ARDHI UNIVERSITY

MAPPING URBAN SPRAWL PATTERN OF ARUSHA REGION

TARIMO, RIGOBERT PATROKIL

22748/ T.2019

GI 472: DISSERTATION

BSc in GEOINFORMATICS (BSc GI)

July, 2023

MAPPING URBAN SPRAWL PATTERN OF ARUSHA REGION

BY

RIGOBERT P TARIMO

A Dissertation submitted to the Department of Geospatial sciences and technology in the partial fulfillment of the requirements for the awards of Bachelor of Science Degree of Geoinformatics (BSc.in Geoinformatics) at Ardhi University

CERTIFICATION AND COPYRIGHT

The undersigned certifies that they have read and hereby recommends for acceptable by Ardhi University, a dissertation entitled, **“Mapping urban sprawl pattern of Arusha region.”** in the fulfillment of the requirements for the degree of Bachelor of Science in Geoinformatics, Ardhi University, Dar es Salaam.

………………………………

Dr. Atupelye Komba

(Supervisor)

Date……………………………

DECLARATION AND COPYRIGHT

I, **Tarimo, Rigobert Patrokil** hereby declare that the contents of this report are the results of my own study and findings and the best of my knowledge. Any reference to work done by any other person or institution or any material obtained from other sources have been duly cited and referenced. I further certify that the research report has not been published or submitted for publication anywhere else.

Signature: ……………………………

TARIMO, RIGOBERT P

Date………………………………….

ACKNOWLEDGEMENT

I extend my sincere gratitude to the Lord Almighty for giving me strength, knowledge and wisdom to carry on throught my journey of education and finally in accomplishing this research.

I am absolutely grateful to my lovely parents, Mr. and Mrs. Patrokil Fidelis Tarimo for their tireless efforts in providing me with best and the little they could afford in all spheres from Kindergarten up to this level. Their efforts and struggles have allowed me to have a key to unlock the mysteries of our world and beyond, my siblings (Catherine, Albert and Ryan) for their unconditional love and support.

I would also like to express my sincere appreciation to my supervisors Dr. Atupelye Komba for her best guidance, time, tireless consultation and advising in ensuring that I complete this dissertation within the limited time Thus, enabled me be equipped with several updated skills in geospatial science.

Also, I would like to thank significant contributions made to staff members of Geospatial Science and Technology Department for their moral and academic support for always being there whenever necessary and equipping us with essential skills in geospatial science and all the necessities that an elite requires.

Special thanks to my lovely brothers and friends Gadiel Mvamba, Clinton Lyimo, George Cyprian, Manaseh Kiunsi, Yusuph Hassan, Mariam Zaid and Frank Mlagara. Their contribution towards the accomplishment of this research is sincerely appreciated and gratefully acknowledged.

I appreciate the love and tireless support from all my colleagues from Geo-Informatics and Geomatics: And whoever contributed in one way or the other in making this research successful. May God's sufficiency be manifested in you all.

DEDICATION

This study is wholeheartedly dedicated to my beloved parents, my father Patrokil Tarimo and my mother Dafrosa Massawe who have been source of inspiration, strength when I thought of giving up, who continually provide their moral, spiritual, emotional and financial support.

To brothers, sisters, relatives, mentor, friends and classmates who shared their words of advice and encouragement.

ABSTRACT

Urbanization is a global phenomenon, and in underdeveloped countries like Tanzania, the ratio of urban sprawl has been rapidly increasing. Non-systematic development, increased migration, and sharp population growth have been predominant factors contributing to this trend. These variables have fueled population growth and led to the rapid expansion of built-up areas in the Arusha region.

This research aims to analyze the pattern and implications of urban sprawl using GIS and remote sensing as an improved approach compared to traditional spatial or cartographic mapping and monitoring methods. The traditional methods lack the effectiveness to analyze and explain the spatial-temporal dimensions of urban sprawl from 2010 to 2022.Remotely sensed data and the random forest algorithm using supervised classification were utilized to investigate urban sprawl, while population data was used in trend analysis to examine the relationship between population and urban sprawl. Landsat images from 2010, 2014, 2018, and 2022 were crucial in providing information and characteristics of urban land cover on spatial and temporal scales.

The results indicate that the total urban built-up areas in the Arusha region were 569 km^2, 616 km^2, 647 km^2, and 715 km^2 with growth rates of 1.51%, 1.64%, 1.72%, and 1.9% respectively in 2010, 2014, 2018, and 2022. Change detection techniques were employed to highlight the major changes detected in the built-up area of Arusha during this 12-year period. The increase in built-up area was 47 km^2, 31 km^2, and 68 km^2 with growth rates of 0.13%, 0.08%, and 0.18% respectively in the intervals of 2010-2014, 2014-2018, and 2018-2022.

The results also revealed a cluster and leap-frog pattern of sprawl primarily in Arusha district and Arusha Municipal, where more than 45% of the total population of the region is concentrated. The research further highlighted that the sprawl patterns were influenced by economic activities such as tourism and mining, rapid urban population growth, increased demand for land, poor urban planning, and social segregation, each with its respective implications.

To control the city's growth and mitigate the effects of widespread urban sprawl, it is essential to study the causes underlying this phenomenon. Proper planning of the area, reduction of potential damage, and the provision of essential services such as roads, hospitals, and schools are recommended to ensure sustainable living in the region.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

GIS	Geographical Information System
LULC	Land Use Land Cover
USGS	United States Geological Survey
UN	United Nations
TM	Thematic Mapper
OLI	Operational Land Imager
TIFF	Tag Image File Format
CSV	Comma Separated Values
ROI	Region Of Interes

CHAPTER ONE

INTRODUCTION

1.1 Background of the Study

Urban planning is a decision-making method aimed at achieving desired objectives within specific resource and time constraints. In the context of addressing socio-economic and physical challenges in a given town (Haimanot, 2009), urban planning assumes a crucial role.

Urban sprawl is a consequence of development patterns that impact cities and their management, resulting in excessive land consumption, increased traffic, and socioeconomic segregation (International Journal of Architecture and Urban Development, Vol. 8, No. 4, Autumn 2018). Carruthers and Ulfarsson (2002) argue that sprawl development imposes a financial burden when it comes to infrastructure needs and public health issues.

With rapid population growth, urbanization has undergone drastic expansion worldwide. For urbanization to have positive outcomes, it must be carefully and effectively planned; otherwise, it can have negative environmental consequences (Morefield et al., 2018).

Migration to urban areas occurs for various reasons in both developing and developed countries. The former experiences population growth, while the latter witnesses population migration (UN Habitat, 2016).

Informal settlements are a significant challenge associated with the urbanization process, affecting urban areas worldwide (UN Habitat, 2016).

Growing countries like Tanzania face challenges such as internal collapse, pollution, traffic congestion, and housing pressure (Mubea et al., 2011). Urban sprawl is a growing concern for governments, environmental organizations, and citizens.

Arusha region has experienced significant population growth, leading to rapid urbanization and subsequent changes in land use and land cover. This has resulted in urban sprawl, driven by the presence of government educational institutions, mining activities, and tourist attractions (Sumari et al., 2019a).

Urban sprawl in the Arusha area manifests in physical and socioeconomic issues such as traffic congestion, loss of open space, segregation, poor infrastructure, inadequate amenities, loss of vegetation, air pollution, and water pollutant runoff (Sumari et al., 2019a).

Advancements in geospatial technology have greatly enhanced the ability of urban planners and managers to research and monitor urban conditions and growth (Sumari et al., 2019a).

Despite the unprecedented rates of urbanization in cities like Arusha across Africa, there is a lack of academic studies examining the patterns and characteristics of urban sprawl in African cities (Couch et al., 2008). This case study aims to fill that gap by mapping and analyzing recent trends in urban land expansion in the Arusha region, with a specific focus on identifying signs of urban sprawl. Geospatial methodologies, including classification and change detection techniques, will provide a comprehensive understanding of urban sprawl patterns (Sumari et al., 2019a).

1.2 Problem Statement

Arusha region is currently facing significant challenges due to an annual population change of 4.26% (UN World Urbanization, 2023). This rapid population growth, resulting from both migration and natural increase, has led to a widespread pattern of urban sprawl in the region. The consequences of this uncontrolled and haphazard growth are evident in the significant environmental, social, and economic impacts it has imposed. The pressure on land resources has negatively impacted the quality of life for residents, with the emergence of unplanned settlements.

Despite the importance of comprehending and addressing this phenomenon, there is a lack of comprehensive research on the spatial patterns and trends of urban sprawl in the Arusha region. It is crucial to map these urban sprawl patterns to understand the extent and distribution of the problem accurately. By doing so, we can develop effective strategies for managing urban growth, ensuring the region's sustainability in the future.

1.3 Research Objectives

1.3.1 Main Objective

The main objective is to map the urban sprawl pattern of Arusha region.

1.3.2 Specific Objectives

i. To produce LULC maps of the of the region for years 2010,2014,2018 and 2022.
ii. Evaluating the Percentage of the urban sprawl across different years by identifying the years with more and less concentration.

iii. To analyze the drivers and impacts of urban sprawl in Arusha region.

1.4 Research Questions

i. What are the LULC maps for the region?

ii. How does the percentage of urban sprawl vary across different years?

iii. What factors and drivers of urban sprawl in Arusha region?

1.5 Importance of the Study

i. This study will provide information concerning with the urban sprawl changes and expansions.

ii. The outputs for this study is important for decision makers and urban planners for proper planning and utilization of natural resources with their management concerning land use.

1.6 Beneficiaries

i. The decision makers should use this information for the decision making and planning processes on the land use and utilization of natural resources with their management and al process concerning land use.

ii. The urban planners from the this current and updated information on the status of the livelihood obtained from this study for proper planning of Arusha City.

iii. The environmental researchers could be able to obtain the information in assist of achieving environmental planning and resource management.

1.7 Description of the Study

Arusha is one of 31 administrative regions found in northern Tanzania. It is located between latitude 2˚30'S and 4˚30'S and longitude 35˚30'E and 38˚30'E with the coverage area of 37576 km^2. It is bordered by Kenya to the north and Kilimanjaro Region to the east. It has an estimated population of over 2.3 million people. The region's capital and largest city is Arusha, which is located at an elevation of 1,400 meters above sea level.

Arusha region has a tropical highland climate with moderate temperatures throughout the year. The average daytime temperatures range from 20-25°C, while the nighttime temperatures typically drop to around 10-15°C. The region experiences two rainy seasons per year, with the heaviest rains occurring from March to May and from November to December. Annual rainfall in the region ranges from 700-1,500 mm per year.

The city has a growing urban population due to economic activities like Tourism due available tourist attraction areas such as Ngorongoro conservation area also mining activities due to Mererani which mines Tanzanite mineral, all these activities has contributed to the urban sprawl also developed services and infrastructure, including hotels, restaurants, and transportation links. Despite its rapid growth, Arusha retains a relaxed and laid-back atmosphere, with a vibrant local culture and a bustling street life.

