ARDHI UNIVERSITY

CHARACTERIZATION OF URBAN SPRAWL PATTERN IN KILOSA DISTRICT

SINGO, REHEMAROSE KOMBO

BSc Geographical Information Systems and Remote Sensing

Dissertation

Ardhi University, Dar es Salaam

July, 2023

CHARACTERIZATION OF URBAN SPRAWL PATTERN IN KILOSA DISTRICT

SINGO REHEMAROSE KOMBO

A Dissertation Submitted to the Department of Geospatial Sciences and Technology in Partially Fulfilment of the Requirements for the Award of Science in Geographical Information Systems and Remote Sensing (BSc. GIS & RS) of Ardhi University

CERTIFICATION

The undersigned certify that they have read and hereby recommend for acceptance by the Ardhi University dissertation titled **“Characterization of Urban Sprawl Pattern in Kilosa District”** in partial fulfillment of the requirements for the award of degree of Bachelor of Science in Geomatics at Ardhi University.

………………………………	………………………………
Dr. Atupelye Komba	Mr. Gadiel Mchau
(Supervisor)	(Supervisor)
Date…………………	Date…………………

DECLARATION AND COPYRIGHT

I, Singo Rehemarose K. hereby declare that, the contents of this dissertation are the results of my own findings through my study and investigation, and to the best of my knowledge they have not been presented anywhere else as a dissertation for diploma, degree or any similar academic award in any institution of higher learning.

…………………………………..

SINGO REHEMAROSE K.

25525/T.2020

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to God for providing me with the opportunity to finish my dissertation. It was forever been sweet to confide in you.

My sincere thanks go to my supervisors, Dr. Atupelye Komba and Mr. Gadiel Mchau for their unwavering support, time, and guidance in completing this dissertation within the allotted time frame. I would like to express my gratitude to the Geospatial Science and Technology Department staff for their academic and moral support. I truly appreciate all of your assistance; May God include the places where you have sacrificed for my sake.

I wouldn't be grateful if I didn't express my gratitude to my friends, Mbonea Kombo Singo, Hemed Lungo, Nelson Shao, Rehema Sendeu, Godwonder Bwemero and Asedi Niwagila for their assistance, encouragement, and support throughout my academic journey and in completing this dissertation. I am so grateful.

DEDICATION

I dedicate this dissertation to my beloved family; my parents Mr. and Mrs. Singo, who have been there for me every step of the way through my university life. These accomplishments are the result of their fervent prayers, love, and assistance with my education. May the almighty father bless them.

ABSTRACT

The characterization of the pattern of urban sprawl in the Kilosa district between the years 2000 and 2022 is the primary focus of this study. The study's objective is to comprehend the land use and population growth as well as the extent and dynamics of urban expansion during this time period.

This study's methodology involved analyzing satellite imagery and geospatial data (Sentinel 2A) using remote sensing methods such as classification, Shannon entropy and urban expansion intensity index and Geographic Information Systems (GIS). The district's urban distribution over the specified time period is identified and mapped using built-up data from the Global Land Analytics and Discovery (GLADs) dataset.

During the study period, urban sprawl was evident in the Kilosa district, according to the analysis's findings. The compact distribution pattern and relatively low density of the urban expansion are characteristics. The district also experiences infilling urban sprawl pattern. Changes in land use have resulted from the rise of built-up areas, including a decrease in the amount of land available for agriculture and vegetation due to the rise of settlements.

Policymakers and urban planners can benefit from this study's findings, which contribute to a better understanding of the urbanization process in the Kilosa district. This study's findings have the potential to aid in decision-making processes pertaining to the management of land use, the creation of infrastructure, and the preservation of natural resources.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS AND ABBREVIATIONS

ERM	Environmental Resources Management
NBS	National Bureau of Statistics
SHP	Shapefile
TIFF	Tag Image File Format
USGS	United States Geological Survey

CHAPTER ONE

INTRODUCTION

1.1 Background

Urban sprawl also known as horizontal spreading or dispersed urbanization. It is the low-density development beyond the edge of service and employment, which separates where people live from where they shop, work, recreate and educate- thus requiring cars to move between zones (Club, 1998). Sprawl development is influenced by improved quality of life which lead to migration from urban to rural areas. Urban sprawl leads to air pollution due to greenhouse emissions, traffic congestion and loss of agricultural capacity.

Urban sprawl is different from urban growth as urban growth refers to the increase in the size and population of cities and urban areas over time. It is a measure of the expansion of urban areas in terms of land use and population density (Auch, 2004). Urban growth focuses on the growth of the city or suburb in form of highway constructions, surfaces pavement, increasing storm water runoff and quickening concentration times while urban sprawl focuses on the increased proportion of the entire population in these areas (Auch, 2004).

Sustainable development goals of 2030 specify on the sustainable cities and communities (Goal 11) aims at making cities and human settlements inclusive, safe, resilient and sustainable (UNDESA , 2015). Sustainable urban development can be obtained by an efficient land use growth and management by implementing proper planning and urban design (Pradhan, 2017).

This study focuses in Kilosa district, a district in Morogoro region which is experiencing urban sprawl as there are different opportunities for business, agriculture and settlement due to the construction of the SGR (Standard Gauge Railway) which was meant to operate by December 2022 but also a release from the town noises and chaos. Kilosa is in the Phase II of the construction as it is constructed in phases from Dar – es – salaam to Mwanza via Isaka, making it a potential for migration. Currently the standard gauge is in the testing phase (ERM, 2019)

The purpose of the study is to comprehend the nature, spatial extent, and distribution of urban growth within the district over a particular time period, typically from 2000 to 2022. By analyzing satellite symbolism, land use information, and other pertinent datasets, the review looks to recognize and measure the progressions in developed regions, survey the effect of urbanization on the scene, and investigate the fundamental variables driving urban sprawl in Kilosa.

1.2 Statement of the problem

As the town encounters development and improvement, there is increase worry about the examples, patterns, and effects of urban sprawl on the locale's scene, framework, and financial texture. Be that as it may, there is an absence of thorough examination and information in regards to the degree, spatial dissemination, and elements of urban sprawl in Kilosa locale. The district's effective strategies for sustainable development, land management, and urban planning are hampered by this knowledge gap. Therefore, a methodical and in-depth investigation into the characteristics of urban sprawl in the Kilosa district is required. This investigation will provide stakeholders, planners, and decision-makers with valuable insights into how to deal with the challenges and take advantage of the opportunities that come with urban growth in a way that is both balanced and sustainable.

1.3 Objectives

1.3.1 Main Objective

i. To map the urban sprawl pattern of Kilosa District from the year 2000-2022.

1.3.2 Specific Objectives

i. To quantify urban expansion intensity of urban growth in various regions of the Kilosa District, thereby gaining insights into the magnitude of expansion in each area from the year 2017 to 2021.
ii. To characterize the type of urban sprawl within the Kilosa District.
iii. To assess urban expansion dynamics by evaluating how the sprawl occurs over time within the Kilosa District from the year 2017 to 2021.

1.4 Research question

i. How can the urban sprawl patterns within the Kilosa District between the years 2000 – 2022 be effectively characterized using the Shannon entropy model and the Urban Expansion Intensity Index?

1.5 Significance

The analysis of urban sprawl pattern will be useful in planning of the district. As the urban sprawl map will indicate the specific pattern that the town has been dispersed or concentrated whether it is infilling, edge-expansion or it spreads spontaneously. These can be used by urban planners and the local government at large to plan the development of the specific location and how to distribute the available resources according to the patterns.

1.6 Beneficiaries

Diverse stakeholders and organizations benefit from characterizing the pattern of urban sprawl in the Kilosa district. The primary beneficiaries include:

i. Urban planners and local government

   The characteristics of urban sprawl in the Kilosa district provide valuable data for urban planners and local government officials. It helps them comprehend the nature, distribution, and spatial extent of the district's urban growth. With this knowledge, you can make better decisions about how to plan land use, build infrastructure, and manage cities.

ii. Policy makers

   Effective policies and guidelines for urban development, sustainable growth, and land management can be supported by the obtained findings and insights. This makes sure that decisions are made in a fair and informed way.

iii. Academics and Researchers

   A useful dataset for further research and investigation is provided by this topic. In the context of Kilosa, it provides an opportunity to investigate the underlying causes, patterns, and effects of urban sprawl. The results have the potential to add to the existing body of knowledge in environmental studies, urban planning, geography, and other related fields.

iv. Conservation and environmental organizations

   Environmental and conservation groups need to be aware of the pattern of urban sprawl in the Kilosa district. It helps them figure out how urbanization affects biodiversity, natural resources, and ecological systems. The creation and implementation of conservation strategies, the identification of ecologically significant regions, and the promotion of practices that promote sustainable development all depend on this data.

v. Local community

The description of urban sprawl can be helpful to the general public who live in or near the Kilosa district. It enables them to participate in discussions, participate in community planning processes, and contribute to sustainable urbanization efforts by raising awareness of the local growth and development dynamics.

1.7 Study Area

The study area is Kilosa district, one of the six districts in Morogoro region. It is bordered by Manyara region in the North, Mvomero in the east, Kilombero in the south and Dodoma in the West. Its lies between latitudes of 6° 7' 17" S to 7° 31' 56" S and longitudes of 36° 21' 19" E to 37° 9' 31" E. Kilosa district covers a total area of 14,918 square kilometers with a population of 555,453 as at 2022 census results (NBS, 2022). The wet season is usually hot and cloudy while the dry season is warm and clear. Its temperature varies from 20.3 C to 29.3 C and is elevated at a range between 5059 feet to 5571 feet. The prominent economic activities are farming of both cash and food crop and livestock keeping. The indigenous tribe are the Kaguru, Sagara and Kidunda (Boudreaux, 2017). The selected region experiences urban sprawl as areas in the region have built areas. Figure 1.1 shows the study area description.

