ARDHI UNIVERSITY

TITLE: **ASSESSMENT OF INFORMAL SETTLEMENT GROWTH IN URBAN AREA USING GEOGRAPHIC INFORMATION SYSTEM AND REMOTE SENSING**

A case study of mwakibete

NAME: MWANJA KENETH S.

REG NO : 25501/T.2020

DECLARATION AND COPYRIGHT

I, Keneth S. Mwanja, affirm that the findings presented in this dissertation are the outcome of my independent research and study. I confirm that this work has not been previously submitted to any other university as part of a diploma, degree, or similar academic qualification.

....

25501/T.2020

Mwanja, keneth S.

Date: …………………………………..

CERTIFICATION

The undersigned certify that they have supervised and proof read the dissertation and recommend for acceptance by The Ardhi University a dissertation document entitled **"Assessment of informal settlement growth in urban area using geographic information system and remote sensing."** In fulfilment of the requirements for the Bachelor of Science Degree in Geographical Information Systems and Remote Sensing.

………………………………………………..

Dr. ATUPELYE KOMBA

(Supervisor)

Date……………………………….

DEDICATION

To my parent my blessed mother Elizabeth and Rose Mwanja, who took their time and energy to raised and educated me since I was a child to become the person I am Today. I am grateful for responsibility, love and everything you have done for me.

You have a power, gratitude and love, and I will never be able to express how much I love you, to my beloved Uncles Dennis Mwanja, and Godfrey Mwanja your presence and support makes me motivated and determined, to my beloved sisters and brothers July, Grace, Jackqueline, Janeth, Ester, Daniel, Herry, Meshack for being there for me when I needed you the most.

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to Almighty God for blessing me with knowledge, good health throughout my study period and the ability to successfully complete my research

I deeply appreciate the support offered by my supervisor; Dr Atupelye Komba whose sincerity, inspiration, guidance, and knowledge have greatly influenced me throughout the entirety of my dissertation

I would like to provide much appreciation to Dr Deus, Mr.Mavura, Mrs. Latifa litwe, Mr. amchau and panel member and the staffs of Geospatial science and Technology

Also, it would not be good if I don't mention my GIS and RS classmates, Charles Byera, Abdurazaq Feisal, kennedy somi, Dennis kumba ,calvin john, Neema Magele and the rest of my colleagues-the company and support offer will not be forgotten, in addition I would like to thank for their assistance, encouragement and support during working in my dissertation

The completion of this research could have been possible without the participation and assistance of many people who might have not been mentioned in this page whose contribution are sincerely appreciated and gratefully acknowledged

ABSTRACT

Mapping informal settlements in urban areas is important for developing and implementing inclusive urban policies and to achieve the millennium sustainable development goals. Informal settlements are often viewed as problematic due to their self-built housing, substandard services, and low resident incomes, which are linked to poverty, marginalization, and irregularity. To improve the resilience of these settlements, it is essential to understand their location, timing, and residents. The aim of this research is to assess the expansion of informal settlement growth in urban area at a case study of Mwakibete in Mbeya city region from 2015 to 2023 by applying random forest algorithm methodology to classify the satellite images and NDBI and NDVI as input data. This study focuses on using machine-learning (ML) methods and satellite imagery from sentinel 2A satellite to map informal areas in Mwakibete, Mbeya. Spectral features which are spectral indices (NDVI and NDBI) and 3 spectral bands from sentinel 2A dataset were derived which were band 4, 8 and 11 and mapped using ML Random Forest approach. The Random forest algorithm produce land cover map of four classes built up, woodland, cropland and bareland with the accuracy assessment of 88.89%, 80.77% and 86.77% respectively, the NDVI and NDBI were use to predict the land cover maps produced as they were inputs for the image classification as band so the random forest algorithm uses the band variable importance to show the importance of the bands to the classification, for NDVI its importance for both years were 42.02%, 50.03% and 71.02% respectively and for NDBI it was 18.02%, 50.09% and 80.5% respectively so from the result NDBI was the best for 2019 and 2023 but not for 2015.

TABLE OF CONTENTS

LIST OF ABBREVIATION

NBS - NATIONAL BUREAU OF STATISTICS

NDBI - NORMALIZED DIFFERENCE BUILT UP INDEX

NDVI - NORMALIZED DIFFERENCE VEGETATION INDEX

RF - RANDOM FOREST

LULC - LAND USE LAND COVER

GEE – GOOGLE EARTH ENGINE

LIST OF TABLE

LIST OF FIGURE

CHAPTER ONE

INTRODUCTION

1.1 Background

Urban growth refers to the expansion of the urban population, regardless of changes in the rural population (Guttenberg, 1960). This growth is closely linked to the process of urbanization, which involves a greater proportion of the overall population residing in cities and their surrounding suburbs (Nabutola, 2011).Urbanization involves an increase in the number of people living in urban areas and is primarily influenced by the natural growth of the population, migration patterns, and the reclassification of neighboring rural regions into urban areas (NBS, 2015).Rapid urbanization and population growth pose significant global development challenges in the twenty-first century (Lintelo et al., 2018). Developing countries experience a surge in population accompanied by economic activities that drive the expansion of construction and land use. Over the past century, the world has witnessed rapid urbanization and the proliferation of slums (Mohammadian et al., 2017).

In Tanzania, cities and towns are experiencing rapid urbanization, driven by a combination of rural-to-urban migration and natural population growth(J. Gwaleba, 2018)Based on census data, the urban population in mainland Tanzania has witnessed a significant increase over the years. In 1967, the urban population accounted for 5.7 percent (685,092 people) of the total population, which rose to 22.6 percent (7.6 million people) in 2002. Moreover, if we consider urban areas based on population density, more than 33 percent of the mainland population, approximately 11 million people, resided in densely populated urban areas in 2002 (“Demogr. Trends Urban.,” 2021). This trend of urbanization is expected to continue at a rapid pace. According to population projections by the United Nations, the percentage of people living in urban areas is projected to increase from 24 percent in 2005 to 38 percent by 2030.The urban population in Tanzania is anticipated to experience a growth rate that is more than twice the rate of the overall population. As a result, it is estimated that by 2030, over 25 million Tanzanians will reside in urban areas.

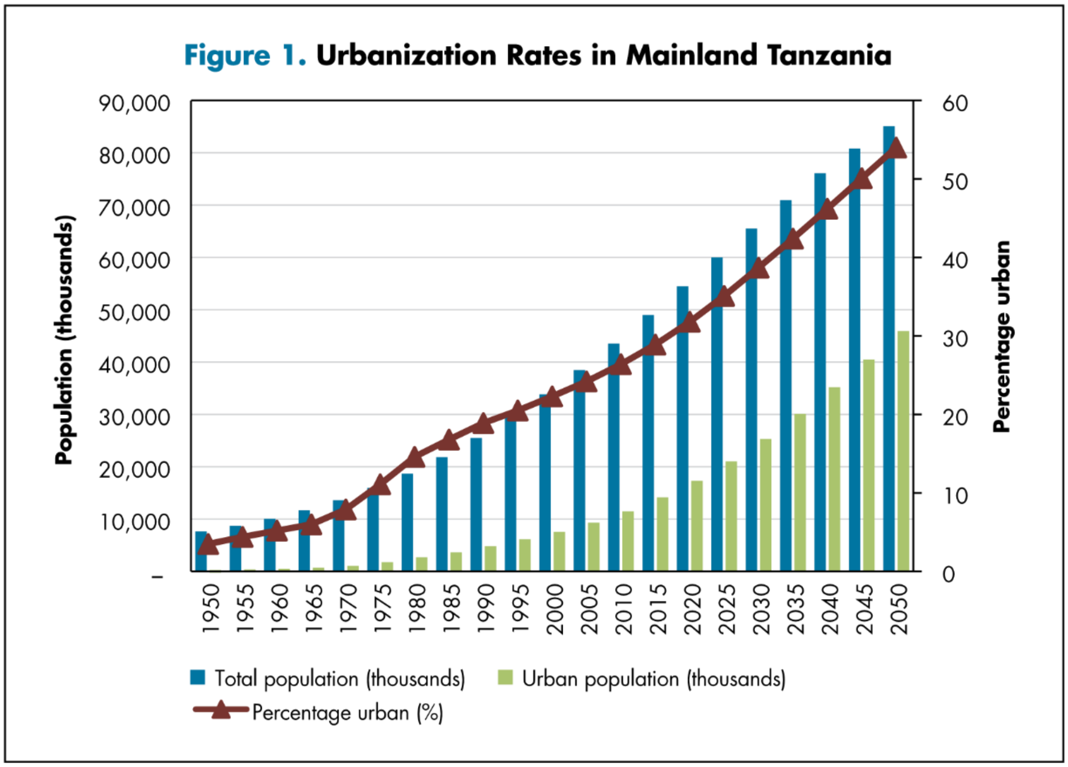


Figure 1. 1 Urbanization Rates in Mainland Tanzania

Informal settlements and urban informality are prevalent issues in third-world countries (Nassar & Elsayed, 2018). Developing countries are characterized by rapid, unstructured, and unplanned development (Asmat & Zamzami, 2012).

"In many regions of the global south, informal settlements and slums have emerged as the dominant form of housing. These settlements arise from the urgent need for shelter, rapid urbanization, and a scarcity of affordable and suitable land and housing options for low-income communities (Kuffer et al., 2016). The residents of informal settlements are particularly vulnerable to natural hazards such as fires and floods. The process of rapid urbanization experienced by numerous countries in the global south in recent decades has presented significant challenges to both the environment and social structures (Gilbert, 2005).

Effective planning for informal settlements must align with the 2030 Agenda for Sustainable Development, as countries are required to report progress towards achieving the Sustainable Development Goals (SDGs) [2]. The 2030 Agenda aims to enhance the living conditions of informal settlement residents by: 1) reducing the proportion of men, women, and children living in multidimensional poverty by at least half, and 2) ensuring access to adequate, safe, and affordable housing, along with basic services, including the upgrading of slums. While there is evidence of a gradual shift in poverty from rural to urban areas, it is evident that more comprehensive information is necessary regarding urbanization patterns, including the spread of informal settlements(United Nations Human Settlement Programme (UN-Habitat), 2003).

Informal settlements are often regarded as problematic due to their self-built nature, inadequate services, and the low incomes of their residents. These settlements are typically associated with poverty, informality, and marginalization, which contribute to challenges such as a lack of secure tenure and limited access to basic infrastructures(Penrose et al., 2010). Consequently, the lives of informal settlers are adversely affected, and urban development is hindered.

