Demographic Change in the Upper Tsitsa Catchment:

The Integration of Census and Land Cover Data

for 2001 and 2011.

A full thesis submitted in fulfilment of the requirements for the degree of

MASTERS OF SCIENCE

In the Department of Geography,

Rhodes University

by

Danuta Lorina Hodgson

June, 2017

Supervisor

Professor R.C. Fox (Department of Geography, Rhodes University)

RHODES UNIVERSITY
Where leaders learn

environmental affairs
Department:
Environmental Affairs
REPUBLIC OF SOUTH AFRICA

ABSTRACT

The purpose of this research was to determine if the integration of census and land cover data could provide evidence of spatial patterns and temporal change for the Upper Tsitsa Catchment. This thesis contributed to academic literature with regards to dasymetric mapping and provided a database for the Ntabelanga and Laleni Ecological Infrastructure Project. The study took place in the Upper Tsitsa Catchment which is located in the north-eastern region of the Eastern Cape and falls within the uMzimvubu Catchment. South Africa National Population Censuses for 2001 and 2011 and the National Land Cover data sets for 2000 and 2013/14 were used to create dasymetric maps depicting demographic changes over time for the catchment area. Spatial statistics were performed on the dasymetric and choropleth map to determine the accuracy of the data that was created. From the results, it was found that although the statistics were skewed, the method was more accurate in displaying the population densities, which was noted during the sampling process of the spatial analysis. It was found that there had been a decrease in the population density within the catchment. This affected the density of several other variables such as population race group, language and employment status statistics as decreases in these values could be due to individuals migrating out of the catchment as well as socio-economic upliftment, such as having better access to services. The use of dasymetric mapping allowed an accurate representation of the population density from the census data to be created. The results of the dasymetric mapping were more accurate as they depicted where the population within the enumeration areas were located, and recognised that some areas were populated while some areas were not. To conclude, it was found that using dasymetric mapping provided reliable and useful data about population density and enables comparison over time.

Keywords: *Dasymetric Mapping, Census, Land Use, Upper Tsitsa Catchment, Spatial Change, South Africa.*

Table of Contents

List of Figures

List of Tables

DECLARATION

I have read and understood the Rhodes University plagiarism policy. This study presents original work by the author and has not plagiarised the works of others. Where use has been made of the works of other authors, it is acknowledged in the text accordingly.

_____

Danuta Lorina Hodgson

Disclaimer: The opinions expressed and conclusions drawn within this research thesis are those of the author, unless otherwise stated. These thoughts may not necessarily represent the views of Rhodes University or the funders.

ACKNOWLEDGEMENTS

I would firstly like to thank my thesis supervisor, Prof. Roddy Fox who introduced me to this field within geography. His door was always open whenever I needed his guidance and advice. Thank you for your continual guidance throughout this research.

I would like to acknowledge the Department of Environmental Affairs – National Research Management for providing the financial support to carry out this research and allowing me the opportunity of a Master's degree.

I would like to express my sincerest gratitude to Mr. Michael Powell for allowing me the opportunity to further my studies and for his continuous motivation and support throughout this study.

I would also like to thank the Catchment Research Group and Prof. Kate Rowntree for their guidance and support in this research.

Many thanks to the following persons who have helped me along the way: Dr. Benjamin van der Waal, Mrs. Karen Milne, Mrs. Glyn Armstrong, Mr. John Landman, Ms. Deidre Fouche, Mr. Sakhile Singata, Mr. Mthunzi Gumede and Mr. Abe Ngoepe.

Finally, I must express my very profound gratitude to my parents for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

Abbreviations

BGIS	Biodiversity GIS
CBD	Central Business District
CD:NGI	Chief Directorate: National Geo-Spatial Information
CRG	Catchment Research Group
DAFF	Department of Agriculture, Forestry and Fisheries
DWA	Department of Water Affairs
DEA	Department of Environmental Affairs
EA	Enumeration Area
FAO	Food and Agriculture Organization of the United Nations
GIS	Geographical Information Systems
LCCS	Land Cover Classification Scheme
MAUP	Minimum Areal Unit Problem
MDB	Municipal Demarcation Board
NLC	National Land Cover
NRM	National Resource Management
NLEIP	Ntabelanga Laleni Ecological Infrastructure Project
SAL	Small Area Layer
SPOT	Satellite Pour l'Observation de la Terre
Stats SA	Statistics South Africa
UN	United Nations
UNEP	United Nations Environment Programme
UN-GGIM	United Nations Committee of Experts on Global Geospatial Information

Chapter One: Introduction

1.1 Context

Understanding linkages and interactions between physical, social and economic environments are important in studies regarding sustainable development and changes in land use and land cover and how people utilise land have become recognised as an important global environmental change (Huby *et. al.,* 2007; Turner, 2002). Anthropogenic impacts on the environment have led to an assortment of ecological problems and benefits existing within a region (Huby *et. al.,* 2007). There is evidence in various academic fields that spatial data are often incompatible due to the various scales and nature in which the data are collected (Huby *et al.,* 2007). Consequently, there have been numerous studies done globally, either using census data or remote sensing data to determine temporal patterns and trends in demographics, land cover and land use, but seldom have these two sources of data been integrated to determine demographic change over time and the impacts that it has on land cover and land use practices.

Environmental data such as land cover are depicted by physiographic or natural boundaries unlike census data which utilise enumeration boundaries that do not conform to the natural environment (Mans, 2011). Land cover is the physical surface of the Earth which is made up of vegetation, soil, rocks, water and man-made elements (CD:NGI, 2009). Land cover can provide information regarding socioeconomic trends, but does not contain characteristic data as found in census data (Martin and Bracken, 1993).

Census data contains valuable information about a country's population in the form of population density, age, sex, education, employment, basic service access, etc. (UN, 2008). The census data are often the main source of statistical information for many countries and allows government, private sectors and researchers to create policies and improve governance, planning and decision-making (Kpedokpo, 1982; Stats SA, 2012a). The census data are commonly mapped as choropleth maps, as it is a geographical standard (Bajat *et.al.*, 2011; Sleeter, 2004; Slocum and Egbert, 1993). Due to changes in the census boundaries, it is difficult to compare changes over time using choropleth maps (Holt *et.al.*, 2004).

Integrating land cover and census data would be beneficial when determining spatial and temporal changes. Dasymetric mapping is the process of disaggregating spatial data to produce a finer distribution of the population (Mennis, 2003). Using the dasymetric mapping approach, census and land cover would be incorporated to accurately predict population densities (Stevens, *et.al*, 2015).

This research focuses on integrating census and remote sensed data using dasymetric mapping to provide evidence of spatial patterns and temporal changes for the Upper Tsitsa Catchment situated in the Eastern Cape, South Africa. The results of this research were in the form of a spatial statistics analysis to compare the accuracy of the dasymetric and choropleth maps and provide detailed dasymetric maps interpreting the spatial and temporal change within the catchment. The results provided additional information for studies that are taking place in the catchment, such as studies of erosion, sustainability and restoration by other DEA-NRM (South African Department of Environmental Affairs – National Resource Management) funded projects for the Ntabelanga Laleni Ecological Infrastructure Project (NLEIP). As the DEA-NRM funded this research, a requirement was to create a digital database that could be used in other research and for decision making within the Upper Tsitsa Catchment.

NLEIP will be utilising the data for decision making and locating where persons are, what level of education they have, common language spoken and employment status within the Upper Tsitsa Catchment. The 2001 and 2011 South African National Population Censuses was used in conjunction with the National Land Cover data sets for 2000 and 2013/14 as the time scale over which the comparisons were made.

Many difficulties were faced when combining these data sets, as they were both captured differently and are structured for different purposes. The census boundaries were inconsistent over time as enumeration boundaries changed and could not be directly compared with one another. The National Land Cover data classes became more refined over time and were not directly comparable with each other, as there were more classes than before. By integrating census and remote sensed data, using dasymetric mapping techniques, a valuable source of information was created to provide more detailed data that could be used for many GIS applications as well as unify the data (Yuan *et.al.,* 1997). It was the intention of this research to determine the extent to which dasymetric mapping was reliable and useful, so that evidence of spatial patterns and temporal change can be identified.

1.2 Research Question

Does dasymetric mapping that integrates census and remote sensing data provide reliable and useful evidence of spatial patterns and temporal change in the Upper Tsitsa Catchment?

1.3 Objectives

1. Locate, collect and extract the relevant data from the 2001 and 2011 censuses.

2. Identify, locate and collect relevant National Land Cover data that correspond as closely as possible to the census data dates.

3. Use dasymetric mapping to integrate the extracted census data and land use/land cover data for 2001 and 2011.

4. Generate a database from the results for the DEA-NRM.

5. Create a spatial analysis of the dasymetric and choropleth data using the 2011 population density data.

6. Analyse the integrated data to identify socioeconomic trends (spatial resolution of 30 m).

1.4 Motivation

This study was conducted in association with the Ntabelanga and Laleni Ecological Infrastructure Project (NLEIP) and the Catchment Research Group (CRG) of Rhodes University. Their aims are to improve the socio-economic condition of the community within the Upper Tsitsa Catchment using green innovations and aid in landscape rehabilitation with the help of the local communities. This catchment has been identified by the Eastern Cape Provincial Government as a priority initiative. By developing this region, social and economic upliftment will be accelerated by building the Ntabelanga dam. NLEIP is linked to the Department of Water Affairs (DWA) Mzimvubu Water Project, which involves building of two dams, namely Ntabelanga and Laleni dams that fall within the initiatives of the Eastern Cape Provincial Government. NLEIP will be making use of the database developed from the dasymetric maps to determine where restoration and other various initiatives could be implemented. Furthermore, the research seeks to investigate the reliability and usefulness of dasymetric mapping in detecting spatial and temporal changes in the Upper Tsitsa Catchment and will contribute to the existing literature of dasymetric mapping within South Africa.

1.5 Structure of the Thesis

The structure adopted for this thesis provides a comprehensive and logical format. The 'Study Area' chapter outlines the study location and contextualises the region with regards to natural, social, economic and historical background of the region, which is supported by relevant literature referring to these sections. Chapter three, focuses on the literature review, which is sub-sectioned into three parts. This chapter critically reviews and evaluates literature on census data, land cover data, dasymetric mapping techniques and its various methodological approaches and accuracy assessments. Case study examples for both South Africa and the international context, were reviewed. The 'Methodology' chapter presents the steps taken to prepare the census and land cover data to create dasymetric data and discusses in depth, the methods implemented for each of the research objectives. The 'Results' chapter presents the quantitative findings, which is then followed by the 'Discussions' chapter which focuses on a general discussion and interpretation of the results. Lastly the overall deductions of the thesis are presented in the 'Conclusions' chapter, along with its limitation and recommendations for future studies is presented.

Chapter Two: Study Area

The following chapter presents the study areas and the natural, social, economic and historical background for this thesis. It will first outline the location of the study area, followed by the natural and economic background. Economic activities for the area are briefly outlined to give a sense of the status of the catchment. Lastly a historical background of the Eastern Cape and the demarcation process of the municipal boundaries is presented to contextualise the research that follows.

2.1 Location

The study area is located in the north-eastern region of the Eastern Cape Province, South Africa and falls within two local municipalities, Elundini Local Municipality and Mhlontlo Local Municipality (Figure 1). The research took place in the Upper Tsitsa Catchment which is a tertiary catchment that is made up of the quaternary catchments T35A to E and covers approximately 2000 km^2. The catchment starts in the Drakensburg mountain range, approximately 15 km from the town of Rhodes and drains for approximately 120 km eastwards into the Mzimvubu River. The catchment boundary will be used as the study area boundary when conducting the dasymetric analysis.

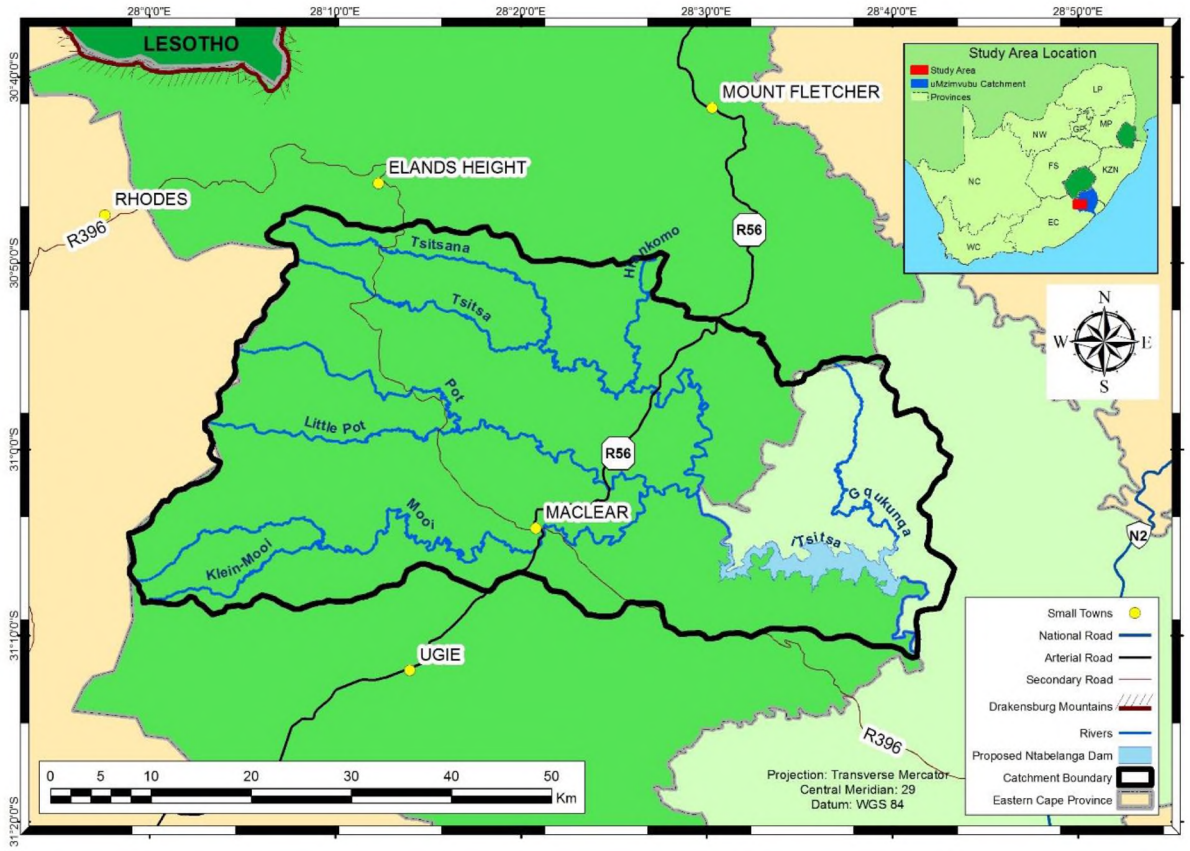


Figure 1: Study area.

2.2 Natural and Socio-Economic Characteristics

To interpret the results created in objective three, which integrates census and land cover data using dasymetric mapping, some background is presented in this section on the region with regards to its demographic, natural and economic activities.

2.2.1 Context

This study area is part of the poorest and least developed regions in South Africa (Calmeyer and Muruven, 2014). Within the study area (Figure 2), 62 percent of the land formed part of the former Cape Province and 38 percent belonged to the former Transkei. The former Cape Province, shown in brown (Figure 2), is located in what is now the Elundini Local Municipality. The 38 percent which was formerly the Transkei, shown in yellow, covers parts of the Elundini Local Municipality and the entire extent of the Mhlontlo Municipality within the eastern part of study area. Most the population found in the catchment is African and the common languages are isiXhosa and Sesotho (Stats SA, 2017a; Stats SA, 2017b)

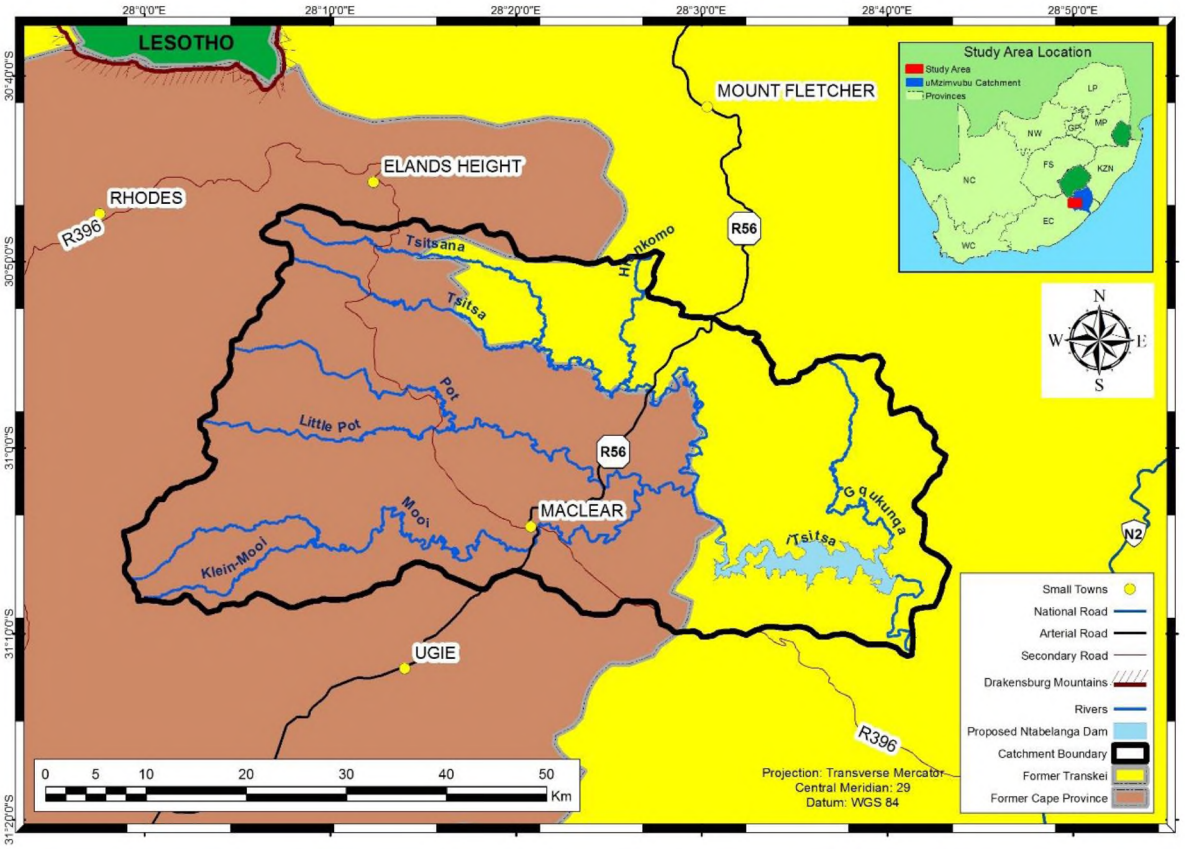


Figure 2: Former Transkei boundary that intersects with the Upper Tsitsa Catchment.

The Elundini Local Municipality has an estimated density of 27 people per square kilometre while Mhlontlo Local Municipality has a population density of 67 people per square kilometre (Stats SA, 2017a; Stats SA, 2017b)

A complete guide created by Yes Media Publishers each year called The Local Government Handbook South Africa, contains information of the municipalities in South Africa that was collected, captured and cross-referenced, where possible, via annual surveys that they conducted (Yes Media, 2015a). The Local Government Handbook (2015b) highlights Elundini Local Municipality as one of the most scenic areas within the province with potential for development due to its fertile soils and high rainfall and therefore has the possibility for economic growth and upliftment for the residents of the municipality as it has an abundance of natural resources available. Elundini tourism is fast growing with interests from both local and international tourists that seek natural attractions as well as cultural and historical heritage sites (Elundini Local Municipality, 2017a).

Mhlontlo Local Municipality, named after chief of the Pondomise people, has been classified as a B4 rural community incorporating Qumbu and Tsolo rural towns (Mhlontlo Local Municipality, 2017). A B4 rural community means there are at most two small towns in the area, communal land tenure is still in practice and the villages are typical of the former homelands (National Treasury, 2011).

2.2.2 Economic Activities

Economic activities for Elundini Local Municipality and Mhlontlo Local Municipality are important for determining development as well as the social and economic status of the catchment. Recognising the significant sectors and new developments in the catchment will be beneficial when interpreting results pertaining to economic and social upliftment.

The Elundini Local Municipality has a relative abundance of natural resources and hosts both commercial and subsistence farming activities in the former White and Black African areas respectively (Yes Media, 2015b). Examples of commercial farmlands with cultivated fields and livestock grazing areas are shown in Figures 3 and 4. Most of the working population is situated in the urban areas, such as Maclear (Figure 5) and Ugie, where they are employed in commercial, service and in the commercial agricultural sector (Elundini Local Municipality, 2017b). The rural settlements of the former Transkei homeland, however, have lower levels of employment as they have a weak economic base and many of the people who live in them still rely on subsistence activities and migrant income (Yes Media, 2015b; Elundini Local Municipality, 2017b).

Figure 3: Aerial image of commercial farming in Elundini Local Municipality (Image: CD:NGI, 2013).

Figure 4: Commercial farm land used for grazing cattle in Elundini Local Municipality (Spring, 2015).

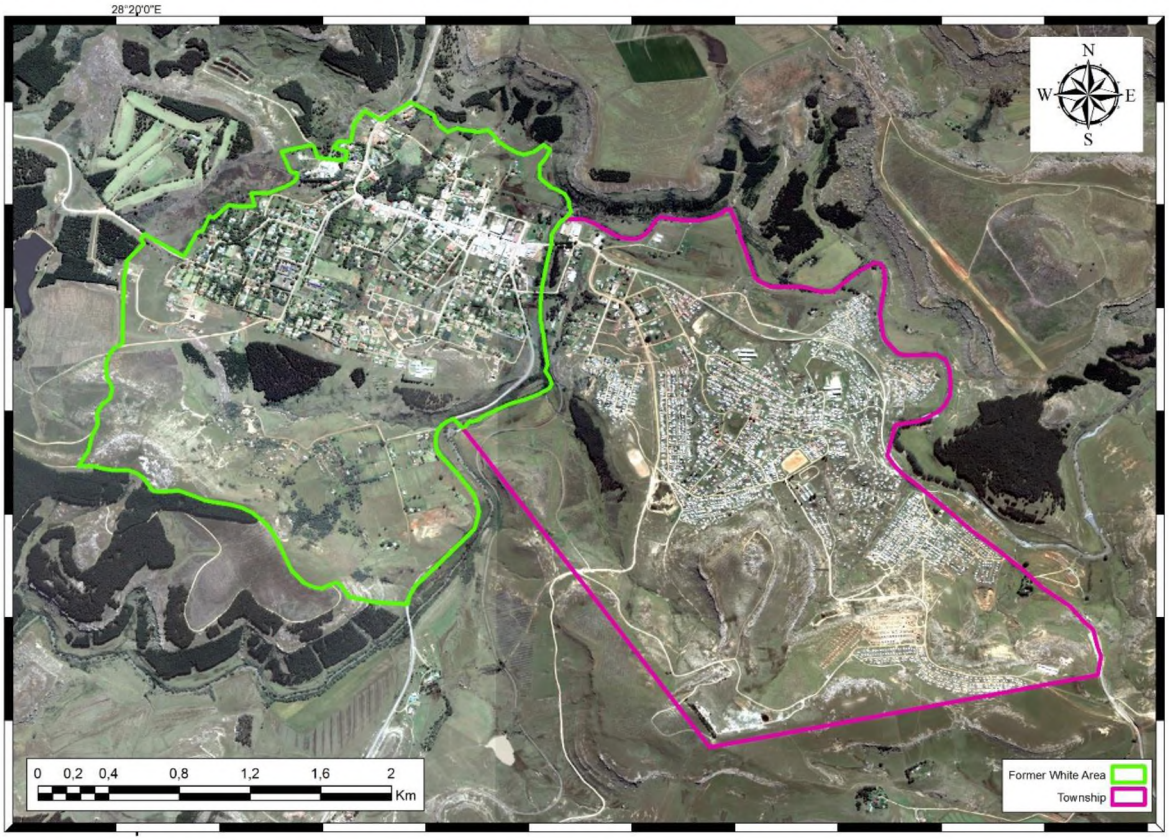


Figure 5: Aerial image of Maclear (Image: CD:NGI, 2013).

The Mhlontlo Local Municipality has very little economic development and is mostly rural with villages or scattered dwellings, with two small towns in the municipality that fall outside of the study area (National Treasury, 2011). KwaMsobomvu village (Figures 6 to 8) is an example of what rural villages look like in the Upper Tsitsa Catchment. Subsistence farming and homestead garden plots commonly occur throughout and stock farming is still of importance within these communities (Yes Media, 2015b). This municipality has two key local economic development projects, namely the Tsitsa Lodge and the Malakhiwe Bakery and has key investment opportunities in industrial development, agriculture, housing development, mining and tourism (Yes Media, 2015b). Due to the dispersed nature of the population in the Elundini and Mhlontlo municipalities, it is difficult for government to provide the required services that people require which causes an increased cost to occur when providing education, water and health care services (Ruiters, 2011).

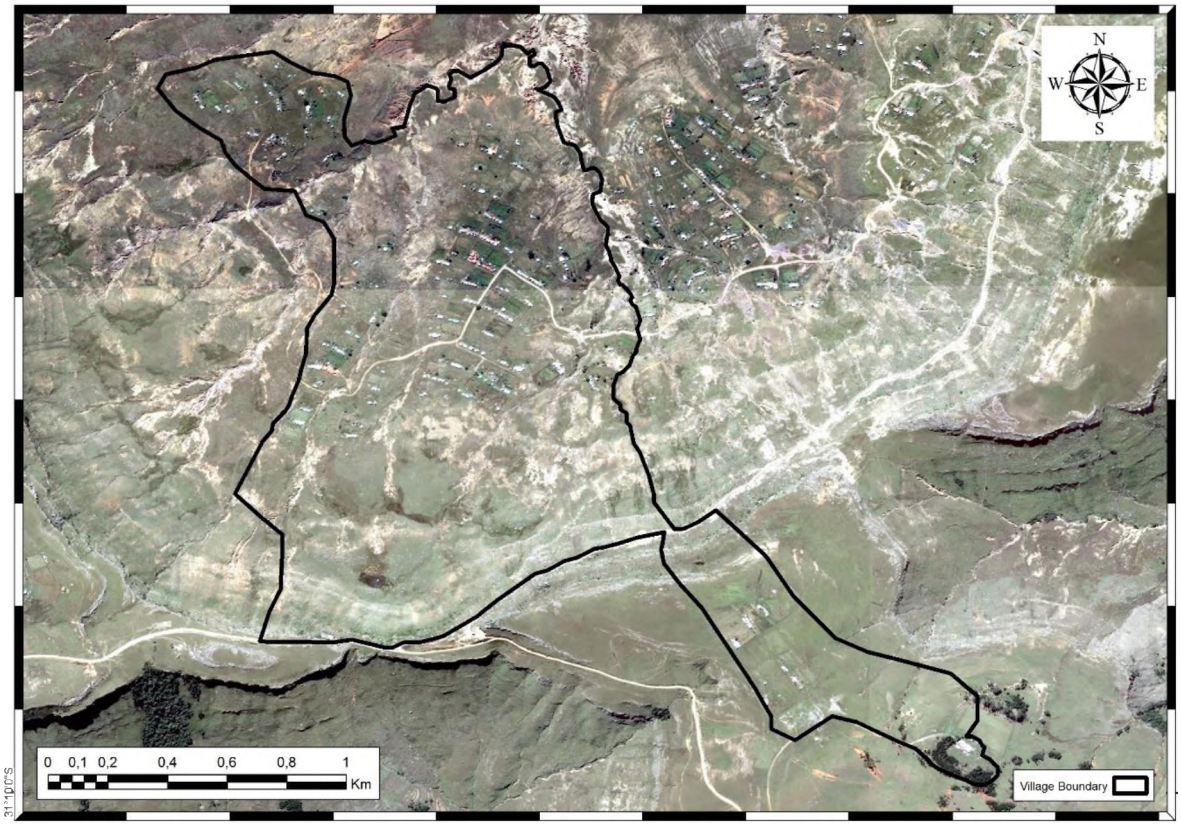


Figure 6: Aerial image of KwaMsobomvu village, Mhlontlo Local Municipality (Image: CD:NGI, 2013).

Figure 7: KwaMsobomvu village, taken from the West (Autumn, 2015).

Figure 8: KwaMsobomvu Village, taken from the East (Winter, 2015).

The proposed Ntabelanga dam site (Figure 1), is located approximately 25 km east of Maclear and is to the north of the R369 road and falls on the boundary between Elundini and Mhlontlo Local Municipality. The Ntabelanga dam is part of the Mzimvubu Water Project, which is a multipurpose project that will provide socio-economic development within the three District Municipalities; Joe Gqabi, OR Tambo and Alfred Nzo (Calmeyer and Muruven, 2014). This project is designed to accelerate the social and economic upliftment of the population in the region from the benefits of the Ntabelanga dam (Calmeyer and Muruven, 2014). The dam is to supply the region with domestic and irrigation water in conjunction with hydropower and will create temporary and permanent employment for both skilled and unskilled labour (Calmeyer and Muruven, 2014).

2.3. Historical Background of the Eastern Cape

With regards to understanding the spatial patterns and trends produced by the dasymetric mapping, one must bear in mind the past apartheid political regime of South Africa, and how it still may have an impact on the results.

The Eastern Cape Province was created as a separate province in 1993 and is composed of parts of the former Cape Province and two Bantustans, Transkei and Ciskei (Figure 9) (Ruiters, 2011). The many decades of spatial, social and economic segregation within South Africa due to Apartheid has left a legacy that has begun to disentangle by the abolishment of segregation and introduction of democracy (Nel and Binns, 2000). Under apartheid laws, the native people were forced to move to, or remain in,

the homelands, which more often than not, exceeded the carrying capacity of the lands (Southall, 1983; Nel and Binns, 2000). The homelands (Figure 9) were peripheral areas within the South African economy and were highly dependent on apartheid South Africa for employment to support their growing population (Lemon, 1976).

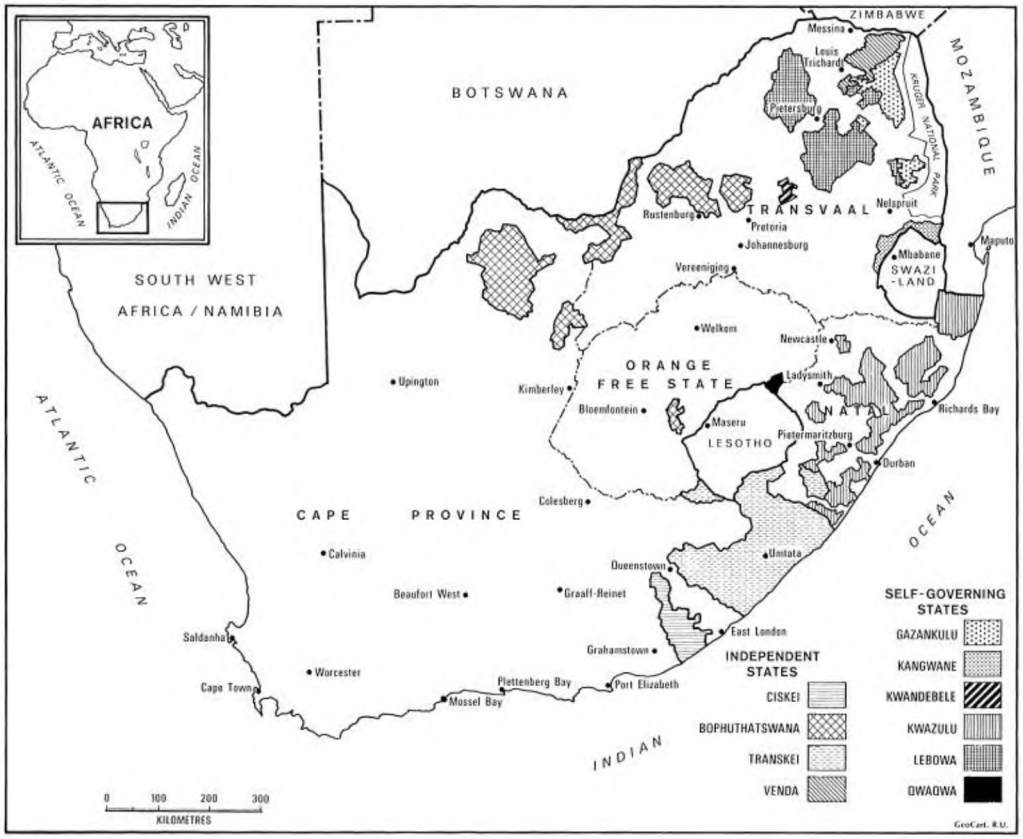


Figure 9: South Africa showing the location of the former homelands (Nel and Binns, 2000).

The homelands depended on the South African treasury for financial support and had limited natural resources with large growing populations which suffered from a shortage of skilled and trained persons at all levels (Lemon, 1976; Munger, 1962). The homelands were also dependant on South Africa for transport facilities and telecommunications, as this was mainly available in the South African cities and White areas of the country only (Lemon, 1976). The homelands were an important source of labour to South Africa's economy and this created a two-way dependence between them (Lemon, 1976).

The Transkei was one of two Xhosa homelands and stretched from the Kei River to the Umtamvuna River in the Eastern Cape Province and became the first "self-governed" homeland territory in South Africa in 1976 (Nel and Binns, 2000, Hamann and Tuinder, 2012). This almost completely African territory was one of the most densely populated and closely settled rural areas in South Africa, where

most of the lands were held on a communal basis and traditional chiefs still had power over portions of the land (Lemon, 1976; Munger, 1962).

The Cape Province which was also known as the Cape of Good Hope Province was one of four strictly governed administrative units in Apartheid South Africa (Ruiters, 2011). When the Cape Colony and British joined into the Union of South Africa in 1910, the Cape Province was established (Hamann and Tuinder, 2012). This was the largest of the four provinces in the Apartheid South Africa and consisted of more than half of the country's total area (Editors of Encyclopaedia Britannica, nd.) The Ciskei and parts of the Transkei and Bophuthatswana were located in the Cape Province but acted independently (Ruiters, 2011). In 1993, the Cape province was divided into three new provinces namely, Northern Cape, Western Cape and Eastern Cape (Ruiters, 2011). The Eastern Cape Province inherited the poorest population with the worst infrastructure in the country (Van Niekerk, 2008). After the democratic elections of 1994, all the homelands were reincorporated into South Africa and the Transkei became part of the new Eastern Cape Province (Hamann and Tuinder, 2012). Due to apartheid policies, along with economic failure and corruption that the Transkei faced, it meant that little economic and infrastructure development had taken place within the homeland which has resulted in inequities in various regions of the Eastern Cape (Hamann and Tuinder, 2012). The Eastern Cape receive little support from the Western Cape with administrative infrastructure thus leading to administrative problems which led to civil strikes in 1995 and high level of corruption (Ruiters, 2011).

2.4 The Municipal Demarcation Board

Before the formation of the national Demarcation Board, municipal boundaries were created by the members of the provincial councils for the 1995-96 local elections (MDB, 2013a). The Municipal Demarcation Board (MDB) was established in 1999 as an independent body that is responsible for the determination of the municipal and ward boundaries for the entire Republic (MDB, 2013a). The MDB's mandate is derived from the Constitution of the Republic of South Africa, 1996 (Act No 108 of 1996) and is mandated in terms of section 155(3)(b) to determine the municipal boundaries independently as well as the Local Government: Municipal Demarcation Act, 1998 (Act No 27 of 1998) (MDB, 2013b).

The function of the MDB is to determine and re-determine municipal boundaries, the delimitation of wards for elections, assess the capacity of the district and local municipality in their ability to perform their functions for the Constitution and Structures Act and lastly to declare and withdraw declarations of district management area (MDB, 2013c). The Cabinet has assigned a leading role for the MDB to assist government departments with aligning their functional service deliveries boundaries with those of the municipal boundaries (MDB, 2013c).

The ward delimitation criteria differ from that of the general criteria (MDB, 2010). The criteria are exclusive to ensure the conditions for local democracy and to represent the voters on the municipal council (MDB, 2010). The ward delimitation criteria has three basic requirements: the voters are spread fairly and equitably, the voting process is not disrupted and voters can access voting stations easily and finally to ensure that residents that identify themselves as a community are able to vote together and participate in ward committees together (MDB, 2010).

2.5 Conclusion

In conclusion, the Upper Tsitsa Catchment is in one of the poorest and least developed regions of South Africa with limited infrastructure, administrative problems and corruption due to many decades of social and economic segregation. This region has the potential of development as it has an abundance of natural resources. The proposed Ntabelanga dam is one such development project that aims to boost the socio-economic development in the area. Most of the employed residents are working in the commercial agriculture sector, while many others are engaged in subsistence farming. Changes in the municipal and electoral ward boundaries have occurred over the years to evenly spread the voters, allow the voting process to not be disrupted and to ensure residents that identify as a community to vote and participate in ward committees together.

Chapter Three: Literature Review

This chapter reviews existing literature pertaining to populations censuses, land cover data, dasymetric mapping and error matrix statistics. It is important to understand how data is collected and created when working with the census data and land cover images.

The first topic discussed in section 3.1 focuses on census data regarding the South African National Population Censuses. Census data contains valuable information regarding a country's population and should be comparable over time, but rarely is. The change in boundaries, however, prohibits the comparison of these resources which could provide spatial and temporal information about the population.

The second section (3.2) introduces land cover data. Land cover is used worldwide as it has been proven to be a vital source of information with regards to the changing environment (Ngcofe and Thompson, 2015). It exhibits a wide range of applications and is used in an ever-growing community comprised from the governmental to academic fields.

The third topic in section 3.3, explores dasymetric mapping techniques. There are many difficulties when combining census and remote sensed data such as land cover. They are captured and structured differently as they serve different purposes. Using dasymetric mapping may solve problems of boundary changes in census data and provide a more useful dataset to compare changes over time.

The last section focuses on accuracy assessments using an error matrix with statistical analysis to determine the accuracy of data created using classification schemes.

3.1 Population Census and Data

Assessing the quality and quantity of human resources at a small area, regional and national level is essential for all governments (UN, 2008). A population and housing census allows statistics pertaining to age, sex, education as well as housing, access to water and availability of essential services to be collected (UN, 2008). The United Nations (2008, 7) defines a population and housing census as "the total process of collecting, compiling, evaluating, analysing and publishing demographic, economic and social data pertaining, at a specific time, to all persons in a country or a well-defined part of a country." A population census is the main source of demographic data in many countries and is often used for statistical purposes (Kpedekpo, 1982). Census data allows government, private sectors, academic and research institutions to formulate policies that will improve governance, planning, monitoring and evaluation for an array of projects and decision-making (Stats SA, 2012a). Using the

census in this research will allow for the appropriate planning, policy implementation and decision making for the Upper Tsitsa Catchment. A population census questionnaire is only able to ask a limited amount of questions and is carried out every ten years as the exercise is costly and requires a vast amount of skilled labour (Kpedekpo, 1982, UN, 2008). Regardless of the long intercensal periods, census data are an extremely rich source of information with a range of characteristics (Colman and Salt, 1992). Censuses hold statistical information regarding populations and are published in the form of well-defined administrative areas, such electoral wards, also known as small areas, in South Africa (Gregory and Ell, 2005). These data sets are able to provide key information about the past and how society has changed over time, however, changes in administrative boundaries have hampered the true potential of this data source (Gregory and Ell, 2005).

The United Nations (2008) estimated that over 190 countries conducted a population census, using mostly the traditional census method, in the early 2000s. The traditional census is comprised of a series of complex operations to collect information from individuals and households that encompass a range of topics which are collected in a certain time frame (UN, 2008). The statistics that are collected during a population census helps determine the approximate size of the population and the race composition, the level of education of the nation, the rate of migration, areas of residence, the employment rate, religious affiliations, the nation's health and welfare status and the economic growth or decline of the country (South African History Online, 2011).

A national census consists of an administrative hierarchy (UN-GGIM, 2010). This system subdivides each level within a country into smaller units to form the next lower level unit until the lowest level is reached (UN-GGIM, 2010). Each country has its own specific administrative hierarchy. Figure 10 below depicts the general subdivisions that are found in a geographic hierarchy, where the country as a whole unit is at the top of the structure and is broken down into small subunits.

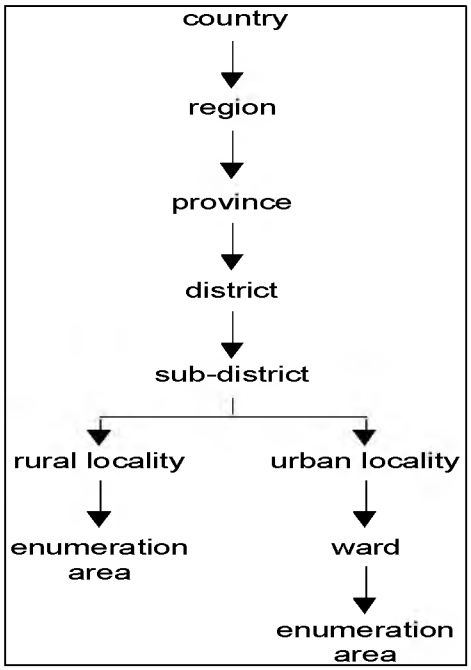


Figure 10: A generic census geographical hierarchy (UN-GGIM, 2010).

Enumeration areas are the smallest geographical units within the hierarchy, which are used when collecting census data and are defined in the census planning phase (UN-GGIM, 2010). Within the hierarchy, there are nested administration areas with defined boundaries that have administrative roles such as provinces, districts and municipalities, while other units may only have statistical purposes such as enumeration areas (Figure 11).

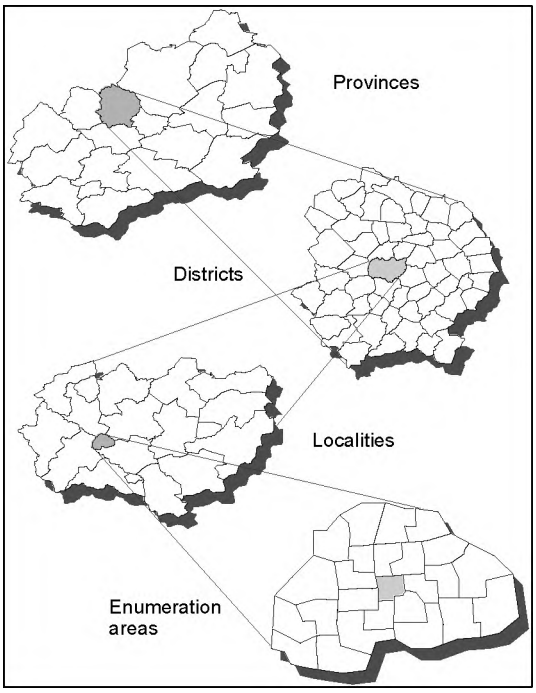


Figure 11: Nested administrative hierarchy (UN-GGIM, 2010).