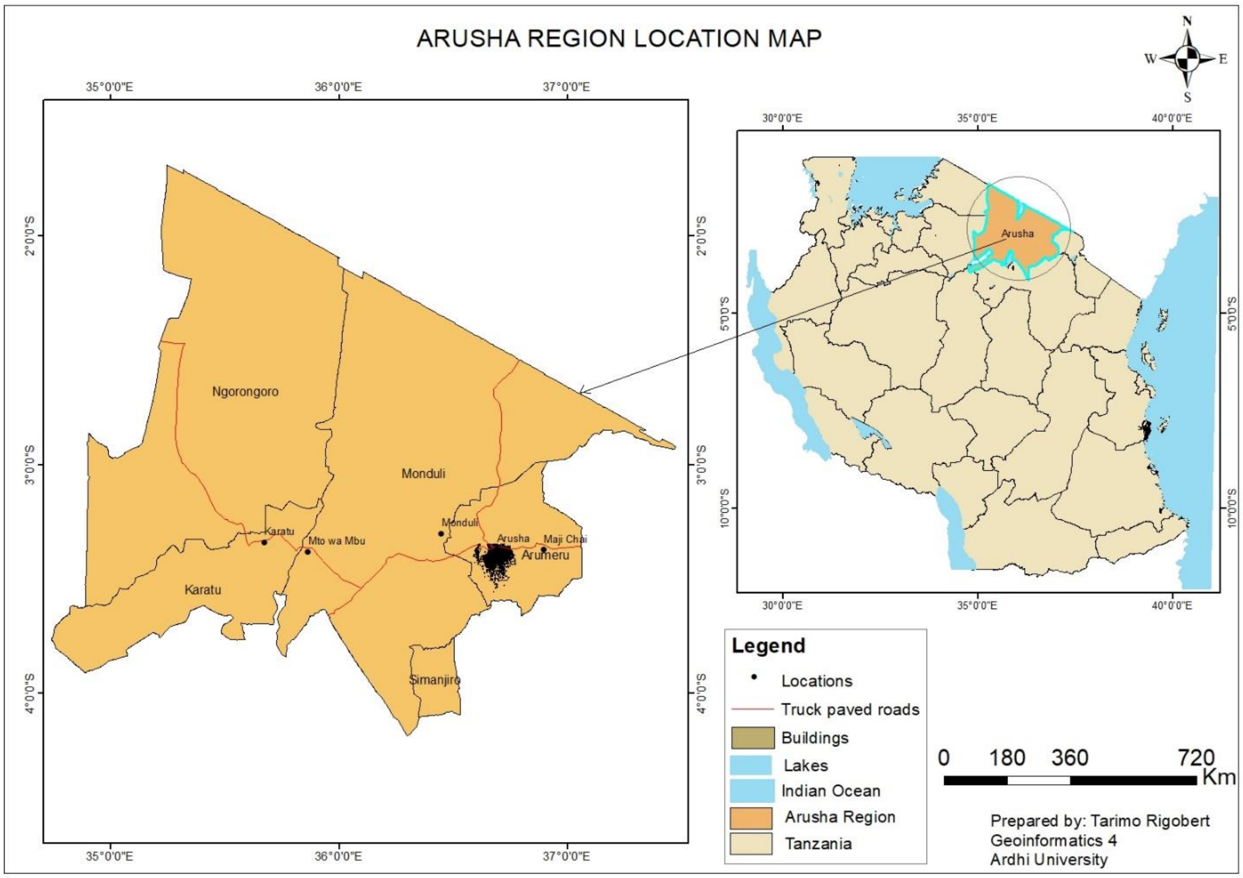

Figure 1.1: *Location of study area*

CHAPTER TWO

LITERATURE REVIEW

2. Overview

In this chapter provides different literature about different relevant topics which were reviewed in the course of conducting this research. Such topics include remote sensing, Landsat sensors, land cover and land use mapping will be discussed;

2.1 Remote Sensing

Remote sensing (RS) involves sensing and recording either reflected or emitted energy from distant objects, the energy is then processed and analyzed in order to obtain useful information for different applications. Remote Sensing is performed from a platform (satellite, airplane or ship) using instruments (sensors) which measure electromagnetic radiation reflected or emitted from the terrain. It consists of analysis and interpretation of measurements of electromagnetic radiation that is reflected from or emitted by targets (Gernt, 2001). Remote sensing has a wide range of applications including mapping urban sprawl pattern.

2.2 Image preprocessing

Images in their raw form contains some deficiencies or errors. Therefore, the process of removing those deficiencies is known as preprocessing. (Koch, 2011), The techniques include geometric and radiometric correction.

2.2.1 Geometric Correction

This involves correction of geometric errors which affect the position of the digital number (DN) value and placing the corrected pixels in their proper reference system. This enable to remove distortions from an image due to earth's rotation and other imaging conditions. Also, it is useful in providing geo referenced images and also in merging different images. It removes both internal and external distortions in an image.

2.2.2 Radiometric Correction

It involves improving the surface spectral reflectance, emittance or backscattered measurement obtained using a remote sensing system, atmospheric and topographic correction. It helps to remove variations within an image, between adjacent images, between bands or between image dates.

2.3 Image Classification

The overall objective of image classification procedures is to automatically categorize all pixels in an image into land cover classes or themes (Kevin et al, 2005). Often this is done using spectral patterns; that is, pixels that share similar combinations of spectral reflectance or emissivity are grouped together in classes that are assumed to represent particular categories of surface features (Thomas, Ralph, & Jonathan, 2002).

Measured reflection values in an image depend on the local characteristics of the earth surface; in other words, there is a relationship between land cover and measured reflection values. In order to extract information from the image data, this relationship must be found. The process to find the relationship is called classification. Classification can be done using a single band, in a process called density slicing, or using many bands (multi-spectral classification).

Image classification is one of the operations performed in remote sensing for analysis and interpretation.

The results form image classification can be used to create thematic maps. There are two types of classification basing on the nature of interaction between the analyst or operator and the computer during classification namely: Supervised and Unsupervised classification.

2.3.1 Supervised Image Classification

With supervised classification the information classes are identified from the image with the need of the training samples. The image processing software system is then used to develop a statistical characterization of the reflectance for each information class, this stage is often called signature analysis and may involve developing a characterization as simple as the mean or the rage of reflectance on each bands, or as complex as detailed analyses of the mean, variance and covariance over all bands. Once a statistical characterization has been achieved for each information class, the image is then classified by examining the reflectance for each pixel and making a decision about which of the signatures it resembles most (Eastman et al, 1993).

2.3.2 Unsupervised Image Classification

Unsupervised image classification is a method which examines a large number of unknown pixels and divides into a number of classes based on natural groupings present in the image values, unlike supervised classification, unsupervised classification does not require analyst specified training data. The basic premise is that values within a given cover type should be close together in the measurement space that is have similar gray levels, whereas data in different classes should be comparatively well separated that is have very different gray levels (Eastman et al, 1993).

The classes that result from unsupervised classification are spectral classed which based on natural groupings of the image values, the identity of the spectral class will not be initially known, must compare classified data to some form of reference data (such as larger scale imagery, maps, or site visits) to determine the identify and informational values of the spectral classes. Thus, in the supervised approach, to define useful information categories and then examines their spectral separability; in the unsupervised approach the computer determines spectrally separable class, and then define their information value (Lillesand and Kiefer, 1994). unsupervised classification is becoming increasingly popular in agencies involved in long term GIS database maintenance. The reason is that there are now systems that use clustering procedures that are extremely fast and require little in the nature of operational parameters. Thus it is becoming possible to train GIS analysis with only a general familiarity with remote sensing to undertake classifications that meet typical map accuracy standards, with suitable ground truth accuracy assessment procedures, this tool can provide a remarkably rapid means of producing quality land cover data on a continuing basis.

2.3.3 Image Classification Approach

The supervised image classification was used. In supervised classification the majority of the effort is done prior to the actual classification. Once the classification is run the output is a map with classes that are labeled and correspond to information classes or land cover types. Supervised classification can be much more accurate than unsupervised classification, but depends heavily on the training sites, the skill of the individual processing the image, and the spectral distinctness of the classes. If two or more classes are very similar to each other in terms of their spectral reflectance (e.g., annual-dominated grasslands versus perennial grasslands), misclassifications will tend to be high. Supervised classification requires close attention to development of training data. If the training data is poor or not representative the classification results will also be poor. Therefore, supervised classification generally requires more times and money compared to unsupervised.

2.4 Classification Algorithms

After the training sample sets have been obtained, classification of the image can be carried out by applying a classification algorithm. Several classification algorithms exist. The choice of the algorithm depends on the purpose of the classification, the characteristics of the image and training data. The operator has to decide if a 'reject' or 'unknown' class is allowed. Here only one algorithm is explained which is Random Forest classifier.

2.4.1 Random Forest Classifier

Random Forest is a popular machine learning algorithm that belongs to the supervised learning technique. It can be used for both Classification and Regression problems in ML. It is based on the concept of ensemble learning, which is a process of combining multiple classifiers to solve a complex problem and to improve the performance of the model. As the name suggests, "Random Forest is a classifier that contains a several number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of the dataset." Instead of relying on a single decision tree, the random forest takes the prediction from each tree and based on the majority votes of predictions, and it predicts the final product (Noida,2021).

The greater number of trees in the forest leads to higher accuracy and prevents the problem of overfitting.

The below diagram explains the working of the Random Forest algorithm**:**

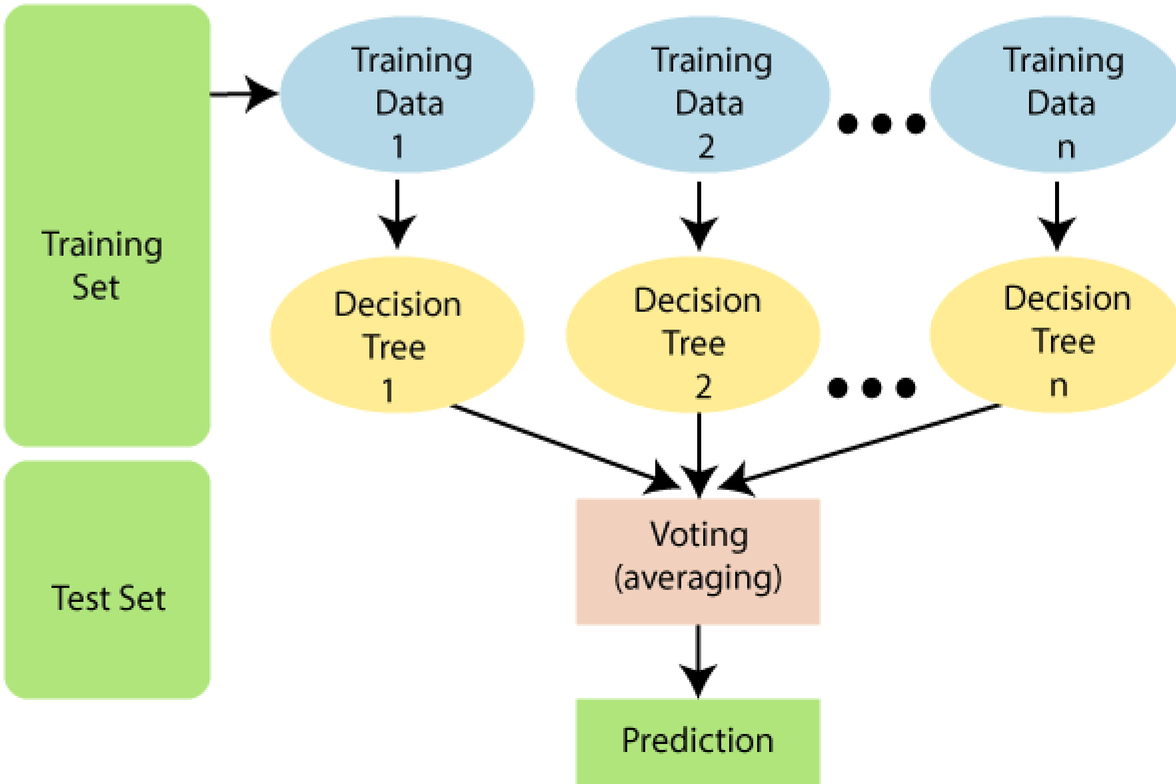

Figure 2.1: *Decision tree algorithm (Breiman, L, 2001)*

2.5 Accuracy Assessment

Accuracy assessment "measures the agreement between a standard assumed to be correct and a classified image of unknown quality." (Campbell, 2007). An integral part of image classification is a validation of the results, again independent data are required. The result of validation process is an error matrix from which different measures of error can be measured (Wim et al, 2009). Precision defines the level of details found within the classification; it is possible to increase the accuracy of classification by decreasing the amount of detail or by generalizing to broad classes rather than very specific ones.