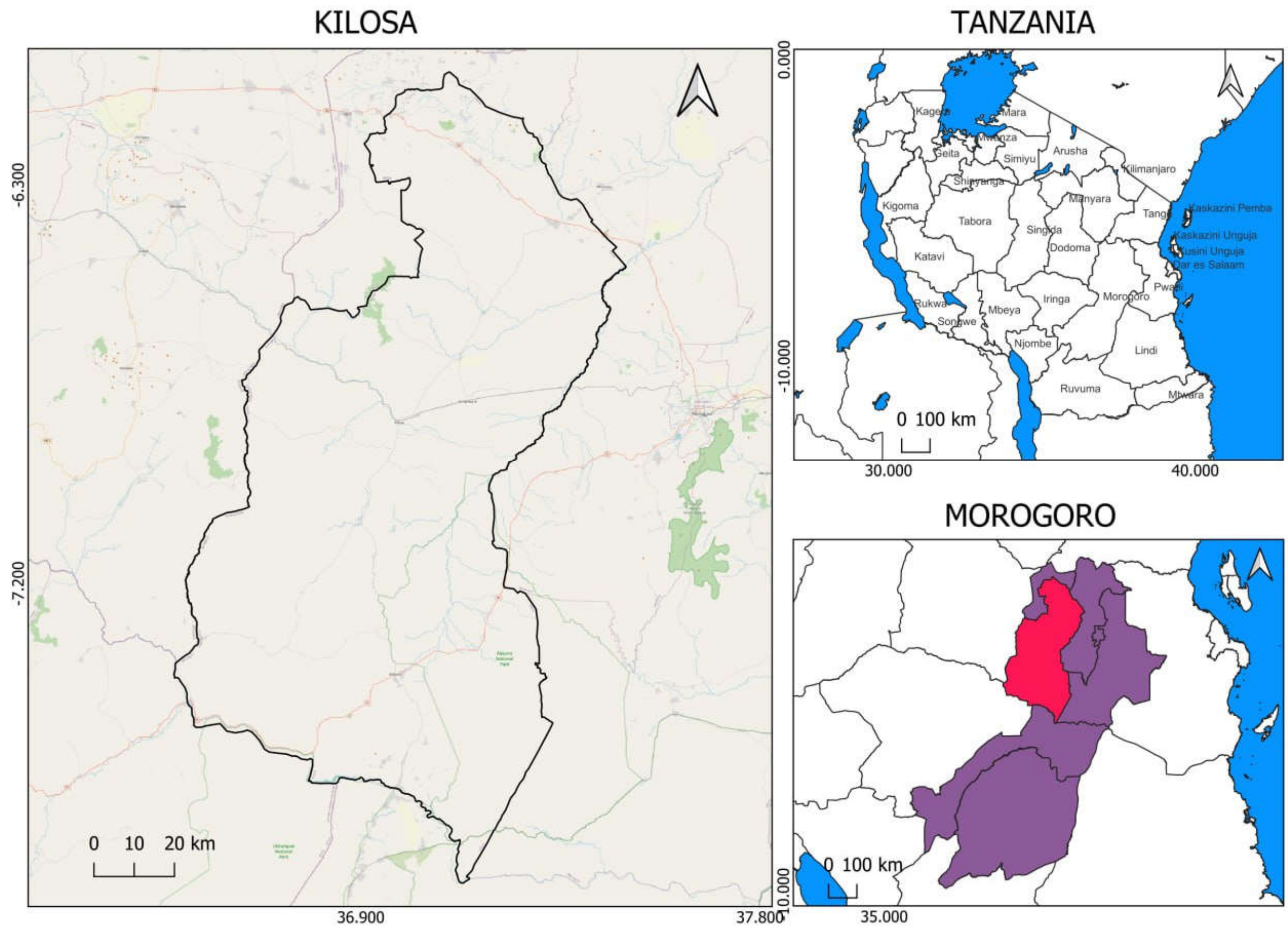


Figure 1.1: Location Map of Kilosa District

1.8 Scope and Limitation

This research covers the area Kilosa district at Morogoro region. The research will base on the analyzation of urban sprawl pattern from the year 2000-2022 using the urban expansion intensity index and Shannon entropy model method. The 20-year time span gives a significant period to catch and examine the drawn-out patterns in Kilosa town. It makes it possible to conduct a comprehensive analysis of how urbanization and shifts in land use have changed over time. The research does not consider the landscape metrics method. It only focuses on the evaluation of the urban spatial expansion difference in Kilosa and the determination of the spatial dispersion or concentration of a town.

.

1.9 Dissertation Organization

This paper comprises of five parts, making sense of exhaustively every one of the techniques, standards, techniques and the outcomes acquired in creating urban sprawl map and measurement of urban expansion and Shannon entropy model metrics for Kilosa District.

Chapter 1 makes sense of the foundation of the review, which brought about the issue. The goals, research questions, importance and recipients of the exploration, along with the portrayal of the review region.

Chapter 2 reviews the study. It provides an explanation of all studies and literature that have been conducted regarding the creation of urban sprawl and the required measurements.

Chapter 3 covers every one of the strategies and procedures associated with the review.

Chapter 4 analyzes and discusses the results. It shows the outcomes got in this study and the outcomes got by past examinations.

Chapter 5 contains the study's conclusion and recommendations.

CHAPTER TWO

LITERATURE REVIEW

2.1 Overview

This chapter reviews important literature on the understanding of urban sprawl through past records. It entails how urban sprawl is explained by other authors through literature. It explains the methods that will be used in the research, it will also establish the research gap. It provides the review of literatures to Urban sprawl.

2.2 Urban Sprawl

The phenomenon urban sprawl is described as " excessive city growth” (Brueckner, 2000). Urban sprawl can also be defined as the unplanned, uncontrolled, and uncoordinated single-use development that does not provide for an attractive and functional mix of uses and/or is not functionally related to surrounding land uses and which variously appears as low density, ribbon on strip, scattered, leapfrog, or isolated development (Pendall, 1999). The literature provides: "Spread" does not necessarily refer to all forms of development, rather, it can be defined as leapfrog development/ unlimited outward extension of development/fragmentation of powers over land use among many small localities/dominance of transportation by private automotive vehicles/lack of centralized planning or control of land uses (Downs, 1999).

They have emphasized particular aspects in the other references in accordance with their perspectives and objectives. Under market conditions, sprawl is defined as the physical pattern of low-density expansion of large urban areas, primarily into the agricultural areas surrounding them (Habibi, 2011). It is a significant problem in many parts of the world, including Kilosa District in Morogoro region situated in Tanzania.

2.3 Characteristics of Urban Sprawl

Sprawl is the forefront of urban growth and indicates that there is little planning oversight over land subdivision. Development is sporadic, dispersed, and prone to discontinuity. It leapfrogs over land, creating agricultural reservation.

Urban sprawl is characterized by rising income inequality, job insecurity, a decline in central cities, rising housing costs, lengthy commutes, environmental issues, the extinction of species, the loss of farmland, a sense of isolation, intolerance, psychological disorientation, and even murder and mayhem (Gordon, 2000). Four sprawl factors that can be measured and examined: densities of housing; mix of homes, businesses, and services; strength of downtowns and activity centers; and the street network's accessibility (Ewing, 2008). In addition, it is believed that inaccessibility and a lack of functional open spaces are the most important indicators (Ewing, 2008).

Sprawl is an ambiguous metaphor (Galster, 2001). Sprawl has been defended on the grounds of choice, equality, and economics while it has been condemned on the grounds of aesthetics, efficiency, equity, and the environment. Density, continuity, concentration, clustering, centrality, nuclearity, mixed uses, and proximity are the eight dimensions of sprawl established. The following are the established urban sprawl characteristics (Angle, 2007);

i. The rise of "endless" cities and a walkable range
ii. The determined decrease in metropolitan densities and the rising utilization of land assets by metropolitan occupants
iii. The ongoing process of suburbanization and the decreasing proportion of the population that lives and works in major cities
iv. The fragmentation of open space within and around cities and the diminished continuity of built-up areas; as the spaces in between cities' finger-like extensions are filled in, they become more compact.

The policies can be developed in a useful way by measuring urban sprawl. The causes and effects of sprawl should be examined as a pattern or method. Sprawl is exemplified by the proportion of urban employment located more than ten miles from the central business district, which shows the degree of employment decentralization in urban area. Sprawl is said to be growing if land is used up faster than the population grows (Fulton, 2001). This idea of sprawl is only related to density (Couch, 2007).

2.4 Causes of Urban Sprawl

There are some things that set sprawl apart from other kinds of urban growth, and different experts have different theories about what causes it. The expansion of income and population, enhancements to transportation systems, a wide range of user options, and competition for land all contribute to urban sprawl (Mieszkowski, 1993). Other indicators, such as high taxes, crime, deteriorated infrastructure and the number of educational facilities in central areas encourage urban sprawl. The so-called "fiscalization of land use" is the basis for a third theory of what might lead to more decentralization in urban areas. Sprawl is brought on by an increase in income, population, and commute costs. As the population grows, so do the requirements for space (Brueckner, 2000). From a different angle, as people's incomes rise, they tend to live in bigger buildings, and as a result, roads and infrastructures will be built. Another factor that has been mentioned in studies is the value of the land. Urban sprawl is caused by four factors: preference among consumers, advancements in technology, subsidies, and public and quasi-public goods (Ewing, 1997).

The following are the general causes of urban sprawl based on different factors.

i. Economic expansion; Economic expansion and rising income Land prices Subsidies
ii. Demographic Housing; A rise in the number of people, more space per person and diversity of choice
iii. Transportation; Possession of a private automobile, low transportation costs and the availability of roads
iv. Inner city problems; Infrastructural damage from high taxes, fewer public facilities, fewer apartments and a lack of open space contribute to social issues.
v. Other; Innovation in technology, public infrastructure, and facilities

2.5 Patterns of Urban Sprawl

Urban sprawl patterns can be described as “edge-expansion sprawl," "infilling sprawl," and "outlying sprawl". It is still possible to differentiate between the three types of urban sprawl despite the fact that the methods used to describe them are generally the same in principle. All the spread type depictions begin from the landscape transformation process (Forman, 1995).

The majority of edge-expansion sprawl occurs outside of cities. It is a newly developed urban area that extends beyond the boundaries of existing urban areas. The process of converting non-construction areas that are surrounded by urban areas into urban land is known as "infilling sprawl," and it usually takes place within the existing built-up areas of cities. Outlying sprawl is defined as newly formed urban areas do not share a direct spatial connection with previously established urban areas (Zhi-qiang, 2011).

2.6 Advantages of Urban Sprawl

Urban sprawl can be very beneficial especially in areas that have only started blooming. The following are the advantages of Urban Sprawl;

i. It results in the preparation of new infrastructure for travel (Habibi, 2011).
ii. Expanded Housing Options: One of the potential benefits of sprawl looked into increased affordability of housing and racial equality in housing opportunities (Kahn, 2001). A greater variety of housing options are frequently made available as a result of urban sprawl. This may include larger residences, suburban communities, and lower housing costs when compared to urban areas that are more centrally located. It gives individuals and families options for the kind, size, and location of their housing (DiPasquale, 1999).
iii. Enhanced Standard of Living: Residents of sprawling urban areas may be able to gain access to larger homes, green spaces, and a more tranquil living environment away from the bustle and noise of urban centers. For some people who would rather live in the suburbs or in the country, this might make them feel like their quality of life is better (Ewing, 1997).
iv. Development and Growth in the Economy: By facilitating the expansion of commercial and industrial activities, urban sprawl can facilitate economic growth. It frequently results in the construction of shopping malls, business parks, and employment opportunities on the outskirts of cities. This has the potential to bring in investments, stimulate the local economy, and create jobs (Gordon, 1997).

2.7 Disadvantages of Urban Sprawl

The majority of viewpoints focus on the negative outcomes of urban sprawl, despite the positive aspects. A portion of these is about issues in the core area. Issues like concentrated poverty, poor educational facilities, and a lack of financial resources. In transportation, sprawl increases commute times, travel times, and congestion. Households are able to travel greater distances to gain access to better housing, educational facilities, and employment opportunities, which is exacerbated by infrastructure access. Costs associated with water, electricity, and parking are a few examples. Additionally, urban sprawl may result in increased energy consumption, pollution, and the loss of land. The psychological and social costs of living in these areas are numerous. There are two distinct categories of psychic costs: access restriction and environmental restriction. There are three distinct levels of cities: new districts, urban areas, and suburban areas. Sprawl has a negative impact on historic centers and cities by bringing about new issues that make them less competitive. It results in an increase in travel distance, the removal of agricultural land, and a decrease in social interactions in new districts. The third level of outcomes that have an impact on urban areas is an increase in pollution, an increase in energy consumption, damage to environmental resources, and increased costs for infrastructure (Habibi, 2011).

2.8 Methods of Urban Sprawl

Urban Sprawl can be measured by the use of geographical information systems (GIS) and Remote sensing by using the following methods;

i. Urban Expansion Intensity Index

   The Urban Expansion Intensity Index, or UEII, can be utilized to quantitatively assess the urban spatial expansion difference. In addition, UEII could be used to identify the preferences of urban growth and compare the rate or intensity of changes in urban land use over a specific time period (Manesha, 2021).

ii. Shannon Entropy Model

   Shannon entropy is frequently used to determine a town's concentration or dispersion in space. The effective entropy model utilizes the integration of ArcGIS software to reveal the configuration and orientation of spatial patterns and spatial variables, as well as assess the distribution of geographical phenomena (Yeh, 2001)

iii. Land scape metrics.

   Using the well-known FRAGSTATS software, a variety of landscape metrics are used to examine the spatial and temporal trends of urban growth patterns and changes in land use (Manesha, 2021).

