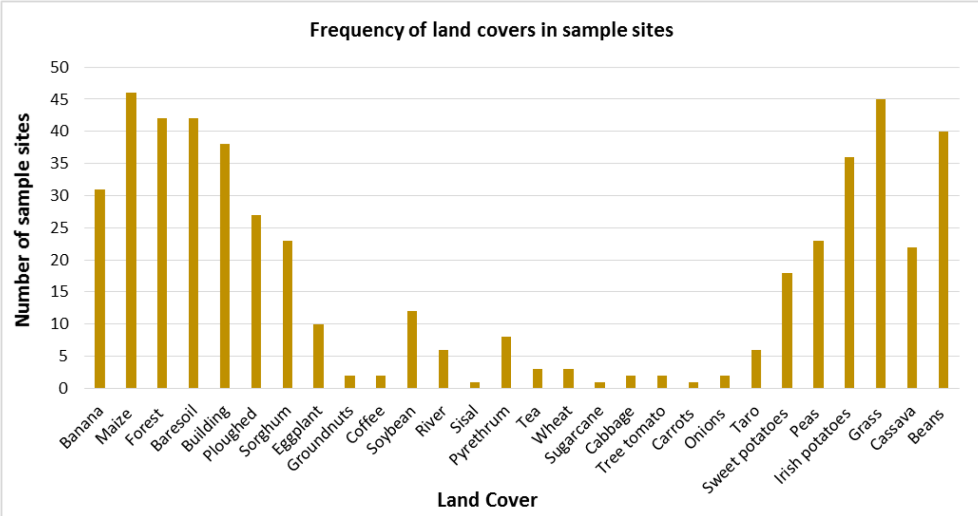


Figure 9: Frequency of identified land covers in sample sites

As presented in figure 9, land covers including maize, forest, grass, bare soil, beans and buildings were the most present in many sample sites. Land covers such as sisal, sugarcane and carrots were observed in not more than one sample site. In total, 28 land covers were identified in all sample sites.

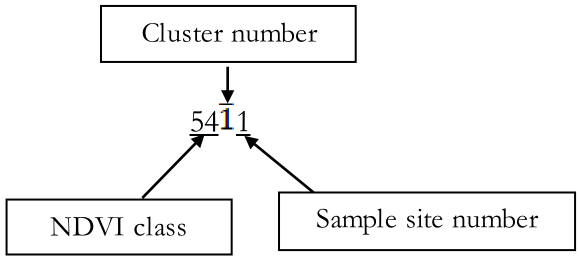
However, the third sample site of NDVI class 24 was not used for the research due to recent abrupt changes in the site. When reached the sample site, it was realized that it was, then, part of an area of about 20 ha which was recently cleared for construction of a new industrial zone in Bugesera district, and, therefore there was no presence of vegetation anymore. The picture in figure 10 below shows the sample site while field work.

Figure 10: Area for new industrial zone in Bugesera district (3rd sample site of class 24)

People on the site explained that they started the land clearing activities in August 2015 (2 months prior to the field work). They also informed that the area had a lot of agriculture including, mainly, cassava, and so, the new construction plan had to compensate the farmers for their land and agricultural activities. Given that the entire area was 100% bare due to these abrupt changes, the sample site was excluded from further analysis. So, the study remained with 47 sample sites for further data analysis.

2.10. New sample sites identification after field work

Before and during the fieldwork, the data were identified following the normal counting order, where, as an example: 5401 was the identity of the first surveyed sample site in NDVI class 54. The last sample site in this class was 5412 (12^{th} sample site in class 54). After field work, it was realized important to rename the sample sites according to their corresponding clusters per NDVI class. For instance, the new identity for sample site 5401 became 5411: with the following meaning:



So, 5411 was the first surveyed sample site in cluster 1 of NDVI class 54 (appendix 1).

2.11. Data analysis

Data analysis was conducted using the data of area covered by main crops in Rwanda agricultural season A: banana, maize, beans and cassava. It was performed in four stages for NDVI classes and two stages for Rwanda strata. First, using one-way Analysis of Variance (ANOVA) (Moore, McCabe, & Craig, 2009), the analysis sought to find out whether the NDVI classes are significantly different between them, to be concluded that they should be separate classes for agricultural survey. Second, Fisher's Least Square Difference (LSD) (Williams & Abdi, 2010) was carried out for NDVI classes pairwise comparison, to evaluate whether they are all pairs of the classes which are significantly different, or if some pairs are not. These two first stages analysed the differences between NDVI classes.

Third, one way ANOVA was carried out for clusters within the same NDVI class, to evaluate whether there are no significant differences within the same class. This ensured that a single NDVI class is homogeneous if there were no significant differences between its clusters, otherwise heterogeneous. Fourth, only in cases where was NDVI class with significant differences between its clusters, Fisher's LSD was carried out in order to identify the specific pairs of clusters with significant differences, and then quantify (in %) the level of homogeneity and heterogeneity of the class.

For Rwanda strata, it appeared that the collected data were only enough for analysis within Rwanda stratum 1 "intensive agriculture with houses". Among the 47 sample sites, 41 were found to be located in the Rwanda stratum 1, 3 sample sites in Rwanda stratum 9 "Forest", 2 sample sites in Rwanda stratum 10 "Tea plantation" and 1 sample site in Rwanda stratum 5 "Cities and towns". Due to insufficient degrees of freedom for statistical analysis, the three Rwanda strata: 5, 9 and 10 were excluded from further analysis. Then, the analysis was made only within Rwanda stratum 1.

The analysis was conducted in two stages to find out whether there are significant differences within Rwanda stratum 1. First, one way ANOVA was performed between clusters of the stratum, in terms of banana, maize, beans and cassava area. Second, only for cases where it appeared to be significant differences between the clusters, Fisher's LSD was carried out to identify specific pairs of clusters with significant differences in order to quantify the level of homogeneity and heterogeneity within the stratum.

3. RESULTS

This chapter presents the results obtained by the present study. The chapter starts by presenting the preliminary results obtained as the output from the hyper-temporal NDVI stratification which resulted in the selection of sample NDVI classes and sample sites for the study. The next section of the chapter presents the results as obtained from the collected field data analysis, through the application of analysis of variance between and within NDVI classes, and through the analysis of variance within Rwanda stratum 1 "intensive agriculture with houses".

3.1. Hyper-temporal NDVI stratification and sample sites in Rwanda

After running the classification of the hyper-temporal NDVI data, separability analysis was applied to find optimal NDVI classes. The best average and minimum separability values of the classification results from 10 to 100 were plotted in Microsoft excel software. The separability analysis revealed that optimal number of classes to distinguish heterogeneous areas in Rwanda from 2004 to 2014 were 95, as indicated by the figure below:

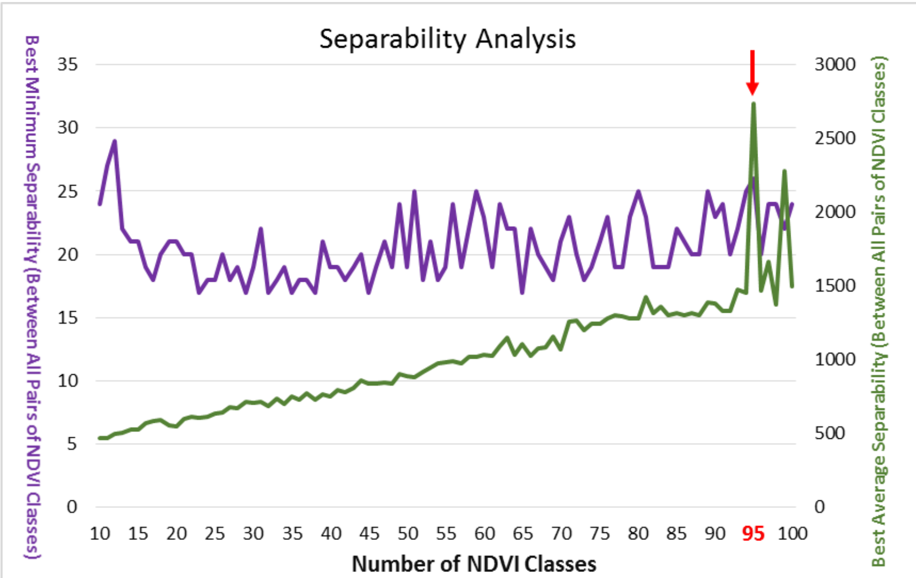


Figure 11: Best separable NDVI classes in Rwanda from 2004-2014 (95 classes)

The peak for both the best minimum and best average separability values in figure 11 indicated clearly that 95 classes were optimum, to be able to separate different areas according to their differences in vegetation cover in Rwanda for the 10 years.

Given that the agricultural classes were the focus of the study, the recent Rwanda land use data (RNRA, 2010) were overlaid with the optimal NDVI classes in order to know the classes dominated with agriculture. The recent Rwanda land use data were up to date till 2010 and contained various land uses including cropland, forestland, wetland, grassland, settlement and other land. The overlay results through the intersection are presented in the following table 5, showing every NDVI class with the amount of dominant land use.

Cropland NDVI Classes (Cropland >= 50%)		**Forestland NDVI Classes (Forest >= 50%)**		**Grassland NDVI Classes (Grassland >= 50%)**		**Wetland NDVI Classes (Wetland >= 50%)**		**Settlement NDVI Classes (Settlements >= 20%)**		**Other land NDVI Classes (Other land >= 20%)**	
Class	**Cropland area %**	**Class**	**Forest area (%)**	**Class**	**Grass Area (%)**	**Class**	**Wetland area (%)**	**Class**	**Settlement are (%)**	**Class**	**Other land area (%)**
20	91	**88**	99	**51**	96	**1**	100	**21**	48	-	-
70	84	**73**	98	**33**	92	**6**	100	**32**	43		
26	80	**92**	97	**31**	88	**2**	100	**38**	31		
50	78	**86**	97	**28**	84	**10**	100	**16**	25		
41	78	**40**	95	**39**	83	**7**	100				
72	77	**78**	94	**37**	79	**3**	100				
55	77	**89**	94	**48**	74	**13**	100				
45	77	**93**	93	**60**	70	**14**	100				
24	76	**91**	92	**49**	65	**12**	100				
34	70	**87**	92	**29**	50	**18**	100				
82	70	**90**	92			**11**	100				
30	70	**80**	92			**5**	100				
52	69	**94**	91			**4**	99				
66	66	**79**	90			**19**	99				
57	66	**84**	77			**9**	99				
58	66	**95**	75			**15**	98				
56	65	**75**	69			**25**	97				
42	63	**81**	65			**22**	97				
44	60	**62**	61			**27**	97				
47	55	**71**	56			**8**	95				
54	55	**65**	54			**17**	88				
61	52	**85**	54			**16**	72				
59	52	**74**	53			**83**	70				
63	50	**76**	50								

Table 5: NDVI classes and dominant land uses in Rwanda

From table 5, cropland NDVI classes were identified to be 24. The map showing the spatial distribution of the 24 cropland NDVI classes was produced, as presented in figure 12 below.

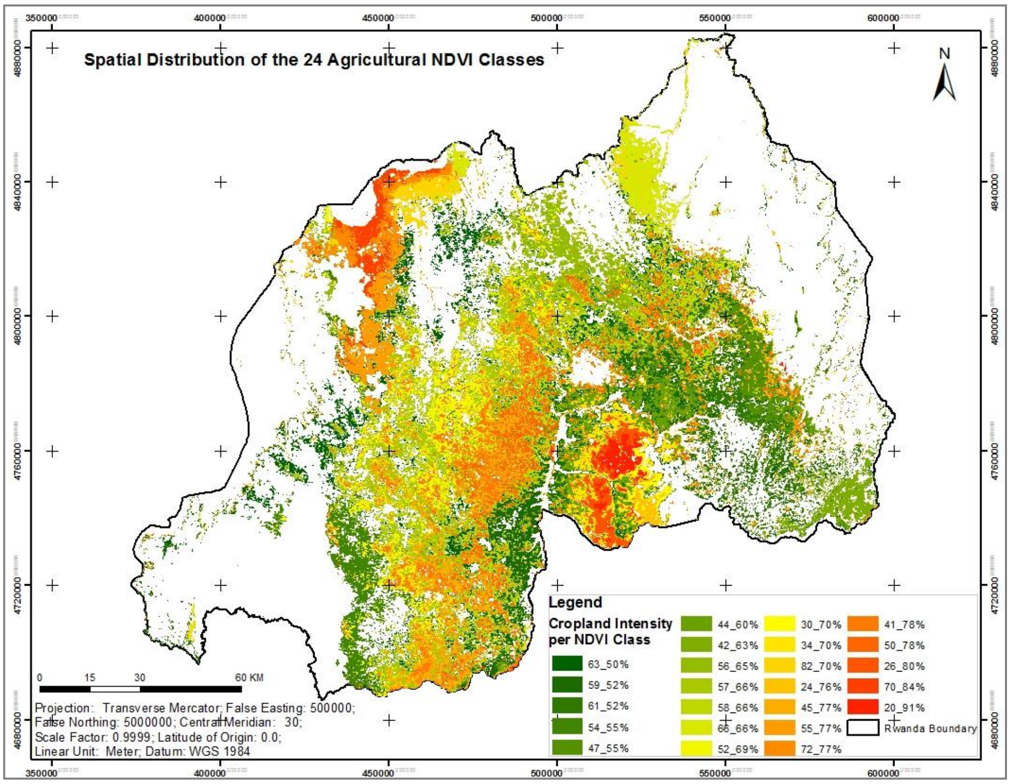


Figure 12: Spatial distribution of cropland NDVI classes in Rwanda

The map in figure 12 shows that, generally, most of the agriculture in Rwanda is located in central part of the country (from North to South with near East and Near West of the country), whereas agriculture is not dominant in far East, far West, and part of North of the country.

In order to observe the temporal behaviour of the 24 cropland NDVI classes over time, the NDVI profiles from 2004 to2014 were plotted as presented in figure 13 below.

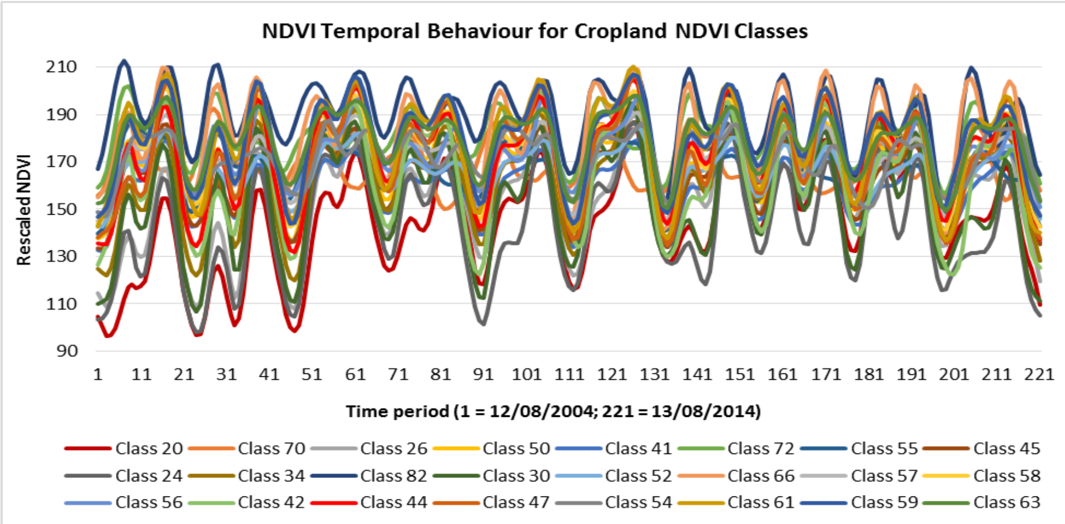


Figure 13: NDVI profiles of 24 cropland classes in Rwanda from 2004 to 2014

The figure 13 indicates that the cropland NDVI classes have been behaving differently, but they all had two main different vegetation growth or agricultural seasons every year, though in some classes, season A seems not to contain much crops' coverage as season B. To have a better overview of the classes' temporal behaviour per year, the following figure 14 presents the annual medians of the classes for 10 years.

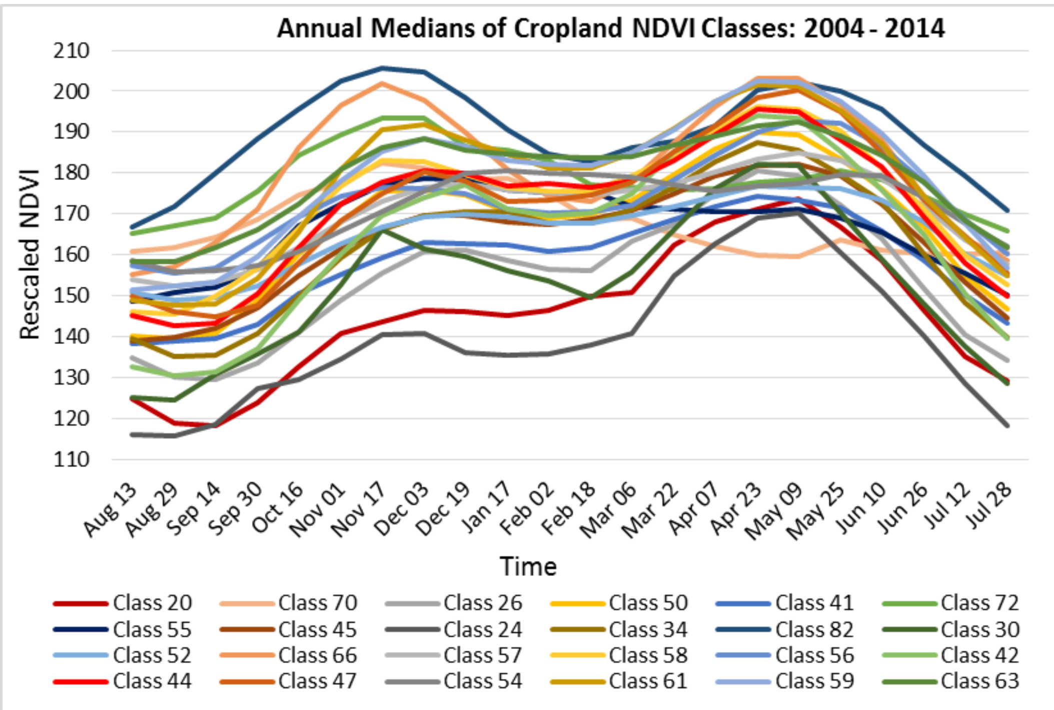


Figure 14: Annual medians of cropland NDVI classes (10 years period)

Figure 14 clearly shows that the cropland classes behaved differently, but with two main seasons every year. Some classes including 24 and 20 had not much of green vegetation as other classes such as 82, but still all were characterized by two different agricultural seasons.

As all the 24 cropland NDVI classes could not be surveyed given time available for the study, four sample NDVI classes were purposively selected given remarkable differences in NDVI profiles in order to represent different categories of classes: from those with less green vegetation over time to those with the highest amount of green vegetation over time. The following are medians NDVI profiles of the four selected sample NDVI classes by the present study.

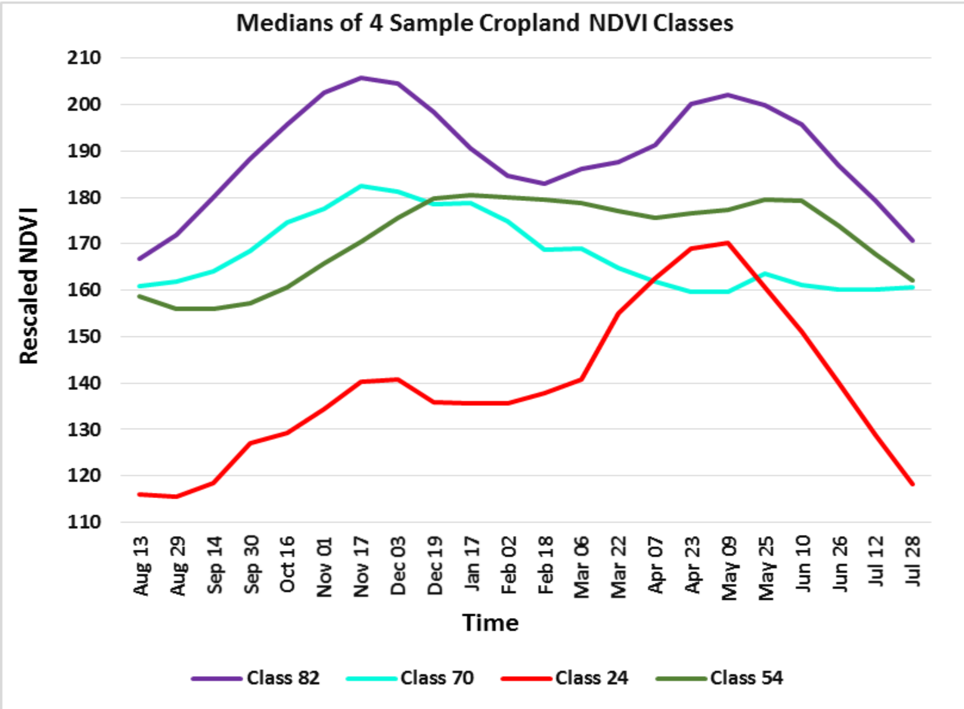


Figure 15: Temporal behaviour of the four sample cropland NDVI classes

The figure 15 shows that the four selected sample NDVI classes were different in terms of their NDVI values medians from 2004 till 2014. NDVI class 24 has been characterized by not much green vegetation in season A, but with much green vegetation in season B. This was confirmed by the farmers on the field who informed that season A is normally not productive in their district, only season B brings much agricultural production. NDVI class 54 looked as there was no difference between season A and season B. NDVI class 70 was almost opposite of class 24, where it had much green vegetation in season A but not much in season B. NDVI class 82 was different from others. It had two very distinctive agricultural seasons, with much green vegetation every season. Given these differences in the four classes, they were considered good representative sample classes from the rest of 20 cropland NDVI classes.

- **Representative sample sites in sample NDVI classes**

Random-representative clustered sampling was applied in order to select representative sample sites. The following figure 16 presents a map showing the location of all 48 selected sample sites in the four sample NDVI classes.

