Examining the Effect of Population Density on Agricultural Land Use and Nutrition Outcomes in Malawi from 1998 to 2008

1. Executive Summary

This project studies the relationship between population growth, agricultural land use, and food security (measured by malnutrition outcomes). In particular, I study the changes between 1998 and 2008 in Malawi's Northern region at the district level. The need for such a study arises from the rapidly growing population in Malawi, particularly around urban areas, which puts pressure on its limited land available for cultivation. This in turn may lead to food insecurity problems if the land is not managed properly, especially in Malawi, where malnutrition among children is very high. For this study, I will consider all the districts around Mzuzu city, most of which are in the Northern region of Malawi.

2. Research Description

2.1. Purpose

This study is aimed at understanding the link between population growth and use of agricultural land near Mzuzu city in Malawi, and the consequent effect on nutrition outcomes from 1998 to 2008. The specific hypothesis to be tested is whether population growth leads to a change in the area of land used for agriculture, and whether this change in agricultural land use is associated with malnutrition outcomes. Depending on the share of agricultural land in 1998, there could be two likely outcomes. If there are no opportunities for agricultural expansion, then as population increases, the area under agriculture will remain the same, which will lead to more intense cultivation of the existing land and lower nutritional outcomes. If there are opportunities for agricultural expansion, then land used for agriculture will grow as population increases. This may or may not improve nutritional outcomes, depending on whether the land is utilized properly. By understanding the link between population growth, agricultural land use, and nutrition outcomes, I attempt to better understand how to solve the problem of food insecurity in Malawi.

2.2. Background

According to USAID (2017), 80% of Malawi's population is engaged in smallholder subsistence farming. The agricultural sector also accounts for one-third of Malawi's gross domestic product, and significantly contributes to food security and nutrition (The World Bank, 2017). The problem of subsistence agriculture failing to meet food demand is further aggravated by the rapid population growth in Malawi. The population density has grown from 112 people per square km in 1987 to 147 people per square km in 2008. This increasing population density is likely to put pressure on agriculture, which could lead to either expansion or increased cultivation on existing land (Malawi Department of Population and Development 2012).

Since 47% of children in Malawi are stunted, there is an increasing focus on agriculture, food security, and nutrition outcomes from the government of Malawi (USAID 2017). The government has developed a National Nutrition Policy and Strategic Plan to address issues of food insecurity through agriculture policy. This research will provide some insight on solving the problem of food insecurity and malnutrition by exploring its link to population density and agricultural land use.

Ricker-Gilbert et al. (2014) explore the link between population density and agricultural intensification, and consequently production in Malawi. They define agricultural intensification in terms of input usage and productivity and one of the ways in which they measure it is by the gross value of crop output per unit of land cultivated. Their research is based on the reasoning that since Malawi has little potential to expand land under cultivation, agricultural production intensifies to meet population growth and agricultural land does not grow much more. I test this theory later in my analysis.

Their methodology also provides a useful way of conceptualizing Malawi's agricultural land when classifying land cover data. The dataset that they use defines arable land in two ways: land that is explicitly under crops, and land which is likely to support conversion to agricultural usage. Since I examine Malawi's land cover data during fallow season, it is difficult to observe the former type of arable land. It is likely that my classification captures the latter, which is the classification that Ricker-Gilbert et al. (2014) use as well.

While there does not seem to be much literature on the link between agricultural pressures and malnutrition outcomes, Orr (2000) explores the link between smallholder agriculture and poverty alleviation in Malawi. The paper also sheds light on the reasons behind the stagnation of productivity from smallholder agriculture. One of the possible reasons is that the increasing population puts pressure on the limited land, thereby decreasing farm size. I will indirectly explore whether this is true between 1998 and 2008, using nutritional outcomes as a proxy for productivity.

The existing literature also provides support for a special focus on urban agriculture in Malawi. For instance, Mkwambisi et al (2010) evaluate how food production in cities affects food security, employment, and income in Malawi. They study two main cities, Lilongwe and Blantyre, and argue for separately examining urban areas because the link between land tenure and food security may not hold for cities. They find that the level of food security in urban farming households varies with education, income level, and size of cultivated land. In my analysis, I focus on the third biggest city after Lilongwe and Blantyre and examine how its population increase affects the districts around it. Mzuzu city is included in the form of a separate district.