1.2 Problem statement

Rapid urbanization has led to the emergence of new types of informal settlements and slums in developing cities. These areas, which are often referred to as contested spaces, face legal and public discourse challenges. The growth of informal settlements is a consequence of urbanization and has resulted in an increase in the number of slum dwellers. According to a report by UN-HABITAT, the number of slum dwellers significantly rose during the 1990s and is projected to reach approximately 2 billion in the next 30 years if no decisive action is taken.

Tanzania, like many other developing countries, is grappling with the issue of informal settlements due to urbanization and population growth. The lack of proper urban land development control by authorities has contributed to the proliferation of informal settlements, particularly among middle- and low-income households. It is estimated that 40 to 75 percent of Tanzania's urban population resides in informal and unplanned settlements, with cities such as Mbeya, Dar es Salaam, Mwanza, and Arusha having significant numbers of such settlements.

The expansion of informal settlements poses challenges to sustainable development efforts. Despite attempts to improve living conditions in urban areas, informal settlements continue to grow rapidly, placing strains on urban planning, infrastructure, and public services. These areas are characterized by high poverty rates, illiteracy, and unemployment. Social problems such as crime, drug addiction, alcoholism, and high rates of mental illness and suicide are prevalent in these informal settlements, as reported by UN-HABITAT in 2016.

The proliferation of informal settlements and slums in developing cities, including Tanzania, presents a pressing concern for sustainable development. The challenges associated with these settlements include inadequate urban planning, insufficient infrastructure, and social problems resulting from high levels of poverty and unemployment. Addressing these issues requires concerted efforts and concrete actions to improve the living conditions and prospects for the residents of informal settlements.

1.3 Research Objectives

1.3.1 Main Objective

To assess the extent of informal settlement expansion in Mbeya city from 2015 to 2023

1.3.2 Specific Objectives

- To use spectral indices as data input for image classification
- To map distribution of informal settlement in Mwakibete from 2015 to 2023

1.4 Research Questions

i. What is the extent and spatial distribution of informal settlements in the urban area?

ii. How has the growth and expansion of informal settlements changed over time?

iii. How accurate random forest algorithm in classification and mapping informal settlements?

1.5 Significance of the research

The research on assessing informal settlement growth in urban areas using Geographic Information System (GIS) and remote sensing holds significant importance as it addresses the challenges posed by rapid urbanization and the growth of informal settlements. By utilizing GIS and remote sensing technologies, the research provides accurate and up-to-date information on the extent, patterns, and dynamics of informal settlements. This data-driven approach enables policymakers and urban planners to make informed decisions, formulate targeted interventions, and allocate resources effectively. The research contributes to sustainable urban development by guiding policy formulation, enhancing disaster risk reduction efforts, and promoting inclusive urban planning. It also advances knowledge in the field of urban studies by providing valuable insights, methodologies, and analytical techniques for researchers and practitioners. Overall, the research has the potential to improve the living conditions of informal settlement residents, address social inequalities, and create equitable and resilient urban environments.

1.6 Beneficiaries of the research

i. Urban planners and policymakers: The research findings can offer valuable insights into the dynamics of informal settlement growth. This information can assist urban planners and policymakers in making well-informed decisions regarding urban development, land use planning, and resource allocation to address the issues related to informal settlements.

ii. Government agencies: Government agencies responsible for urban development, housing, and infrastructure can benefit from the research outcomes. The research can enhance their understanding of the spatial distribution and characteristics of informal settlements, enabling them to formulate appropriate policies and implement targeted interventions to improve living conditions in these settlements.

iii. Researchers and academics: The research can serve as a valuable resource for researchers and academics in the fields of urban studies, remote sensing, and GIS. It can contribute to their understanding of methodologies, techniques, and data sources used in assessing informal settlements, while expanding the existing knowledge on urbanization dynamics and challenges associated with informal settlements.

iv. Community-based organizations and informal settlement residents: The research outcomes can empower community-based organizations working with informal settlement residents by providing them with accurate and current information on settlement growth and characteristics. This knowledge can support their advocacy efforts, community mobilization, and engagement with relevant stakeholders to address the needs and rights of informal settlement residents.

1.7 Scope and limitations of the research

The research focuses on using GIS and remote sensing to assess informal settlement growth in urban areas. The scope includes analyzing spatial and temporal changes in informal settlements, using geospatial data like satellite imagery. By utilizing GIS and remote sensing technologies, the research provides accurate and up-to-date information on the extent, patterns, and dynamics of informal settlements. This data-driven approach enables policymakers and urban planners to make informed decisions, formulate targeted interventions, and allocate resources effectively. The research contributes to sustainable urban development by guiding policy formulation, enhancing disaster risk reduction efforts, and promoting inclusive urban planning. It also advances knowledge in the field of urban studies by providing valuable insights, methodologies, and analytical techniques for researchers and practitioners. The research aims to provide valuable insights for urban planning and policy-making. However, limitations exist, such as data availability and quality, potential classification inaccuracies, and the need for specialized expertise. Despite these challenges, the research contributes to understanding and managing informal settlements, offering valuable information for urban development.

CHAPTER TWO

LITERATURE REVIEW

2.1 Overview

This chapter intends to has reviewed past findings, experience and perceptions revealed by various researches on similar topic.

Rapid urbanization and population growth are among the key global challenges in the 21st century (Lintelo et al., 2018), particularly in developing countries where economic activities drive land expansion for construction (Guan et al., 2011). The proliferation of slums and informal settlements has been a common trend in the 20th and 21st centuries (Mohammadian et al., 2017) with informality being a major issue in third-world countries (Nassar & Elsayed, 2018). Developing countries often experience rapid, unstructured, and unplanned development (Asmat & Zamzami, 2012) and the combination of rapid population growth and unemployment in rural areas has led to the formation of informal settlements in major cities, with more than half of the population residing in such settlements(Payne, 2005).

According to the United Nations (UN), people living in settlements that fulfill any of the following conditions are considered to be residing in informal settlements (Urban & Edition, 2013):

- The residents do not have legal ownership or a secure right to use the land or housing they occupy, and this can range from squatting to living in informal rental housing.
- These neighborhoods typically lack access to basic services and urban infrastructure, or are physically isolated from them.
- The housing structures may not comply with current building and planning regulations, and are often located in areas that pose environmental or geographical risks.

Around one billion people, which is one-seventh of the world's population and one-third of the world's urban population, reside in informal settlements (Urban & Edition, 2013). Projections indicate that by 2050, approximately 66 percent of the global population will be living in urban areas, up from 54 percent in 2014, due to the current rate of urban population growth (Abunyewah et al., 2018). Studies suggest that this growth is happening mainly in small and medium-sized cities (West & Orr, 2007), where informal settlements are expanding in areas that are prone to hazards such as floodplains, valleys, marshlands, and watercourses (Abunyewah et al., 2018). As one of the factor, urbanization has great impact to the expansion of informal settlements to many developing countries like Tanzania

2.2 Informal settlement

Informal settlements are often viewed as problematic due to their self-built housing, substandard services, and low resident incomes, which are linked to poverty, marginalization, and irregularity (Lombard, 2014). The lack of affordable housing and economic opportunities drives the formation of informal settlements, where residents often lack secure tenure and access to basic infrastructure (Penrose et al., 2010), leading to adverse impacts on their quality of life and urban development. Development of unplanned informal settlements poses negative effect to the community and environment at all for example increase in poverty level, crimes, poor sanitation conditions (Chadchan & Shankar, 2012).The management of urban growth and development is a highly intricate task due to its formidable challenges, not only in terms of utilities (Banda & Mwale, 2018), but also in its far-reaching impact on the surrounding environment, natural resources, health conditions, social cohesion, and individual rights (Abebe et al., 2019). Numerous studies have been conducted to examine the effects of informal settlements, primarily focusing on how urban informality influences residents and local development. The conclusions drawn from previous studies (Abunyewah et al., 2018; Lejano and Bianco, 2018; te Lintelo et al., 2018; Guan et al., 2011; Payne, 2005) suggest that the occurrence of informal urban settlements is primarily driven by factors such as rapid urbanization, population growth, economic status, and unplanned development activities.

To control growth of informal settlements or slums Tanzania established clearance programe aimed at solving the problem of informal settlements growth this was in early 1960s (Kyessi & Sekiete, 2014). The project was implemented so as to clear urban squatters and to improve the livelihood of the poor people living in the informal area, to do so the Government of Tanzania decided to formulate National Housing Corporation [NHC] under National Housing Corporation Act (1962).The implementation of this project was unsuccessful and ultimately abandoned due to the significant socio-economic costs involved, rendering it unsustainable (World Bank, 2002).

After the failure of the clearance project the government adopted new project emphasize more on squatter upgrading than clearance in the 1970s. This project was originally devised as a nationwide strategy to address the growth of unplanned and informal settlements by providing services and infrastructure. Implemented during the Ujamaa era (the 1970s-1990s), it experienced a surge in squatters and the emergence of new informal settlements. Consequently, the previously established services and infrastructure began to deteriorate, and the authorities struggled to keep up with the rapid pace of urbanization(Magina et al., 2020).In the 2000s, several initiatives were undertaken, including the implementation of the Community Infrastructure Programme (CIUP) and the Regularization and Formalization of Informal Settlements. The CIUP was a component of the Local Government Support Program (LDSP), which focused on improving infrastructure and services in unplanned settlements across the country to enhance the quality of life for urban residents in Dar es Salaam. The urban infrastructure and services upgrading project (2005-2012) consisted of two phases. Phase I encompassed 16 informal settlements, benefiting over 160,000 individuals (Magina et al., 2020).

2.3 Google Earth Engine Platform

Google Earth Engine (GEE) has the potential to address the challenges associated with handling large volumes of data by providing advanced data processing and analytic tools, high computational power, and vast storage capacities (Method et al., 2019). GEE's extensive archives of imagery and data products, such as Landsat-8, Sentinel-1 and -2, and MODIS (Amani et al., 2019), eliminate the need to download large datasets to local directories. The ability to integrate different feature sets and parallelize image processing through effective script writing has enabled timely outputs at great speeds (Patel et al., 2015). GEE provides cloud resources that allow for fast processing and analysis of images, rendering traditional software and desktop-based image analysis obsolete (Forest et al., 2022). Researchers have leveraged GEE cloud-computing for mapping purposes at various scales, from global (X. Liu et al., 2018), continental(Forest et al., 2022), to country scale (Composition, 2020).