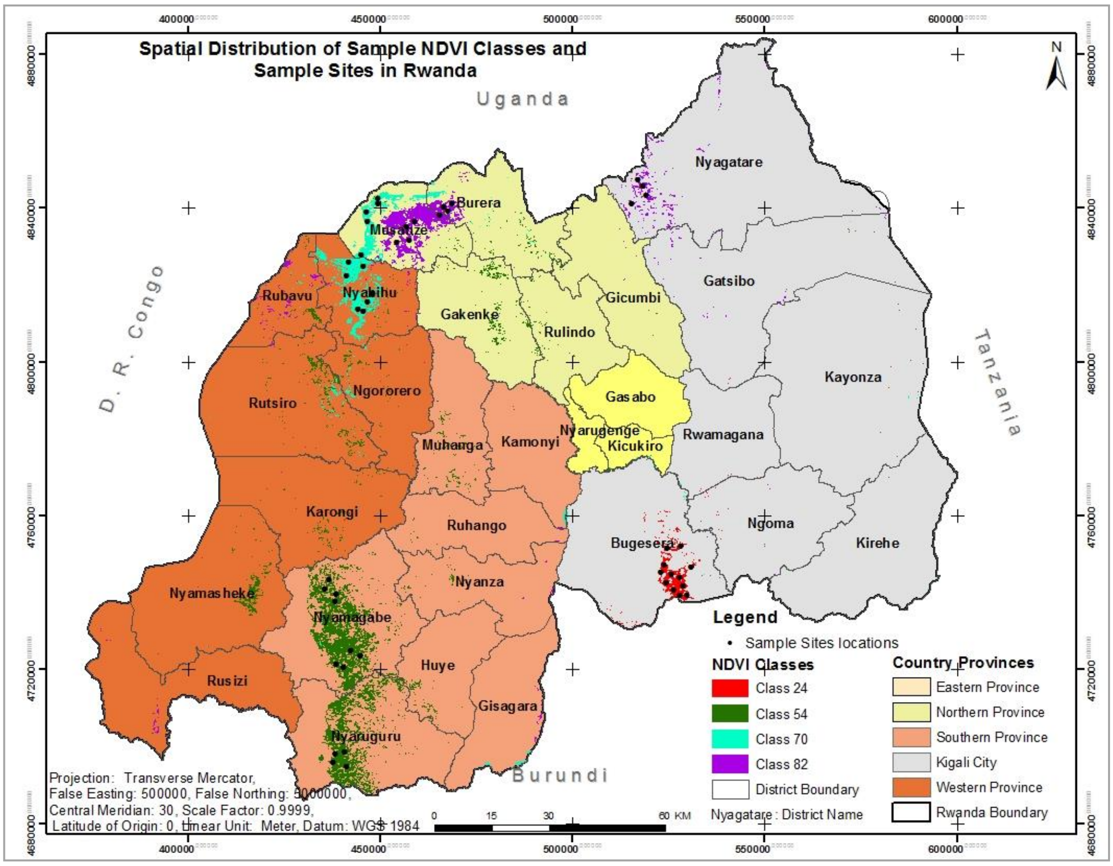


Figure 16: Spatial distribution of sample NDVI classes and sample sites in Rwanda

The map in figure 16 shows that the sample sites were grouped into three clusters per sample NDVI class and that they cover main parts of a class (South, Centre and North, or West, Centre and East of a class). This indicated that the applied methodology resulted in representative sample sites covering the main parts of the sample NDVI classes.

Fascinatingly, the NDVI classes selected based on remarkable temporal differences, were also remarkably different in spatial distribution. NDVI class 24 was mainly located in South East of Rwanda, class 54 was mainly located in South West, class 70 was mainly located in North West, and class 82 was mainly located in North and North East of the country. In fact, the remarkable differences in temporal behaviour had also affected the remarkable spatial differences of the classes.

❖ **Spatial relationship between NDVI classes and administrative boundaries**

Looking at the spatial coverage of NDVI classes in terms of administrative boundaries, the sample NDVI classes were intersected with Rwanda districts shapefiles. The results are presented in the following table 6, indicating coverage (in ha and %) of every sample NDVI class in Rwanda districts.

No	District	District Area (ha)	Class 24 Coverage		Class 54 Coverage		Class 70 Coverage		Class 82 Coverage	
			%	ha	%	ha	%	ha	%	ha
1	Bugesera	129,056.08	4.738	6115	0.082	106	0.193	249	0.024	31
2	Burera	64,455.94	0	0	1.422	917	1.117	720	4.173	2690
3	Gakenke	70,406.26	0	0	3.175	2235	0	0	0.127	89
4	Gasabo	42920.657	0.005	2	0.009	4	0	0	0	0
5	Gatsibo	158232.06	0.021	33	0.049	78	0.003	5	0.277	438
6	Gicumbi	82951.697	0.023	19	0.17	141	0	0	0.075	62
7	Gisagara	67919.689	0	0	0	0	0.486	330	0.362	246
8	Huye	58152.704	0	0	1.486	864	0.004	2	0	0
9	Kamonyi	65553.082	0	0	0.149	98	0.252	165	0.061	40
10	Karongi	99303.223	0	0	1.112	1104	0	0	0.038	38
11	Kayonza	193496.35	0.014	27	0.021	41	0.017	33	0.012	23
12	Kicukiro	16670.518	0	0	0.339	57	0.384	64	0	0
13	Kirehe	118485.01	0.009	11	0.005	6	0.013	15	0.014	17
14	Muhanga	64771.464	0	0	1.945	1260	0.016	10	0.008	5
15	Musanze	53038.067	0	0	0.363	193	11.512	6106	20.465	10854
16	Ngoma	86774.276	0.135	117	0.031	27	0	0	0.036	31
17	Ngororero	67898.574	0	0	1.134	770	0.71	482	0.031	21
18	Nyabihu	53149.709	0	0	0.409	217	19.33	10274	0.801	426
19	Nyagatare	192011.26	0	0	0.003	6	0	0	1.86	3571
20	Nyamagabe	109035.95	0	0	22.329	24347	0.059	64	0	0
21	Nyamasheke	117399.3	0.009	11	2.034	2388	0.005	6	0.023	27
22	Nyanza	67214.337	0	0	0.7	471	0	0	0.143	96
23	Nyarugenge	13395.006	0	0	1.224	164	0.053	7	0	0
24	Nyaruguru	101026.78	0	0	17.331	17509	0.01	10	0	0
25	Rubavu	38833.897	0	0	1.872	727	0.657	255	3.65	1417
26	Ruhango	62677.777	0	0	0.165	103	0	0	0.102	64
27	Rulindo	56698.292	0	0	0.718	407	0.007	4	0.009	5
28	Rusizi	95859.223	0	0	0.011	11	0	0	0.396	380
29	Rutsiro	115729.1	0	0	2.335	2702	0.587	679	0.055	64
30	Rwamagana	68196.249	0.085	58	0	0	0.027	18	0	11

Table 6: Sample NDVI classes coverage per district in Rwanda

Table 6 shows that NDVI class was not limited to a single administrative boundary. Instead, it extended to different districts (as long as there were similarities in terms of vegetation cover). This showed that by using NDVI stratification, similar areas are represented wherever they are in the country regardless their administrative boundaries, which would lower the bias in final agricultural estimates.

3.2. Land covers and their area coverage in NDVI classes

The following graph visualizes all data: land covers and their area (in %) per sample site per cluster per NDVI class, as collected from the field.

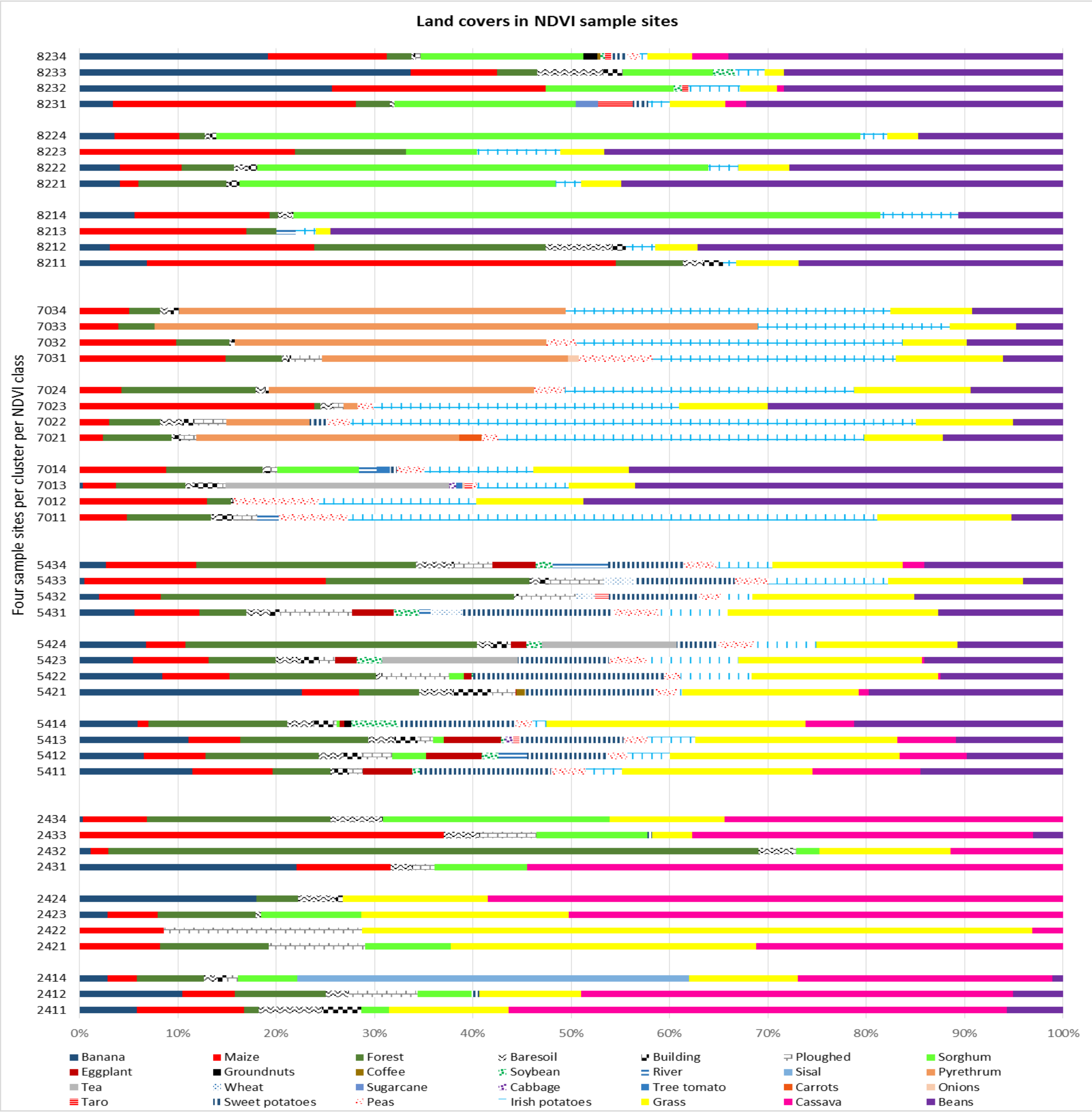


Figure 17: Visualization of land cover data in the 47 sample sites per cluster per NDVI class

From the data visual impression in figure 17, there are notable patterns in the data in accordance with the respective NDVI classes and clusters. There are differences in terms of crops data from one NDVI class to another, and similarities within the same NDVI class. Looking at the dominant land covers, for instance, cassava was present and remarkably dominant in class 24, whereas it was hardly present in other classes. Grass was present and abundant in class 24 and 54, but few in class 70, and very few in class 82. Sorghum was found in all the classes, but was dominant in class 82, and few in class 24, very few in class 54 and only in one sample site in class 70.

Beans were dominant in class 82 and 70. For class 70, beans were much more dominant in cluster 1, and cluster 2 and 3 had almost similar proportions of the crop. While on the field, local farmers from last three sample sites of cluster 1 in class 70, informed that their plots are located in beans model plots area. Indeed, the area was covered by large plots with beans of similar growth stage.

Banana was dominant in NDVI class 82, especially cluster 3 was mostly dominated by banana than any other previously surveyed area. This might be due to effect of crop intensification policy, as the local farmers informed that their region is a one of the regions allowed to grow banana, by crop regionalization programme.

Looking at maize coverage and distribution, it had almost similar proportions in all NDVI classes, except for class 82 where was much maize compared to other classes. Nevertheless, one sample site (2433) from NDVI class 24 was the site with the highest amount of maize, though the crop was not dominant in the class. The distribution of maize looks similar as that of forest, which was present in all NDVI classes, but dominant in class 24 and 54, and with the highest proportion in only one sample site (2432) of class 24.

Further, there were some crops present and dominant in only some classes. They include pyrethrum which was present and very dominant only in NDVI class 70 but limited to cluster 2 and 3 of the class. Another crop was irish potatoes which were present in class 54, 70 and 82, but only dominant in class 70. Actually, the area where NDVI class 70 is located is one of the main areas in Rwanda assigned for the production of irish potatoes by the crop intensification policy, because of rainfall and soil conditions favourable for the crop (Fané et al., 2006). This reflected a very traditional saying in Rwanda, that "much and very delicious irish potatoes are from *Ruhengeri*" (*Ruhengeri* is a former name for current Musanze district: a district where is located NDVI class 70 and part of class 82). Further, looking at sweet potatoes, they were present with dominance in NDVI class 54, and very few in some sample sites of class 24 and class 82.

There were some other crops and land covers that were very few, and only found in some sample sites and classes. These include onions only found in two sample sites of class 70, carrots only in one sample site of class 70, sisal only in one sample site of class 24 but with quite large area, cabbage in one sample site of class 54 and another one sample site of class 70, sugarcane only in one sample site of class 82, wheat in three sample sites of class 54, tea in two sample sites of class 54 and one sample site of class 70, coffee in one sample site of class 54 and another one sample site of class 82, groundnuts in one sample site of class 54 and another one sample site of class 82. Peas were present in class 54 and 70 with relatively small size, and also present in only one sample site of class 82. Other of these very few crops and land covers include eggplant, river, tree tomato, taro, and soybean.

Looking at the content of the land covers, some land covers were classified as the same. These include shrubs, patches of trees and big forests that were classified as forest. Paths, roads, streets and rocky land were classified as bare soil. The covers such as rivers and buildings were unique, but rivers were found in very limited sample sites, while the buildings were present in all NDVI classes but they could not dominate any class due to their small size within a sample site.

To better capture the differences in spatial distribution and coverage of the collected land cover data per sample sites per cluster per NDVI class, the following figure 18 presents a map of land covers spatial distribution in sample sites of cluster 1 of NDVI class 54 as an example. The rest of the maps are presented from appendix 2 to appendix 12.

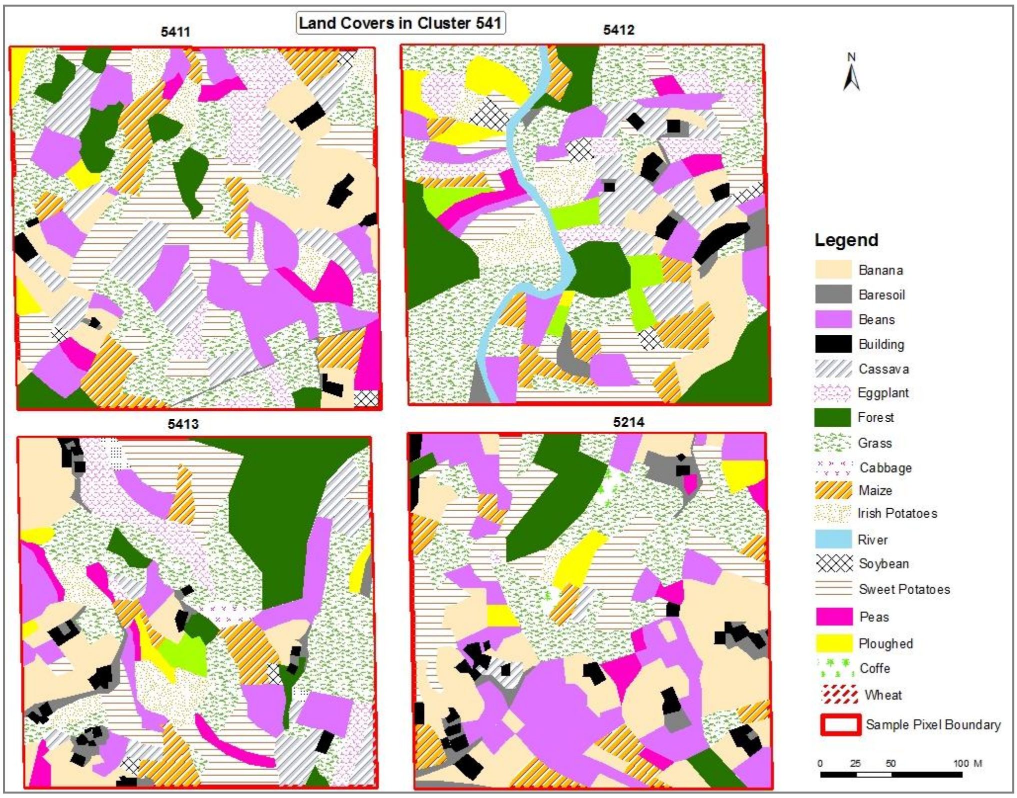


Figure 18: Land covers spatial coverage in cluster 1 of NDVI class 54

The differences in land cover areas in different NDVI classes were statistically investigated using analysis of variance and Fisher's LSD, to find out whether they are significant or not, and then be able to make conclusions about "between NDVI classes heterogeneity", and "within NDVI classes homogeneity" level.

3.3. Statistical analysis of variance of crops cover area in NDVI classes

To investigate whether NDVI class is homogeneous, the collected crops area coverages were analysed using ANOVA in terms of areas (in ha) covered by the four main crops of season A in Rwanda: banana, maize, beans and cassava, in the sample sites of the NDVI classes as illustrated by table 7 below.

Crops	Banana				Maize				Beans				Cassava			
Sample Site No	Class 24	Class 54	Class 70	Class 82	Class 24	Class 54	Class 70	Class 82	Class 24	Class 54	Class 70	Class 82	Class 24	Class 54	Class 70	Class 82
11	0.31	0.61	0	0.37	0.58	0.43	0.26	2.54	0.31	0.77	0.28	1.44	2.70	0.59	0	0
12	0.56	0.35	0	0.16	0.29	0.34	0.69	1.11	0.27	0.52	2.60	1.98	2.35	0.37	0	0
13	-	0.59	0.02	0	-	0.28	0.18	0.90	-	0.58	2.33	3.98	-	0.32	0	0
14	0.15	0.32	0	0.30	0.16	0.06	0.47	0.73	0.06	1.13	2.36	0.57	1.38	0.26	0	0
21	0	1.21	0	0.22	0.44	0.31	0.13	0.10	0	1.06	0.65	2.40	1.67	0.05	0	0
22	0	0.45	0	0.22	0.46	0.36	0.16	0.33	0	0.66	0.27	1.49	0.17	0.01	0	0
23	0.15	0.29	0	0	0.27	0.41	1.28	1.17	0	0.75	1.60	2.49	2.68	0.01	0	0
24	0.96	0.36	0	0.19	0	0.21	0.23	0.35	0	0.57	0.50	0.79	3.12	0	0	0
31	1.18	0.30	0	0.18	0.51	0.35	0.79	1.32	0	0.68	0.33	1.72	2.91	0	0	0.12
32	0.06	0.11	0	1.37	0.10	0.33	0.53	1.16	0	0.81	0.52	1.51	0.61	0	0	0.04
33	0	0.03	0	1.80	1.98	1.31	0.21	0.47	0.16	0.22	0.25	1.51	1.85	0	0	0
34	0.02	0.15	0	1.02	0.35	0.49	0.27	0.65	0	0.75	0.49	1.81	1.84	0.12	0	0.20
ANOVA results	**F (3.51) > Fcrit (2.82) P-value (0.02) < α (0.05)**				**F (2.92) > Fcrit (2.82) P-value (0.04) < α (0.05)**				**F (13.73) > Fcrit (2.82) P-value (2.02 E-6) < α (0.05)**				**F (45.34) > Fcrit (2.82) P-value (2.25 E-13) < α (0.05)**			
Classes are significantly different	**Yes**				**Yes**				**Yes**				**Yes**			

Table 7: ANOVA for sample NDVI classes in terms of Rwanda season A main crops

ANOVA results in table 7 show that the four sample NDVI classes were significantly different. To investigate specific significant differences between the classes, pairwise comparison was performed using Fisher's LSD, and the results are as follow:

Banana				
NDVI classes pairs	**Fisher's LSD**	**Absolute value of means' difference**	**Results**	**Pairs are significantly different**
(24; 54)	0.33	0.09	0.09 < 0.33	No
(24; 70)	0.33	0.31	0.31 < 0.33	No
(24; 82)	0.33	0.18	0.18 < 0.33	No
(54; 70)	0.32	4.74	4.74 > 0.32	Yes
(54; 82)	0.32	0.10	0.10 < 0.32	No
(70; 82)	0.32	0.49	0.49 > 0.32	Yes
Maize				
(24; 54)	0.40	0.06	0.06 < 0.40	No
(24; 70)	0.40	0.03	0.03 < 0.40	No
(24; 82)	0.40	0.44	0.44 > 0.40	Yes
(54; 70)	0.39	0.03	0.03 < 0.39	No
(54; 82)	0.39	0.50	0.50 > 0.39	Yes
(70; 82)	0.39	0.47	0.47 >0.39	Yes
Beans				
(24; 54)	0.56	0.63	0.63 > 0.56	Yes
(24; 70)	0.56	0.94	0.94 > 0.56	Yes
(24; 82)	0.56	1.74	1.74 > 0.56	Yes
(54; 70)	0.54	0.31	0.31 < 0.54	No
(54; 82)	0.54	1.10	1.10 > 0.54	Yes
(70; 82)	0.54	0.79	0.79 > 0.54	Yes
Cassava				
(24; 54)	0.39	1.79	1.79 > 0.39	Yes
(24; 70)	0.39	1.93	1.93 > 0.39	Yes
(24; 82)	0.39	1.90	1.90 > 0.39	Yes
(54; 70)	0.39	0.14	0.14 < 0.39	No
(54; 82)	0.39	0.11	0.11 < 0.39	No
(70; 82)	0.39	0.03	0.03 < 0.39	No

Table 8: Fisher's LSD results for the NDVI classes in terms of Rwanda season A main crops' area

The least significant differences results in table 8 indicate that some pairs of NDVI classes were significantly different while others were not. Nevertheless, for pairs with "no significant differences", there was no one which was not significantly different in terms of all four main crops. For instance, NDVI class 24 was not significantly different from class 54 in terms of banana and maize, but the same classes were significantly different in terms of beans and cassava. This made the classes relevant separately, for effective crops representativeness in an agricultural survey.

To evaluate homogeneity within NDVI classes, one way ANOVA was carried out for clusters within the same NDVI class (table 9 and 10). If there was "no significant differences" between the clusters of the same class, the class was considered homogeneous, otherwise heterogeneous.