2.3. Methods

2.3.1. Country and Time Period

This project examines change in population density and area of land being used for agriculture in Malawi's Mzuzu city between two census years, 1998 and 2008. The population density data has been derived from the censuses conducted by the National Statistics Office of Malawi, the agricultural data is be derived by classifying land cover from Landsat 5 satellite imagery, and the malnutrition data comes from the Demographic Health Survey (DHS). I chose this particular time period because it was the most recent instance of census and DHS coinciding.

2.3.2. Region of Interest

The region of interest is the area around Mzuzu city in the Northern region of Malawi, including a few enumeration areas from Central Malawi, namely the districts Chitipa, Karonga, Nkhata Bay, Rumphi, Mzimba, Kasungu, and Nkhotakota. Mzuzu city is considered as a separate district, thereby amounting to a total of 8 districts serving as enumeration areas as is reported in the 2008 census of Malawi. Figure 1 (Annex A) displays a map of Malawi's districts and major cities. As seen from my population density calculations, Mzuzu city is highly populated and provides an opportunity for examining the effect of this urban population pressure on neighboring districts.

2.3.3. Enumeration Areas and Census Data

Malawi has 3 major administrative levels: regions, districts, and traditional authorities, in the order of hierarchy. According to the online statistical database, Statoids, a few older districts were subdivided to form smaller districts over time. Thus, I will consider the districts as they existed in 2008, along with Mzuzu city as a separate enumeration area in order to isolate urban and rural areas. In order to aggregate 1998 data to 2008 enumeration areas, I use census data that is available for that year at a lower enumeration level (traditional authority). For the most part, these traditional authorities are contained within the 2008 districts. Thus, I spatially join the data from 1998 to the 2008 polygons and compare population across the years accordingly.

2.3.4. Land Cover Data

For each year, I use two satellite images from July taken on the same day and mosaic them together in order to better cover the area of interest. As previously mentioned, the satellite imagery is obtained from Landsat 5 from the website Earthdata. The land cover of Malawi is primarily comprised of cultivated land, public land, hills, steep slopes, wetland, and protected areas. I chose to focus on cultivated land, of which 30,000 agricultural estate farms cover around 1.2 million ha, and 1.8–2 million smallholder farms cover about 4.5 million ha (Ricker-Gilbert et al 2014).

In order to obtain land cover classes, I tried both unsupervised classification with k-means clustering and supervised classification based on 40 training polygons throughout the area of interest. In both these classifications, I used the Normalized Difference Vegetation Index (NDVI) of the image as an additional layer. Based on ground-truthing, I found that unsupervised classification worked better with a kappa statistic of 0.76 (Annex B, Tables 1-3). Thus using this classification, the final land cover classes were: agriculture, water, forest, urban/barren, and other (clouds are classified and excluded; other class includes grasslands, shrubs, etc.). Since the focus of my project is on agriculture, I did not try to distinguish between different types of forests or further separate the urban/barren class into urban and barren land.

2.3.5. Malnutrition Data

The malnutrition data is obtained from Malawi's Demographic Health Surveys (DHS) in 2000 (corresponding to 1998 census) and 2010 (corresponding to 2008 census). This survey provides household level information health and economic conditions and contains information about the "clusters" in which households are interviewed. Each cluster consists of a group of households. In order to use this malnutrition data for my analysis, I first aggregate the raw survey data by

DHS cluster to obtain the average values of the three types of nutrition scores for children under 5: height-for-age (indicates stunting, long term malnutrition), weight-for-age (indicates wasting, current malnutrition), and weight-for-height (composite of height-for-age and weight-for-age, good measure of nutritional status). Each of these scores was also used to generate the percentage of children who were under two standard deviations of these scores, which indicates the percentage of the children interviewed that are stunted, wasted, or underweight, respectively. This percentage is referred to as prevalence in the following sections.

This cluster-level aggregated data was then spatially joined to the 2008 districts and then aggregated by district. Thus, I the average scores and the percentage of children who were malnourished according to the score in each of the three cases for each district. It is important to note that some districts did not have malnutrition data associated with them, either due to lack of DHS clusters or the contained clusters not containing data about children's nutrition. Thus, there are 5 districts left in the sample: Kasungu, Mzimba, Nkhata Bay, Nkhotakota, and Rumphi. For each DHS year, there were about 3000 children in the sample in the area of interest.