GEE has been used in urban environments for mapping purposes, as shown in several studies (Goldblatt et al., 2018). For example, researchers in (Shafizadeh-Moghadam et al., 2021) integrated spectral, textural, and topographical features in land use and land cover (LULC) mapping in the Tigris-Euphrates basin. In another study, (Shafizadeh-Moghadam et al., 2021) utilized spectral and textural indices for object-based LULC classification in Trasimeno Lake, in Umbria, Central Italy. In (Computing et al., n.d.; Forest et al., 2022), Landsat 7 and Landsat 8 bands, vegetation indices, and textural features were used to obtain a land cover map in Mozambique.

2.3 Image classification

Image classification is the process the pixels are assigned to the specified classes by the use of computer algorithm, image classification seeks to assist and support the interaction that occurs between human vision and computer or machine vision. This is accomplished by training computers to analyze data and sort images into pre-defined categories based on their visual content. Essentially, image classification involves identifying the category to which an image belongs, based on its visual features (Krishna et al., 2018).

Spectral indices are measured or calculated to show the relationship that exist between an object reflectance at two or more wavelengths of the electromagnetic spectrum. Spectral indices are measured by combination of pixels from two or more spectral bands from multispectral image to depict pixels that indicates or emphasize the area that is highly covered by a certain land cover type or lack of a certain land-cover type of interest in an image (Viswambharan & Lenhardt, 2019).As for this study spectral indices derived are NDVI and NDBI used as input parameters to extract features.

Normalized difference vegetation index, NDVI is the most common spectral index that is used to show the area highly covered by vegetation by comparing the different reflectance values of the red and near-infrared bands [normalized such that the minimum value is -1.0, and the maximum is +1.0](Viswambharan & Lenhardt, 2019). NDVI pixel's values range from -1 to 1: water, clouds, and snow have larger reflectance in Red than in NIR, so these features present negative index values; rock and bare soil areas have similar reflectance in the two bands and produce values near zero; because of their relatively high near-infrared reflectance and low red reflectance, vegetated areas generally return high values (Ali & Jassem, 2015). The common range for green vegetation

is ranged from 0.2 to 0.8 (Hasegawa, 1976).The NDVI was calculated using spectral band 8 which is near infrared (NIR) and red (R) band which is band 4 in sentinel 2A dataset.

$$\textbf{NDVI} = \frac{NIR-RED}{NIR+RED} \quad \text{---------------------} \textbf{ equation 1}$$

Normalized difference built up index, NDBI the index have been previously incorporated for the extraction of built-up areas (Kaimaris & Patias, 2016). The NDBI value lies between -1 to +1 where by the negative values represent water bodies or other land cover classes like vegetation, forest, cropland, shrubs and the higher value of NDBI represent built up areas. The NDBI was calculated using spectral band 11 which is short wave infrared (SWIR) and near infrared (NIR) band which is band 8 in sentinel 2A dataset.

$$\textbf{NDBI} = \frac{\boldsymbol{SWIR-NIR}}{\boldsymbol{SWIR+NIR}} \quad \textbf{--------------------- equation 2}$$

2.4 Random Forest classification

Random forest is a machine learning algorithm widely used for remote sensing classification, which incorporates elements of self-classification and decision trees (CART) (Engine, 2021) (Pal, 2005). It utilizes bootstrap sampling to extract a specific number of samples from the original dataset, creating a new training dataset. Each tree within the forest is allowed to grow to its maximum extent without pruning, and the random sampling process helps prevent overfitting (Id & Im, 2018) (Rodriguez-Galiano et al., 2012). Random forest offers several advantages, including superior classification accuracy and the ability to handle high-dimensional data without feature selection. In this study, a value of 50 was chosen for nTree to ensure accuracy and avoid overfitting, while the default value of the square root of the input feature data was used for Mtry (Tian et al., 2019). A supervised RF (Random Forest) machine learning algorithm was utilized for pixel-based classification in this study. Previous applications of RF classifiers to Sentinel-2A imagery in GEE have demonstrated successful mapping of built-up areas(Rudiastuti et al., 2021), human settlements(Trianni et al., 2014) and specifically, informal settlements(Tingzon et al., n.d.). The choice of RF classifier was based on its ability to handle urban area classification with high-dimensional feature spaces and its robustness in mapping informal settlements in complex environments (Matarira et al., 2023). RF classifiers also provide information about the contribution of each variable to the classification output, which is crucial for variable evaluation (Teluguntla et al., 2018). The entire classification process was conducted in GEE using the "ee.Classifier" package. Seven sets of constructed features were used as inputs for the classification models. Following (Rudiastuti et al., 2021) an RF model with 100 trees was created, and the number of variables per split was set to the square root of the total number of variables(Computing et al., n.d.). Training samples were chosen as small polygons to ensure homogeneity within each land cover class and minimize the influence of spatial correlation (Composition, 2020). The classification was performed using 899 training samples and 282 testing samples.The model employed a random sampling strategy, where approximately 70% of the training samples were randomly selected from the original dataset to generate a decision tree for each training sample separately. The remaining 30% of training samples were used as validation data for internal cross-validation to evaluate the classification accuracy of the random forest(Forest et al., 2022).

CHAPTER THREE

METHODOLOGY

3.0 Overview

This chapter presents the systematic procedures applied towards achieving the main objective. It includes the data collection, data preprocessing, processing techniques and data analysis methods that were used in this study.

3.1 Study area

Mwakibete is an administrative ward in the Mbeya Urban district of the Mbeya Region of Tanzania. The study area, which is 15.8274457911 km2 wide, is located in Southern part of Tanzania (8° 56′ 0″ S, 33° 29′ 0″ E), (8˚ 58′ 20″S, 33˚ 30′ 0″E). The area is characterized by heterogeneous landscape with abundant Built-up lands, wooded areas, Cropland and less vegetation and water surface. The study area was chosen within the tropical region (where the imageries are available in GEE) to effectively design and test the methodology as shown on figure 3.2. the informal settlement in Mwakibete in this study were defined as buildings or built-up areas over the area. The area is growing informally since there are area which were planned as mental retardation and cerebral, cemetery area but they are not as planned now as shown on the figure 3.1.

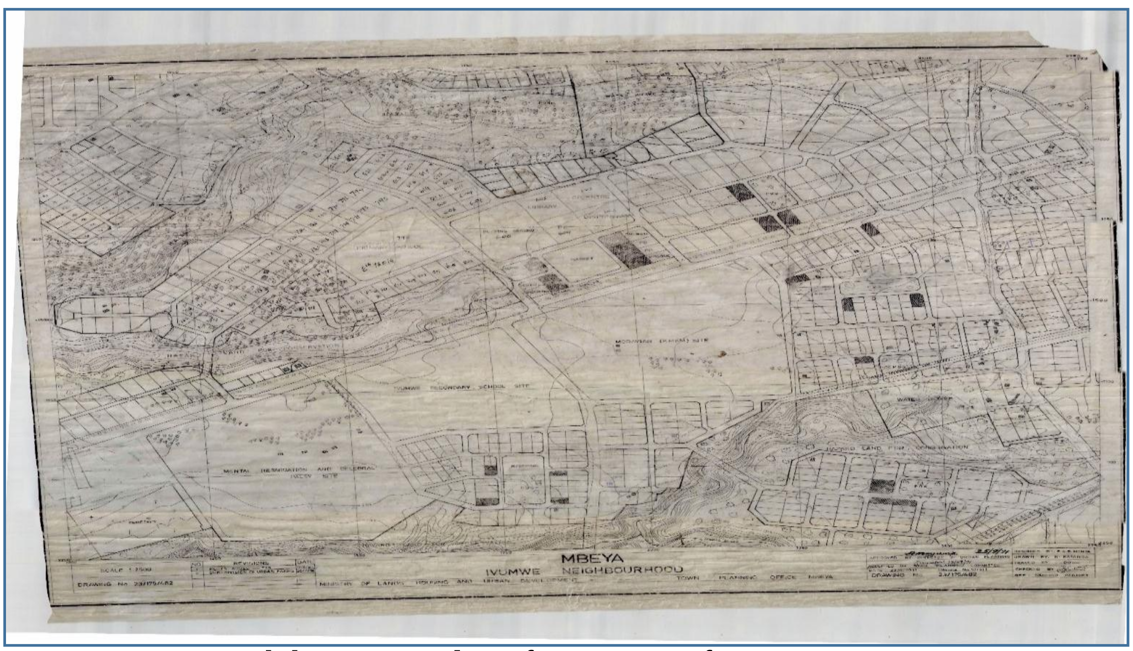
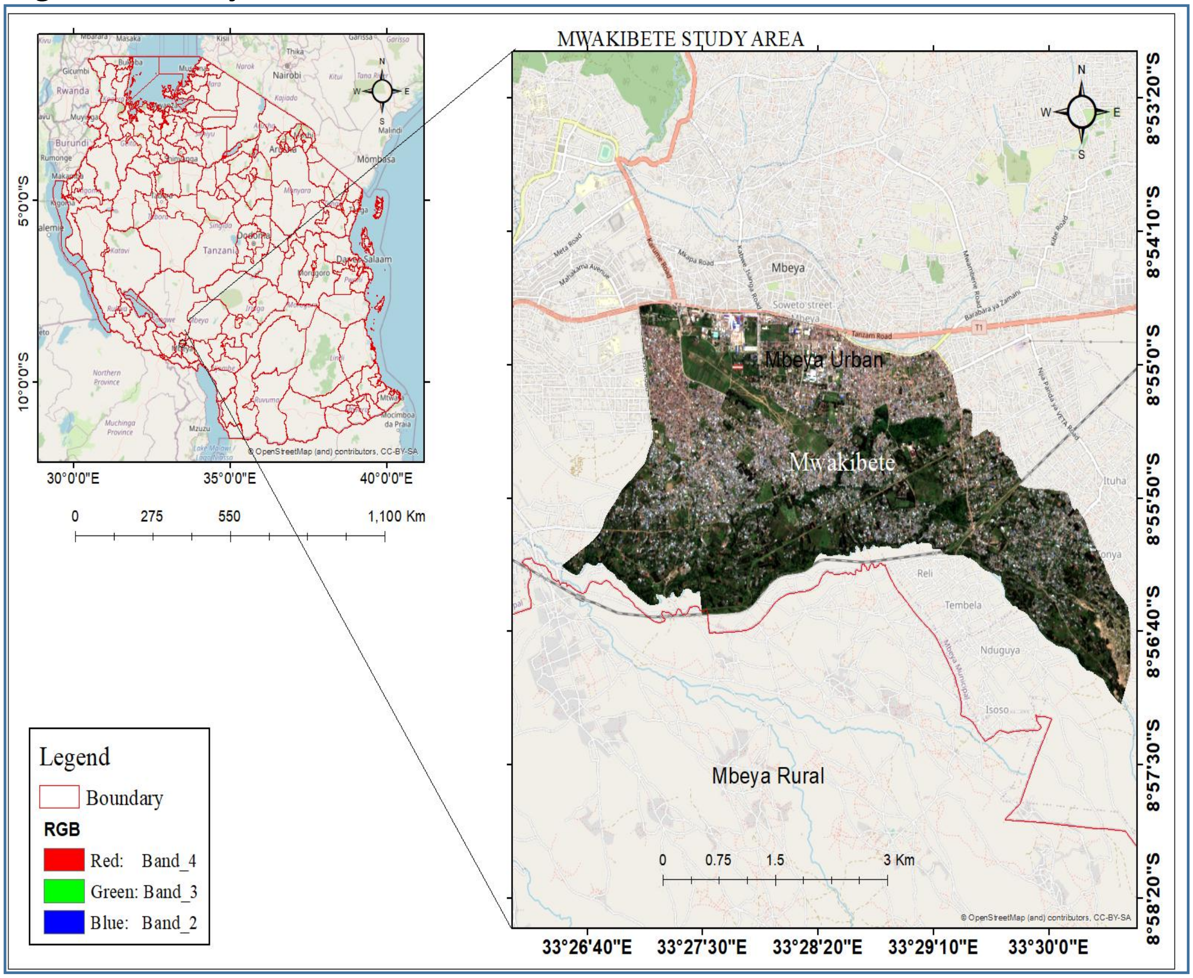