For example, a scheme which generally categorizes trees versus crops presents less opportunity for classification error than one that distinguishes many types of trees and many types of crops. In this case, lower precision provides the potential for higher accuracy. However, the user of the map that offers only general classes cannot make precise statements about any given point on the map. Classification error occurs when a pixel (or feature) belonging to one category is assigned to another category. Errors of omission occur when a feature is left out of the category being evaluated; errors of commission occur when a feature is incorrectly included in the category being evaluated. An error of omission in one category will be counted as an error in commission in another category. Accuracy assessment is performed by comparing the map created by remote sensing analysis to a reference map based on a different information source. It is important to know why the remote sensing analysis is needed if the reference map to compare it to already exists. One of the primary purposes of accuracy assessment and error analysis in this case is to permit quantitative comparisons of different interpretations.

Classification done from images acquired at different times, classified by different procedures, or produced by different individuals can be evaluated using a pixel by pixel, point-by-point comparisons. The results must be considered in the context of the application to determine which is the "most correct" or "most useful" for a particular purpose. In order to be compared, both the map to be evaluated and the reference map must be accurately registered geometrically to each other. They must also use the same classification scheme they should have been classified at the same level of details. One simple method of comparison is to calculate the total area assigned to each category in both maps and to compare the overall figures.

This type of assessment is called non-site-specific-accuracy. On other hand, site-specific accuracy is based on a comparison of the two maps at specific locations (i.e., individual pixels in two digital images). In this type of comparison, it is obvious that the degree to which the pixel one image spatially align with the pixels in the second image contributes to the accuracy assessment result. Errors in classification should be distinguished from errors in registration or positioning of boundaries. Another useful form of site-specific accuracy assessment is to compare field data or training data at a number of locations within the image, similar to the way spatial accuracy assessment using ground checkpoints is performed for digital orthophotos and terrain models.

Accuracy of image classification is most often reported as a percentage correct. The customer's accuracy is computed using the number of correctly classified pixels to the total number of pixels assigned to a particular category. It takes errors of commission into account by telling the consumer that, for all areas identified as category X, a certain percentage are actually correct. The producer's accuracy informs the image analyst of the number of pixels correctly classified in a particular category as a percentage of the total number of pixels actually belonging to that category in the image. Producer's accuracy measures errors of in omission.

2.6 Landsat Images

The Landsat program is the longest-running enterprise for acquisition of satellite imagery of Earth. On July 23, 1972 the Earth Resources Technology Satellite was launched. This was eventually renamed to Landsat. The most recent, Landsat 8, was launched on February 1 1, 2013. The instruments on the Landsat satellites have acquired millions of images. The images, archived in the United States and at Landsat receiving stations around the world, are a unique resource for global change research and applications in agriculture, cartography, geology, forestry, regional planning, surveillance and education, and can be viewed through the USGS 'Earthexplorer' website. Landsat 7 data has eight spectral bands with spatial resolutions ranging from 15 to 60 meters; the temporal resolution is 16 days.

Types of Landsat instruments are as follows Landsat 1, Landsat 2, Landsat 3, Landsat 4, Landsat 5, Landsat 6, Landsat 7, Landsat 8 etc.

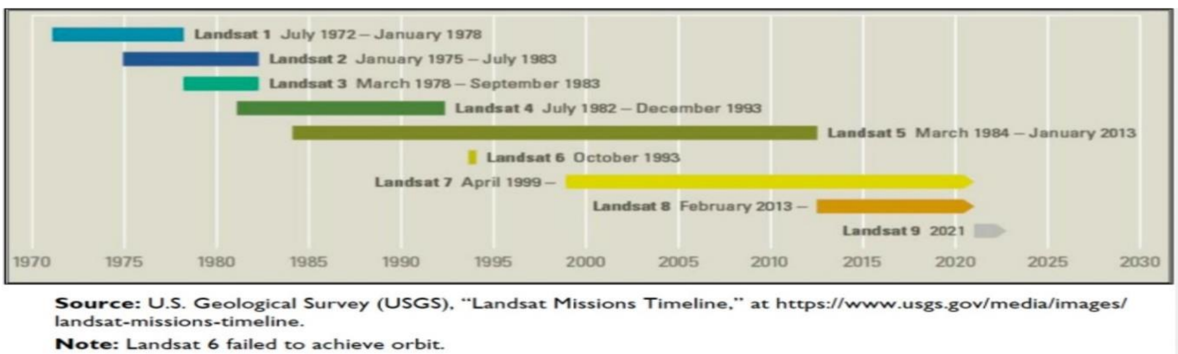

Figure 2.2: *Timeline of Landsat Satellites (USGS, 2021)*

2.7 Land Cover and Land Use Mapping

Although the terms land cover and land use are often used interchangeably, their meaning are quite distinct. Land cover refers to the surface cover on the ground like water, bare soil, vegetation, forest etc., while Land use refers to the purpose the land serves such as recreation, cultivation, fishing, grazing etc.

The properties measured with remote sensing techniques relate to land cover, from which land use can be inferred, particularly with ancillary data or a priori knowledge. Land use applications of remote sensing include the following: natural resource management, wildlife habitat protection, mapping for GIS input, urban expansion / encroachment, routing and logistics planning for seismic / exploration / resource extraction activities, damage delineation (tornadoes, flooding, volcanic, seismic, fire), legal boundaries for tax and property evaluation, target detection - identification of landing strips, roads, clearings, bridges, land/water interface.

2.8 Statistical analysis

Statistical analysis is the process of collecting and analyzing large volumes of data in order to identify trends and develop valuable insights. In the professional world, statistical analysts take raw data and find correlations between variables to reveal patterns and trends to relevant stakeholders. Working in a wide range of different fields, statistical analysts are responsible for new scientific researches such as mapping urban sprawl pattern (Coursera, 2023).

2.8.1 Trend analysis

Trend analysis is the practice that gives us the ability to look at data over time for a long-running survey. Horizontal trends can be helpful in comparing quiz or test scores (see an increase in knowledge over the course if you manage the same survey multiple times over the matter of a few weeks or months) or identifying a trend in data sets for a regularly distributed satisfaction survey. Methods for trend analysis includes parametric and non-parametric, where parametric method rely on the assumption about the nature of distribution. The method supposes a fundamental distribution or normal distribution for the variables of interest an example of this method is least squares linear regression. Non parametric method does not rely on the assumption that data should have a normal distribution (Yue and Wang, 2004). Hydro-meteorological time series data are characterized substantial departure from normality. For such data the non-parametric methods are preferred for detecting monotonic trends because they have higher power than parametric methods (Muthoni, et al., 2018).

The Non-parametric methods for detecting monotonic trends in time series data include Man Kendall, Spearman’s Rho and cumulative rank difference tests. A study by (Gebrechorkos, et al.,2019) and (Yue and Wang, 2004) suggests that Mann-Kendall test is one of the most popular and widely used to detect long-term seasonal and annual trends in and climate datasets. Mann–Kendall test.

CHAPTER 3

METHODLOGY

3. Overview

The methodological part of this study is divided into three major steps; whereas the first part covers the remote sensing data acquisition and pre-processing. The second part covers the image classification and validation and finally, the Spatial patterns identified using classified results of land cover changes. The methodology starts with Data acquisition, data preparation, then data processing and finally data analysis as shown in the schematic diagram in Figure 3.1.

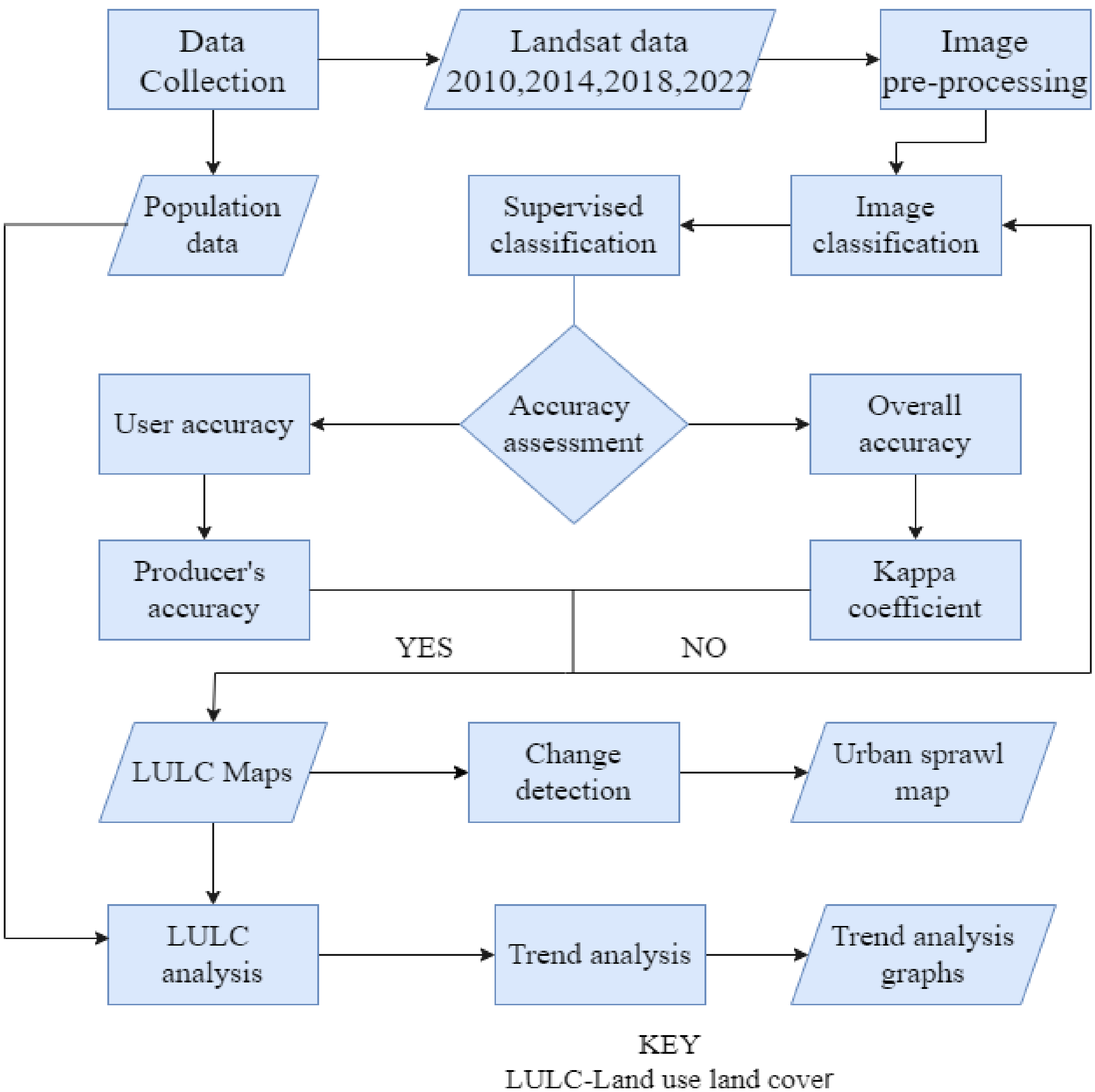

Figure 3.1: *Methodology workflow*

3.1 Data acquisition

The first step in this research was data acquisition, which involved obtaining two types of datasets: multi-spectral satellite imageries with a 30m spatial resolution in raster format. The dataset used in this study included satellite imagery data of the Arusha region. The satellite images covering the study area were downloaded from Google Earth Engine. The downloaded images consisted of Landsat 7 TM and Landsat 8 OLI for the years 2010, 2014, 2018, and 2022, with a 4-year interval, captured in September, and covering the entire Arusha region. The spatial resolution of the images was 30m.