This dissertation will look at Urban Expansion Intensity Index method and the Shannon Entropy Model as a measurement of Urban sprawl. These two methods were picked because of their:

i. Relevance to Urban Sprawl: The Shannon entropy model is a generally involved measure in concentrating on urban sprawl as it gives experiences into the dissemination and scattering of land use designs. It aids in quantifying a region's degree of urbanization and spatial heterogeneity. The study can effectively examine the extent and characteristics of urban sprawl in the Kilosa district using this model (Yeh, 2001).

ii. Validity of previous research: The utilization of the Shannon entropy model and the Urban Expansion Intensity Index in past examinations has laid out their legitimacy and unwavering quality in catching urban sprawl elements. The study can contribute to the body of literature on urban sprawl characterization by adopting these established measures (Yeh, 2001).

2.9 Urban Expansion Intensity Index (UEII)

Due to the rule of urban deriving factors (such as road network, population density, slope, and economics) and their spatial impacts during the urbanization process, urban expansions will differ in each region and in each direction, a phenomenon known as preference of urban growth. Urban Expansion Intensity Index can be utilized to quantitatively survey and dissect the urban spatial development contrasts. In addition, Urban Expansion Intensity Index can be utilized to identify the preference for urban growth during a particular time period. The Urban Expansion Intensity Index compares the speed or intensity of urban land use change over various time periods and reflects the probable future direction and potentials of urban expansion (Abubakr, 2014).

The UEII standard is divided as follow:

i. 0 to 0.28 is slow development
ii. 0.28 to 0.59 is low-speed development
iii. 0.59-1.05 is medium-speed development
iv. 1.05-1.92 is high-speed development
v. >1.92 is very high-speed development.

The Urban Expansion Intensity Index for overall study area, each temporal span and each zone is calculated using formula as following:

$$\text{UEII}=\left[\frac{ULA_{ib}-ULA_{ia}}{t}\right] / TLA_i \times 100 \tag{2.1}$$

Where "UEII_{it} is the annual average urban expansion intensity index of (*i*th) zone in time period (t) ULA_{ia} and ULA_{ib} are the quantity of built-up area at time periods *a* and b in (*i*th) spatial zone, respectively. TLA_{i} is the total area of (*i*th) spatial zone" (Al-Sharif, 2017)

2.10 Shannon Entropy Model

Urban sprawl phenomena are extensively studied using this method. Any geographical variable can be analyzed and evaluated using Shannon's entropy model, which is a useful measure of spatial concentration or dispersion. This method can analyze spatial variables within a GIS and reveal the configuration and orientation of spatial patterns (Sudhira, 2004); (Yeh, 2001).

The entropy value indicates the degree of urban sprawl. The value of relative entropy ranges from 0 to 1. The compact distribution of urban areas is represented by a value of zero, while the dispersed distribution of urban areas is represented by values close to one. In this way, higher entropy values show higher spread events. The following equation (2.2) is used to determine the values of the relative Shannon entropy in this study:

$$Hn = \sum_{i}^{n} P_i \log\left(\frac{1}{P_i}\right) / \log(n) \tag{2.2}$$

Hn=, where n is the total number of zones and Pi is the probability or percentage of the variable occurring within zone i (i.e., percentage of urban area in ith zone determined by urban area in ith zone/zone area).

2.11 Applications of Remote Sensing in Urban Sprawl Studies

Unfortunately, the estimation of urban sprawl requires expensive and time-consuming conventional surveying and mapping methods, and such data is unavailable for the majority of urban centers, particularly in developing nations. Consequently, GIS and remote sensing-based mapping and monitoring of urban sprawl and growth are receiving increased research attention (Epstein, 2002).

Because of its low cost and advanced technology, remote sensing is increasingly being used to study urban sprawl (Haack, 2006); (Sudhira, 2004). Urban change detection using remotely sensed images has been the subject of extensive research for nearly three decades (Haack, 2006); (Yeh, 2001); (Gomarasca, 1993).Either a post-classification or image-to-image comparison has been used to back up these studies.

Most of the time, the impervious (built-up) area is used to measure urban sprawl (Epstein, 2002); (Torrens, 2000). In this context, the term "impervious area" refers to the area that is comprised of residential, commercial, and industrial complexes, as well as paved roads, markets, and other infrastructure. The impervious area, which can be derived from either a physical survey or remotely acquired data, are used as the primary criterion for quantifying urban sprawl.

The area of impervious surfaces can be estimated or measured using a variety of methods. Using heads-up digitizing, impervious surface features can be manually extracted from remote sensing images, which is the most labor-intensive and expensive method but also the most accurate. Despite the fact that digitizing takes longer and is less accurate, point sampling can be used instead. Supervised, unsupervised, and knowledge-based expert system approaches to remote sensing pattern recognition (Stefanov, 2001); (Greenberg, 1997). Measurements of impervious area and urban sprawl (Mundia, 2005) have recently been made. These necessitate expertise in processing and analyzing remote sensing data of moderate to high resolution. Many local planners and decision-makers lack access to these data and analytical capabilities, particularly in developing nations.

Numerous studies on urban sprawl have utilized statistical methods, remote sensing, and geographic information systems (Sudhira, 2004); (Cheng, 2003); (Lo, 2001). Numerous developed nations have attempted urban growth studies (Epstein, 2002); (Torrens, 2000); (Batty, 1999). The relationship between the percentage of impervious area and various urban development parameters like road density, population density, land use type, and development unit size has been determined using statistical techniques like multivariate regression (Sudhira, 2004); (Cheng, 2003); (Lo, 2001).The urban sprawl phenomenon has been quantified, monitored, modeled, and predicted as a result of the convergence of

. GIS and database management systems. The pattern of urban sprawl must be identified, quantified, and statistically summarized in order to be characterized. The appropriate spatial unit used in this analysis is the appropriate scale of urban sprawl characterization. The representation of this phenomenon in terms of statistical parameters and indices like Shannon entropy, Patchiness, and so on is referred to as the statistical summarization of urban growth pattern. There are now metrics that can describe the pattern of the landscape and quantify urban growth and its spatial distribution. Forest patches are studied using the landscape pattern metric (Civco, 2002); (Trani, 1999) and identifying the pattern of urban sprawl in clusters of villages (Sudhira, 2004). Because there are only a few primary measurements that can be made from patches (patch type, area, edge, and neighbor type), the majority of the indices are correlated with each other. From these primary metrics, all other metrics are then derived. In order to characterize landscape properties in terms of their spatial distribution and change, GIS aids in the calculation of landscape metrics like patchiness and density at the landscape level (Sudhira, 2004); (Civco, 2002); (Trani, 1999).

2.12 Image Classification

The process of sorting pixels into a limited number of distinct classes, or data categories, based on their data file values is known as classification (Lillesand, 2008) or the process of extracting information from multi-band raster images like satellite images is known as image classification. Density slicing, which uses just one band, or multi-spectral classification, which employs multiple bands, are both methods of classification.

The decision of strategy for picture characterization generally relies upon the research targets, the idea of the image, and the degree of detail or exactness expected for a particular application (Lillesand, 2008). There are extensively two strategies for picture characterization, which are:

i. Supervised Classification
ii. Unsupervised classification

2.12.1 Supervised Classification

The supervised classification is based on external knowledge of the area of interest because it is dependent on their prior knowledge of the location and identity of the land cover in the image (Mather, 2004). Following the user's input, the supervised classification is a commonly used classification.

There are a number of advantages to using the supervised classification method for classification, which has been applied in this study. First, because the analyst has complete control over the classes that will be used in the final classification, it is easier to compare it to other classifications because they use the same classes. Second, the analyst does not have to deal with the problem of matching spectral classes to informal classes because it is already addressed during the selection of training areas, specific areas on the image of known identity. Third, the resulting classification is tied to specific areas on the image of known identity (Campbell, 2002).

2.12.2 Unsupervised Classification

Due to the fact that it is based on a computed algorithm, this kind of classification does not require prior knowledge. In this case, the user only needed to choose the number of classes and the computer will perform accordingly. Because it uses an unbiased geographical evaluation of pixels, this method is useful for evaluating areas where you have little or no prior knowledge of the site. However, it is not affected by covariation in object spectral signature variation.

2.12.3 Machine-learning Algorithm

Machine learning algorithms like Support Vector Machine (SVM), neural networks, K-Nearest Neighbors (KNN), and random forest are utilized in image classification tasks. In this research, the random forest was used for the classification.

Random Forest – It is a combination of machine learning algorithm which are then combined with a series of tree classifiers. Each tree casts a unit vote for the class that is most popular, and the final sort result is obtained by combining these results. The classification accuracy of Random Forest is high, it can handle noise and outliers well and it has never been overfitted. One of the most widely used research techniques in the data mining and biological information fields is Random Forest (Yanli, 2012).

Algorithms for machine learning have a number of advantages over conventional/ traditional methods (Yanli, 2012):

i. Capability to handle relationships that are complex and nonlinear: They succeed in catching complex and nonlinear relationships in data. They can detect patterns and interactions that may not be easily perceivable using traditional methods. This flexibility allows for more accurate and robust predictions or classifications in various domains.
ii. Capability of Adapting to Massive and High-Dimensional Datasets: They deal with enormous and high-layered datasets proficiently. It can process and dissect huge measures of information, including assorted elements or factors, without critical computational limits. This adaptability empowers specialists to extract important bits of knowledge from huge information sources.
iii. Capability of Adapting to Changing Conditions: It can adjust and refresh their models as new information opens up. They can learn from constant feedback,

which helps them perform better and adjust to new situations. This adaptability is especially helpful in dynamic settings where conventional methods may have trouble keeping up with changing trends or patterns.

CHAPTER THREE

METHODOLOGY

3.1 Overview

In this chapter, the methods used, techniques applied and data collected in the study are explained. The chapter describes all the processes done from data collection, data manipulation to creation of urban sprawl map and analysis by Shannon-entropy and urban expansion. The whole process is summarized step by step in figure 3.1.

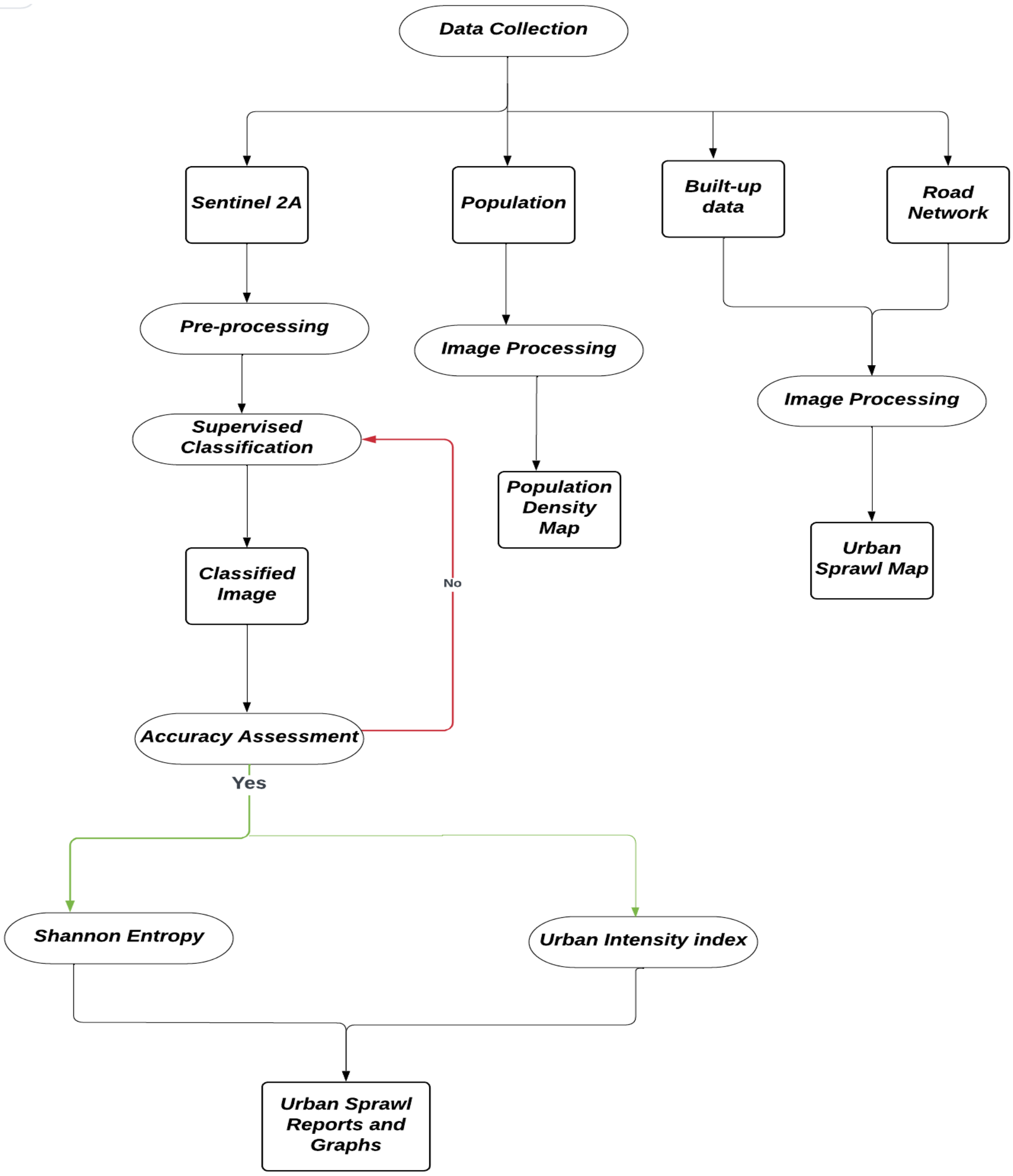


Figure 3.1: Urban Sprawl Flow Chart

3.2 Software Used

The software used were selected based on their availability and required licenses. Table 3.1 shows software package needed for creation of urban sprawl map and analysis using Shannon entropy model and urban expansion.