A geographical framework (Figure 12) was created to describe the hierarchy for each of the geographic units that are within the South African population census data (Stats SA, 2012a). These spatial divisions facilitate the data processing, analysis and results of the population census data (Stats SA, 2012a). The hierarchy ensures that units at lower levels fit within boundaries at higher levels (Stats SA, 2012a). The hierarchy of South Africa in 2001 consists of nine provinces that are divided into six metropolitan and 46 district municipalities, but changed in 2011 to eight metropolitan areas and 44 district municipalities, with each consisting of one or more local municipalities (Stats SA, 2012c; Goodman and Medani,2016). There are two metropolitan areas and six district municipalities in the Eastern Cape, of which two district municipalities; Joe Gqabi and OR Tambo, fall within the study. At the local municipal level, there are two municipalities; Elundini and Mholontlo each from different district municipalities. The geographical framework for South Africa has remained the same for both the 2001 and 2011 censuses, however the boundaries and numbers of theses geographies have changed.

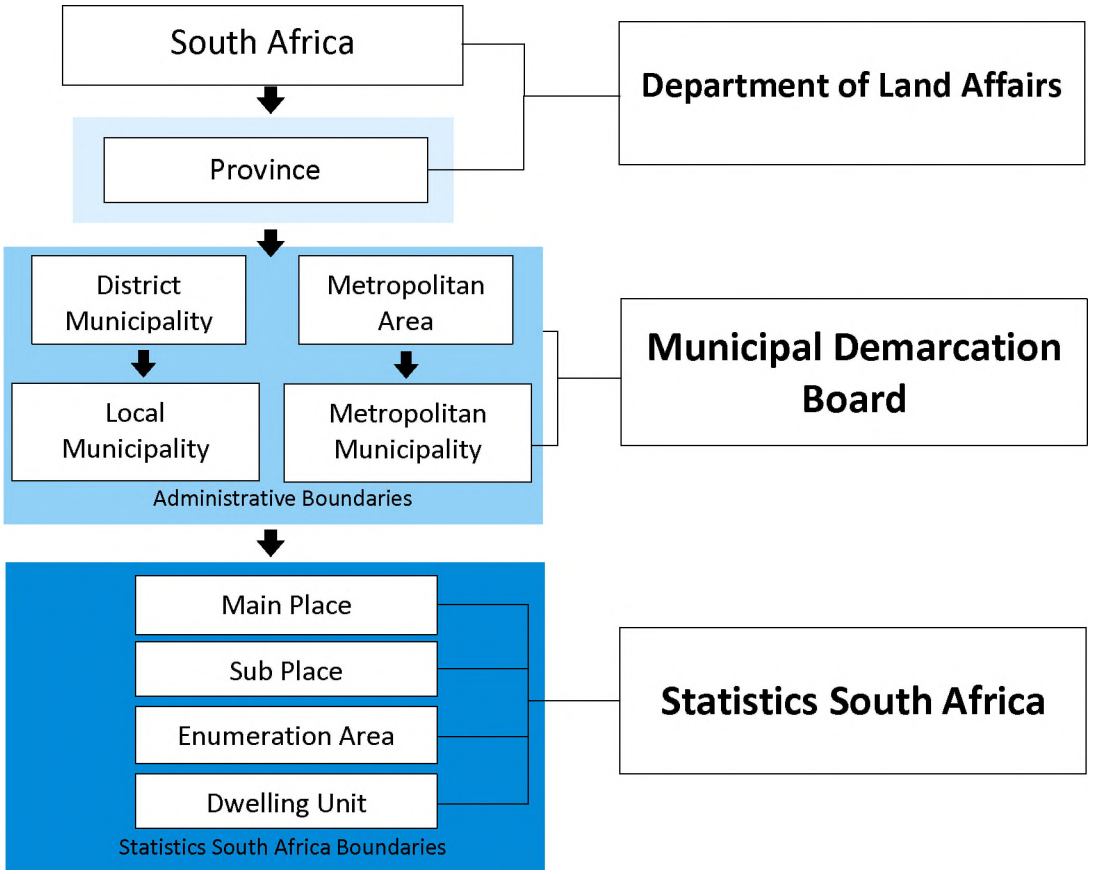


Figure 12: Geographical framework showing the hierarchy of the South African Population Census and responsible government bodies (Adapted from: Stats SA, 2012a).

Statistics South Africa created smaller work areas known as enumeration areas that are manageable for the enumerator to enumerate within a few days within the local municipality (Stats SA, 2015). At the enumeration level, the population census is carried out and the data is captured (Stats SA, 2015). Before each population census, the enumeration areas are updated and the boundaries changed due to provincial and municipal boundary demarcations after 2001 (Stats SA, 2012b; Stats SA, 2015). Each of the enumeration areas consisted of approximately 180 households and the enumerators conducted a face-to-face interview with each of the respondents available while completing the census questionnaire (Stats SA, 2012a). Households were given the chance to enumerate themselves and were supplied with guidelines on how to complete the questionnaire (Stats SA, 2012a).

The census questionnaire is based on four dimensions: accuracy, relevance, interpretability and coherence (Stats SA, 2012a). South Africa uses the *de facto* population and housing census, where all individuals are counted at the place where they spent the night of the census (Stats SA, 2012a). Three sets of census questionnaire were developed, which were administered according to an individual's location and focused on household, collective living quarters and the population in transit (Stats SA, 2012a). This study will focus on the household census questionnaire with variables such as employment status, income, education and language.

Statistics South Africa created a small areas layer (SAL) which are based on the electoral wards from enumeration areas that had less than 500 persons but also has the same sub-place name (Stats SA, 2016). There were 56 255 small areas for the 2001 population census and 84 907 small areas for the 2011 population census for South Africa. Of these small areas, 142 areas for 2001 and 161 areas for 2011 were used for the analysis of the Upper Tsitsa Catchment. Due to the demarcation changes in provincial and municipal boundaries, alignment is required to compare the 2001 and 2011 census data sets (Stats SA, 2012b). Figures 13 and 14 below depict the differences in the census boundaries for the Upper Tsitsa Catchment between 2001 and 2011. To overcome the boundary changes, dasymetric mapping was used (section 3.3).

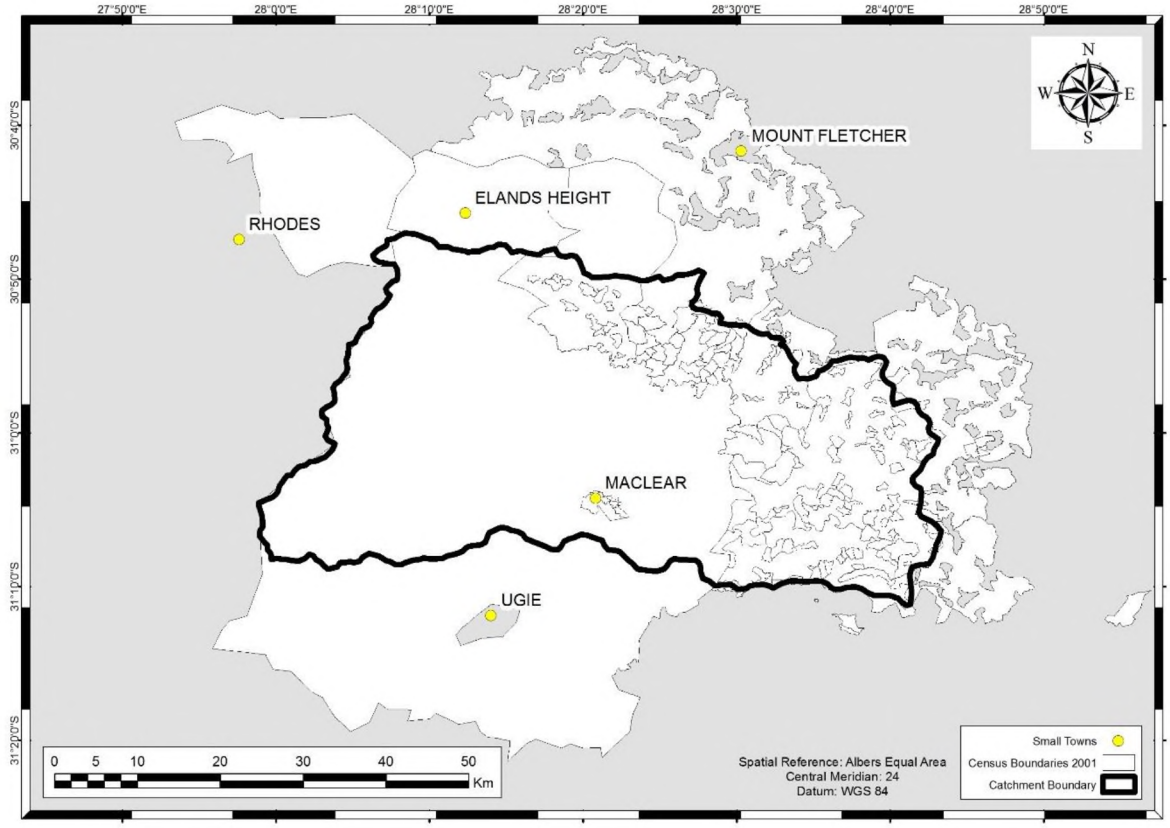


Figure 13: South Africa National Census Boundaries for the Upper Tsitsa Catchment, 2001.

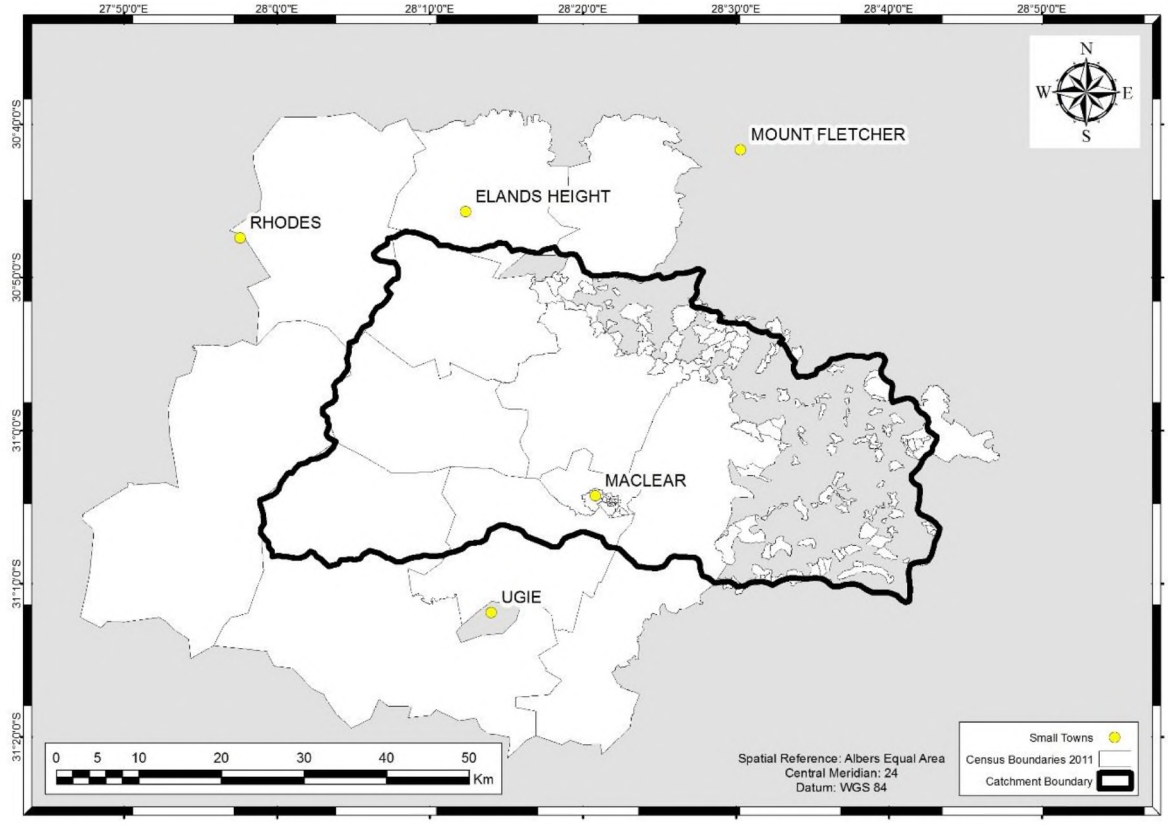


Figure 14: South Africa National Census Boundaries for the Upper Tsitsa Catchment, 2011.

Prior to the 1994 democratic elections, there was no reliable statistics for South Africa (Stats SA, 2015). Census data were collected during the apartheid era however, it is unreliable with regards to the Black African population (South African History Online, 2011). In 1996, the census included all population race groups to be counted, whereas in the past, the numerical size of each population race group was estimated (South African History Online, 2011). Since the 1996 census, the race groups have been classified as Black African, Coloured, Indian, White and Unspecified. (South African History Online, 2011). The post-apartheid government held its first 'complete' census in 1996 and its second in 2001 (South African History Online, 2011). Due to limited resources in the Department of Statistics South Africa, the interval period of census collection was changed to 10 years and therefore the third census took place in 2011 (South African History Online, 2011). For this study, 2001 and 2011 census data will be utilised for the dasymetric comparison.

Statistics South Africa mandate states that their aim is to provide statistics to the State with regards to the economic, demographic, social and environmental situation within the country (Stats SA, 2015). This mandate is in line with the Statistics Act (Act No. 6) of 1999 and follows the Fundamental Principles of Official Statistics that has been described by the United Nations (Stats SA, 2015). Collecting data is the fundamental stage during a census (Stats SA, 2012a).

Demographic data sets and socioeconomic information are commonly mapped as statistical surfaces, but are displayed using choropleth mapping techniques in the form of vector data (Bajat *et. al.*, 2011; Sleeter, 2004). Most census data are displayed using choropleth maps as it is seen as a geographic standard and is almost always used by all geographers and many other professions (Slocum and Egbert, 1993). According to Holt *et. al.* (2004), census data are often reported using the enumeration areas, which is often designed for data collection and has little or no meaning when representing the data or determining changes over time. Environmental data are mostly depicted by physiographic boundaries, while the census makes use of boundaries that do not conform to environment and constantly change (Mans, 2011). Figures 15 and 16 are examples of choropleth maps showing population density for the Upper Tsitsa Catchment for 2001 and 2011. Comparing these data would be difficult as the boundaries have changed between 2001 and 2011 and the data seems to be homogeneous for each of the small areas. Mapping population density has one of two purposes; either to cartographically depict the extent to which the population is spread across the study area or two, to portray quantitative estimates of the population's density which can be used in analytical models (Langford, 2003). To calculate the population density, the census enumeration areas need to be standardised (Holt *et. al.*, 2004). The most common way to calculate the density is by dividing the population by the given area of an enumeration area (Holt *et. al.*, 2004).

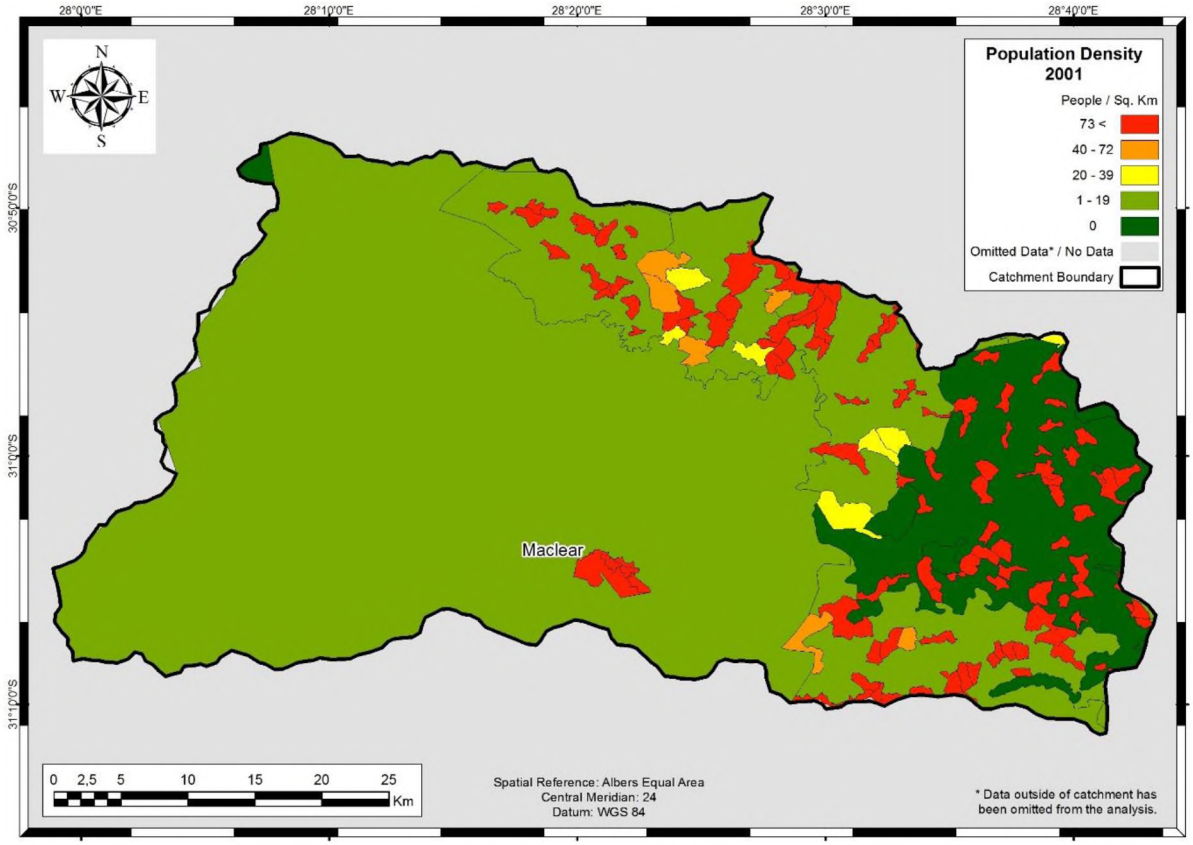


Figure 15: Population density for 2001.

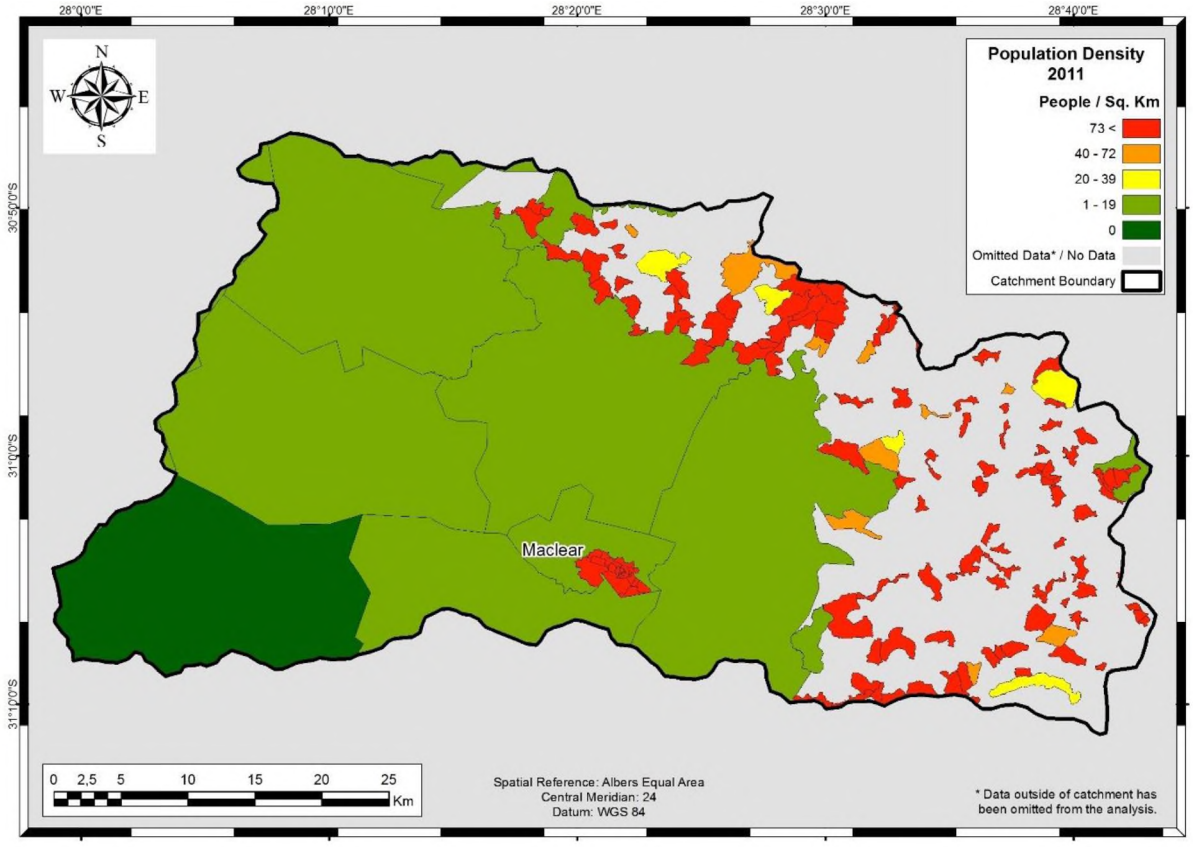


Figure 16: Population density for 2011.

There are a number of limitations when using choropleth maps to represent population density. First, these maps are exposed to the Modifiable Areal Unit Problem (MAUP), therefore potential variations within the data are not regarded, as the maps spatially aggregate data into its various geographic locations or use the areal administrative units such as small area layer (SAL) and represents a homogeneously distributed data set at a set scale. (Openshaw, 1984; Dorling, 1993). Second, the data can be interpreted to have abrupt changes at the boundaries of the enumeration area, while presenting the population density as a continuous variable across the area even in places where the data are irrelevant, giving the impression that the population is evenly distributed in the region, even if portions of the region are uninhibited (Holt *et. al.*, 2004; Sleeter, 2004). When mapping variables using choropleth techniques, the results often reveal more information about the size and the shape of the enumeration area than information about the people living in the areas. And lastly, overestimation occurs for areas that are under-populated, while areas that are over-populated are under-estimated (Holt *et. al.*, 2004). These drawbacks can be seen between 2001 and 2011 in Figures 15 and 16 above, where the boundaries have changed and the data for the enumeration areas seem to be homogenous.

Changes in the census enumeration boundaries that are nested within the hierarchy (Figure 12) makes it difficult to compare data over time, therefore the dasymetric mapping technique is one of the possible approaches that can be used to solve the problem of changed enumeration boundaries and will allow a more accurate representation of the data in a temporal comparison (Holloway *et. al.*, 1999; Bajat *et. al.*, 2011). Choropleth maps and dasymetric maps have become more differentiated, with choropleth maps being more popular of the two in cartography (Eicher and Brewer, 2003).

3.2 Land Use and Land Cover Data

Countries worldwide experience changing environments due to increases in agricultural production, population growth, climate change and higher demands on natural resources which may result in the natural environment becoming degraded (Martino and Zommers, 2007). Land cover is the physical surface of the Earth, that is made up of a number of vegetation types, soil types, exposed rock, water bodies and man-made elements such as agriculture and the built up environment (CD:NGI, 2009). Land use is the means by which people utilise the land cover present, or the activities which directly relate to the land (CD:NGI, 2009).

Land cover provides information about a location's socioeconomic trends, but does not present any characteristics of the data which is present in the census data (Martin and Bracken, 1993). By documenting land use and land cover change, significant contributions are being made at a regional and global scale (Lambin *et. al.*, 2003). Land use and land cover information is vital for an extensive

range of applications in academic, economic, government and various other fields at a variety of levels at both national and international levels (Ngcofe and Thompson, 2015). Policy users require knowledge on the status of the earth's land cover to develop sustainable development policies and strategies that range from local to global scale (Mayaux *et. al.*, 2004).

With an immense and growing user community, consistent and continuous land cover data is required for a variety of applications (Hubald and Di Gregorio., 2009). These data sets are essential as input information to conduct analysis in a range of fields such as environmental problem solving, management policies and determining spatial patterns (Ngcofe and Thompson, 2015). The integration of land cover with census data would improve the accuracy of data that is used to produce existing population mapping products (Tatem *et. al.*, 2007). This study will utilise land cover data to accurately determine where people are in the Upper Tsitsa Catchment.

There are many global and regional data sets utilising developed classification schemes, so that large scale land cover monitoring can be standardised (Hubald and Di Gregorio, 2009). The UN land cover classification scheme (LCCS) is considered the most comprehensive, flexible and internationally accepted approach for land cover classification and was developed by the FAO (Food and Agricultural Organisation) and UNEP (United Nations Environment Programme) (Hubald and Di Gregorio, 2009). This scheme provides a common language and is based on an *a priori* system with classes that are defined before data collection occurs, preventing inconsistencies and providing standardisation (Hubald and Di Gregorio, 2009). This system is hierarchically arranged by land cover type and uses dichotomous and modular-hierarchal phases (Figure 17) to distinguish eight major land types and provide certainty with higher levels of detail (Hubald and Di Gregorio, 2009). The dichotomous phase is split into vegetated and non-vegetated which is split further into terrestrial and aquatic or regularly flooded. These classes are then broken down into the eight major land types (Figure 17). The modular-hierarchical phase breaks the eight major classes into more detailed sub-classes to provide detailed descriptions for the user.

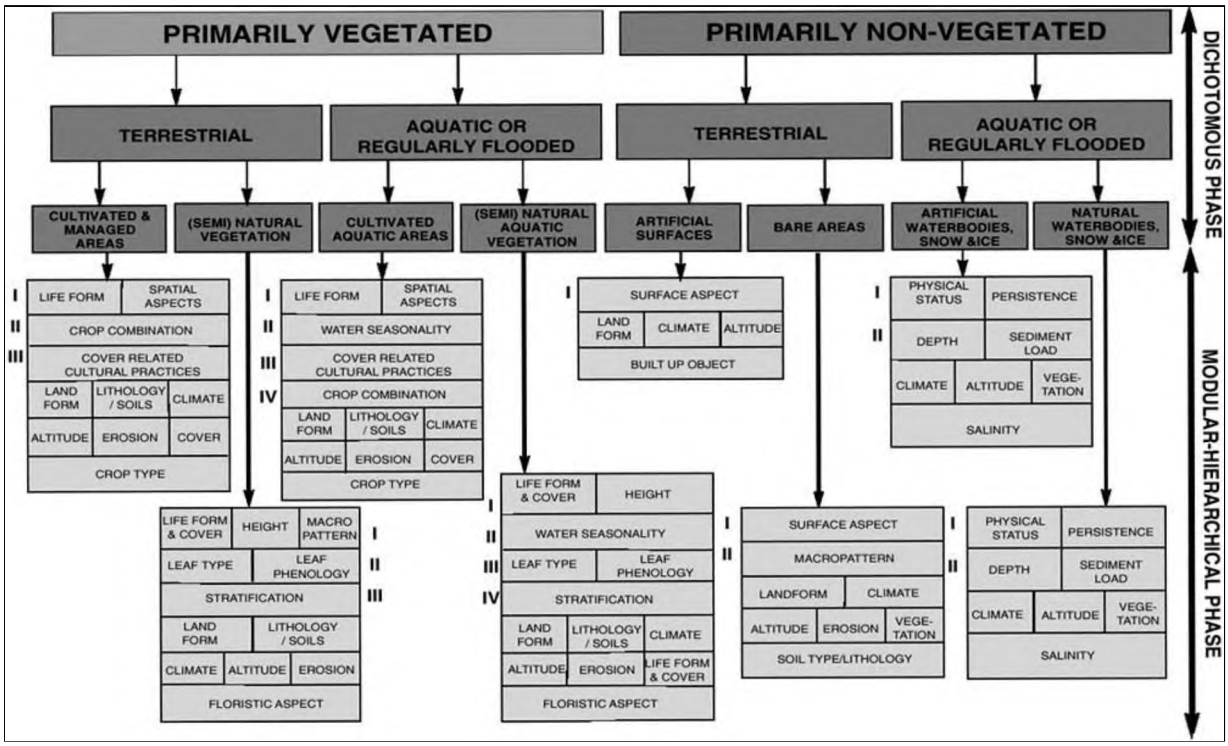


Figure 17: An overview of the land cover classification system depicting the dichotomous and modular-hierarchal phases (from Di Gregorio and Jansen, 2000).

In the past, South Africa underwent considerable changes with regards to policy, legislation and service delivery (Schoeman *et. al.*, 2013). The use of a national land cover (NLC) map allows quantitative assessment of the impacts of these changes to the environment and planning for future developments (Schoeman *et. al.*, 2013). South African land cover mapping using LandSat imagery can be traced back to 1994, where governmental organisations, along with the private sector at both provincial and national scale, have collaborated to create national level land cover data (Ngcofe and Thompson, 2015). The NLC 2000 and NLC 2013 data sets are reviewed below, as these data sets were used when conducting the dasymetric mapping for this research.

Before the development of the NLC database, there were no standards set for land cover classification in South Africa (Fairbanks *et. al.*, 2000). Through participating in mapping with the Africover project initiated by the UN and FAO a standardised hierarchical framework was created that suited the South African environment (Thompson, 1996; Ngcofe and Thompson, 2015). The land cover classification scheme (LCCS) was based on the LCCS created by the FAO and UNEP as described above and is an *a priori* ranking scheme, which is divides broad generic classes into more specific subclasses (Ngcofe and Thompson, 2015). This structure has three hierarchical levels (Table 1), level one consists of 12 categories and level 23 categories, whilst level three is user defined subcategories (Ngcofe and Thompson, 2015). The first two levels in the LCCS were adopted in 2003 as South Africa's official landcover standard (Ngcofe and Thompson, 2015). The classes in level one are similar to the eight

major classes from the FAO and UNEP land cover classification scheme. This level is split into types of vegetation, waterbodies and non-vegetated types. Level two describes how the 12 classes have been broken down to provide more detailed land cover data, which coincides with the modular-hierarchical phase in Figure 17. Level three, which is the user defined sub-categories, (Figures 18 and 19) shows that NLC2000 has 49 classes and NLC 2013/14 has 72 classes (Table 2).

Table 1: Standard land cover classification for land cover (Thompson, 1996).

Level I	Level II
1. Forest and woodland	Forest Woodland Wooded Grassland
2. Thicket, bushland, scrub forest and high fynbos	Thicket Scrub forest Bushland Bush clumps High heathland (high fynbos)
3. Shrubland and low fynbos	Shrubland Low fynbos (heathland)
4. Herbland	
5. Grassland	Unimproved grassland Improved grassland
6. Forest Plantation	Pine species Eucalypt species Wattle / other species Indigenous species
7. Waterbodies	
8. Wetlands	
9. Barren lands	Bare rock / soil Degraded land
10. Cultivated lands	Permanent crops Temporary crops
11. Urban / built up land	Residential Commercial Industrial / transport
12. Mines and quarries	

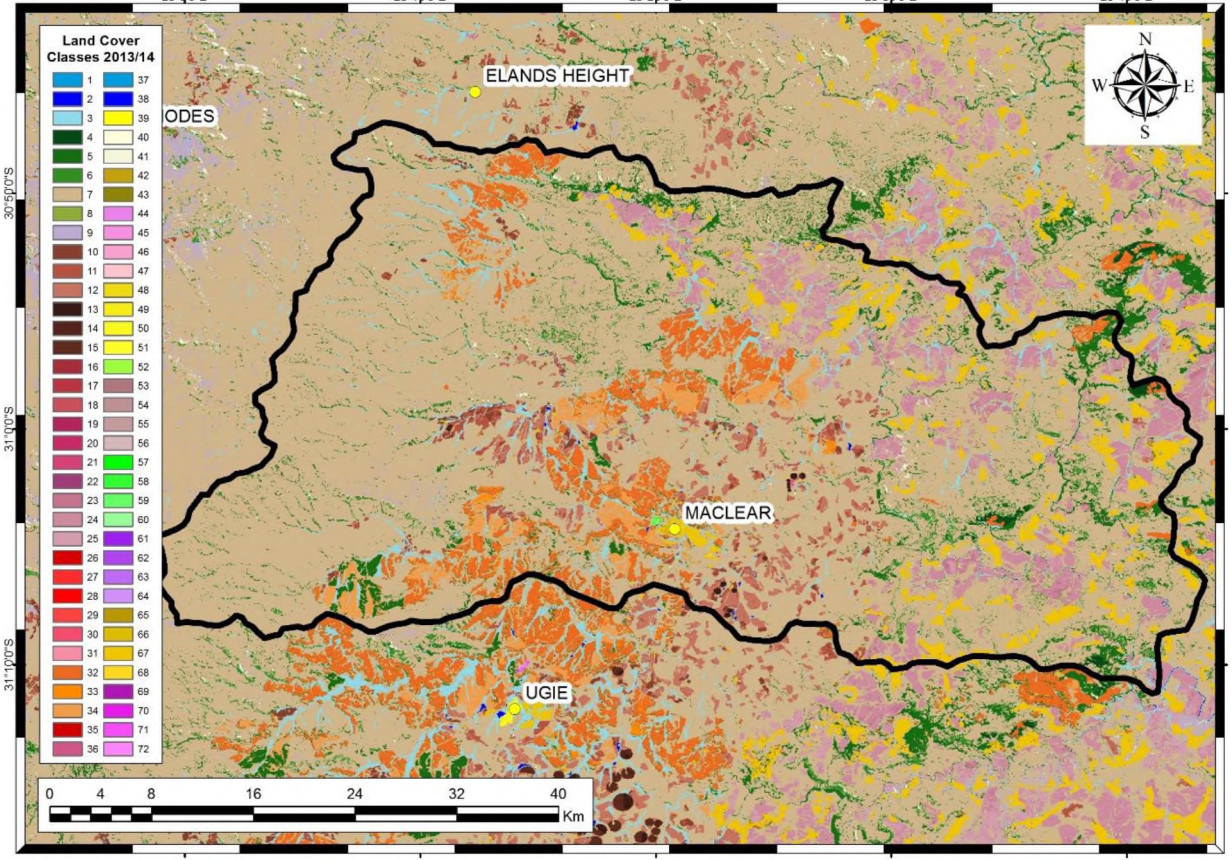


Figure 19: Land cover classes for NLC 2013/14.

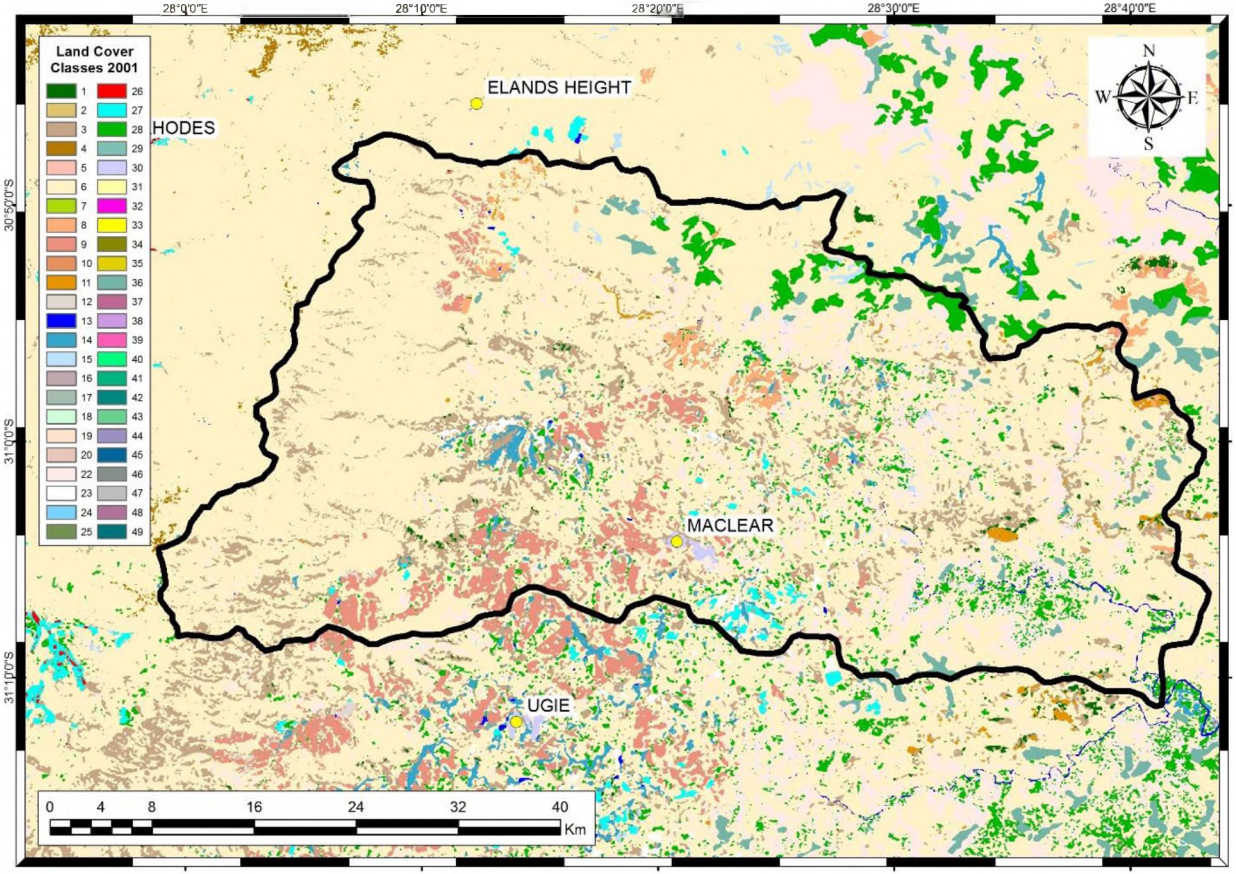


Figure 18: Land cover classes for NLC 2000.

Table 2: Description of NLC 2000 and NLC 2013/14 land cover classes.

NLC 2000 ID	NLC 2013/14 ID
1 Forest (indigenous) 2 Woodland (previously termed Forest and Woodland) 3 Thicket, Bushveld, Bush Clumps, High Fynbos 4 Shrubland and Low Fynbos 5 Herbland 6 Unimproved (natural) Grassland 7 Improved Grassland 8 Forest Plantations (Eucalyptus spp) 9 Forest Plantations (Pine spp) 10 Forest Plantations (Acacia spp) 11 Forest Plantations (Other / mixed spp) 12 Forest Plantations (clear felled) 13 Waterbodies 14 Wetlands 15 Bare Rock and Soil (natural) 16 Bare Rock and Soil (erosion: dongas / gullies) 17 Bare Rock and Soil (erosion: sheet) 18 Degraded Forest and Woodland 19 Degraded Thicket, Bushland, etc. 20 Degraded Shrubland and Low Fynbos 21 Degraded Herbland 22 Degraded Unimproved (natural) Grassland 23 Cultivated, permanent, commercial, irrigated 24 Cultivated, permanent, commercial, dryland 25 Cultivated, permanent, commercial, sugarcane 26 Cultivated, temporary, commercial, irrigated 27 Cultivated, temporary, commercial, dryland 28 Cultivated, temporary, subsistence, dryland 29 Cultivated, temporary, subsistence, irrigated 30 Urban / Built-up (residential) 31 Urban / Built-up (rural cluster) 32 Urban / Built-up (residential, formal suburbs) 33 Urban / Built-up (residential, flatland) 34 Urban / Built-up (residential, mixed) 35 Urban / Built-up (residential, hostels) 36 Urban / Built-up (residential, formal township) 37 Urban / Built-up (residential, informal township) 38 Urban / Built-up (residential, informal squatter camp) 39 Urban / Built-up (smallholdings, woodland) 40 Urban / Built-up (smallholdings, thicket, bushland) 41 Urban / Built-up (smallholdings, shrubland) 42 Urban / Built-up (smallholdings, grassland) 43 Urban / Built-up (commercial, mercantile) 44 Urban / Built-up (commercial, education, health, IT) 45 Urban / Built-up (industrial / transport: heavy) 46 Urban / Built-up (industrial / transport: light) 47 Urban / Built-up (underground / subsurface mining)	1 Water seasonal 2 Water permanent 3 Wetlands 4 Indigenous Forest 5 Thicket / Dense bush 6 Woodland / Open bush 7 Grassland 8 Shrubland fynbos 9 Low shrubland 10 Cultivated comm fields (high) 11 Cultivated comm fields (med) 12 Cultivated comm fields (low) 13 Cultivated comm pivots (high) 14 Cultivated comm pivots (med) 15 Cultivated comm pivots (low) 16 Cultivated orchards (high) 17 Cultivated orchards (med) 18 Cultivated orchards (low) 19 Cultivated vines (high) 20 Cultivated vines (med) 21 Cultivated vines (low) 22 Cultivated permanent pineapple 23 Cultivated subsistence (high) 24 Cultivated subsistence (med) 25 Cultivated subsistence (low) 26 Cultivated cane pivot - crop 27 Cultivated cane pivot - fallow 28 Cultivated cane commercial - crop 29 Cultivated cane commercial - fallow 30 Cultivated cane emerging - crop 31 Cultivated cane emerging - fallow 32 Plantations / Woodlots mature 33 Plantation / Woodlots young 34 Plantation / Woodlots clear felled 35 Mines 1 bare 36 Mines 2 semi-bare 37 Mines water seasonal 38 Mines water permanent 39 Mine buildings 40 Erosion (dongas) 41 Bare non-vegetated 42 Urban commercial 43 Urban industrial 44 Urban informal (dense tree / bush) 45 Urban informal (open tree / bush) 46 Urban informal (low veg / grass) 47 Urban informal (bare) 48 Urban residential (dense tree / bush) 49 Urban residential (open tree / bush) 50 Urban residential (low veg / grass) 51 Urban residential (bare) 52 Urban school and sports grounds 53 Urban smallholding (dense tree / bush)

NLC 2000 ID	NLC 2013/14 ID
48 Urban / Built-up (surface-based mining) 49 Urban / Built-up (mine tailings, waste dumps)	54 Urban smallholding (open tree / bush) 55 Urban smallholding (low veg / grass) 56 Urban smallholding (bare) 57 Urban sports and golf (dense tree / bush) 58 Urban sports and golf (open tree / bush) 59 Urban sports and golf (low veg / grass) 60 Urban sports and golf (bare) 61 Urban township (dense tree / bush) 62 Urban township (open tree / bush) 63 Urban township (low veg / grass) 64 Urban township (bare) 65 Urban village (dense tree / bush) 66 Urban village (open tree / bush) 67 Urban village (low veg / grass) 68 Urban village (bare) 69 Urban built-up (dense tree / bush) 70 Urban built-up (open tree / bush) 71 Urban built-up (low veg / grass) 72 Urban built-up (bare)

Land use and land cover classifications used in the past in South Africa were created for a specific objective with regards to the study and was largely influenced by geographical location of the research area and the available technology, therefore these land cover data sets were incomparable with each other with regards to the classes used, scale and geographical coverage (Fairbanks *et. al.*, 2000). The National Land Cover (NLC) project created a standardised digital land cover database that covers South Africa, Lesotho and Swaziland and contains base-line data on landcover for South Africa (Fairbanks *et. al.*, 2000). The broad classification classes that are applicable to southern Africa were used for NLC classification, which can be further adapted by the user to suit their requirements (Fairbanks *et. al.*, 2000). For this study, the classes for NLC 2000 and NLC 2013/14 were simplified into five major categories; residential, commercial/industrial land, agriculture, natural and water.