Figure 3. 1; Mwakibete town plan of 1991 part of

Figure 3. 2 Study Area

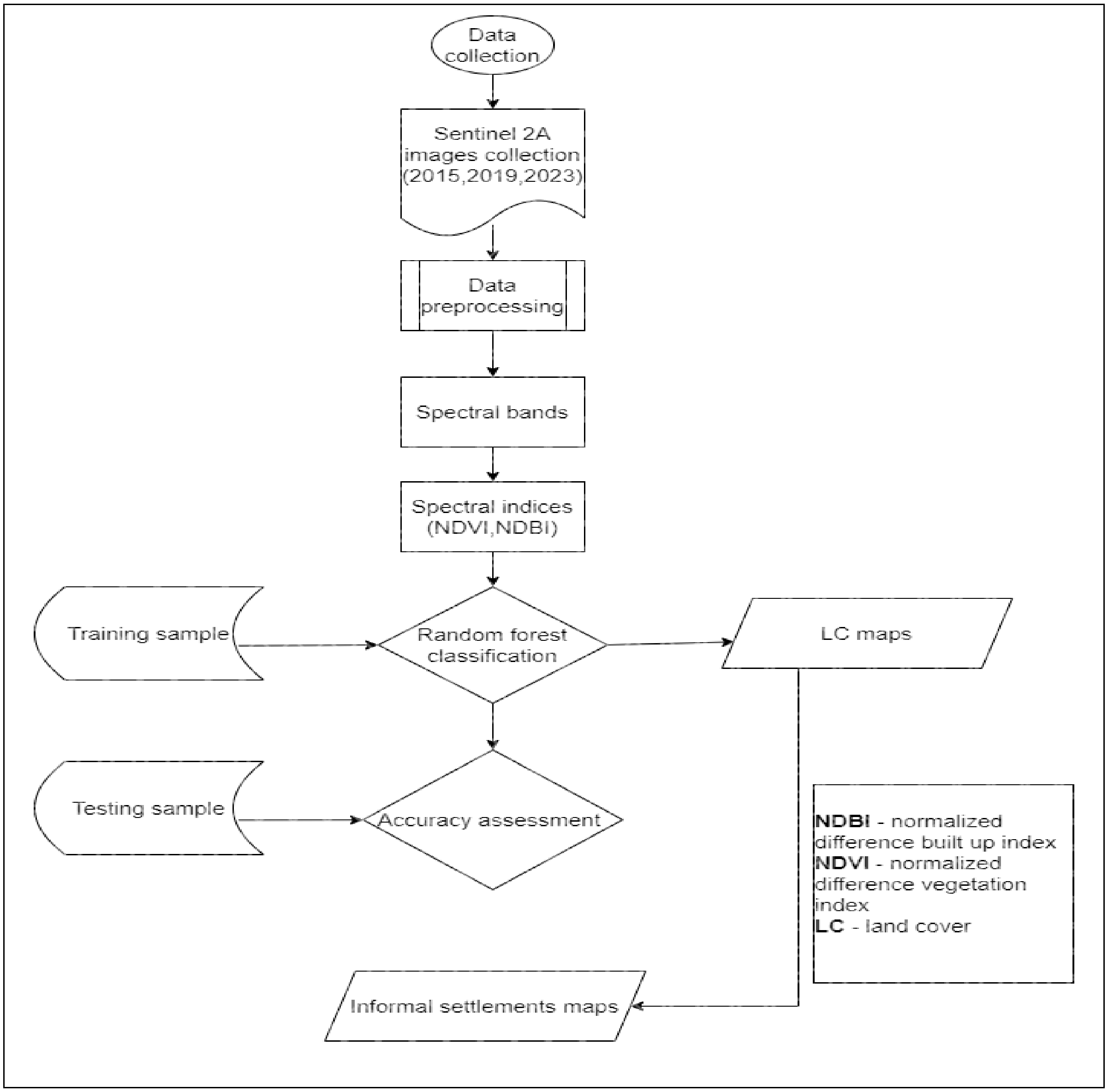


3.2 Data Collection

For this research Sentinel 2A data was used to implement informal settlement in Mbeya with resolution of 10m, as shown in Table 3.1.

Table 3. 1 Data collection

Data	Format	Resolution	Source	Use
Sentinel 2A	JP2	10m	GEE	Toclassify informal settlement area
Boundary	SHAPEFILE		DIVA www.Diva Gis.com	To show the area of interest



3.3 Dataset

The analysis utilized the Sentinel-2A image collection, specifically the COPERNICUS/S2_SR surface reflectance dataset. This image collection consists of 13 bands covering the visible, near-infrared, and shortwave infrared (SWIR) wavelengths. It includes four bands at a spatial resolution of 10 m, six bands at 20 m, and three bands at 60 m.

3.4 Feature extraction

In this research spectral features, 10 spectral bands and spectral indices were used as input parameters for image classification, as shown in Table 3.2.

Table 3. 2:Image feature sets that were extracted from sentinel 2A imagery

Image feature	Name	No of features
Spectral Bands	Bands (B4, B8 and B11)	3
Spectral indices	NDVI and NDBI	2

The mathematical equations used for the calculation of the mentioned indices are presented in Table 3.3.

Table 3. 3:Spectral indices selected for research

Spectral indices	Equation
NDVI	$\frac{B8-B4}{B8+B4}$ ----equation 3
NDBI	$\frac{B11-B8}{B11+B8}$ -------equation 4

3.5 Image Classification

Image classification based on supervised image classification is accomplished through the application of random forest (RF) machine learning classifier, the processes were achieved in GEE which uses the bootstrap aggregation to create multiple decision trees (DTs). The sample points were splitted into two part training sample and validation or testing sample where by 70% sample point were trained for training by the classifier and the remaining 30% for validation or testing. A graph was plotted using the variable importance included in the output from the “explain” command to visualize features’ relevance in each classification. This graph was used for each approach to empirically analyze and iteratively select the more crucial bands for the LULC classification.

3.6 Accuracy Assessment

The confusion matrix (CM) technique, a commonly used methodology based on comparing classification outputs with ground truth data, was used to assess the accuracy of the various approaches(C. Liu et al., 2007). The matrix was used to produce specific accuracy measures: Overall Accuracy (OA), Producer’s Accuracy (PA), and User’s Accuracy (UA), according to the following formulas:

$$\frac{OA = Number\ of\ correct\ predictions}{Total\ number\ of\ predictions}$$ --------- equation 5

$$\frac{PA = Number\ correctly\ identified\ in\ a\ given\ map\ class}{Number\ actually\ in\ that\ reference\ class}$$ -------------- equation 6

$$\frac{UA = Number\ correctly\ identified\ in\ a\ given\ map\ class}{Number\ claimed\ to\ be\ in\ that\ map\ class}$$ --------------- equation7

The OA is commonly expressed as a percentage, with 100% accuracy indicating that all reference sites were correctly classified (Al-Saady et al., 2015).

CHAPTER FOUR

RESULT, ANALYSIS AND DISCUSSION

4.1 Overview

This chapter presents all the results and analysis obtained in this research; it will also provide a detailed discussion of all results obtained.