	ANOVA between clusters in terms of banana area											
	Class 24			Class 54			Class 70			Class 82		
Sample site No	**Cluster 1**	**Cluster 2**	**Cluster 3**	**Cluster 1**	**Cluster 2**	**Cluster 3**	**Cluster 1**	**Cluster 2**	**Cluster 3**	**Cluster 1**	**Cluster 2**	**Cluster 3**
1	0.31	0	1.18	0.61	1.21	0.30	0.00	0.00	0.00	0.37	0.22	0.18
2	0.56	0	0.06	0.35	0.45	0.11	0.00	0.00	0.00	0.16	0.22	1.37
3	-	0.15	0	0.59	0.29	0.03	0.02	0.00	0.00	0	0	1.80
4	0.15	0.96	0.02	0.32	0.36	0.15	0.00	0.00	0.00	0.30	0.19	1.02
ANOVA Results	F (0.02) < Fcrit (4.46) P-value (0.98) > α (0.05)			F (2.78) < Fcrit (4.26) P-value (0.11) > α (0.05)			F (1) < Fcrit (4.26) P-value (0.40) > α (0.05)			F (6.57) > Fcrit (4.26) P-value (0.02) < α (0.05)		
Clusters are significantly different	No			No			No			Yes		
	ANOVA between clusters in terms of maize area											
1	0.58	0.44	0.51	0.43	0.31	0.35	0.26	0.13	0.79	2.54	0.10	1.32
2	0.29	0.46	0.10	0.34	0.36	0.33	0.69	0.16	0.53	1.11	0.33	1.16
3	-	0.27	1.98	0.28	0.41	1.31	0.18	1.28	0.21	0.90	1.17	0.47
4	0.16	0	0.35	0.06	0.21	0.49	0.47	0.23	0.27	0.73	0.35	0.65
ANOVA Results	F (0.76) < Fcrit (4.46) P-value (0.50) > α (0.05)			F (1.67) < Fcrit (4.26) P-value (0.24) > α (0.05)			F (0.02) < Fcrit (4.26) P-value (0.98) > α (0.05)			F (1.95) < Fcrit (4.26) P-value (0.20) > α (0.05)		
Clusters are significantly different	No			No			No			No		
	ANOVA between clusters in terms of beans area											
1	0.31	0	0	0.77	1.06	0.68	0.28	0.65	0.33	1.34	2.40	1.72
2	0.27	0	0	0.52	0.66	0.81	2.60	0.27	0.52	1.98	1.49	1.51
3	-	0	0.16	0.58	0.75	0.22	2.33	1.60	0.25	3.98	2.49	1.51
4	0.06	0	0	1.13	0.57	0.75	2.36	0.50	0.49	0.57	0.79	1.81
ANOVA Results	F (6.07) > Fcrit (4.46) P-value (0.02) < α (0.05)			F (0.43) < Fcrit (4.26) P-value (0.67) > α (0.05)			F (4.76) > Fcrit (4.26) P-value (0.04) < α (0.05)			F (0.11) < Fcrit (4.26) P-value (0.89) > α (0.05)		
Clusters are significantly different	Yes			No			Yes			No		

Table 9: ANOVA for clusters within NDVI class in terms of banana, maize and beans area

	ANOVA between clusters in terms of cassava area											
	Class 24			Class 54			Class 70			Class 82		
Sample Site ID	**Cluster 1**	**Cluster 2**	**Cluster 3**	**Cluster 1**	**Cluster 2**	**Cluster 3**	**Cluster 1**	**Cluster 2**	**Cluster 3**	**Cluster 1**	**Cluster 2**	**Cluster 3**
1	2.70	1.67	2.91	0.59	0.05	0	0.00	0.00	0.00	0.00	0.00	0.12
2	2.35	0.17	0.61	0.37	0.01	0	0.00	0.00	0.00	0.00	0.00	0.04
3	-	2.68	1.85	0.32	0.01	0	0.00	0.00	0.00	0.00	0.00	0.00
4	1.38	3.12	1.84	0.26	0	0.12	0.00	0.00	0.00	0.00	0.00	0.20
ANOVA Results	F (0.09) < Fcrit (4.46) P-value (0.91) > α (0.05)			F (21.5) > Fcrit (4.26) P-value (4 E-4) < α (0.05)			-			F (3.98) < Fcrit (4.26) P-value (0.06) > α (0.05)		
Clusters are significantly different	No			Yes			No			No		

Table 10: ANOVA for clusters within NDVI class in terms of cassava area

From tables 9 and 10, it was realized that only considering the maize crop, there was significant homogeneity within all NDVI classes at 95% confidence level. On another hand, considering banana crop, NDVI classes 24, 54 and 70 had significant homogeneity but not class 82. In terms of beans, NDVI classes 54 and 82 had significant homogeneity, but not class 24 and 70. In terms of cassava, NDVI class 24, 70 and 82 had significant homogeneity, but not class 54. For classes with significant differences, pairwise comparison was carried between their clusters using Fisher's LSD to find specific differences. The results are as follow:

Banana in NDVI class 82				
Clusters pairs	**Fisher's LSD**	**Absolute value of means differences**	**Results**	**Pairs are significantly different**
(1; 2)	0.66	0.50	0.50 < 0.66	No
(1; 3)	0.66	0.89	0.89 > 0.66	Yes
(2; 3)	0.66	0.94	0.94 > 0.66	Yes
Beans in NDVI class 24				
(1; 2)	0.15	0.21	0.21 > 0.15	Yes
(1; 3)	0.15	0.17	0.17 > 0.15	Yes
(2; 3)	0.14	0.04	0.04 < 0.14	No
Beans in NDVI class 70				
(1; 2)	1.14	1.13	1.13 < 1.14	No
(1; 3)	1.14	1.49	1.49 > 1.14	Yes
(2; 3)	1.14	0.36	0.36 < 1.14	No
Cassava in NDVI Class 54				
(1; 2)	0.14	0.36	0.36 > 0.14	Yes
(1; 3)	0.14	0.36	0.36 > 0.14	Yes
(2; 3)	0.14	0.01	0.01 < 0.14	No

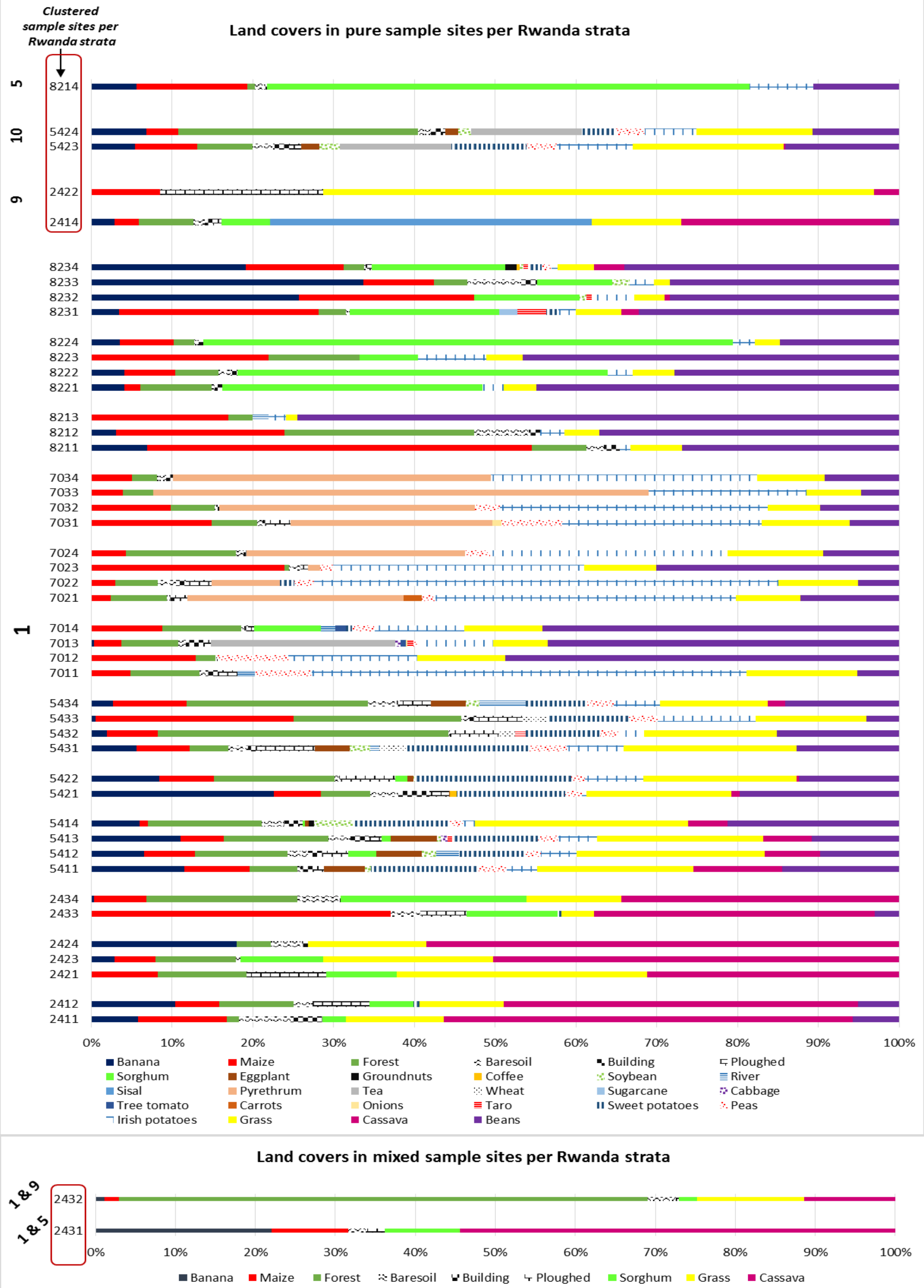
Table 11: Fishers' LSD results within NDVI classes in terms of season A main crops

The pairwise analysis results in table 11 show that: for banana in class 82, cluster 1 and 2 were not significantly different, but cluster 3 was significantly different from the rest. For beans in class 24, cluster 1 was significantly different from the rest, whereas cluster 2 and 3 were not significantly different. In class 70, only cluster 1 was significantly different from 2 in terms of beans. In terms of cassava, cluster 1 was significantly different from other clusters in NDVI class 54, the rest were not significantly different.

To summarize, among 48 pairs of the clusters within the four NDVI classes in terms of banana, maize, beans and cassava, 85% (41 pairs) were not significantly different and 15% (7 pairs) were significantly different. So, there was 85% of significant homogeneity within the four NDVI classes to effectively survey banana, maize, beans and cassava area, whereas there was 15% of significant heterogeneity suggesting further land stratification in order to maximize the homogeneity within the classes in terms of the four main crops' area.

3.4. Land covers and their area coverage in Rwanda strata

To compare NDVI classes and Rwanda strata, the two land strata were spatially overlaid. The overlay showed that 40 sample sites were fully located in Rwanda stratum 1 "intensive cropland with houses", 1 sample site fully in Rwanda stratum 5 "cities and towns", 2 sample sites fully in Rwanda stratum 9 "forest", 2 sample sites in Rwanda stratum 10 "tea plantation", 1 sample site partly in Rwanda stratum 1 and partly in Rwanda stratum 9, and 1 sample site partly in Rwanda stratum 1 and partly in Rwanda stratum 5, as presented in the next figure 19.



Meaning: - 1 = Intensive agriculture with (50 - 100% cultivated)
- 5 = Cities and towns (5 - 15% cultivated)
- 9 = Rwanda stratum 9 = Forest
- 10 = Tea plantation

Note: Sample sites in frame were not enough to represent their strata. They were excluded from further analysis

Figure 19: Land covers in full and mixed sample sites per Rwanda strata

As presented in figure 19, two sample sites were found to be mixed, and then excluded from further statistical analysis. In addition, sample sites for strata 5, 9 and 10 were not enough to represent their strata in statistical analysis, so, they were also excluded from further analysis, resulting in analysing variability differences only within Rwanda stratum 1.

3.5. Statistical analysis of variance for crops cover area within Rwanda stratum 1

Considering NDVI clusters within Rwanda stratum 1 as separate groups, ANOVA was carried out at 95% confidence level, to find out whether there are significant differences between the clusters (heterogeneous stratum) or whether there are no significant differences (homogeneous stratum). First, table 12 below presents sample sites and clusters to which they belong within Rwanda stratum 1.

Sample sites per cluster in Rwanda stratum 1											
Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7	Cluster 8	Cluster 9	Cluster 10	Cluster 11	Cluster 12
2411	2421	2433	5411	5421	5431	7011	7021	7031	8211	8221	8231
2412	2423	2434	5412	5422	5432	7012	7022	7032	8212	8222	8232
	2424		5413		5433	7013	7023	7033	8213	8223	8233
			5414		5434	7014	7024	7034		8224	8334

Table 12: Clusters and their sample sites within Rwanda stratum 1

Second, as ANOVA was conducted on the data of area covered by the season A main crops: banana, maize, beans and cassava, the following table 13 shows values of the area (ha) covered by every crop in every sample site per clusters within Rwanda stratum 1.

Banana area in Rwanda stratum 1's sample sites											
Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7	Cluster 8	Cluster 9	Cluster 10	Cluster 11	Cluster 12
0.31	0	0	0.61	1.21	0.3	0	0	0	0.37	0.22	0.18
0.56	0.15	0.02	0.35	0.45	0.11	0	0	0	0.16	0.22	1.37
	0.96		0.59		0.03	0.02	0	0	0	0	1.8
			0.32		0.15	0	0	0		0.19	1.02
Maize area in Rwanda stratum 1's sample sites											
0.58	0.44	1.98	0.43	0.31	0.35	0.26	0.13	0.79	2.54	0.1	1.32
0.29	0.27	0.35	0.34	0.36	0.33	0.69	0.16	0.53	1.11	0.33	1.16
	0		0.28		1.31	0.18	1.28	0.21	0.9	1.17	0.47
			0.06		0.49	0.47	0.23	0.27		0.35	0.65
Beans area in Rwanda stratum 1's sample sites											
0.31	0	0.16	0.77	1.06	0.68	0.28	0.65	0.33	1.44	2.4	1.72
0.27	0	0	0.52	0.66	0.81	2.6	0.27	0.52	1.98	1.49	1.51
	0		0.58		0.22	2.33	1.6	0.25	3.98	2.49	1.51
			1.13		0.75	2.36	0.5	0.49		0.79	1.81
Cassava area in Rwanda stratum 1's sample sites											
2.7	1.67	1.85	0.59	0.05	0	0	0	0	0	0	0.12
2.35	2.68	1.84	0.37	0.01	0	0	0	0	0	0	0.04
	3.12		0.32		0	0	0	0	0	0	0
			0.26		0.12	0	0	0		0	0.2

Table 13: Season A main crops' area coverage data per clusters within Rwanda stratum 1

The following table 14 presents results of ANOVA for the four main crops area, between clusters within Rwanda stratum 1.

ANOVA between clusters within Rwanda stratum 1																
	Banana				Maize				Beans				Cassava			
Cluster	*Count*	*Sum*	*Average*	*Variance*	*Count*	*Sum*	*Average*	*Variance*	*Count*	*Sum*	*Average*	*Variance*	*Count*	*Sum*	*Average*	*Variance*
Cluster 1	2	0.87	0.44	0.03	2	0.87	0.44	0.04	2	0.58	0.29	0.00	2	5.05	2.53	0.06
Cluster 2	3	1.11	0.37	0.27	3	0.71	0.24	0.05	3	0.00	0.00	0.00	3	7.47	2.49	0.55
Cluster 3	2	0.02	0.01	0.00	2	2.33	1.17	1.33	2	0.16	0.08	0.01	2	3.69	1.85	0.00
Cluster 4	4	1.87	0.47	0.02	4	1.11	0.28	0.02	4	3.00	0.75	0.08	4	1.54	0.39	0.02
Cluster 5	2	1.66	0.83	0.29	2	0.67	0.34	0.00	2	1.72	0.86	0.08	2	0.06	0.03	0.00
Cluster 6	4	0.59	0.15	0.01	4	2.48	0.62	0.22	4	2.46	0.62	0.07	4	0.12	0.03	0.00
Cluster 7	4	0.02	0.01	0.00	4	1.60	0.40	0.05	4	7.57	1.89	1.17	4	0.00	0.00	0.00
Cluster 8	4	0.00	0.00	0.00	4	1.80	0.45	0.31	4	3.02	0.76	0.34	4	0.00	0.00	0.00
Cluster 9	4	0.00	0.00	0.00	4	1.80	0.45	0.07	4	1.59	0.40	0.02	4	0.00	0.00	0.00
Cluster 10	3	0.53	0.18	0.03	3	4.55	1.52	0.80	3	7.40	2.47	1.79	3	0.00	0.00	0.00
Cluster 11	4	0.63	0.16	0.01	4	1.95	0.49	0.22	4	7.17	1.79	0.65	4	0.00	0.00	0.00
Cluster 12	4	4.37	1.09	0.47	4	3.60	0.90	0.16	4	6.55	1.64	0.02	4	0.36	0.09	0.01
ANOVA Results	F (4.81) > Fcrit (2.15) P-value (4 E-4) < α (0.05)				F (2.01) < Fcrit (2.15) P-value (0.07) > α (0.05)				F (5.23) > Fcrit (2.15) P-value (2 E-4) < α (0.05)				F (61.27) > Fcrit (2.15) P-value (1 E-16) < α (0.05)			
Clusters are significantly different	Yes				No				Yes				Yes			

Table 14: ANOVA results between clusters within Rwanda stratum 1

ANOVA results in table 14 show that there were significant differences (heterogeneity) between clusters within Rwanda stratum 1 in terms banana, beans and cassava area at 95% confidence level. Only the stratum had no significant differences in terms of maize crop. To investigate further and find out the specific significant differences, the pairwise comparison using Fisher's LSD was carried out as shown in appendices 15, 16 and 17 for banana, beans and cassava respectively. The pairwise comparison revealed that, in total, among all 264 pairs existing within Rwanda stratum 1 in terms of the four crops, 31% were significantly different, and 69% had no significant differences at 95% confidence level. In other words, within Rwanda stratum 1 was 31% significant heterogeneity, and 69% significant homogeneity in terms of the four main crops' area at 95% confidence level.

Next, it was assessed whether it is possible to use NDVI stratification to improve homogeneity within Rwanda stratum 1. To do so, sample areas that were found misclassified by Rwanda stratification were assigned to the true stratum according to the data collected from the field, as detailed in the following section.

3.6. Use of NDVI stratification to improve homogeneity within Rwanda stratum 1

The data collected from the field revealed that there were areas misclassified by Rwanda stratification. For instance, sample site 2431 had been classified as 82% forest and 18% as agriculture, while actually no forest was found in the site while on the field (figure 19). Instead, about 98% of the site was agriculture and about 2% was bare soil (appendix 14). Another sample site (2432) had been classified as 50% forest and 50% agriculture but the area was found to contain 66% of forest (which is part of its 79% non-agriculture), and agriculture was 21% (appendix 14). Sample site 8214 was classified as cities and towns, but, buildings and bare soil together did not even contribute to 2% of the site. Instead, almost 98% of the site was cropland. Also, sample sites 2414 and 2422 had been classified as forest, but it was found that forest does not even make 10% of the sites.

Nevertheless, some sample sites had been well classified by Rwanda stratification. These include the 40 sample sites located in stratum 1 “intensive agriculture with houses”, and 2 sample sites (5423 and 5424) located in stratum 10 “tea plantation”. The two sample sites were found to contain more than 10% of tea, and so they remained as tea plantation.

In order to improve the homogeneity of Rwanda strata, the misclassified sample sites were corrected according to the field observation. Sample sites 2414, 2422 and 8214 were assigned to Rwanda stratum 1, sample site 2432 was considered forest (Rwanda stratum 9) and sample sites 5423 and 5424 were still considered tea plantation (Rwanda stratum 10), as presented in the following figure 20.

Rwanda strata coded	- 1 : Intensive cropland with houses (50-100% cultivation) - 5 : Cities and towns (0-15% cultivated) - 9 : Forest - 10 : Tea plantation
Other symbols	- ✓ : 100% location within single Rwanda stratum (pure sample sites) - Red: Mixed sample sites. 2431 was adapted to Rwanda stratum 1, and 2432 was considered Rwanda stratum 9 - Sample sites in yellow (✓) were well classified but not enough to represent their stratum. So, they were excluded from further analysis - → Sample sites containing > 50% of agriculture but misclassified by Rwanda stratification. They were adapted to Rwanda stratum 1

Figure 20: Land covers in sample sites within Rwanda strata, with misclassified sample sites corrected

To carry out statistical analysis, first, the sample sites for strata other than Rwanda stratum 1 were considered not representative, and then, they were excluded from further analysis. The analysis was carried out only within Rwanda stratum 1.

3.7. Statistical analysis of variance for crops cover area within corrected Rwanda stratum 1

The following table 15 presents clusters and their sample sites (with corrected sample sites highlighted in green) within Rwanda stratum 1. ANOVA was carried out for the clusters to find out whether the number of significant differences can decrease compared to the previous original Rwanda stratum 1.

Sample sites per cluster within Rwanda stratum 1 (corrected sample sites in green)											
Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7	Cluster 8	Cluster 9	Cluster 10	Cluster 11	Cluster 12
2411	2421	2431	5411	5421	5431	7011	7021	7031	8211	8221	8231
2412	2422	2433	5412	5422	5432	7012	7022	7032	8212	8222	8232
2414	2423	2434	5413		5433	7013	7023	7033	8213	8223	8233
	2424		5414		5434	7014	7024	7034	8214	8224	8334

Table 15: Rwanda stratum 1 clusters with corrected sample sites

The next table 16 presents the crops area data for the correct sample sites per cluster within Rwanda stratum 1.

Banana area in corrected Rwanda stratum 1's sample sites											
Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7	Cluster 8	Cluster 9	Cluster 10	Cluster 11	Cluster 12
0.31	0.00	1.18	0.61	1.21	0.30	0.00	0.00	0.00	0.37	0.22	0.18
0.56	0.00	0.00	0.35	0.45	0.11	0.00	0.00	0.00	0.16	0.22	1.37
0.15	0.15	0.02	0.59		0.03	0.02	0.00	0.00	0.00	0.00	1.80
	0.96		0.32		0.15	0.00	0.00	0.00	0.30	0.19	1.02
Maize area in corrected Rwanda stratum 1's sample sites											
0.58	0.44	0.51	0.43	0.31	0.35	0.26	0.13	0.79	2.54	0.1	1.32
0.29	0.46	1.98	0.34	0.36	0.33	0.69	0.16	0.53	1.11	0.33	1.16
0.16	0.27	0.35	0.28		1.31	0.18	1.28	0.21	0.9	1.17	0.47
	0		0.06		0.49	0.47	0.23	0.27	0.73	0.35	0.65
Beans area in corrected Rwanda stratum 1's sample sites											
0.31	0	0	0.77	1.06	0.68	0.28	0.65	0.33	1.44	2.4	1.72
0.27	0	0.16	0.52	0.66	0.81	2.6	0.27	0.52	1.98	1.49	1.51
0.06	0	0	0.58		0.22	2.33	1.6	0.25	3.98	2.49	1.51
	0		1.13		0.75	2.36	0.5	0.49	0.57	0.79	1.81
Cassava area in corrected Rwanda stratum 1's sample sites											
2.70	1.67	2.91	0.59	0.05	0	0	0	0	0	0	0.12
2.35	0.17	1.85	0.37	0.01	0	0	0	0	0	0	0.04
1.38	2.68	1.84	0.32		0	0	0	0	0	0	0
	3.12		0.26		0.12	0	0	0	0	0	0.2

Table 16: Season A main crops' area coverage in corrected sample sites per cluster in Rwanda stratum 1

The following table 17 presents ANOVA results between clusters, with corrected sample sites within Rwanda stratum 1.