The two land cover data types that are of interest for this thesis are the NLC 2000 and NLC 2013 data sets. The NLC 2000 data was the second national land cover map created for South Africa using the standardised LCCS describe by Thompson (1996). It was produced using the LandSat 5 imagery for the years 2000 and 2001 and was later published in 2005 (Ngcofe and Thompson, 2015). This data set was the first to be created from digital imagery and the data set was captured as a digital raster with a minimum mapping unit of two hectares (Schoeman *et. al.*, 2013). It contained 49 land cover classes, which is the third user-specified level and the final map accuracy for the NLC 2000 data was 65.8% (Schoeman *et. al.*, 2013).

Ngcofe and Thompson (2015), recognised that frequent and up to date land cover data is important to conduct a variety of analyses as the physical environment is consistently changing, therefore GeoTerraImage Pty Ltd., which is a private company, produced the National Land Cover 2013/14 (NLC 2013/14) for the years 2013-2014. This data set utilised the LandSat 8 imagery which consists of 72 land cover user defined classes as discussed above, covering both natural and man-made features and is the most up to date land cover map for South Africa to date (Ngcofe and Thompson, 2015). This data set was created using a semi-automated procedure which is based on a series of repeatable and standardised modelling routines, that was specifically created for this data set (GeoTerraImage, 2015). The scale for this data set is based on 30 x 30 m raster cells and is suitable for mapping at a 1: 75 000 to 1: 250 000 scale and has an accuracy of 82 percent (GeoTerraImage, 2015).

The following case studies describe how land cover data sets were created at an international and a local region. The first paper is an example of how the FAO LLCS was used to create a uniform land cover data set for Africa as existing data ranged over various years and could not be compiled into a single data set, therefore illustrating the importance of using a standardised LCCS. The second study is an example of how the Thompson (1996) classification (Table 1) was used to create land cover data for a study site in Limpopo predating existing land cover data for South Africa.

In the paper titled "A new land-cover map of Africa for the year 2000," by Mayaux *et. al.*, (2004) a land cover map was produced for Africa with a spatial resolution of one kilometer. These data sets were created from four sensors that were on-board satellites observing the Earth's surface. According to the authors, numerous areas have been mapped using data collected from the LandSat satellites, but the land cover information derived from the data came from different years. The categories that were to be used for the land cover classification data needed to meet the international requirements as set by the UN and FAO. The LCCS was based on the LCCS created by the FAO which allowed for analysis as well as cross-referencing the land cover descriptions (Di Gergorio and Jansen, 2000). The categories within the hierarchal structure were generalised at the global scale to place Africa into a global context. The GLC 2000 map of Africa presents reliable thematic context with regards to spatial detail and permanent features, however it does not allow accurate calculations and estimation of the land cover classes (Mayaux *et.al.*, 2004).

Giannecchini *et. al.*, (2007), in their article entitled: "Land-cover change and human-environment interaction in a rural cultural landscape in South Africa," presented a study on the historical analysis of land cover change for three villages in the former homelands of the Limpopo province, South Africa to assess land cover change with remote sensing and socioeconomic data. For this study, the authors had to create a temporal analyse of the land cover. Black and white aerial photographs were used to

determine the land cover for the villages for the period of 1974 to 1997. A land cover classification was developed using Thompson (1996). The classification scheme was modified to match the vegetation ecology that was used in previous studies for the region (Giannecchini *et. al.*, 2007). Land cover images where then created from the aerial photographs using the created classification scheme to analyse the change in vegetation for the villages of interest.

The articles above are examples of how land cover data sets are utilised and demonstrate the usefulness of using a standardised land cover classification schemes. By using the standardised method of land cover classification, comparison between studies and future studies can be made to determine spatial and temporal changes. Utilising the land cover data created by Thompson (1996) and GeoTerraImage (2015), was useful in this research as it will help locate where there are urban and rural areas and economic activities such as agriculture and forestry that will be used as ancillary data for dasymetric mapping.

3.3 Dasymetric Mapping

Geographic Information Systems has renewed the interest in dasymetric mapping, however there is a lack in standardisation in the methods used, which prevents most researchers from using this method (Eicher and Brewer, 2003). There are various approaches to dasymetric mapping, with the most accepted method by Jeremy Mennis (2003) (Mann and Chandra, 2013). Sections 3.1 and 3.2 have shown that the integration of census and land cover would be beneficial when determining spatial and temporal demographic changes. Intercensal boundary changes cause spatial discontinuities and compromise the reliability of comparison of the time series data as shown in Figures 13 and 14 depicting changes in population density between 2001 and 2011 (Blake *et. al.*, 2000). Using the dasymetric approach, census information is incorporated with ancillary data to accurately predict population densities (Stevens *et. al.*, 2015). Dasymetric mapping may be able to provide a solution to comparing population census with different enumeration boundaries for the same region, as it generates a surface-based data set where the data is modelled as a continuous field that is independent of partitioning into areal units (Mennis, 2003). It is possible to create this accurate and detailed data of the population for low income nations using cheap and freely available data (Tatem *et. al.*, 2007). The dasymetric method accounts for the spatial distribution of a population within the administrative or census boundary, yielding a more accurate population density estimate than choropleth maps (Holt *et. al.*, 2004). Spatial and temporal distribution of the population can be represented in a meaningful way, which is essential for the health, economic, political and environmental disciplines (Stevens *et. al.*, 2015). These spatial and temporal changes for the Upper Tsitsa Catchment will be identified using dasymetric mapping.

Mennis (2003), defines dasymetric mapping as a process of disaggregation of the spatial data to produce a finer distribution of the population. The dasymetric process involves transforming data from the statistical zones of the aggregated data set to depict the underlying statistical surface (Mans, 2011). A dasymetric map is therefore a result of intersecting data layers such as the census data and NLC data using various rules to predict where the actual concentration of the variable, such as the populations, would be (Eicher and Brewer, 2003)

Dasymetric maps have a number of differences to choropleth maps (Eicher and Brewer, 2003). The first difference is the zonal boundaries that are used in dasymetric mapping. They are based on sharp changes in the statistical surface created from the census data being mapped and combined with the land cover units to describe the amount of habitation within a unit. (Eicher and Brewer, 2003; Sleeter, 2004). The second is identifying individual dasymetric zones which are originally homogeneous, but are not defined by the data in a choropleth map, which are illustrated in the appraisals below. Third, choropleth maps have become standardised while the various methods of dasymetric mapping have not yet been thoroughly researched, therefore many of the techniques used for dasymetric mapping are not clearly represented in the literature (Eicher and Brewer, 2003).

Dasymetric mapping is an uncommon mapping technique and is most popular for mapping population density, although it has become more popular in the last two decades (Eicher and Brewer, 2003; Mennis 2009). Demographic data such as census data are associated with analytical errors due to the use of areal unit partitioning (Mann and Chandra, 2013). These problems occur predominantly because of Modifiable Areal Unit Problem (MAUP) which is the modifying of boundaries or the scale at which the data are collected and in turn affects the results of the spatial data analysis as previously described (Openshaw, 1983; Mann and Chandra, 2013). The advantage to using this method is that data can be aggregated to any desired areal unit (Bracken, 1993). The surface representation uses cells (raster) as a unit of display that is uniform in size across the area, creating a more accurate cartographic representation of the variable being mapped (Sleeter, 2004). In this thesis, the output dasymetric data is in the form of a raster image with a cell size of 30 m x 30 m and the colour of the cell represents the number of persons present with regards to population density, education and services they have access to.

There are three core methods for dasymetric mapping according to Eicher and Brewer (2003). First, the Grid and Polygon Binary Method. The advantage of this method is that it is simple. Land use data, such as the National Land Cover is classified into two classes of inhabited area and uninhabited area based on the land use classification scheme used and knowledge of the area being mapped. All land

uses that are not urban, agricultural or woodlands are labelled as zero to mark the polygons as uninhabited. A grid (raster) or polygon can be used for this method.

The second, is the polygon and grid three-class method. This method of dasymetric mapping uses a weighting scheme to assign population data to three different land use classes within the study area. This is a functional relationship between the area-class map categories and the statistical surface is created using a percentage basis. Mennis (2003) used three classes; urban, agricultural or woodland and forestry and water. Each of the area-class categories is assigned a percentage of population and is subjectively chosen until all the categories sum up to 100 percent (Mennis, 2009). Water was classified as no data and therefore zero percent is assigned to this class. During the dasymetric mapping process, the choropleth data is redistributed as a proportion to the dasymetric zones per the assigned percentages. Although this method is called the Three-class method, this technique can be applied to any number of categories. This study will be utilising five classes, which are based on adapted methods created by Duh (2005) described in section 4.3. One of the limitations of this method is that it does not account for the area of each land use within the study area. A grid (raster) based method was devised from the above polygon method, but has the same weaknesses.

Third, the limiting variable method. This method assigns data by using areal weighting to all the inhabited land polygons in the study area. The three inhabitable lands are each set to have an equal proportion of the population. A maximum threshold density is then set and applied to the study area. For the final step the threshold values are used to adjust the data distribution within the study area. This method produces maps with the least error out of all the methods and can be implemented in a GIS to create accurate dasymetric maps even though it is over a half-century old (Wright, 1936). Steps on how to create this data could not be found and therefore the more common three-class method was used.

The article by Bajat *et. al.*, (2011) titled "Dasymetric mapping of spatial distribution of population in Timok region," focuses on choropleth and a dasymetric mapping technique applied to the region of Timočka Krajina situated in the Province of Vojvodina, Republic of Serbia. The paper shows the differences between choropleth maps and dasymetric maps. The 2002 census of Serbia was used to create the choropleth map (Figure 20 left), where the mapping units are polygons that represent the census areas. On Figure 20 (right), a dasymetric map was created for the same region, but indicates a realistic representation of the population's distribution. The dasymetric map was created using the 2002 census data with 2000 CORINE Land Cover data as the ancillary data. Bajat *et. al.*, 2011, concluded that the red areas, such as in the North East on the choropleth map are unpopulated spaces

with sparsely distributed settlements although it shows the area has a population density of 51 to 100 people per square kilometer.

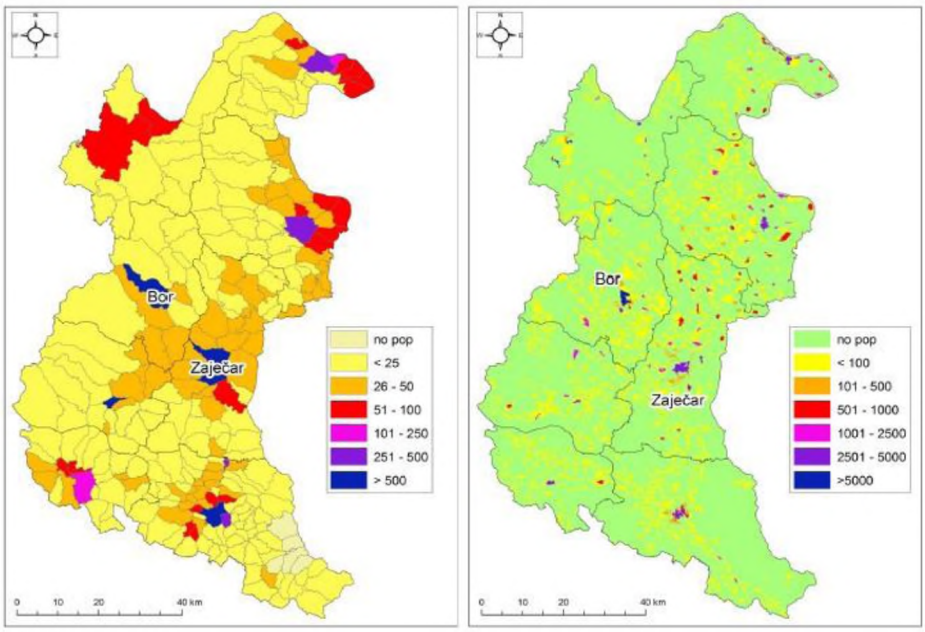


Figure 20: Population Density of Timočka Krajina; Choropleth Map (left) and Dasymetric Map (right) (Bajat et. al., 2011).

A large scale representation of the towns of Bor and Zaječar (Figure 21) are shown to gain better insight of the results obtained from the two methods used. A more detailed population density for the two municipal centres are visible in the dasymetric map (Figure 21 right), unlike the choropleth map (Figure 21, left) which shows the whole region as being densely populated.

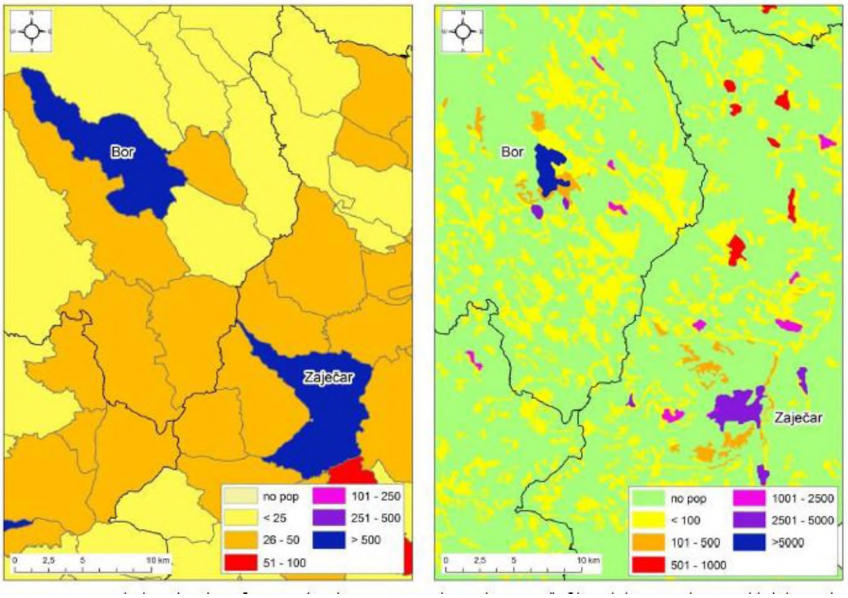


Figure 21: Population Density of Bor and Zaječar Towns; Choropleth Map (left) and Dasymetric Map (right) (Bajat et. al., 2011).

Mennis, (2003) paper on "Generating surface models of population using dasymetric mapping," used a raster population surface generation methodology. The three-class method utilised heterogeneous population distribution data and the urban land use as the ancillary data. Figure 22 (left) shows the map Mennis (2003) produced as an accurate surface model representing the population in rural settlements and in the urban areas. The urban land cover data or urbanisation data polygons were classified into three classes, high-density urban, low-density urban and non-urban area. Figure 22 compares the resulting dasymetric map (left) to the vector data in a choropleth map (right). The dasymetric map concentrated the population into sub-block regions giving reliable representation of the population distribution.

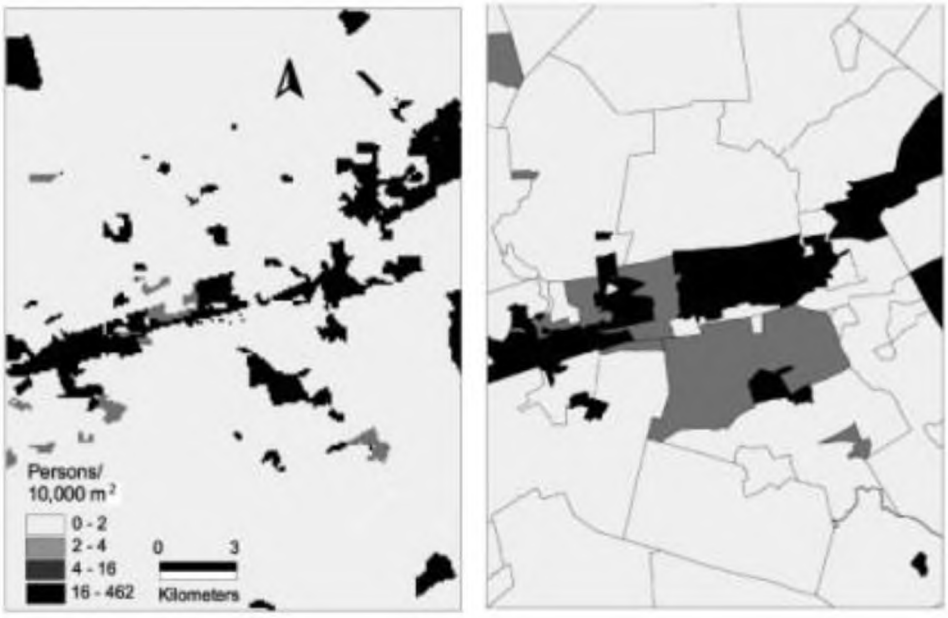


Figure 22: Detail of a Rural Area Showing the Difference Between Raster Population Surface (left) and Vector Enumeration Areas (right) Representing Population Density (Mennis, 2003).

For this study, the use of Mennis' (2003) three class method was utilised, as it is the most recognised in the available literature.

3.4 Accuracy Assessment

Accuracy is used to express the correctness of a thematic map or classification (Foody, 2002). Accuracy assessments are used to quantify data so that the user can determine the usefulness of the data for its intended purpose (Stehman and Czaplewski, 1998). Due to the complexity of digital classification, one needs to assess the reliability of the results created (Congalton, 1991). It is fundamental that the analysis technique is understood to realise the importance of the various factors that are part of an accuracy assessment (Congalton, 1991).

To do an accuracy assessment, the reference source needs to be identified and is usually in the form of pixels or polygons (Stehman and Czaplewski, 1998). In the case of this research the reference source were polygons. The classified map is compared to the reference classifications (Stehman and Czaplewski, 1998), and took the form of the dasymetric map. Point sampling units are commonly used as it is considered as to be continuous across the area (Stehman and Czaplewski, 1998). From the point sampling, an error matrix is created as a starting point to generate a series of statistical evaluations of the data (Congalton, 1991).

An error matrix (Figure 23) is a common way of representing classification accuracies of remote sensed data and has been recommended by many researchers as a reporting convention (Congalton, 1991; Stehman and Czaplwski, 1998). An error matrix is a table of rows and columns which depict the number of samples units that fall within a particular category relative to the existing category as verified in the field (Congalton, 1991). The column represents the reference data and the rows represent the classified data (Congalton, 1991). This method it effective in representing data accuracies within each category and outline both commission errors and omission errors in the classification (Congalton, 1991).

Reference Data

	D	C	BA	SB	row total
D	65	4	22	24	115
C	6	81	5	8	100
BA	0	11	85	19	115
SB	4	7	3	90	104
column total	75	103	115	141	434

Land Cover Categories

D = deciduous
C = conifer
BA = barren
SB = shrub

OVERALL ACCURACY = 321/434 = 74%

PRODUCER'S ACCURACY

D = 65/75 = 87%
C = 81/103 = 79%
BA = 85/115 = 74%
SB = 90/141 = 64%

USER'S ACCURACY

D = 65/115 = 57%
C = 81/100 = 81%
BA = 85/115 = 74%
SB = 90/104 = 87%

Figure 23: An example of an error matrix table (Congalton, 1991).

Once the table has been created and populated, an overall accuracy can be calculated by adding all the correctly classified sample point together and then dividing them by the total number of sample points (Congalton, 1991). The omission error or producer's accuracy, which is how accurately an area has been classified is calculated by taking the total number of correct sample points within a category and dividing it by the total number of sample point for that category giving the column total (Congalton, 1991). These results indicate the probability of the sample point being correctly classified (Congalton, 1991). When the correct sample points for a category are divided by the total number of

points that were classified into that category, the commission error or user's accuracy is determined, which indicated the probability of the point classified being the same as that in the field (Story and Congalton1986; Congalton, 1991).

3.5 Conclusion

The quality and quantity of population data at small area, regional and national level is an essential key for providing information regarding the past and how society has changed over time. Although the intercensal periods are long, the data sets are rich in information regarding the characteristics of a country. However, changes in administrative boundaries has hindered the potential of these data sources. Most census data are displayed using choropleth maps as it is familiar with most map users and is the easiest method to use. Changes in the census enumeration boundaries makes it difficult to compare data over time, however, integrating land cover and population census can create data sets of greater significance.

There are difficulties in combining census and remote sensed data, as the methods used to capture the data differ immensely. The case studies have shown that by integrating land cover with census data, more accurate population density can be produced. The dasymetric mapping technique enables the integration of population census and remote sensed data such as land cover or land use and combining these data sets, socioeconomic changes and land use changes can be linked together. The South African national population censuses in conjunction with the National Land Cover data NLC 2000 and NLC 2013 was used for dasymetric mapping of the Upper Tsitsa Catchment.

An error matrix is a simple yet effective way to evaluate the overall, producer and user accuracy of classified data. This method was used to describe the patterns and the correctness of the dasymetric data for the population density in the Upper Tsitsa Catchment.

Chapter Four: Methodology

The following chapter is divided into six sections and discusses the methods followed for each of the objectives that has been outlined in Chapter One.

Before the objectives could be achieved, the tertiary catchment boundary layer was extracted from a map package that contained data created by CD:NGI (National Geo-Spatial Information) and DAFF (Department of Agriculture, Forestry and Fisheries) (Wannenburgh, 2014). All data that were used and created were set to the WGS84 ellipsoid datum and projected to the Universal Transverse Mercator 35S as this was the prescribed projection settings of the landcover data sets used when creating the dasymetric maps.

4.1 Objective One

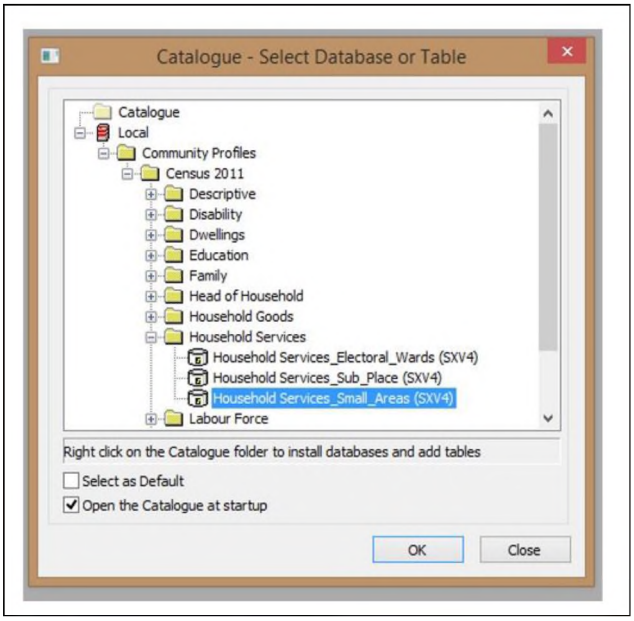
Collect and extract the relevant data from the 2001 and 2011 census.

Census data for the Upper Tsitsa Catchment needed to be extracted from the South African National Census databases for the years 2001 and 2011. These data sets were originally sourced from the South African Department of Statistics and was readily available from the Rhodes University Geography Department and Rhodes University Main Library. The 2001 census was in the form of a single shapefile at the small area layer (SAL) level (Stats SA, 2002), while the 2011 population census was available on 2011 Community Profiles Database CD-ROM as a SuperCross database (Stats SA, 2012d).

The 2001 South African census data shapefile showed the SAL (Figure 13, section 3.1.) as described in the literature, with the census results attached as an attribute table. The shapefile was loaded in ArcMap 10.3 along with the shapefile of the Upper Tsitsa Catchment boundary and all polygons of the 2001 census data that fell within the catchment or intersected the catchment boundary were selected and exported into a new ESRI Shapefile in preparation for dasymetric mapping. The extracted data for 2001 consisted of 142 small areas and contained 139 fields.

The above procedure was then applied to the 2011 national census data set to create a shapefile of the data required for 2011. The shapefile was extracted from the Community Profiles Database CD-ROM (Stats Sa, 2012d). The 2011 census small areas for the Upper Tsitsa Catchment were recorded from the attribute table of the new shapefile using ArcMap 10.3 to identify the 161 areas in the SuperCross database that comprises of 13 community profiles. SuperCross, is cross-tabulation software that allows the user to design tables, charts and view basic choropleth maps of the 2011

census data. The 2011 South African Population Census data were extracted at the SAL scale into Excel spreadsheets using the SuperCross tabulation software from the 2011 Community Profiles Database. This software was installed onto a computer in order to extract the data needed for this research and for the DEA-NRM database. The first step was to open SuperCross once installed, then from the catalogue (Figure 24) in SuperCross, each database was singularly selected, for example



Household Service_Small_Area (SXV4) (Figure 24).

Figure 24: Catalogue of databases available in SuperCross.

Within each database, geography (Figure 25, left) was selected from the list and the small areas identified from the attribute data extracted for the study area were chosen under field value (Figure 25, right).

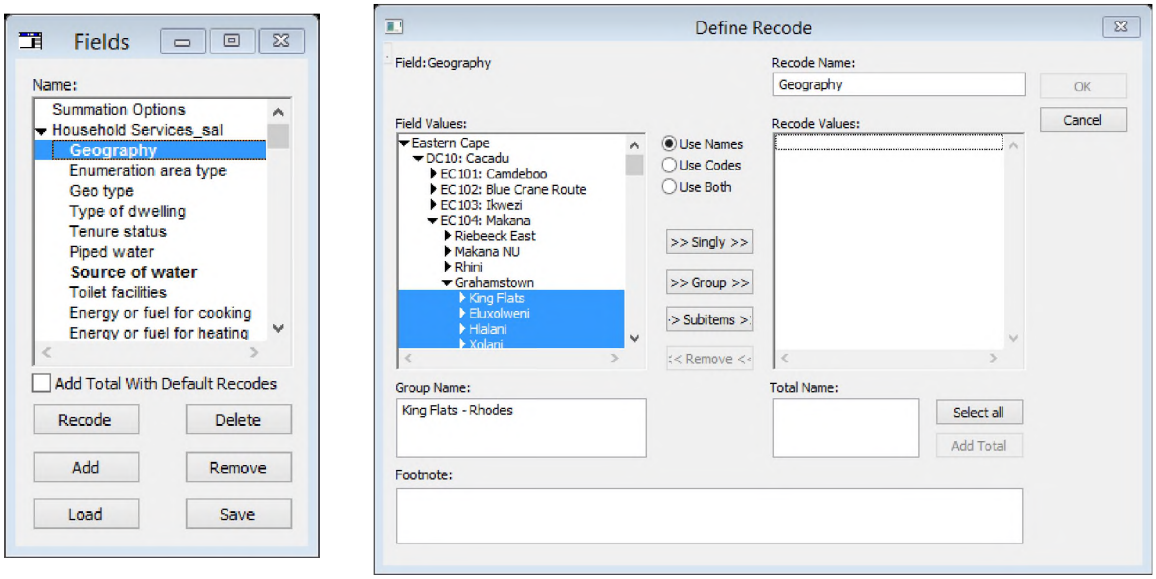


Figure 25: Field selection (left) and field values (right).

A second variable such as source of water (Figure 26) and the various sub-field values such as regional/local water schemes, borehole, etc. were then chosen to be cross-tabulated with geography to gain data for the variable of interest.

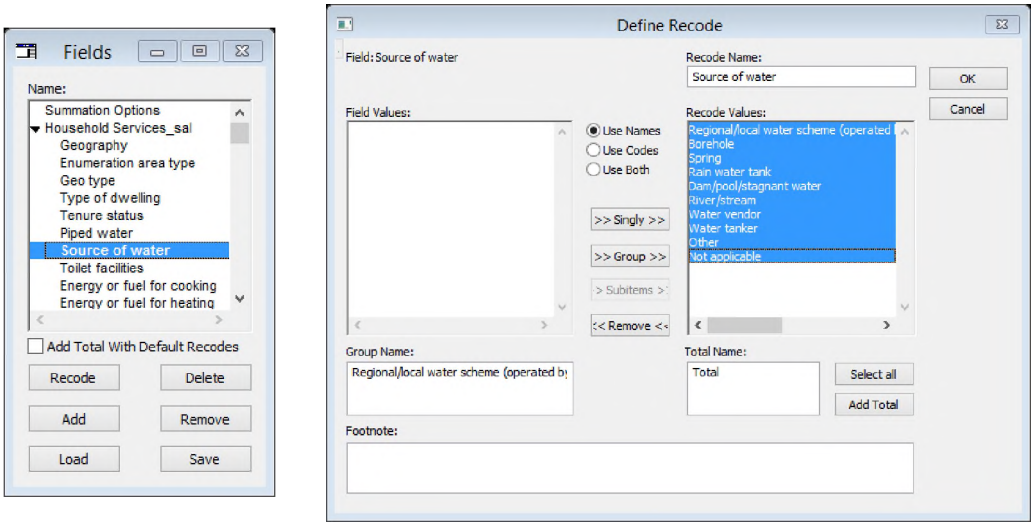


Figure 26: Field Selection (left) and Field Values (right).

Once the geography and a variable had been selected, an empty table (Figure 27) was then shown in the SuperCross window.

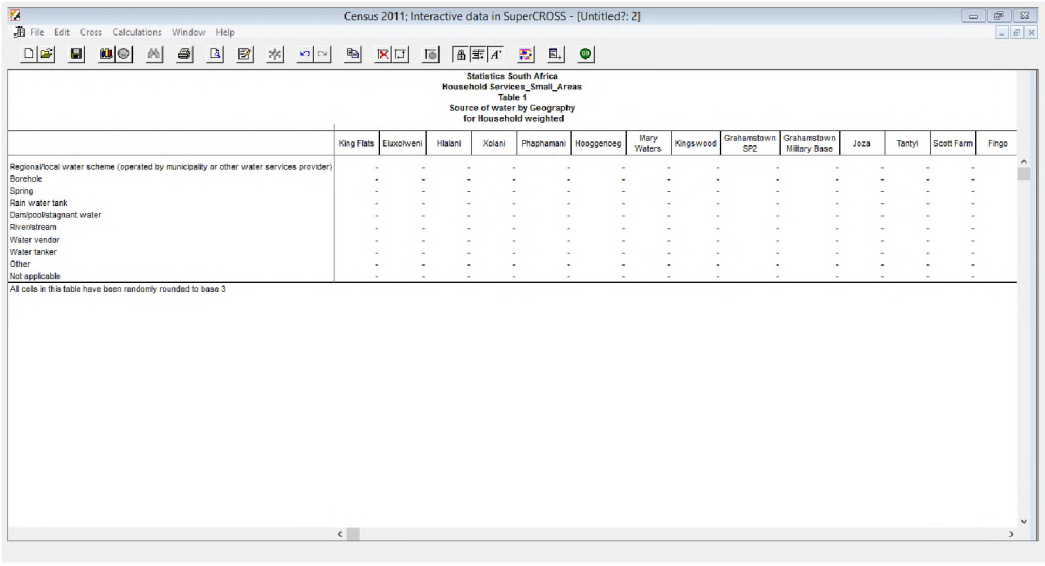


Figure 27: Table showing fields chosen from the database.

Under the cross tab, cross-tabulate was chosen, which then cross-tabulated the data and presented it in the table. This table was then exported into Excel workbooks. This process was repeated until all the fields from all 13 databases in the community profiles were cross tabulated with the geography field and extracted into ten Excel workbooks.

The data in the tabulated form within SuperCross can be viewed using ArcExplorer, which is installed with SuperCross. ArcExplorer, however, only allows the user to view the data as a basic choropleth map (Figure 28) and does not allow the user to manipluate the data.

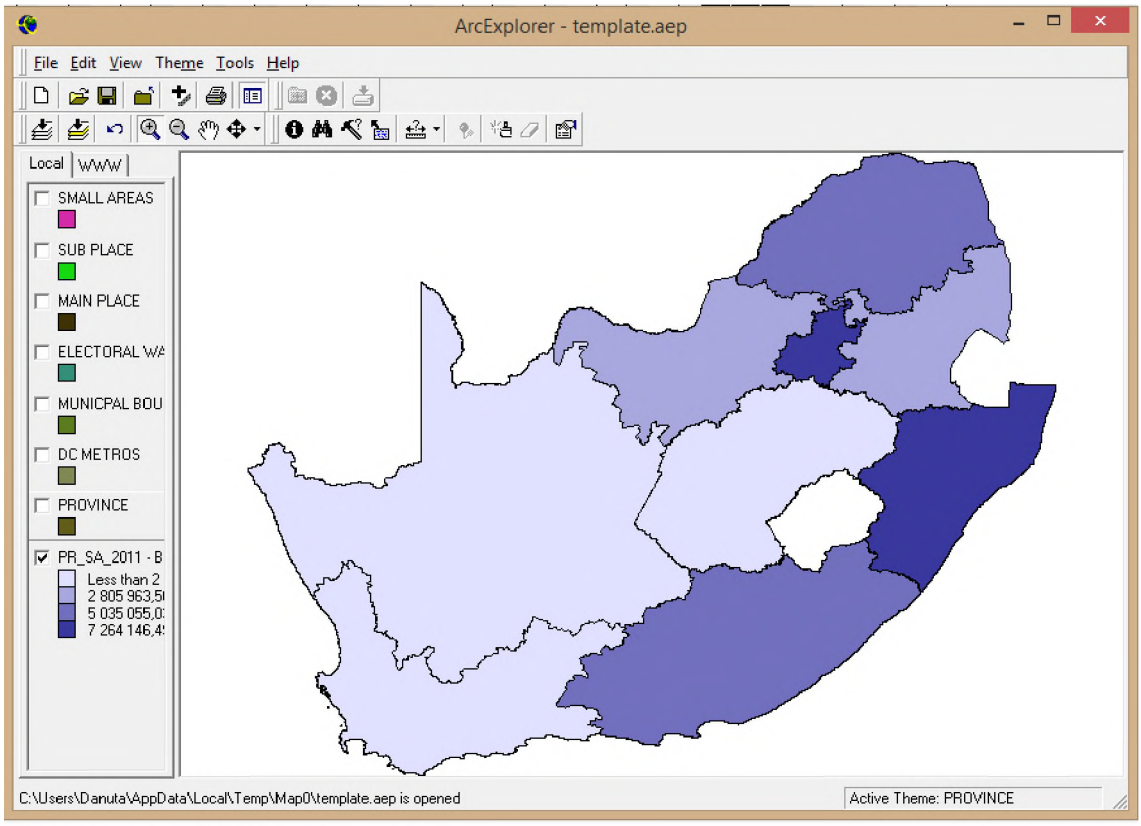


Figure 28: Example of a map using ArcExplorer with SuperCross.

Ten Excel workbooks were created from the exported data. The relevant data consists of 163 small areas, each with 858 fields (columns) of data in total. Due to the amount of data that were collected in the 2011 population census, not all the data could simultaneously be attached into one attribute table. To overcome this, only data that were being analysed (18 categories) were joined to the shapefile, the rest of the variables were placed into the DEA-NRM database, as the funders wanted all the data for the region readily available for them to utilise.

The 18 comparable categories that were to be mapped, were placed into an Excel spreadsheet, so that bar graphs could be created and be used when interpreting the trends and patterns of the maps (objective 3).

4.2 Objective Two

Collect the relevant National Land Cover data that corresponds to the census dates.

The 2000 land cover data copyrighted by CSIR (Council for Scientific and Industrial Research) and ARC (Agricultural Research Council), (2005) was sourced from the Rhodes University Geography Department. The 2013/14 National land cover (72 class) was created by Thompson, (2015) and was downloaded from the Biodiversity GIS (BGIS) website.

The 2013/14 land cover data were chosen as it was the closest available data to the 2011 population census. The land cover data sets came in the form of a raster image, which were loaded into ArcGIS 10.3 along with the Upper Tsitsa Catchment boundary shapefile. Using ArcMap 10.3, the 2000 and 2014 national land cover raster data was clipped to the catchment boundary to match the extracted census data sets in preparation for dasymetric mapping. These land cover data sets are shown in Figures 18 and 19 in section 3.2.1 above. The NLC 2000 data has 49 classes while the NLC 2013/14 data has 72 classes.

4.3 Objective Three

Use dasymetric mapping to integrate the extracted census data and land use/land cover data for 2001 and 2011 and generate a database.

For this objective, the method used was the Grid Three-Class method outlined by Mennis (2003), which has been described in section 3.3 in the literature review. The steps that were taken to produce the dasymetric maps results, closely follows the procedure that was designed by Duh (2005) that follows the methods of Mennis (2003) and Holloway *et. al.* (1999). The instructions presented by Duh (2005) on how to create dasymetric maps is an adapted method of Mennis (2003) and uses five land use classes instead of three to produce more accurate results, therefore this research used five land use classes instead of three as described in the Grid Three-class method. Duh's (2005) methodology was further adapted to match the South Africa Census data, as it was created for the United States Census. These adaptions were minor, for example in the United States uses "blockgroups," while South Africa uses small place codes in the census.

Eighteen categories were mapped using this method, creating 36 maps in total. The categories mapped included population density, male and female density, population race groups, languages, levels of schooling, employment status and access to water and electricity.

Dasymetric mapping technique chosen uses the following basic equation below which was described by Holloway *et. al.* (1999) and has been simplified by Duh (2005):

$$P = R_A \times N \times \text{Pixel height} \times \text{Pixel length} / (E * \text{Total})$$

Where:

- P represents the population of the cell (pixel),
- R_A is the relative density of a specific land cover within a cell,
- N represents the actual population in the SAL,
- E is the expected population per SAL polygon, which is calculated using relative land cover density (R_A) and
- Total is the sum of all the relative densities,

to calculate population density.

This equation was used for the raster mapping approach. Due to the land cover being in the form of a raster, the raster mapping approach was chosen to provide accurate results, as converting polygon datasets were easier to convert with less error being introduced. This method requires eleven steps to create a dasymetric map. Steps one through eleven were followed for 2001 and were repeated in a new ArcMap document for 2011. The following flowchart (Figure 29) presents the process of how the dasymetric maps were created.

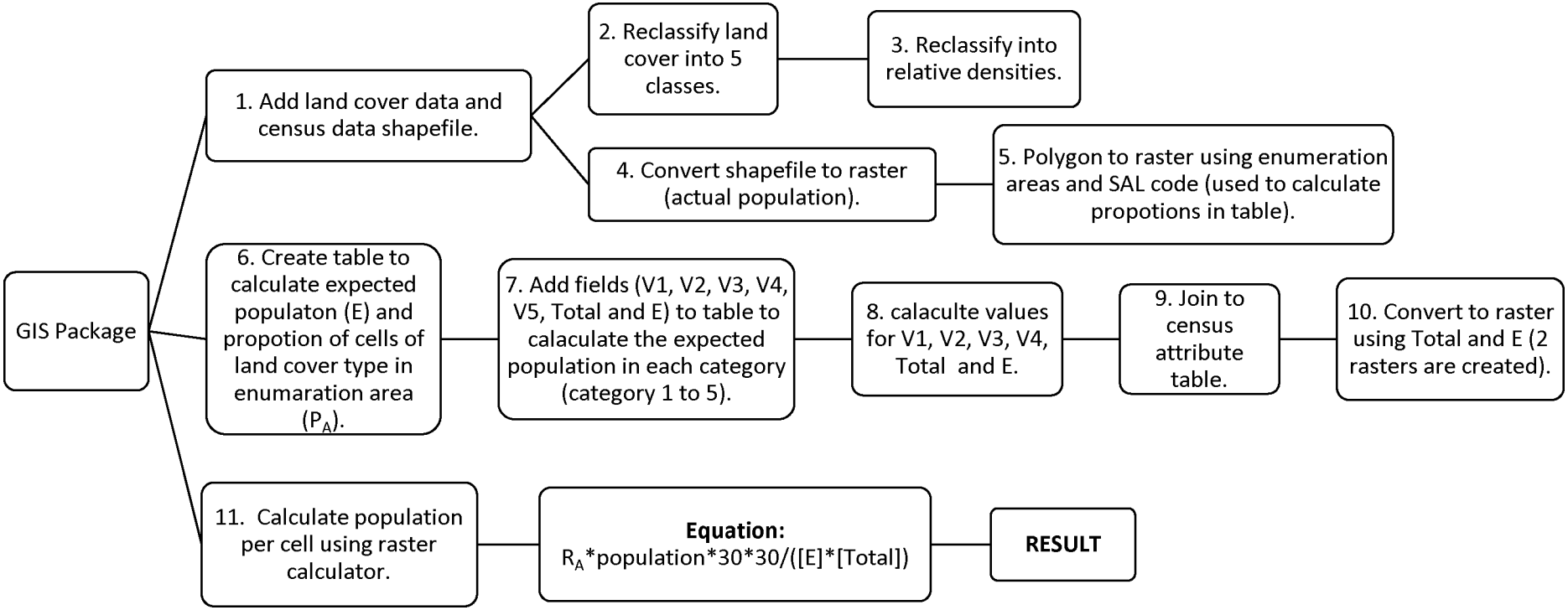


Figure 29: Flowchart describing process of dasymetric mapping.

Step one. Land cover datasets and the census data shapefiles were loaded into a new map document. The environment was set to allow all processes to run under the same conditions.

Step two required that the land cover data be reclassified into five classes, per Duh's (2005) manual. Table 3 below shows how NLC 2000 and NLC 2013/14 classes were reclassified into the five classes to be comparable and to simplify the data to calculate R_A. Figures 30 and 31 below show the reclassified land cover for the NLC 2000 and NLC 2013/14 land cover. data

Table 3: Reclassification groups for land cover data.