Table 3.1: Software package

S/N	SOFTWARE/TOOLS	FUNCTION
1	Quantum GIS	Creation of Urban Sprawl Map
2	R Studio	Creation of urban sprawl map
3	ArcGIS	Calculation of Shannon's entropy model and urban expansion
4	Blender	Creation of 3-D Population Density Map

3.3 Data Collection

The data used to create the outputs came from a variety of sources. Global Land Analytics and Discovery (GLADs) provided the built-up data for the years 2000 to 2022. OpenStreetMap provided the road network data, while the National Bureau of Statistics provided the population data. The Humanitarian Data Exchange provided the data for the administrative boundaries. Furthermore, Sentinel information was downloaded from the US Land Study (USGS). Table 3.2 provides a summary of the data used in this study.

Table 3.2: Characteristics of the data used in this study

DATA TYPE	FORMAT	SOURCE	YEAR OF PRECISION	USE
Administrative boundaries	SHP	Humanitarian Data Exchange	2022	To show the boundary of the region (mapping component)
Built up	TIFF	GLADs	2000-2022	To create an urban sprawl map
Population	SHP	Kontur dataset	2022	To show the population of Kilosa district
Road network	SHP	Open Street Map	2022	To show regions accessing transportation
Sentinel 2A	TIFF	USGS	2017-2022	Classification and measurement of Shannon entropy model and urban expansion

3.4 Image- Preparation

The sentinel data were improved by preserving undesirable details and enhancing essential characteristics for further processing. It included the redresses of Sentinel images in order to increment picture goal, stacking/mosaicking of the groups and projection.

3.4.1 Layer stacking

Using the layer stack tool in Quantum Gis software, five bands from Sentinel 2A were combined into a single band image to improve classification accuracy. By layer stacking, the study can take advantage of the complementary information present in different spectral bands. This can enhance the accuracy of classification algorithms by providing more discriminative features for distinguishing between different land cover classes. Sentinel 2A is equipped with band 2,3,4,8,11.

3.4.2 Atmospheric corrections

Removal of haze and noise was conducted utilizing the haze reduction and noise reduction in QGIS respectively . This was done to eliminate the atmosphere's scattering and absorption effects. The main purpose is to facilitate the surface reflectance characterization of surface properties.

3.4.3 Re-projection

In order to match the sentinel images' coordinates with those of the Area of Interest (AOI), the corrected stacked sentinel images from 2017, 2018, 2019, 2020, 2021 and 2022 were transformed into the local coordinate system of UTM zone 37S from WGS 84 using datum Arc 1960. Reprojection of Sentinel-2A images makes sure that spatial analysis, integration with other datasets, and visualization are consistent, compatible, and accurate. This makes it possible to do a more accurate and comprehensive analysis of the study area.

3.5 Training Samples

The shapefile of the study area was used to link the QGIS software with Google Earth to collect the training samples. For each class about 100 training samples were collected. The Land Cover Classification Scheme, which reflects a particular kind of land cover based on the main features, was the basis for the use of four land cover classes.

i. Built-up area
ii. Vegetation
iii. Bare land
iv. Water

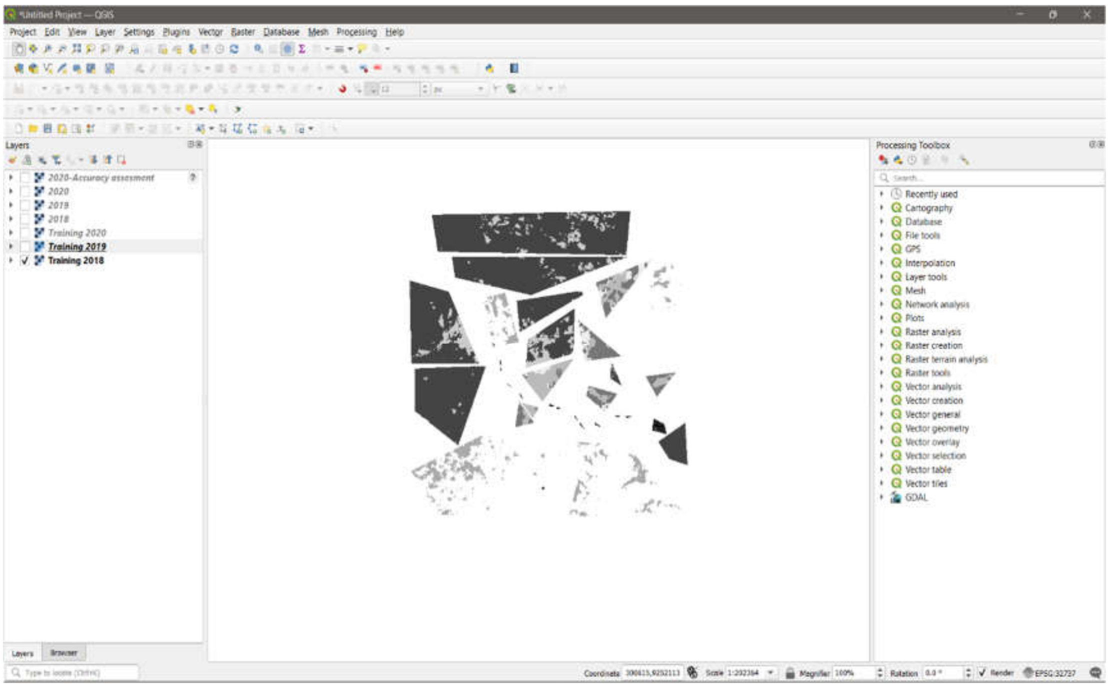


Figure 3.2: Selection of Training Samples for the year 2018

3.6 Image classification

The technique utilized for the characterization of the review region was supervised classification and the calculation utilized was the random forest. This process was finished in QGIS. The classes utilized in arrangement were in portrayed in the determination of preparing, are;

i. Built-up area
ii. Vegetation
iii. Bare land
iv. Water

The classes were then reclassified to make it easier to study urban sprawl. This led to the creation of a built-class. In order to accurately examine and comprehend the patterns of urban expansion, this classification was necessary.

3.7 Accuracy assessment

Accuracy assessment of Kilosa District and kappa values for the images for 2017, 2018. 2019, 2020, 2021 and 2022 was done after being reclassified in order to survey the viability of the pixels that were tested into right land cover classes. The method used to obtain these results was error-matrix. When assessing the reclassified images, the coordinates extracted from Google Earth served as the reference coordinates. The software QGIS was used throughout.

3.8 Preparation of Maps

i. Urban sprawl map

Data was imported in the QGIS software and processing took place. The data required was built up and road networks. Creation of the buffer around the map and cropping it to the best fit. In addition, the map was enriched with urban features such as roads and paths from OSM data.

ii. Population density map

A dataset of hexagonal grids was imported into Blender software to produce a three-dimensional population density map. The dataset was used to create a three-dimensional visual representation of the population density map within Blender. The map was exported from Blender and imported into QGIS after the desired effect was achieved. The final population density map was created using QGIS, incorporating Blender-created three-dimensional elements.

3.9 Calculation of Urban index intensity

The selected classes were narrowed down to the most required class, “built- up class” for all the years, 2017, 2018, 2019, 2020, 2021 and 2022. For a number of reasons, the years 2017, 2018, 2019, 2020, 2021 and 2022 were chosen for the research’s Shannon entropy and urban expansion intensity index calculations. By choosing these years, it was possible to use data from relatively recent years, giving a more current understanding of the patterns and trends of urban sprawl in the Kilosa district. Over a sufficient amount of time, the four-year period was able to capture and analyze changes in urbanization and land development. Also, by looking at back-to-back years, it was feasible to survey the worldly elements of endless suburbia, recognize any arising patterns or examples, and comprehend the movement of urbanization in the Kilosa region. The years 2017, 2018, 2019,2020, 2021 and 2022 were chosen for the analysis because of its availability (satellite imagery and related datasets). The selection of built-up class is shown in figure 3.3.

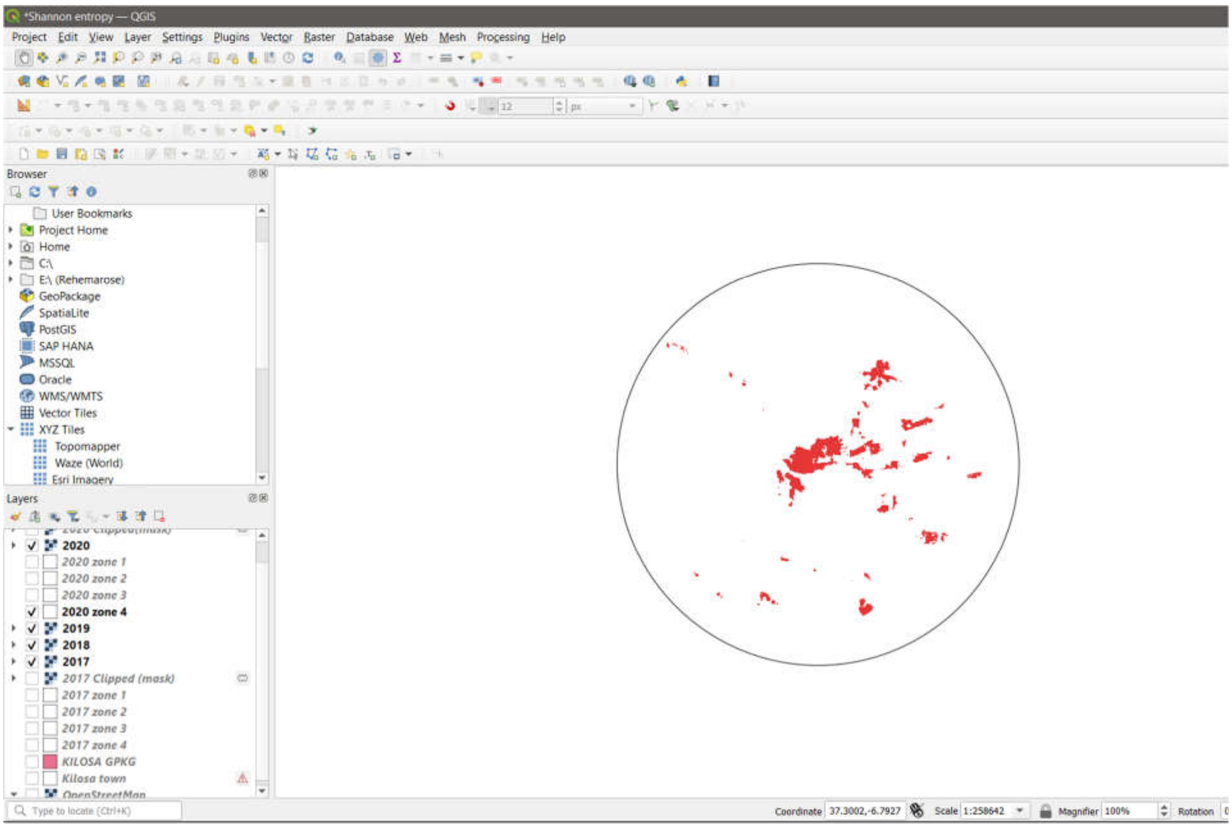
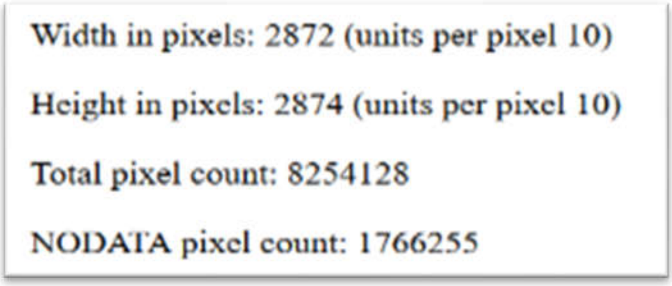