The cloud cover for the sub-scenes was all less than 50%, and their respective path was 169 with row 64. This ensured that the entire area of interest was covered, ensuring accuracy in studying the changes in the urban extent of the Arusha region. Table 3.1 provides a description of the acquired data used in this research.

Table 3.1: Data description

DATA	FORMAT	RESOLUTION	SOURCE	PURPOSE
Landsat images	TIFF	30m	-USGS	Land use and Land cover classification.
Administrative boundary (Arusha Region)	Shape file	-	-Diva GIS (https://www.diva-gis.og)	Study area extraction.
Ground truthing coordinates	CSV	-	-Area of interest	Accuracy assessment.
Population data	CSV	-	National Bureau Statistics (NBS)	Temporal change analysis

3.2 Software used

This includes a list of software used during the execution of this study Table 3.2 summarizes soft wares that were used for data pre-processing, processing, obtaining the expected outputs and finally analyzing the outputs.

Table 3.2: Software used in the study

SOFTWARE	VERSION	PURPOSE
ArcGIS	10.8	-Mapping the classification outputs - For statistical analysis - Change detection
R studio	4.2.2	-Image classification
Erdas Imagine	2014	-Image mosaicking -Image pre-processing
Microsoft Excel	2016	-For plotting graphs

3.3 Data preparation

This includes stages which all datasets were prepared for processing purpose. It covers all the necessary steps in preparation for all Landsat satellite images as follows;

3.3.1 Data cleaning

Landsat images for various intervals were obtained downloaded from United States Geological Survey (USGS), through the link (https://earthexplorer.usgs.gov and preprocessed using for image correction and enhancement of poor quality images and enhancement of low resolution images. Data cleaning is a process which deals with detecting and removing errors and inconsistencies from data in order to improve the quality of data and remove redundancy. The dataset was processed to remove incorrect records, empty values and converted into a format suitable for analysis. This process was done using Erdas Imagine software.

3.3.2 Layer Stacking

This refers to the process of combining more than one Image bands with similar characteristics to form a single composite multispectral image file. Layer stacking was performed as a step of the pre-processing stage; whereas, it was applied to all the Landsat images by using the Layer stacking tools. Erdas Imagine 2014 software was used to perform this activity, layer stacking function was applied in performing Layer stacking of the acquired Landsat images. The purpose of doing layer stacking was getting the required band combination for assessing lake extent changes.

3.3.3 Image subset

The mosaicked images were clipped to obtain the area of interest in order reduce the processing time and storage for other processes. The clipping extent was the Arusha region and its surrounding in Shape file extension. This process was done in the ArcGIS software.

3.4 Data processing

This stage covers all the procedures followed in obtaining the desired outputs from the already prepared data. The aim of doing so is to produce the outputs with respect to the research objectives so as to analyze them and come up with useful information. The procedures are as follows,

3.4.1 Classification

The resulting pre-processing images were classified using supervised classification in R studio software and analyzed for determination of the urban sprawl pattern of the region. The following three main steps were followed in carrying out classification.

i. Training samples were collected from the area of interest as per image obtained after doing image pre-processing. These were selected randomly for each class; whereas five classes were contained in the area of interest i.e. water, bare land, built-up area, vegetation and forest.
ii. Supervised classification of the study area was achieved by using Random forest algorithm which normally considers calculates a response variable by creating many decision trees and then putting each object to be modeled down each of the decision trees.
iii. Accuracy assessment was performed to check the accuracy of the classified Landsat images. This was achieved by using higher resolution Google earth pro images of the respective years being classified. The higher resolution images were used as validation

data to collect the sample points from the respective years. R studio was used in performing accuracy assessment.

Random regional of interest (ROI) for each class of the classified image were selected randomly from higher resolution Google earth pro image and then all selected ROI were named the same names of classified land classes. Confusion matrix using ground truthing ROI on the post classification toolbox was used to match all selected ground truthing. ROI with the land cover classes to check the accuracy of the classified images for all years.

3.5 Data analysis

This presents how the data were analyzed in order to come up with the overall expected output so as to meet the main goal of the study. In this study it mainly covers the statistical analysis basing on the trend analysis of the outputs obtained from land cover maps obtained and change detection was conducted to produce the Arusha urban sprawl.

3.5.1 Trend analysis

Trend analysis is a technique used to examine and predict movements of an item based on current and historical data. In this study, trend analysis was used for the purpose of determining the trend that exist in the Land cover maps of 2010, 2014, 2018 and 2022 with population data for the purpose of studying the changes occurred for the years mentioned. Area and percentage of the built-up area were computed from the land cover data in order to find the relationship between built-up area extents and the population change of Arusha region.

3.5.2 Change Detection

The Land cover maps for the years 2010, 2014, 2018 and 2022 were used as input to detect mainly the built-up area changes. The process was done in the ArcGIS software where the detection was conducted in the span of four years and this led to production of three change detection maps which were used in mapping of urban sprawl pattern of Arusha region. As seen in the figure 4.5.

CHAPTER FOUR

RESULT AND ANALYSIS

4. Overview

This chapter involves result and analysis of the research findings which obtained through the implementation of the research methodology according to the intended objective of this research which was to map the urban sprawl pattern of Arusha region.

4.1 Land Cover Maps

The LULC maps were classified using supervised method with the Random forest classifier algorithm and class that were used are vegetation, forest, water bodies, bare land and built up. LULC maps produced were of 2010, 2014, 2018 and 2022 as shown in the figures 4.1, 4.2 and 4.3, 4.4.

4.1.1 Land Cover Map of 2010

The Land cover map of 2010 seems the area is more occupied with vegetation and bare land as seen in the figure 4.1 below, which has small number of built up areas.

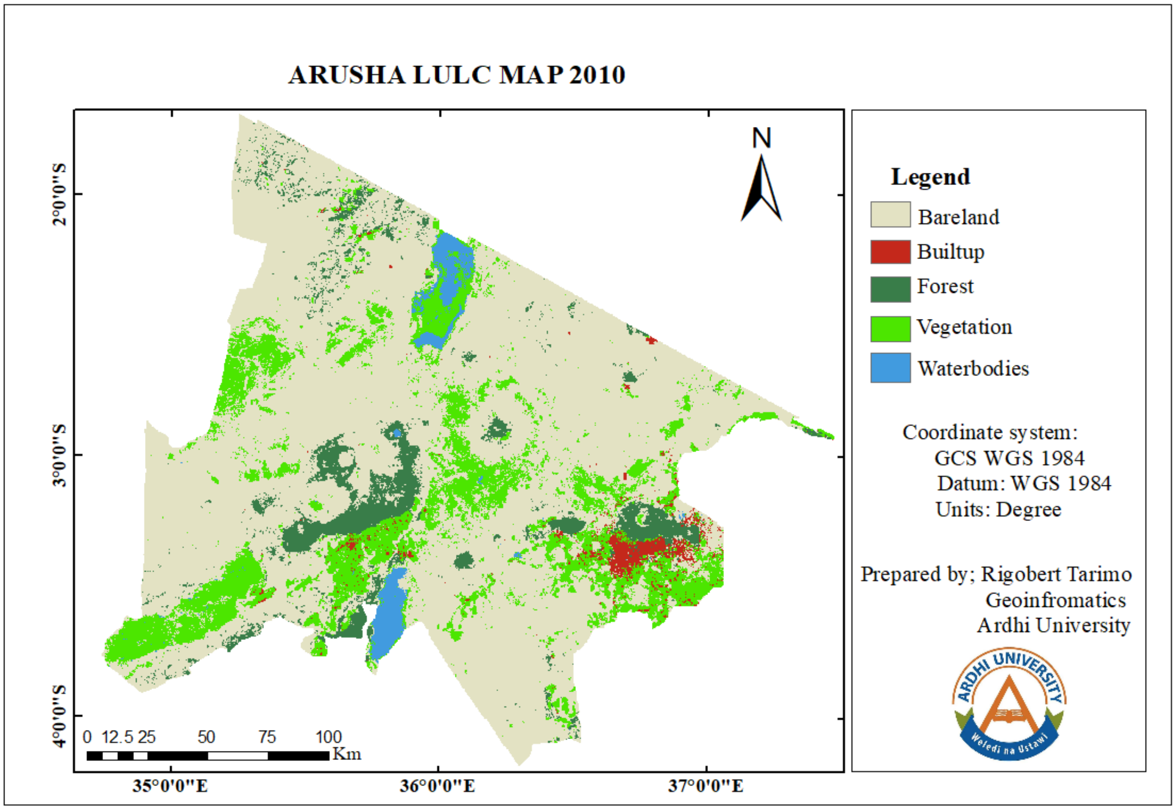

Figure 4.1: Land cover map of Arusha region 2010

4.1.2 Land Cover Map of 2014

The Land cover map of 2014 seems the area is more occupied with vegetation and bare land with the water bodies and some built-up areas as seen in the figure 4.2 below, which has small number of built up areas.