ANOVA between clusters within corrected Rwanda stratum 1																
	Banana				Maize				Beans				Cassava			
Cluster	*Count*	*Sum*	*Average*	*Variance*	*Count*	*Sum*	*Average*	*Variance*	*Count*	*Sum*	*Average*	*Variance*	*Count*	*Sum*	*Average*	*Variance*
Cluster 1	3	1.02	0.34	0.04	3	1.03	0.34	0.05	3	0.64	0.21	0.02	3	6.43	2.14	0.47
Cluster 2	4	1.11	0.28	0.21	4	1.17	0.29	0.05	4	0.00	0.00	0.00	4	7.64	1.91	1.71
Cluster 3	3	1.20	0.40	0.46	3	2.84	0.95	0.81	3	0.16	0.05	0.01	3	6.60	2.20	0.38
Cluster 4	4	1.87	0.47	0.02	4	1.11	0.28	0.02	4	3.00	0.75	0.08	4	1.54	0.39	0.02
Cluster 5	2	1.66	0.83	0.29	2	0.67	0.34	0.00	2	1.72	0.86	0.08	2	0.06	0.03	0.00
Cluster 6	4	0.59	0.15	0.01	4	2.48	0.62	0.22	4	2.46	0.62	0.07	4	0.12	0.03	0.00
Cluster 7	4	0.02	0.01	0.00	4	1.60	0.40	0.05	4	7.57	1.89	1.17	4	0.00	0.00	0.00
Cluster 8	4	0.00	0.00	0.00	4	1.80	0.45	0.31	4	3.02	0.76	0.34	4	0.00	0.00	0.00
Cluster 9	4	0.00	0.00	0.00	4	1.80	0.45	0.07	4	1.59	0.40	0.02	4	0.00	0.00	0.00
Cluster 10	4	0.83	0.21	0.03	4	5.28	1.32	0.69	4	7.97	1.99	2.09	4	0.00	0.00	0.00
Cluster 11	4	0.63	0.16	0.01	4	1.95	0.49	0.22	4	7.17	1.79	0.65	4	0.00	0.00	0.00
Cluster 12	4	4.37	1.09	0.47	4	3.60	0.90	0.16	4	6.55	1.64	0.02	4	0.36	0.09	0.01
ANOVA Results	F (3.67) > Fcrit (2.10) P-value (0.002) < α (0.05)				F (1.77) < Fcrit (2.10) P-value (0.10) > α (0.05)				F (4.81) > Fcrit (2.10) P-value (2 E-4) < α (0.05)				F (13.33) > Fcrit (2.10) P-value (5 E-9) < α (0.05)			
Clusters are significantly different	Yes				No				Yes				Yes			

Table 17: ANOVA results between clusters with corrected sample sites within Rwanda stratum 1

Table 17 shows that, even after correcting Rwanda stratum 1, there were still significant differences between its clusters in terms of banana, beans and cassava at 95% confidence level. Only in terms of maize crop, the stratum remained with no significant differences. Nevertheless, the F values had lowered and p-values increased compared to the original Rwanda stratum 1. To investigate further the specific differences after correction, the pairwise comparison was carried out using Fisher's LSD as presented by appendices 18, 19 and 20 for banana, beans and cassava respectively.

The Fisher's LSD results revealed that, after correcting the Rwanda stratum 1 using NDVI classification, the significant differences became 26% between the total pairs within the stratum, and 74% of the left pairs had no significant differences at 95% confidence level. In other words, homogeneity within the stratum increased by 5%, in terms of banana, maize, beans and cassava area data at 95% confidence level, after correcting the stratum using NDVI stratification.

4. DISCUSSION

This chapter discusses key findings of the present study in line with research objectives. It starts discussing the output from hyper-temporal NDVI data stratification for Rwanda and its representativeness of different areas for agricultural surveys in the country. Further, the chapter discusses differences found between and within NDVI classes, differences found in Rwanda strata, and then compare NDVI classes' differences and those within Rwanda strata, and shows how the NDVI classification was used to improve Rwanda stratification. Finally, the chapter discusses key differences and similarities between the two land stratification methods to give insights about both their advantages and drawbacks.

4.1. NDVI classes and their representativeness for agricultural surveys

The hyper-temporal NDVI stratification resulted in different NDVI classes, from which through integration with recent Rwanda land use data, the cropland NDVI classes were identified. The four selected sample NDVI classes were selected based on their remarkable difference in temporal behaviour from 2004-2014. After observing their spatial distribution, it was realized that the sample NDVI classes were not only different in their temporal behaviour but also in their spatial distribution from West to East of the country.

The remarkable differences in spatial distribution could be related to differences in areas in terms of altitude and rainfall patterns in Rwanda as elaborated by Clay and Dejaegher, 1987; Twagiramungu (2006). NDVI class 24 was mainly located in South East of Rwanda; a region known for low altitude and very low rainfall. NDVI class 54 was mainly located in South West; a region with medium altitude and medium rainfall. NDVI class 70 was mainly located in North West of the country; a region known for the highest altitudes in the country with the highest rainfall throughout a year. NDVI class 82 was mainly located in North and North East of the country; a region known for high altitudes (mainly in the North) and medium rainfall. This revealed that the remarkable differences in temporal behaviour, also resulted in remarkable spatial differences between the sample NDVI classes. This emphasised the ability to use hyper-temporal NDVI stratification to produce different strata, effectively representing different areas for agricultural surveys.

The results also showed the possibility of aggregating the agricultural statistics collected according to NDVI classes, to administrative boundaries. Actually, agricultural statistics are typically collected and then aggregated to administrative boundaries (Cotter et al., 2010). In this regard, the present study related the NDVI classes to Rwanda districts. The districts are head of four tiers of the local government in Rwanda (CLGF, 2011), and they are the administrative level to which statistics are initially aggregated. By relating the NDVI classes to the districts through the intersection, it was clarified that the NDVI classes were not limited to, or guided by administrative boundaries. The NDVI classes extended to any area where there are similar areas in terms of vegetation cover (NASA, 2015; Tucker & Sellers, 1986). It was found that the four sample NDVI classes had portions in almost all districts of the country (table 6) indicating similar areas in terms of vegetation cover in different districts of the country. This would be the basis for the major activity of agricultural survey after data collection, of aggregating the agricultural statistics at the country level, and would be expected to result in less biased output with small error, given that the NDVI classes maximize internal homogeneity, and minimize similarities with neighbouring and other NDVI classes.

4.2. Differences between and within NDVI classes

Differences between the sample NDVI classes were assessed in terms of Rwanda season A main crops: banana, maize, beans and cassava area. The results revealed that the NDVI classes were generally significantly different at 95% confidence level. Looking at specific "no significant differences", for instance, NDVI class 24 was not significantly different from NDVI class 54 in terms of banana and maize, but the same classes were significantly different in terms of beans and cassava. So, given that all the NDVI classes had significant differences between either all or some of their pairs (table 8), they were generally significantly different, and so, they had to be separate classes for effective representativeness in surveying banana, maize, beans and cassava crops area.

Looking at homogeneity within the classes, a class should be homogeneous if it contains no significant differences in the data collected from its various clusters. This is supported by the statement by Clay & Dejaegher (1987), that regions for agricultural surveys should maximize intra-regional homogeneity regarding agricultural aspects, and make their inter-regional homogeneity minimal. The analysis showed that the NDVI class 24's clusters were not significantly different in terms of area covered by banana, maize and cassava, at 95% confidence level. Nevertheless, in terms of beans, cluster 1 was significantly different from the rest, but also, there were no significant differences between cluster 2 and 3.

The significant differences within NDVI class 24 might have been due to abrupt changes brought by crop regionalization program under crops intensity policy in Rwanda, which started recently in 2012 (Cantore, 2013; Kathiresan, 2012). The programme determines areas and suitable districts for different crops in the country. NDVI class 24 is located in Bugesera district, one of the districts selected to grow mainly cassava (FAO, 2015a) with small zones for other crops, which is likely to bring discrepancies in areas covered by crops other than cassava. This might have affected cluster 1 of class 24 to have few beans grown, and no beans in cluster 2 at all, and very few in cluster 3.

Next, NDVI class 54's clusters had no significant differences in terms of banana, maize and beans, at 95% confidence level. However, the cluster 1 was significantly different from other 2 clusters, but there were no significant differences between cluster 2 and 3. The significant differences might have been also due to the crop regionalization programme.

NDVI class 70's clusters had no significant differences in terms of banana, maize and cassava. Nonetheless, cluster 1 and 3 were significantly different in terms of beans, but the rest had no significant differences. The cause of the significant differences within this class might have been another recent programme "*agasozi indatwa*" (crops model plots). The programme is part of model village policy (MININFRA, 2009): which started initially focusing on human settlement, but later extended to focus on agriculture as well. The programme selects a small area for growing a certain specific and unique crop, as a model for the rest of the region for that crop (Kathiresan, 2012). While the field wok, local farmers from sample sites 7012, 7013 and 7014 informed that they are located in beans crop model plots. A big area of the sample sites had uniquely beans of the same type and similar growth stage. In other sample sites of class 70, there was also beans, but with similar size as other different crops, which might have resulted in disparities in beans cover area within the class.

Finally, NDVI class 80's clusters had no significant differences in terms of maize, beans and cassava area. But, there appeared to be significant differences in terms of cassava between cluster 3 and the rest, but no significant differences between cluster 1 and 2. Also, it is more likely that this class had undergone effect of crop regionalization programme, because 8 first surveyed sample sites were located in the North of the country; a region determined mainly for cultivation of beans and irish potatoes given rainfall and soil conditions favourable for the crops (Fané et al., 2006), whereas other last 4 sample sites were located in North East of the country; a region determined mainly for cultivation of banana and beans. The crop regionalization programme only considers administrative boundary (mainly districts) (Cantore, 2013), whereas NDVI classes are in accordance with natural behaviour of the area (Nemani, 2014), which resulted in the same NDVI class 82 to be located in different areas of the country; subsequently with different crops allocation by the crop regionalization programme.

In summary, the NDVI classification resulted in different NDVI classes, also covering a big range of Rwanda natural differences in different areas, which is important for agricultural surveys. Furthermore, the NDVI classes were generally significantly different in terms of Rwanda season A main crops. Nevertheless, the homogeneity within the classes was not maximal. It was found that 85% of total clusters pairs within the four sample NDVI classes were not significantly different in terms of season A main crops' areas at 95% confidence level, but the rest 15% was significantly different. Nonetheless, the significant differences are likely due to the effect of recent agricultural policies in Rwanda such as; crops regionalization programme under crops intensification policy (Cantore, 2013; Kathiresan, 2012), *agasozi indatwa* (model plots) initiative extended from model village policy (MININFRA, 2009). The programmes determine areas suitable for various crops, and the areas have to grow only the determined suitable crops. Though the policies aim at boosting agriculture production in the country, sometimes they result in abrupt changes in different agricultural areas of the country, as most of the time farmers do not adapt to the new changes so quickly.

4.3. Differences in Rwanda strata

Rwanda stratification for agricultural survey classified the country into 10 different strata, among which 3 were of interest for the agricultural survey. By overlaying the 47 surveyed sample sites by the present study with the Rwanda strata, it was found that majority of the sample sites (40 sites) were fully located within the stratum 1 "intensive cropland with few houses", 1 sample site within stratum 5 "cities and towns", 2 sample sites within stratum 9 "forest", 2 sample sites within Rwanda stratum 10 "tea plantation", and other 2 sample sites were found mixed. It was noticed that there could not be performed statistical analysis of differences between all the strata, given the insufficient sample sites to represent stratum 5, 9 and 10. Also, the 2 mixed sample sites were excluded from the analysis. So, the analysis was only carried within Rwanda stratum 1.

Analysis of variance was carried out between clusters of Rwanda stratum 1 to examine the level of significant differences within the stratum. The results revealed that within the stratum, there are significant differences between the clusters in terms of banana, beans and cassava area, at 95% confidence level. Only there were no significant differences in terms of maize crop area. Furthermore, the pairwise comparison showed that, specifically, there were 31% of clusters pairs with significant differences, while 69% of the rest had no significant differences.

These significant differences, which were higher than any of the sample NDVI classes, might have been the effect of two main aspects: first, Rwanda land classification used one time data (2008 aerial photographs) (MINIRENA & RNRA, 2009) to delineate the strata, which might not only result in bias due to the use of outdated data, but also might not have effectively considered the changes that happened in different areas over time from 2008 till the agricultural survey time (2012-2013). Second, the strata delineation was carried out using visual interpretation technique, which is subjective to the operator's professional experience, guidelines provided and data quality (Baks et al., 2013), whereas the data collected from the field depend on area's natural behaviour and applied farming techniques. Beside the two main effects, the recent agricultural policies in Rwanda might have also played a role, as possible for the NDVI classes. Nevertheless, there was a possibility to improve the homogeneity of the stratum, by using data according to NDVI stratification.

4.4. Use of NDVI stratification to improve homogeneity within Rwanda stratum 1

The overlay of sample sites by NDVI classes with Rwanda strata revealed that some sample areas had been misclassified by Rwanda stratification. Then, the areas were corrected by classifying them according to the reality as found from the field. First, the sample site 2431 which was initially classified as 82% forest, and 18% was then classified as intensive cropland given that from the field there was no forest at all in the site, and agriculture was dominant. Sample site 2432 was initially classified as 50% forest and 50% cropland, but it contained 66% of forest according to the field data, so, the site was removed from further analysis. Sample sites 2414 and 2422 were initially classified as forest while they contained more than 50% of agriculture according to the ground truth, so, they were then classified as intensive cropland. Sample site 8214 was initially classified as cities and towns, but it contained 98% of agriculture according to the ground truth, so, it was then classified as intensive cropland. Two sample sites: 5423 and 5434, were well initially well classified as tea plantation, and field data also showed that they contained a high amount of tea, so, they were excluded from further analysis because of insufficiency to represent the stratum. Afterwards, the analysis was carried out within Rwanda stratum 1, with 44 correctly classified sample sites.

The analysis showed that, then, among all the clusters pairs within Rwanda stratum in terms of banana, maize, beans and cassava, there were 26% with significant differences, and other 74% had no significant differences, at 95% confidence level. So, the significant differences decreased by 5%, from using NDVI classification to correct Rwanda stratum 1. This improvement might have been a result from two main aspects: first: NDVI classes are defined using data rich in temporal dimension in order to consider temporal variabilities while defining classes. Second, the classification applies unsupervised classification looking at the minimal spectral distance between point data (Ball & Hall, 1965), which does not only reduce possible bias from the operator but also increases the chance of grouping classes according to the maximum homogeneity in terms of vegetation cover.

Nevertheless, though the NDVI classification increased the homogeneity within Rwanda stratum 1, the homogeneity was not maximal. This might have been the effect of, mainly, big size the stratum which required more separate NDVI classes to cover the variabilities within the stratum. The stratum covered a big part of the country from West to East, with 86% coverage of entire national agricultural survey area. However, these parts of the country could not be classified as homogeneous into one class, given the country's differences from West to East, in terms of topography, hydrography, soil types and other natural factors (Clay & Dejaegher, 1987; Twagiramungu, 2006).

4.5. Further factors of dissimilarities between NDVI classes and Rwanda strata

NDVI classes and currently used area frames have aspects in common such as spatial representatives, use of geo-information for land stratification, and representative sampling for a survey. Nevertheless, they also have different aspects which might be the result of differences in homogeneity of the delineated strata. This section provides insights into different aspects of dissimilarities which might be the result of significant homogeneity differences between NDVI classes and Rwanda strata.

- **Types of data used for land stratification**

a. Data types and level of up-to-date

Wigton and Bormann (1978), and Cotter et al. (2010) indicated that the currently used area frame sampling uses any recent geo-data including maps such as satellite imagery, aerial photographs, topographic and/or land use maps to define homogeneous strata. Most of these data including topographic maps and aerial photographs took quite long time to update, mainly due to the cost involved in data production. To exemplify this, after 8 years of data acquisition, Rwanda aerial photographs (MINIRENA & RNRA, 2009) are still in use for area frame sampling for agricultural surveys in the country. The land use maps in Rwanda are updated after 10 years, and the recent ones are from 2010 (RNRA, 2010). Given that agricultural land use is the most rapidly changing land use over time (Roser, 2015), the use of outdated spatial data for agricultural surveys may result in biased sample selection thus erroneous samples representativeness, which shall propagate in data collection and becomes huge in overall agricultural estimates.

Differently, hyper-temporal NDVI land stratification uses remote sensing NDVI data to define homogeneous strata. The NDVI classes are defined based on the amount of greenness present in a specific area over time (NASA, 2015). This indicates that, no matter how the vegetation land cover of an area is rapidly changing, if it changes in a similar pattern throughout a year and repeatedly over time, the area should be considered same class. This makes it an important tool in agricultural surveys, which deals with a rapidly changing environment. Furthermore, most of the NDVI data are available up-to-date. An example is of MODIS NDVI data, where a new image data is available every two days since 1999 (NASA, 2002). Other data include, for example, NOAA-HVRR NDVI data which are available every day since 1981 (Tucker, 2009), but with coarse spatial resolution, which is actually suitable for surveys over a big size, of a continental level for instance. Additionally, most of the hyper-temporal NDVI data are open and of free access.

b. Spatial and temporal richness of the data

Rwanda strata were defined using spatially rich data: aerial photographs with very high spatial resolution (25 cm) (MINIRENA & RNRA, 2009). However, the technique used for the design of the homogeneous strata "visual interpretation" might have introduced new errors depending on personal perception and professional experience of the interpreter, data and provided guidelines (Baks et al., 2013).

Differently, NDVI data are not very spatially rich as they aggregate a certain spatial level into one unit with same reflectance value "a pixel" (Liew, 2001), but they are very rich in temporal dimension. Nonetheless, the spatial richness is further claimed back, as the NDVI data are integrated with other existing land use data with high spatial richness. To exemplify, the present study used MOD13Q1 data which aggregate the spatial entity of 250 m^2 into one pixel. Later, the data were integrated with recent Rwanda land use data from 2010 (RNRA, 2010), in order to identify the amount of every land cover existing in each NDVI class.

- **Land stratification process and sampling**

The land stratification process in currently used area frame sampling in Rwanda, is done through visual interpretation technique. This is not only increasing the level of bias in stratification and sampling design but also might ignore consideration of important natural phenomena, including possible abrupt changes in the natural system. This is exemplified by the area in figure 16; where agricultural activities were cleared in August 2015 in order to build new industrial zone. As the changes cannot be seen on the 2008 aerial images (MINIRENA & RNRA, 2009), the high probability is that the area will continue to be classified as agriculture, whereas there is even no vegetation anymore.

On contrast, NDVI data consider all events occurring in nature according to similarities in areas' vegetation photosynthetic activity (Tucker & Sellers, 1986). The area currently under industrial zone construction in figure 16 as an example was classified as agriculture by the present study because the images used did not include those for 2015; a year when the changes took place. However, if a survey has to take place next year (2016), images data of 2015 should be included, giving a chance for the recent changes to be considered for new NDVI stratification. Hence, the representativeness shall be improved by the use of up-to-date data. Additionally, the land classification process in done with the use of computer system after user's instructions, through unsupervised classification. This reduces the bias that might be subject to the operator, and thus enhance the replicability of the work.

- **Time frame for land stratification validity**

Cotter et al. (2010) indicated that the currently designed area frames can be used for 15 to 20 years without a need for replacement. For the case of Rwanda, the currently designed area frames are valid for 5 years, as informed by staff in charge of national agricultural surveys in Rwanda. However, as proven by the present study, the possible error in spatial representativeness shall accumulate over time.

Contrastingly, in order to incorporate temporal changes, NDVI classes should be updated every year. This will ensure that the recent spatial-temporal heterogeneity (Ali et al., 2012; Turner & Gardner, 1994) is incorporated in new stratification every year.

- **Further possible improvement of homogeneity within NDVI classes**

Homogeneity within NDVI classes can still be improved, considering inter-annual variability in landscape (Ali et al., 2012). Given that Rwanda has three agricultural seasons with different agricultural activities each season (NISR, 2015a, 2015b, 2015c), consideration of seasons differently in classification shall give confidence for the land classification to results in much better homogeneous NDVI classes. In this case, only NDVI data for months making an agricultural season of interest shall be considered while stratification. So, there would be three land stratification per year, in Rwanda. This will ensure that specific heterogeneity existing in an area due to different seasonal agricultural activities are considered, which would lead to more improvement in the homogeneity within the NDVI classes.

5. CONCLUSIONS AND RECOMMENDATIONS

This section presents conclusions from the results and discussion sections by the present study, a summary of answers to the study's research questions, and finally, recommendations for further studies.

❖ Conclusions

From the results as found by the present study, the following conclusions are drawn:

- The NDVI stratification resulted not only in classes different in terms of temporal behaviour, but they were also different in spatial distribution. The four sample NDVI classes selected looking at their remarkable differences in temporal behaviour, were also remarkably different in spatial distribution. The classes were distributed according to natural differences in Rwanda in terms of, mainly, topography and rainfall patterns from East to West and varying from South to North of the country. This showed that the hyper-temporal NDVI stratification provides classes which effectively represent spatial differences of areas; a very important aspect for agricultural surveys.

- The sample NDVI classes were generally significantly different in terms of Rwanda season A main crops: banana, maize, beans and cassava, at 95% confidence level. So, it was proven that the classes deserved to be classified as separate (different strata) for effective representativeness of the crops area in different areas for an agricultural survey.

- The two land stratification methods: one using NDVI and the currently used area frame sampling in Rwanda, have differences but also similarities. Regarding the strata designed by the two methods; the most important difference lies in the homogeneity. NDVI classes were found to have a higher level of homogeneity compared to the current Rwanda strata in terms of Rwanda season A main crops: banana, maize, beans and cassava at 95% confidence level. The present study found that, with the same data collected in sample agricultural areas in Rwanda, there was 15% of significant differences within NDVI classes, whereas there was 31% of significant differences within Rwanda stratum 1.