Figure 3.3: Selection of built-up class

3.10 Total Pixel count calculation

For the analysis, the total area affected by urban sprawl, expressed as square meters, was gathered. With the available built-up class in meter square, the count and class were used to calculate the unbuilt area. For each of the four years under consideration, this procedure was carried out once more. Figure 3.4 provides a visual representation of this procedure.


Width in pixels: 2872 (units per pixel 10)

Height in pixels: 2874 (units per pixel 10)

Total pixel count: 8254128

NODATA pixel count: 1766255

Figure 3.4: Total pixel count(m^2) of urban sprawl in Kilosa District

The pixel data was then transferred to Excel for further analysis, including the urban index calculation. This required locating parts of the dataset that had not been affected by urban sprawl for the previous four years. Table 3.3 provides an illustration of this procedure.

Table 3.3: Embedding of total pixel count in excel

Year	**Built-up area(m^2)**	**Total Pixel Count(m^2)**	**Non-built-up area(m^2)**
2017	18429900	825412800	806982900
2018	18418500	825412800	806994300
2019	18418500	825412800	806994300
2020	21174300	825412800	804238500
2021	23665800	825412800	801747000

Where; $Non - built\ up\ area = Total\ Pixel\ Count - Built - up\ area$

3.11 Calculation of Shannon Entropy Model

Around the Kilosa District, the area of interest, specific zones were designated for the selected years. Each year, four zones were chosen to represent both urban and non-urban sprawl areas. The Shannon entropy model's calculation was made easier by these selected zones, which were instrumental in capturing the characteristics of both urban and non-urban expansion. The circles represent the selected zones. The smallest circle is zone 1 and the largest one is zone 4. Figure 3.5 provides a visual representation of this procedure.

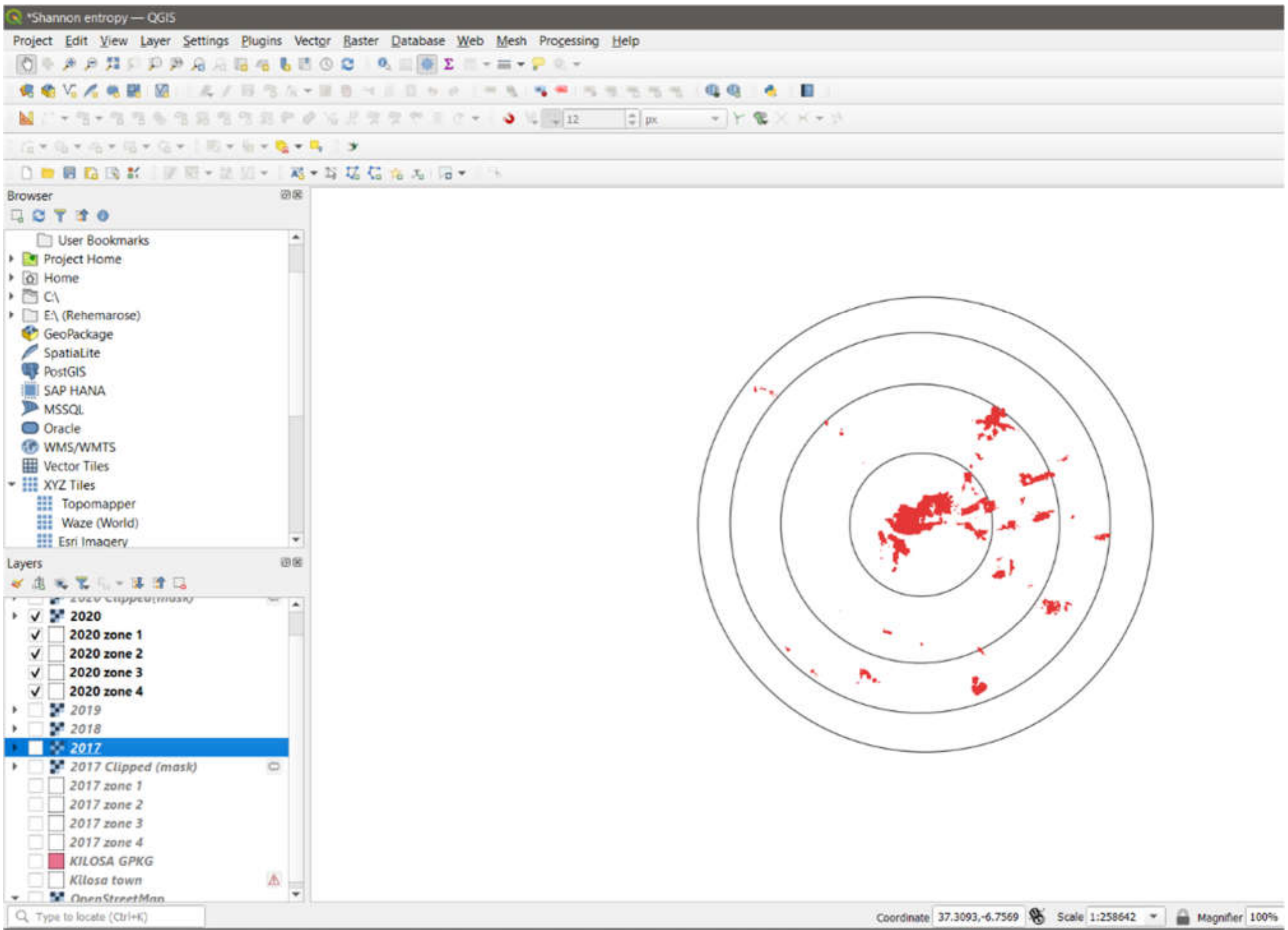


Figure 3.5: Creation of zones for Shannon entropy calculation

The zones where then clipped by mask layer to obtain the built-up square meters of each region. This process is shown in figure 3.6 and figure 3.7.

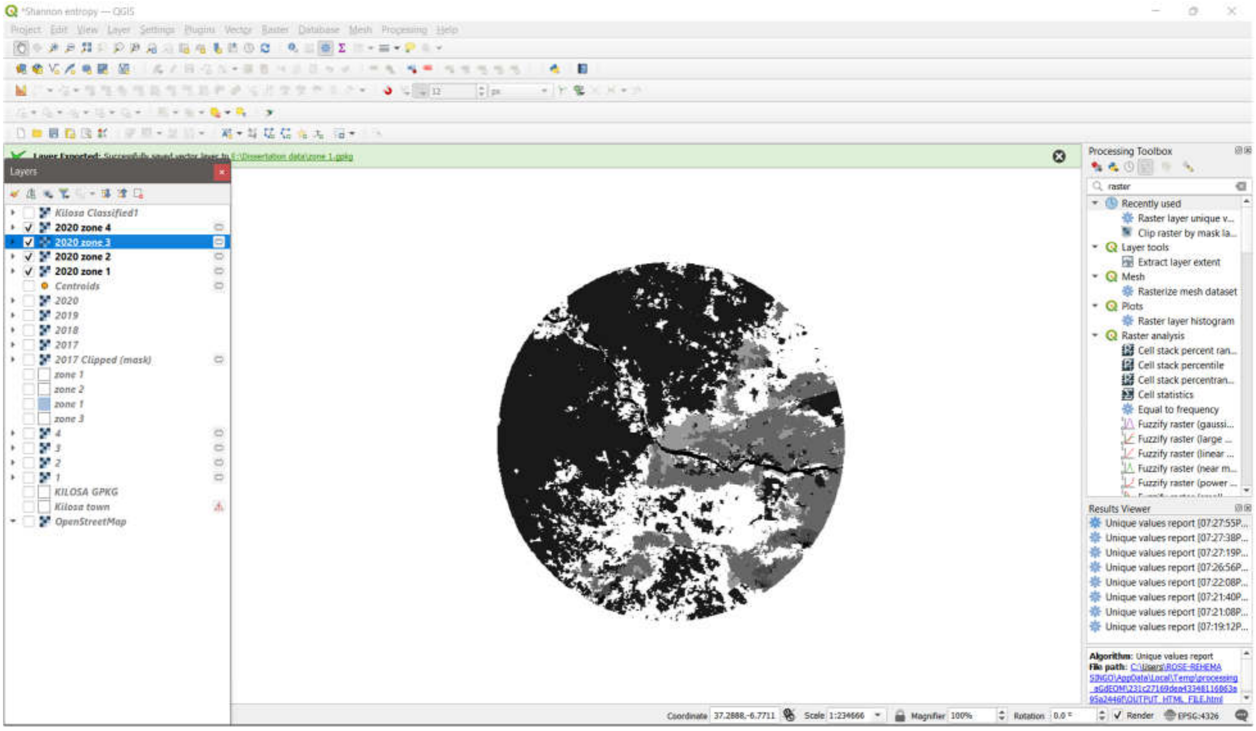


Figure 3.6: Clipping by mask layer for the year 2017

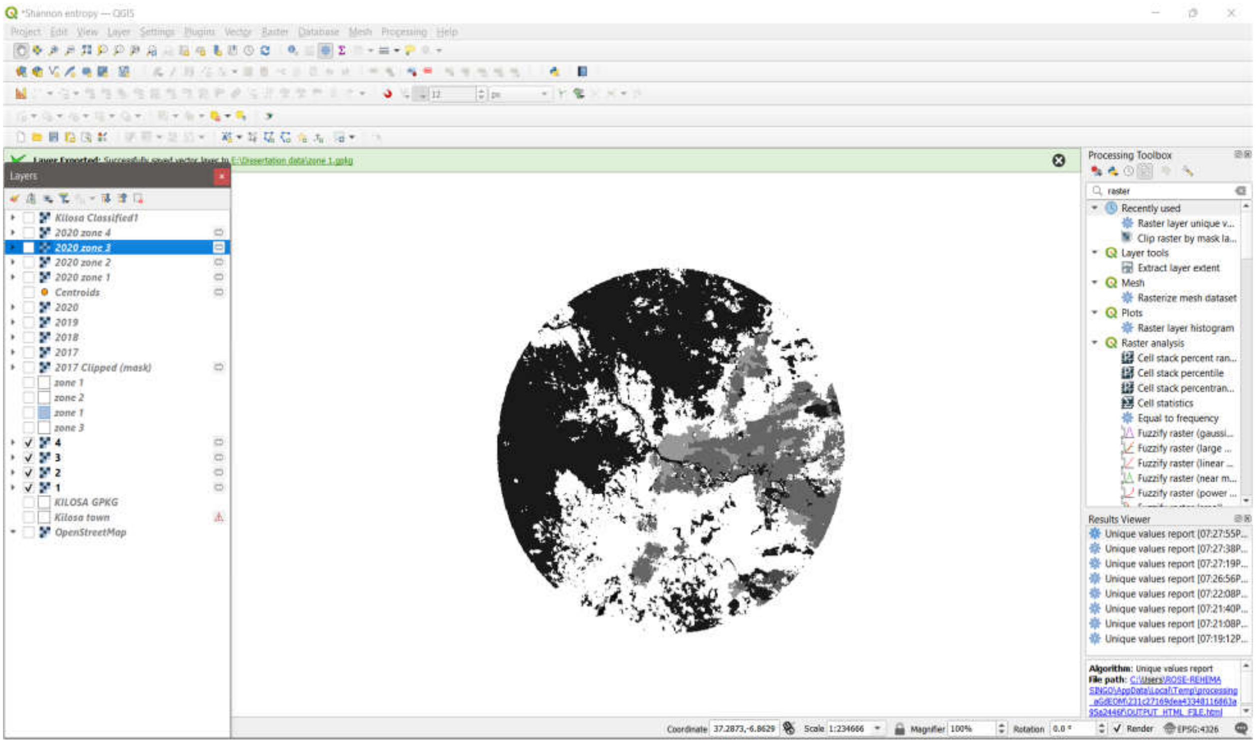


Figure 3.7: Clipping by mask layer for the year 2020

The built-up square meters where then obtained for each zone in each year by the use of raster layer unique value.