Reclassified land cover code	Description (Class)	NLC 2000 ID	NLC 2013/14 ID
1	Residential / Houses	30 Urban / Built-up (residential) 31 Urban / Built-up (rural cluster) 32 Urban / Built-up (residential, formal suburbs) 33 Urban / Built-up (residential, flatland) 34 Urban / Built-up (residential, mixed) 35 Urban / Built-up (residential, hostels) 36 Urban / Built-up (residential, formal township) 37 Urban / Built-up (residential, informal township) 38 Urban / Built-up (residential, informal squatter camp) 39 Urban / Built-up (smallholdings, woodland) 40 Urban / Built-up (smallholdings, thicket, bushland) 41 Urban / Built-up (smallholdings, shrubland) 42 Urban / Built-up (smallholdings, grassland)	44 Urban informal (dense tree / bush) 45 Urban informal (open tree / bush) 46 Urban informal (low veg / grass) 47 Urban informal (bare) 48 Urban residential (dense tree / bush) 49 Urban residential (open tree / bush) 50 Urban residential (low veg / grass) 51 Urban residential (bare) 53 Urban smallholding (dense tree / bush) 54 Urban smallholding (open tree / bush) 55 Urban smallholding (low veg / grass) 56 Urban smallholding (bare) 61 Urban township (dense tree / bush) 62 Urban township (open tree / bush) 63 Urban township (low veg / grass) 64 Urban township (bare) 65 Urban village (dense tree / bush) 66 Urban village (open tree / bush) 67 Urban village (low veg / grass) 68 Urban village (bare) 69 Urban built-up (dense tree / bush) 70 Urban built-up (open tree / bush) 71 Urban built-up (low veg / grass) 72 Urban built-up (bare)

Reclassified land cover code	Description (Class)	NLC 2000 ID	NLC 2013/14 ID
2	Commercial / Industrial Education / Sports Fields	43 Urban / Built-up (commercial, mercantile) 44 Urban / Built-up (commercial, education, health, IT) 45 Urban / Built-up (industrial / transport: heavy) 46 Urban / Built-up (industrial / transport: light) 47 Urban / Built-up (underground / subsurface mining) 48 Urban / Built-up (surface-based mining) 49 Urban / Built-up (mine tailings, waste dumps)	35 Mines 1 bare 36 Mines 2 semi-bare 37 Mines water seasonal 38 Mines water permanent 39 Mine buildings 42 Urban commercial 43 Urban industrial 52 Urban school and sports grounds 57 Urban sports and golf (dense tree / bush) 58 Urban sports and golf (open tree / bush) 59 Urban sports and golf (low veg / grass) 60 Urban sports and golf (bare)
3	Agriculture	8 Forest Plantations (Eucalyptus spp) 9 Forest Plantations (Pine spp) 10 Forest Plantations (Acacia spp) 11 Forest Plantations (Other / mixed spp) 12 Forest Plantations (clear felled) 23 Cultivated, permanent, commercial, irrigated 24 Cultivated, permanent, commercial, dryland 25 Cultivated, permanent, commercial, sugarcane 26 Cultivated, temporary, commercial, irrigated 27 Cultivated, temporary, commercial, dryland 28 Cultivated, temporary, subsistence, dryland 29 Cultivated, temporary, subsistence, irrigated	10 Cultivated comm fields (high) 11 Cultivated comm fields (med) 12 Cultivated comm fields (low) 13 Cultivated comm pivots (high) 14 Cultivated comm pivots (med) 15 Cultivated comm pivots (low) 16 Cultivated orchards (high) 17 Cultivated orchards (med) 18 Cultivated orchards (low) 19 Cultivated vines (high) 20 Cultivated vines (med) 21 Cultivated vines (low) 22 Cultivated permanent pineapple 23 Cultivated subsistence (high) 24 Cultivated subsistence (med) 25 Cultivated subsistence (low) 26 Cultivated cane pivot - crop 27 Cultivated cane pivot - fallow 28 Cultivated cane commercial - crop 29 Cultivated cane commercial - fallow 30 Cultivated cane emerging - crop 31 Cultivated cane emerging - fallow 32 Plantations / Woodlots mature 33 Plantation / Woodlots young 34 Plantation / Woodlots clear felled

Reclassified land cover code	Description (Class)	NLC 2000 ID	NLC 2013/14 ID
4	Natural	1 Forest (indigenous) 2 Woodland (previously termed Forest and Woodland) 3 Thicket, Bushveld, Bush Clumps, High Fynbos 4 Shrubland and Low Fynbos 5 Herbland 6 Unimproved (natural) Grassland 7 Improved Grassland 15 Bare Rock and Soil (natural) 16 Bare Rock and Soil (erosion: dongas / gullies) 17 Bare Rock and Soil (erosion: sheet) 18 Degraded Forest and Woodland 19 Degraded Thicket, Bushland, etc. 20 Degraded Shrubland and Low Fynbos 21 Degraded Herbland 22 Degraded Unimproved (natural) Grassland	4 Indigenous Forest 5 Thicket / Dense bush 6 Woodland / Open bush 7 Grassland 8 Shrubland fynbos 9 Low shrubland 40 Erosion (dongas) 41 Bare non-vegetated
5	Water / Wetland	13 Waterbodies 14 Wetlands	1 Water seasonal 2 Water permanent 3 Wetlands

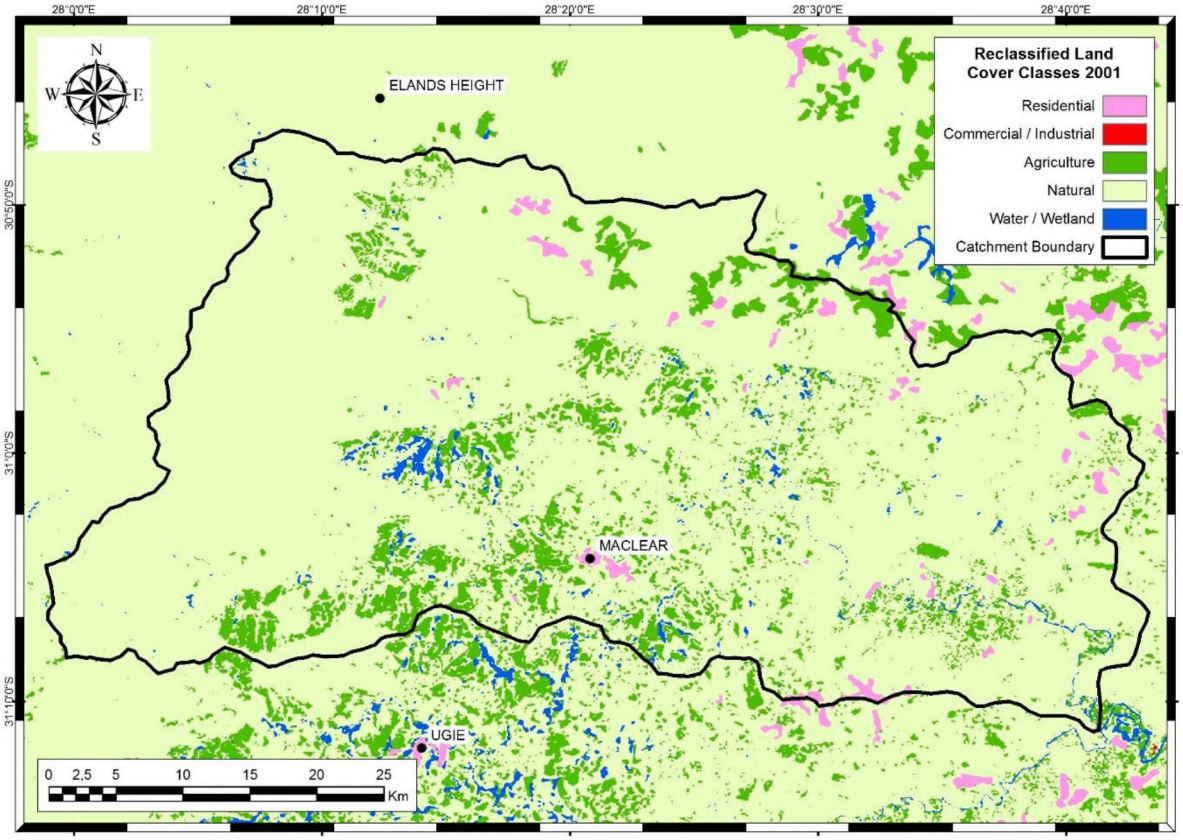


Figure 30: Reclassified NLC 2000.

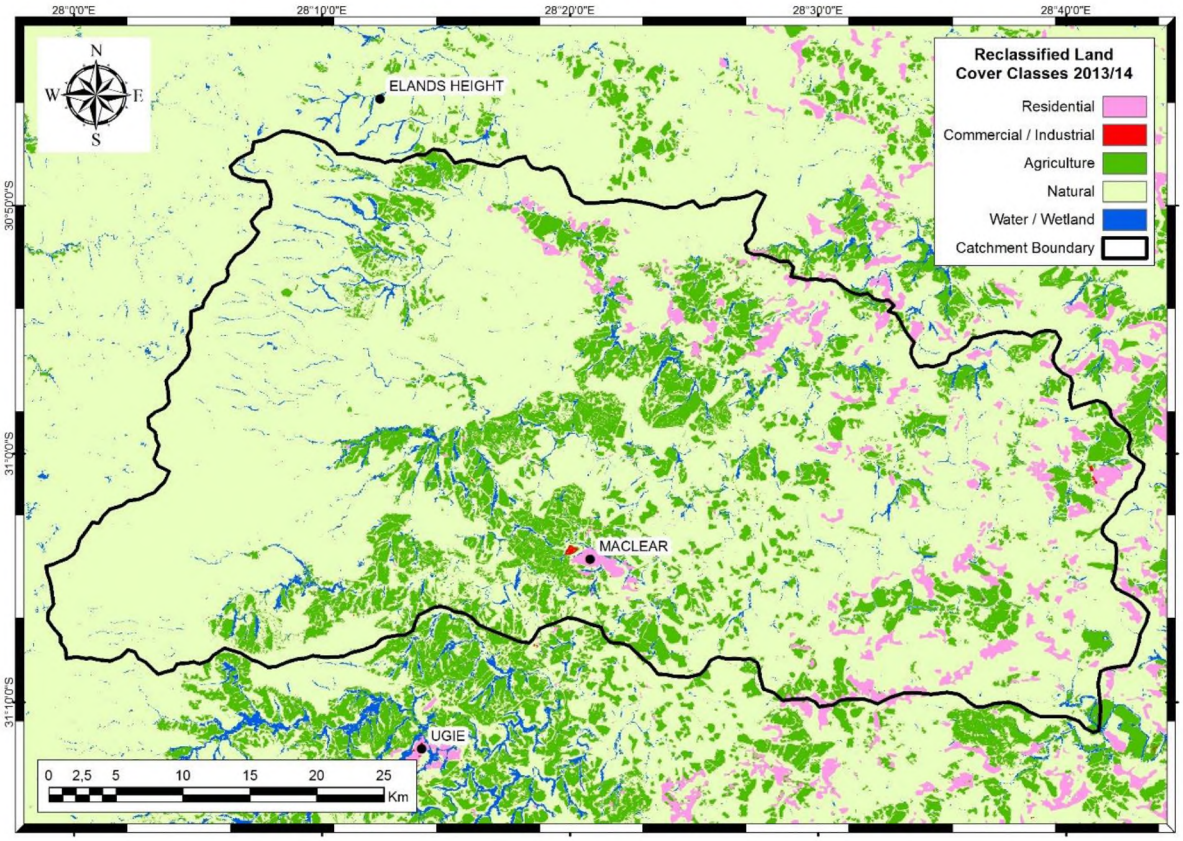


Figure 31: Reclassified NLC 2013/14.

Step three required creating the R_A (relative density) raster layer, by reclassifying the land cover data that were created in step two. Each of the five land cover classes were assigned a relative density that were based on assumption from field observations in the Upper Tsitsa Catchment. This is an accepted method as per the literature, as each geographic area is unique and comparisons cannot easily be made (Duh, 2005; Mennis, 1993).

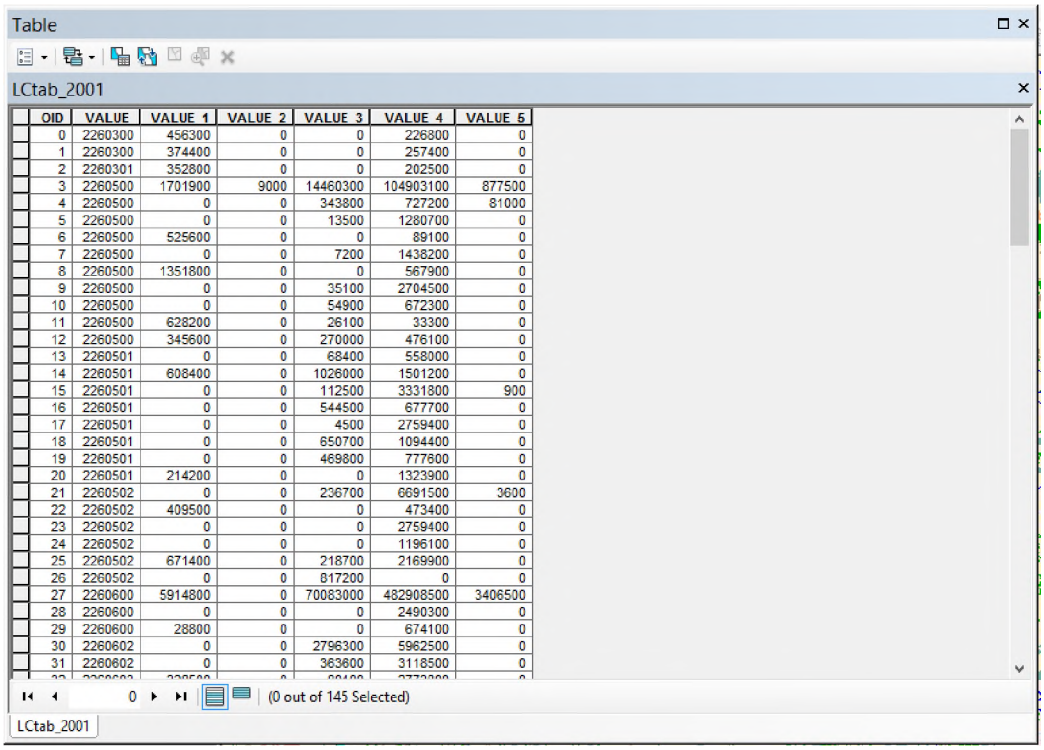
Table 4: Relative densities set for land cover type.

Land cover code	Description	Relative Density (RA) in percentage
1	Residential	70 %
2	Commercial/Industrial	5 %
3	Agricultural	15 %
4	Natural	10 %
5	Water/Wetland	0 %

Step four requires the calculation of N, the actual population or any value that needs to be mapped e.g. employment or access to water. These were actual values present in the census data. For this step, the data needed to be converted from a polygon layer to raster layer. The input for the field was set to the appropriate variable e.g. total population. This step was repeated 36 times for all the categories to be mapped in the 2001 and 2011 census.

Step five calculated E, the expected population for each cell. This step used the relative densities calculated in step three. E is decided by the proportion of land cover type in all the small areas. The SP_code (small place code that is one of the fields in the census data) was used when creating this layer as each of these values are unique for each small place and this layer was used to calculate the proportions of land cover type. E was converted from a polygon to a raster layer and the SP_Code was set as the field.

Step six. A table was created to calculate E and P_A for each of the land cover types. The SP_code layer that was created in step five was selected as the input raster (zonal data) and the reclassified NLC 2000 land cover from step two was set as the feature class data. Figure 32 shows the table that was created for 2001.



OID	VALUE	VALUE_1	VALUE_2	VALUE_3	VALUE_4	VALUE_5
0	2260300	456300	0	0	226800	0
1	2260300	374400	0	0	257400	0
2	2260301	352800	0	0	202500	0
3	2260500	1701900	9000	14460300	104903100	877500
4	2260500	0	0	343800	727200	81000
5	2260500	0	0	13500	1280700	0
6	2260500	525600	0	0	89100	0
7	2260500	0	0	7200	1438200	0
8	2260500	1351800	0	0	567900	0
9	2260500	0	0	35100	2704500	0
10	2260500	0	0	54900	672300	0
11	2260500	628200	0	26100	33300	0
12	2260500	345600	0	270000	476100	0
13	2260501	0	0	68400	558000	0
14	2260501	608400	0	1026000	1501200	0
15	2260501	0	0	112500	3331800	900
16	2260501	0	0	544500	677700	0
17	2260501	0	0	4500	2759400	0
18	2260501	0	0	650700	1094400	0
19	2260501	0	0	469800	777600	0
20	2260501	214200	0	0	1323900	0
21	2260502	0	0	236700	6691500	3600
22	2260502	409500	0	0	473400	0
23	2260502	0	0	0	2759400	0
24	2260502	0	0	0	1196100	0
25	2260502	671400	0	218700	2169900	0
26	2260502	0	0	817200	0	0
27	2260600	5914800	0	70083000	482908500	3406500
28	2260600	0	0	0	2490300	0
29	2260600	28800	0	0	674100	0
30	2260602	0	0	2796300	5962500	0
31	2260602	0	0	363600	3118500	0

Figure 32: Table using Tabulate Area Tool.

Step seven needed fields to be added to the table created in step six to complete calculations that determined the population expected for each of the landcovers in each of the small areas. In the table in Figure 32 above, each value corresponds to one of the five land covers described in Table 3 above; VALUE_1 was residential, VALUE_2 was commercial, VALUE_3 was agricultural, VALUE_4 was natural and VALUE_5 was water or wetlands.

In the table, seven fields were added; Total, P1 to P4 and E. P1 to P4 were the proportions for each category, example, P1 was the proportion of residential area for the total area.

Step eight calculated the values for six of fields created in the previous step. The following formula was used to calculated the values for Total.

$$value\_1 + value\_2 + value\_3 + value\_4$$

Value_5 which is water and wetlands was not included as its relative population density was set to zero. People tend not to live in wetlands and water and therefore its relative density was zero. When determining its proportion, value_5 would be divided by the total and multiplied by zero (relative density) and give an answer of zero.

P1 to P4 (proportion) were then calculated in the Field Calculator using the following formula:

$$P1 = value\_1/Total$$

All the required data to calculate E (expected population) had now been created. The following formula calculated E:

$$P1*70 + P2*5 + P3*15 + P4*10$$

P1 to P4 was multiplied by its R_A shown in Table 3 above, and added together to determine E.

Step nine created raster maps from the total and E calculated in the previous step. The table from step eight was joined to the census data attribute table.

Step ten followed the same procedure as step four, where the census data were converted into a raster layers. Two raster layers were created in this step, the first was a raster layer using the total field and the second using the E value calculated.

The final step, eleven, calculated the population of the cell using a Raster Calculator with the following equation:

$$[relative\ density] * [population] * 30 * 30 / ([E] * [Total])$$

The output of this equation resulted in the dasymetric maps. This process was followed to create dasymetric maps for both 2001 and 2011 covering population, race, language, schooling, employment and basic services. A total of 18 map sets (36 maps) were created using the above method.

The raster data that were created has its symbology set so that visual comparison for patterns could be identified between 2001 and 2011. The 2001 dasymetric map data was divided into five quantiles and then converted into kilometres squared and saved as layer files. The layer files that were then applied to the matching 2011 map so that a comparison could be made. The spatial statistics also utilised the symbology there the 5 quantiles were converted to no people, low, medium, high and very high population density to perform the error matrix.

All data that fell outside of the study area boundary was omitted for the analysis. Grey regions within the study area were areas that were not included in the census areas and therefore, no data exists in these regions.

4.4 Objective Four

Generate a database from the results for the DEA-NRM.

This objective required a database to be created for the DEA-NRM that contained all the above created data, that could be used when planning their various projects in the catchment.

The DEA-NRM database was divided into six categories as seen in Figure 33 below. The first two folders, 2001 and 2011, contains the data that were used to create the dasymetric maps and data created from the dasymetric process. The third folder contains the legends for the data created for the user to recreate the data with the same legends. The maps folder contains the already created maps as presented in this thesis in Chapter five below. Tsitsa Shapefiles folder contains extra data such as the catchment boundary, roads, rivers, towns etc. The final folder, Documents, contains the metadata, dasymetric mapping instructions for the user to recreate or create new dasymetric data. This manual was the adopted version of Duh's (2005), dasymetric mapping procedures that was used in objective 3. Two map package (.mpk) files were included, the first, is to be used for any data that needs to be converted into dasymetric data and the second, is to view already created maps and data, which are included in the Documents file. This database was then loaded onto a USB flash drive and handed over to the funder

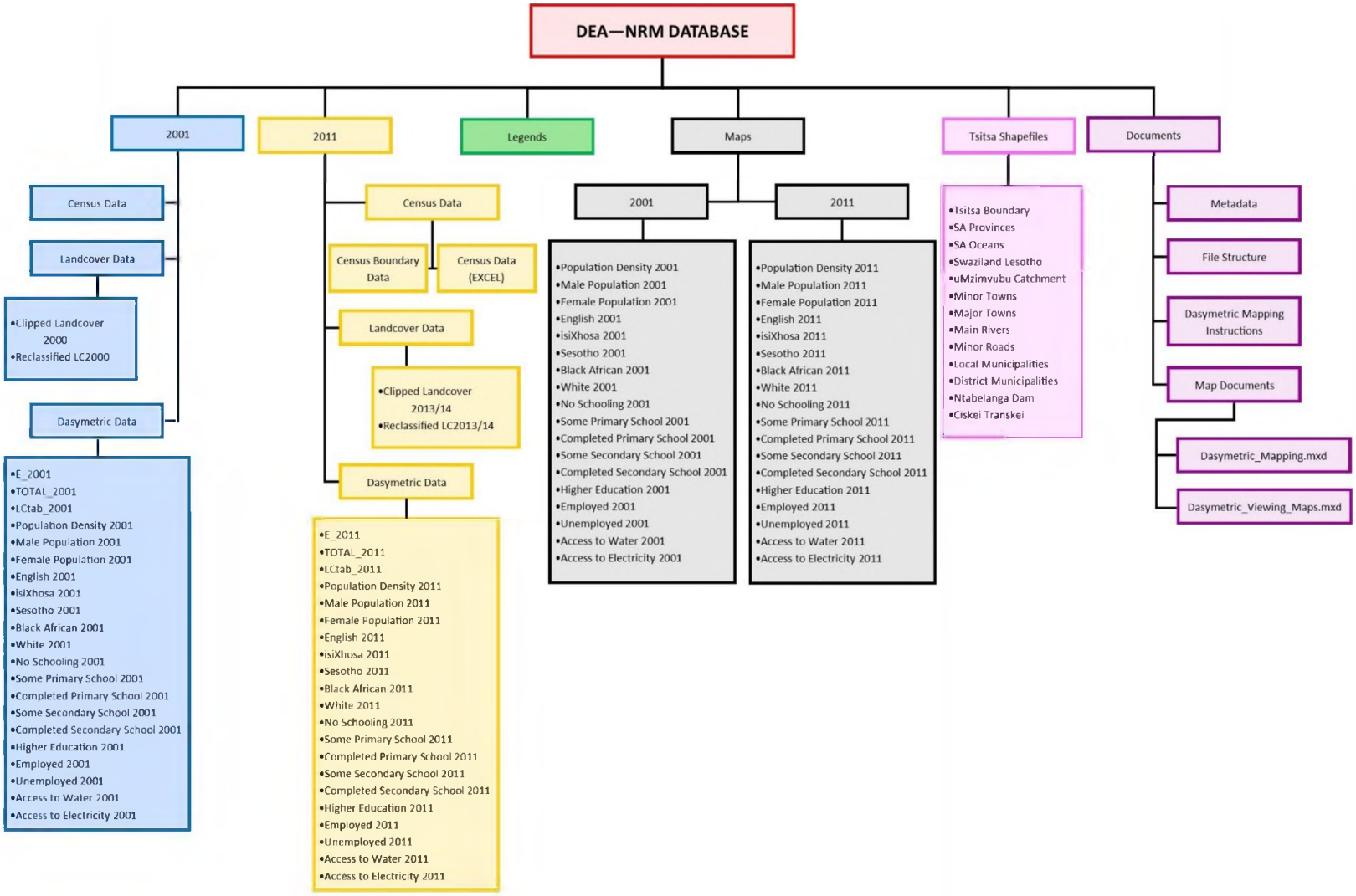


Figure 33: DEA-NRM Database File Structure.

4.5 Objective Five

Create a spatial analysis of the dasymetric and choropleth data using the 2011 population density data.

The spatial analysis was undertaken in the form of an error matrix (Figure 34) and producer, user and overall accuracy was calculated. A fishnet was used to create a set of systematic sampling points. There were 2019 sampling points in total for the study area (Figure 35). The categories no data, no people, low, medium, high and very high were used to populate the error matrix table. At a sample point, both categories for the population density choropleth map and dasymetric map were noted down. This process was done manually.

		Choropleth Map (reference source)							
		No Data	No People	Low	Medium	High	Very High	Totals	User Accuracy %
Dasymetric Map (classified map)	No Data								
	No People								
	Low								
	Medium								
	High								
	Very High								
	Totals								
	Producer Accuracy %								Total Accuracy:

Figure 34: Error matrix table used for spatial analysis.

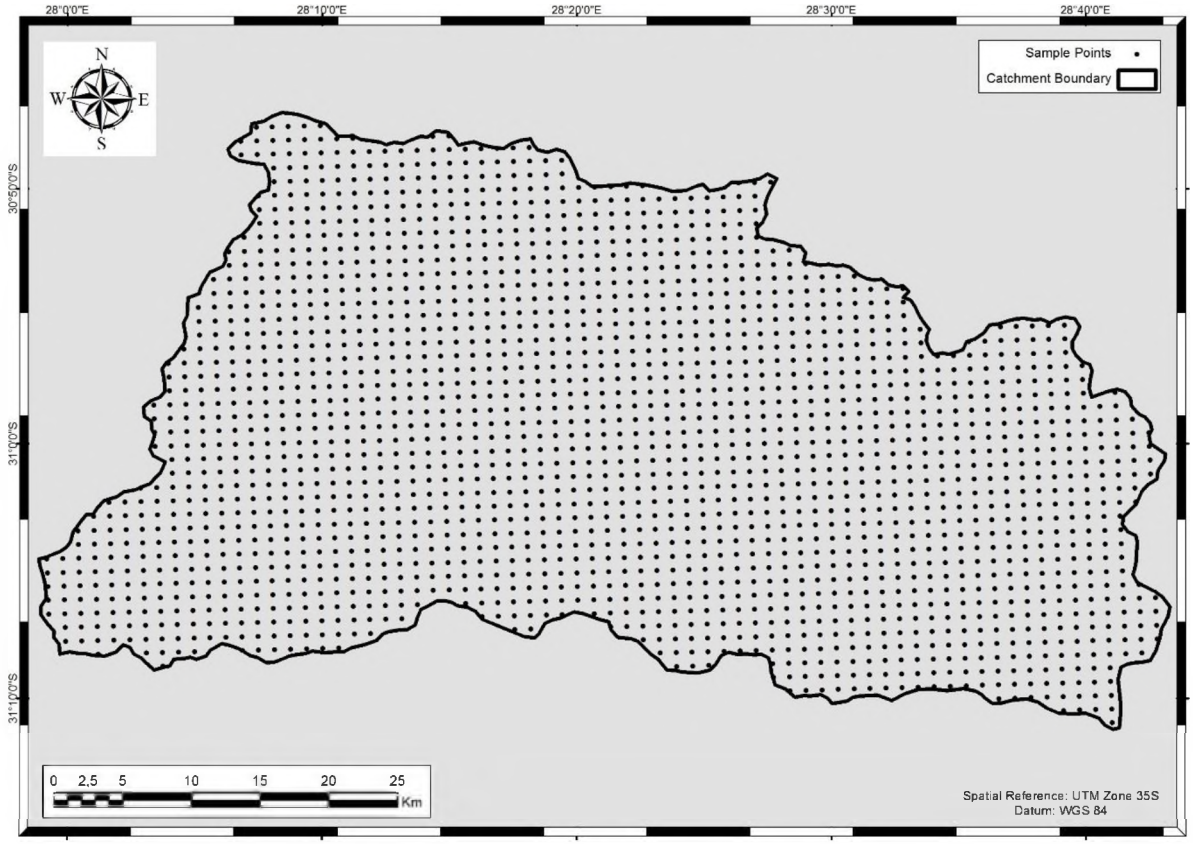


Figure 35: Map showing sample points for spatial analysis.

4.5 Objective Six

Analyse the integrated data to identify socioeconomic trends (spatial resolution of 30 m).

The 18 map sets created using the dasymetric technique described in objective 3 were compared visually to determine socioeconomic trends between 2001 and 2011 with the aid of bar graphs that summarises the census data statistics for the catchment. The map sets were divided into six categories; demographics, population race group, language, level of schooling, employment status and basic services which illustrate socio-economic trends.

Chapter Five: Results

This chapter is divided into six sections. The first section focuses on the results of the spatial statistics that is in the form of an error matrix. Sections 5.2 to 5.6 interpreted the spatial patterns and trends from the dasymetric maps. Demographics, population race group, language, level of schooling, employment status and basic services were the categories used in determining trends for the Upper Tsitsa Catchment described in Chapter one.

The results are in the form of an error matrix and the dasymetric maps show the differences between the 2001 and 2011 census data, accompanied by bar graphs for each category. It needs to be noted that the 2001 dasymetric data is presented as five quantiles of 20 percent, as this best represented the data, therefore each category represents 20 percent of the catchments population. The categories that were created from the quantile statistics were then applied to the 2011 dasymetric data so that the same classes could be used during comparison.

5.1 Spatial Statistics

Spatial statistics were performed on the 2011 population density dasymetric map. These data were compared with the population density of the 2011 population density choropleth map (Figures 36 and 37). The error matrix displayed the accuracy of the dasymetric data compared to choropleth data. Due to manually creating the error matrix table (Figure 38), patterns were observed that would not have been noted if the process was automated.

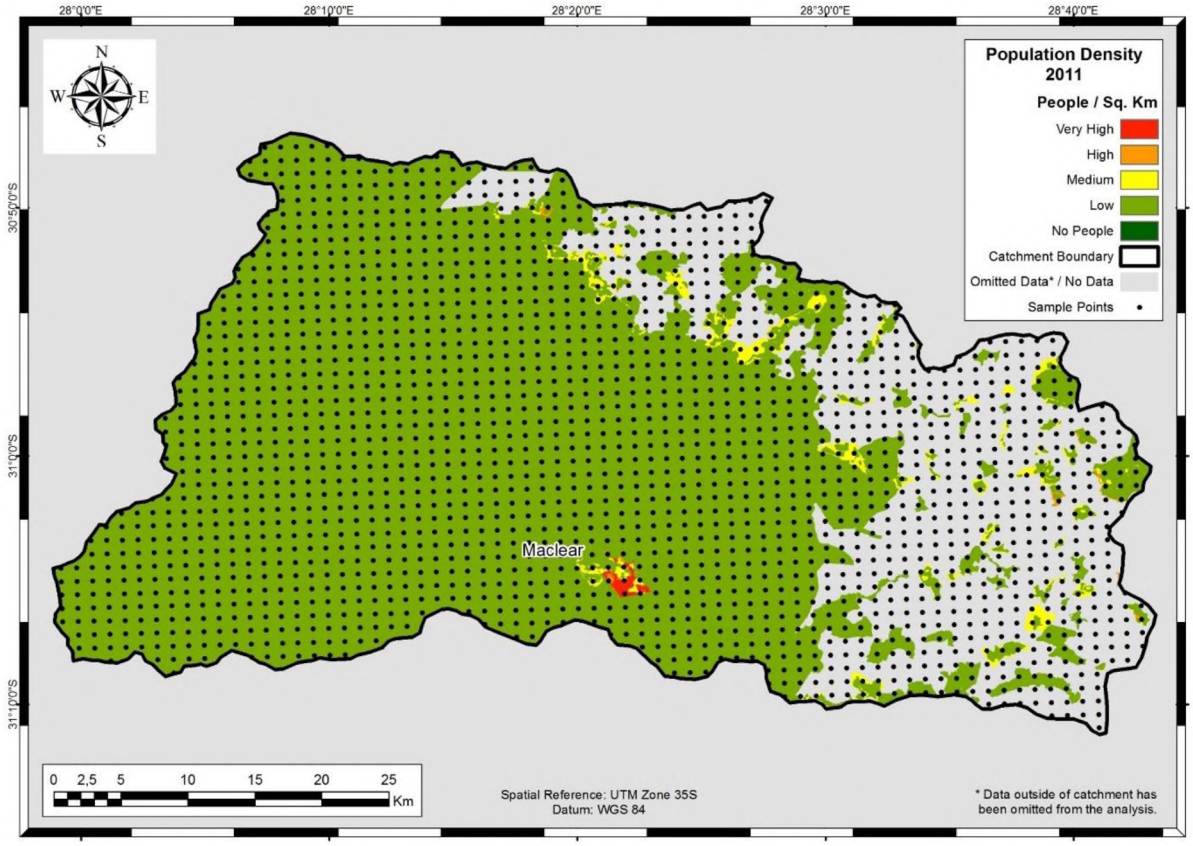


Figure 36: Dasymetric map of 2011 population density with sample points.

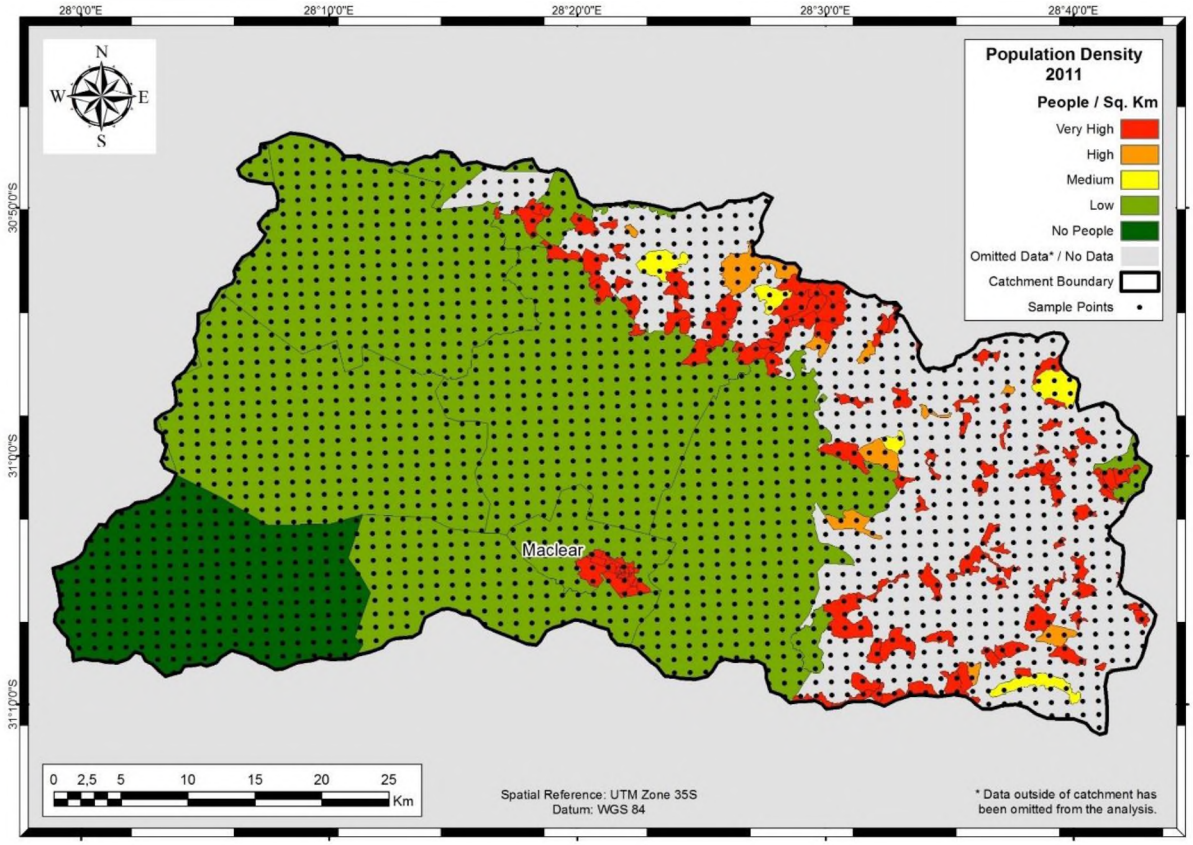


Figure 37: Choropleth map of 2011 population density with sample points.

From the error matrix (Figure 38), the overall accuracy of the dasymetric map was 82.32 %, however, these results were skewed as majority of the points fell within the 'no data' and 'low' population density categories. All the points that had medium or high densities on the choropleth maps were mapped as low population density on the dasymetric maps. Only three sample points were correct for the very high category, while none of the sample points matched in the medium and high categories. The choropleth values for very high and medium dominantly fell within the low category of the dasymetric data.

		Choropleth Map (reference source)							
		No Data	No People	Low	Medium	High	Very High	Totals	User Accuracy %
Dasymetric Map (classified map)	No Data	508	0	0	0	0	0	**508**	100,00
	No People	0	0	0	0	0	0	**0**	0,00
	Low	0	189	1151	23	18	94	**1475**	78,03
	Medium	0	0	0	0	0	31	**31**	0,00
	High	0	0	0	0	0	2	**2**	0,00
	Very High	0	0	0	0	0	3	**3**	100,00
	Totals	**508**	**189**	**1151**	**23**	**18**	**130**	**2019**	
	Producer Accuracy %	100,00	0,00	100,00	0,00	0,00	2,31		**Total Accuracy: 82,32%**

Figure 38: Error matrix of choropleth and dasymetric data for 2011 population density.

From these statistics and visual observations made while compiling the error matrix data, it was noted that the choropleth data gave less detailed information about the small area layer than the dasymetric data. The choropleth data had a homogeneous value for the small area layer (Figure 39), while the dasymetric data had a variety of values within one small area (Figure 40) as was noted by Holt *et. al.* (2004) and Sleeter (2004) in the literature review (Section 3.3).

The choropleth data (Figure 39), illustrated that all the small areas for Maclear had a very high population density, while the dasymetric map had up to three categories within one small area (Figure 40).

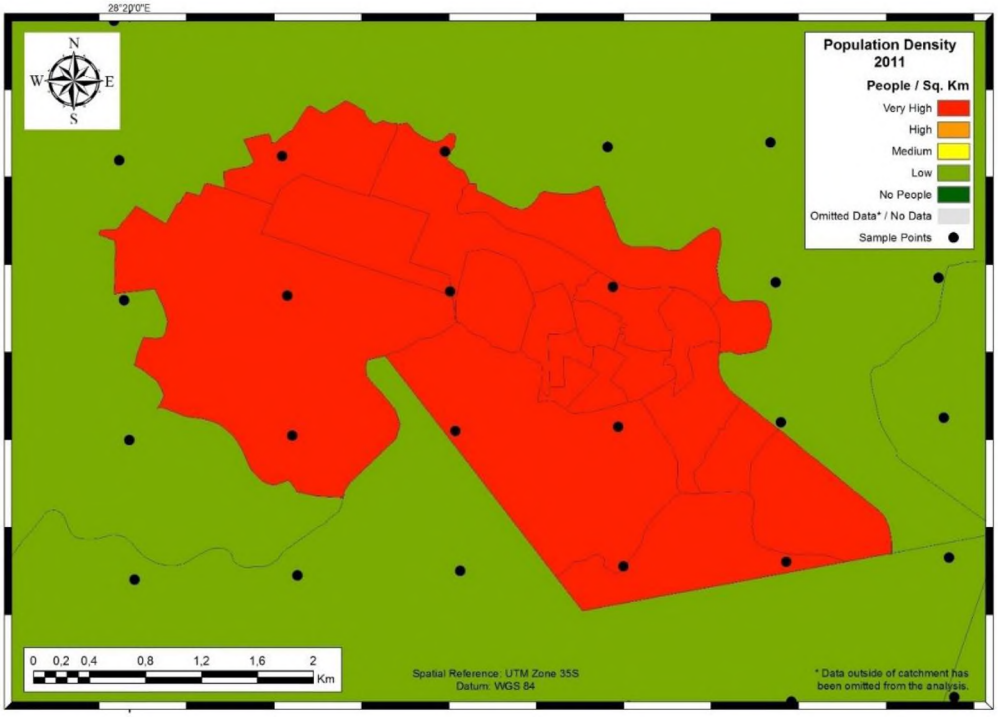


Figure 39: Choropleth data for Maclear, illustrating population density.

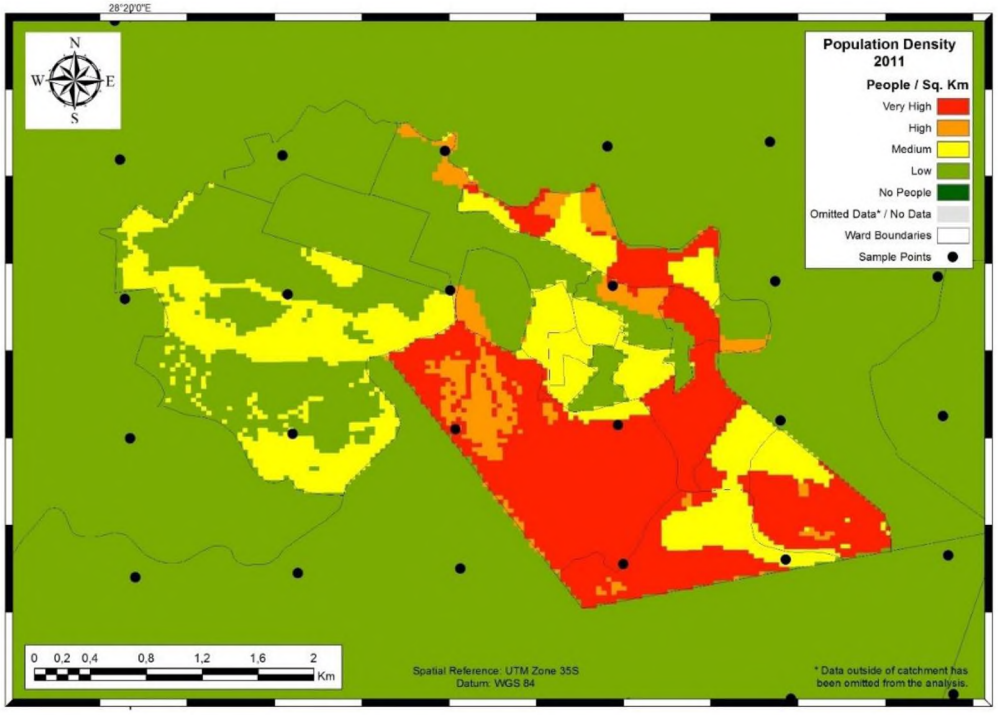


Figure 40: Dasymetric data of Maclear, illustrating population density.

The village of Makgetheng, situated in the northern region of the catchment illustrated a similar pattern to that of Maclear above, where the choropleth map had only the very high category for population density, therefore all four sample points had the same value (Figure 41). The dasymetric map had two categories; low and medium present in the same village, but three sample points indicated low and one indicated medium (Figure 42).

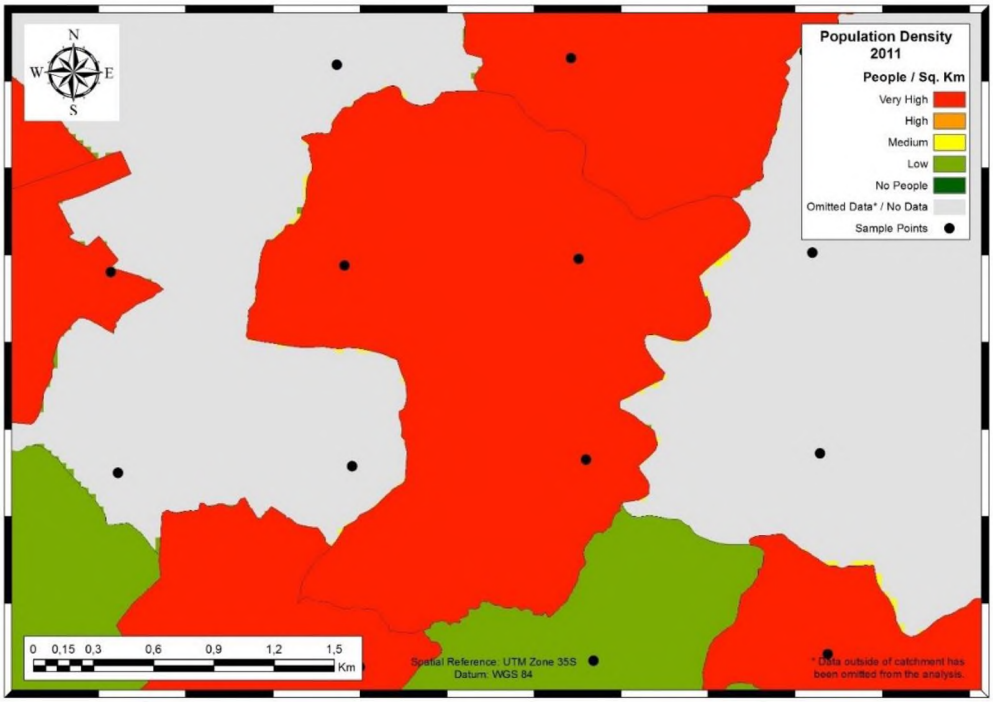


Figure 41: Choropleth data for Makgetheng village, illustrating population density.