4.2 Spectral indices

In this study NDVI and NDBI of 2015, 2019 and 2023 was determined using band 4, 8 and 11 of sentinel 2A, where by these spectral indices then used as input in image classification. Where by the NDVI were calculated and they were classified using NDVI pixel's value assigned to different classes as well as NDBI were measured to show built up area with high value of +1, also the indices were used to compare with the land cover maps to see how the built-up area distributed in the area. Here the NDVI shows classes where by built up are presented in negative value while vegetated areas generally return high values as shown on figure 4.1

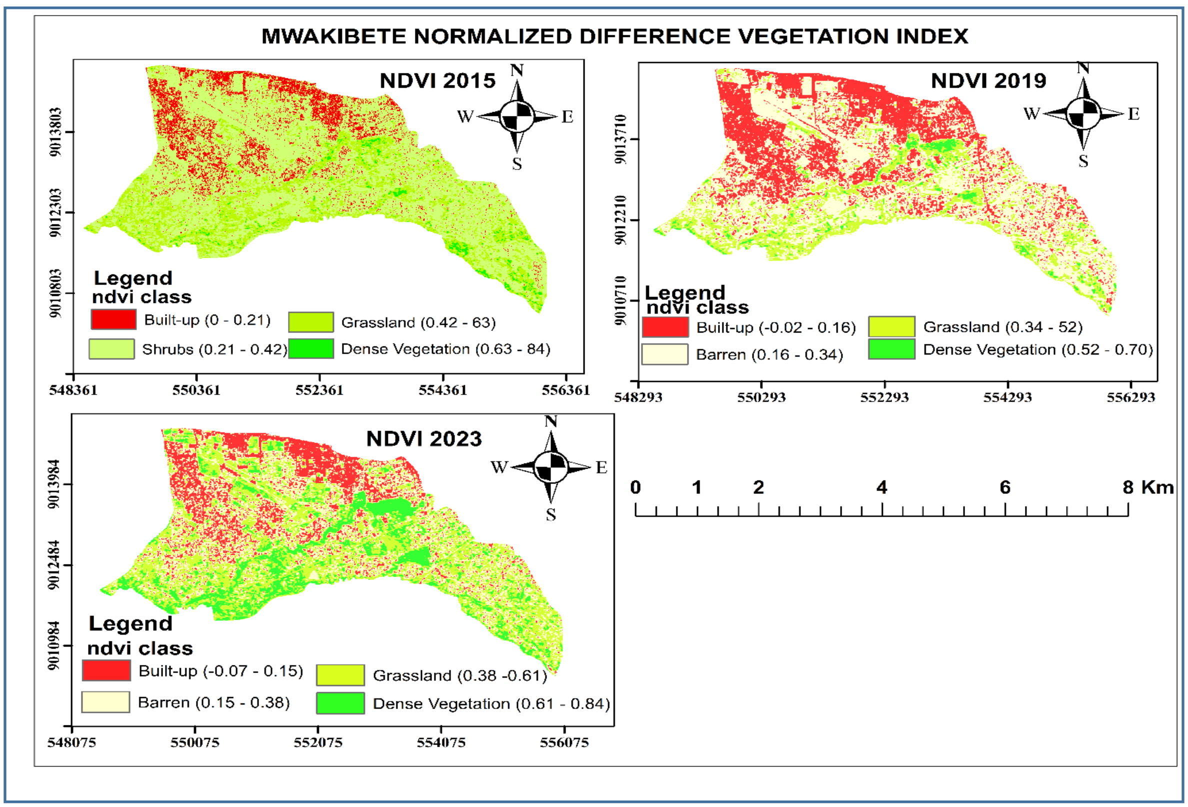


Figure 4. 1: Mwakibete NDVI map

In normalized difference built up index (NDBI) shows that built up area are more concentrated in high level represented in red color with the positive values in 2019 and 2023 but negative in 2015 while other classes like bare land, cropland vegetation and water area represented in low value with negative value and it show that the high value which is positive value is increasing from 2015 to 2023 which indicate that the built-up area increasing as shown on figure 4.2

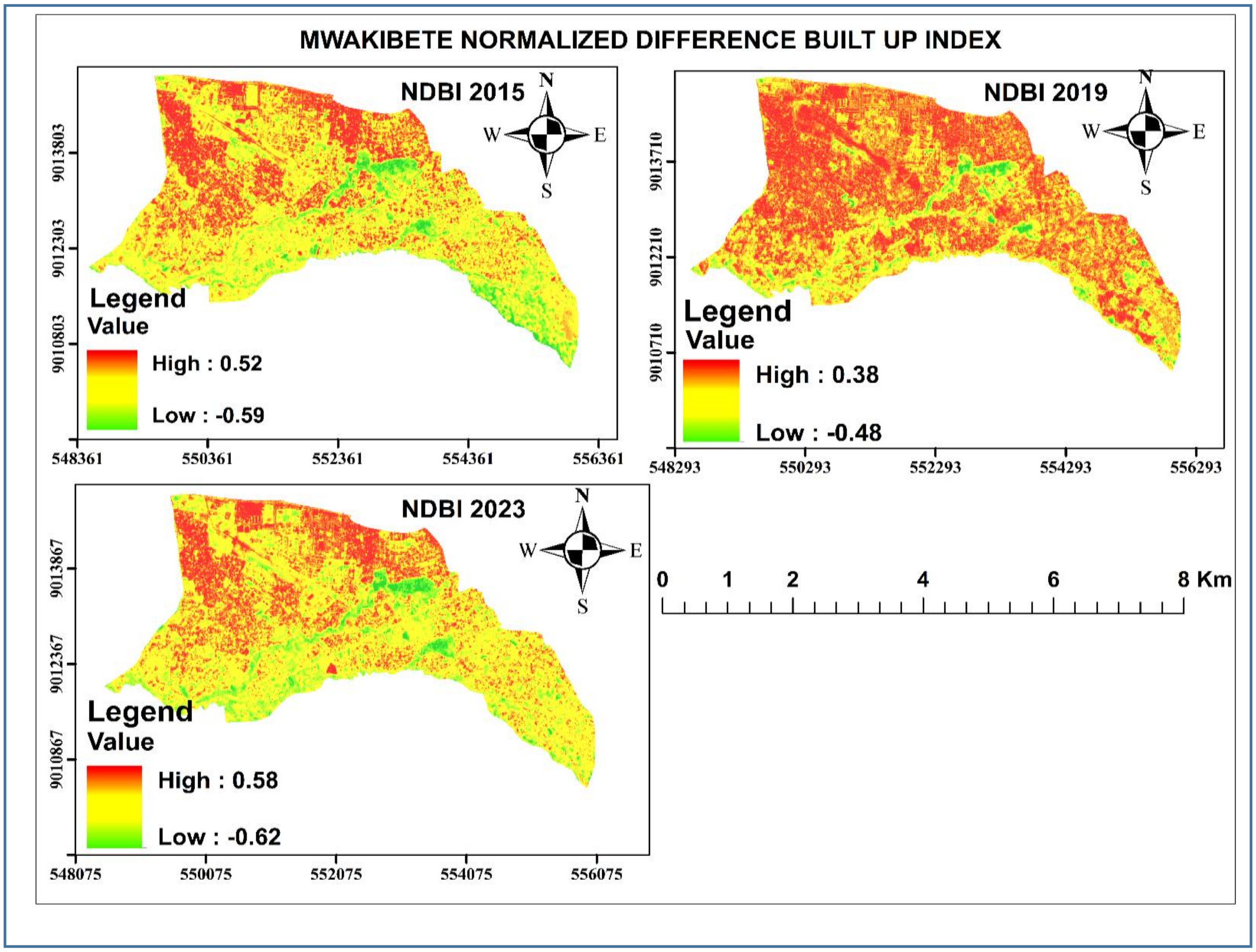


Figure 4. 2: Mwakibete NDBI map

4.3 Classified images

Classified images of Mwakibete from 2015 to 2023 from sentinel were used to extract and map informal settlement area from classified land cover accomplished through the use of random forest algorithm, classified images comprise of four classes which are built up, woodland, cropland and bare land as shown on the figure 4.3.

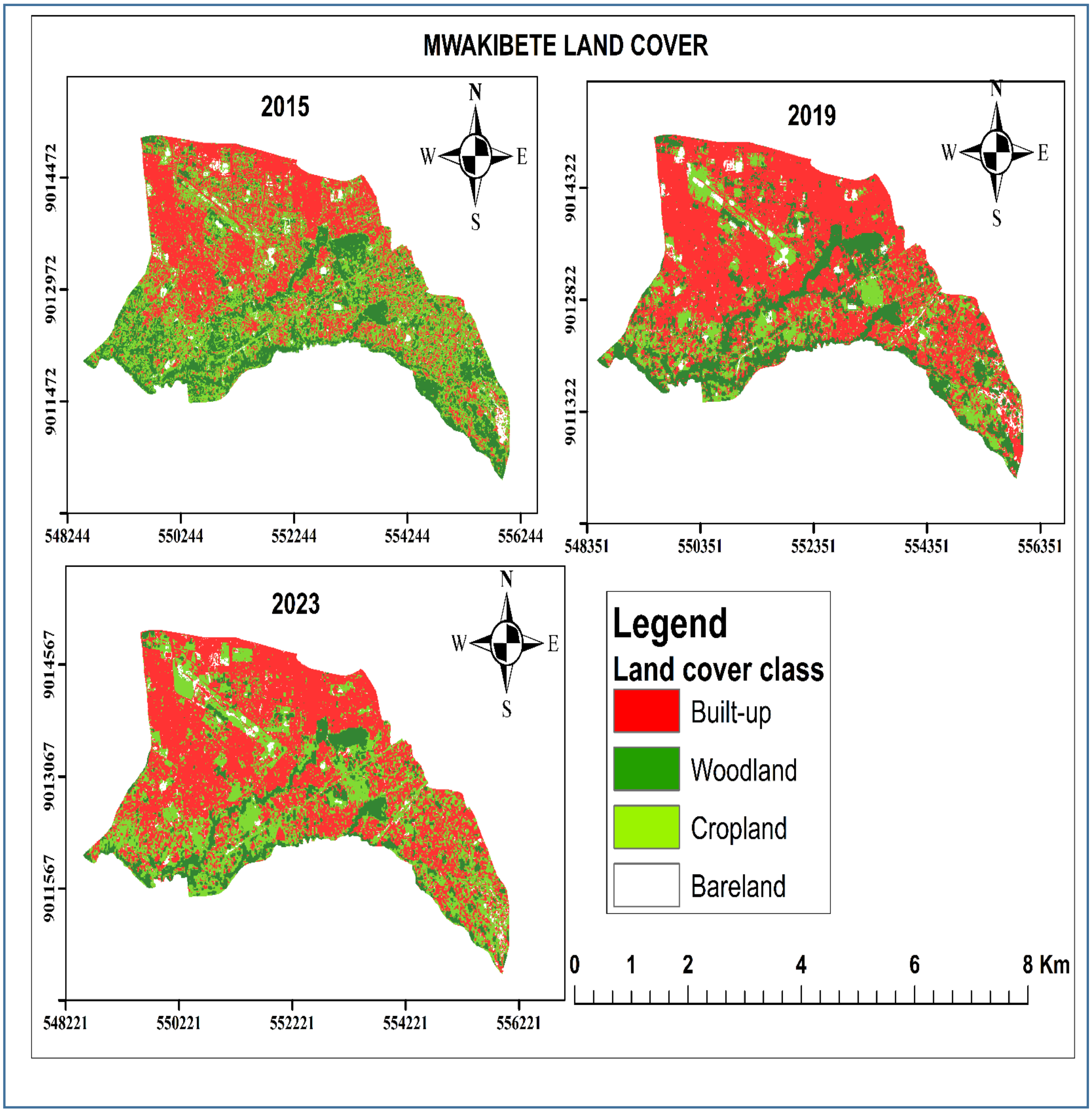


Figure 4. 3: land cover map of mwakibete

Form the classified images the area in hectare was calculated in each class for all years as shown on the table 4.1 were by area cover by built up is increasing from 2015 to 2023 by 264.8784 hectares while the other classes their area are decreasing and the change in area for each class for all years as illustrated from the graph shown on the figure 4.4 and 4.5