- The present study also found that: by using NDVI stratification to better redefine areas belonging to current Rwanda strata, the level of significant differences lowered by 5% within Rwanda stratum 1. This showed that NDVI classes may not only be used as independent strata for agricultural surveys in order to maximize different areas' homogeneity but also can be used to improve the homogeneity of the currently used area frames.

❖ Answers to research questions and hypotheses verification

The following section presents summary of answers to the present study's questions, based on the results and discussion in chapter 3 and chapter 4:

Question 1: Are there significant differences in crops area coverage data between NDVI classes?

Answer 1: Yes, the present study found that the sample NDVI classes: class 24, 54, 70 and 82 were generally significantly different in terms of areas covered by Rwanda season A main crops: banana, maize, beans and cassava at 95% confidence level. Looking specifically at differences between pairs of the classes, there appeared to be some pairs that were not significantly different in terms of some crops, but the same pairs were significantly different in terms of other crops. An example is of NDVI class 24 which was not significantly different from NDVI class 54 in terms of banana and maize, but the same classes were significantly different in terms of beans and cassava. So, the classes were generally significantly different, to effectively represent all the four main crops of Rwanda season A in a survey.

Question 2: What is the significance level of differences in crops area coverage data from sample sites within the same NDVI class?

Answer 2: The present study found that, taking an example of the area covered by four main crops in agricultural season A in Rwanda (banana, maize, beans and cassava), there was 85% of the clusters pairs with “no significant differences” within the sample NDVI classes, and the rest 15% had significant differences at 95% confidence level. In other words, there was 85% of significant homogeneity within NDVI classes, considering area covered by Rwanda season A main crops at 95% confidence level. The rest 15% contained significant heterogeneity which required further stratification to reach maximal homogeneity within NDVI classes.

Question 3: What is the significance level of differences in crops area coverage data from sample sites within the same Rwanda stratum?

Answer 3: It was only possible to carry out the analysis within Rwanda stratum 1, due to data insufficiency to represent other strata. The present study found that, in terms of banana, maize, beans and cassava, there was 69% of clusters pairs with “no significant differences” within Rwanda stratum 1, and the rest 31% had significant differences at 95% confidence level. Nevertheless, by correcting Rwanda stratum 1 using NDVI classification the homogeneity increased by 5%, having 74% of clusters pairs with “no significant differences” within the stratum and 26% with significant differences.

Question 4: How is the significance level of differences in crops area coverage data within NDVI classes compared to the significance level of differences in crops area coverage data within Rwanda strata?

Answer 4: Considering Rwanda season A main crops area coverage data, the significant differences were different but smaller for NDVI classes and higher for Rwanda stratum 1, at 95% confidence level. Regarding NDVI classes separately, the smallest significant differences were 8% between clusters within NDVI class 24, and the highest were 17% between clusters within NDVI classes 54, 70 and 82. For Rwanda stratum 1, first, the significant differences were 31% between clusters within the stratum and lowered to 26% after correcting the stratum using NDVI stratification. This indicates that, even though it appeared to remain some degree of heterogeneity within NDVI classes, but the highest significant differences within the classes (17%) was even still lower compared to the significant differences within Rwanda stratum 1.

❖ Recommendations for further studies

The present study investigated and reached conclusions in comparing NDVI classes and the currently used area frames for Rwanda agricultural surveys. The focus in terms of agricultural statistics was limited to area coverage data of Rwanda season A main crops: banana, maize, beans and cassava, and four sample NDVI classes, given time available for the research. In this regard, the following recommendations are suggested for further studies in order to reach the overall and definite conclusions about comparing the two land stratification methods for agricultural surveys:

1. Future researchers may consider comparing the NDVI classes with the currently used areas frames, with consideration of other various agricultural statistics such as crops production and agricultural practices data.

2. Identified best separable NDVI strata for Rwanda were 95, among which 24 were dominated by agriculture. However, only 4 NDVI classes from the 24 were used as samples for the present study. Further researchers may consider other 20 NDVI strata for similar studies, in order to investigate overall effectiveness of the improved land stratification using NDVI for agricultural survey in the entire country where agriculture dominates.

3. Expectedly, there would not have been significant differences in the agricultural data from the same NDVI stratum. For the 15% of significant differences that was identified, there is high chance they were due to abrupt changes in agriculture in Rwanda because of the recent policies for agricultural development. In this respect, further researchers with similar studies may consider social and economic aspects and assess their possible impacts on the designed homogeneous area frames for agricultural surveys.

4. Future researchers with similar studies may consider also the Rwanda stratification while sampling design and samples selection.

5. Future researchers may consider NDVI land stratification using hyper-temporal NDVI data for months specific to an agricultural season of interest, in order to incorporate specific inter-annual variabilities of different agricultural seasons, which would contribute to maximal homogeneity within NDVI class for agricultural surveys.

LIST OF REFERENCES

Ali, A., de Bie, C. A. J. M., Scarrott, R. G., Ha, N. T. T., & Skidmore, A. K. (2012). Comparative Performance Analysis of a Hyper-Temporal NDVI Analysis Approach and a Landscape-Ecological Mapping Approach. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *I-7*(1), 105–110. http://doi.org/10.5194/isprsannals-I-7-105-2012

Baks, W., Janssen, L., Schetselaar, K., & Tolpekin, V. (2013). Image Analysis. In V. Tolpekin & A. Stein (Eds.), *The Core of GIScience: a Systems - based Approach* (ITC Educat). Enschede: The International institute for Geo-Information Science and Earth Observation (ITC).

Ball, G. H., & Hall, D. J. (1965). *ISODATA, A Novel Method of Data Analysis and Pattern Classification*. Menlo Park, California.

Barbosa, H. A., Huete, A. R., & Baethgen, W. E. (2006). A 20-Year Study of NDVI Variability over the Northeast Region of Brazil. *Journal of Arid Environments*, *67*(2), 288–307. http://doi.org/10.1016/j.jaridenv.2006.02.022

Benedetti, R., Bee, M., Espa, G., & Piersimoni, F. (Eds.). (2010). *Agricultural Survey Methods*. *Agricultural Survey Methods*. Chichester, UK: John Wiley & Sons, Ltd. http://doi.org/10.1002/9780470665480

Benedetti, R., & Rossini, P. (1993). On the Use of NDVI profiles as a Tool for Agricultural Statistics: The Case Study of Wheat Yield Estimate and Forecast in Emilia Romagna. *Remote Sensing of Environment*, *45*(3), 311–326. http://doi.org/10.1016/0034-4257(93)90113-C

Bolton, D. K., & Friedl, M. A. (2013). Forecasting Crop Yield using Remotely Sensed Vegetation Indices and Crop Phenology Metrics. *Agricultural and Forest Meteorology*, *173*, 74–84. http://doi.org/10.1016/j.agrformet.2013.01.007

Bovolo, F., & Bruzzone, L. (2007). A Theoretical Framework for Unsupervised Change Detection Based on Change Vector Analysis in the Polar Domain. *IEEE Transactions on Geoscience and Remote Sensing*, *45*(1), 218–236. http://doi.org/10.1109/TGRS.2006.885408

Boyko, E., & Hill, C. (2009). *An Introduction to Agriculture Statistics*. Maputo. Retrieved from http://siteresources.worldbank.org/SCBEXTERNAL/Resources/Introduction_to_Agriculture_Statistics.pdf

Cantore, N. (2013). *The Crop Intensification Program in Rwanda: a Sustainability Analysis*. Kigali.

Clay, D. C., & Dejaegher, Y. M. J. (1987). Agro-ecological Zones: The Development of a Regional Classification Scheme for Rwanda. *Tropicultura*, *5*(4), 153–159. Retrieved from http://www.tropicultura.org/text/v5n4/153.pdf

CLGF. (2011). Rwanda. Retrieved December 22, 2015, from http://www.clgf.org.uk/userfiles/1/file/Rwanda_Local_Government_Profile_2013_CLGF.pdf

Cotter, J., Davies, C., Nealon, J., & Roberts, R. (2010). Area Frame Design for Agricultural Surveys. *Agricultural Survey Methods*, (June), 169–192. http://doi.org/10.1002/9780470665480.ch11

de Bie, C. A. J. M. (2015). Remote Sensing and Geographic Information Systems Analysis Methods to Support Food and Water Security Studies (2014-2016) (Class Lecture Notes). University of Twente, Enschede.

de Bie, C. A. J. M., Khan, M. R., Toxopeus, A. G., Venus, V., & Skidmore, A. K. (2008). Hypertemporal Image Analysis for Crop Mapping and Change Detection. *ISPRS Journal of Photogrammetry and Remote Sensing*. Retrieved from http://www.isprs.org/proceedings/XXXVII/congress/7_pdf/5_WG-VII-5/13.pdf

de Bie, C. A. J. M., Nguyen, T. T. H., Ali, A., Scarrott, R., & Skidmore, A. K. (2012). LaHMa: a Landscape Heterogeneity Mapping Method using Hyper-temporal Datasets. *International Journal of Geographical Information Science*, *26*(11), 2177–2192. http://doi.org/10.1080/13658816.2012.712126

Enrique, M. (2010). *Moderate Resolution Imaging Spectroradiometer (MODIS): Data Products, Data Ordering, and Data Pre-processing*.

Eurostat. (2015). Agricultural Census Introduction. Retrieved July 29, 2015, from http://ec.europa.eu/eurostat/statistics-explained/index.php/Agricultural_census_introduction

Fané, I., Kribes, R., Ndimurwango, P., Nsengiyumva, V., Oyono, C. N., & Wennink, B. (2006). Linking Actors for Potato Production and Marketing. ROPARWA and IMBARAGA Initiatives in North-Western Rwanda. Kigali.

FAO. (1996a). *Multiple Frame Agricultural Surveys* (No. 7). Rome.

FAO. (1996b). *World Food Summit*. Rome.

FAO. (2015a). GIEWS Country Brief: Rwanda. Retrieved February 1, 2016, from http://www.fao.org/giews/countrybrief/country/RWA/pdf/RWA.pdf

FAO. (2015b). Statistics at FAO. Retrieved July 29, 2015, from http://www.fao.org/statistics/en/

FAO. (2016). *Enhancing Early Warning Capabilities and Capacities for Food Safety* (Training H). Rome: FAO.

Farooq, U. (2013). What is Census Method of Data Collection, Advantages and Disadvantages? Retrieved July 29, 2015, from http://www.studylecturenotes.com/social-research-methodology/what-is-census-method-of-data-collection-advantages-disadvantages

Funk, C., & Budde, M. E. (2009). Phenologically-tuned MODIS NDVI-based production anomaly estimates for Zimbabwe. *Remote Sensing of Environment*, *113*(1), 115–125. http://doi.org/10.1016/j.rse.2008.08.015

Gillan, J. (2013). Normalized Difference Vegetation Index. Retrieved August 24, 2015, from http://wiki.landscapetoolbox.org/doku.php/remote_sensing_methods:normalized_difference_vegetation_index

Hamad, A. A. (2010). *The Potential of Hyper-temporal NDVI Data to Assess Vegetation Condition and Grazing Intensity*. University of Twente. Retrieved from http://www.itc.nl/library/papers_2010/msc/nrm/hamad.pdf

Hiron, M., Berg, Å., Eggers, S., Berggren, Å., Josefsson, J., & Pärt, T. (2015). The Relationship of Bird Diversity to Crop and Non-Crop Heterogeneity in Agricultural Landscapes. *Landscape Ecology*, *30*(10), 2001–2013. http://doi.org/10.1007/s10980-015-0226-0

Holben, B. N. (1986). Characteristics of Maximum-value Composite Images from Temporal AVHRR Data. *International Journal of Remote Sensing*, *7*(11), 1417–1434. http://doi.org/10.1080/01431168608948945

Huete, A., Didan, K., & Van Leeuwen, W. (1999). *MODIS Vegetation Index (MOD 13): Algorithm Theoretical Basis Document, Version 3.1*. Arizona.

Hulley, G., Veraverbeke, S., & Hook, S. (2014). Thermal-based techniques for land cover change detection using a new dynamic MODIS multispectral emissivity product (MOD21). *Remote Sensing of Environment*, *140*, 755–765. http://doi.org/10.1016/j.rse.2013.10.014

INTERGRAPH. (2013). *ERDAS Field Guide: Security, Government and Infrastructure*. Huntsville, Alabama. Retrieved from http://www.hexagon-solutions.com.cn/libraries/misc_docs/erdas_fieldguide_pdf_intergraph_brand.sflb.pdf

Kathiresan, A. (2012). *Farm Land Use Consolidation in Rwanda: Assessment from the Perspectives of Agriculture Sector*. Kigali. Retrieved from http://www.minagri.gov.rw/fileadmin/user_upload/documents/agridocs/Farm_Land_Use_Consolidation_in_Rwanda.pdf

Kennel, T. (2008). Area Frame. Encyclopedia of Survey Research Methods. Sage Publications, Inc. http://doi.org/http://dx.doi.org/10.4135/9781412963947

King, M. D., Closs, J., Wharton, S., & Myers, M. (Eds.). (2003). *EOS Data Products Handbook (Volume 1)*. Greenbelt, Maryland: NASA Goddard Space Flight Center.

Kiregyera, B., Megill, D., José, B., & Eding, D. (2007). *National Agricultural Information System in Mozambique*. Maputo.

Kiss, O., Tokody, B., Deák, B., & Moskát, C. (2016). Increased Landscape Heterogeneity Supports the Conservation of European Rollers (Coracias garrulus) in Southern Hungary. *Journal for Nature Conservation*, *29*, 97–104. http://doi.org/10.1016/j.jnc.2015.12.003

Liew, S. C. (2001). Principles of Remote Sensing. Retrieved February 2, 2016, from http://www.crisp.nus.edu.sg/~research/tutorial/rsmain.htm

LP DAAC. (2014). Vegetation Indices 16-Day L3 Global 250m. Retrieved June 23, 2015, from https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mod13q1

Lund Research Ltd. (2013). Measures of Central Tendency. Retrieved February 11, 2016, from https://statistics.laerd.com/statistical-guides/measures-central-tendency-mean-mode-median.php

Lunetta, R. S., Knight, J. F., Ediriwickrema, J., Lyon, J. G., & Worthy, L. D. (2006). Land Cover Change Detection using Multi-Temporal MODIS NDVI Data. *Remote Sensing of Environment*, *105*(2), 142–154. http://doi.org/10.1016/j.rse.2006.06.018

Luo, J., Ying, K., & Bai, J. (2005). Savitzky–Golay Smoothing and Differentiation Filter for Even Number Data. *Signal Processing*, *85*(7), 1429–1434. http://doi.org/10.1016/j.sigpro.2005.02.002

Maningas, R. V., Villagonzalo, S. O., & Macaraig, A. J. T. (2004). *Electronic Information Dissemination through the Farmers' Information and Technology Services (FITS)/Techno Pinoy Program: Bringing Information and Technology Within the Reach of the Farmers*. Laguna. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.202.7673&rep=rep1&type=pdf

Matsushita, B., Yang, W., Chen, J., Onda, Y., & Qiu, G. (2007, November 1). Sensitivity of the Enhanced Vegetation Index (EVI) and Normalized Difference Vegetation Index (NDVI) to Topographic

Effects: A Case Study in High-Density Cypress Forest. *Sensors (Basel, Switzerland)*. Multidisciplinary Digital Publishing Institute (MDPI). Retrieved from /pmc/articles/PMC3965234/?report=abstract

MBA Official. (2015). What is Sampling? What are its Characteristics, Advantages and Disadvantages? Retrieved July 29, 2015, from http://www.mbaofficial.com/mba-courses/research-methodology/what-is-sampling-what-are-its-characteristics-advantages-and-disadvantages/

McCloy, K. R. (2005). *Resource Management Information Systems: Remote Sensing, GIS and Modelling, Second Edition - CRC Press Book*. Florida: CRC Press. Retrieved from https://www.crcpress.com/Resource-Management-Information-Systems-Remote-Sensing-GIS-and-Modelling/McCloy/9780415263405

Memarsadeghi, N., Mount, D. M., Netanyahu, N. S., & Le Moigne, J. (2007). A Fast Implementation of the ISODATA Clustering Algorithm. *International Journal of Computational Geometry & Applications*, *17*(01), 71–103. http://doi.org/10.1142/S0218195907002252

Menza, C., Caldow, C., Jeffrey, C., & Monaco, M. (2008). *Analysis of Sample Frames and Subsampling Methods for Reef Fish Surveys*. Retrieved from http://ccma.nos.noaa.gov/publications/NCCOSTM72.pdf

MINECOFIN, & NISR. (2014). *Fourth Population and Housing Census, Rwanda 2012*. Kigali.

MININFRA. (2009). *Updated Version of the National Human Settlement Policy in Rwanda*. Kigali.

MINIRENA, & RNRA. (2009). Rwanda Aerial Photographs (25*25 cm).

Moore, D. S., McCabe, G. P., & Craig, B. A. (2009). *Introduction to the Practice of Statistics*. (S. Burke & A. Scanlan-Rohrer, Eds.) (Sixth Edit). New York: W.H. Freeman and Company.

NASA. (2002). *MODIS: Moderate-resolution Imaging Spectrometer*. Greenbelt, Maryland. Retrieved from http://modis.gsfc.nasa.gov/about/media/modis_brochure.pdf

NASA. (2015). Measuring Vegetation (NDVI & EVI): Normalized Difference Vegetation Index (NDVI). Retrieved July 29, 2015, from http://earthobservatory.nasa.gov/Features/MeasuringVegetation/measuring_vegetation_2.php

Nemani, R. (2014). AVHRR NDVI3g.

Neteler, M. (2011). *Spatio-temporal Reconstruction of Satellite-based Temperature Maps and Their Application to the Prediction of Tick and Mosquito Disease Vector Distribution in Northern Italy*. BoD – Books on Demand. Retrieved from https://books.google.com/books?id=kQC6DLPgLtIC&pgis=1

Nicholson, S. E., Davenport, M. L., & Malo, A. R. (1990). A Comparison of the Vegetation Response to Rainfall in the Sahel and East Africa, using Normalized Difference Vegetation Index from NOAA AVHRR. *Climatic Change*, *17*(2-3), 209–241. http://doi.org/10.1007/BF00138369

NISR. (2012a). *EICV 3 District Profile: East-Bugesera*. Kigali.

NISR. (2012b). *The System of Agricultural Statistics in Rwanda (Improvement Action Plan)*. Kigali. Retrieved from http://statistics.gov.rw/system/files/user_uploads/files/books/SAS_Progress_PapeR_FINAL.pdf

NISR. (2013). *Seasonal Agricultural Survey Report: Rwanda*. Kigali.

NISR. (2015a). *Seasonal Agricultural Survey (SAS) - 2015 Season A*. Kigali.

NISR. (2015b). *Seasonal Agricultural Survey 2013 (Version 2)*. Kigali.

NISR. (2015c). *Seasonal Agricultural Survey 2014*. Kigali.

NIST/SEMATECH. (2013). *Engineering Statistics Handbook*. NISTR/SEMATECH. Retrieved from http://www.itl.nist.gov/div898/handbook/index.htm

RNRA. (2010). Rwanda 2010 Classification Scheme 1. Kigali: Unpublished.

Roser, M. (2015). Land Use in Agriculture. Retrieved July 29, 2015, from http://ourworldindata.org/data/food-agriculture/land-use-in-agriculture/#note-3

Rouse, J. W. J. (1974). Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation. Retrieved from http://ntrs.nasa.gov/search.jsp?R=19740022555

Sisilana, K. P. (2008). *Area Frame Sampling : A Comparison for Crop Area Estimation of Frames Based on Corine Land Cover Map versus Frames Derived from SPOT - NDVI Images : A case Study for Major Crops Grown in Andalucia, Spain (MSc Thesis)*. Enschede.

SUPARCO. (2015). Pakistan Agriculture Information System. Retrieved July 29, 2015, from http://dwms.fao.org/~test/geo_portal_en.asp

The IDB Project. (2015). Index Database (IDB): A Database for Remote Sensing. Retrieved August 24, 2015, from http://www.indexdatabase.de/db/i.php

The Republic of Pakistan. (2008). *Brief on Use of Administrative Data for Statistical Purposes*. Pakistan. Retrieved from unstats.un.org/unsd/dnss/docViewer.aspx?docID=2443

Tuanmu, M.-N., & Jetz, W. (2015). A global, Remote Sensing-based Characterization of Terrestrial Habitat Heterogeneity for Biodiversity and Ecosystem Modelling. *Global Ecology and Biogeography*, *24*, n/a–n/a. http://doi.org/10.1111/geb.12365

Tucker, C. (2009). NOAA Advanced Very High Resolution Radiometer NDVI Data Set Comparison. Washington, D.C.: NASA/GSFC.

Tucker, C. J., & Sellers, P. J. (1986). Satellite Remote Sensing of Primary Production. *International Journal of Remote Sensing*, *7*(11), 1395–1416. http://doi.org/10.1080/01431168608948944

Turner, M. G., & Gardner, R. H. (1994). *Quantitative Methods in Landscape Ecology: The Analysis and Interpretation of Landscape Heterogeneity*. Retrieved from https://books.google.nl/books/about/Quantitative_Methods_in_Landscape_Ecolog.html?id=8AuMildBNvIC&pgis=1

Twagiramungu, F. (2006). *Environmental Profile of Rwanda*. Kigali. Retrieved from http://www.vub.ac.be/klimostoolkit/sites/default/files/documents/rwanda-environmental-profile.pdf

University of Wisconsin-Madison. (2015). MODIS Today. Retrieved August 6, 2015, from http://ge.ssec.wisc.edu/modis-today/index.php?viewcal=true&satellite=t1

Väisänen, P. (2009). *Integration of Registers and Survey-based Data in the Production of Agricultural and Forestry Economics Statistics*. Retrieved from http://ec.europa.eu/eurostat/documents/1001617/4398397/S7P4-NEW-INTEGRATION-OF-REGISTERS-AND-SURVEY-BASED-DATA.pdf

Vidanapathirana, N. P. (2012). Agricultural Information Systems and their Applications for Development of Agriculture and Rural Community, a Review Study. Retrieved July 29, 2015, from https://www.academia.edu/2980504/Agricultural_information_systems_and_their_applications_for_development_of_agriculture_and_rural_community_a_review_study

Walker, P. A., & Mallawaarachchi, T. (1998). Disaggregating Agricultural Statistics Using NOAA-AVHRR NDVI. *Remote Sensing of Environment*, *63*(2), 112–125. http://doi.org/10.1016/S0034-4257(97)00130-2

Wigton, W. H., & Bormann, P. (1978). *A Guide to Area Sampling Frame Construction Utilizing Satellite Imagery*. New York.