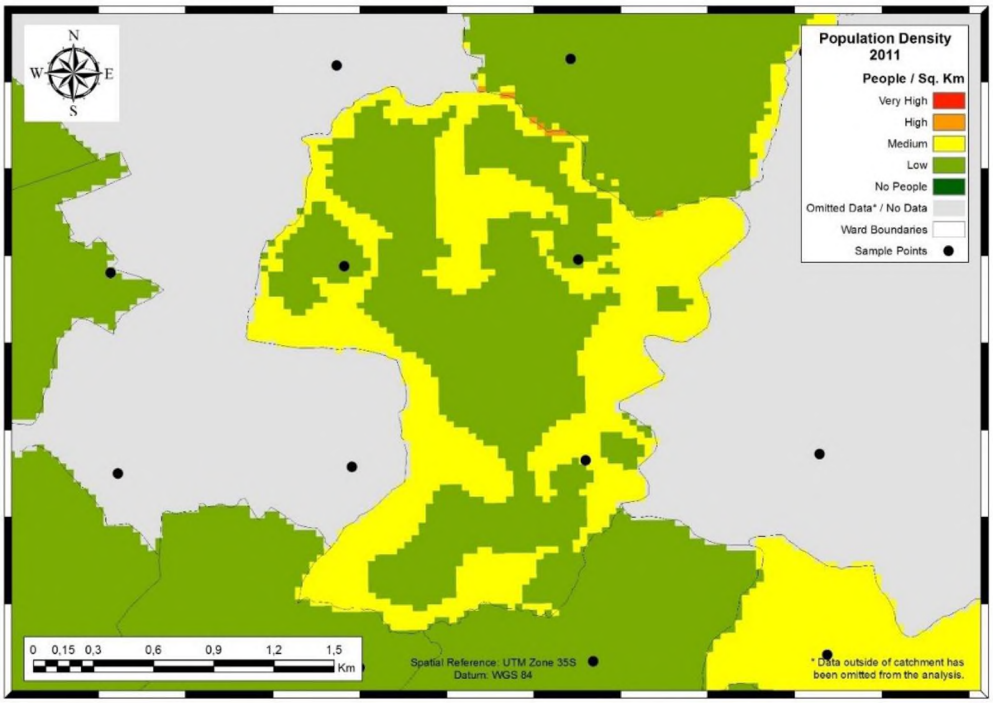


Figure 42: Dasymetric data for Makgetheng village, illustrating population density.

5.2 Demographics

This section examines the basic demographic information about the population in the Upper Tsitsa Catchment with regards to the total population as well as male and female population statistics. There are six maps in this section displaying dasymetric data with regards to the total population, male and female density.

From the bar graph below (Figure 43), a decrease in the total population was seen for the catchment. The total population decreased by 10 754 (19%) persons. The male population decreased by 4 330 (17%) persons while the female population decreased by 6 409 (22%) persons over the last ten years. The decrease in female population was greater than that of the male population. The sex ratio for the total population of the catchment was calculated to determine the ratio of males to females in a population. During the 2001 census, there were 87 men to 100 women which increases in 2011 to 92 men to 100 women.

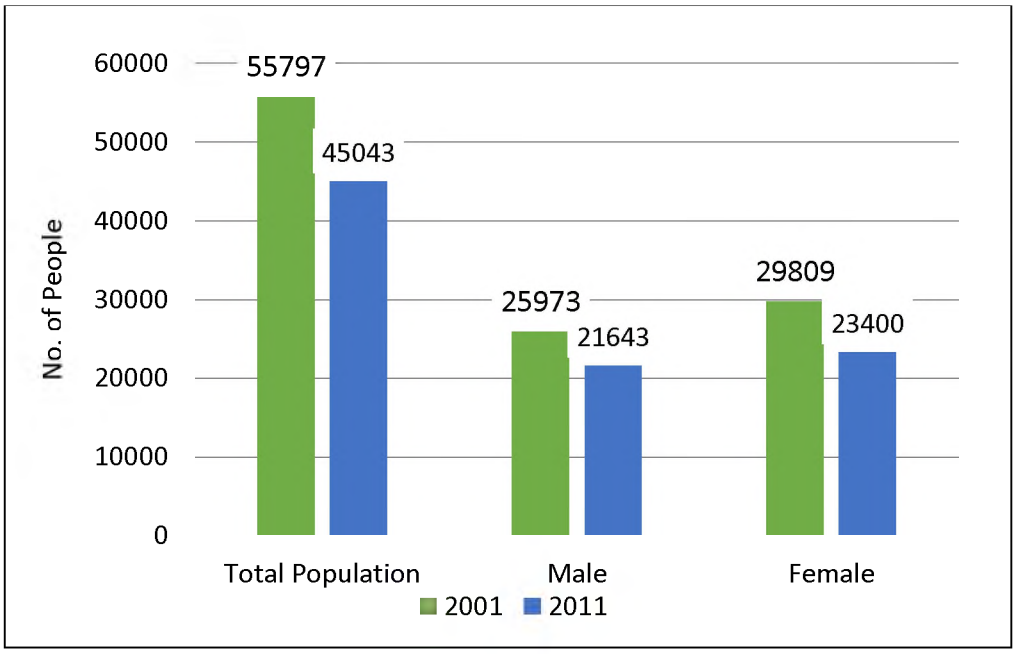


Figure 43: Demographics of the Upper Tsitsa Catchment.

The majority of the population was found in the rural villages to the East of the catchment. These villages fell within the former Transkei as shown in Figure 2, Chapter two. The rural villages had a population density of 20 to 73 or more persons per square kilometer.

For 2001 (Figure 44), 60 percent of the catchments population resided in the rural villages and urban area of Maclear, 20 percent of the population resided in the rest of the catchment, predominantly on commercial farms and the last 20 percent of the catchment had no inhabitants (zero persons per square kilometer). Maclear had a higher population density, with most of the urban area having 73

or more persons per square kilometer. The central business district, which lies between the suburbs and township had a lower density of 20 to 39 persons per square kilometer.

The 2011 dasymetric maps illustrated that the population density had decreased since 2001 (Figure 45). There were significantly less areas with a population density of 73 people per square kilometer. Forty percent of the catchments population had a density of between one and 39 people per square kilometer. The township situated in Maclear remained constant since 2001, while the density decreased in the suburbs situated in the west. The trends displayed in the maps above showed the same trends as the bar chart in Figure 42, with the average decrease of 20 percent in population density.

The male population for 2001 (Figure 46) had 60 percent of the catchment's population within the rural villages and Maclear. Twenty percent of the catchment had a male density of 35 or more men per square kilometer. The villages that were once very highly populated by men demonstrated a decrease to one to 17 men per square kilometer in 2011 (Figure 47).

Eighty percent of the catchments female population had a density of 10 to 42 females per square kilometer (Figure 48). The decrease in female population density within the catchment between 2001 and 2011 was noticeable from the maps. The township in Maclear had a female population density of 42 or more people per square kilometer during 2001, while in 2011 the density had decreased to one to 22 people per square kilometer.

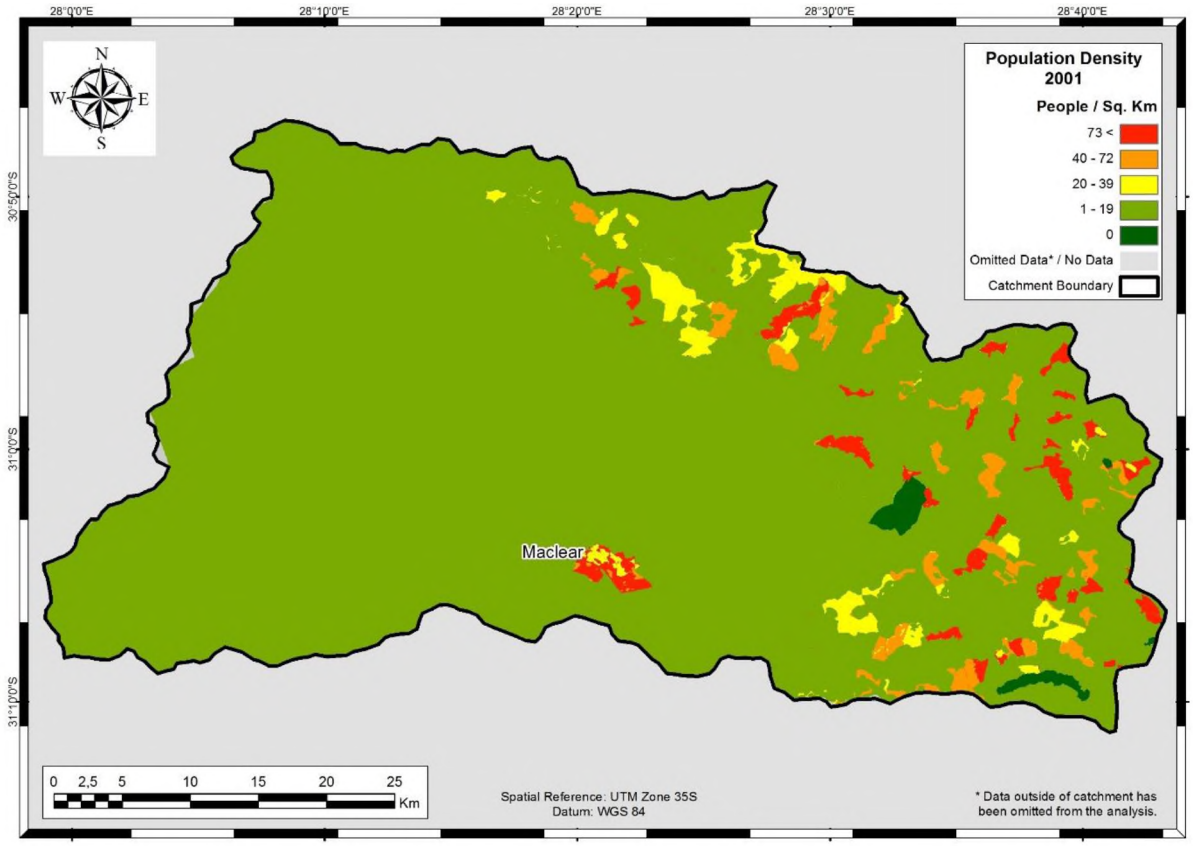


Figure 44: Population density for 2001.

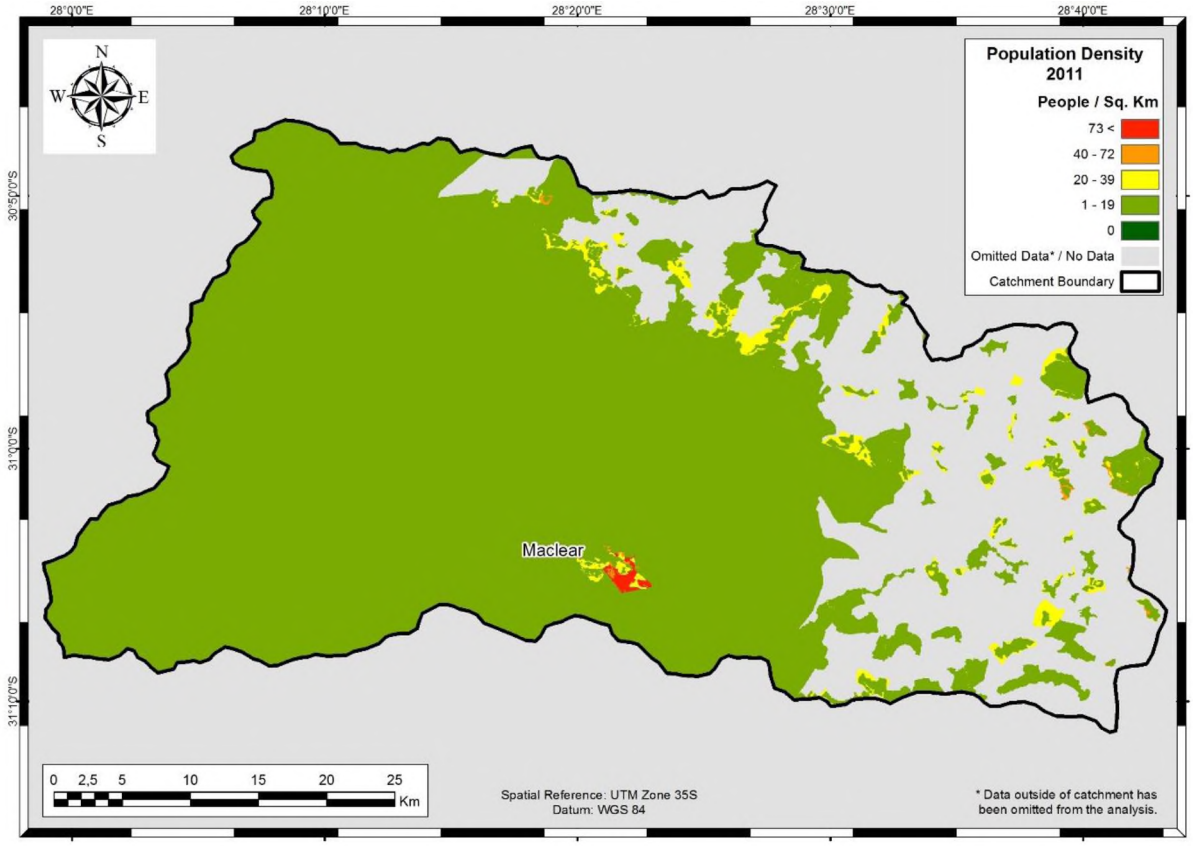


Figure 45: Population density for 2011.

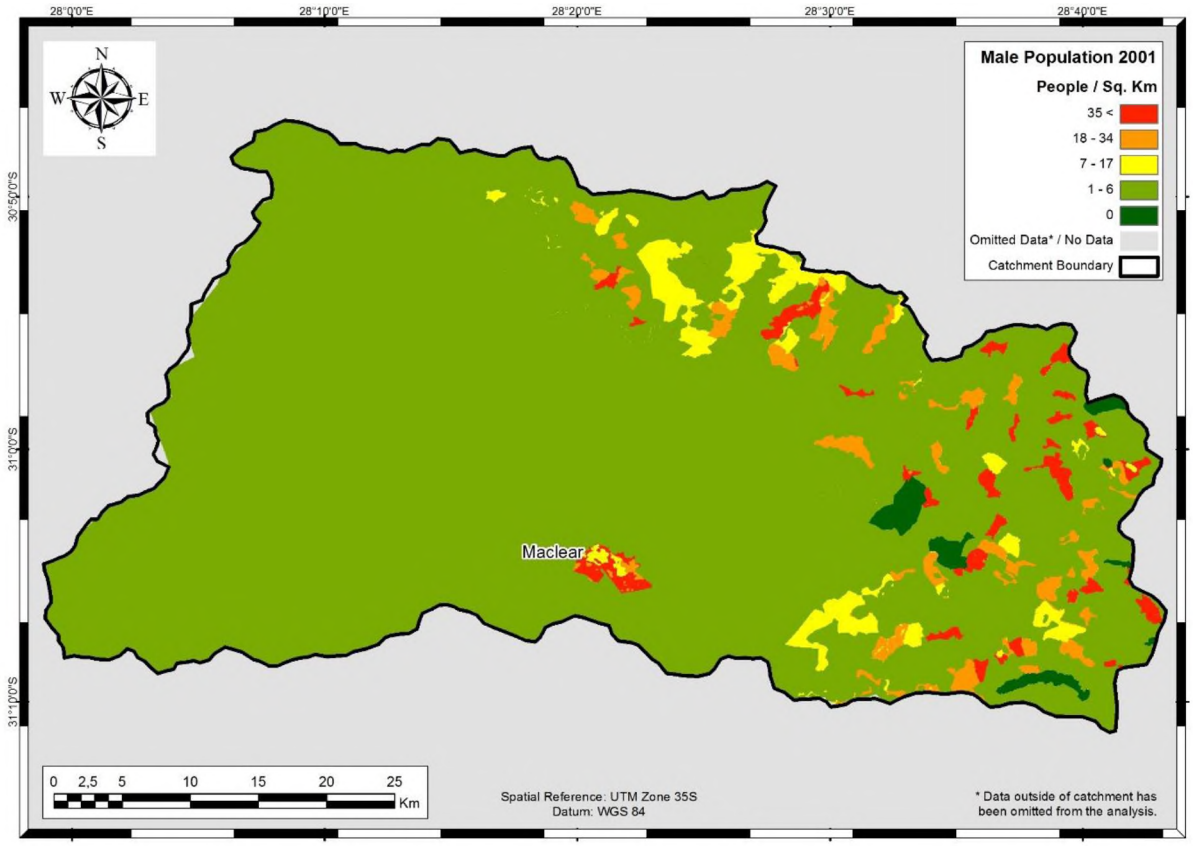


Figure 46: Male population density, 2001.

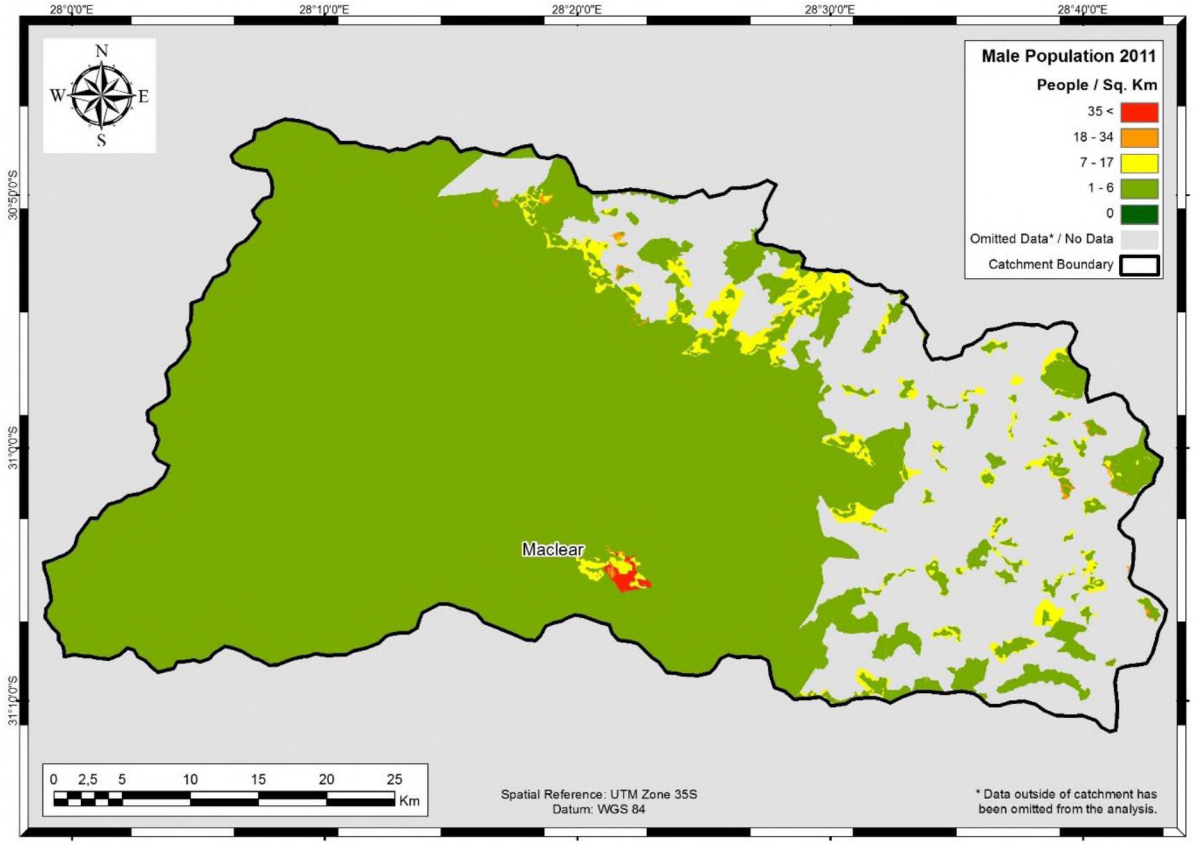


Figure 47: Male population density, 2011.

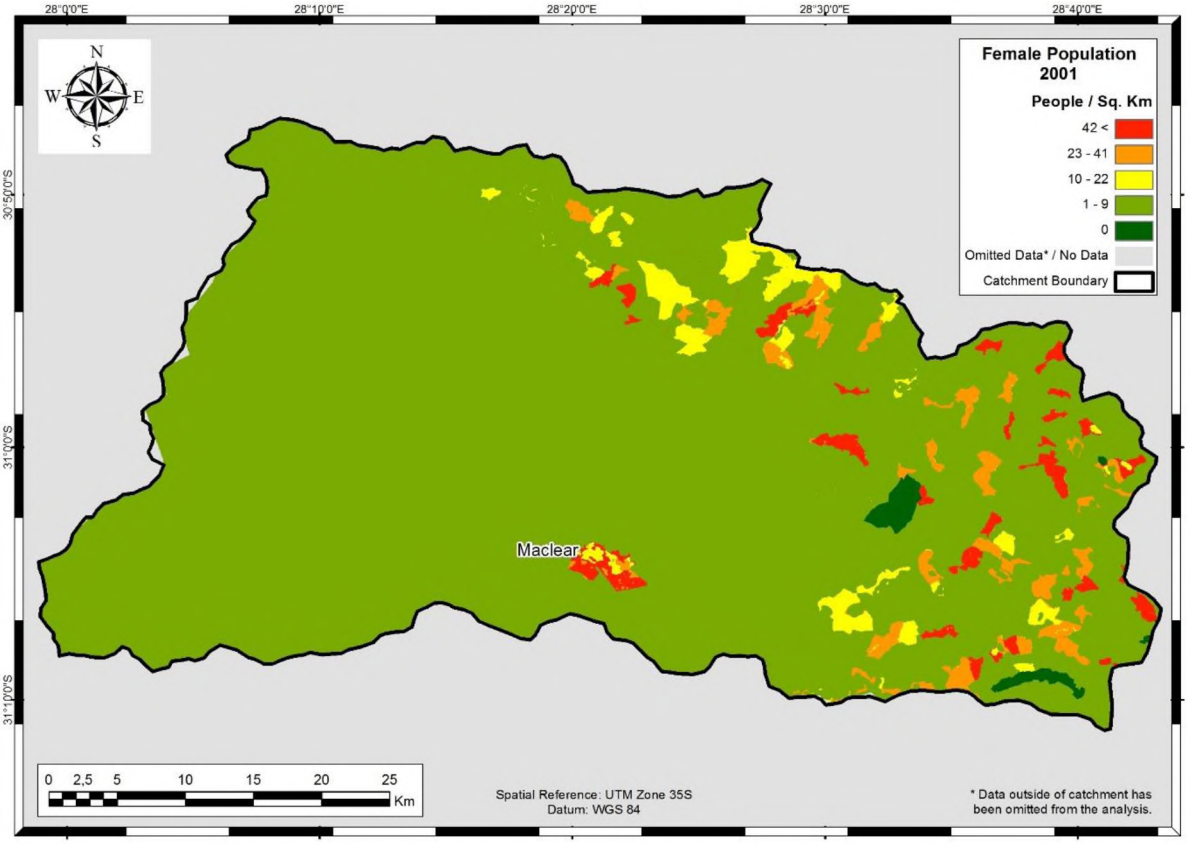


Figure 48: Female population density, 2001.

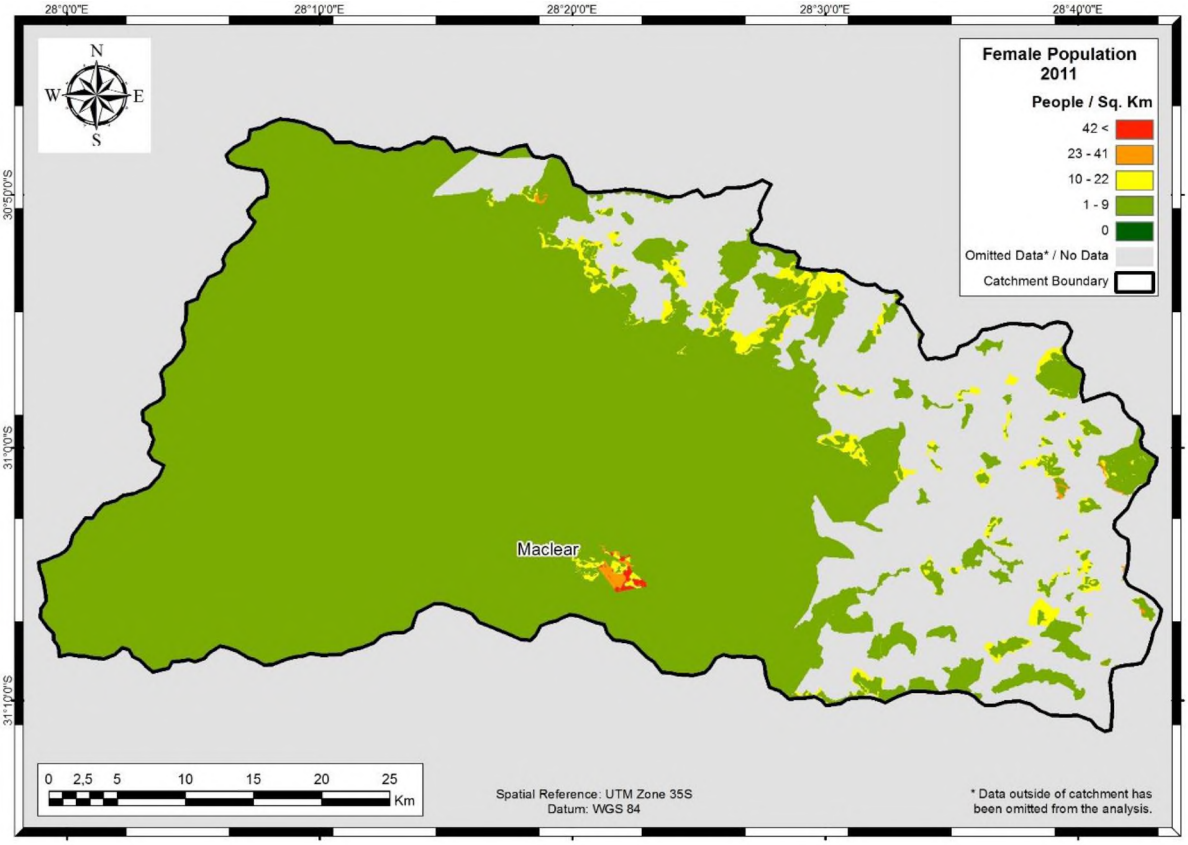


Figure 49: Female population density, 2011

5.3 Population Race Groups

The following section examines the two dominant population race groups in the Upper Tsitsa Catchment. There are four maps presented for this topic.

The Black African and White race groups were the two prevailing population race groups within the catchment (Figure 50). The catchment was mostly dominated by the Black African race group. There was a decrease in both race groups, which could be attributed to the decrease in the total population within the catchment as seen in section 5.2. The Black African population had decreased by approximately 10 621 (20%) persons and the White population by 182 (24%) persons.

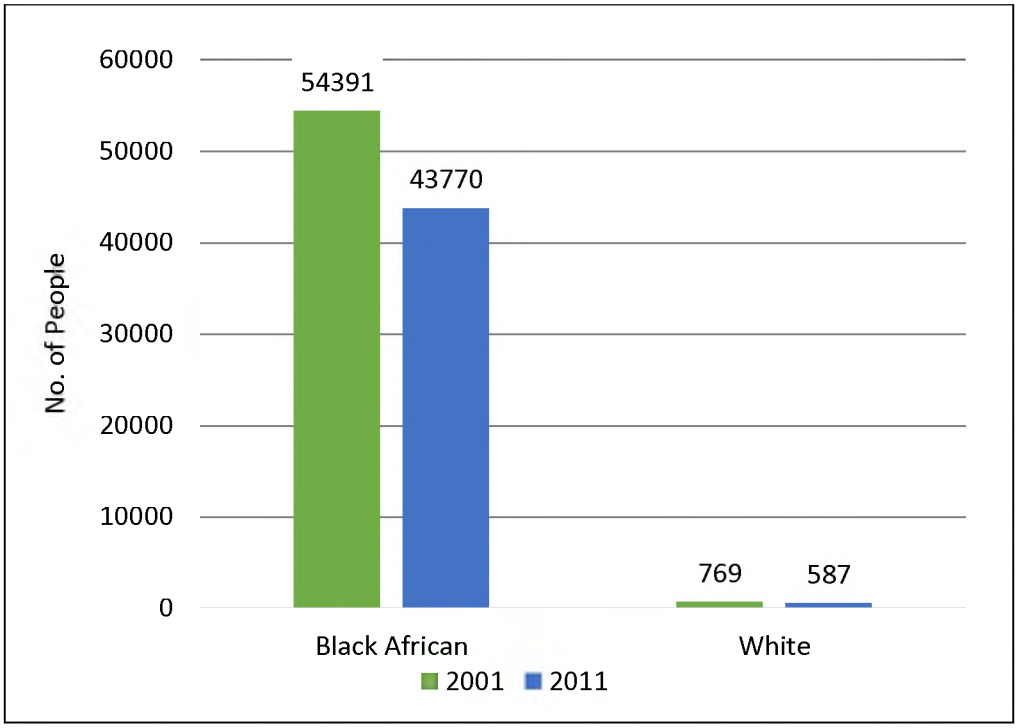


Figure 50: Population race groups for the Upper Tsitsa Catchment.

Sixty percent of the catchment's Black African population has a density between 18 and 75 or more persons within the eastern region where the rural villages are situated for 2001. The highest density for this population race group occurred in Maclear and the rural villages. Majority of the Maclear township showed a population density of 75 or more Black Africans for 2001 (Figure 51). In the township, the Black African population had slightly decreased and covered a larger area with 18 to 41 persons per square kilometer since 2001. The western suburbs of Maclear has decreased from 18 to 75 or more persons to one to 41 persons between 2001 and 2011. There was a significant decrease within the rural villages where the density reduced to one to 41 people per square kilometer (Figure 52).

The White population within the catchment, lived either in the suburbs of Maclear or on the commercial agricultural farms (Figure 3 and 4). During 2001, majority of the White population lived in the suburbs located in the west of Maclear, with a population density of 8 or more white persons per square kilometer (Figure 53). The CBD had a density of two to five persons per square kilometer which reduced to one person per square kilometer in 2011 (Figure 54).

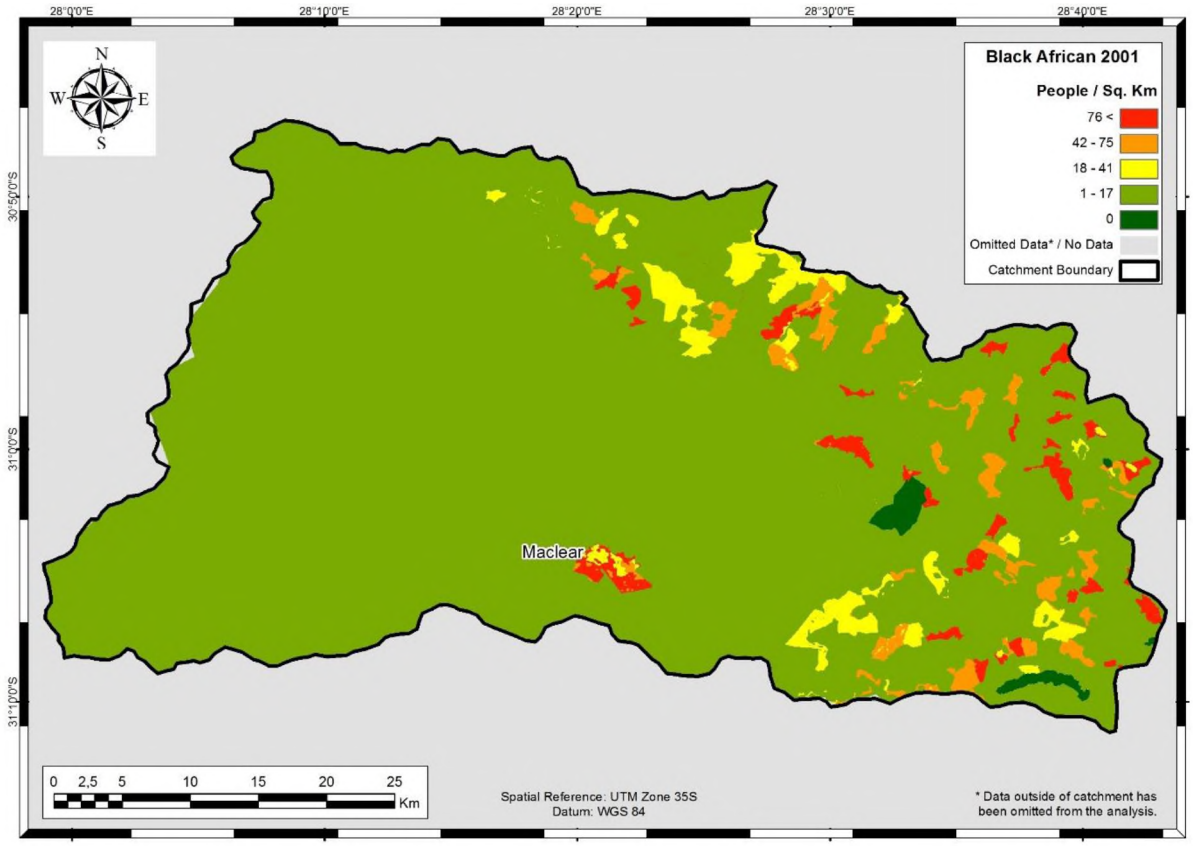


Figure 51: Black African population density, 2001.

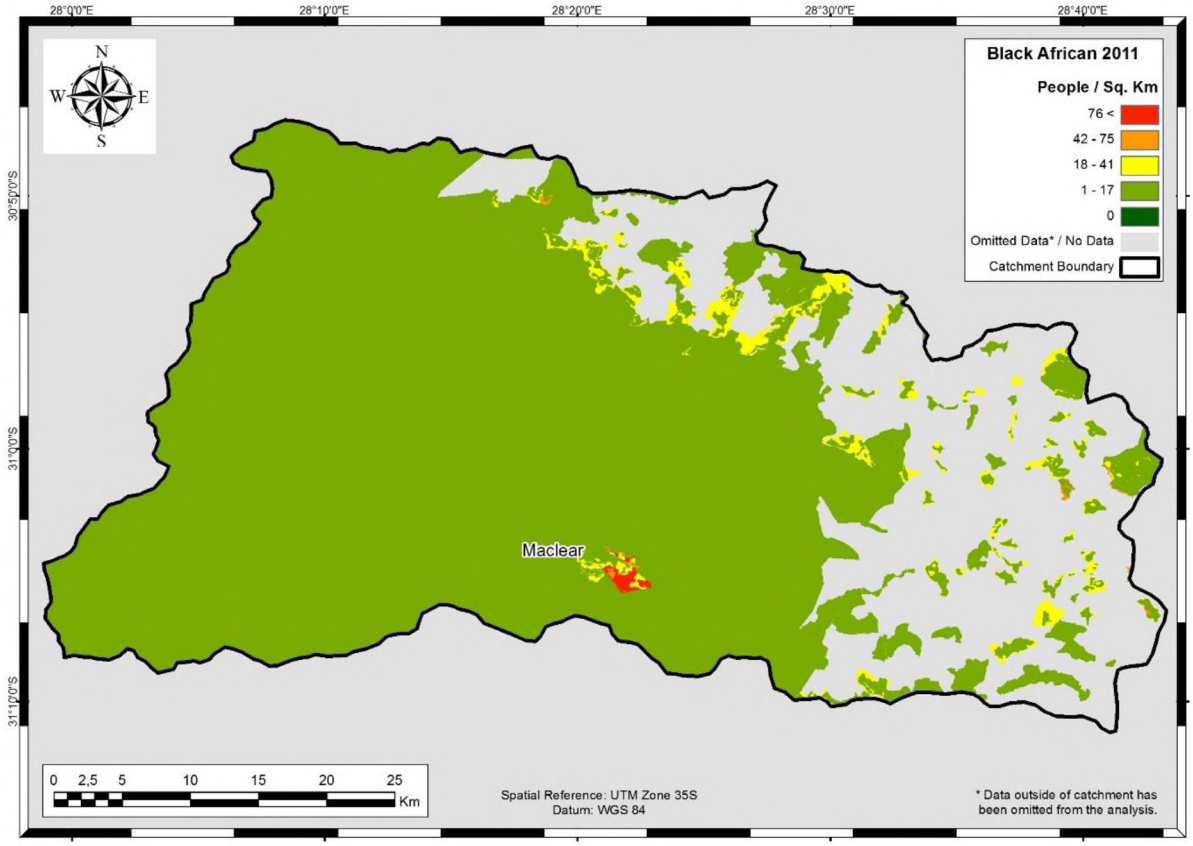


Figure 52: Black African population density, 2011.

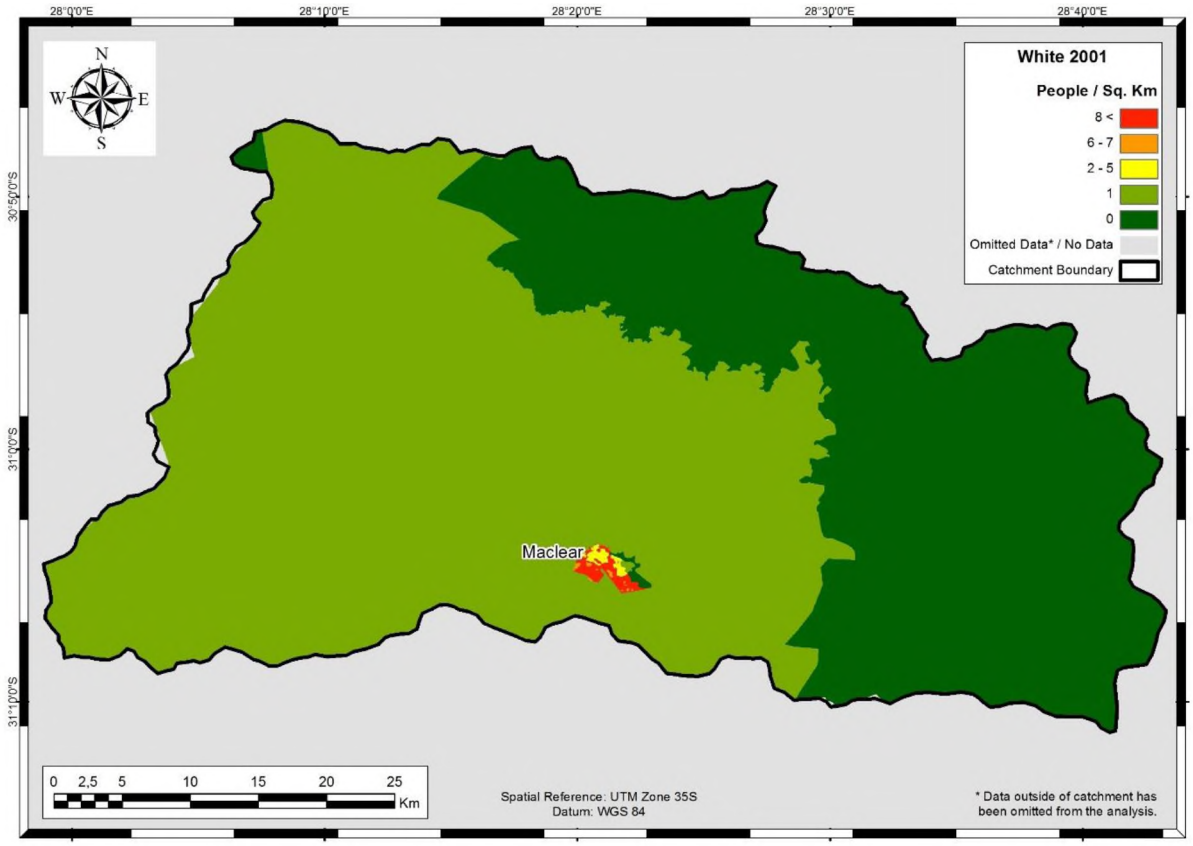


Figure 53: White population density, 2001.

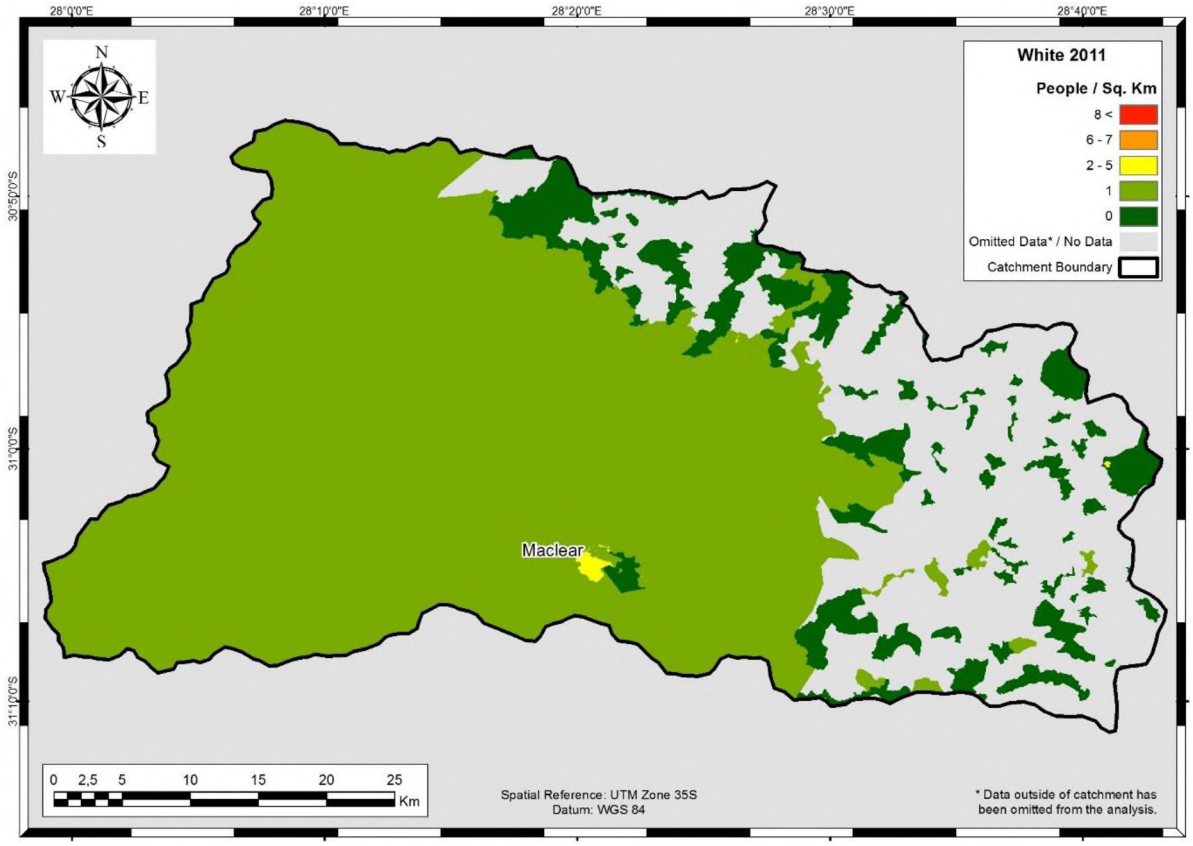


Figure 54: White population density, 2011.