Table 4. 1: Land cover area

	Area(ha)		
Class/year	**2015**	**2019**	**2023**
Built up	593.625479	890.237661	858.503832
Woodland	349.076342	338.138944	190.422975
Cropland	610.773973	311.591615	508.851818
Bareland	41.671663	55.18932	37.432325

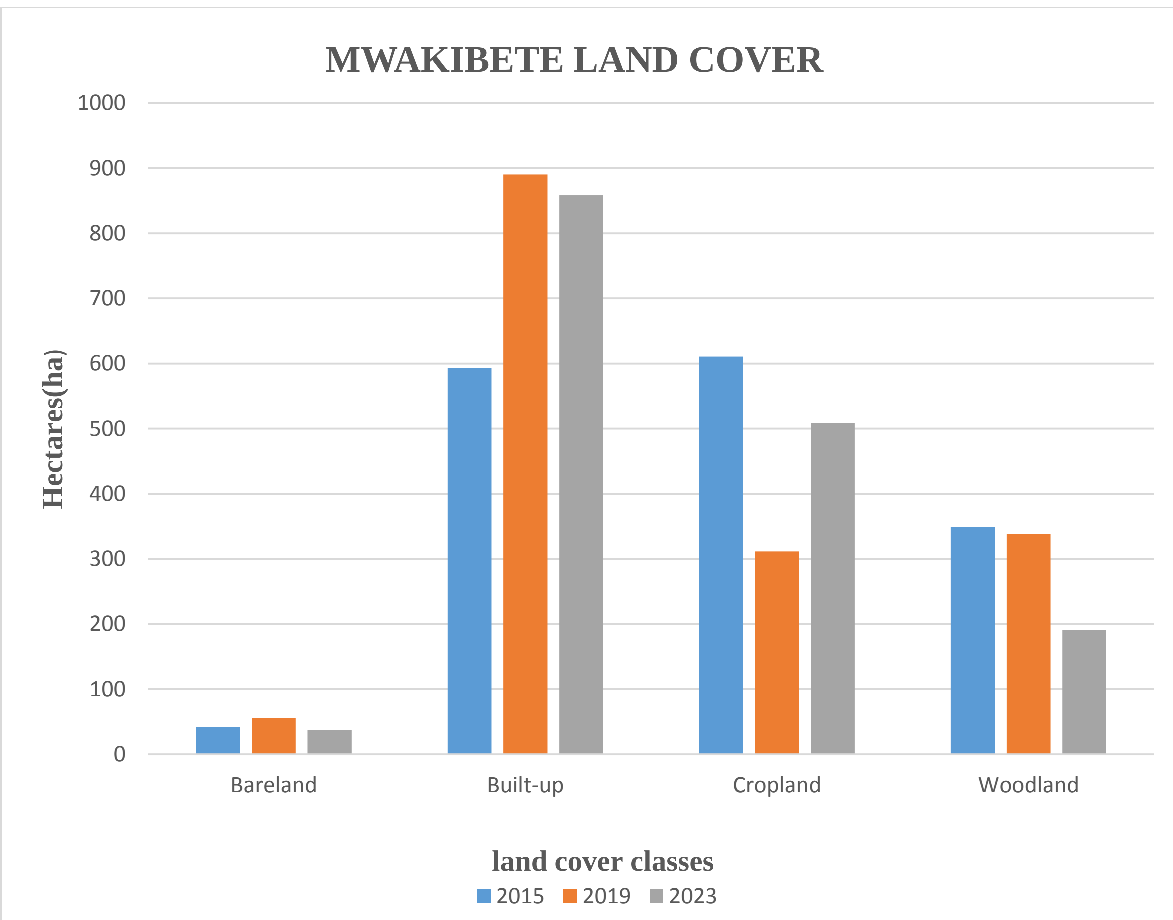


Figure 4. 4: Graph of mwakibete land cover area in hectares

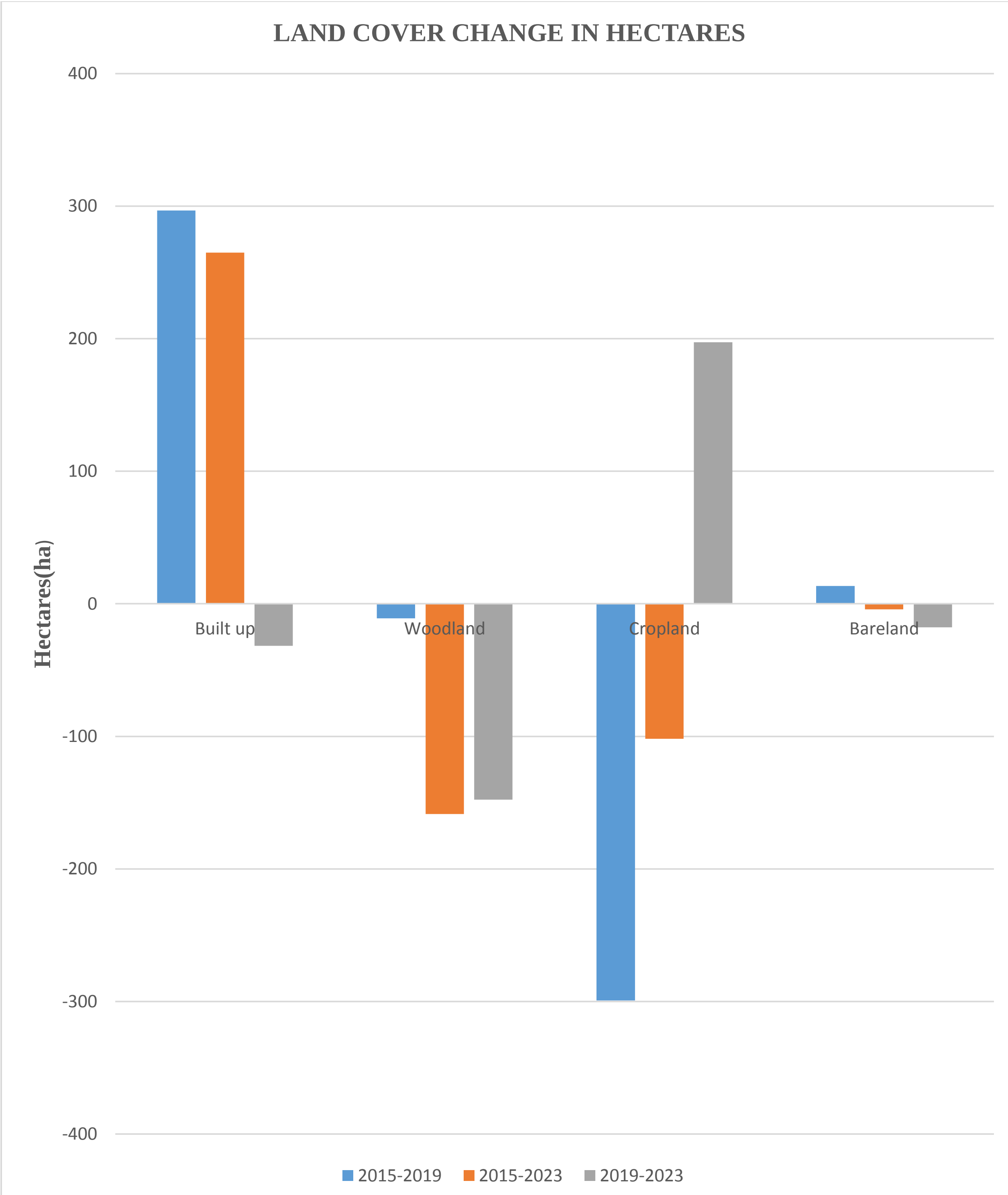


Figure 4. 5: Graph of mwakibete showing land cover change in hectares

4.4 Extracted built up (informal settlement area)

Built up area were extracted from the classified land cover map to show the informal settlement area within the area of the interest, built up (informal settlement) are shown on the figure 4.6