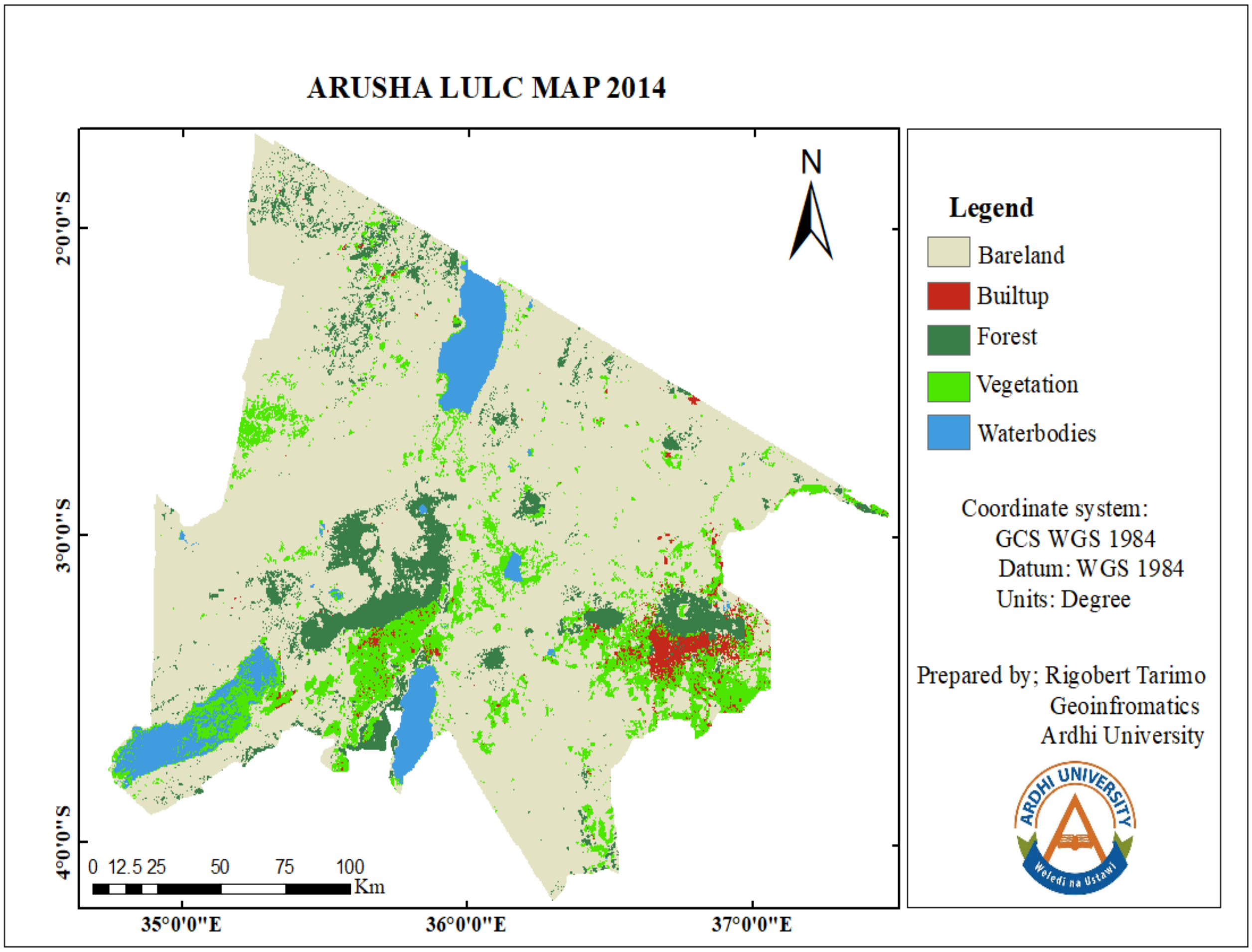

Figure 4.2: *Land cover map of Arusha region 2014*

4.1.3 Land Cover Map of 2018

The Land cover map of 2018 there is increase in forest maybe due to increase in rainfall and built up area and as seen in the figure 4.3 below, which has small number of vegetation and bare land.

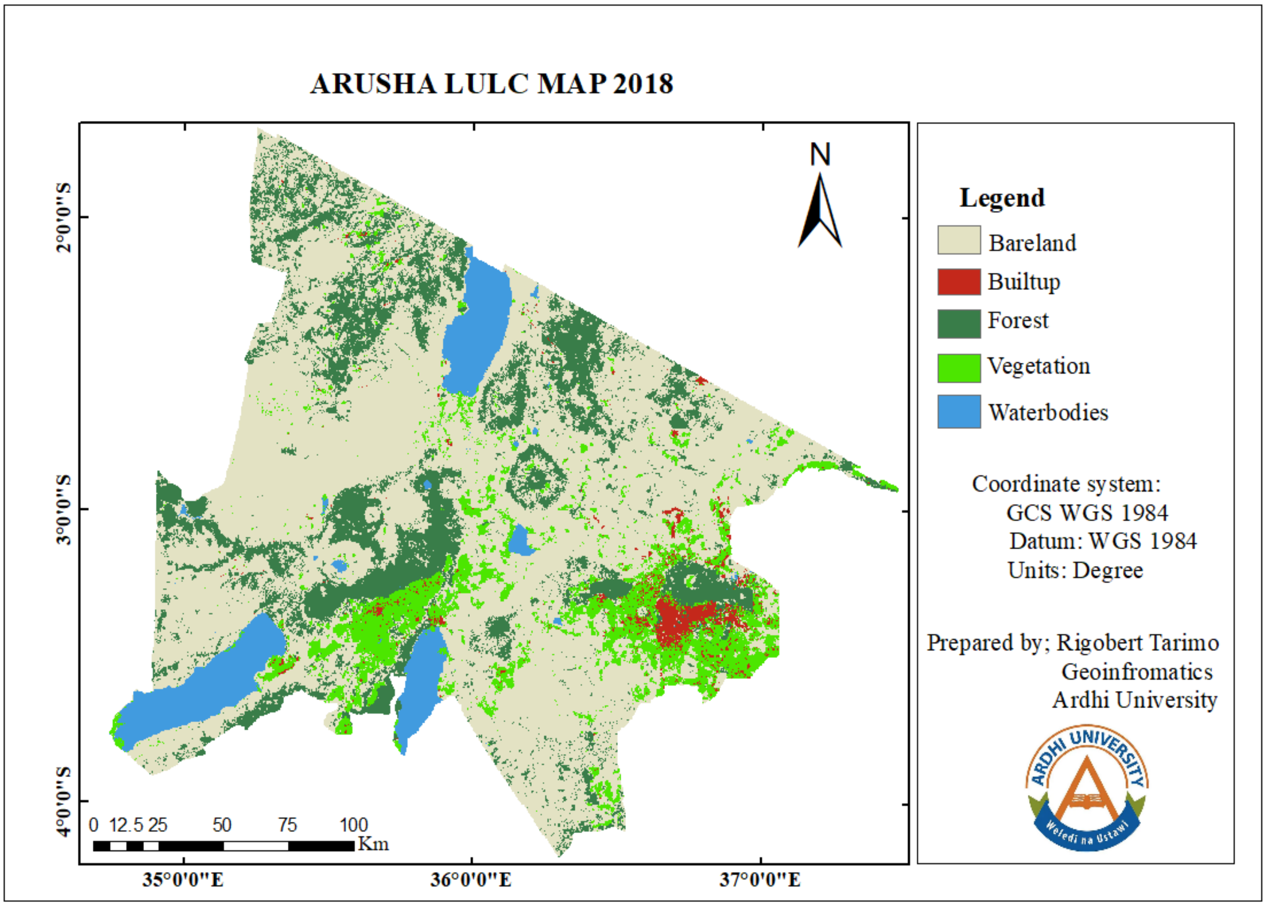

Figure 4.3: *Land cover map of Arusha region 2018*

4.1.4 Land Cover Map of 2022

In the land cover map of 2022 there is also increase in built up area mostly in Arusha district and bare land which facilitate the decrease in vegetation and forest due to maybe increase in temperature in the region, see in the figure 4.4 below

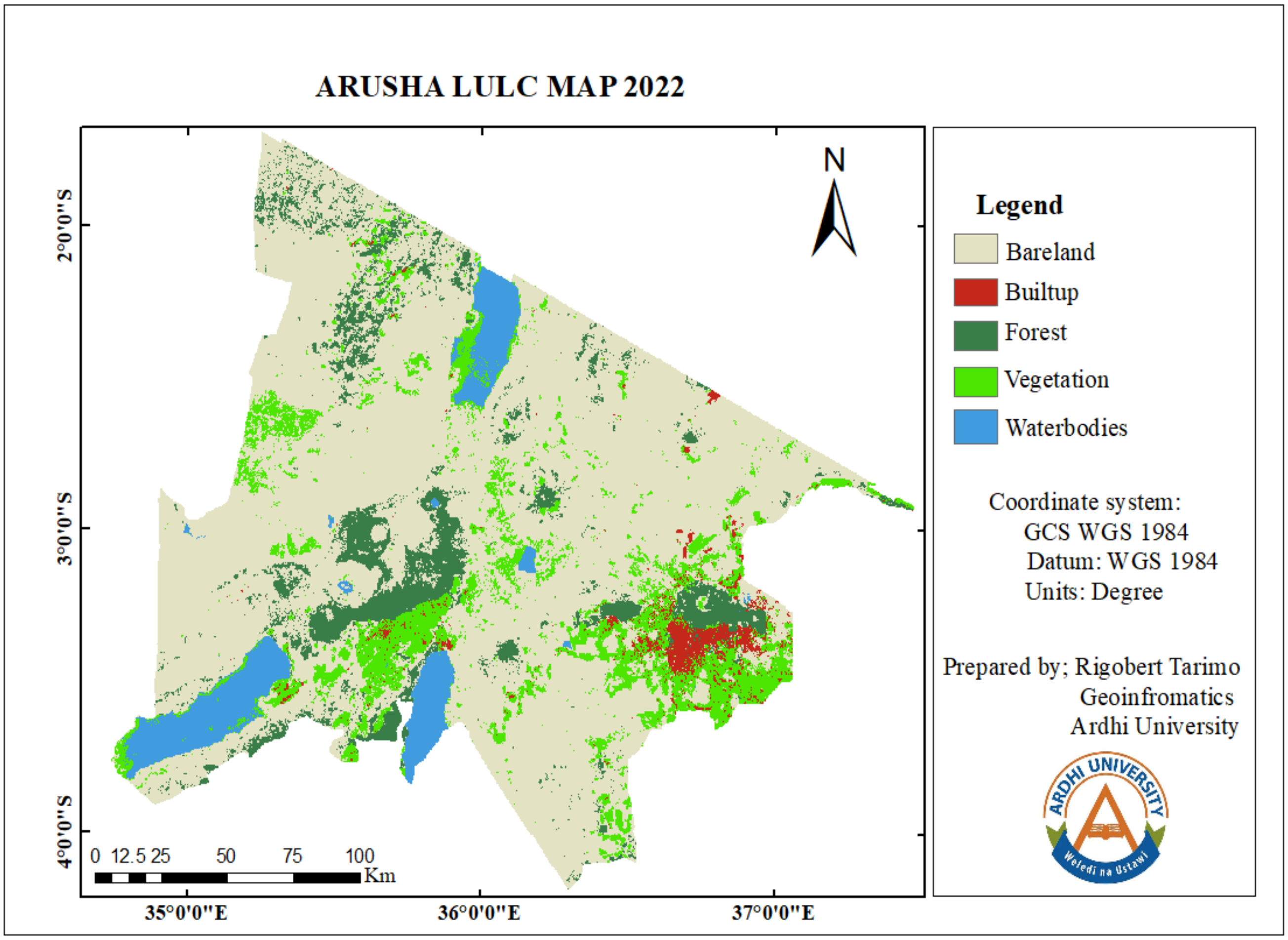

Figure 4.4: *Land cover map of Arusha region 2022*

4.1.5 Accuracy assessment summary

A summary user and producer's classification accuracies as well as the overall accuracies of the classification for the respective years are provided in table 4.1. The table summarizes the accuracy assessment of land covers classification for 2010, 2014, 2018 and 2022 which were 87.84%, 85.76%, 91.89%, and 90.2% respectively and kappa coefficients were 0.84, 0.82, 0.89 and 0.88 as shown in the table 4.1.

Table 4.3: Accuracy assessment summary

Class	2010		2014		2018		2022	
	Producer. Acc. (%)	User Acc. (%)	Producer. Acc. (%)	User Acc. (%)	Producer. Acc. (%)	User Acc. (%)	Producer. Acc. (%)	User Acc. (%)
Vegetation	95.65%	100%	97.1%	94.37%	98.55%	100%	91.30%	95.45%
Forest	100%	100%	100%	98.46%	100%	100%	94.44%	100%
Water bodies	82.61%	74.51%	67.34%	73.81%	91.3%	85.71%	89.13%	80.39%
Bare land	75.44%	78.18%	71.93%	73.21%	73.68%	87.5%	84.21%	92.31%
Built up	81.67%	81.67%	84.75%	81.97%	93.33%	83.58%	86.67%	81.25%
Overall Kappa Statistics	0.84		0.82		0.89		0.88	
Overall accuracy	87.84%		85.76%		91.89%		90.2%	

4.2 Arusha Region urban sprawl

The urban area extent of the Arusha region exhibited variations between 2010 and 2022, reflecting the changing nature of urbanization over the years. These variations demonstrate an overall increase in the urban sprawl of the Arusha region. Notably, the urban sprawl is more prominent in the Arusha district and Arusha municipal, as depicted in Figure 4.5, which integrates remote sensing images and utilizes change detection techniques. The results indicate that the east southern part of the region are more developed and congested, following a leapfrog pattern of expansion.