The width, height, total pixel count and no data pixel count were obtained so as to calculate the Shannon entropy model for Kilosa district. Table 3.4 displays the above information.

Table 3.4: Requirements for the calculation of Shannon Entropy Model

Year	**Zone**	**Width (Units per pixel 10)**	**Height (Units per pixel 10)**	**Total Pixel count(m^2)**	**No Data pixel count**
2017	**1**	589	589	346921	73403
	4	3146	3148	9903608	3430786
2020	**1**	589	589	346921	73403
	4	3146	3148	9903608	3430786

The above information were then used for the calculation of Shannon entropy model for Kilosa District.

Table 3.5: Unbuilt region in meters for the year 2017

2017	**Built-up (m^2)**	**Total pixel count (m^2)**	**Non-Built up (m^2)**
Zone 1	8063100	34692100	26629000
Zone 2	11033200	231344100	220310900
Zone 3	18392700	706762200	688369500
Zone 4	18429900	990360800	971930900

Where; $Unbuilt\ area = Total\ pixel\ count - Built\ up\ area$

Table 3.6: Unbuilt area in meters square for the year 2020

2020	Built-up (m^2)	Total pixel count (m^2)	Non-Built up (m^2)
Zone 1	9069100	34692100	25623000
Zone 2	13599100	231344100	217745000
Zone 3	21005600	706762200	685756600
Zone 4	21174300	990360800	969186500

The above information was then used to obtained the absolute and relative Shannon entropy for the year 2017 and 2020.

The constituted values used in calculation of Shannon entropy are defined as follows;

Pi is defined by Total area
Built-up area

1/Pi is defined by I/ Total area
Built-up area

3.11.1 Calculation of absolute Shannon entropy

Table 3.7: Calculation used to obtain absolute Shannon entropy for the year 2017

Zone	Total area	Built-up (m^2)	Pi	1/Pi	Ln(1/Pi)	Pi*Ln(1/Pi)
1	34692100	8063100	0.232419	4.302576	1.459214	0.339148
2	231344100	11033200	0.047692	20.968000	3.042997	0.145125
3	706762200	18392700	0.026024	38.426230	3.648740	0.094954
4	990360800	18429900	0.018609	53.736630	3.984095	0.074141
					SUM	**0.653370**

From the table 3.7 above;

i. 0.653370 is the maximum for the absolute Shannon entropy for all the zones in the year 2017.

Table 3.8: Calculation used to obtain absolute Shannon entropy for the year 2020

Zone	Total area(m^2)	Built-up (m^2)	Pi	1/Pi	Ln(1/Pi)	Pi*Ln(1/Pi)
1	34692100	9069100	0.261417	3.825308	1.341639	0.350727
2	231344100	13599100	0.058783	17.011720	2.833903	0.166585
3	706762200	21005600	0.029721	33.646370	3.515905	0.104496
4	990360800	21174300	0.021380	46.771830	3.845281	0.082214
					SUM	**0.704022**

From the table 3.8 above;

i. 0.704022 is the maximum for the absolute Shannon entropy for all the zones in the year 2020.

3.11.2 Calculation of Relative Shannon Entropy

Relative Shannon entropy was calculated as shown in Table 3.9 and 3.10 for both the years.

Table 3.9: Relative Shannon entropy value obtained for the year 2017

Zone	Total area(m^2)	Built-up(m^2)	Pi	1/Pi	Ln(1/Pi)	Pi*Ln(1/Pi)	Pi*Ln(1/Pi)/ln(n)
1	34692100	8063100	0.232419	4.302576	1.459214	0.339149	0.244644
2	231344100	11033200	0.047620	20.968000	3.042997	0.145126	0.104686
3	706762200	18392700	0.026024	38.426230	3.648740	0.094954	0.068495
4	990360800	18429900	0.018610	53.736630	3.984095	0.074141	0.053482
						SUM	**0.471307**

From the above table 3.9;

i. 0.471307 is the maximum value obtained after summing all the relative Shannon entropies for the four (4) zones.

Table 3.10: Relative Shannon entropy value obtained for the year 2020

Zone	Total area (m2)	Built-up(m2)	Pi	1/Pi	Ln(1/Pi)	Pi*Ln(1/Pi)	Pi*Ln(1/Pi)/ln (n)
1	34692100	9069100	0.261417	3.825308	1.341639	0.350727	0.252996
2	231344100	13599100	0.058783	17.011720	2.833903	0.166585	0.120166
3	706762200	21005600	0.029721	33.646370	3.515905	0.104496	0.075378
4	990360800	21174300	0.021380	46.771830	3.845281	0.082214	0.059305
						SUM	**0.507844**

From the table 3.10;

i. 0.507844 is the maximum value obtained after summing all the relative Shannon entropies for the four (4) zones.

CHAPTER FOUR

RESULTS, ANALSYSIS AND DISCUSSION

4.1 Overview

The results of this examination are exhausted in this section. The use of maps, graphs, and tables provides support for these outcomes. The final product of the urban sprawl map, the determination of the Urban Expansion Intensity Index, the Shannon entropy model, the landscape metrics, and finally the urban sprawl graph are comprised in this chapter.

4.2 Urban Sprawl Map

The finished map depicts the growth of urban sprawl in the Kilosa district over a 20-year period. The color blue is used to represent areas on the map that have not changed in terms of their built-up areas during this time. On the other hand, areas that have seen the development of new built-up areas as a result of urban sprawl in the past two decades are highlighted in yellow as shown in figure 4.1.

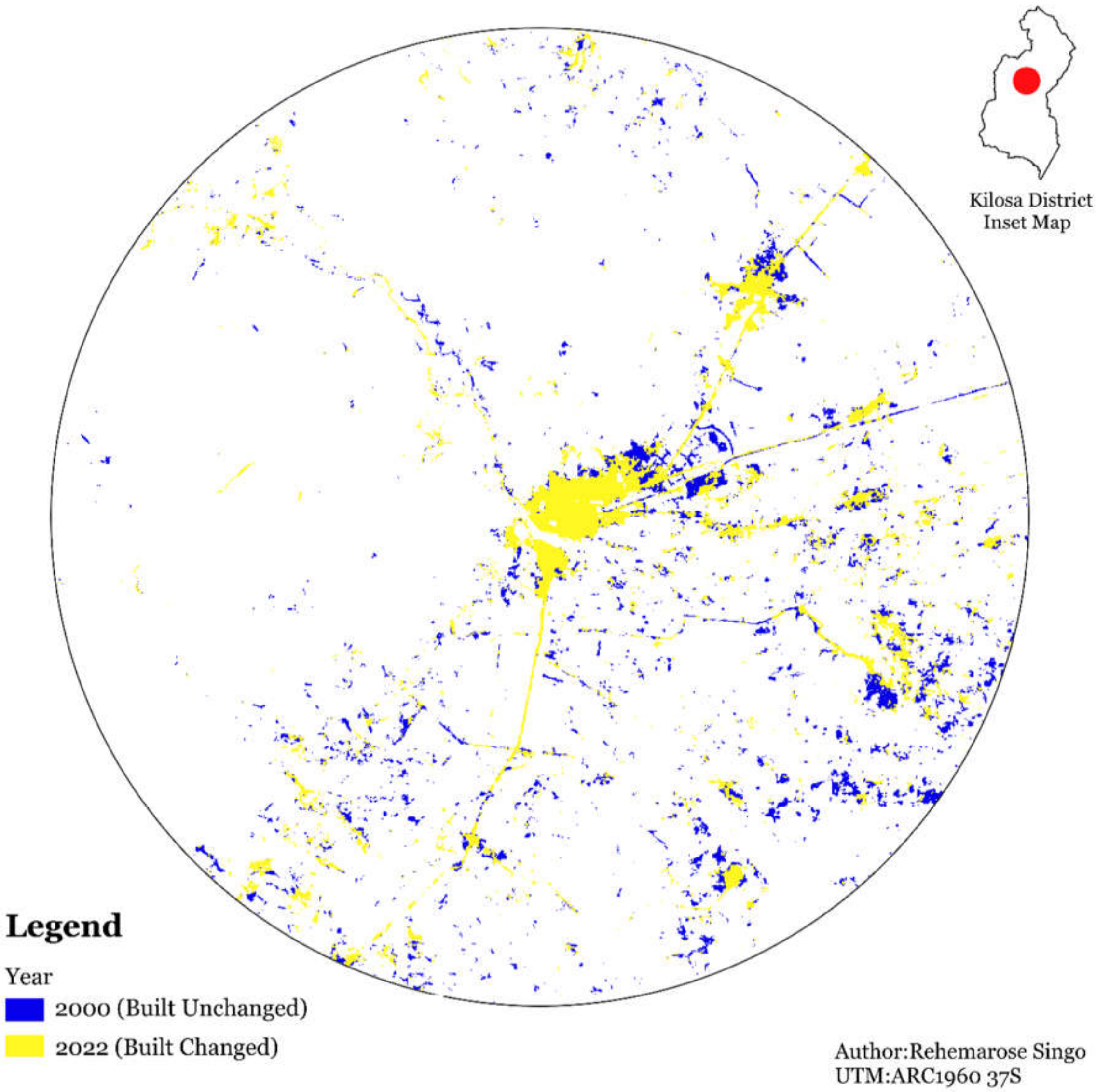


Figure 4.1: Kilosa Urban Sprawl Map between 2000 and 2022

Kilosa district clearly exhibits infilling patterned sprawl according to the provided map. This refers to the transformation of non-construction areas into urban land that is surrounded by urbanized areas. The red dot on the inset map shows the extent of Kilosa district that was used in this dissertation.

4.3 Population Density Map of Kilosa District

The total number of people living in a given location is represented by each hexagon on the map. The number of people represented by the color assigned to each hexagon varies from hexagon to hexagon as seen in figure 4.2.