5.4 Language

This section examines the distribution of the three dominant language groups in the Upper Tsitsa Catchment. There are six maps presented for this section.

The three prevailing languages in the catchment were isiXhosa, Sesotho and English. Majority of the catchment spoke isiXhosa while the minority spoke English (Figure 55). There was an increase in the number of English speakers, while a decrease in isiXhosa and Sesotho speakers in 2011. Decrease in isiXhosa and Sesotho can be linked to a decrease in the population as mentioned in section 5.2. English may have increased as more people in the catchment have attended school, which will be discussed in section 5.6. There has been an increase of 701 (230%) persons speaking the English language, while isiXhosa decreased by 10 842 (21%) and Sesotho decreased by 1 312 (42%).

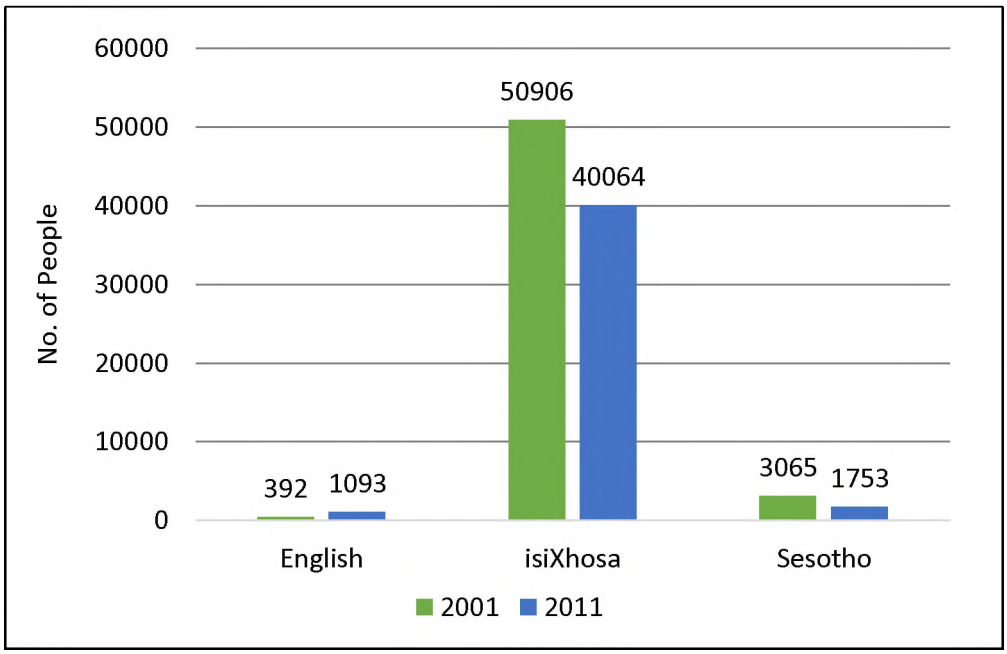


Figure 55: Languages in the Upper Tsitsa Catchment.

Three of these villages had three to six people per square kilometer, while one village has nine or more English speaking persons per square kilometer. English speakers in Maclear coincided with the location of White people within Maclear in 2001 (Figure 56) and had a distribution of nine or more persons speaking English. There was a substantial increase in English speakers since 2001 and most of the villages had between three and nine or more persons who spoke English in 2011 (Figure 57). Western (White area) Maclear was still dominantly English speaking with nine or more persons per square kilometer, although there was a decrease to seven to eight persons on the southern side of this region in 2011. The township of Maclear shows that majority of the area had three to six people

speaking English, which has declined since 2001. This could be due to predominantly isiXhosa people moving from the rural villages to the urban area.

There were between 77 and 301 persons per square kilometer within the rural villages that spoke isiXhosa (Figure 58). Maclear had between 151 and 301 or more persons speaking isisXhosa in 2001. The rural villages were still densely populated by isisXhosa speakers, although the villages had become more nucleated as people have moved closer together since 2001. The township region of Maclear exhibits the same patterns as in 2001 (Figure 58). Western Maclear has declined in isiXhosa speakers since 2001, with majority of the region having one to 150 persons per square kilometer.

Majority of Sesotho speakers were located in the rural villages to the North of the catchment for both 2001 and 2011 (Figure 60 and 61). The northern villages had between two or 11 or more Sesotho speakers in 2001 (Figure 60). Sesotho speakers had a density of two to six persons per square kilometer in Maclear and lived in the same areas as the Whites (Figure 53). There has been decline in Sesotho speakers since 2001 (Figure 61). The rural villages in the north, which had a high density of Sesotho speakers had between one and six persons per square kilometer by 2011. Sesotho speakers were only found in the township of Maclear and covered less area than in 2011, although the density of two to six persons per square kilometer remained the same.

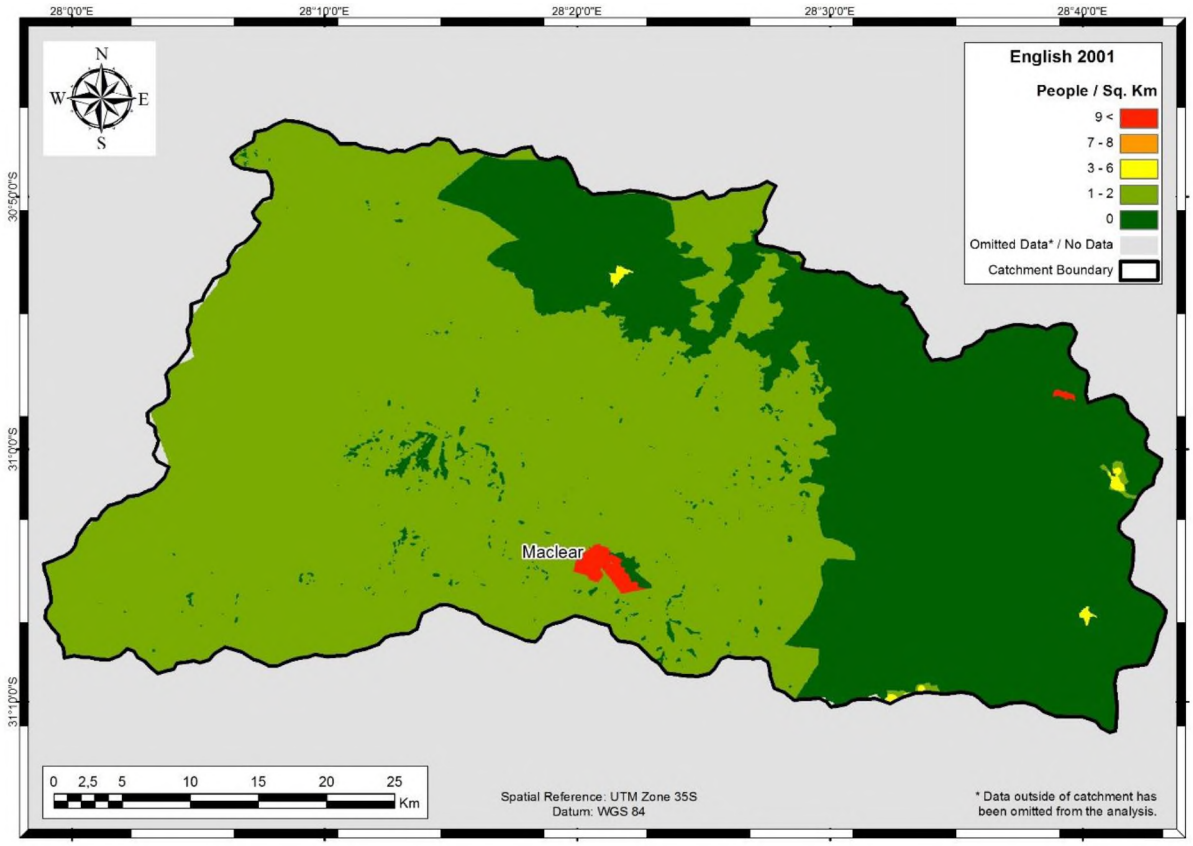


Figure 56: English language distribution, 2001.

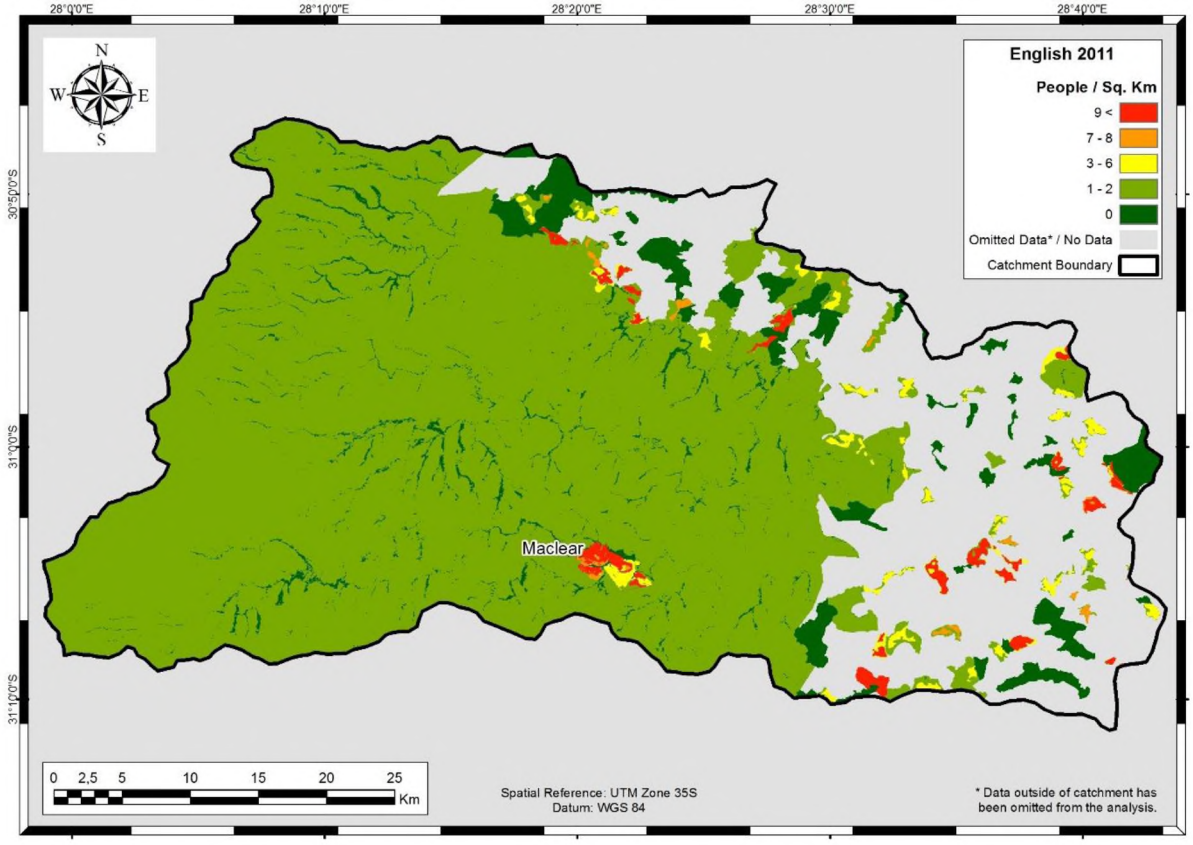


Figure 57: English language distribution, 2011.

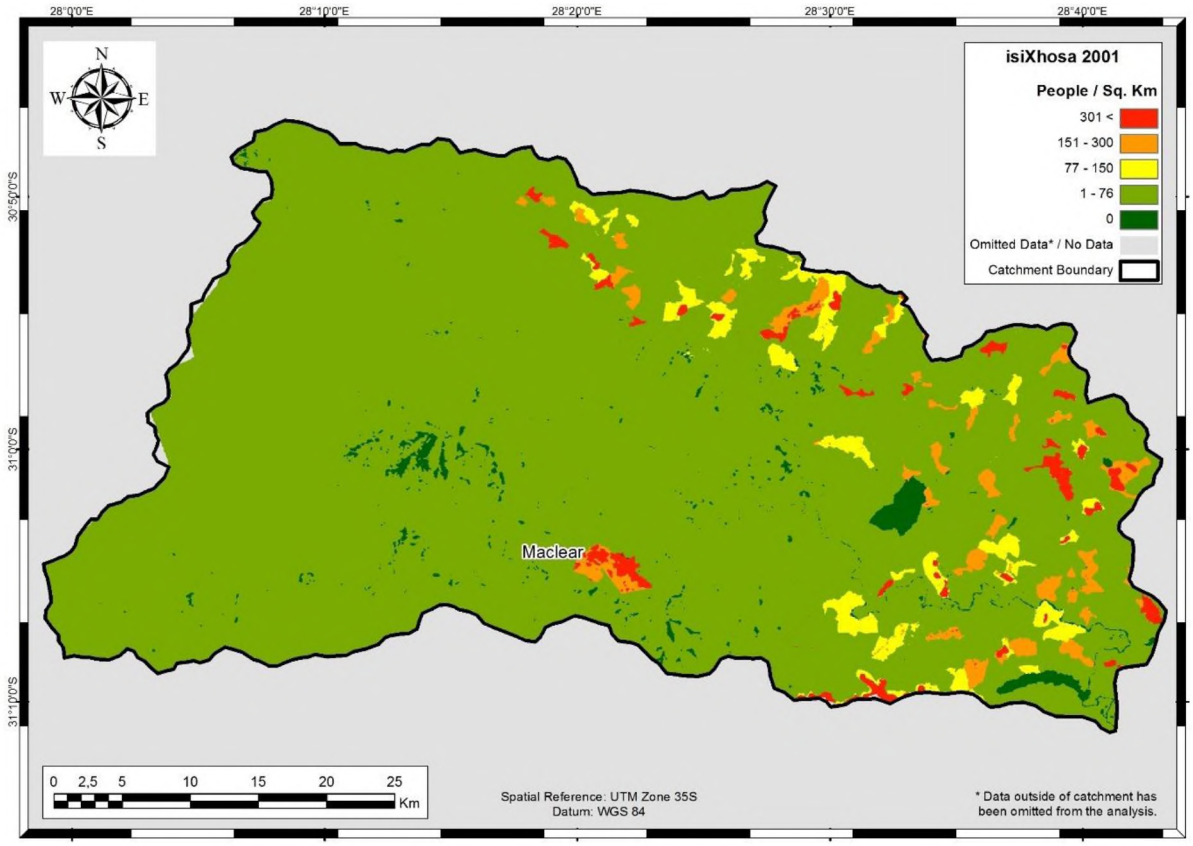


Figure 58: isiXhosa language distribution, 2001.

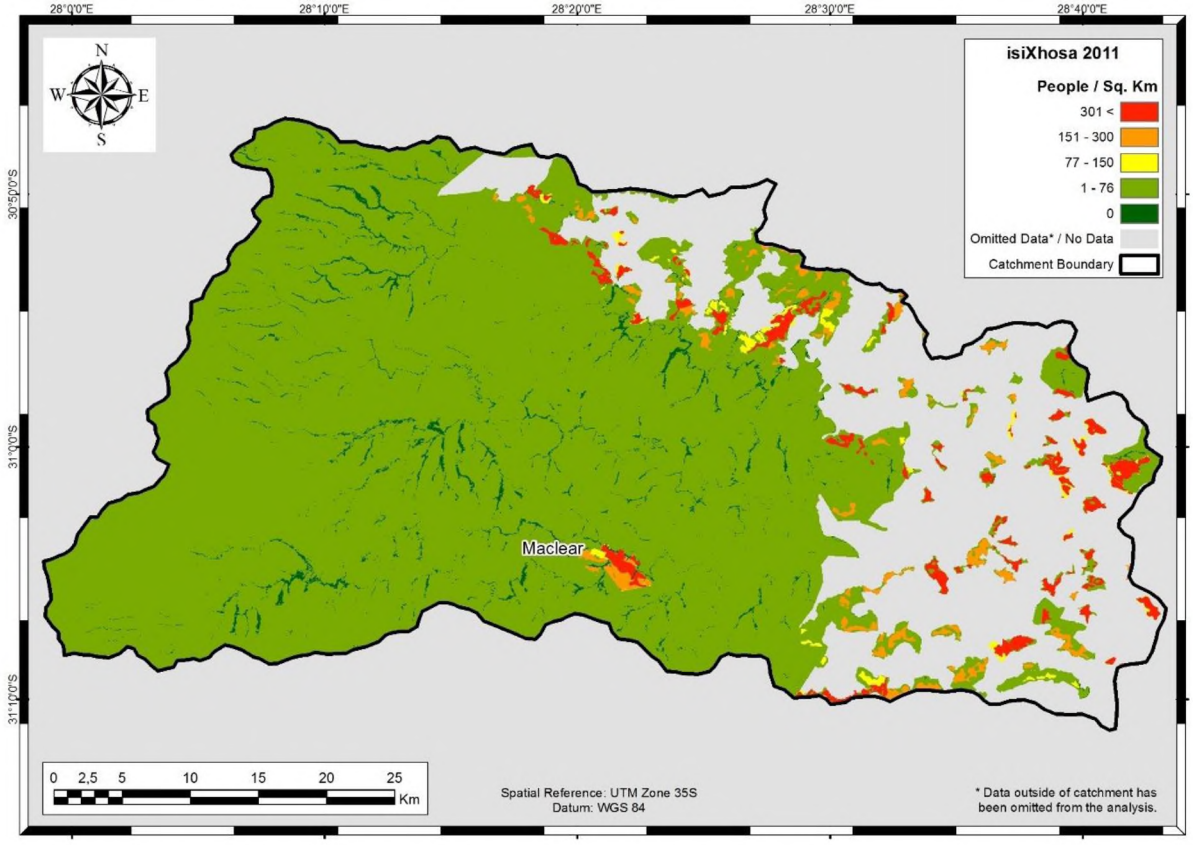


Figure 59: isiXhosa language distribution, 2011.

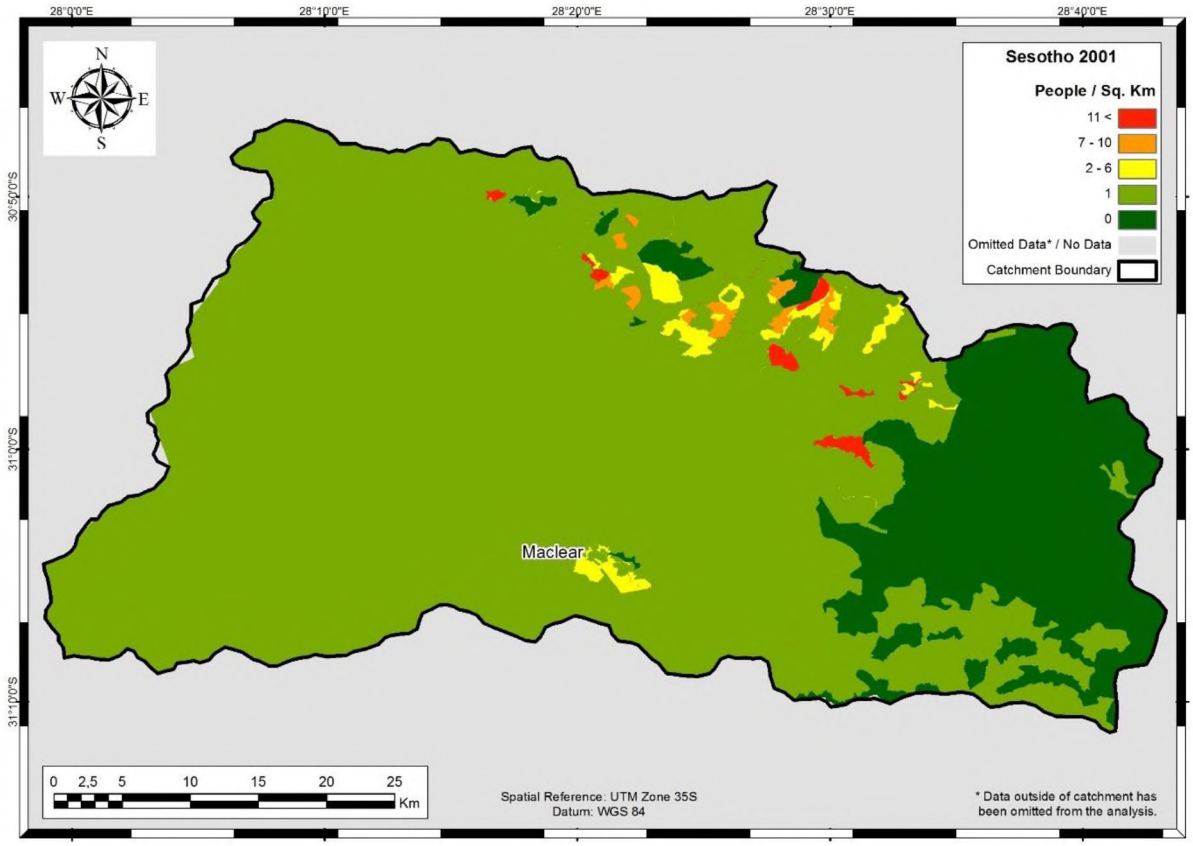


Figure 60: Sesotho language distribution, 2001.

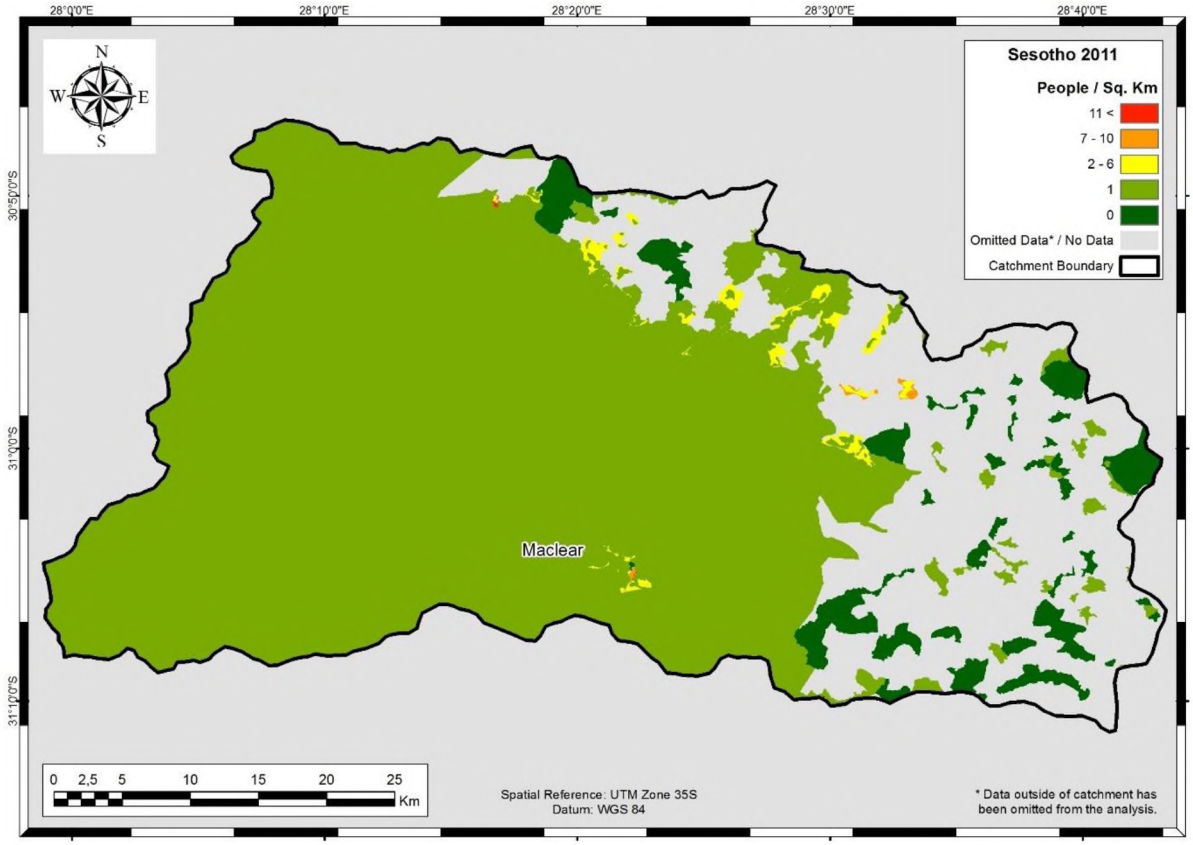


Figure 61: Sesotho language distribution, 2011.

5.5 Level of Schooling

This section examines the level of schooling in the Upper Tsitsa Catchment. There are 12 maps presented in this section.

Education has increased in the catchment since 2001. From Figure 62 below, it is evident that persons with no schooling has decreased significantly by 3 273 (39%) people since 2001. There had been an increase of 9 203 (137%) persons with some primary schooling, which could be linked to patterns such as an increase in the use of English as discussed in the previous section. A possible reason for the increase in schooling, irrespective at which level, may be due to better access with more schools being built in the catchment area. Newly established schools were seen when in the field. Attaining some secondary schooling had increased by 5 088 (88%) persons, while completing secondary school had only improved by 1 242 (45%) persons since 2001. Many learners may not be completing schooling as they leave school so that they can help support families at home with herding or looking after younger relatives. Secondary schools could be too far from home and therefore learners do not attend. During field observations, secondary schools were fewer and further apart. Although higher education has increased by almost 141 (12%) persons in the last ten years, the number of persons with any form of higher education is very low in this region. Persons that obtain higher education will be inclined to leave the catchment in search of better employment.

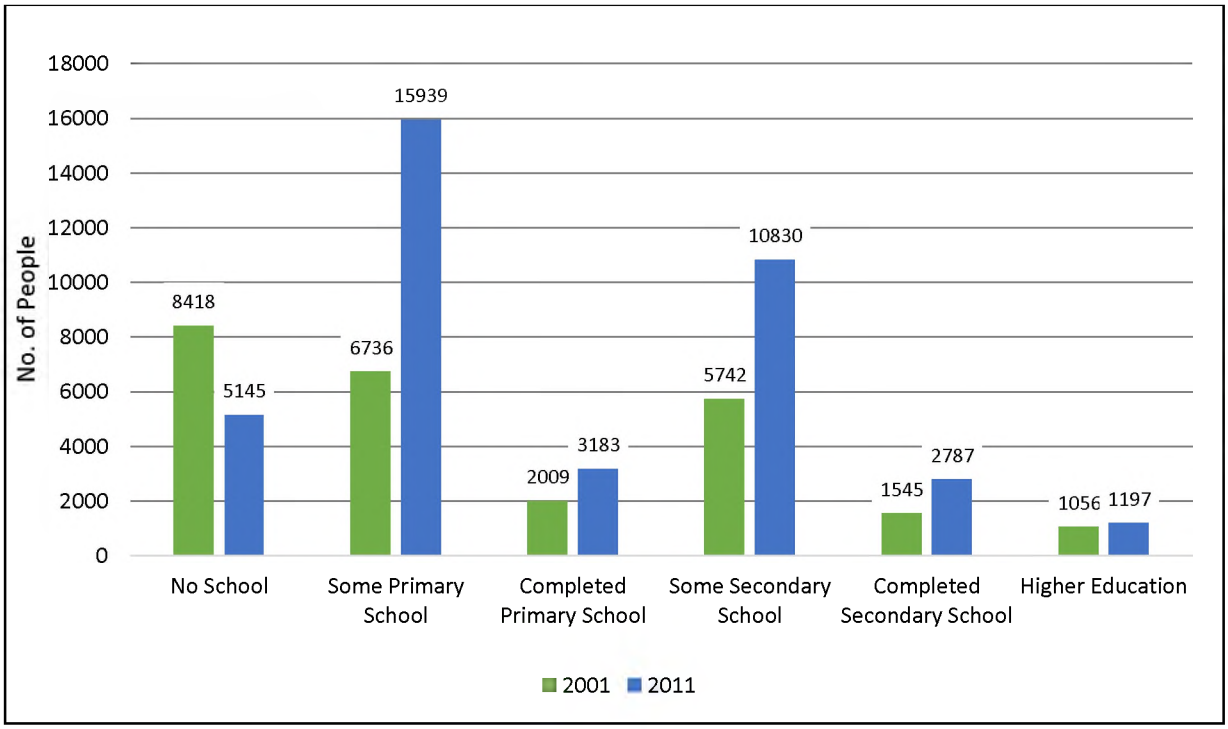


Figure 62: Level of schooling in the Upper Tsitsa Catchment.

Persons with no schooling had decreased across the catchment since 2001 (Figure 63 and 64). This could be due to some of the population leaving the catchment. The majority of the rural villages had between 17 and 57 or more persons per square kilometer with no schooling (Figure 63). In the southern side Maclear, there were between 25 and 56 persons with no schooling, which covered the same region as the White race (Figure 53). The northern areas of Maclear had 57 or more persons per square kilometer with no schooling. In 2011, rural villages had decreased to 25 to 56 persons with no schooling per square kilometer. The villages in the south east of the catchment have significantly decreased to one to 16 persons with no schooling. This decrease could be due to persons leaving the villages to move to the urban area or leaving the catchment area to seek new employment opportunities of better standard of living. The northern region of Maclear had high densities of persons with no schooling with 57 or more persons, while the southern area had decreased to 17 to 27 persons per square kilometer. This decrease could be due to more schools available in the region, as newly established schools were seen in the field.

There had been a considerable increase in some primary schooling, within the rural villages. The rural villages had 15 to 37 people per square kilometer with some primary schooling in 2001 (Figure 65). Maclear followed the trend where the northern region had 38 or more persons with some primary schooling, while the southern area has 15 to 37 persons per square kilometer in 2001. Some primary schooling had increased in the catchment since 2001 (Figure 66). There was a significant increase in some primary schooling within the villages to 18 to 38 or more persons per square kilometer. The township had 38 or more persons per square kilometer and western Maclear had between 15 to 37 persons per square kilometer with some primary schooling.

Rural villages had between four and 18 persons per square kilometer that completed primary schooling (Figure 67). Northern Maclear had 19 or more and between nine and 18 persons per square kilometer were found on the southern side of Maclear. This pattern had been observed previously in multiple categories in this chapter. There was a considerable increase over the past ten years with regards to completing primary school. Most of the rural villages had 19 or more persons per square kilometer that completed primary school. Most of the township in Maclear had completed primary school with 19 or more persons per square kilometer. Western Maclear had decreased to four to eight persons per square kilometer, which could be due to out-migration occurring in the catchment.

The majority of the rural villages had between seven and 17 persons per square kilometer that has some secondary schooling (Figure 69). In the northern areas of Maclear, there were 41 or more persons per square kilometer, while the southern area had between 18 and 40 persons per square kilometer with some secondary schooling. There was a significant increase by 2011 with regards to

persons having some secondary schooling (Figure 70). The rural villages increased to 41 persons per square kilometers, although on the perimeters of the villages, there were seven to 17 persons per square kilometer with some secondary schooling. This could be due to the learners living too far from secondary schools or being involved with family and herding responsibilities, which is common in rural areas of South Africa. The township of Maclear had 41 or more persons per square kilometer with some secondary schooling, which could be due to more schools or bigger schools being present in the urban area. Western Maclear had decreased to seven to 17 persons per square kilometer, which could be linked to out migration as described above.

Most of the rural villages had between zero and eight persons per square kilometer that completed secondary school (Figure 71). Maclear had an overall of 19 or more persons per square kilometer, with a few small areas on the north with three to eight persons per square kilometer. Since 2001, there had been an increase in persons who had completed secondary school. Most of the rural villages had increased to having three to 19 or more persons per square kilometer that completed secondary school (Figure 72). The township of Maclear had remained constant with 19 or more persons per square kilometer with secondary education since 2001. The small areas that had three to eight persons per square kilometer in the township had improved to nine to 18 persons per square kilometer.

Lastly, higher education status of the catchment between 2001 and 2011 is presented. Only a few of the rural villages in the catchment had persons with higher education (Figure 73). There were zero to seven people per square kilometer in the rural villages that had obtained higher education. One village on the far east of the catchment has 20 or more persons with higher education in 2001. Maclear had eight to more than 20 persons per square kilometer with higher education. There was an increase in number of persons with higher education since 2001 (Figure 74). Most of the rural villages had between one and 19 persons with higher education. The township in Maclear shows that there was a decrease in the number of persons with higher education, while the western suburbs had an increase in the area cover by individuals with higher education. Most of the people that had obtained a tertiary education resided in Maclear, where there are more relevant employment opportunities. Western Maclear showed that there were 20 or more persons per square kilometer.

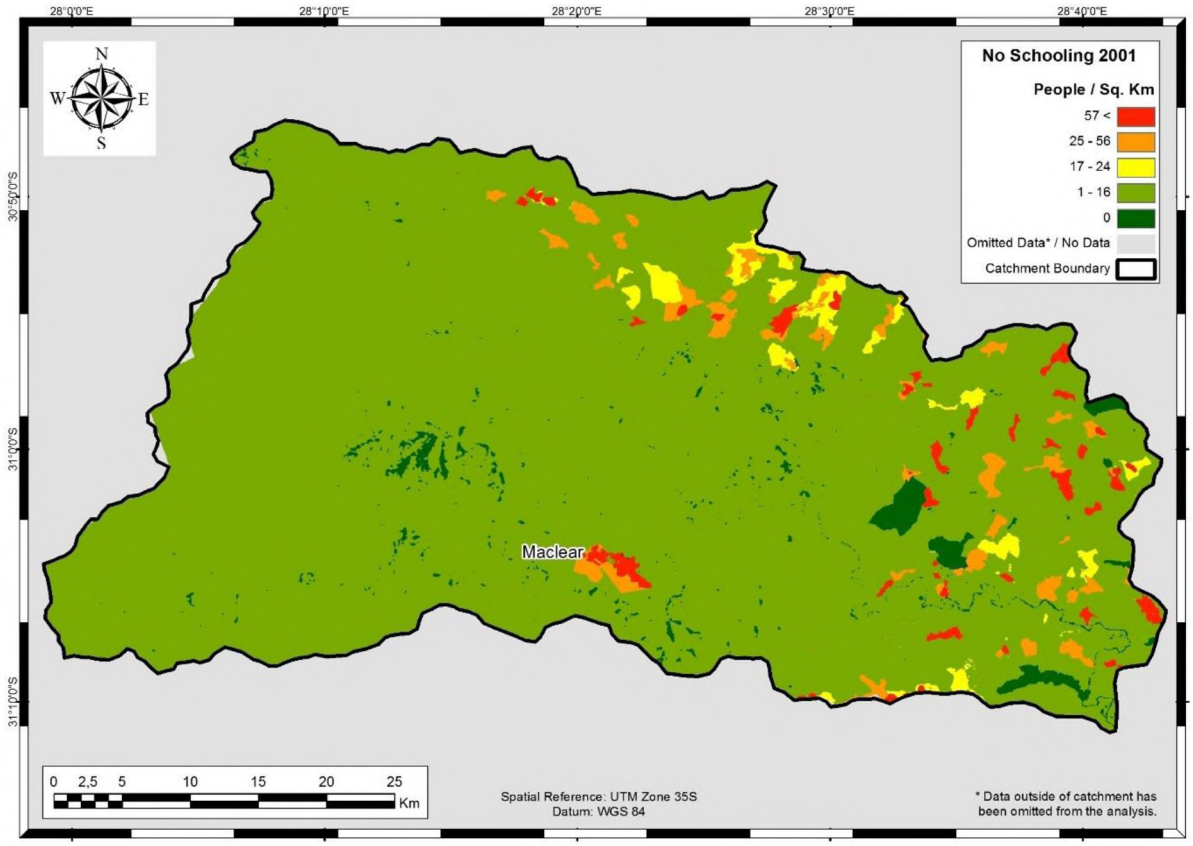


Figure 63: Persons with no schooling, 2001.

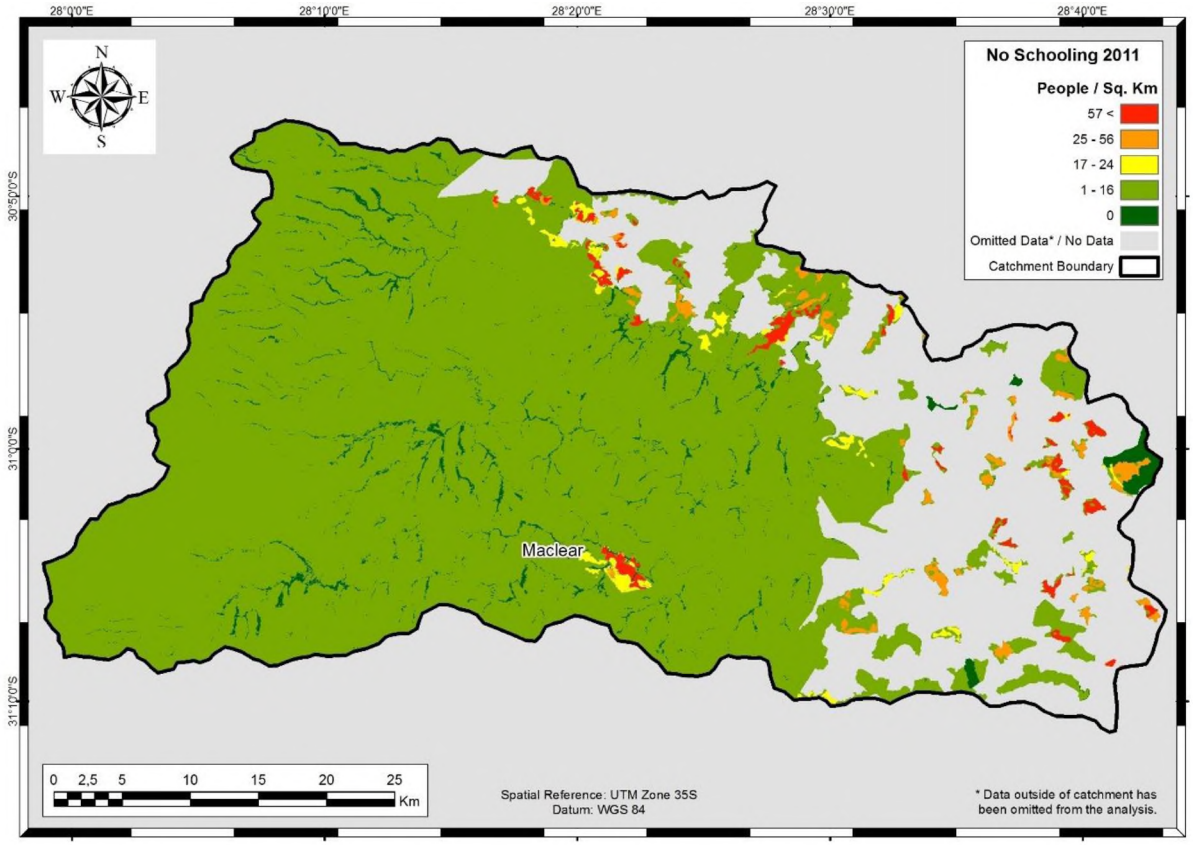


Figure 64: Persons with no schooling, 2011.

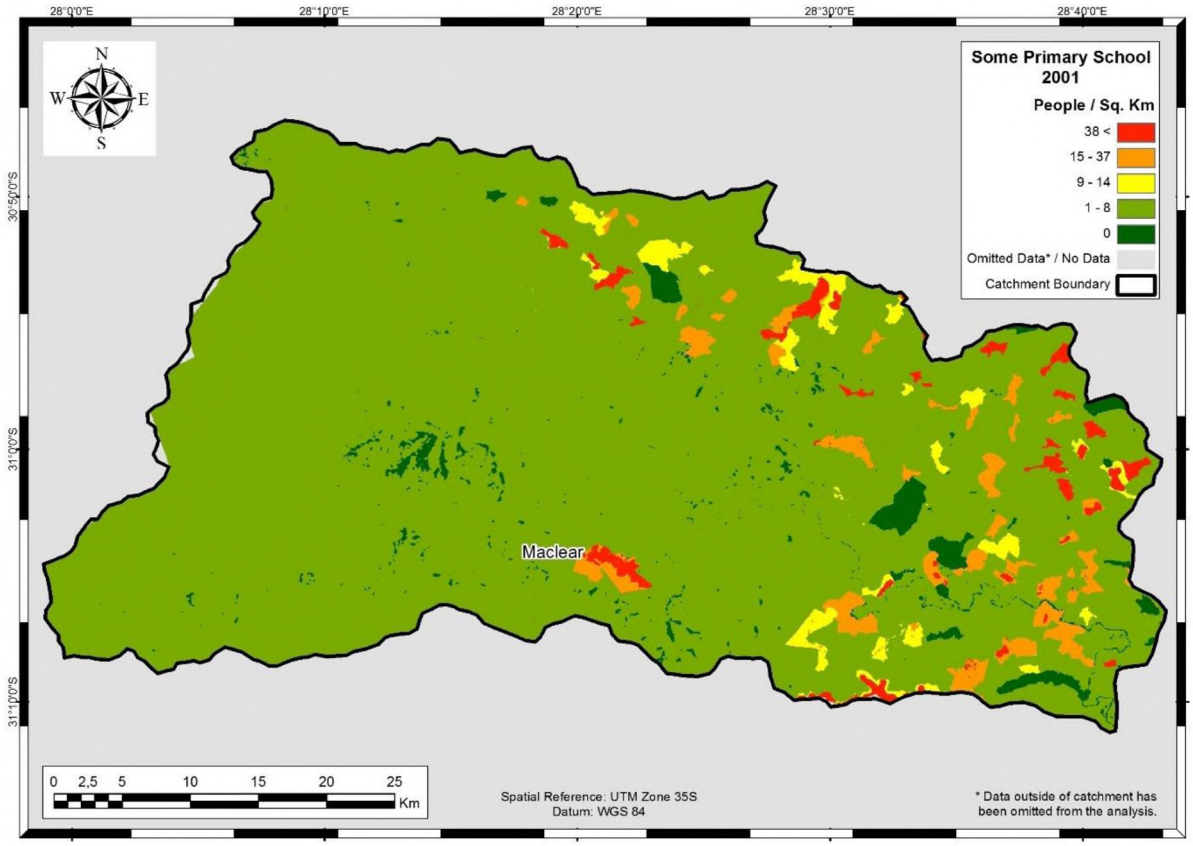


Figure 65: Persons with some primary schooling, 2001.