This pattern suggests that services and infrastructure, such as piped water supply, schools, healthcare facilities, and solid waste management, are primarily concentrated in the city center, gradually diminishing as one moves outward from the center within the region.

The urban map of Arusha illustrates the significant growth of the urban area, particularly in the Arusha district and Arusha municipal, characterized by a cluster expansion and leapfrog sprawl pattern. Figure 4.6 clearly demonstrates the unchecked expansion of urban areas into undeveloped land, leading to urban sprawl within the region. The map specifically highlights the selected areas of Arusha district and Arusha municipal, where more than 45% of the region's population is concentrated.

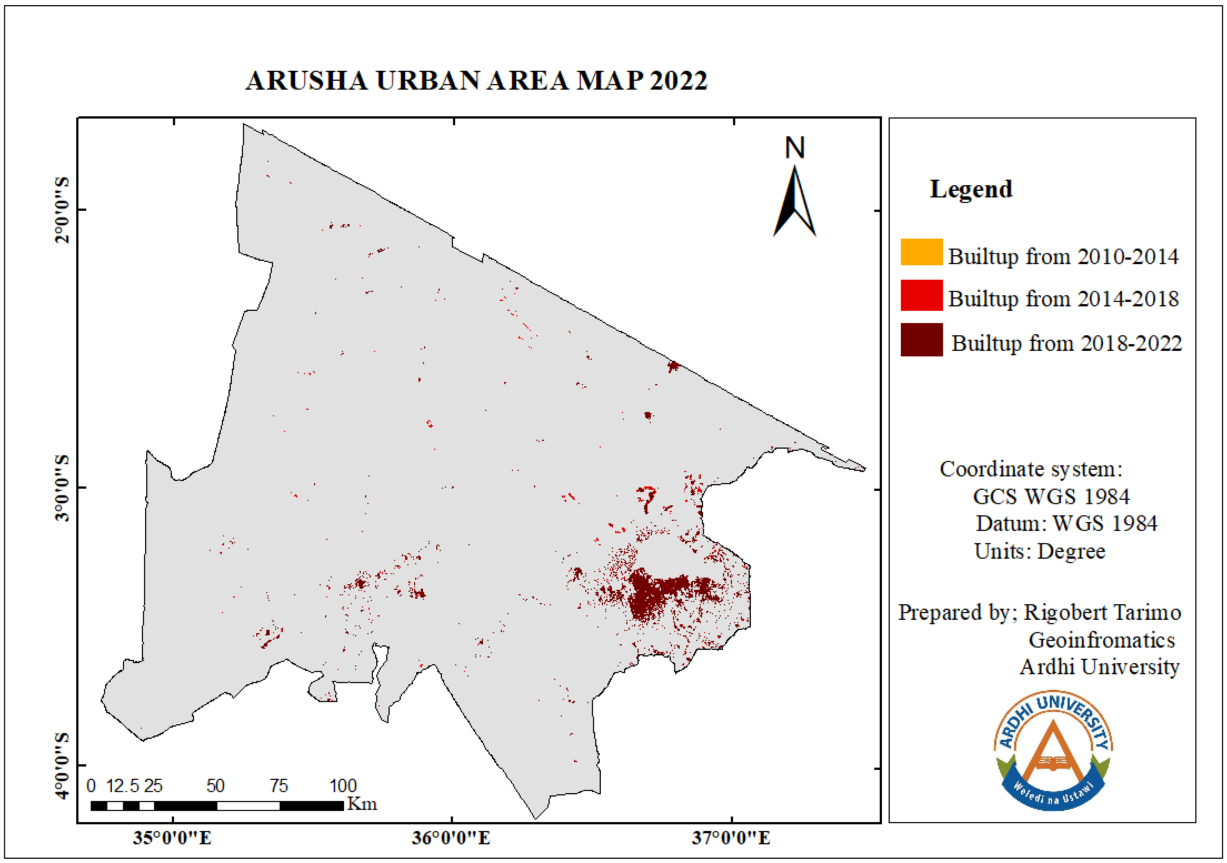

Figure 4.5: *Arusha region urban area map*

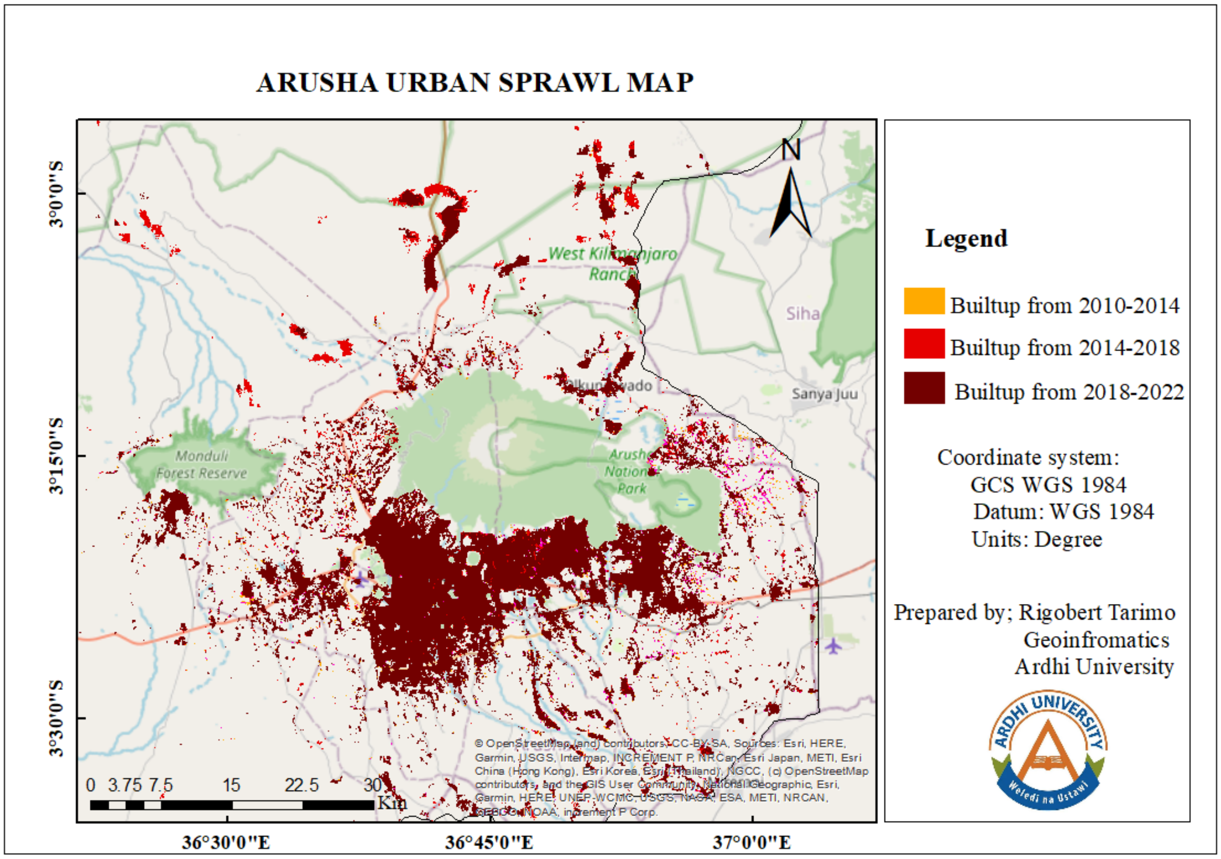

Figure 4.6: Arusha urban sprawl map

4.3 Statistical analysis results of Arusha urban sprawl

The statistical results from the calculation of trend analysis for changes in land use and land cover indicate that the built-up area has been growing in relation to the population for the years 2010, 2014, 2018, and 2022, as presented in Table 4.2.

Table 4.4: Proportion of Built-up change for the years 2010, 2014, 2018 and 2022

Year	2010	2014	2018	2022
Non built up area	37,007 km^2	36,960 km^2	36,929 km^2	36,861 km^2
Built up area	569 km^2	616 km^2	647 km^2	715 km^2
Total area of the region	37,576 km^2	37,576 km^2	37,576 km^2	37,576 km^2

Percentage (%) of non-built up area	98.49%	98.36%	98.28%	98.10%
Percentage (%) of built up area	1.51%	1.64%	1.72%	1.9%

4.3.1 Trend analysis for rapid urban population between 2010-2022

Table 4.5: Population analysis between 2010-2022

Interval	Population increase	Built up increase (%)	Remarks
2010-2014	1,665,000-1,788,318	1.51%-1.64%	The population of the Arusha urban area increased from 1,665,000 people to a projected population of 1,788,318 people during this period. The Arusha district stands out as the most densely populated area..
2014-2018	1,788,318-1,999,907	1.64%-1.72%	Significant increase was recorded at this stage from 1,788,318 people to 1,999,907 people as projected is also as a continuous increase maybe in tourism activities and other economic activities.
2018-2022	1,999,907-2,356,255	1.72%-1.9%	There was increase in urban expansion during this period because most expansion in urban land was much within the seven districts in Arusha region especially Arusha municipal.

4.3.2 Population density analysis of Arusha region

The generated urban sprawl map of the Arusha region, the population data for the year 2022 was utilized to generate the population density map. This analysis aimed to examine the areas with higher concentrations within the region and establish the relationship between urban sprawl and population density. Figure 4.6, which represents the population density map, illustrates that the Arusha municipal and Arusha district are the most densely populated areas compared to other areas in the region.

Based on the analysis, the total population of the Arusha region is 2,356,255, with 45.3% of the population residing in the Arusha municipal and district. The remaining percentage is distributed among other parts of the region. This population density is crucial in urban planning since it influences decisions related to land use, housing and transportation so these two districts are in need of compact and efficient urban designs so as to accommodate larger population within limited spaces.