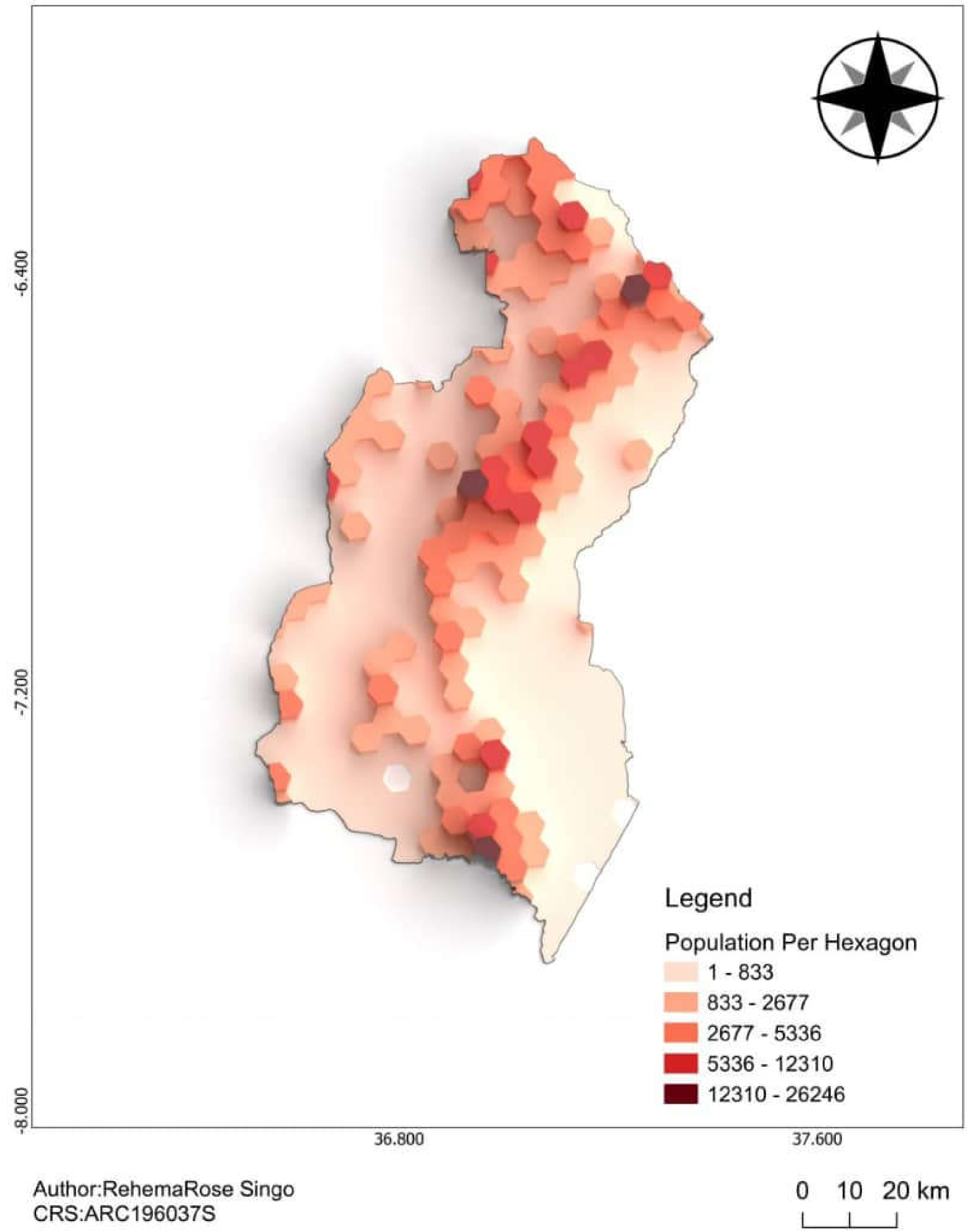


Figure 4.2: Population Density Map of Kilosa District.

The three-dimensional map shows that the Kilosa District's central region, as well as some parts in the north and south, have the highest population densities. This statement highlights how the urban sprawl map is complemented by the fact that the majority of the population resides in the central region.

4.4 Urban Expansion Intensity Index

The evaluation the urban spatial expansion difference was done by the means of calculating the Urban Expansion Intensity Index as follows;

Where as

ULAia is the urban sprawled area for the year 2017.

ULAib is the urban sprawled area for the year 2022

t is the time between the selected years.

TLAi is the average total area in meters square.

V represents the change of the urban sprawled areas in square meters between the selected years.

UEII is the urban expansion intensity index.

$$\text{UEII}=\left[\frac{ULA_{ib}-ULA_{ia}}{t}\right] / TLA_i \times 100 \tag{4.3}$$

Table 4.1: Calculation of Urban Expansion Intensity Index

ULAib(m²)	**ULAia(m²)**	**ULAib-ULAia(V) (m²)**	**t**	**$\frac{V}{t}$**	**TLAi(m²)**	**UEII =(v/t) / TLAi* 100**
23665800	18429900	5235900	5	1047180	825412800	0.126867

A UII value between 0 and 0.3 indicates a relatively low level of urbanization. This suggests that non-urban features like vegetation, agricultural land, or natural areas cover a greater portion of the area's land than urban development or built-up land.

A UII value between 0.7 and 1.0, indicates a relatively high level of urbanization. This would imply that a significant amount of the land is highly urbanized and characterized by built-up areas, infrastructure, and urban characteristics.

The Urban Expansion Intensity Index (UII) for the Kilosa district is 0.126867, as determined by the aforementioned analysis. This suggests that the area is developing at a slower rate and has undergone relatively little urbanization. In simpler terms, the findings suggest that Kilosa has a relatively low proportion of urbanized areas and a relatively slow rate of infrastructure growth.

4.5 Shannon Entropy Model

The evaluation of shannon entropy model used the absolute and relative entropies to obtained the actual model as shown in table 4.2 as shannon entropy is given by;

$$Shannon\ Entropy = Relative\ Shannon\ Entropy - Absolute\ Shannon\ Entropy$$

Table 4.2: Results of Shannon entropy model

Year	Absolute	Relative	Shannon Entropy
2017	0.653370	0.471307	0.182063
2020	0.704022	0.507844	0.196177

Compact distribution is evident when Shannon entropy is near 0 and dispersed when it is near 1. It is evident from the categorization and findings that the distribution pattern in the Kilosa district is compact. The Shannon entropy results for both years indicate a value close to zero, leading to this conclusion. To put it more clearly, the research suggests that the urban development in the Kilosa district is concentrated and does not have a great deal of dispersion.

4.6 Accuracy assessment results

The year 2018 had the overall accuracy of 83.2944. An overall accuracy of 80 percent or higher means the classification algorithm has successfully identified the majority of land cover classes.

The year 2018 had the Kappa value of 0.6938 which shows a moderate to significant degree of understanding between the classification results and the reference data.

The year 2019 had the overall accuracy of 80.5460. An overall accuracy of 80 percent or higher means the classification algorithm has successfully identified the majority of land cover classes.

The year 2019 had the Kappa value 0.6381 which shows a moderate to significant degree of understanding between the classification results and the reference data.

The year 2020 had the overall accuracy of 78.9334. An overall accuracy of 78% suggests that the classification algorithm achieved a relatively good level of accuracy in assigning class labels to the pixels.

The year 2020 had Kappa value 0.6138 which shows a moderate to significant degree of understanding between the classification results and the reference data.

4.7 Urban Sprawl Graphs

The urban sprawl that has taken place in the Kilosa district is clearly and conclusively depicted in Figure 4.3. The vertical line on the graph shows how many built-up areas there are in meter square, and the horizontal axis shows the years that go with it.

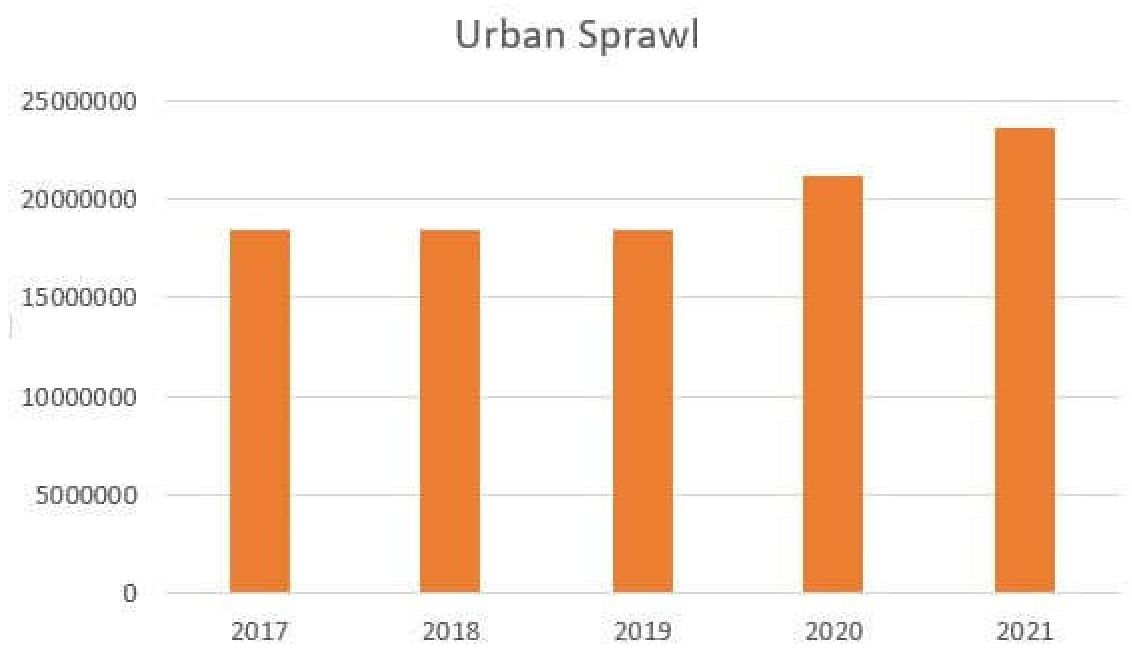


Figure 4.3: Urban Sprawl Graph

The five year time range shows that urban sprawl has been occuring subtly from year 2017 to the year 2019, from the year 2020 to 2021 there has been a modest increase in the area of built up area in meter square this can be due to;

Migration: Economic opportunities, employment prospects, or social factors may have prompted an influx of people migrating to Kilosa from other areas.

Urbanization: People from rural areas may have moved to the district in search of better living conditions, access to services, and employment opportunities as a result of the district's rapid urbanization.

Projects for improvement: During that time, Kilosa may have seen an increase in population as a result of investments or development projects that attracted individuals and families.

Governing Practices: The population growth may have been influenced by policies, programs, or incentives provided by the government to encourage people to settle in Kilosa.

4.8 Discussion

The urban sprawl analysis used built-up class to calculate the Shannon entropy and the Urban Expansion Intensity Index to examine urban sprawl in the Kilosa District. The findings of Urban Expansion Intensity Index show that urban sprawl exists in Kilosa, at a relatively low density and for Shannon entropy the findings shows that urban sprawl happens in a compacted pattern. Kilosa district clearly exhibits infilling patterned sprawl according to the provided map. The same pattern was experienced in Oregon, USA (Richard, 2012) who’s development was in vacant or underutilized spaces within already developed urban areas, rather than expanding outward into previously undeveloped areas. A study done in Libya, Tripoli by the use of Shannon entropy model confirm that the model, remotely sensed data, and Geographic Information System (GIS) can be used to identify patterns and general trends in urban sprawl and growth. According to the analysis, Tripoli's urbanization has spread out over the three time periods. Tripoli is an appealing location for development and urbanization due to its geographical location, close proximity to the Mediterranean coast, and accessibility to transportation routes. Urban sprawl is also aided by the availability of land for expansion and natural resources, also Libya's economic and administrative capital is Tripoli. People migrate from rural areas in search of better livelihoods, which further contributes to urbanization and urban sprawl as the city attracts investment, job opportunities, and services. Due to the influence of economic activities and the geographical location of a given area, the extent and characteristics of urban sprawl vary. Because of this discrepancy, it is possible that the degree of urban sprawl varies significantly depending on a region's economic development and specific geographic characteristics. Urban sprawl patterns and intensities vary widely across locations as a result of the interaction between economic factors and geographical considerations (Abubakr, 2014).