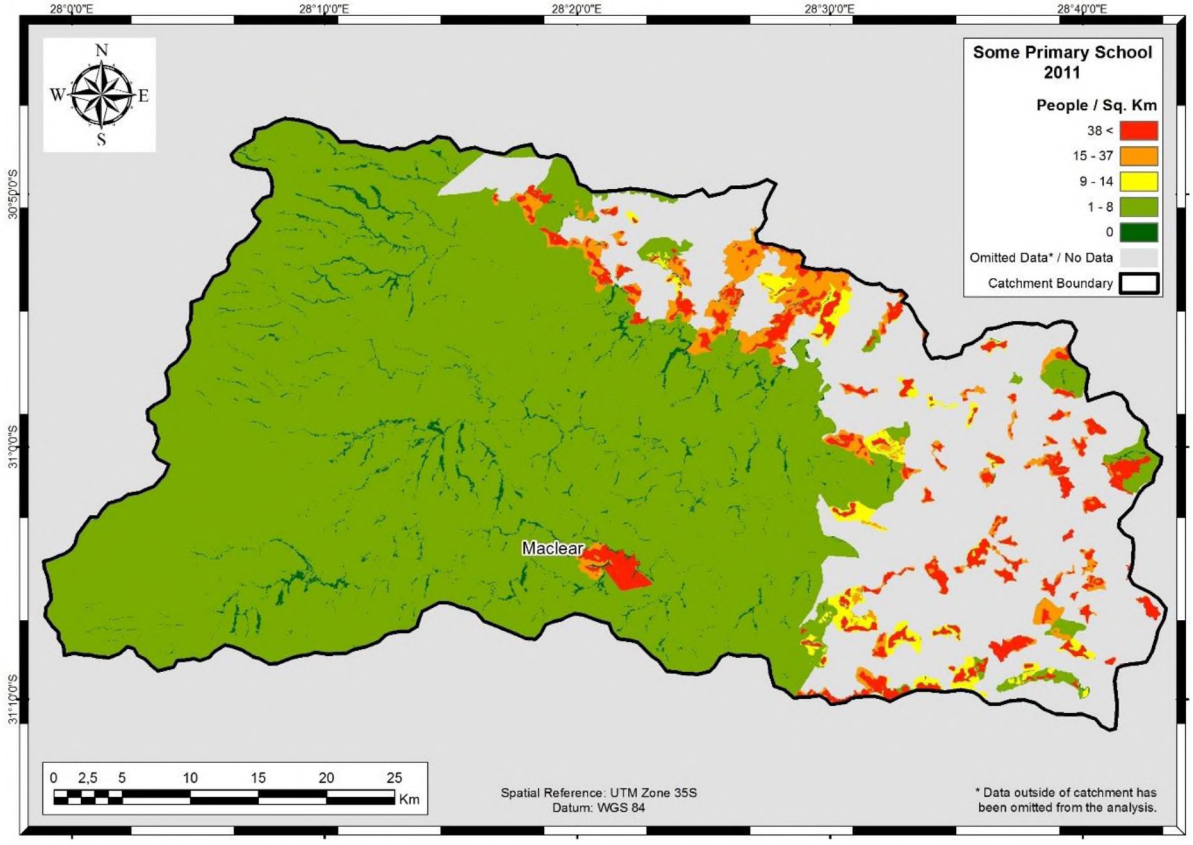


Figure 66: Persons with some primary schooling, 2011.

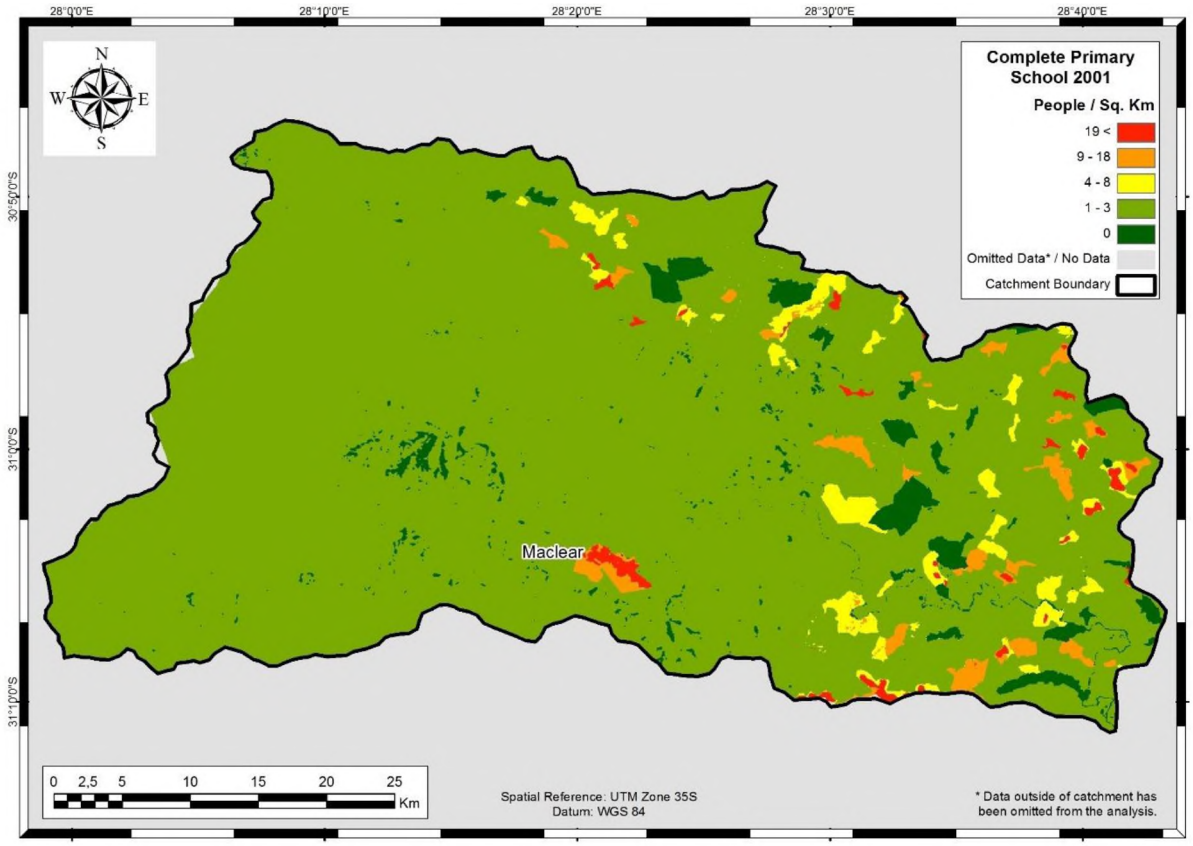


Figure 67: Persons that completed primary schooling, 2001.

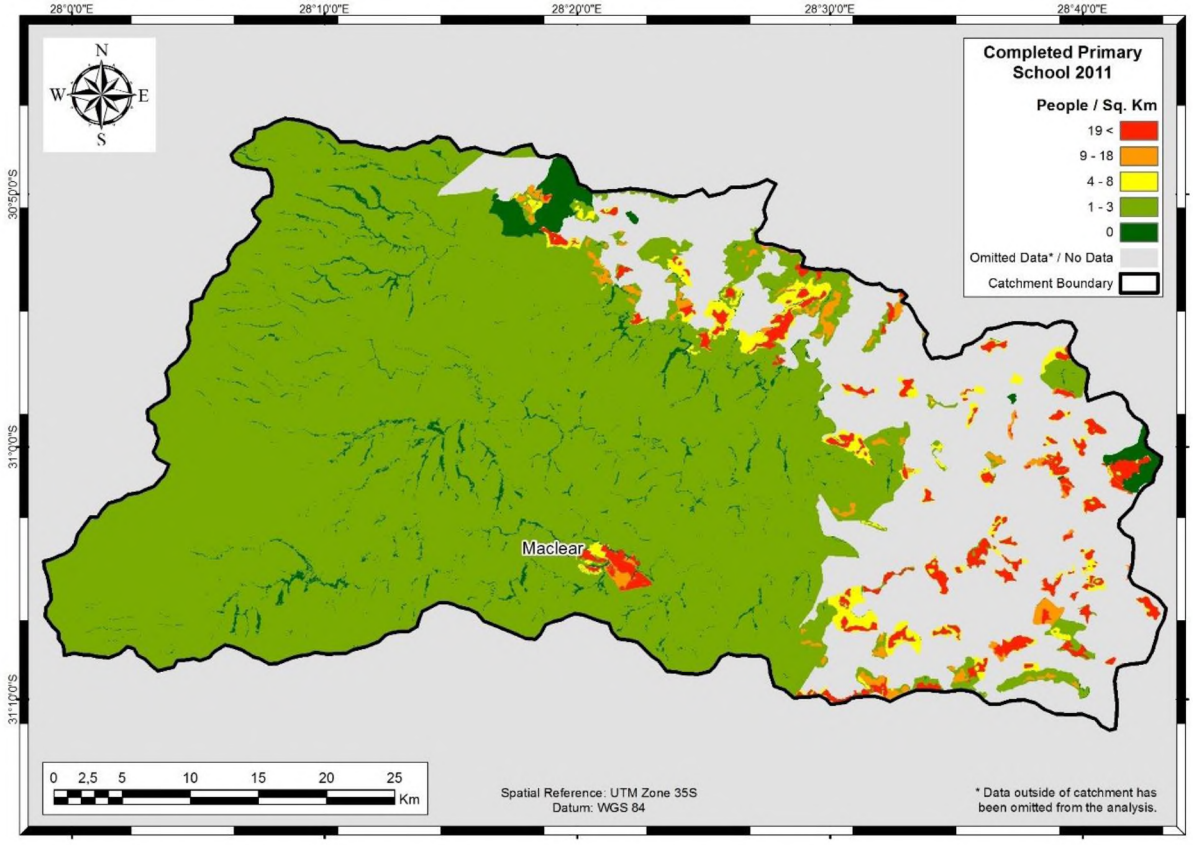


Figure 68: Persons that completed primary schooling, 2011.

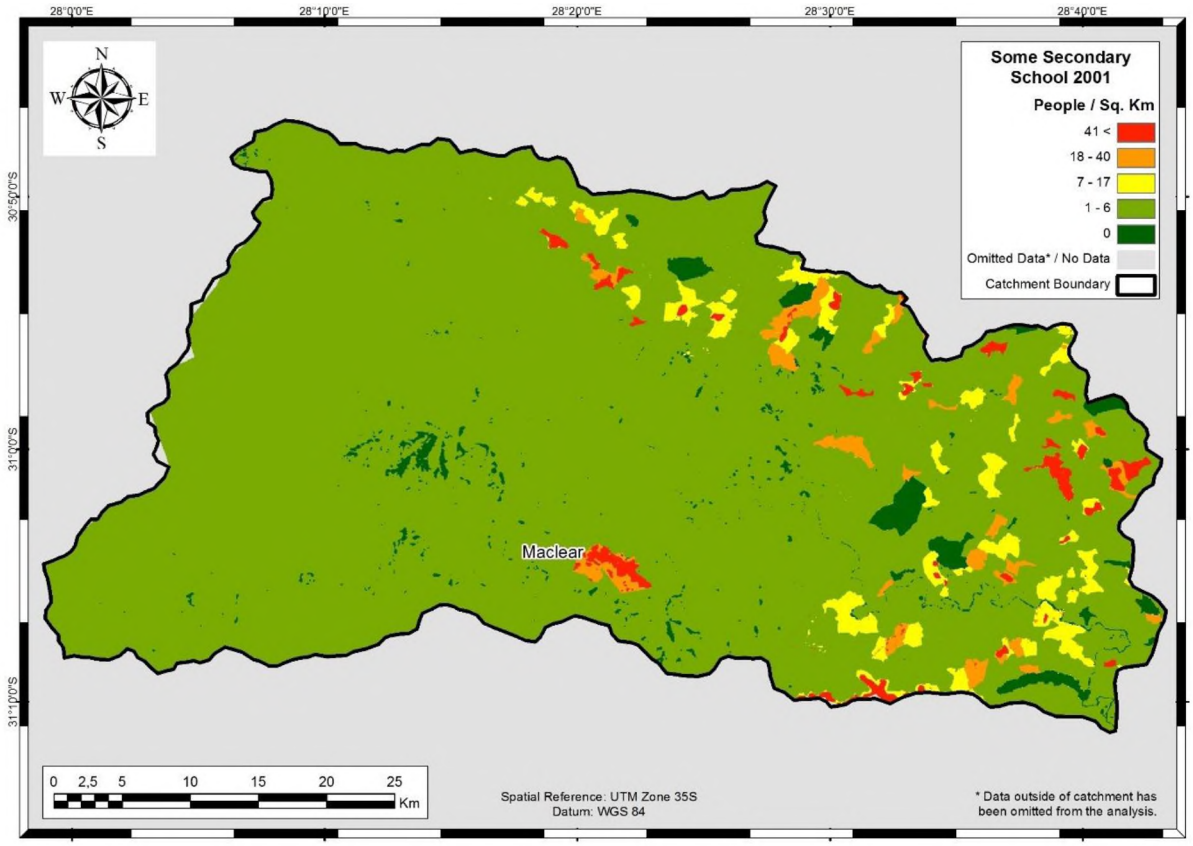


Figure 69: Persons that have some secondary schooling, 2001.

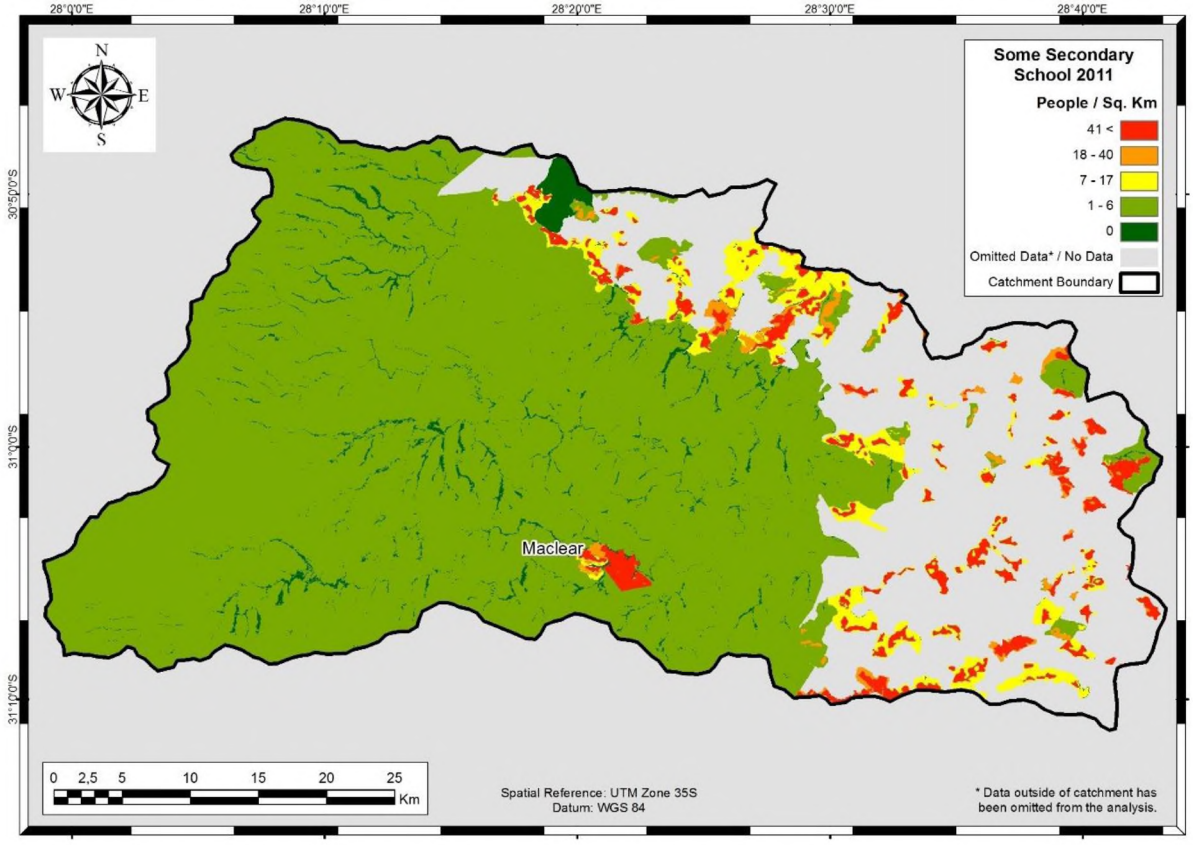


Figure 70: Persons that have some secondary schooling in 2011.

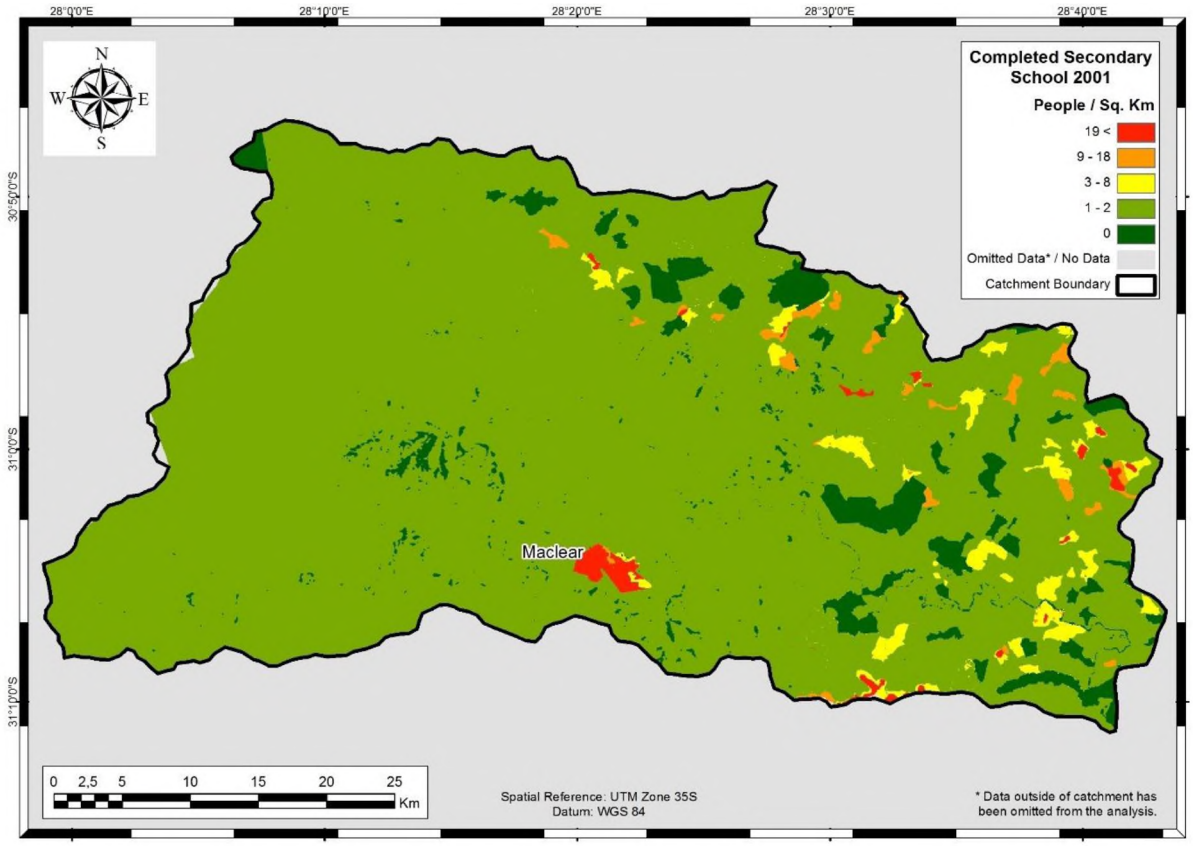


Figure 71: Persons that completed secondary schooling, 2001.

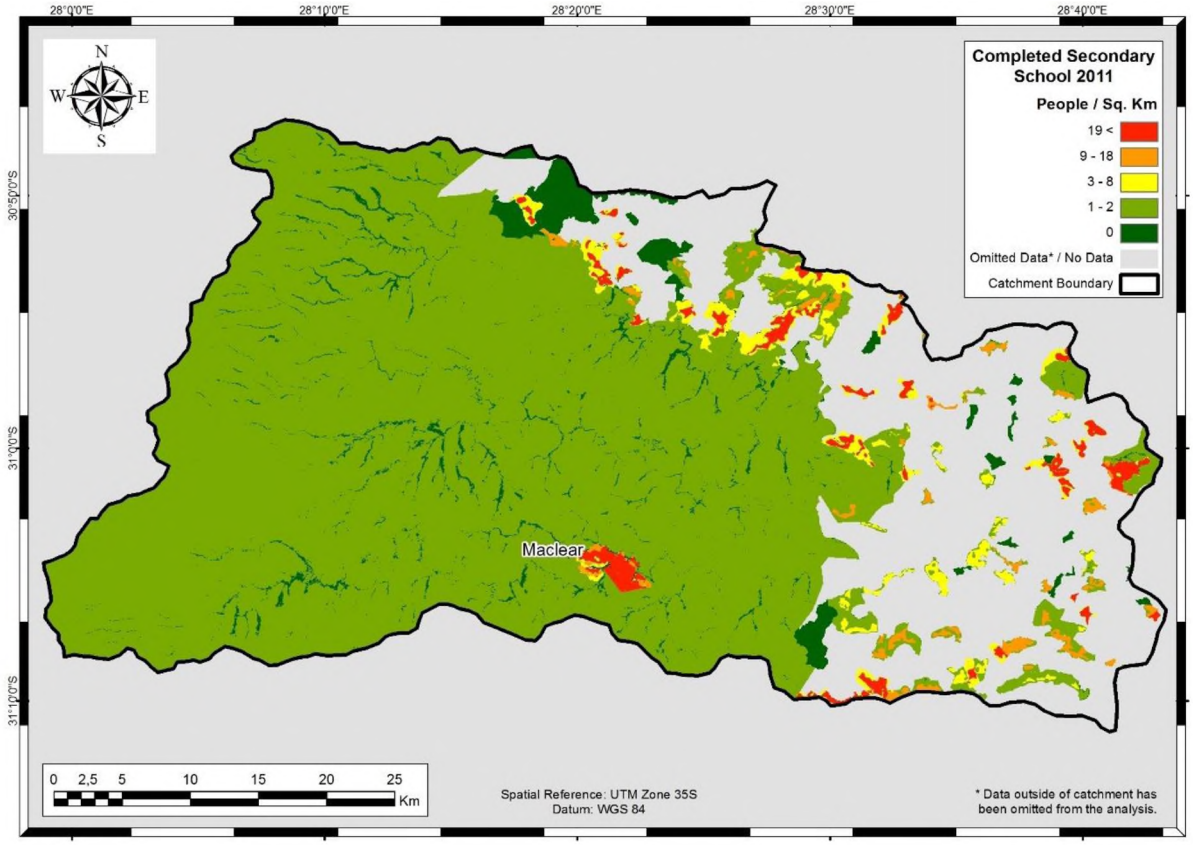


Figure 72: Persons that completed secondary schooling, 2011.

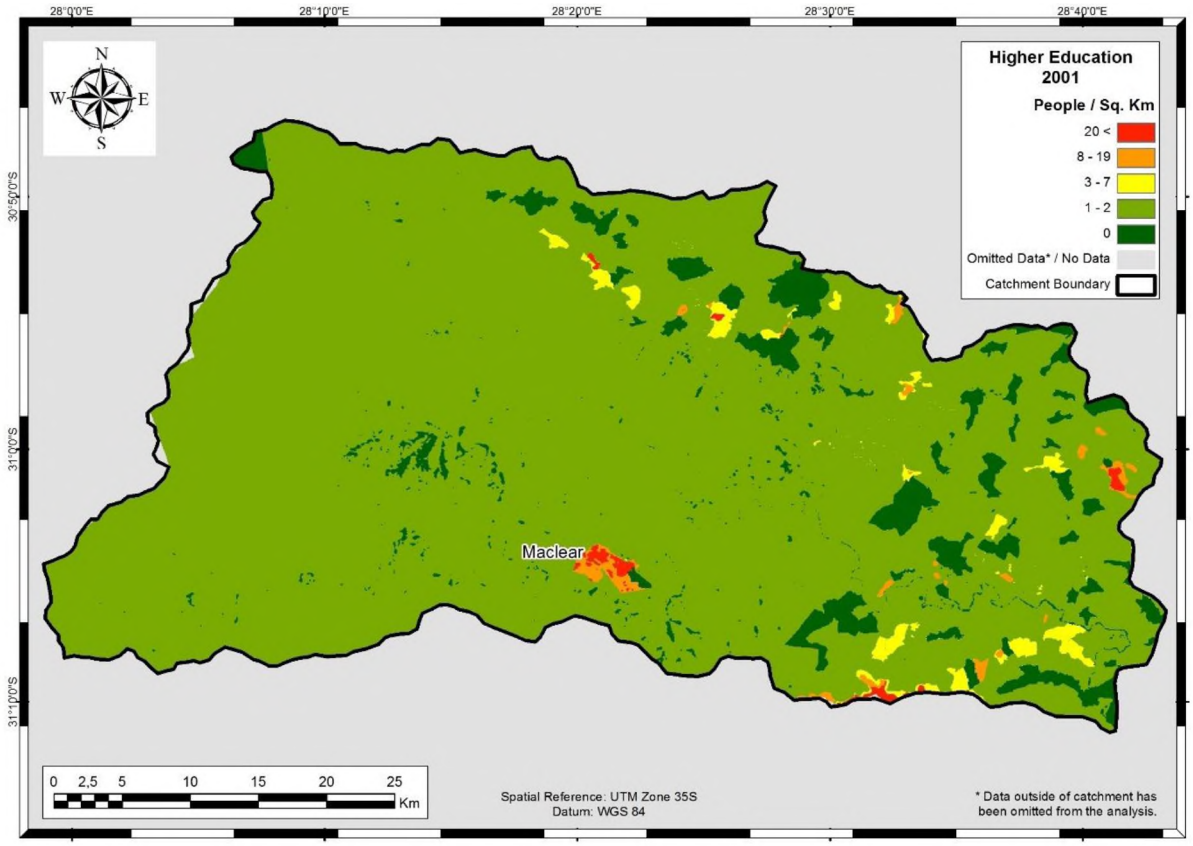


Figure 73: Persons with higher education, 2001.

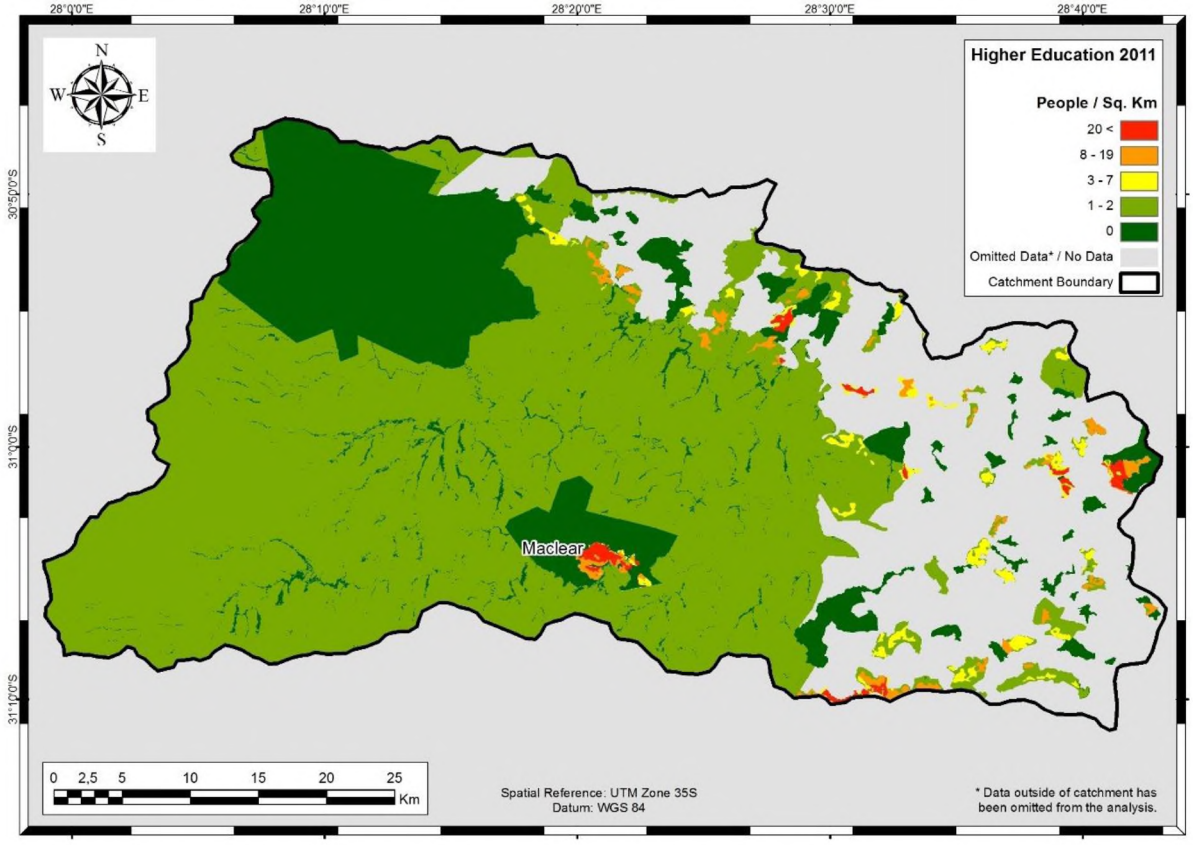


Figure 74: Persons with higher education, 2011.

5.6 Employment Status

This section examines the employment status in the Upper Tsitsa Catchment. There are four maps presented in this section. Employment refers to persons that are economically active, while unemployed are persons who are not employed.

The bar graph below (Figure 75) presents the changes in employment status between 2001 and 2011. Employment had increased by 427 (8%) persons since 2001 and unemployment has decreased by 2 571 (41%) persons. The decrease in unemployment could be due to people migrating out of the catchment for job opportunities and being employed by governmental initiatives such as: Working for Water, Working for Wetlands or private projects such as chicken farming as observed in the catchment.

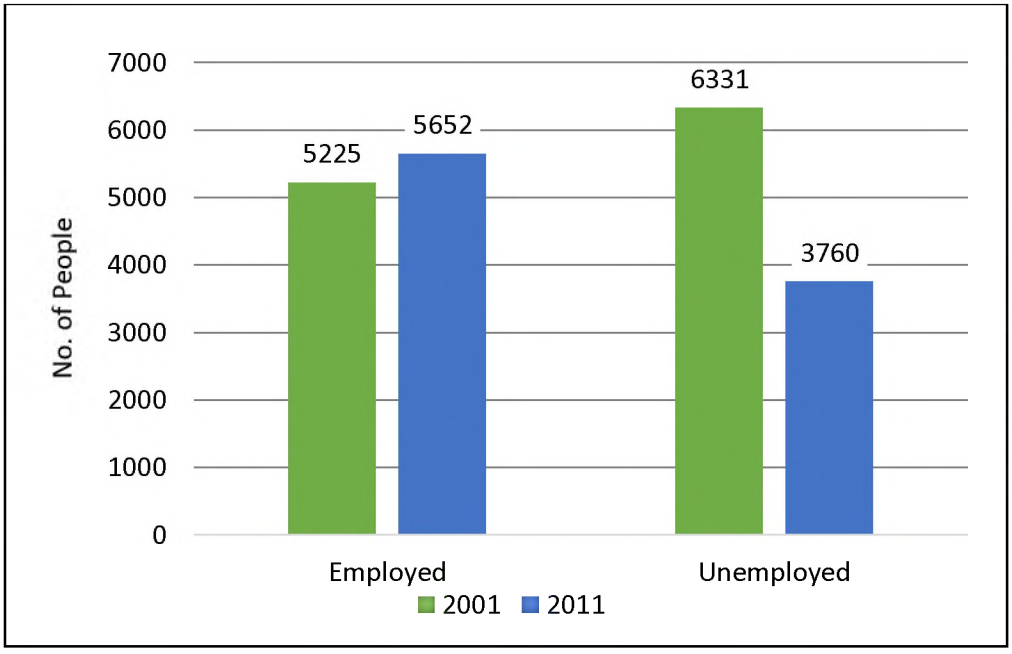


Figure 75: Employment status in the Upper Tsitsa Catchment.

In most of the rural villages in 2001, one to 18 people per square kilometer were employed (Figure 76). A couple of the villages, found to the east and north-east of the catchment had between 19 and 36 people employed per square kilometer. Two villages, Sulenkama and Sitishini, to the far east showed that it had up to 124 employed persons per square kilometer. This is the same village that was observed in Figure 74 with high density of persons with higher education. There is a police station and hospital in Sitishini, which provides employment and requires persons with higher education. The southern parts of Maclear had a density between 37 and 123 employed persons, while the northern region had 124 or more employed persons per square kilometer. In 2001, the rural villages' employment had increased to 19 to 123 persons being employed per square kilometer (Figure 77).

This could be due to the decrease in population in the catchment and governmental and private initiatives within the catchment that are creating employment opportunities for the local communities. The number of people that were employed in western Maclear had decreased to 19 to 123 persons per square kilometer, which may have been caused by out-migration as discusses in earlier sections. The township seemed to remain consistent with the number of employed persons since 2001.

The rural villages had between eight and 64 or more unemployed persons per square kilometer in 2011 (Figure 78). The northern areas in Maclear had 64 or more unemployed persons per square kilometer. The southern areas of Maclear had between 27 and 63 unemployed persons per square kilometer for 2001. Since 2001, unemployment had decreased in the rural areas to one to 63 unemployed persons and very few villages with more than 64 unemployed persons per square kilometer (Figure 79). The township of Maclear had a high density of unemployed persons but with more in the 27 to 63 unemployed person category than in the more than 64 unemployed person's category when compared to 2001. The western suburbs of Maclear had significantly decreased in unemployed persons, from 27 to 63 down to one to 26 unemployed persons per square kilometer. The general trend in decreased number of unemployed persons could be due to the out-migration of the people that lived in the catchment.

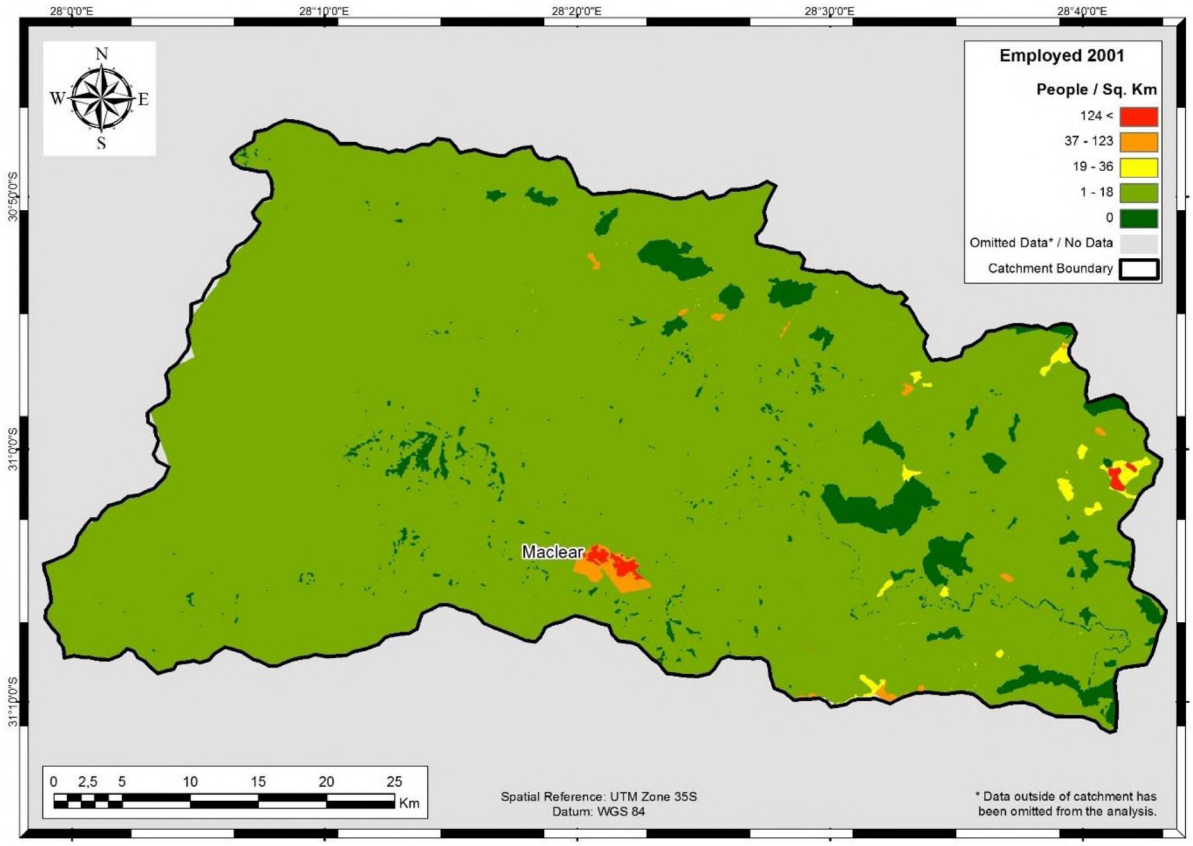


Figure 76: Persons who were employed in 2001.

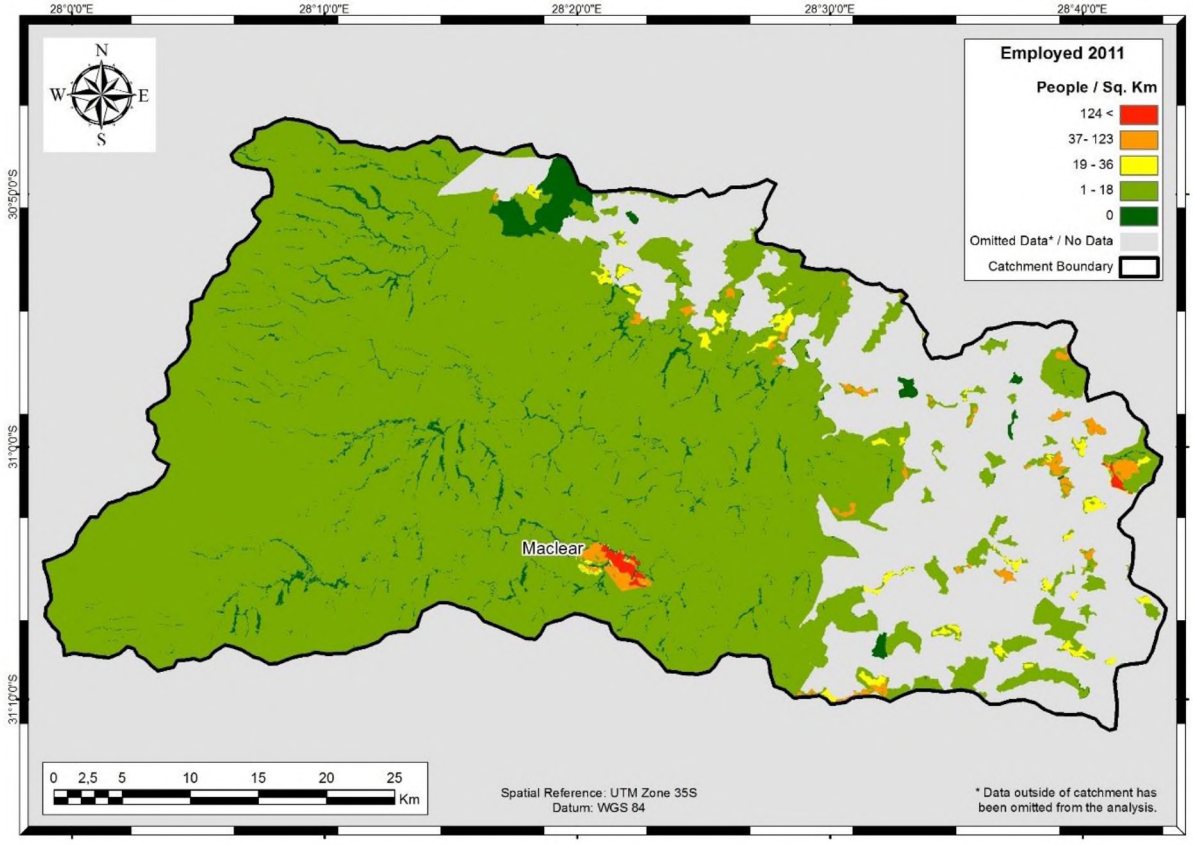


Figure 77: Persons who were employed in 2011.

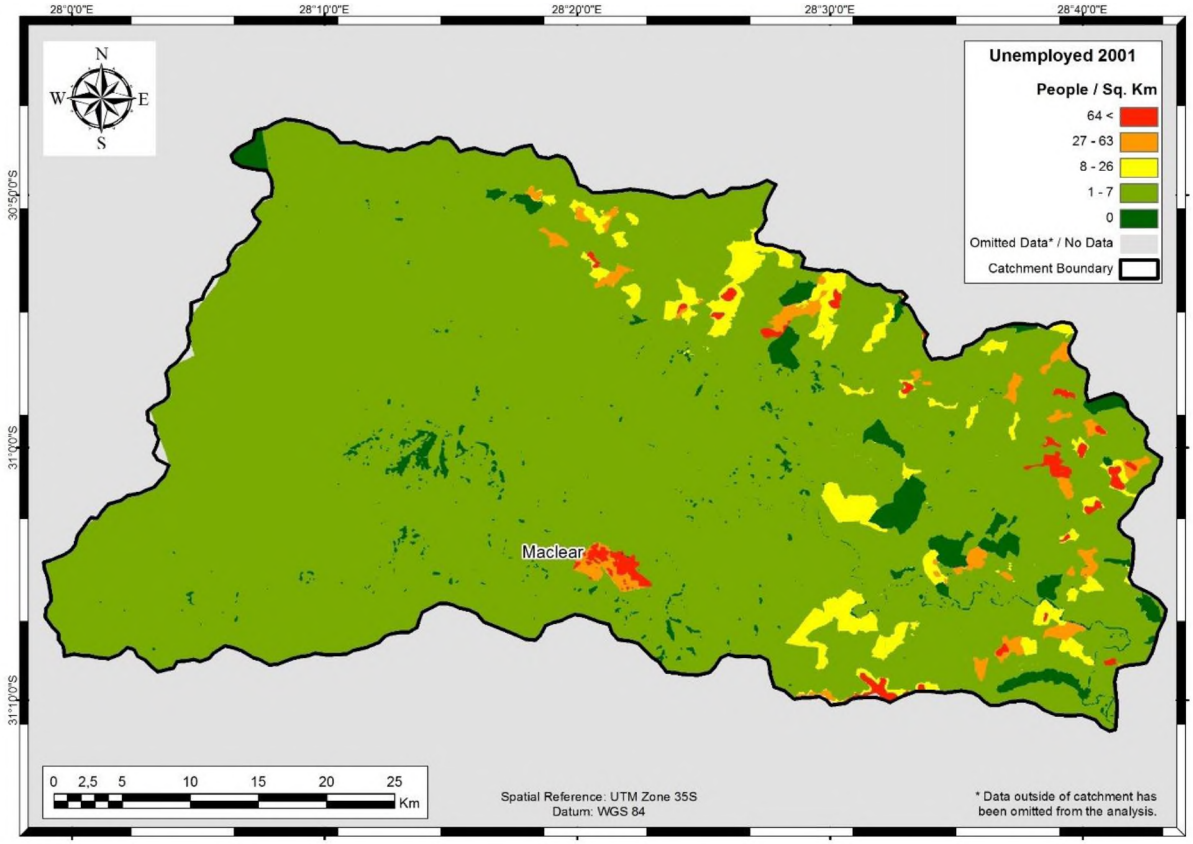


Figure 78: Persons who were unemployed in 2001.