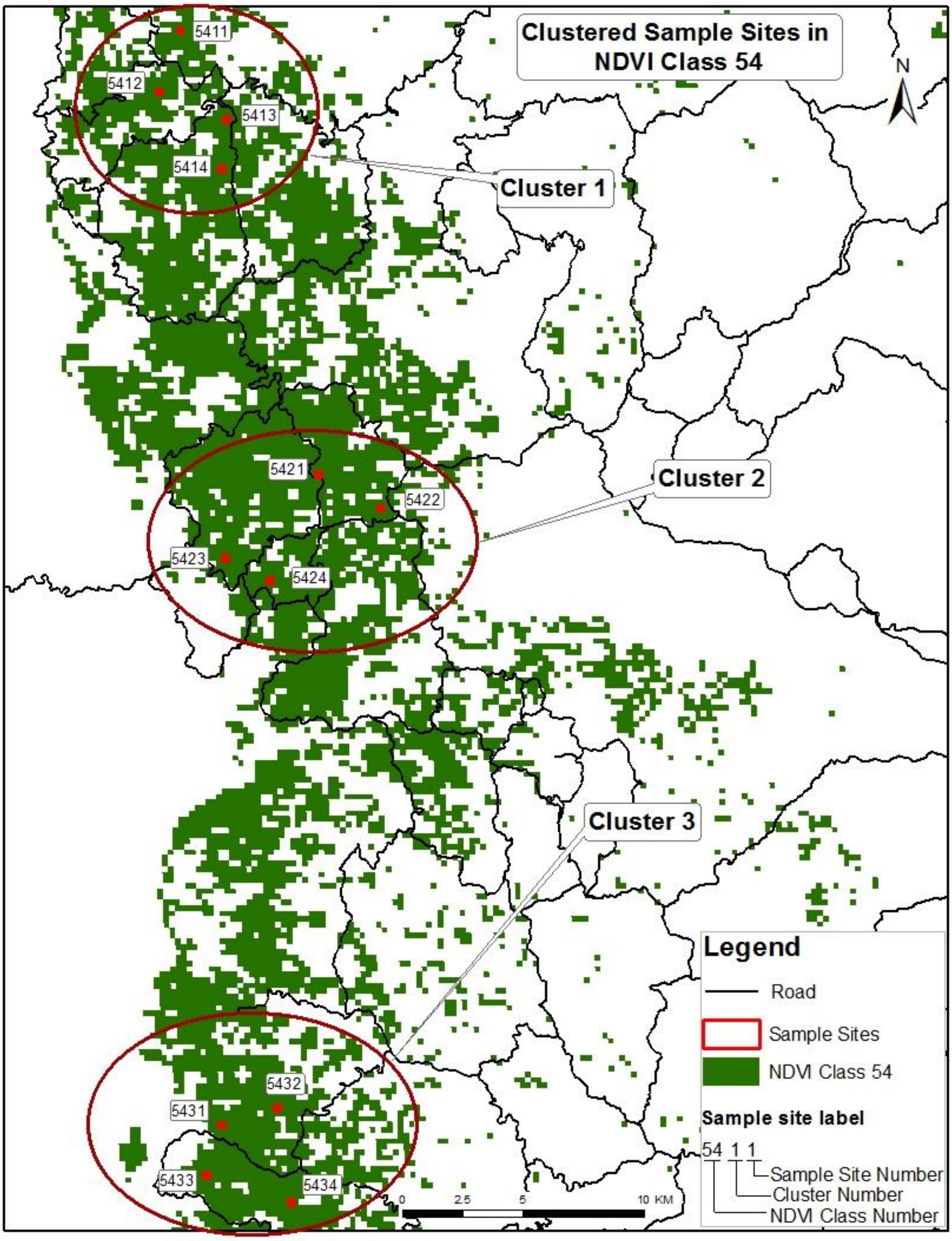
Williams, L. J., & Abdi, H. (2010). Fisher's Least Significant Difference (LSD) Test. *Encyclopedia of Research Design*.

Worldometers. (2015). Current World Population. Retrieved June 1, 2015, from http://www.worldometers.info/world-population/wpc.php?utm_expid=4939992-7.scuhn054Q5WXvFD9uRG9Xw.2&utm_referrer=http://www.worldometers.info/wm2c.php

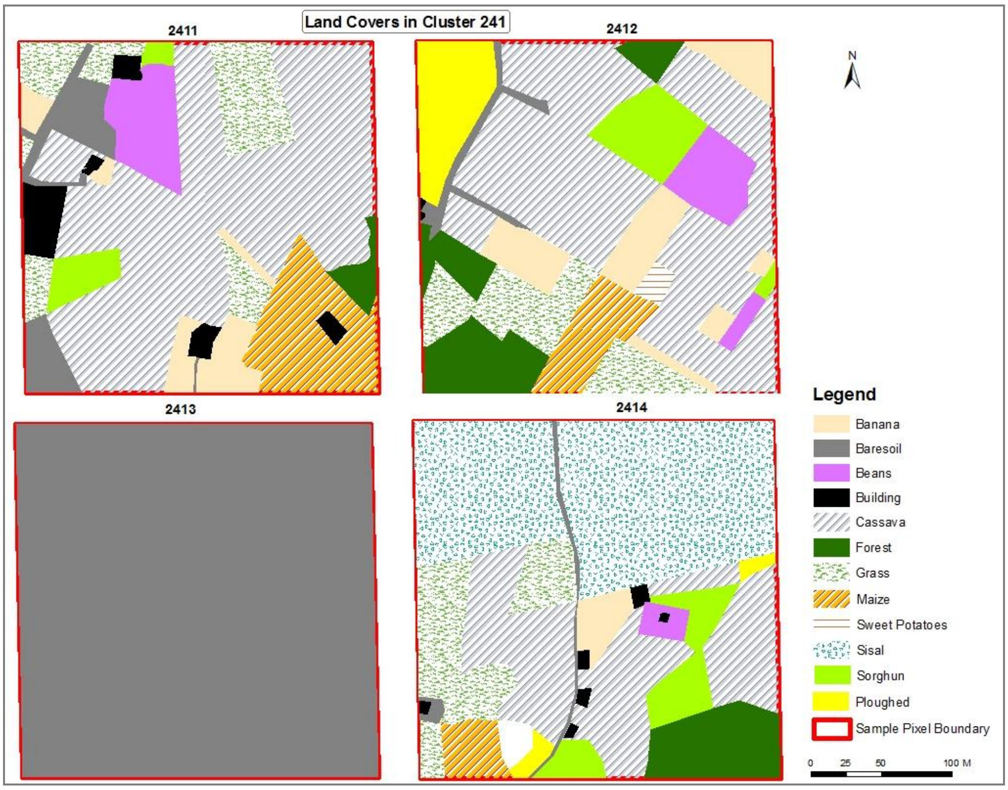
Yuan, Y., Lv, H., & Lu, X. (2015). Semi-supervised change detection method for multi-temporal hyperspectral images. *Neurocomputing*, *148*, 363–375. http://doi.org/10.1016/j.neucom.2014.06.024

APPENDICES

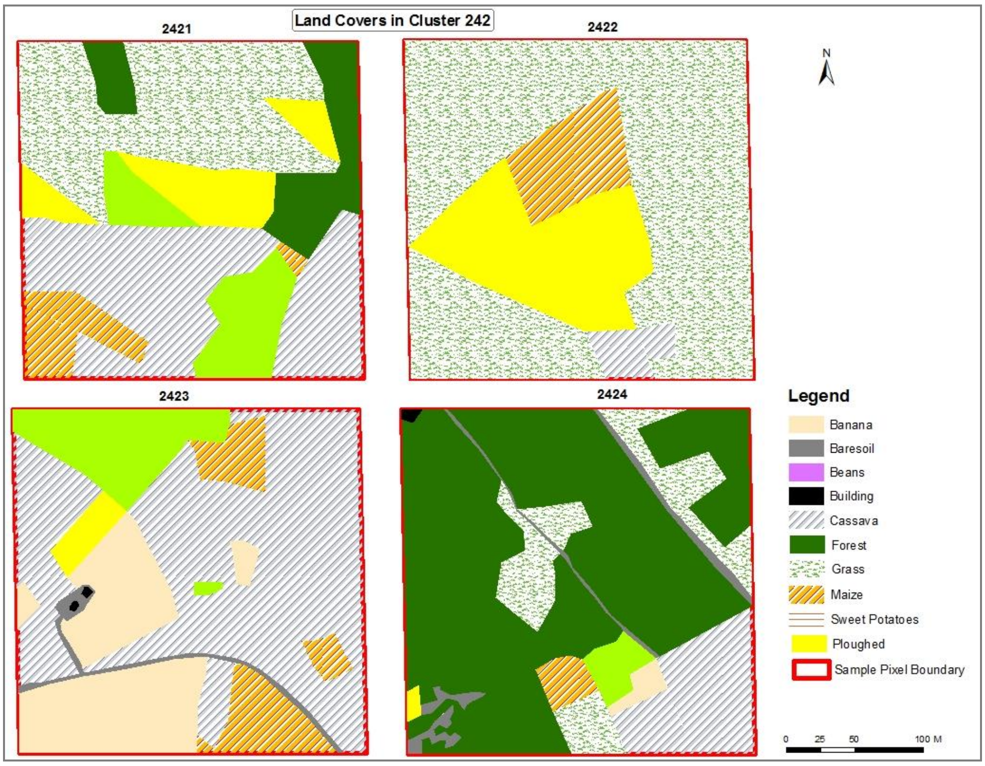
Appendix 1: Sample sites clusters per NDVI class (Example of NDVI class 54)



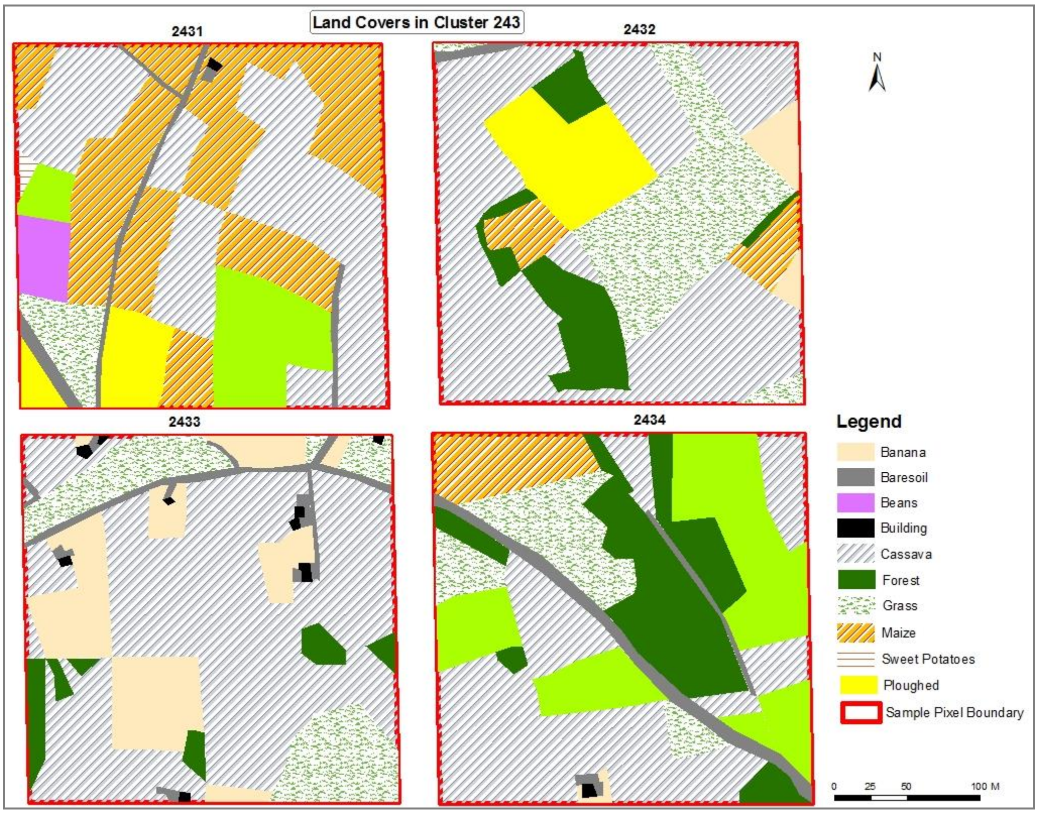
Appendix 2: Land covers in cluster 1 of NDVI class 24



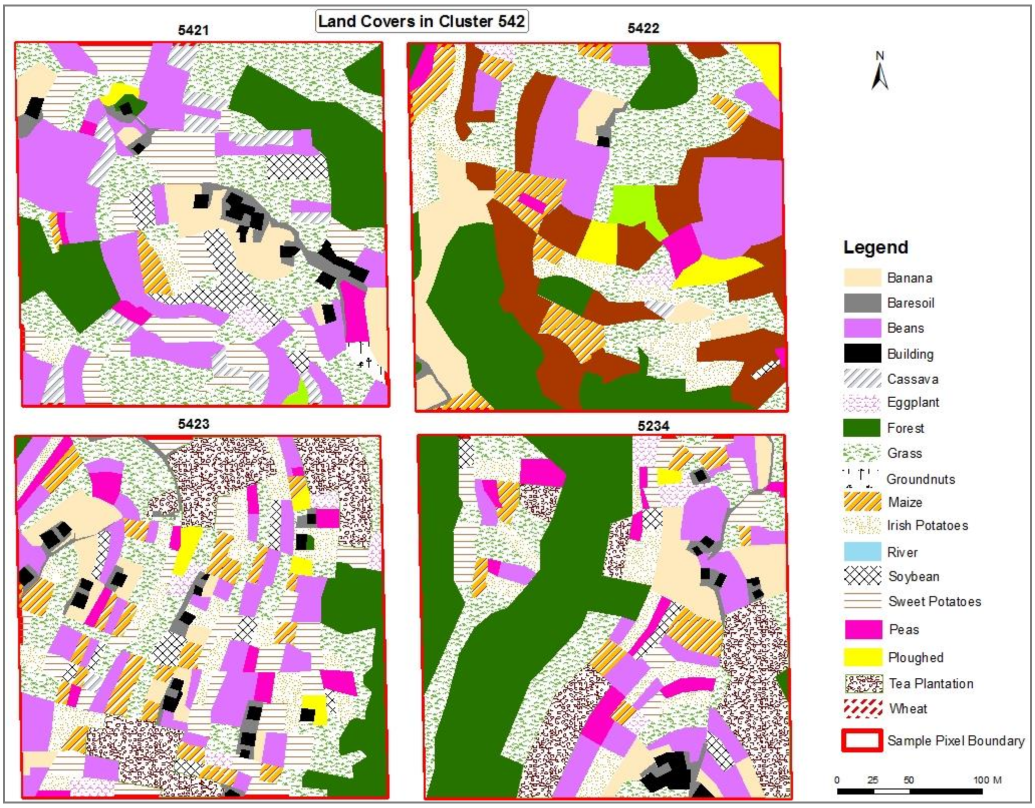
Appendix 3: Land covers in cluster 2 of NDVI class 24



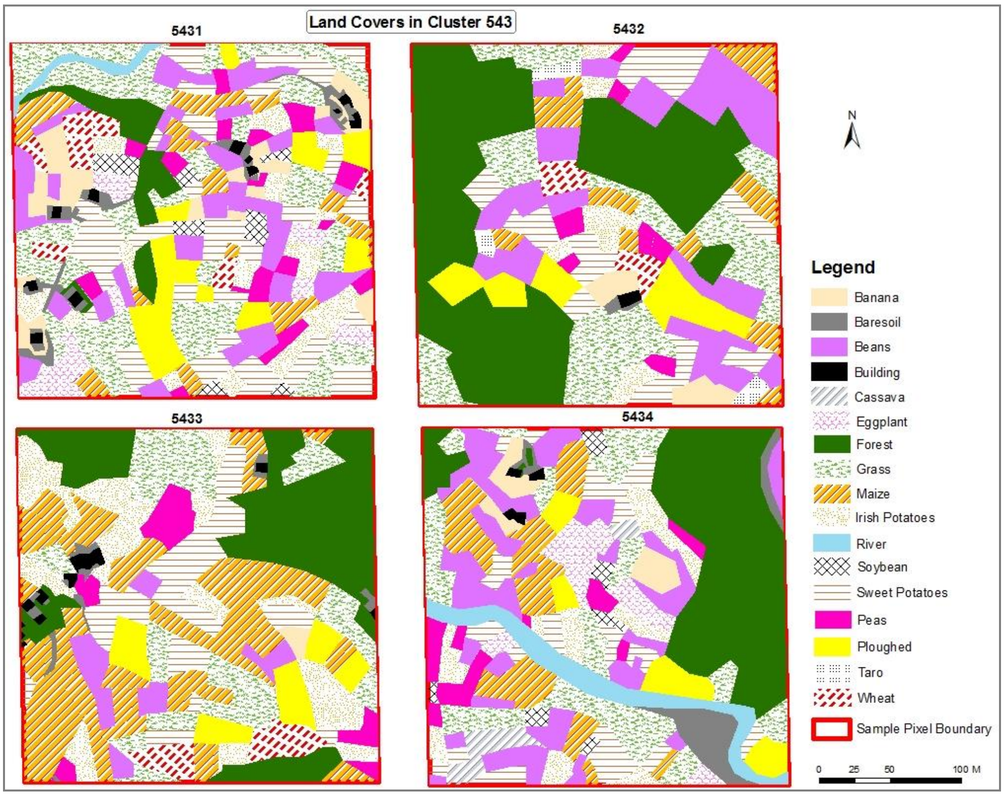
Appendix 4: Land covers in cluster 3 of NDVI class 24



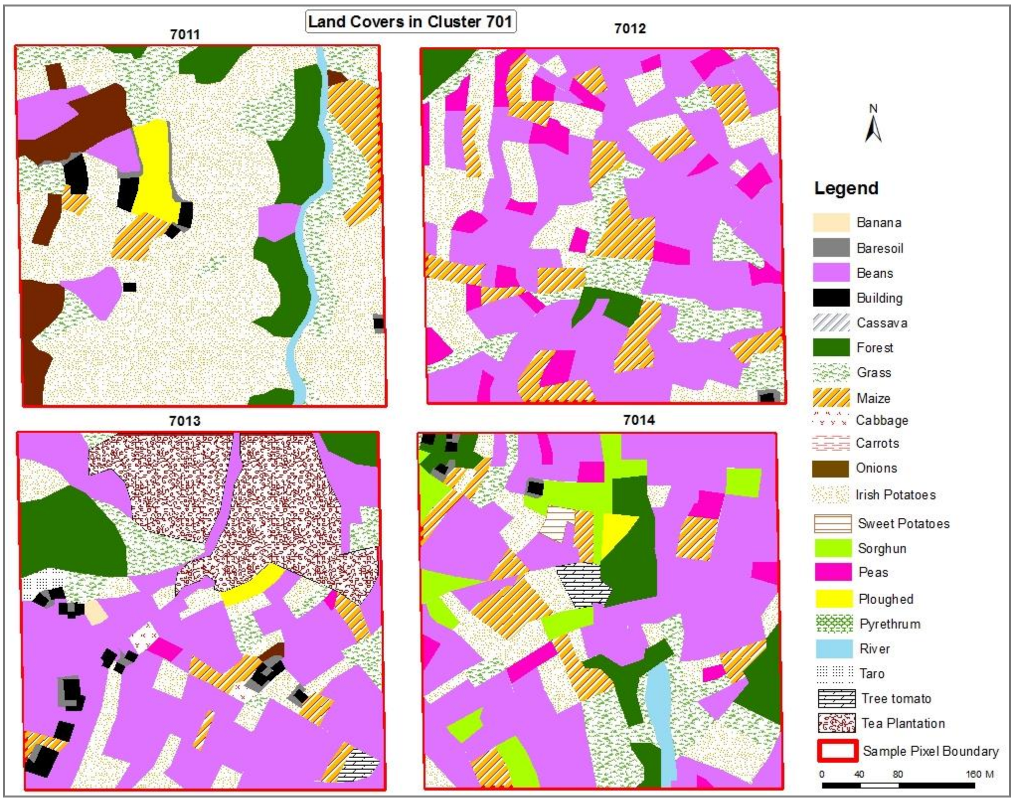
Appendix 5: Land covers in cluster 2 of NDVI class 54



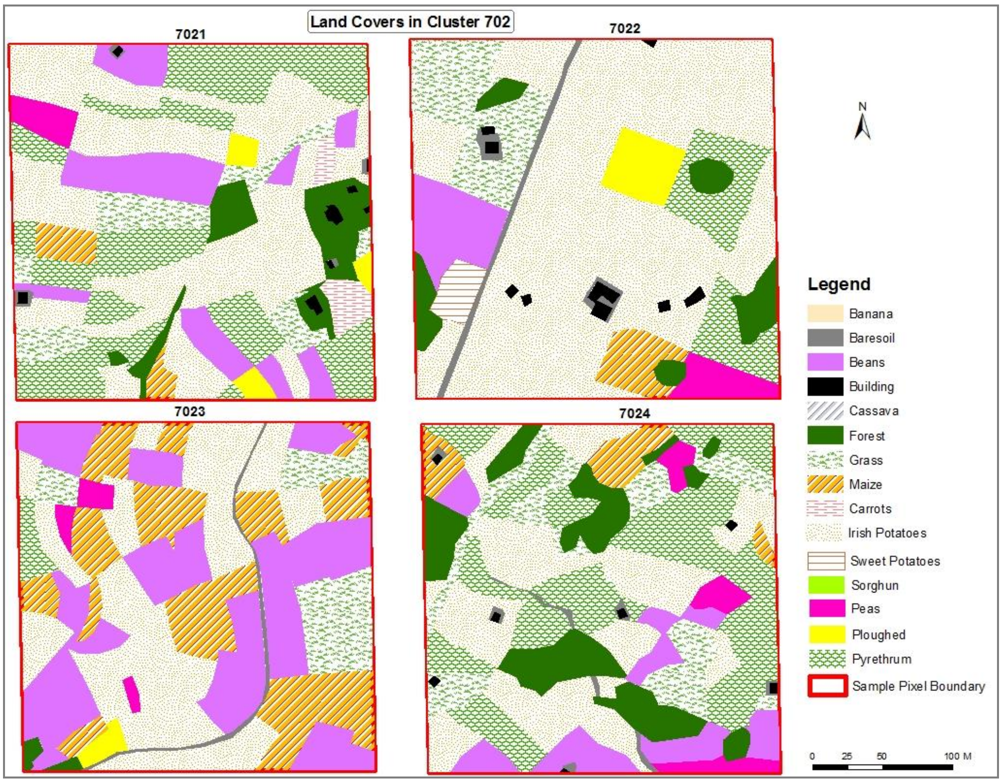
Appendix 6: Land covers in cluster 3 of NDVI class 54