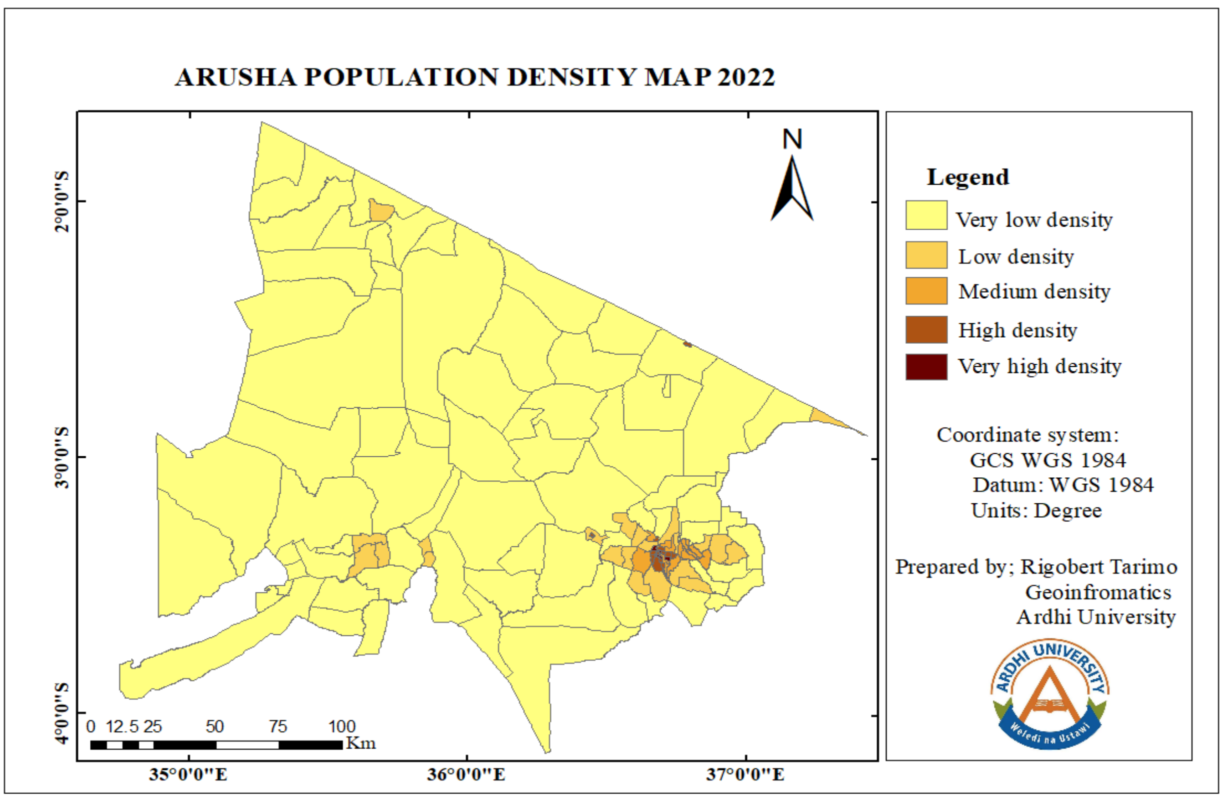

Figure 4.7: Population density map 2022

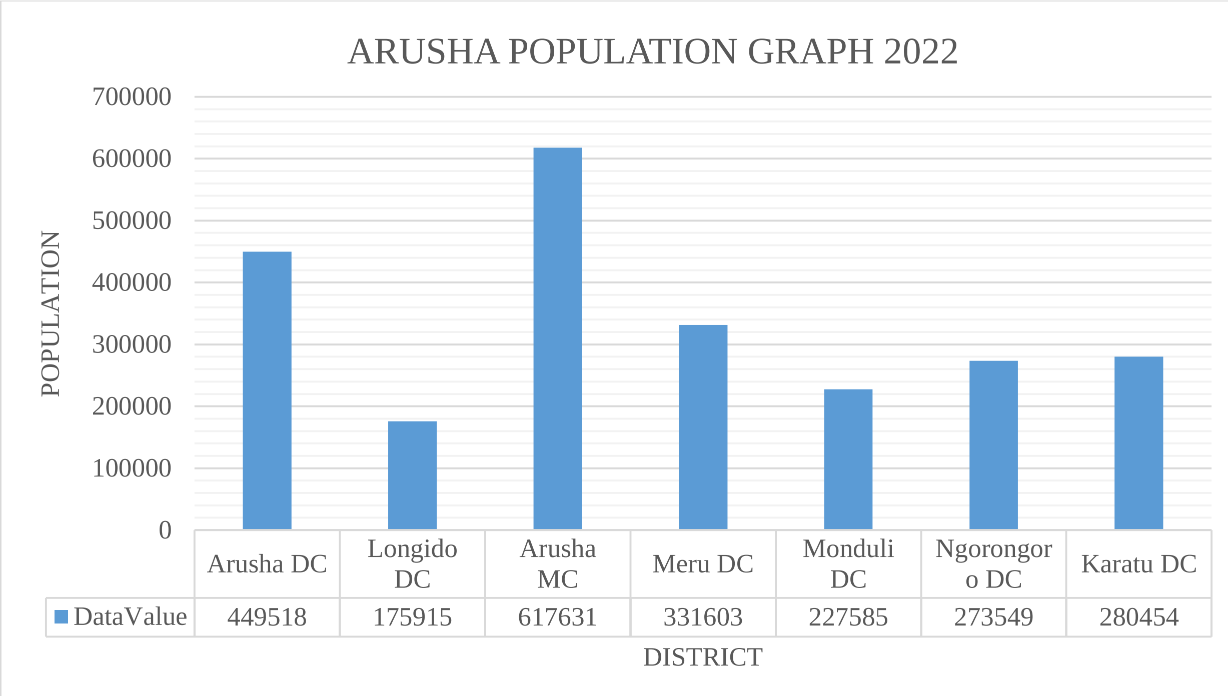

Figure 4.8: *Arusha region population graph 2022*

4.3.3 Trend analysis for built-up area between 2010-2022

Table 6.4: Trend analysis for Built-up area between 2010-2022

Interval	Built-up increase (km^2)	Built-up increase (%)	Remarks
2010-2014	$569km^2$-$616km^2$	1.51%-1.64%	Since the base year of 2010, the Arusha urban area has experienced a cluster growth pattern primarily concentrated in Arusha Municipal and District, where the built-up area expanded from approximately 569 km^2 to 616 km^2.
2014-2018	$616km^2$-$647km^2$	1.64%-1.72%	There was a limited increase in urban

			expansion during this period, characterized by leapfrog growth in the Arusha region and cluster growth in the Arusha district.
2018-2022	647km^2-715km^2	1.72%-1.9%	A significant increase was recorded during this period, which can be attributed to further outward expansion from Arusha town and contributes to a leapfrog pattern in the Arusha municipal and district areas.

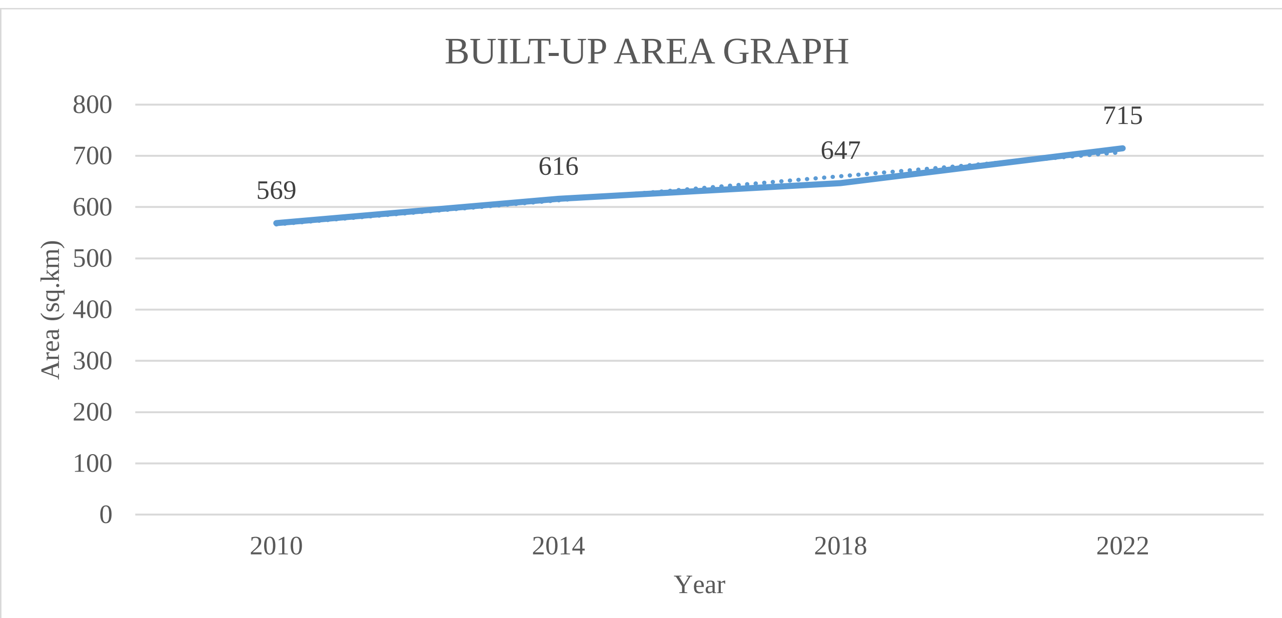

Figure 4.9: *Built-up area graph*

4.3.4 Rate and magnitude of change

Table 4.7: Rate and Magnitude of Change

Interval	Built-up increase (km^2)	Built-up increase (%)	Trend Percentage Change (Observed change/ Sum of change × 100)
2010-2014	47 km^2	0.13%	32.19%
2014-2018	31 km^2	0.08%	21.23%
2018-2022	68 km^2	0.18%	46.58%
TOTAL	**146 km^2**	**0.39%**	**100%**

4.4 Temporal change analysis

The temporal change analysis was conducted to examine and interpret the changes in the extent of the built-up area and population. The analysis revealed an upward trend in population growth for the years 2010, 2014, 2018, and 2022, along with an increase in the built-up area, which measured 569 km^2, 616 km^2, 647 km^2, and 715 km^2 respectively for those years. The analysis concluded that an increase in the region's population directly correlates with an expansion in the extent of the built-up area, as they are directly proportional.

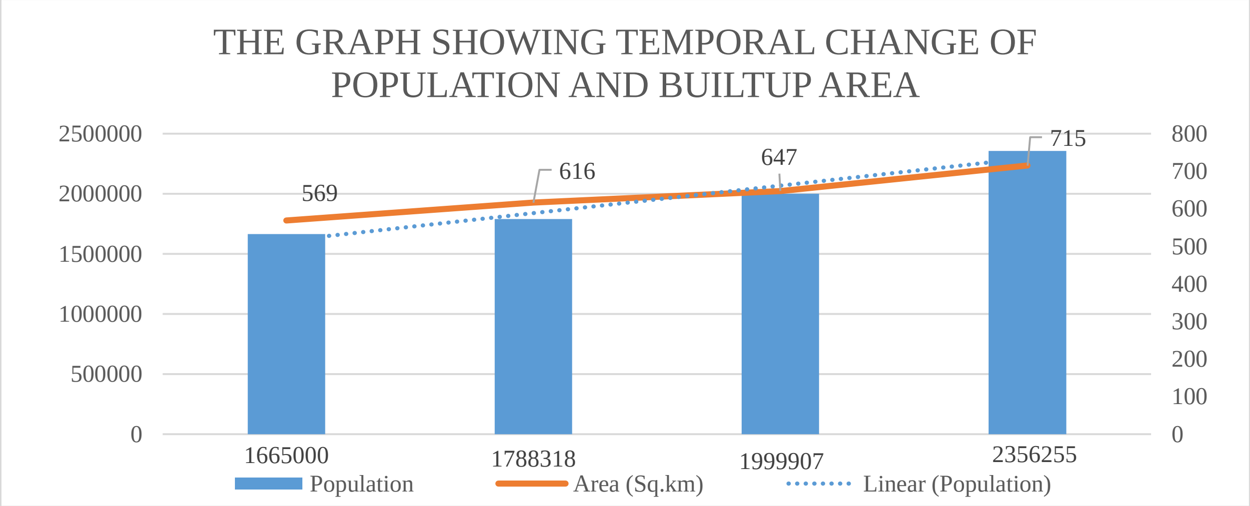

__Figure 4.10:__ Population and built-up area temporal change graph

4.5 Discussion of the results

Computing the rate of urban growth for Arusha was achieved in this research by using the mapping, measure area, and map calculator modules in GIS to compute the area statistics of the built-up area extent of the Arusha region for the four epochs. The difference in the built-up extent area between each epoch and the preceding one was computed, and based on the statistics, the growth rate for the Arusha region area is estimated to be 636.75 km^2 annually. Furthermore, the change detection technique successfully extracted the extent of Arusha's urban sprawl, and the region's built-up area continuously increased its extent over a period of twelve years from 2010 to 2022. The percentage increase was calculated as 0.13%, 0.08%, and 0.18% for the periods 2010-2014, 2014-2018, and 2018-2022, respectively, with an annual increase of 0.13%.