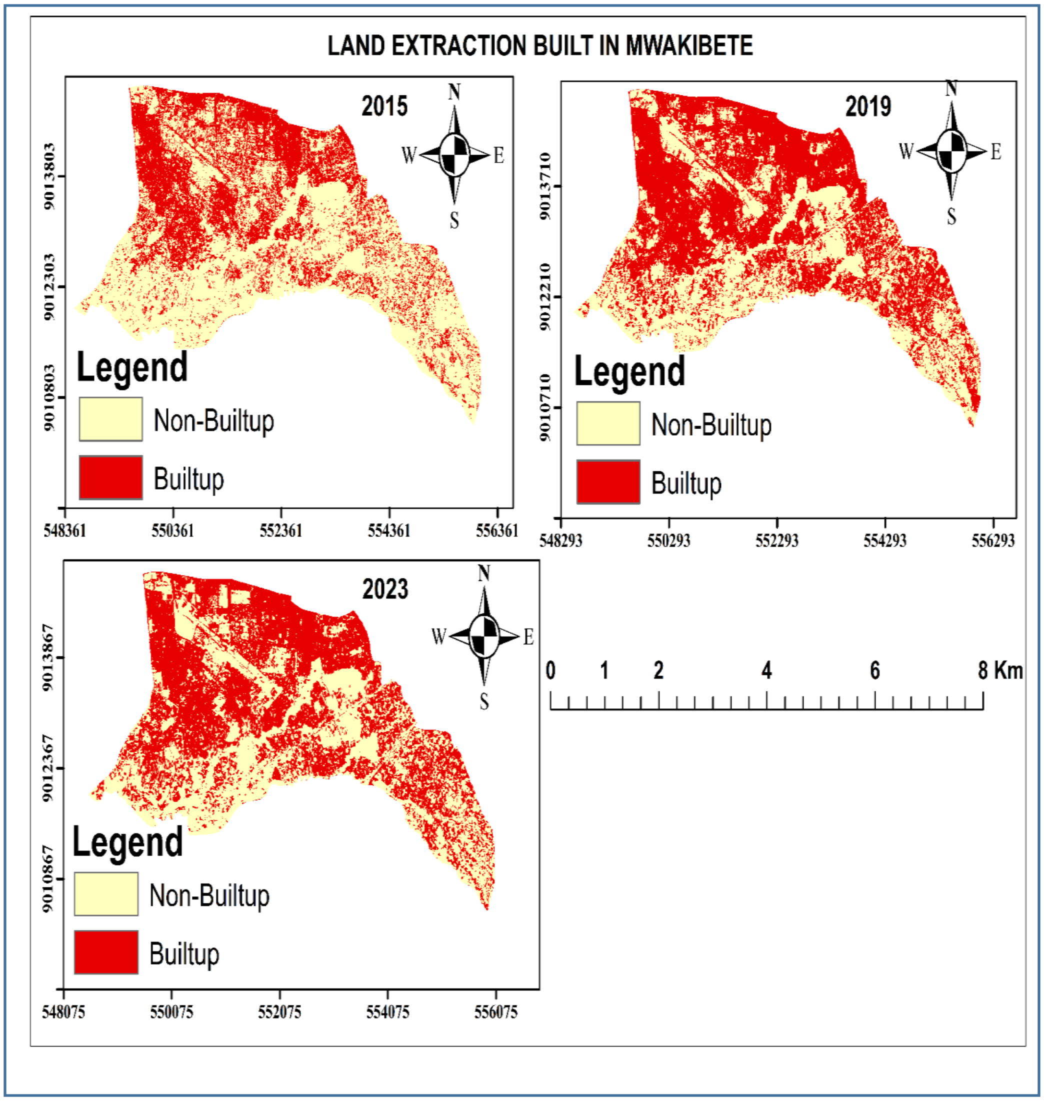


Figure 4. 6: Map showing mwakibete built up area and non built up area

4.5 Informal settlement assessment

The informal settlements were assessed through built up land cover as shown on the figures above, the classified map help to extract informal settlements area and able to analyze the extent of the expansion or growth of the informal settlements through the calculated area in hectares from 2015 to 2023 where by the area of built up in 2015 was 593.625479 hectares while in 2023 area measured for built area was 858.503832 hectares. The findings show that the area is increased by 264.8784 hectares.

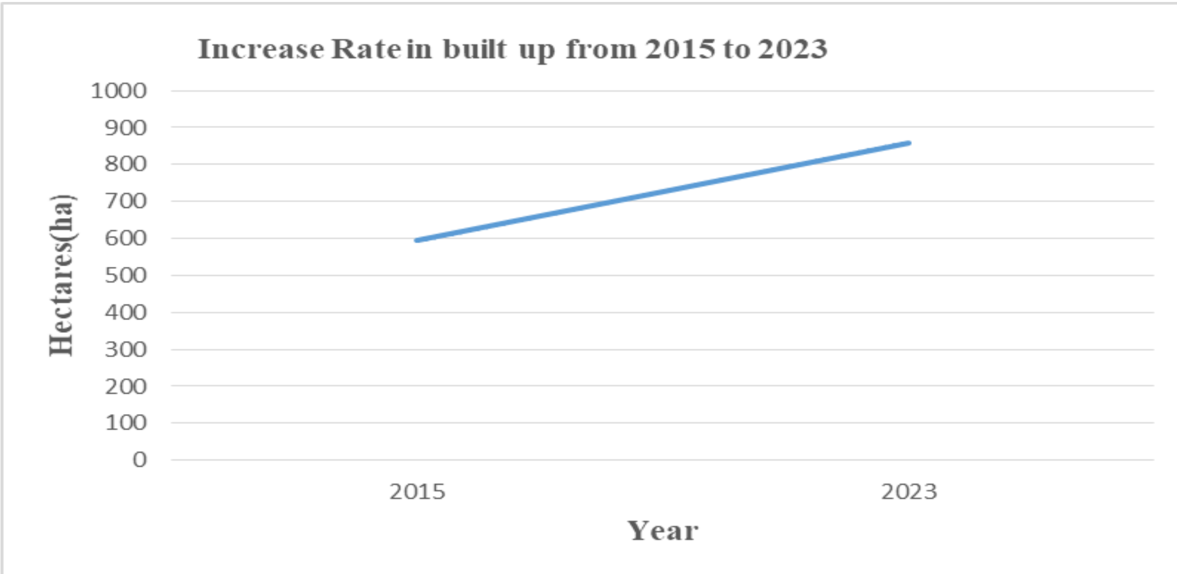


Figure 4.7: Graph showing the rate increasing in built up from 2015 to 2023

4.6 Classification accuracy

Images classification were accomplished through random forest algorithm to produce land cover maps of 2015, 2019 and 2023. The overall accuracy and kappa coefficient derived by random forest algorithm in GEE by calculating the percentages of both producer's accuracy and user's accuracy from error confusion matrix and the overall accuracy derived were 88.89%, 80.77% and 86.47% as shown on table 4.2, 4.3 and 4.4.

Table 4. 2 Accuracy assessment of land cover 2015

	Accuracy Assessment of land cover 2015							
		Reference				Total	commision error	user's accuracy(%)
		Built up	Woodland	Cropland	Bareland			
classification	Built up	46	0	6	0	52	11.54%	88.46%
	Woodland	0	47	1	0	48	2.08%	97.92%
	Cropland	6	0	23	2	31	25.81%	74.19%
	Bareland	0	0	2	20	22	9.09%	90.91%
	Total	52	47	32	22	153		
ommision error		11.54%	0.00%	28.13%	9.09%			
producer's accuracy(%)		88.46%	100.00%	71.88%	90.91%			
Overall accuracy	88.89%							
kappa coefficient=p(a)-p(r)/1-p(r)	84.68%							
p(a)	0.888888889							
p(r)	0.27493699							

Table 4. 3 Accuracy assessment of land cover 2019

	Accuracy Assessment of land cover 2019							
		Reference				Total	commision error	user's accuracy(%)
		Built up	Woodland	Cropland	Bareland			
classification	Built up	100	0	7	1	108	7.41%	92.59%
	Woodland	3	39	19	0	61	36.07%	63.93%
	Cropland	11	1	48	1	61	21.31%	78.69%
	Bareland	6	1	0	23	30	23.33%	76.67%
	Total	120	41	74	25	260		
ommision error		16.67%	4.88%	35.14%	8.00%			
producer's accuracy(%)		83.33%	95.12%	64.86%	92.00%			
Overall accuracy	80.77%							
kappa coefficient=p(a)-p(r)/1-p(r)	72.27%							
p(a)	0.807692308							
p(r)	0.30658284							

Table 4. 4 Accuracy assessment of land cover 2023

	Accuracy Assessment of land cover 2019							
		Reference				Total	commision error	user's accuracy(%)
		Built up	Woodland	Cropland	Bareland			
classification	Built up	150	0	10	7	167	10.18%	89.82%
	Woodland	0	49	3	0	52	5.77%	94.23%
	Cropland	8	2	45	2	57	21.05%	78.95%
	Bareland	8	0	1	18	27	33.33%	66.67%
	Total	166	51	59	27	303		
ommision error		9.64%	3.92%	23.73%	33.33%			
producer's accuracy(%)		90.36%	96.08%	76.27%	66.67%			
Overall accuracy	86.47%							
kappa coefficient=p(a)-p(r)/1-p(r)	78.34%							
p(a)	0.864686469							
p(r)	0.375409818							

CHAPTER FIVE

CONCLUSION AND RECOMMENDATION

5.1 Overview

This chapter elaborate the difficulties that were faced during the entire work of the research, the conclusion from the findings or results obtained from the research and also recommend on what to be improved from the end of this research to be starting point from the other research.

5.2 Conclusion

This study has employed techniques of GIS and remote sensing to assess and analyze the expansion of informal settlement in Mwakibete from 2015 to 2023.it provides description changes of informal settlement through land cover map produced and define built up as informal settlement, the study use built up class from land cover map to define and assess the extent of informal settlements in hectares and use NDVI and NDBI for comparison. From the year 2015 to 2023 it has been observed that there is an increase in built up area where the area calculated was 858.50 ha (53.81%), 890.237661ha (55.80%) and 593.625479ha (37.21%) respectively. From the result the informal areas are increasing by 264.8784 hectares which is 16.60% from 2015 to 2023 and assessment uses NDVI and NDBI together with land cover to analyze the growth of informal settlement such that NDBI has contribute more to the classification of the image than NDVI.The assessment done from this study can have important impact in urban planning, since the city is still growing and according to sustainable development goals good infrastructure and good city or urban development are one of the goals so this study can be of helpful to the government. Therefore, to assess informal settlements should be done continuously as measure of problem caused by informal settlement.

5.4 Recommendation

The study show the expansion of informal settlement through identified built up from land cover and it show an increase of 16% from 2015 to 2023 this information can be used reduce the expansion of informal area by structuring again the town plan so that the area expand as planned thus improve the living standard of people and maintain good sanitation condition and environment at all.

REFERENCE

Abebe, M. S., Derebew, K. T., & Gemeda, D. O. (2019). Exploiting temporal-spatial patterns of informal settlements using GIS and remote sensing technique: a case study of Jimma city, Southwestern Ethiopia. *Environmental Systems Research*, *8*(1). https://doi.org/10.1186/s40068-019-0133-5

Abunyewah, M., Gajendran, T., & Maund, K. (2018). ScienceDirect ScienceDirect ScienceDirect Profiling Informal Settlements for Disaster Risks Profiling Informal Settlements for a Disaster Risks. *Procedia Engineering*, *212*(2017), 238–245. https://doi.org/10.1016/j.proeng.2018.01.031

Al-Saady, Y., Merkel, B., Al-Tawash, B., & Al-Suhail, Q. (2015). Land use and land cover (LULC) mapping and change detection in the Little Zab River Basin (LZRB), Kurdistan Region, NE Iraq and NW Iran. *FOG - Freiberg Online Geoscience*, *43*(November), 1–32.