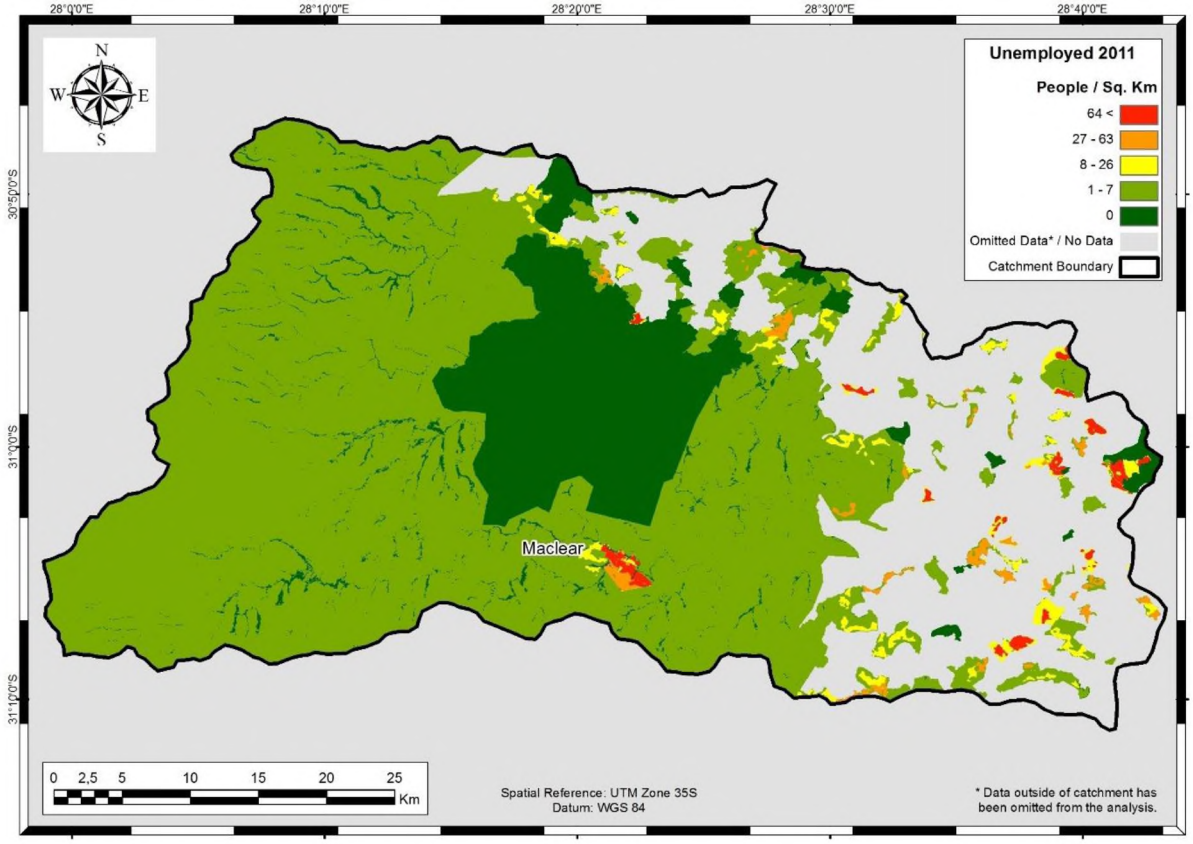


Figure 79: Persons who were unemployed in 2011.

5.7 Basic Services

This section examines the basic services available in the Upper Tsitsa Catchment. There are four maps in this section displaying data with regards to access to water and electricity. These data sets were mapped according to household use and not according to individuals as in the preceding sections.

From the bar graph below (Figure 80), a decrease in access to water was observed. This decrease of 1 120 (8%) households having access to water could be linked to population decrease or availability of water in the catchment or new dwellings that have not received any of the basic services. When in the field, it was noted that many of the communal taps were not functioning. Access to water considered the way a household was accessing water, be it from a river, borehole, communal tap or taps within the home. Access to electricity had increased by 4 637 (229%) households between 2001 and 2011. To determine access to electricity for 2011, 'electricity for lighting' was used in the 2011 census data, while the 2001 census only had one category for electricity.

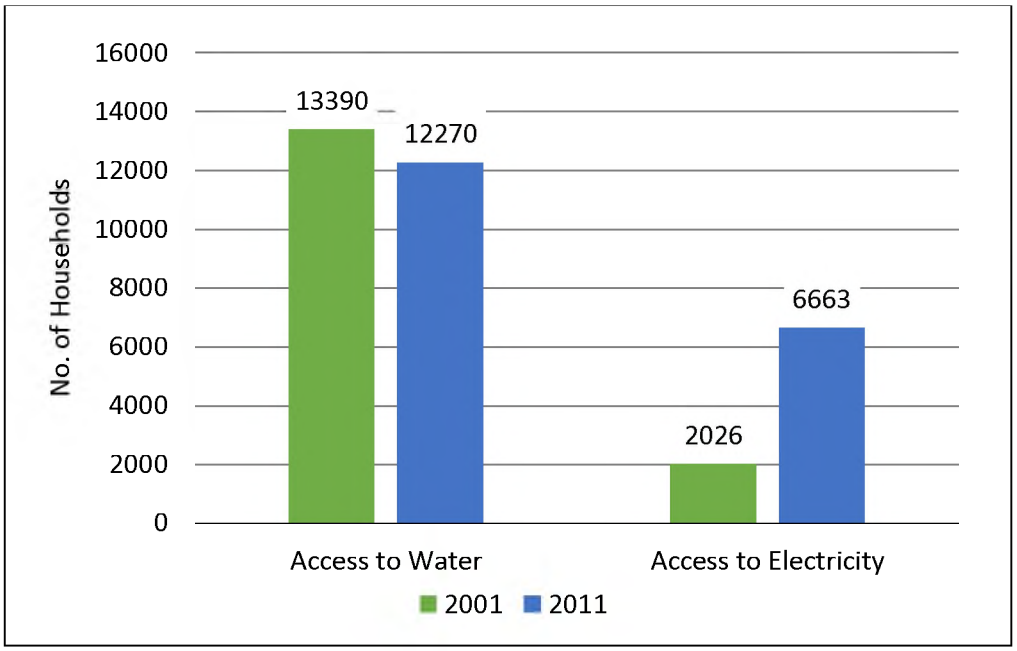


Figure 80: Basic service access in the Upper Tsitsa Catchment.

Between 13 and 73 or more households per square kilometer in the catchment had access to water (Figure 81). The northern areas of Maclear had 73 or more households with access to water while the southern areas of Maclear had 38 to 72 households per square kilometer with access to water. There was an increase from the 38 to 72 households within the rural villages to 73 or more households per square kilometer having access to water (Figure 82). The township of Maclear had 38 to more than 73 households per square kilometer with access to water while the western suburbs had decreased from 38 to 72 households down to 13 to 37 households per square kilometer having access to water.

Only two rural villages had access to electricity in 2001. The first village had a density of two to 10 household per square kilometer with access to electricity (Figure 83). The second village had two to 15 households per square kilometer with access to electricity. The northern areas of Maclear had mostly between two and 10 households while the southern areas of Maclear had more than 16 households per square kilometer with access to electricity. A change in access to electricity, where the rural villages which had none or one household in 2001, have moved into the category where two to 10 households have electricity (Figure 84). At the time the 2011 census was conducted, the northern villages had no access to electricity, but used alternative energies such as candles and paraffin. There had been a decrease in the number of households per square kilometer with access to electricity in western Maclear moving down into the category of two to 10 households per square kilometer. The decrease could be due to households changing to alternative energy sources as was seen in the field, where households were going off the grid and using paraffin and gas instead. The township seems to have decreased on its northern boarders to two to 10 households with access to electricity.

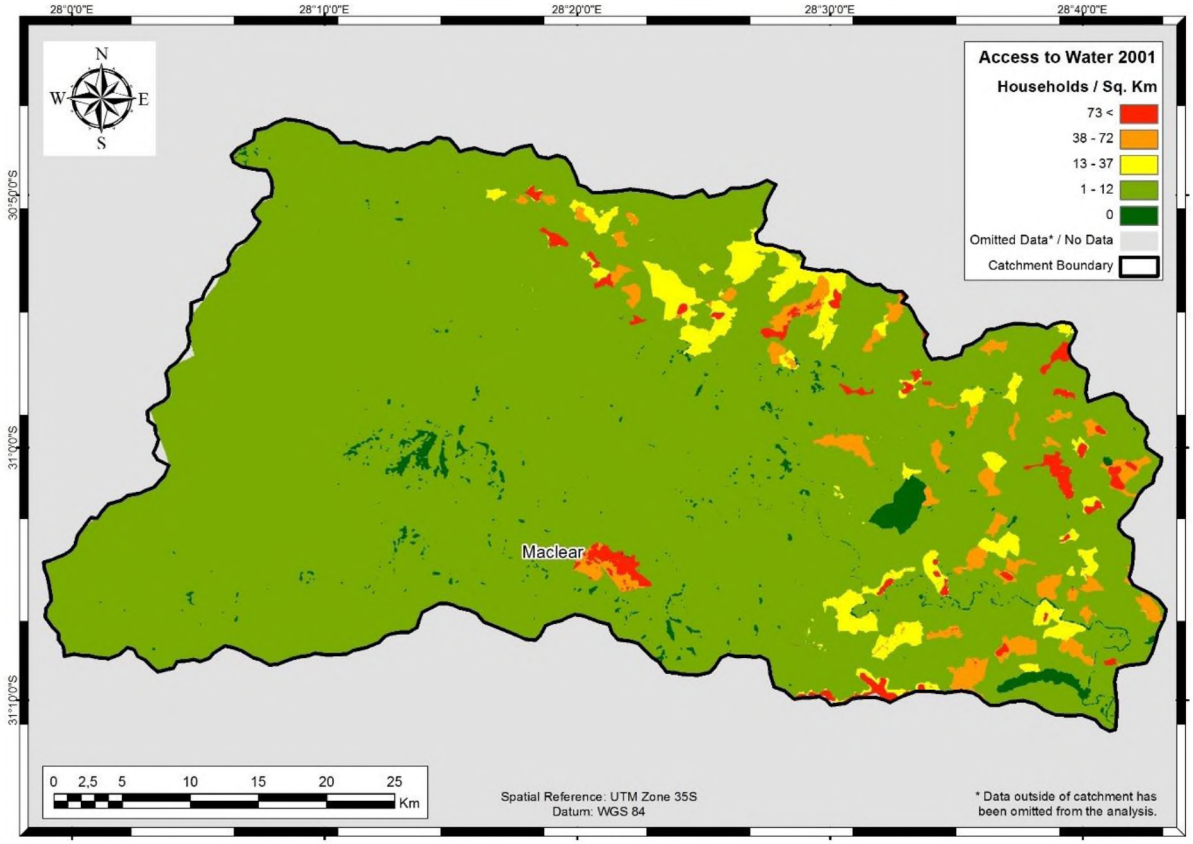


Figure 81: Households with access to water, 2001.

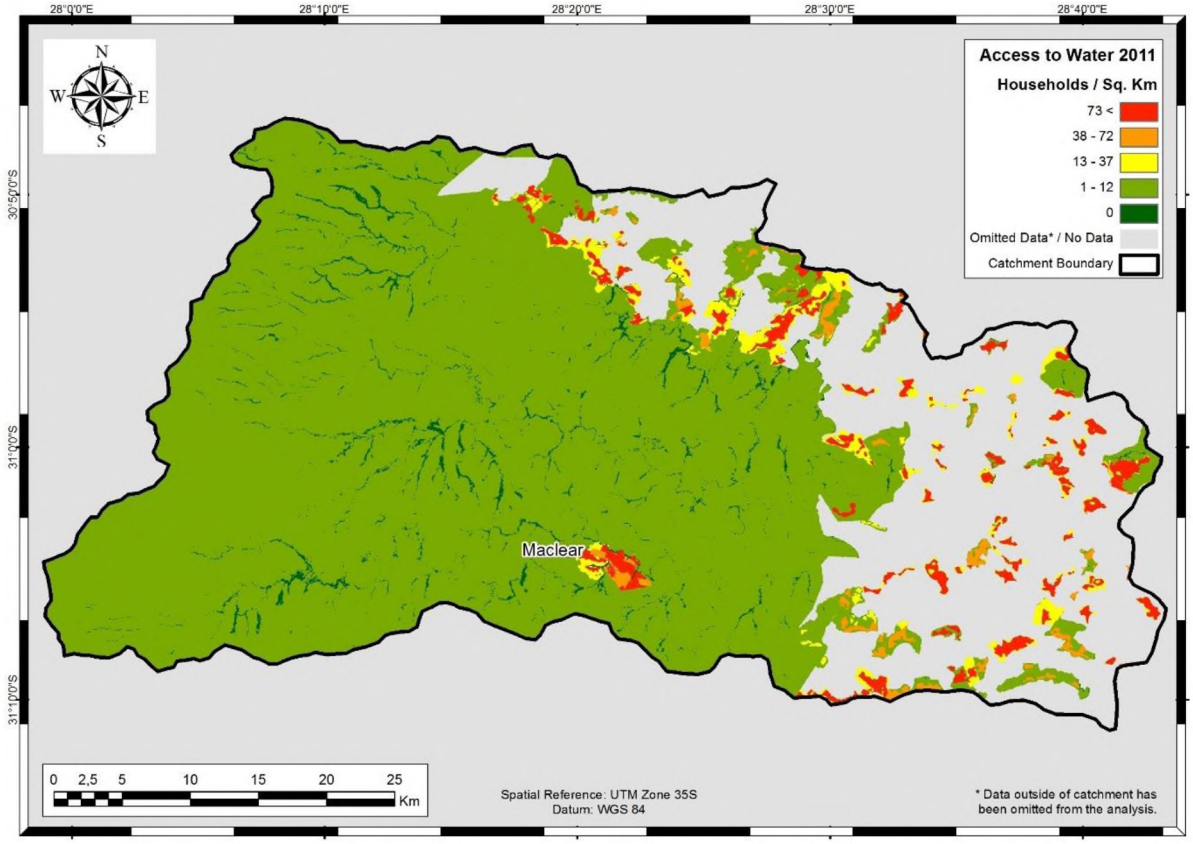


Figure 82: Households with access to water, 2011.

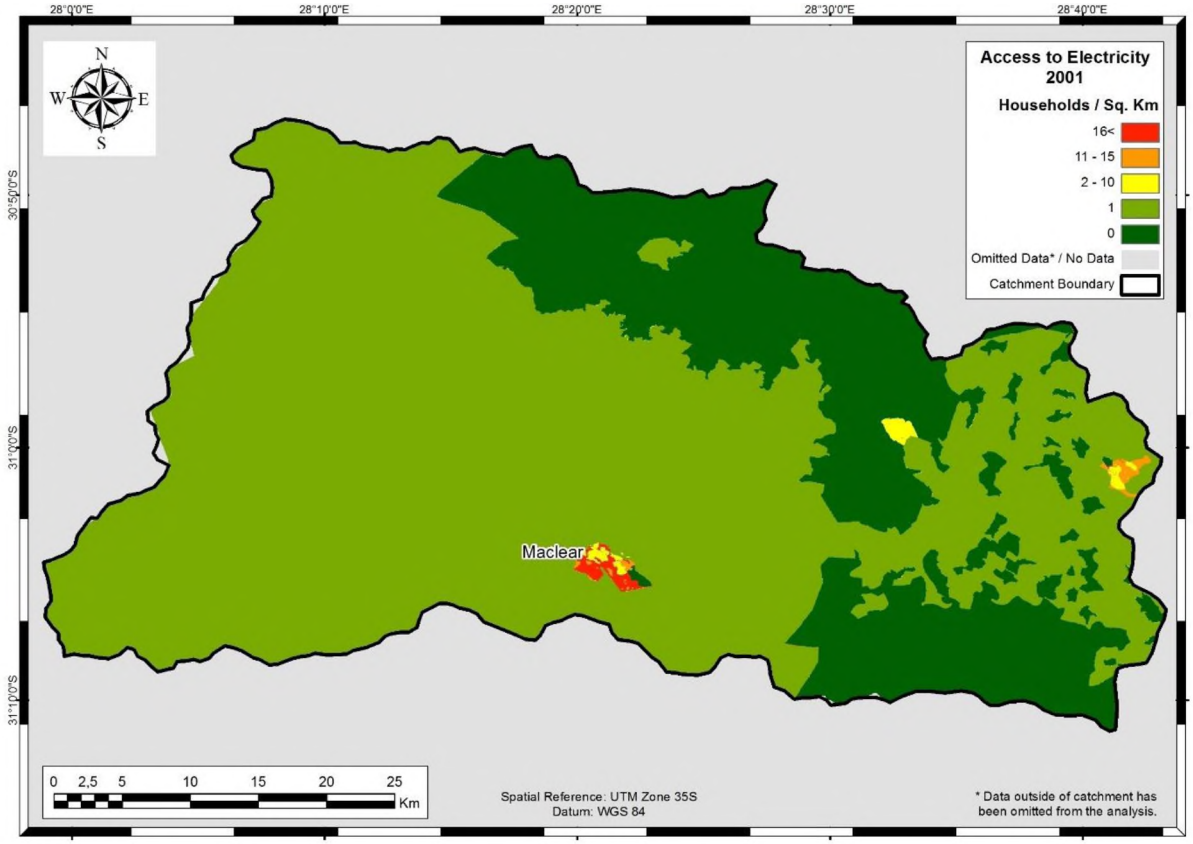


Figure 83: Households with access to electricity, 2001.

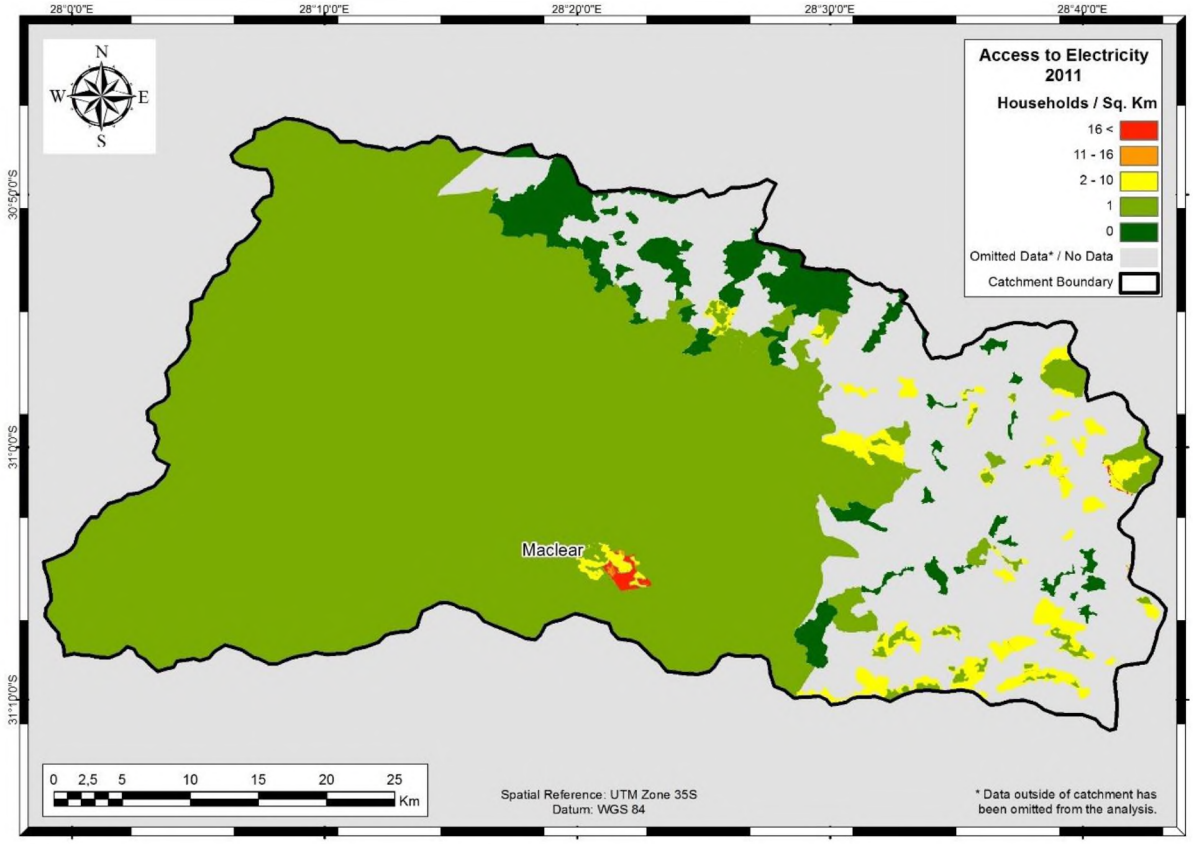


Figure 84: Households with access to electricity, 2011.

5.8 Conclusion

From the statistical results that were in the form of an error matrix, the overall accuracy was skewed due to most of the sample points lying in the no data and low categories. Manually collecting the data gave better insight on how the data sets differed in accuracy and the dasymetric data provided more detailed data than the choropleth data showing that dasymetric data provided more accurate data as shown in section 5.1.

From the results presented in sections 5.2 to 5.7, there has been significant decrease in the population present in the Upper Tsitsa Catchment. This trend can be seen not just in the population density, but in the population race groups, languages, employment and household services. Rural villages in the catchment became nucleated, suggesting people moved closer to each other as people moved to the urban centre, Maclear, or migrated out of the catchment. Although isiXhosa and Sesotho speakers decreased in the catchment, the number of English speakers increased. This was linked to the decrease in population, but also to the increase of persons having higher levels of schooling in the catchment. It was evident that persons with no schooling decreased in the study area, which could be due to people migrating out of the catchment and better access to schools. Persons with some primary schooling had doubled since 2001. School attendance has increased in the study areas from attending primary school to obtaining higher education. Individuals that were employed in the catchment increased, while unemployment decreased, which may be linked to people leaving the catchment. Households in the Upper Tsitsa Catchment decreased with regards to access to water, which may be linked to the population density decrease. However, households with access to electricity in the catchment substantially increased.

Chapter Six: Discussion

The following chapter discusses the reliability of dasymetric maps to choropleth maps that were presented in Chapter five.

Overall, this study found that integrating census data with appropriate, high resolution land cover data provided accurate and useful data to determine spatial and temporal changes over time. Thirty-six dasymetric maps with a spatial resolution of 30 m x 30 m were created for the Upper Tsitsa Catchment and formed part of the data for the DEA-NRM socio-economic database for NLEIP. By eliminating the MAUP from the data, more significant results were displayed on the map, compared to the homogenous data of choropleth maps. Using the NLC data sets also ensures that the results of the dasymetric maps can be compared to surrounding regions. The integration of the census and land cover data has allowed for statistical and visual interpretation of spatial and temporal patterns of the 18 census variables mapped for the Upper Tsitsa Catchment. These statistics, data and the results on spatial and temporal changes presented in Chapter five, can aid in decision making by the NLEIP managers with regards to socio-economic, ecological and environmental policies and development plans, by providing insight into what human resource are available and where these resources are.

From the census data sets 142 areas for the 2001 census and 161 areas from the 2011 census were extracted for the Upper Tsitsa Catchment. Due to the municipal boundary changes, it was decided that dasymetric mapping technique would be used to remove the choropleth boundaries and produce an accurate distribution of the census data to determine socio-economic trends within the catchment. According to Holt *et. al.* (2004), census data is often reported using the enumeration areas, which is designed for data collection and has no meaning when representing or determining changes over time. Usually, census data are mapped using choropleth mapping techniques. The most common way to determine the population density is by dividing the population by the area of the small areas, then presenting these results as choropleth maps (Holt *et. al.,* 2004). Using the choropleth mapping method presents several problems. First, the data is homogenously distributed within the small areas as described by Dorling (1993). There are abrupt changes at the boundaries of the small areas, while the data represents a continuous variable across the area which was identified by Holt *et. al.,* 2004 and Sleeter 2004.

To overcome the boundary changes, dasymetric mapping was used. This technique integrated the census data with the NLC data for the catchment, which accurately predicted the population densities for the catchment (Stevens *et.al.*, 2015). Holt *et. al.* (2004), noted that dasymetric maps reflected the underlying geography of the area, compared to choropleth maps which only displayed the statistical

data about the population with regard the enumeration boundary. The results of these variables determined where people with these characteristics were in the catchment, unlike choropleth maps which reported the data as continuous statistical data, with abrupt changes at the enumeration boundaries.

Following the adopted version of Mennis', created by Duh (2005), which uses the Grid Three-class dasymetric method, results were produced illustrating the density of persons in the catchment depicting several variables. The adapted version used five classes instead of three. Mennis (2009) stated that this technique could be applied to any number of classes. The five classes that were described in Table 4, Chapter four, suited the Upper Tsitsa Catchment as these were the dominant land covers in the catchment. The data were created in raster format, therefore the colour of each of the pixels indicated the number of persons within the 30 m by 30 m pixel. The resolution of the dasymetric map data were determined by the landcover data, which had a high spatial resolution of 30 m by 30 m. Certain areas within the enumeration boundary were populated while other areas were not (Wright, 1936). This allows dasymetric maps to produce maps that display a better visualisation of the data set as seen in Chapter 5.

The statistical analysis in the form of an error matrix determined the accuracy of the dasymetric mapping method. The high total accuracy of 82% is largely the result of the majority if the sampling points falling within the 'no data' and 'low' categories. Where the sample points for the choropleth maps were 'medium' or 'high', they were low on the dasymetric maps. Only three of the 2000 sample points were correct for the 'very high' category. This was due to the enumeration areas not covering the entire catchment for the 'no data' and the large commercial agricultural sector had low population densities. From working with the data while completing the error matrix, showed how the data points within the same polygon could have a variety of categories, where the choropleth maps only show one category.

By using dasymetric maps, changes and movement of the population was visually interpreted. Using the high spatial resolution satellite imagery in the form of land cover allowed the output data to produce more appropriate visualisations of the population, whereby spatial and temporal patterns could be identified, as stated by Holt *et. al.* (2004). The pronounced decrease in individuals within the catchment could be identified. The colour of the pixels determined the density of the people, while the shape of high density areas determined how individuals had moved in the area. The size of the high-density areas determined if population had increased or decreased as well as demonstrated nucleation of persons in the catchment. The combination of the colour, shape and size of population density categories, provides an accurate visual representation of data on the map compared to

choropleth maps (Sleeter, 2004). Besides population density, other variables in the census data such as language, race and education were mapped using the same principles (Chapter five). Bajat *et. al.* (2001) and Mennis (2003) found that dasymetric mapping provided more realistic representations of the population's distribution and provided a more detailed and accurate population density.

Chapter Seven: Conclusion

7.1 Research Overview

This thesis evaluated the accuracy and usefulness of dasymetric maps in determining spatial and temporal trends for the Upper Tsitsa Catchment using the 2001 and 2011 census data integrated with corresponding National Land Cover data. Using dasymetric mapping allowed the census to be comparable over time by removing the small area layer boundaries. Detailed and accurate demographic and socio-economic data about the Upper Tsitsa population was created. Patterns of accuracy and detail of dasymetric maps that were shown from the results aligned with the results of the authors Bajat *et. al.* (2001) and Mennis (2003). Spatial statistics conducted presented the accuracy of the dasymetric data to that of the choropleth data. Dasymetric data had finer details to that of choropleth maps. The extraction of the census data for objective one and creation of dasymetric maps following the eleven steps described in section 4.3 was a timely exercise. None of the literature on dasymetric mapping stated that the process was time consuming. Creating dasymetric data sets are worth the time as it provides more detailed data about the study area. This data will aid in decision-making when implementing various polices and development plans within the study area. It has been concluded that dasymetric mapping is useful in determining accurate and reliable spatial and temporal trends in the Upper Tsitsa Catchment.

7.2 Research Limitations

Some of the limitations faced for this research with the limited availability of instructions on the various dasymetric techniques such as the limiting variable method. Time was another limitation, as some data such as income requires standardisation before it can be mapped and compared and was therefore left out. Census questionnaire for 2001 was more limited than for 2011, therefore not all variables or more detailed data could be mapped e.g. Types of water access.

7.3 Recommendations for Potential Future Studies

First, this technique can be applied to the neighbouring regions, where projects such as the Laleni dam will be build and restoration will continue. Second, creating dasymetric map data with regards to financial economics for the Upper Tsitsa Catchment may be beneficial as well as in neighbouring catchments. To map individual and household income, the data needs to be manipulated so that the monetary values between 2001 and 2011 can equate. Third, community survey data could be used to create intermediate data as opposed to of having a ten-year interval. Lastly future censuses can be

mapped and compared to this data to determine if there have been any changes in the catchment from the NLEIP with regards to economic upliftment.

Reference List

Bajat, B., Krunić, N. and Kilibarda, M., 2011. Dasymetric mapping of spatial distribution of population in Timok Region. In Proceedings *of international scientific conference professional practice and education in geodesy and related fields*, Kladovo, Serbia. P. 30-34.

Blake, M., Bell, M. and Rees, P. 2000. Creating a Temporally Consistent Spatial Framework for the Analysis of Inter-Regional Migration in Australia. *International Journal of Population Geography.* 6: 155-174.

Bracken, I. 1993. An Extensive Surface Model Database for Population-Related Information: Concept and Application. *Environment and Planning B: Planning and Design*. 29: 13-27.

Calmeyer, T. and Muruven. L. 2014. Environmental Impact Assessment for the Mzimvubu Water Project. Scoping Report. *Directorate Options Analysis.*

Chief Directorate: National Geo-Spatial Information (CD:NGI). 2009. Development Methods of a Methodology for National Land Use Mapping for Chief Directorate: Geo-Spatial Information. Report.

Chief Directorate: National Geo-Spatial Information (CD:NGI). 2013. Aerial Imagery. [Online]. Available: http://geog-mappa-gnas1.storage.ru.ac.za/scripts/body_ftp02.php. [06/06/2015].

Colman, D. and Salt. J. 1992. *The British Population.* Oxford: Oxford University Press.

Congalton, R.G. 1991. A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data. *Remote Sensing Environment*. 37: 35-46.

CSIR & ARC (2005). *National Land-cover Database 2000.* Council for Scientific and Industrial Research and the Agricultural Research Council.

Di Gregorio. A. and Jansen, L. 2000. Land Cover Classification System (LCCS): Classification Concepts and User manual for software version 1.0. *Food and Agriculture Organisation of the United Nations*. Rome.

Dorling, D. 1993. Map Designs for Census Mapping. *Cartographic Journal.* 30: 167-183.

Duh, G. 2005. *Lab 4: Dasymetric Mapping.* [Online]. Available: http://web.pdx.edu/~jduh/courses/geog492w11/lab04.html. [31/05/2016].

Editors of the Encyclopaedia Britannica. n.d. *Cape Province.* [Online]. Available: http://www.britannica.com/place/Cape-Province. [31/05/2017].

Eicher, C.L. and Brewer, C.A. 2003. Dasymetric Mapping and Areal Interpolation: Implementation and Evaluation. *Cartography and Geographic Information Systems.* 28(2): 125-138.

Elundini Local Municipality. 2017a. *Tourism Elundini.* [Online]. Available: http://elundini.org.za/index.php/tourism2. [31/05/2017].

Elundini Local Municipality. 2017b. *About Us.* [Online]. Available: http://elundini.org.za/. [31/05/2017].

Fairbanks, D.H.K, Thompson, M.W., Vink, D.E., Newby, T.S., van den Berg, H.M., Everard, D.A. 2000. The South African Land Cover Characteristics Database: A Synopsis of the Landscape. *South African Journal of Science*. 96: 69-82.

Foody, G.M. 2000. Status of Land Cover Classification Accuracy Assessment. *Remote sensing of Environment.* 80:185-201.

GeoTerraImage: 2015. 2013-2014 South African National Land-Cover. Report. *A Commercial Data Product Prepared for Department of Environmental Affairs.*

Giannecchini, M., Twine, W. and Vogel, C. 2007. Land-Cover Change and Human-Environment Interaction in a Rural Cultural Landscape in South Africa. *The Geographic Journal.* 173(1): 26-42.

Goodman, L. and Medani, S. 2016. *Inside South Africa – Municipality Analysis.* [Online]. Available: http://www.marketdecisions.co.za/south_africa/municipal.php. [31/05/2017].

Gregory, I.N and Ell, P.S. 2005. Breaking the Boundaries: Geographical Approaches to Integrating 200 Years of the Census. *Journal of the Royal Statistical Society.* 168 (2): 419-437.

Hamann, M. and Tuinder, V. 2012. *Introducing the Eastern Cape: A Quick Guide to Its History, Diversity and Future Challenges.* Stockholm University: Stockholm Resilience Centre.

Holloway, S., Schumacher,J. and Redmond, R. 1999. Chapter 6: People and Places. Dasymetric Mapping using Arc/Info. *In* Morain, S.A. (ed.), *GIS Solutions in Natural Resource Management: Balancing the Technical-Political Equation.* Santa Fe: OnWord Press. P. 238 - 291.

Holt, J.B., Lo, C.P and Hodler, T.W. 2004. Dasymetric Estimations of Population Density and Areal Interpolation of Census Data. *Cartography and Geographic Information Systems*. 31(2): 103.121.

Hubald, M.H. and Di Gregorio, A. 2009. Translating and Evaluating Land Cover Legend using UN Land Cover Classification System (LCCS). Report No. 43 *Global Observation of Forest and Land Cover Dynamics (GOFC GOLD).*

Huby, M., Own, A and Cinderby, S. 2007. Reconciling Socio-Economic and Environmental Data in a GIS Context: An Example from Rural England. *Applied Geography*. 27: 1-13.

Kpedekpo, G.M.K. 1982. *Essentials of Demographic Analysis for Africa.* London: Heinemann Educational Books Ltd.

Lambin, E.F., Geist, H. and Lepers, E. 2003. Dynamic Land Use and Cover Change in Tropical Regions. *Annual Review of Environment and Resources*. 28: 205-241.

Langford, M. 2003. Refining Methods for Dasymetric Mapping using Satellite Remote Sensing. In Mesev, V. (ed). *Remotely Sensed Cities*. New York: Taylor and Francis. 137-156.

Lemon, A. 1976. *Apartheid: A geography of Separation.* Westmead: Saxon House.

Mann, D. and Chandra, O. B. 2013. *Creating a Gridded Population Data of India: A Comparison and Evaluation Study of Dasymetric Mapping Methods.* ESRI India 14th User Conference 2013.

Mans, G. 2011, Developing a Geo-Data Frame using Dasymetric Mapping Principles to Facilitate Data Integration. In *AfriGeo Conference: Developing Geomatics for Africa*.

Martin, D. and Bracken, I. 1993. The Integration of Socioeconomic and Physical Resource Data for Applied Land Management Information Systems. *Applied Geography*. 13: 45-53.

Martino, D. and Zommers, Z. 2007. *Global Environmental Outlook*. [Online]. Available: http://web.unep.org/geo/assessments/global-assessments/global-environment-outlook-4. [17/10/2016].

Mayaux, P., Bartholomé, E., Fritz, S. Belward, A. 2004. A New Land-Cover Map of Africa for the Year 2000. *Journal of Biogeography.* 31: 861-877.

Mennis, J. 2003. Generating Surface models of Population Using Dasymetric Mapping. *The Professional Geographer.* 55(1): 31-42.

Mennis, J. 2009. Dasymetric Mapping for Estimating Population in Small Areas. *Geography Compass*. 3(2): 727-745.

Mhlontlo Local Municipality. 2017. *Introduction to Mhlontlo Local Municipality.* [Online]. Available: http://www.mhlontlom.gov.za/about. [31/05/2017].

Munger, E.S. 1962. *American Universities Field Staff Report Service.* Central and Southern Africa Series vol. X no. 3.

Municipal Demarcation Board (MDB). 2010. *About Demarcation.* [Online]. Available: http://warwickchapman.com/wp-content/uploads/2010/05/Municipal-Demarcation-Board-Website-About.pdf. [08/04/2017].

Municipal Demarcation Board (MDB). 2013a. *About Demarcation.* [Online]. Available: http://www.demarcation.org.za/site/about-demarcation/. [08/04/2017].

Municipal Demarcation Board (MDB). 2013b. *Mandate.* [Online]. Available: http://www.demarcation.org.za/site/mandate/. [08/04/2017].

Municipal Demarcation Board (MDB). 2013c. *Legislation Mandate.* [Online]. Available: http://www.demarcation.org.za/site/legislation/. [08/04/2017].

National Treasury. 2011. *The Local Government Budget and Expenditure Review: 2006/07 – 2012/13.* Cape Town: FormeSet Printers.

Nel, E. and Binns, T. 2000. Rural Self-Reliance Strategies in South Africa: Community Initiatives and External Support in the Former Black Homelands. *Journal of Rural Studies.* 16(2000): 367-377.

Ngcofe, L. and Thompson, M. 2015. The Status of Land Cover Mapping in South Africa: 1994-2015. *PositionIT.*11:49-52.

Openshaw, S. 1983. *The Modifiable Areal Unit Problem.* Concepts and Techniques in Modern Geography, vol 38. Norwich: Geobooks.

Ruiters, G. 2011. *Inventing Provinces: Situating the Eastern Cape.* In: Ruiters, G. The Fate of the Eastern Cape: History, Politics and Social Policy. Scottsville: University of KwaZulu-Natal Press. 19-41.

Schoeman, F., Newby, T.S., Thompson, M.W., van den Berg, E.C. 2013. South Africa Natinal Land-Cover Change Map. *South African Journal of Geomatics.* 2(2): 94-105.

Sleeter, R. 2004. Dasymetric Mapping Techniques for the San Francisco Bay Region, California. *In Urban and Regional Information Systems Association Annual Conference Proceedings.* 7-10.

Slocum, T.A. and Egbert, S.L. 1993. Knowledge Acquisition from Choropleth Maps. *Cartography and Geographic Information Systems.* 20(2): 83-95.

South African History Online. 2011. *Census in South Africa.* [Online]. Available: http://www.sahistory.org.za/print/social-histories/census-south-africa. [26/03/2015].

Southall, R. 1983. *South Africa's Transkei: The Political Economy of an "Independent" Bantustan.* New York: Monthly Review Press.

Statistics South Africa (Stats SA). 2002. CD-ROM. SuperCross Community Profiles 2001. Pretoria.

Statistics South Africa (Stats SA). 2012a. *Census 2011 Metadata*. Pretoria: Stats SA.

Statistics South Africa (Stats SA). 2012b. *Census 2011 Statistical Release – P0301.4*. Pretoria: Stats SA.

Statistics South Africa (Stats SA). 2012c. *About Statistics South Africa.* [Online]. Available: http://mapserver2.statssa.gov.za/geographywebsite/about.html#geoFrame. [17/11/2016].

Statistics South Africa (Stats SA). 2012. CD-ROM. SuperCross Community Profiles 2011. Pretoria.

Statistics South Africa (Stats SA). 2015. *Introduction.* [Online]. Available: http://www.statssa.gov.za/census2011/intro.asp. [26/03/2015].

Statistics South Africa (Stats SA). 2016. *2001 Small Area Statistics.* [Online]. Available: http://www.statssa.gov.z?/?page_id=4086. [23/10/2016].

Statistics South Africa (Stats SA). 2017a. *Statistics by Place: Elundini.* [Online]. Available: http://www.statssa.gov.za/?page_id=993&id=elundini-municipality. [08/04/2017].

Statistics South Africa (Stats SA). 2017b. *Statistics by Place: Mhlontlo*. [Online]. Available: http://www.statssa.gov.za/?page_id=993&id=mhlontlo-municipality. [08/04/2017].

Stephman, S.V and Czaplewski, R.L. 1998. Design and Analysis for Thematic Map Accuracy Assessment: Fundamental Principles. *Remote Sensing Environment.* 64: 331-344.

Stevens, F.R., Gaughan, A.E., Linard, C., Tatem, A.J. 2015. Disaggregation Census Data for Population Mapping Using Random Forests with Remotely Sensed and Ancillary Data. *PLOS ONE* 10(2): e0107042. doi: 10.1371/journal.pone.0107042.

Story, M. and Congalton, R.G. 1986. Accuracy Assessment: A User's Perspective. *Photogrammetric Engineering and Remote Sensing.* 52: 397-399.

Tatem, A.J., Noor, A.M., von Hagen, C., Di Gregorio, A., Hay, S.I. 2007. High Resolution Population Maps for Low Income Nations: Combing Land Cover and Census in East Africa. *PLOS ONE* 2(12): e1298. doi: 10.1371/journal.pone.0001298.

Thompson, M. 1996. A Standard Land-Cover Classification Scheme for Remote -Sensing Applications in South Africa. *South African Journal of Science*. 92: 34-42.

Thompson, M. 2015. National Land Cover 2013/14. [Online]. Available: http://bgis.sanbi.org/DEA_Landcover/project.asp. [11/04/2015].

Turner, B.L. 2002. Toward Integrated Land-Change Science: Advances in 1.5 Decades of Sustained International Research on Land-Use and Land-Cover Change. *In* Steffen. W, Jäger. J., Carson, D. J., Bradshaw, D. (eds.) *Challenges of a Changing Earth*. Berlin: Springer. P. 21-26.

United Nations (UN). 2008. *Principles and Recommendations for Population and Housing Censuses Revision 2.* New York: United Nations.

United Nations (UN-GGIM). 2010. *Definition of National Census Geography*. [Online].Available: http://ggim.un.org/docs/meetings/Regional%20Workshop_Jordan/Session%202-Definition%20of%20National%20Census%20Geography.pdf. [17/10/2016].

Van Niekerk, R. 2008. *Social Policy in South Africa*. PhD. Thesis. Oxford University.

Wannenburgh, A. 2014. *Ntabelanga Dam Catchments*. [Online]. Available: http://www.arcgis.com/home/item.html?id=5e293e6560bd49cdaca588342e41d70a. [11/04/2015].

Wright, J.K. 1936. A Method of Mapping Densities of Population with Cape Cob as an Example. *Geographic Review.* 26(1): 103-110.

YesMedia. 2015a. *Methodology, Source and Acknowledgements.* [Online]. Available: http://localgovernment.co.za/source. [31/05/2017].

YesMedia. 2015b. *The Local Government Handbook: South Africa 2015. A Complete Guide to Municipalities in South Africa* (5e). Mowbray: YesMedia.

Yuan, Y., Smith, R. M. and Limp, F. 1997. Remodelling Census Population with Spatial Information From LandSat TM Imagery. *Computer, Environment. and Urban Systems.* 21 (3/4): 245-258.