This excessive physical growth rate is likely to result in further sprawl, along with its associated implications, as there is no reason to suggest that growth will occur in any other form or pattern.

The research also analyzed the population growth for the years 2010, 2014, 2018, and 2022, as it played a significant role in influencing the urban sprawl of the Arusha region. From the research, the population density map for the year 2022 was established, revealing that the majority of the population is concentrated in Arusha Municipal and Arusha District. These two districts account for more than 1 million people, representing over 45% of the population in the region. The constructed urban sprawl map and population map indicate a correlation between them. This surge in population has led to an increased demand for housing and commercial spaces, resulting in outward expansion to accommodate the growing urban population. Additionally, the presence of economic opportunities such as mining and tourism, which provide employment opportunities in the city center, has attracted migrants and further fueled urban sprawl. This relationship is depicted in Figure 4.9.

CHAPTER FIVE

CONCLUSION AND RECOMMENDATION

5.1 Conclusion

Urban sprawl has increasingly become a major issue in the global trend towards urbanization. It is faced not only by developed countries but also by developing countries, large urban centers, and medium and small cities alike. Urban sprawl raises social and environmental concerns while exhibiting a multiplicity of divergent trajectories that somehow defy the dominance of homogeneous characteristics around the world.

This research investigates evidence of urban sprawl in the Arusha region using geospatial approaches, including GIS and remote sensing data. The research contributes to the existing literature on spatial analysis of urban land expansion and sprawl studies from three perspectives. First, this study uses Landsat images to extract land cover data for the region and population estimates. The use of population data allows for a more authentic understanding of population distribution within the region. Second, the study examines the extent of the built-up area in relation to the population data for the chosen estimated years. Third, in addition to examining the physical expansion of urban land from 2010 to 2022, the analysis extends further to investigate the most affected areas with sprawl in the region.

The results of this research indicate concrete evidence of sprawling development in the Arusha region from 2010 to 2022, demonstrated by a substantial increase in the built-up area and a significant increase in the region's population. The research shows that most of the built-up area is concentrated in the Arusha district and Arusha municipal areas, which correlates with the population distribution in those areas. It is evident that most of the population is concentrated in those districts compared to other districts in the region. This concentration of people in specific areas leads to various issues, such as unplanned settlements caused by the population increase.

Furthermore, the research reveals a changing trend in the built-up areas. From 2010 to 2014, the increase was 32.19%. From 2014 to 2018, it was 21.23%, and from 2018 to 2022, it was 46.58%. These increases correspond to the population growth within the region. The research also provides the percentages of the built-up area for the years 2010, 2014, 2018, and 2022, which were 1.51%, 1.64%, 1.72%, and 1.9%, respectively, for each determined year. This equates to an approximate overall increase of about 7% in the built-up area from 2010 to 2022, calculated by summing up all the percentage changes in the built-up area in the region.

5.2 Recommendation

The research recommends using the obtained findings as a basis for further research in order to supplement the existing spatial information and address various spatial-related problems. Urban planners should utilize the current and updated information about the region to identify the most affected areas, which, in this research, are primarily the Arusha district and Arusha municipal areas. This knowledge will facilitate proper planning of these areas, reducing potential damage and ensuring the provision of essential services such as roads, hospitals, and schools, thereby promoting sustainable living.

Additionally, future studies should focus on providing information about both formal and non-formal settlements across the region. This will enable the identification of a comprehensive profile of sprawl encompassing both formal and informal residential development in the Arusha region.

In general, there is ample scope for future studies to ensure that cities like Arusha experience growth and are planned with the aim of improving the quality of life for all residents, regardless of whether they reside in planned or informal settlements (Penrose et al., 2010). Such endeavors are crucial and relevant, as the existing literature recognizes the importance of governance, planning, and land management in achieving sustainable and proper urbanization while preventing or minimizing the negative impacts of sprawl. This is particularly pertinent for cities in developing countries like the Arusha region, which are experiencing rapid growth and require enhanced planning and government interventions.

REFERENCES

Abaje, Ati, and Iguisi (2012), Climate Variability and Change, Impacts and Adaptation Strategies in Dutsin-Ma Local Government Area of Katsina State, Nigeria.

Abimbola, T. (2008). The Planning Implications of Urban Sprawl in Akure, 44th ISOCARP Congress.

Aguda, A., and S. Adegboyega. (2013). *Evaluation of spatio-temporal dynamics of urban sprawl in Osogbo*, Nigeria using satellite imagery & GIS techniques. International Journal of Multidisciplinary and Current Research 1:60–73.

Bhatta, B. (2009). Analysis of urban growth pattern using remote sensing and GIS: a case study of Kolkata, India. *International Journal of Remote Sensing* 30 (18):4733–46. doi:10.1080/0143116080265196.

Bullard, R.D. (2000). Sprawl City: Race, Politics and Planning in Atlanta. Island Press, Aug. 1. Carruthers & Ulfarsson, (2002), Fragmentation and sprawl: Evidence from Interregional Analysis. Growth and Change, Vol . 33. Issue 3. Downs, A. (1999) "Some Realities about Sprawl and Urban Decline". Housing Policy Debate 10 (4): 955-974. Duany, A., Plater-Zyberk, E., & Speck, J. (2000). [BOOK REVIEW] Suburban nation, the rise of sprawl and the decline of the American Dream. American Prospect, 11(13), 46-49.

Carruthers & Ulfarsson, (2002), Fragmentation and sprawl: Evidence from Interregional Analysis. Growth and Change, Vol . 33. Issue 3.

Cervero, R., & Kockelman, K (1997). *Travel demand and the 3Ds: Density, diversity and design. Transportation Research Part D*: Transport and Environment, 2(3), 199-219.

Checkoway, B. 1980. Large builders, federal housing programs, and postwar suburbanization. International Journal of Urban and Regional Research 4 (1):21–45. doi:10.1111/j.1468-2427.1980.tb00350.x. Cobbinah, P. B., and H. N. Aboagye. 2017. A Ghanaian twist to urban sprawl. Land Use Policy 61:231–41. doi:10.1016/j. landusepol.2016.10.047.

Haimanot, M. (2009). Temporal urban growth analysis and potential site identification for high rising buildings in Bahir Dar City.

Jiang, L.,&Li, X. (2009). *Mapping urban sprawl using GIS and remote sensing*: a review of the recent literature. Progress in physical Geography, 33(1), 73-94.

Kalabamu, F. T. 1992. Tanzania: Developing urban residential land. Journal of Urban Affairs 14 (1):61–78. doi:10.1111/ j.1467-9906.1992.tb00275.x.

Kironde, J. L. 1995. Access to land by the urban poor in Tanzania: some findings from Dar es Salaam. Environment and Urbanization 7 (1):77–96. doi:10.1177/095624789500700111.

Kironde, J. L. 2000. Understanding land markets in African urban areas: the case of Dar es Salaam, Tanzania. Habitat international 24 (2):151–65. doi:10.1016/S0197-3975(99)00035-1.

Kleemann, J., J. N. Inkoom, M. Thiel, S. Shankar, S. Lautenbach, and C. Furst. € 2017. Peri-urban land use pattern and its relation to land use planning in Ghana, West Africa. Landscape and Urban Planning 165:280–294. doi:10.1016/j. landurbplan.2017.02.004.

Kombe, W. J. 1994. The demise of public urban land management and the emergence of informal land markets in Tanzania: A case of Dar-es-Salaam city. Habitat International 18 (1):23–43. doi:10.1016/0197-3975(94)90037-X.

Kombe, W. J. 2005. Land use dynamics in peri-urban areas and their implications on the urban growth and form: the case of Dar es Salaam, Tanzania. Habitat International 29 (1):113–35. doi:10.1016/S0197-3975(03)00076-6.

Laidley, T. 2016. Measuring Sprawl. Urban Affairs Review 52 (1):66–97. doi:10.1177/1078087414568812.

Linard, C., A. J. Tatem and M. Gilbert. 2013. Modelling spatial patterns of urban growth in Africa. Applied Geography 44:23–32. doi:10.1016/j.apgeog.2013.07.009. Mercer, C. 2017. Landscapes of extended ruralisation: postcolonial suburbs in Dar es Salaam, Tanzania. Transactions of the Institute of British Geographers 42 (1):72–83. doi:10.1111/tran.12150.

Mbise, R. J., Balira, R. S., & Niwagaba, C. B. (2018). *Urban sprawl and its environmental impacts in Dar es Salaam, Tanzania*. International Journal of sustainable Built Environment, 7(1), 45 53.

Penrose, K., M. C. de Castro, J. Werema, and E. T. Ryan. 2010. Informal urban settlements and cholera risk in Dar es Salaam, Tanzania. PLoS Negl Trop Dis 4 (3):e631. doi:10.1371/journal.pntd.0000631.

Sumari, N. S., H. Tanveer, Z. Shao, and E. Simon 2019b. “Geospatial Distribution and Accessibility of Primary and Secondary Schools : A Case of Abbottabad City, Pakistan.” Proceedings of the International Cartographic Association 29th International Cartographic Conference (ICC 2019), Tokyo, Japan, July, 15–20. doi:10.5194/icaproc-2-125-2019.

Sumari, N. S., Z. Shao, M. Huang, C. A. Sanga, and J. L. Van Genderen. 2017. “Urban Expansion: A Geo-spatial Approach for Temporal Monitoring of Loss of Agricultural LAND.”

International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives 42 XLII-2/W7 (2W7): 1349–1355. doi:10.5194/isprs-archives-XLII-2-W7-1349-2017.

UN-HABITAT, 2010. *Informal Settlements and Finance in Dar Es Salaam, Tanzania.* Nairobi: UN-HABITAT

Yalcin, R., Ozkan, K., & Akkemik, U. (2015). *Mapping urban sprawl using remote sensing data*: a case study of Instabul, Turkey.

Yin, L., Q. Cheng, Z. Wang, and Z. Shao. 2015. “Big Data for Pedestrian Volume : Exploring the Use of Google Street View Images for Pedestrian Counts.” Applied Geography 63:: 337–345. doi:10.1016/j.apgeog.2015.07.010.

Yuan, T.; Lee, H.; Jung, H.C. Toward estimating wetland water level changes based on hydrological sensitivity analysis of PALSAR backscattering coefficients over different vegetation fields. Remote Sens. 2015, 7, 3153–3183

Zhang, G., X. Guo, D. Li, and B. Jiang. 2019c. “Evaluating the Potential of LJ1-01 Nighttime Light Data for Modeling Socio-economic Parameters.” Sensors 19: 6. doi:10.3390/s19061465. Zhang, H., X. Ning, Z. Shao, and H. Wang. 2019a. “Spatiotemporal Pattern Analysis of China’s Cities Based on High-resolution Imagery from 2000 to 2015.” In ISPRS International Journal of Geo-Information 8: 241. doi:10.3390/ijgi8050241.

Zhang, L., Q. Weng, and Z. Shao. 2017. “An Evaluation of Monthly Impervious Surface Dynamics by Fusing Landsat and MODIS Time Series in the Pearl River Delta, China, from 2000 to 2015.” Remote Sensing of Environment 201 (August): 99–114. doi:10.1016/j.rse.2017.08.036.

Zhang, L., Z. Shao, J. Liu, and Q. Cheng. 2019b. “Deep Learning Based Retrieval of Forest Aboveground Biomass from Combined LiDAR and Landsat 8 Data.” Remote Sensing 11: 1459. doi:10.3390/rs11121459.