Ali, S. M., & Jassem, D. R. (2015). *Monitoring Vegetation Areas by using Remote Sensing Techniques*. *September*. https://doi.org/10.13140/RG.2.1.3691.8244

Amani, M., Mahdavi, S., Afshar, M., Brisco, B., Huang, W., Mohammad, S., Mirzadeh, J., White, L., Banks, S., Montgomery, J., & Hopkinson, C. (2019). *Canadian Wetland Inventory using Google Earth Engine : The First Map and Preliminary Results*. 1–20. https://doi.org/10.3390/rs11070842

Asmat, A., & Zamzami, S. Z. (2012). Automated House Detection and Delineation using Optical Remote Sensing Technology for Informal Human Settlement. *Procedia - Social and Behavioral Sciences*, *36*(June 2011), 650–658. https://doi.org/10.1016/j.sbspro.2012.03.071

Banda, E. M., & Mwale, F. D. (2018). Utility performance in supplying water to informal settlements: A case study from Malawi. *Utilities Policy, 55*, 151–157. https://doi.org/https://doi.org/10.1016/j.jup.2018.09.009

Chadchan, J., & Shankar, R. (2012). An analysis of urban growth trends in the post-economic reforms period in India. *International Journal of Sustainable Built Environment, 1*(1), 36–49. https://doi.org/10.1016/j.ijsbe.2012.05.001

Composition, I. (2020). *Land Cover Classification using Google Earth Engine and Random Forest Classifier — The Role of Image Composition*.

Computing, C., Mananze, S., & Isabel, P. (n.d.). *Mapping and Assessing the Dynamics of Shifting Agricultural Landscapes Using Google Earth Engine*. 1–23.

Demographic Trends and Urbanization. (2021). *Demographic Trends and Urbanization*. https://doi.org/10.1596/35469

Engine, E. (2021). *Accuracy Improvements to Pixel-Based and Object-Based LULC Classification with Auxiliary Datasets from Google Earth Engine*.

Forest, R., Using, C., & Information, T. (2022). *Google Earth Engine for Informal Settlement Mapping : A Random Forest Classification Using Spectral and Textural Information*.

Gilbert, A. (2005). Book Review: The challenge of slums: global report on human settlements 2003. In *Progress in Human Geography* (Vol. 29, Issue 1). https://doi.org/10.1177/030913250502900122

Goldblatt, R., Deininger, K., & Hanson, G. (2018). Utilizing publicly available satellite data for urban research: Mapping built-up land cover and land use in Ho Chi Minh City, Vietnam. *Development Engineering*, *3*(March), 83–99. https://doi.org/10.1016/j.deveng.2018.03.001

Guttenberg, A. Z. (1960). Urban Structure and Urban Growth. *Journal of the American Institute of Planners*, *26*(2), 104–110. https://doi.org/10.1080/01944366008978392

Hasegawa, S. (1976). Metabolism of limonoids. Limonin d-ring lactone hydrolase activity in pseudomonas. *Journal of Agricultural and Food Chemistry*, *24*(1), 24–26. https://doi.org/10.1021/jf60203a024

Id, S. P., & Im, J. (2018). *Classification and Mapping of Paddy Rice by Combining Landsat and SAR Time Series Data*. 1–22. https://doi.org/10.3390/rs10030447

J. Gwaleba, M. (2018). Urban Growth in Tanzania: Exploring Challenges, Opportunities and Management. *International Journal of Social Science Studies*, *6*(12), 47. https://doi.org/10.11114/ijsss.v6i12.3783

Kaimaris, D., & Patias, P. (2016). Identification and Area Measurement of the Built-up Area with the Built-up Index (BUI). *International Journal of Advanced Remote Sensing and GIS*, *5*, 1844–1858. https://doi.org/10.23953/cloud.ijarsg.64

Krishna, M. M., Neelima, M., Harshali, M., & Rao, M. V. G. (2018). Image classification using Deep learning. *International Journal of Engineering and Technology(UAE)*, *7*(August), 614–617. https://doi.org/10.14419/ijet.v7i2.7.10892

Kuffer, M., Pfeffer, K., Sliuzas, R., & Baud, I. (2016). Extraction of Slum Areas From VHR Imagery Using GLCM Variance. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *9*(5), 1830–1840.

https://doi.org/10.1109/JSTARS.2016.2538563

Kyessi, A., & Sekiete, T. (2014). Formalising Property Rights in Informal Settlements and Its Implications on Poverty Reduction: The Case of Dar es Salaam, Tanzania. *Journal of Business and Economics*, *5*, 2352–2370. https://doi.org/10.15341/jbe(2155-7950)/12.05.2014/016

Lintelo, D. J. H., Gupte, J., Mcgregor, J. A., Lakshman, R., Jahan, F., & Asia, S. (2018). Wellbeing and urban governance : Who fails , survives or thrives in informal settlements in Bangladeshi cities ? *Cities*, *72*(October 2017), 391–402. https://doi.org/10.1016/j.cities.2017.10.002

Liu, C., Frazier, P., & Kumar, L. (2007). Comparative assessment of the measures of thematic classification accuracy. *Remote Sensing of Environment*, *107*, 606–616. https://doi.org/10.1016/j.rse.2006.10.010

Liu, X., Hu, G., Chen, Y., Li, X., Xu, X., Li, S., Pei, F., & Wang, S. (2018). No Title. *High-Resolution Multi-Temporal Mapping of Global Urban Land Using Landsat Images Based on the Google Earth Engine Platform." Remote Sensing of Environment*, *209*(null), 227.

Lombard, M. (2014). Constructing ordinary places : Place-making in urban informal settlements in Mexico. *Progress in Planning*, *94*, 1–53. https://doi.org/10.1016/j.progress.2013.05.003

Magina, F. B., Kyessi, A. G., & Kombe, W. J. (2020). *Informal Settlements Regularisation in Tanzania : the Case of Magengeni in Dar Es Salaam City Benefits and Burdens Accrued in Community-Led*. *December*, 1–24.

Matarira, D., Mutanga, O., & Naidu, M. (2023). *Engine Using the Integration of Sentinel-1 , Sentinel-2 , and*. 1–17.

Method, P., Yin, Q., Liu, M., Cheng, J., Ke, Y., & Chen, X. (2019). *Mapping Paddy Rice Planting Area in Northeastern China Using Spatiotemporal Data Fusion and*.

Mohammadian, H., Tavakoli, J., & Khani, H. (2017). Monitoring land use change and measuring urban sprawl based on its spatial forms The case of Qom city. *The Egyptian Journal of Remote Sensing and Space Sciences*, *20*(1), 103–116. https://doi.org/10.1016/j.ejrs.2016.08.002

Nabutola, W. L. (2011). *Planning and Managing Urbanization in Kenya Following the New Constitutional Dispensation Promulgated in August 2010*. *2010*(May), 18–22.

Nassar, D. M., & Elsayed, H. G. (2018). From Informal Settlements to sustainable communities.

Alexandria Engineering Journal, *57*(4), 2367–2376. https://doi.org/10.1016/j.aej.2017.09.004

Pal, M. (2005). Random forest classifier for remote sensing classification. *International Journal of Remote Sensing*, *26*(1), 217–222. https://doi.org/10.1080/01431160412331269698

Patel, N., Angiuli, E., Gamba, P., Gaughan, A., Lisini, G., Stevens, F., Tatem, A., & Trianni, G. (2015). Multitemporal settlement and population mapping from Landsat using Google Earth Engine. *International Journal of Applied Earth Observation and Geoinformation*, *35*, 199–208. https://doi.org/10.1016/j.jag.2014.09.005

Payne, G. (2005). Getting ahead of the game: A twin-track approach to improving existing slums and reducing the need for future slums. *Environment and Urbanization*, *17*(1), 135–146. https://doi.org/10.1177/095624780501700114

Penrose, K., De Castro, M. C., Werema, J., & Ryan, E. T. (2010). Informal urban settlements and cholera risk in Dar es Salaam, Tanzania. *PLoS Neglected Tropical Diseases*, *4*(3), 1–11. https://doi.org/10.1371/journal.pntd.0000631

Rodriguez-Galiano, V., Ghimire, B., Rogan, J., Chica-Olmo, M., & Rigol-Sanchez, J. (2012). An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, *67*, 93–104. https://doi.org/10.1016/j.isprsjprs.2011.11.002

Rudiastuti, A., Farda, N., & Ramdani, D. (2021). *Mapping built-up land and settlements: a comparison of machine learning algorithms in Google Earth engine*. https://doi.org/10.1117/12.2619493

Shafizadeh-Moghadam, H., Khazaei, M., Alavipanah, S. K., & Weng, Q. (2021). Google Earth Engine for large-scale land use and land cover mapping: an object-based classification approach using spectral, textural and topographical factors. *GIScience & Remote Sensing*, *58*(6), 914–928. https://doi.org/10.1080/15481603.2021.1947623

Teluguntla, P., Thenkabail, P. S., Oliphant, A., Xiong, J., Krishna, M., Congalton, R. G., Yadav, K., & Huete, A. (2018). ISPRS Journal of Photogrammetry and Remote Sensing A 30-m landsat-derived cropland extent product of Australia and China using random forest machine learning algorithm on Google Earth Engine cloud computing platform. *ISPRS Journal of Photogrammetry and Remote Sensing*, *144*(July), 325–340. https://doi.org/10.1016/j.isprsjprs.2018.07.017

Tian, F., Wu, B., Zeng, H., & Zhang, X. (2019). *Efficient Identification of Corn Cultivation Area with Multitemporal Synthetic Aperture Radar and Optical Images in the Google Earth Engine Cloud Platform*. https://doi.org/10.3390/rs11060629

Tingzon, I., Dejito, N., Guzman, R. De, Villaveces, J., & Rubio, D. (n.d.). *Mapping New Informal Settlements using Machine Learning and Time Series Satellite Images : An Application in the Venezuelan Migration Crisis*.

Trianni, G., Angiuli, E., Lisini, G., & Gamba, P. (2014). Human settlements from Landsat data using Google Earth Engine. In *International Geoscience and Remote Sensing Symposium (IGARSS)*. https://doi.org/10.1109/IGARSS.2014.6946715

UN-Habitat. (2016). Slum Almanac 2015/2016. Tracking improvement in the lives of slum dwellers. *Participatory Slum Upgrading Programme*, 98.

United Nations Human Settlement Programme (UN-Habitat). (2003). Slums of the World : The face of urban poverty in the new millennium? In *UN-Habitat*.

Urban, W., & Edition, F. (2013). *STATE OF THE WORLD ' S CITIES STATE OF THE WORLD ' S CITIES*.

Using Google Earth Engine to Map Complex Shade-Grown Coffee Landscapes in Northern Nicaragua. (2018). https://doi.org/10.3390/rs10060952

Viswambharan, B. V., & Lenhardt, J. (2019). *Introducing the Spectral Index Library in ArcGIS*. 8–9.

West, D., & Orr, M. (2007). Race, Gender, and Communications in Natural Disasters. *Policy Studies Journal*, *35*, 569–586. https://doi.org/10.1111/j.1541-0072.2007.00237.x