CHAPTER FIVE

CONCLUSION AND RECOMMENDATION

5.1 Conclusion

By characterizing urban sprawl in the Kilosa district using GIS and remote sensing methods, the research achieved its goals. The map of urban sprawl, which covers the years 2000 to 2022, was successfully created. This made it easier to determine the town's spatial dispersion or concentration and to evaluate quantitative differences in urban expansion. The findings show that there is a compacted distribution pattern in the central part of Kilosa District. The compact distribution suggests that development is happening slowly. The town is expanding at a fairly moderate rate. Stakeholders can effectively plan for the town's future land use and keep track of Kilosa's ongoing growth with the help of the data analysis. Additionally, it gives Kilosa recognition and visibility by making the general public aware that it is one of the towns experiencing significant growth.

5.2 Recommendation

This research implemented Shannon entropy and Urban expansion intensity index for characterization of urban sprawl patterns in the Kilosa district from 2000 to 2022. Additional research is recommended in order to acquire a comprehensive comprehension of the spatial and temporal dynamics of urban sprawl, this may necessitate the application of complementary methods like landscape metrics. It is important to carefully examine how urban sprawl affects land use, transportation, and environmental resources. The negative effects of urban sprawl can be mitigated by employing an integrated planning strategy that takes into account sustainable development principles like compact city design, mixed land use, and effective transportation systems. Interventions of policy in Kilosa district which creates and implement policies that control how land is used and how the city grows.

Future research on the characterization of patterns of urban sprawl has the potential to contribute to informed decision-making, sustainable urban planning, and the promotion of resilient and livable cities by addressing these challenges and recommendations.

REFERENCES

Abubakr. (2014). Quantitative analysis of urban sprawl in Tripoli using Pearson's Chi-Square statistics and urban.

Al-Sharif. (2017). Urban Sprawl Assessment. Spatial Modeling and Assessment of Urban form: Analysis of Urban Growth: From Sprawl to Compact Using Geospatial Data.

Angle. (2007). Urban sprawl metrics: An analysis of global urban expansion using GIS, in Annual Conf. Tampa, Florida.

Auch. (2004). *Urban Growth In American Cities.*

Batty. (1999). The Dynamics of Urban Sprawl. London.: Centre for Advanced Spatial Analysis, University College.

Boudreaux. (2017). *COMMUNITY LAND CONFLICTS: HOW LOCAL LAND DISPUTES AFFECT THE PRIVATE SECTOR INVESTMENTS AND DEVELOPMENT PROJECTS.*

Brueckner. (2000). Urban Sprawl: Diagnosis and Remedies. International Regional Science Review.

Campbell. (2002). *Introduction to Remote Sensing(3rd ed.).* Newyork: Guilford Press.

Cheng. (2003). Urban Growth Pattern Modelling: A Case Study of Wuhan City, PR China. Landsc. Urban Plan.

Civco. (2002). Quantifying and Describing Urbanizing Landscapes in the Northeast United States. Photogr. Eng. Remote Sens.

Club. (1998). The Dark Side of the American Dream: The Costs and Consequences of Suburban Sprawl. College Park, MD:Challenge to Sprawl Campaign.

Couch. (2007). The Consequences of Urban Sprawl: A Case Study of Wirral, Merseyside, Liverpool. John Moores University.

DiPasquale. (1999). Incentives and Social Capital: Are homeowners better citizens? Journal of Urban Economics.

Downs. (1999). Some realities about sprawl and urban decline. Housing Policy Debate 10(4).

Epstein. (2002). Techniques for mapping suburban sprawl. Photogr. Eng. Remote Sens.

ERM. (2019). *Standard Gauge Railway Line(SGR) Project.*

Ewing. (1997). Is Los Angeles-Style sprawled desirable? Journal of the American Planning Association.

Ewing. (2008). Measuring Sprawl and its Impact. Washington: Smart Growth America.

Forman. (1995). Land mosaics: The ecology of landscapes and regions. Cambridge: Cambridge University Press.

Fulton. (2001). Who sprawls most? How growth patterns differ across the U.S..Center for urban and metropolitan policy. Washington, D.C.: the booking institution.

Galster. (2001). Wrestling sprawl to the ground: Defining and measuring an elusive concept. Housing Policy Debate.

Gomarasca. (1993). One century of land use changes in the metropolitan area of Milan. Int. J. Remote Sens.

Gordon. (1997). Are compact cities a desirable planning goal? Journal of the American Planning Association.

Gordon. (2000). Critiquing sprawl's critics. Policy Analysis.

Greenberg. (1997). Analyzing the urban–wildland interface with GIS. J. Forestry.

Haack. (2006). Urban growth analysis and modelling in the Kathmandu valley, Nepal. Habitat International.

Habibi. (2011). Causes, results and methods of controlling urban sprawl.

Kahn. (2001). Does sprawl reduce the black/white housing consumption gap? Housing Policy Debate.

Lillesand. (2008). *Remote sensing and images interpretation.*

Lo. (2001). Modeling the population of China using DMSP operational Linescan system nighttime data. Photogr. Eng. Remote Sens.

Manesha. (2021). Measuring urban sprawl of small and medium towns using GIS and remote sensing techniques: A case study of Sri Lanka.

Mather. (2004). *Computer Processing of Remotely Sensed Images.*

Mieszkowski. (1993). The causes of metropolitan suburban. Journal of Economic Perspectives.

Mundia. (2005). Analysis of land use/cover changes and urban expansion of Nairobi city using remote sensing and GIS. Int. J. Remote Sens.

NBS. (2022). *National Bureau of statistics*. Retrieved from Tanzania National Bureau of Statistics: https://www.nbs.go.tz/index.php/en/

Pendall. (1999). Do land use controls cause sprawl? Environamental and Planning.

Pradhan. (2017). *Spatial Modelling and Assessment of Urban Form.*

Richard. (2012). Infill development that benefits neighborhoods and the environment.

Stefanov. (2001). Monitoring urban land cover change: An expert system approach to land cover. Remote Sens. Environ.

Sudhira. (2004). Urban sprawl: Metrics, dynamics and modelling using GIS. S. Int. J. Appl.Earth Observ. Geoinform.

Torrens. (2000). Measuring sprawl. Working paper no. 27. Centre for Advanced Spatial Analysis, University College, London.

Trani. (1999). An analysis of deforestation: Metrics used to describe pattern change. Forest Ecol. Manage.

UNDESA . (2015). Retrieved 02 2023, from UnitedNationsEnable: https://www.un.org/development/desa/disabilitites/envision2030.html#menu-header-menu

Yanli. (2012). New Machine Learning Algorithm: Random Forest.

Yeh. (2001). Measurement and monitoring of urban in rapidly growing region using in entropy. Photogr. Eng. Remote Sens.

Zhi-qiang. (2011). Evaluation of urban sprawl and urban landscape pattern in a rapidly developing region.

APPENDICES

Appendix 1: Classification Accuracy Report of Sentinel 2A the year 2018

ErrMatrixCode	Reference	Classified	PixelSum
1	1	1	3337
3	1	2	5973
10	1	4	144
2	2	1	108
5	2	2	1067022
9	2	3	1047
13	2	4	118165
8	3	2	85559
12	3	3	83408
15	3	4	28189
7	4	1	295
11	4	2	126583
16	4	4	666139

> ERROR MATRIX (pixel count)
> Reference

V_Classified	1	2	3	4	Total
1	3337	108	0	295	3740
2	5973	1067022	85559	126583	1285137
3	0	1047	83408	0	84455
4	144	118165	28189	666139	812637
Total	9454	1186342	197156	793017	2185969

> AREA BASED ERROR MATRIX
> Reference

V_Classified	1	2	3	4	Area	Wi
1	0.0067	0.0002	0.0000	0.0006	3344500.0000	0.0075
2	0.0027	0.4901	0.0393	0.0581	263811200.0000	0.5903
3	0.0000	0.0005	0.0379	0.0000	17160700.0000	0.0384
4	0.0001	0.0529	0.0126	0.2982	162583200.0000	0.3638
Total	0.0095	0.5437	0.0898	0.3570	446899600.0000	
Area	4239056	242987282	40151129	159522133	446899600	
SE	0.0001	0.0002	0.0002	0.0002		
SE area	23322	108628	67062	99161		
95% CI area	45711	212910	131441	194356		
PA [%]	70.3958	90.1433	42.2104	83.5455		
UA [%]	89.2246	83.0279	98.7603	81.9725		
Kappa hat	0.8912	0.6280	0.9864	0.7197		

Overall accuracy [%] = 83.2944
Kappa hat classification = 0.6938

Appendix 2: Classification Accuracy Report of Sentinel 2A the year 2019

ErrMatrixCode	Reference	Classified	PixelSum
1	1	1	2693
3	1	2	5806
10	1	4	955
2	2	1	166
5	2	2	1089320
9	2	3	2975
13	2	4	93881
8	3	2	89048
12	3	3	83239
15	3	4	24869
7	4	1	258
11	4	2	207649
14	4	3	417
16	4	4	584693

> ERROR MATRIX (pixel count)
> Reference

V_Classified	1	2	3	4	Total
1	2693	166	0	258	3117
2	5806	1089320	89048	207649	1391823
3	0	2975	83239	417	86631
4	955	93881	24869	584693	704398
Total	9454	1186342	197156	793017	2185969

> AREA BASED ERROR MATRIX
> Reference

V_Classified	1	2	3	4	Area	Wi
1	0.0063	0.0004	0.0000	0.0006	3267000.0000	0.0073
2	0.0027	0.5017	0.0410	0.0956	286459500.0000	0.6410
3	0.0000	0.0015	0.0407	0.0002	18919800.0000	0.0423
4	0.0004	0.0412	0.0109	0.2568	138253300.0000	0.3094
Total	0.0094	0.5448	0.0926	0.3532	446899600.0000	
Area	4205003	243449416	41387588	157857594	446899600	
SE	0.0001	0.0003	0.0002	0.0002		
SE area	26158	116075	67900	107663		
95% CI area	51269	227506	133084	211019		
PA [%]	67.1247	92.0929	43.9238	72.6976		
UA [%]	86.3972	78.2657	96.0845	83.0061		
Kappa hat	0.8627	0.5226	0.9568	0.7372		

Overall accuracy [%] = 80.5460
Kappa hat classification = 0.6381

Appendix 3: Classification Accuracy Report of Sentinel 2A the year 2020

ErrMatrixCode	Reference	Classified	PixelSum
1	1	1	4342
3	1	2	2233
10	1	4	2879
2	2	1	1110
5	2	2	1099069
9	2	3	2573
13	2	4	83590
4	3	1	2
8	3	2	75852
12	3	3	87602
15	3	4	33700
7	4	1	2151
11	4	2	251246
14	4	3	942
16	4	4	538678

> ERROR MATRIX (pixel count)
> Reference

V_Classified	1	2	3	4	Total
1	4342	1110	2	2151	7605
2	2233	1099069	75852	251246	1428400
3	0	2573	87602	942	91117
4	2879	83590	33700	538678	658847
Total	9454	1186342	197156	793017	2185969

> AREA BASED ERROR MATRIX
> Reference

V_Classified	1	2	3	4	Area	Wi
1	0.0088	0.0023	0.0000	0.0044	6902400.0000	0.0154
2	0.0010	0.4907	0.0339	0.1122	284985900.0000	0.6377
3	0.0000	0.0012	0.0418	0.0004	19428400.0000	0.0435
4	0.0013	0.0385	0.0155	0.2481	135582900.0000	0.3034
Total	0.0111	0.5326	0.0912	0.3650	446899600.0000	
Area	4978836	238037640	40749332	163133791	446899600	
SE	0.0001	0.0003	0.0001	0.0003		
SE area	41773	118627	66094	117116		
95% CI area	81875	232509	129545	229547		
PA [%]	79.1522	92.1198	45.8386	67.9525		
UA [%]	57.0940	76.9441	96.1423	81.7607		
Kappa hat	0.5661	0.5067	0.9576	0.7128		

Overall accuracy [%] = 78.9334
Kappa hat classification = 0.6138