2.3.6. Analysis

I started with calculating population density in 1998 and 2008 by districts, and also the simple yearly population change across those years, as shown in Figures 2, 3, and 4 (Annex B).Then, I used land cover classes to aggregate agricultural land use for the years 1998 and 2008 by districts in the area of interest and examined the growth in area of land being used for agriculture as demonstrated in Figures 6 and 7 (Annex B). I examined the relationship between population growth and agricultural land use (Annex B, Figure 8). Finally, I used the data on malnutrition to examine the link between change in population density, agricultural land use, and malnutrition outcomes (Annex C, Table 6, Figures 9-11).

3. Results

3.1. Population Change (Annex A)

From Figure 2, I see that in 1998, only Muzu city and Likoma had a very high population density. In Figure 3 for 2008, more neighboring districts show a high population density, especially Karonga, Kasungu, and Nkhotakota. Figure 4 shows that Mzuzu city had a very high population growth rate of about 10% per year, and was followed by the districts Chitipa, Karonga, and Kasungu, which had growth rates close to 5% a year.

3.2. Land Cover Change (Annex B)

Figure 5 and 6 display an exploratory analysis of land cover change based on the unsupervised k-means classification. It can be seen that agricultural land cover increases from 1998 to 2008. Table 5 and Figure 7 show that most of the increase in agricultural land has taken place in Rumphi, Nkhotakota, and Mzuzu City at the rate of about 16% per year. As can be seen from Table 5, these districts started from a considerably low level of agricultural land use in 1998 as compared to districts like Mzimba and Chitipa which were already somewhat saturated and did not see as big an increase in agricultural land usage in 2008. This shows that agricultural expansion was occurring between 1998 and 2008, and Table 4 demonstrates that this increase in agricultural land is obtained by clearing forests and using urban/barren land (the water class is likely capturing clouds by error).

As for the relationship between population growth and change in area under agricultural usage, they are found to be positively correlated with R2 = 0.18 (Annex B, Figure 8). It is also important to note that the increase in agricultural land use may not result from population increase within the district, but from population increase in a neighboring district or major city, which causes pressure on surrounding food sources. This may be the case for Rumphi, which shows a very high agricultural land use increase. Even though it did not have a high population increase between 1998 and 2008, it is surrounded by the districts Chitipa and Mzuzu city, which saw a boom in population (Annex A, Figure 4; Annex B, Figure 7).

3.3. Malnutrition Outcomes (Annex C)

From Table 6 and Figures 9-11, it seems that there is a positive correlation between increase in agricultural area and reduction in malnutrition. However, it is difficult to say so with confidence due to the small number of data points. On a case-by-case basis, Rumphi and Nkhotakota, which have the highest increase in agricultural area also show the highest reduction in the prevalence of stunting, wasting, and underweight children.

4. Discussion

Overall, I see that population growth is associated with an increase in agricultural land use through expansion, which is in turn associated with better nutritional outcomes. According to The Ecologist (2015), the increase in population in Malawi leads to people expanding their homes on their fields, which thus creates a need to develop farmland elsewhere in non-residential areas. This is also supported by Table 4 (Annex B), which shows a majority of new agricultural land in 2008 being formed by what were forests and bare land in 1998.

The link between agricultural land increase and better nutritional outcomes is harder to explain. There could be a multitude of confounding factors that lead to the presence of this correlation. A potential explanation is government policies that are introduced to support the agricultural sector to meet growing population demands. One such policy is the Malawi Growth and Development Strategy (MGDS) that was introduced in 1998. In addition to increasing proper agricultural land use, this policy is aimed at poverty alleviation and food security (Food and Agriculture Organization, 2015). Thus, it is possible that the observed trend is a result of such government policies.

Other than a small number of data points and the high likelihood of confounding of effects, it is important to recognize the other limitations of this research. Due to the large preparation and processing time of satellite imagery, one two granules were used for each year. Thus, the district Lilongwe is not available for land cover analysis, and only a small portion of Kasungu and Nkhotakota districts is present in the land cover dataset. Thus, I am cautious about drawing conclusions from them.