Appendix 7: Land covers in cluster 1 of NDVI class 70

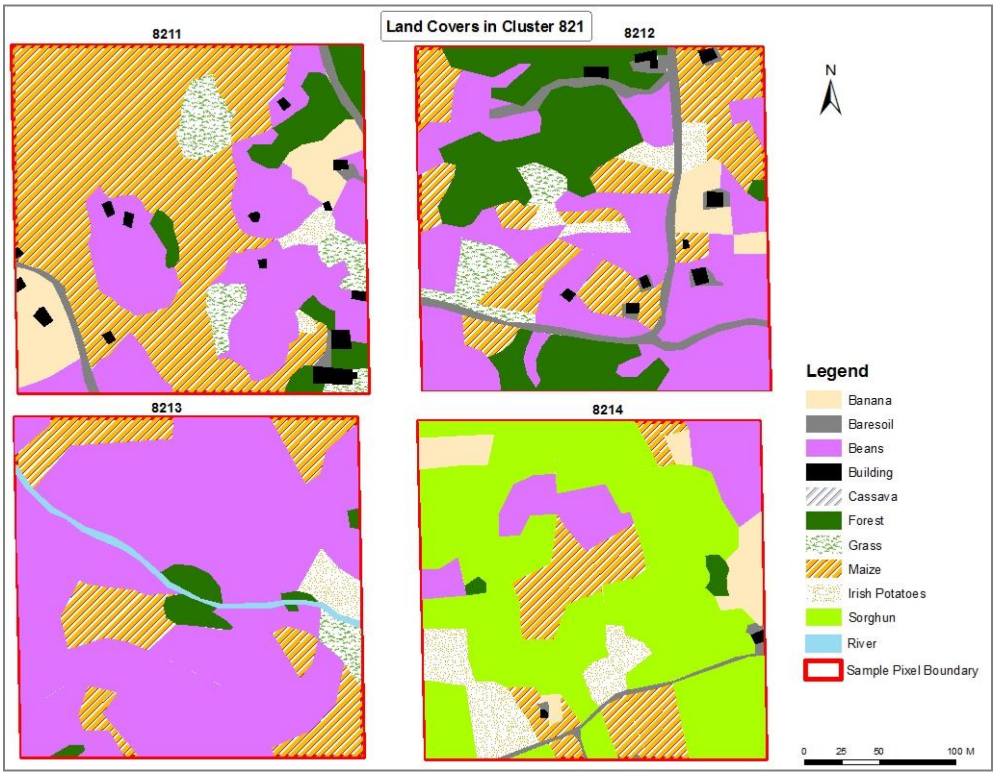


Appendix 8: Land covers in cluster 2 of NDVI class 70

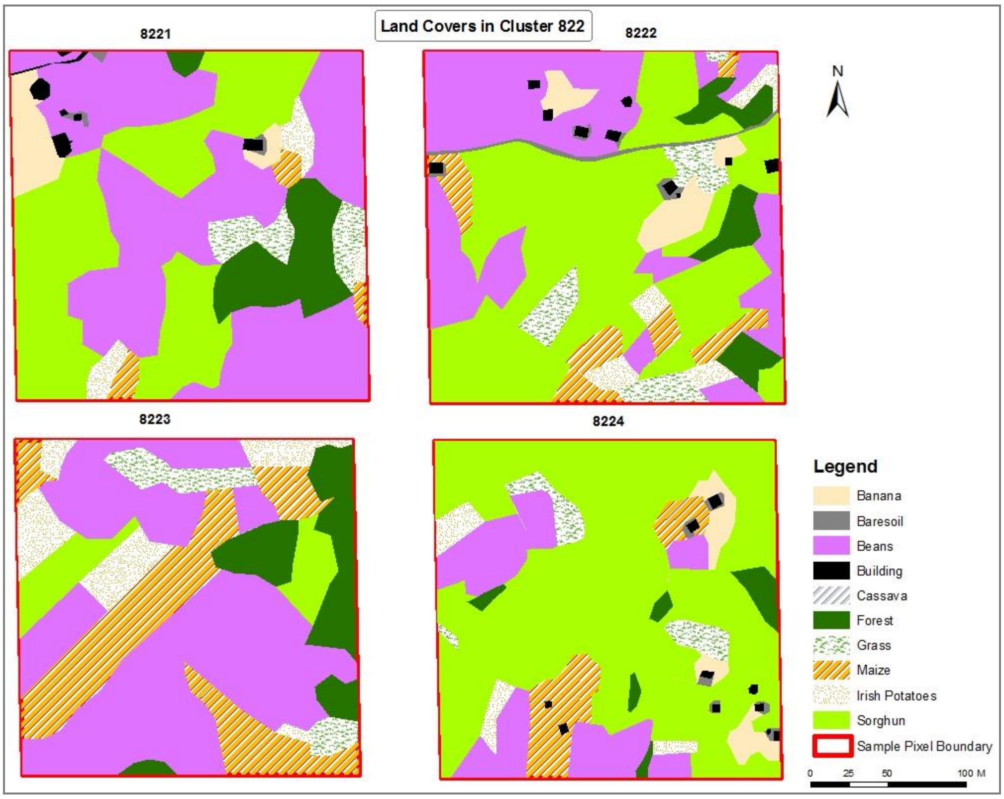


Appendix 9: Land covers in cluster 3 of NDVI class 70

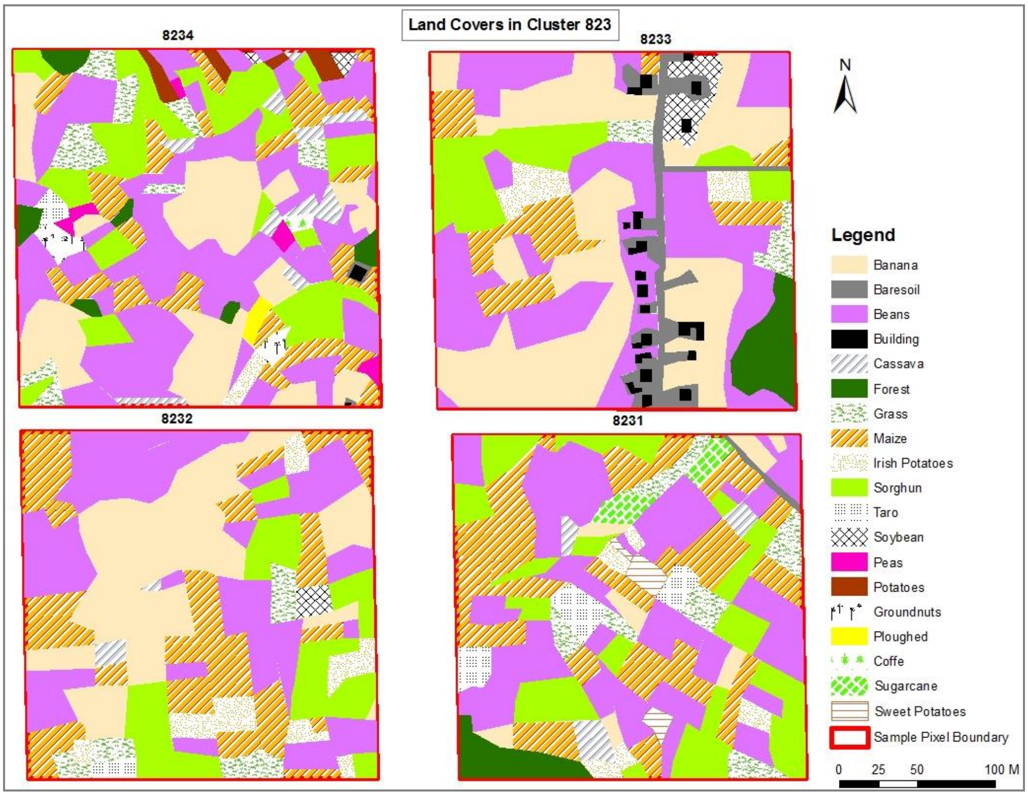
Appendix 10: Land covers in cluster 1 of NDVI class 82



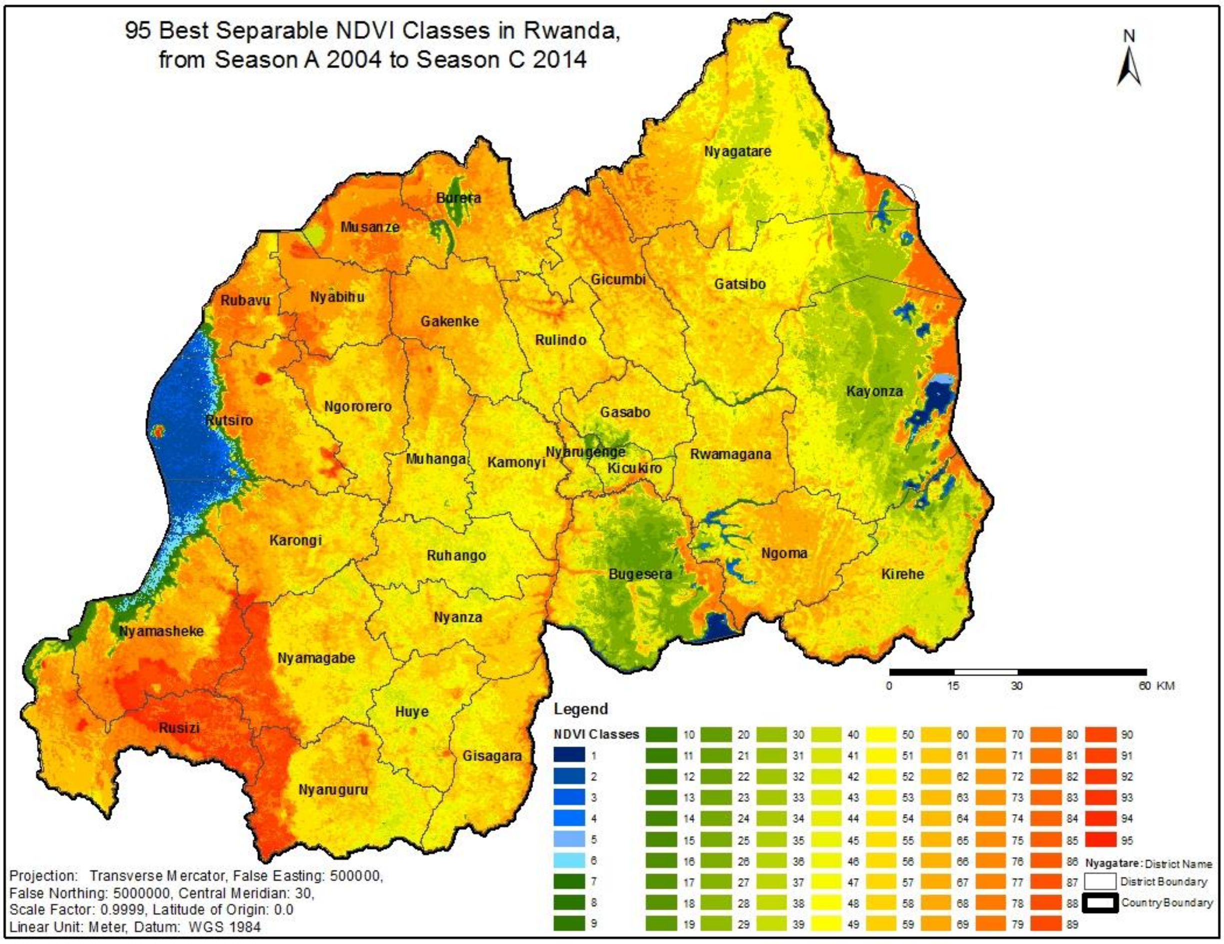
Appendix 11: Land covers in cluster 2 of NDVI class 82



Appendix 12: Land covers in cluster 3 of NDVI class 82



Appendix 13: 95 best separable NDVI classes in Rwanda, from Season A 2004 to Season C 2014



Appendix 14: Agricultural land area (ha and %) per each sample site, according to the field data

Sample site ID	Agricultural land area (ha)	Agricultural land (%)	Non-agricultural land area (ha)	Non-agricultural land area (%)
2411	4.25	79.58	1.09	20.42
2412	4.33	81.17	1.01	18.83
2414	4.44	83.09	0.90	16.91
2421	3.59	67.26	1.75	32.74
2422	2.79	52.27	2.55	47.73
2423	3.99	74.74	1.35	25.26
2424	4.32	80.87	1.02	19.13
2431	5.22	97.70	0.12	2.30
2432	**1.12**	**21.00**	**4.22**	**79.00**
2433	4.99	93.43	0.35	6.57
2434	3.62	67.80	1.72	32.20
5411	4.61	86.37	0.73	13.63
5412	3.96	74.24	1.38	25.76
5413	4.05	75.84	1.29	24.16
5414	3.92	73.32	1.42	26.68
5421	4.33	81.14	1.01	18.86
5422	4.20	78.74	1.14	21.26
5423	4.44	83.08	0.90	16.92
5424	3.36	62.92	1.98	37.08
5431	4.50	84.32	0.84	15.68
5432	3.13	58.68	2.21	41.32
5433	3.91	73.18	1.43	26.82
5434	3.42	64.02	1.92	35.98
7011	4.43	82.90	0.91	17.10
7012	5.02	94.02	0.32	5.98
7013	4.69	87.78	0.65	12.22
7014	4.52	84.72	0.82	15.28
7021	4.79	89.69	0.55	10.31
7022	4.72	88.40	0.62	11.60
7023	5.09	95.33	0.25	4.67
7024	4.35	81.45	0.99	18.55
7031	4.81	90.06	0.53	9.94
7032	4.92	92.10	0.42	7.90
7033	5.03	94.29	0.31	5.71
7034	4.93	92.40	0.41	7.60
8211	4.65	87.13	0.69	12.87
8212	3.58	67.04	1.76	32.96
8213	5.05	94.51	0.29	5.49
8214	5.21	97.56	0.13	2.44
8221	4.73	88.56	0.61	11.44
8222	4.84	90.71	0.50	9.29
8223	4.66	87.33	0.68	12.67
8224	5.09	95.31	0.25	4.69
8231	5.04	94.37	0.30	5.63
8232	5.28	98.88	0.06	1.12
8233	4.62	86.60	0.72	13.40
8234	5.11	95.76	0.23	4.24
Total	**205.67**	**81.95 (Average %)**	**45.31**	**18.05 (Average %)**

Appendix 15: Pairwise analysis by Fisher's LSD, for differences in banana area in Rwanda stratum 1

Banana in Rwanda stratum 1										
Clusters pairs	Fisher's LSD	Absolute value of clusters means difference	Results	Significantly different		Clusters pairs	Fisher's LSD	Absolute value of clusters means difference	Results	Significantly different
(1 ; 2)	0.56	0.07	0.07 < 0.56	No		(4 ; 8)	0.43	0.47	0.47 > 0.43	Yes
(1 ; 3)	0.56	0.43	0.43 < 0.56	No		(4 ; 9)	0.43	0.47	0.47 > 0.43	Yes
(1 ; 4)	0.53	0.03	0.03 < 0.53	No		(4 ; 10)	0.47	0.29	0.29 < 0.47	No
(1 ; 5)	0.61	0.40	0.40 < 0.61	No		(4 ; 11)	0.43	0.31	0.31 < 0.43	No
(1 ; 6)	0.53	0.29	0.29 < 0.53	No		(4 ; 12)	0.43	0.63	0.63 > 0.43	Yes
(1 ; 7)	0.53	0.43	0.43 < 0.53	No		(5 ; 6)	0.53	0.68	0.68 > 0.53	Yes
(1 ; 8)	0.53	0.44	0.44 < 0.53	No		(5 ; 7)	0.53	0.83	0.83 > 0.53	Yes
(1 ; 9)	0.53	0.44	0.44 < 0.53	No		(5 ; 8)	0.53	0.83	0.83 > 0.53	Yes
(1 ; 10)	0.56	0.26	0.26 < 0.56	No		(5 ; 9)	0.53	0.83	0.83 > 0.53	Yes
(1 ; 11)	0.53	0.28	0.28 < 0.53	No		(5 ; 10)	0.56	0.65	0.65 > 0.53	Yes
(1 ; 12)	0.53	0.66	0.66 > 0.53	Yes		(5 ; 11)	0.53	0.67	0.67 > 0.53	Yes
(2 ; 3)	0.56	0.36	0.36 < 0.56	No		(5 ; 12)	0.53	0.26	0.26 < 0.53	No
(2 ; 4)	0.47	0.10	0.10 < 0.47	No		(6 ; 7)	0.43	0.14	0.14 < 0.43	No
(2 ; 5)	0.56	0.46	0.46 < 0.56	No		(6 ; 8)	0.43	0.15	0.15 < 0.43	No
(2 ; 6)	0.47	0.22	0.23 < 0.47	No		(6 ; 9)	0.43	0.15	0.15 < 0.43	No
(2 ; 7)	0.47	0.37	0.37 < 0.47	No		(6 ; 10)	0.47	0.03	0.03 < 0.47	No
(2 ; 8)	0.47	0.37	0.37 < 0.47	No		(6 ; 11)	0.43	0.01	0.01 < 0.43	No
(2 ; 9)	0.47	0.37	0.37 < 0.47	No		(6 ; 12)	0.43	0.95	0.95 > 0.43	Yes
(2 ; 10)	0.50	0.19	0.19 < 0.50	No		(7 ; 8)	0.43	0.01	0.01 < 0.43	No
(2 ; 11)	0.47	0.21	0.21 < 0.47	No		(7 ; 9)	0.43	0.01	0.01 < 0.43	No
(2 ; 12)	0.47	0.72	0.72 > 0.47	Yes		(7 ; 10)	0.47	0.17	0.17 < 0.47	No
(3 ; 4)	0.53	0.46	0.46 < 0.53	No		(7 ; 11)	0.43	0.15	0.15 < 0.43	No
(3 ; 5)	0.61	0.82	0.82 > 0.61	Yes		(7 ; 12)	0.43	1.09	1.09 > 0.43	Yes
(3 ; 6)	0.53	0.14	0.14 < 0.53	No		(8 ; 9)	0.43	0.00	0.00 < 0.43	No
(3 ; 7)	0.53	0.01	0.01 < 0.53	No		(8 ; 10)	0.47	0.18	0.18 < 0.47	No
(3 ; 8)	0.53	0.01	0.01 < 0.53	No		(8 ; 11)	0.43	0.16	0.16 < 0.43	No
(3 ; 9)	0.53	0.01	0.01 < 0.53	No		(8 ; 12)	0.43	1.09	1.09 > 0.43	Yes
(3 ; 10)	0.56	0.17	0.17 < 0.56	No		(9 ; 10)	0.47	0.18	0.18 < 0.47	No
(3 ; 11)	0.53	0.15	0.15 < 0.53	No		(9 ; 11)	0.43	0.16	0.16 < 0.43	No
(3 ; 12)	0.53	1.08	1.08 > 0.53	Yes		(9 ; 12)	0.43	1.09	1.09 > 0.43	Yes
(4 ; 5)	0.53	0.36	0.36 < 0.53	No		(10 ; 11)	0.47	0.02	0.02 < 0.47	No
(4 ; 6)	0.43	0.32	0.32 < 0.43	No		(10 ; 12)	0.47	0.92	0.92 > 0.47	Yes
(4 ; 7)	0.43	0.46	0.46 > 0.43	Yes		(10 ; 11)	0.43	0.94	0.94 > 0.43	Yes

Appendix 16: Pairwise analysis by Fisher's LSD, for differences in beans area in Rwanda stratum 1

Beans in Rwanda stratum 1										
Clusters pairs	Fisher's LSD	Absolute value of clusters means difference	Results	Significantly different		Clusters pairs	Fisher's LSD	Absolute value of clusters means difference	Results	Significantly different
(1 ; 2)	1.16	0.29	0.29 < 1.16	No		(4 ; 8)	0.90	0.01	0.01 < 0.90	No
(1 ; 3)	1.27	0.21	0.21 < 1.27	No		(4 ; 9)	0.90	0.35	0.35 < 0.90	No
(1 ; 4)	1.10	0.46	0.46 < 1.10	No		(4 ; 10)	0.97	1.72	1.72 > 0.97	Yes
(1 ; 5)	1.27	0.57	0.57 < 1.27	No		(4 ; 11)	0.90	1.04	1.04 > 0.90	Yes
(1 ; 6)	1.10	0.33	0.33 < 1.10	No		(4 ; 12)	0.90	0.89	0.89 < 0.90	No
(1 ; 7)	1.10	1.60	1.60 > 1.10	Yes		(5 ; 6)	1.10	0.25	0.25 < 1.10	No
(1 ; 8)	1.10	0.47	0.47 < 1.10	No		(5 ; 7)	1.10	1.03	1.03 < 1.10	No
(1 ; 9)	1.10	0.11	0.11 < 1.10	No		(5 ; 8)	1.10	0.11	0.10 < 1.10	No
(1 ; 10)	1.16	2.18	2.18 > 1.16	Yes		(5 ; 9)	1.10	0.46	0.46 < 1.10	No
(1 ; 11)	1.10	1.50	1.50 > 1.10	Yes		(5 ; 10)	1.16	1.61	1.61 > 1.16	Yes
(1 ; 12)	1.10	1.35	1.35 > 1.10	Yes		(5 ; 11)	1.10	0.93	0.93 < 1.10	No
(2 ; 3)	1.16	0.08	0.08 < 1.16	No		(5 ; 12)	1.10	0.78	0.78 < 1.10	No
(2 ; 4)	0.97	0.75	0.75 < 0.97	No		(6 ; 7)	0.90	1.28	1.28 > 0.90	Yes
(2 ; 5)	1.16	0.86	0.86 < 1.16	No		(6 ; 8)	0.90	0.14	0.14 < 0.90	No
(2 ; 6)	0.97	0.62	0.62 < 0.97	No		(6 ; 9)	0.90	0.22	0.22 < 0.90	No
(2 ; 7)	0.97	1.89	1.89 > 0.97	Yes		(6 ; 10)	0.97	1.85	1.85 > 0.97	Yes
(2 ; 8)	0.97	0.76	0.76 < 0.97	No		(6 ; 11)	0.90	1.18	1.18 > 0.90	Yes
(2 ; 9)	0.97	0.40	0.40 < 0.97	No		(6 ; 12)	0.90	1.02	1.02 > 0.90	Yes
(2 ; 10)	1.03	2.47	2.47 > 1.03	Yes		(7 ; 8)	0.90	1.14	1.14 > 0.90	Yes
(2 ; 11)	0.97	1.79	1.79 > 0.97	Yes		(7 ; 9)	0.90	1.50	1.50 > 0.90	Yes
(2 ; 12)	0.97	1.64	1.64 > 0.97	Yes		(7 ; 10)	0.97	0.57	0.57 < 0.97	No
(3 ; 4)	1.10	0.67	0.67 < 1.10	No		(7 ; 11)	0.90	0.10	0.10 < 0.90	No
(3 ; 5)	1.27	0.78	0.78 < 1.27	No		(7 ; 12)	0.90	0.26	0.26 < 0.90	No
(3 ; 6)	1.10	0.54	0.54 < 1.10	No		(8 ; 9)	0.90	0.36	0.38 < 0.90	No
(3 ; 7)	1.10	1.81	1.81 > 1.10	Yes		(8 ; 10)	0.97	1.71	1.71 > 0.97	Yes
(3 ; 8)	1.10	0.68	0.68 < 1.10	No		(8 ; 11)	0.90	1.04	1.04 > 0.90	Yes
(3 ; 9)	1.10	0.32	0.32 < 1.10	No		(8 ; 12)	0.90	0.88	0.88 < 0.90	No
(3 ; 10)	1.10	2.39	2.39 > 1.10	Yes		(9 ; 10)	0.97	2.07	2.07 > 0.97	Yes
(3 ; 11)	1.16	1.71	1.71 > 1.16	Yes		(9 ; 11)	0.90	1.40	1.40 > 0.90	Yes
(3 ; 12)	1.10	1.56	1.56 > 1.10	Yes		(9 ; 12)	0.90	1.24	1.24 > 0.90	Yes
(4 ; 5)	1.10	0.11	0.11 < 1.10	No		(10 ; 11)	0.97	0.67	0.67 < 0.97	No
(4 ; 6)	0.90	0.14	0.14 < 0.90	No		(10 ; 12)	0.97	0.83	0.83 < 0.97	No
(4 ; 7)	0.90	1.14	1.14 > 0.90	Yes		(10 ; 11)	0.90	0.16	0.16 < 0.90	No

Appendix 17: Pairwise analysis by Fisher's LSD, for differences in cassava area in Rwanda stratum 1

Cassava in Rwanda stratum 1										
Clusters pairs	Fisher's LSD	Absolute value of clusters means difference	Results	Significantly different		Clusters pairs	Fisher's LSD	Absolute value of clusters means difference	Results	Significantly different
(1 ; 2)	0.40	0.04	0.04 < 0.40	No		(4 ; 8)	0.31	0.39	0.39 > 0.31	Yes
(1 ; 3)	0.43	0.68	0.68 > 0.43	Yes		(4 ; 9)	0.31	0.39	0.39 > 0.31	Yes
(1 ; 4)	0.38	2.14	2.14 > 0.37	Yes		(4 ; 10)	0.33	0.39	0.39 > 0.31	Yes
(1 ; 5)	0.43	2.50	2.50 > 0.43	Yes		(4 ; 11)	0.31	0.39	0.39 > 0.31	Yes
(1 ; 6)	0.38	2.50	2.50 > 0.38	Yes		(4 ; 12)	0.31	0.30	0.30 < 0.31	No
(1 ; 7)	0.38	2.53	2.53 > 0.37	Yes		(5 ; 6)	0.38	0.00	0.00 < 0.31	No
(1 ; 8)	0.38	2.53	2.53 > 0.38	Yes		(5 ; 7)	0.38	0.03	0.03 < 0.38	No
(1 ; 9)	0.38	2.53	2.53 > 0.38	Yes		(5 ; 8)	0.38	0.03	0.03 < 0.38	No
(1 ; 10)	0.40	2.53	2.53 > 0.40	Yes		(5 ; 9)	0.38	0.03	0.03 < 0.38	No
(1 ; 11)	0.38	2.53	2.52 > 0.40	Yes		(5 ; 10)	0.40	0.03	0.03 < 0.40	No
(1 ; 12)	0.38	2.44	2.44 > 0.38	Yes		(5 ; 11)	0.38	0.03	0.03 < 0.38	No
(2 ; 3)	0.40	0.65	0.65 > 0.40	Yes		(5 ; 12)	0.38	0.06	0.06 < 0.38	No
(2 ; 4)	0.33	2.11	2.11 > 0.33	Yes		(6 ; 7)	0.31	0.03	0.03 < 0.31	No
(2 ; 5)	0.40	2.46	2.46 > 0.40	Yes		(6 ; 8)	0.31	0.03	0.03 < 0.31	No
(2 ; 6)	0.33	2.46	2.46 > 0.33	Yes		(6 ; 9)	0.31	0.03	0.03 < 0.31	No
(2 ; 7)	0.33	2.49	2.46 > 0.33	Yes		(6 ; 10)	0.33	0.03	0.03 < 0.33	No
(2 ; 8)	0.33	2.49	2.49 > 0.33	Yes		(6 ; 11)	0.31	0.03	0.03 < 0.31	No
(2 ; 9)	0.33	2.49	2.49 > 0.33	Yes		(6 ; 12)	0.31	0.06	0.06 < 0.31	No
(2 ; 10)	0.35	2.49	2.49 > 0.35	Yes		(7 ; 8)	0.31	0.00	0.00 < 0.31	No
(2 ; 11)	0.33	2.49	2.49 > 0.33	Yes		(7 ; 9)	0.31	0.00	0.00 < 0.31	No
(2 ; 12)	2.40	0.33	2.40 > 0.33	Yes		(7 ; 10)	0.33	0.00	0.00 < 0.33	No
(3 ; 4)	0.38	1.46	1.46 > 0.38	Yes		(7 ; 11)	0.31	0.00	0.00 < 0.31	No
(3 ; 5)	0.43	1.46	1.46 > 0.43	Yes		(7 ; 12)	0.31	0.09	0.09 < 0.31	No
(3 ; 6)	0.38	1.82	1.82 > 0.38	Yes		(8 ; 9)	0.31	0.00	0.00 < 0.31	No
(3 ; 7)	0.38	1.82	1.82 > 0.38	Yes		(8 ; 10)	0.33	0.00	0.00 < 0.33	No
(3 ; 8)	0.38	1.85	1.85 > 0.38	Yes		(8 ; 11)	0.31	0.00	0.00 < 0.31	No
(3 ; 9)	0.38	1.85	1.85 > 0.38	Yes		(8 ; 12)	0.31	0.09	0.09 < 0.31	No
(3 ; 10)	0.40	1.85	1.85 > 0.40	Yes		(9 ; 10)	0.33	0.00	0.00 < 0.33	No
(3 ; 11)	0.38	1.85	1.85 > 0.38	Yes		(9 ; 11)	0.31	0.00	0.00 < 0.31	No
(3 ; 12)	0.38	1.85	1.85 > 0.38	Yes		(9 ; 12)	0.31	0.09	0.09 < 0.31	No
(4 ; 5)	0.38	0.36	0.36 < 0.38	No		(10 ; 11)	0.33	0.00	0.00 < 0.33	No
(4 ; 6)	0.31	0.36	0.36 > 0.31	No		(10 ; 12)	0.33	0.09	0.09 < 0.33	No
(4 ; 7)	0.31	0.39	0.39 > 0.31	No		(10 ; 11)	0.31	-0.09	0.09 < 0.31	No

Appendix 18: Pairwise analysis by Fisher's LSD, for differences in banana area in corrected Rwanda stratum 1

Banana										
Clusters pairs	Fisher's LSD	Absolute value of clusters means difference	Results	Significantly different		Clusters pairs	Fisher's LSD	Absolute value of clusters means difference	Results	Significantly different
(1 ; 2)	0.52	0.06	0.06 < 0.52	Yes		(4 ; 8)	0.48	0.47	0.47 < 0.48	Yes
(1 ; 3)	0.52	0.06	0.06 < 0.52	Yes		(4 ; 9)	0.48	0.47	0.47 < 0.48	Yes
(1 ; 4)	0.52	0.13	0.13 < 0.52	Yes		(4 ; 10)	0.48	0.26	0.26 < 0.48	Yes
(1 ; 5)	0.62	0.49	0.49 < 0.62	Yes		(4 ; 11)	0.48	0.31	0.31 < 0.48	Yes
(1 ; 6)	0.52	0.19	0.19 < 0.52	Yes		(4 ; 12)	0.48	0.63	0.63 > 0.48	No
(1 ; 7)	0.52	0.34	0.33 < 0.52	Yes		(5 ; 6)	0.59	0.68	0.68 > 0.59	No
(1 ; 8)	0.52	0.34	0.34 < 0.52	Yes		(5 ; 7)	0.59	0.83	0.83 > 0.59	No
(1 ; 9)	0.52	0.34	0.34 < 0.52	Yes		(5 ; 8)	0.59	0.83	0.83 > 0.59	No
(1 ; 10)	0.52	0.13	0.13 < 0.52	Yes		(5 ; 9)	0.59	0.83	0.83 > 0.59	No
(1 ; 11)	0.52	0.18	0.18 < 0.52	Yes		(5 ; 10)	0.59	0.62	0.62 > 0.59	No
(1 ; 12)	0.52	0.75	0.86 > 0.59	No		(5 ; 11)	0.59	0.67	0.67 > 0.59	No
(2 ; 3)	0.52	0.12	0.12 < 0.52	Yes		(5 ; 12)	0.59	0.26	0.26 < 0.59	Yes
(2 ; 4)	0.48	0.19	0.19 < 0.48	Yes		(6 ; 7)	0.48	0.14	0.14 < 0.48	Yes
(2 ; 5)	0.59	0.55	0.55 < 0.59	Yes		(6 ; 8)	0.48	0.15	0.15 < 0.48	Yes
(2 ; 6)	0.48	0.13	0.13 < 0.48	Yes		(6 ; 9)	0.48	0.15	0.15 < 0.48	Yes
(2 ; 7)	0.48	0.27	0.27 < 0.48	Yes		(6 ; 10)	0.48	0.06	0.06 < 0.48	Yes
(2 ; 8)	0.48	0.28	0.28 < 0.48	Yes		(6 ; 11)	0.48	0.01	0.01 < 0.48	Yes
(2 ; 9)	0.48	0.28	0.28 < 0.48	Yes		(6 ; 12)	0.48	0.95	0.95 > 0.48	No
(2 ; 10)	0.48	0.07	0.07 < 0.48	Yes		(7 ; 8)	0.48	0.01	0.01 < 0.48	Yes
(2 ; 11)	0.48	0.12	0.12 < 0.48	Yes		(7 ; 9)	0.48	0.01	0.01 < 0.48	Yes
(2 ; 12)	0.48	0.82	0.82 > 0.48	No		(7 ; 10)	0.48	0.20	0.20 < 0.48	Yes
(3 ; 4)	0.52	0.07	0.07 < 0.52	Yes		(7 ; 11)	0.48	0.15	0.15 < 0.48	Yes
(3 ; 5)	0.62	0.43	0.43 < 0.62	Yes		(7 ; 12)	0.48	1.09	1.09 > 0.48	No
(3 ; 6)	0.52	0.25	0.25 < 0.52	Yes		(8 ; 9)	0.48	0.00	0.00 < 0.48	Yes
(3 ; 7)	0.52	0.40	0.40 < 0.52	Yes		(8 ; 10)	0.48	0.21	0.21 < 0.48	Yes
(3 ; 8)	0.52	0.40	0.40 < 0.52	Yes		(8 ; 11)	0.48	0.16	0.16 < 0.48	Yes
(3 ; 9)	0.52	0.40	0.40 < 0.52	Yes		(8 ; 12)	0.48	1.09	1.09 > 0.48	No
(3 ; 10)	0.52	0.19	0.19 < 0.52	Yes		(9 ; 10)	0.48	0.21	0.21 < 0.48	Yes
(3 ; 11)	0.52	0.24	0.24 < 0.52	Yes		(9 ; 11)	0.48	0.16	0.16 < 0.48	Yes
(3 ; 12)	0.52	0.69	0.69 > 0.52	No		(9 ; 12)	0.48	1.09	1.09 > 0.48	No
(4 ; 5)	0.59	0.36	0.36 < 0.59	Yes		(10 ; 11)	0.48	0.05	0.05 < 0.48	Yes
(4 ; 6)	0.48	0.32	0.32 < 0.48	Yes		(10 ; 12)	0.48	0.89	0.89 > 0.48	No
(4 ; 7)	0.48	0.46	0.46 < 0.48	Yes		(10 ; 11)	0.48	0.94	0.94 > 0.48	No

Appendix 19: Pairwise analysis by Fisher's LSD, for differences in beans area in corrected Rwanda stratum 1

Beans in Rwanda stratum 1										
Clusters pairs	Fisher's LSD	Absolute value of clusters means difference	Results	Significantly different		Clusters pairs	Fisher's LSD	Absolute value of clusters means difference	Results	Significantly different
(1 ; 2)	1.01	0.21	0.21 < 1.01	No		(4 ; 8)	0.94	0.01	0.01 < 0.94	No
(1 ; 3)	1.08	0.16	0.16 < 1.08	No		(4 ; 9)	0.94	0.35	0.35 < 0.94	No
(1 ; 4)	1.01	0.54	0.54 < 1.01	No		(4 ; 10)	0.94	1.24	1.24 > 0.94	Yes
(1 ; 5)	1.21	0.65	0.65 < 1.21	No		(4 ; 11)	0.94	1.04	1.04 > 0.94	Yes
(1 ; 6)	1.01	0.40	0.40 < 1.01	No		(4 ; 12)	0.94	0.89	0.89 < 0.94	No
(1 ; 7)	1.01	1.68	1.68 > 1.01	Yes		(5 ; 6)	1.15	0.25	0.25 < 1.15	No
(1 ; 8)	1.01	0.54	0.54 < 1.01	No		(5 ; 7)	1.15	1.03	1.03 < 1.15	No
(1 ; 9)	1.01	0.18	0.18 < 1.01	No		(5 ; 8)	1.15	0.11	0.11 < 1.15	No
(1 ; 10)	1.01	1.78	1.78 > 1.01	Yes		(5 ; 9)	1.15	0.46	0.46 < 1.15	No
(1 ; 11)	1.01	1.58	1.58 > 1.01	Yes		(5 ; 10)	1.15	1.13	1.13 < 1.15	No
(1 ; 12)	1.01	1.42	1.42 > 1.01	Yes		(5 ; 11)	1.15	0.93	0.93 < 1.15	No
(2 ; 3)	1.01	0.05	0.05 < 1.01	No		(5 ; 12)	1.15	0.78	0.78 < 1.15	No
(2 ; 4)	0.94	0.75	0.75 < 0.94	No		(6 ; 7)	0.94	1.28	1.28 > 0.94	Yes
(2 ; 5)	1.15	0.86	0.86 < 1.14	No		(6 ; 8)	0.94	0.14	0.14 < 0.94	No
(2 ; 6)	0.94	0.62	0.62 < 0.94	No		(6 ; 9)	0.94	0.22	0.22 < 0.94	No
(2 ; 7)	0.94	1.89	1.89 > 0.94	Yes		(6 ; 10)	0.94	1.38	1.38 > 0.94	Yes
(2 ; 8)	0.94	0.76	0.76 < 0.94	No		(6 ; 11)	0.94	1.18	1.18 > 0.94	Yes
(2 ; 9)	0.94	0.40	0.40 < 0.94	No		(6 ; 12)	0.94	1.02	1.02 > 0.94	Yes
(2 ; 10)	0.94	1.99	2.00 > 0.94	Yes		(7 ; 8)	0.94	1.14	1.14 > 0.94	Yes
(2 ; 11)	0.94	1.79	1.79 > 0.93	Yes		(7 ; 9)	0.94	1.50	1.50 > 0.94	Yes
(2 ; 12)	0.94	1.64	1.64 > 0.93	Yes		(7 ; 10)	0.94	0.10	0.10 < 0.94	No
(3 ; 4)	1.01	0.70	0.70 < 1.01	No		(7 ; 11)	0.94	0.10	0.10 < 0.94	No
(3 ; 5)	1.21	0.81	0.81 < 1.21	No		(7 ; 12)	0.94	0.26	0.26 < 0.94	No
(3 ; 6)	1.01	0.56	0.56 < 1.01	No		(8 ; 9)	0.94	0.36	0.36 < 0.94	No
(3 ; 7)	1.01	1.84	1.84 > 1.01	Yes		(8 ; 10)	0.94	1.24	1.24 > 0.94	Yes
(3 ; 8)	1.01	0.70	0.70 < 1.01	No		(8 ; 11)	0.94	1.04	1.04 > 0.94	Yes
(3 ; 9)	1.01	0.34	0.34 < 1.01	No		(8 ; 12)	0.94	0.88	0.88 < 0.94	No
(3 ; 10)	1.01	1.94	1.94 > 1.01	Yes		(9 ; 10)	0.94	1.60	1.60 > 0.94	Yes
(3 ; 11)	1.01	1.74	1.74 > 1.01	Yes		(9 ; 11)	0.94	1.40	1.40 > 0.94	Yes
(3 ; 12)	1.01	1.58	1.58 > 1.01	Yes		(9 ; 12)	0.94	1.24	1.24 > 0.94	Yes
(4 ; 5)	1.15	0.11	0.11 < 1.15	No		(10 ; 11)	0.94	0.20	0.20 < 0.94	No
(4 ; 6)	0.94	0.14	0.14 < 0.94	No		(10 ; 12)	0.94	0.36	0.36 < 0.94	No
(4 ; 7)	0.94	1.14	1.14 > 0.94	Yes		(10 ; 11)	0.94	0.16	0.16 < 0.94	No

Appendix 20: Pairwise analysis by Fisher's LSD, for differences in cassava area in corrected Rwanda stratum 1

Cassava in Rwanda stratum 1										
Clusters pairs	Fisher's LSD	Absolute value of clusters means difference	Results	Significantly different		Clusters pairs	Fisher's LSD	Absolute value of clusters means difference	Results	Significantly different
(1 ; 2)	0.73	0.23	0.23 < 0.73	No		(4 ; 8)	0.67	0.39	0.39 < 0.67	No
(1 ; 3)	0.78	0.06	0.07 < 0.78	No		(4 ; 9)	0.67	0.39	0.39 < 0.67	No
(1 ; 4)	0.73	1.76	1.76 > 0.73	Yes		(4 ; 10)	0.67	0.39	0.39 < 0.67	No
(1 ; 5)	0.87	2.11	2.11 > 0.87	Yes		(4 ; 11)	0.67	0.39	0.39 < 0.67	No
(1 ; 6)	0.73	2.11	2.11 > 0.73	Yes		(4 ; 12)	0.67	0.30	0.30 < 0.67	No
(1 ; 7)	0.73	2.14	2.14 > 0.73	Yes		(5 ; 6)	0.82	0.00	0.00 < 0.82	No
(1 ; 8)	0.73	2.14	2.14 > 0.73	Yes		(5 ; 7)	0.82	0.03	0.03 < 0.82	No
(1 ; 9)	0.73	2.14	2.14 > 0.73	Yes		(5 ; 8)	0.82	0.03	0.03 < 0.82	No
(1 ; 10)	0.73	2.14	2.14 > 0.73	Yes		(5 ; 9)	0.82	0.03	0.03 < 0.82	No
(1 ; 11)	0.73	2.14	2.14 > 0.73	Yes		(5 ; 10)	0.82	0.03	0.03 < 0.82	No
(1 ; 12)	0.73	2.05	2.05 > 0.73	Yes		(5 ; 11)	0.82	0.03	0.03 < 0.82	No
(2 ; 3)	0.73	0.29	0.29 < 0.73	No		(5 ; 12)	0.82	0.06	0.06 < 0.82	No
(2 ; 4)	0.67	1.53	1.53 > 0.67	Yes		(6 ; 7)	0.67	0.03	0.03 < 0.67	No
(2 ; 5)	0.82	1.88	1.88 > 0.82	Yes		(6 ; 8)	0.67	0.03	0.03 < 0.67	No
(2 ; 6)	0.67	1.88	1.88 > 0.67	Yes		(6 ; 9)	0.67	0.03	0.03 < 0.67	No
(2 ; 7)	0.67	1.91	1.91 > 0.67	Yes		(6 ; 10)	0.67	0.03	0.03 < 0.67	No
(2 ; 8)	0.67	1.91	1.91 > 0.67	Yes		(6 ; 11)	0.67	0.03	0.03 < 0.67	No
(2 ; 9)	0.67	1.91	1.91 > 0.67	Yes		(6 ; 12)	0.67	0.06	0.06 < 0.67	No
(2 ; 10)	0.67	1.91	1.91 > 0.67	Yes		(7 ; 8)	0.67	0.00	0.00 < 0.67	No
(2 ; 11)	0.67	1.91	1.91 > 0.67	Yes		(7 ; 9)	0.67	0.00	0.00 < 0.67	No
(2 ; 12)	0.67	1.82	1.82 > 0.67	Yes		(7 ; 10)	0.67	0.00	0.00 < 0.67	No
(3 ; 4)	0.73	1.82	1.82 > 0.72	Yes		(7 ; 11)	0.67	0.00	0.00 < 0.67	No
(3 ; 5)	0.87	1.82	1.82 > 0.87	Yes		(7 ; 12)	0.67	0.09	0.09 < 0.67	No
(3 ; 6)	0.73	2.17	2.17 > 0.72	Yes		(8 ; 9)	0.67	0.00	0.00 < 0.67	No
(3 ; 7)	0.73	2.17	2.17 > 0.72	Yes		(8 ; 10)	0.67	0.00	0.00 < 0.67	No
(3 ; 8)	0.73	2.20	2.20 > 0.72	Yes		(8 ; 11)	0.67	0.00	0.00 < 0.67	No
(3 ; 9)	0.73	2.20	2.20 > 0.72	Yes		(8 ; 12)	0.67	0.09	0.09 < 0.67	No
(3 ; 10)	0.73	2.20	2.20 > 0.72	Yes		(9 ; 10)	0.67	0.00	0.00 < 0.67	No
(3 ; 11)	0.73	2.20	2.20 > 0.72	Yes		(9 ; 11)	0.67	0.00	0.00 < 0.67	No
(3 ; 12)	0.73	2.20	2.20 > 0.72	Yes		(9 ; 12)	0.67	0.09	0.09 < 0.67	No
(4 ; 5)	0.82	0.36	0.36 < 0.82	No		(10 ; 11)	0.67	0.00	0.00 < 0.67	No
(4 ; 6)	0.67	0.36	0.36 < 0.67	No		(10 ; 12)	0.67	0.09	0.09 < 0.67	No
(4 ; 7)	0.67	0.39	0.39 < 0.67	No		(10 ; 11)	0.67	0.09	0.09 < 0.67	No