Another issue is that the time of the satellite image's acquisition is in July, which is in fallow season. Malawi's maize growing season is from December to February, but the presence of a large number of clouds in the satellite imagery made it less preferable to use images from this time for analysis. Due to this, there are very sparse crops visible in satellite imagery, which makes it difficult to distinguish between grasslands, shrubs, agriculture plots, and even bare land.

Additionally, the large number of subsistence farming plots are difficult to differentiate without vegetation since their boundaries are not as clearly marked. This is potentially the reason why unsupervised classification performs better than supervised classification. It is possible that the training polygons selected for supervised classification were not appropriate for differentiating between agriculture and other land.

5. Conclusions

My findings directly address some theories outlined in the literature. For instance, the Ricker-Gilbert et al. (2014) paper states that an increase in population should increase agricultural density based on their research on Malawi using household level panel data from 2003 to 2008. My findings contradict this hypothesis since I observe very high growth rates of agricultural land, which are greater than the population growth rates. This suggests that there is a need to do more research in this area using different techniques to assess the impact of population change on agricultural density in Malawi.

Even though the results from this study are more exploratory than conclusive, they provide a useful framework for considering population, agriculture, and malnutrition in Malawi. The results from this study could be used to inform future projects such as a causal analysis of Malawi's agriculture-focused government programs. The same analysis as in this project can be repeated for an appropriately chosen area of interest to generate a dataset for a difference-in-difference study. A future study should also include finer temporal resolution, smaller enumeration areas, and a careful consideration of confounding factors such as economic growth, immigration, migration across districts, etc.

6. References

Crouch, Marc. "Sustainable Agriculture in Malawi: a desperate struggle." *Ecologist*, 17 April 2015, https://theecologist.org/2015/apr/17/sustainable-agriculture-malawi-desperate-struggle.

Department of Population and Development. *Why Population Matters to Malawi's Development: Managing Population Growth for Sustainable Development*. Ministry of Economic Planning and Development, 2012, p. 24. *Zotero*, http://www.prb.org/pdf12/malawi-population-matters.pdf.

Food and Agriculture Organization of the United Nations. "Country Fact Sheet on Food and Agriculture Policy Trends." *Food and Agriculture Policy Decision Analysis*, March 2015, http://www.fao.org/3/a-i4491e.pdf.

Mkwambisi, David D., et al. "Urban Agriculture and Poverty Reduction: Evaluating How Food Production in Cities Contributes to Food Security, Employment and Income in Malawi." *Journal of International Development*, vol. 23, no. 2, Mar. 2011, pp. 181–203. *CrossRef*, doi:10.1002/jid.1657.

Orr, Alastair. "'Green Gold'?: Burley Tobacco, Smallholder Agriculture, and Poverty Alleviation in Malawi." *World Development*, vol. 28, no. 2, Feb. 2000, pp. 347–63. *ScienceDirect*, doi:10.1016/S0305-750X(99)00127-8.

Ricker-Gilbert, Jacob, et al. "How Does Population Density Influence Agricultural Intensification and Productivity? Evidence from Malawi." *Food Policy*, vol. 48, Oct. 2014, pp. 114–28. *ScienceDirect*, doi:10.1016/j.foodpol.2014.02.006.

The World Bank. "New Policies to Help Transform Malawi's Agriculture Sector." *World Bank*, 31 Jan. 2017, http://www.worldbank.org/en/news/feature/2017/01/31/new-policies-to-transform-malawi-agriculture-sector.

USAID. *Agriculture and Food Security | Malawi | U.S. Agency for International Development*. 1 Aug. 2017, https://www.usaid.gov/malawi/agriculture-and-food-security.

Annex A: Population Data and Change Analysis

Metadata

- Two polygon shape files used so far:
  - 2008 population: A shape file with Malawi's districts (major cities counted as districts) with population data from the 2008 census.
  - 1998 population: A shape file with census enumeration tracts and detailed attributes such as male population, female population, which traditional authority and district the tract lies in, etc from the 1998 census.
- Both data sources obtained from MASDAP (Malawi Spatial Data Platform) website. Links:
  - 2008 population: http://www.masdap.mw/layers/geonode%3Acensus_dist, published on Oct 31, 2013.
  - 1998 population: http://www.masdap.mw/layers/geonode%3Aeas_bnd, published on Nov 28, 2013.
- Owner of these datasets listed as geonode on MASDAP, which is a geospatial content management system, a platform for the management and publication of geospatial data.
- Contact information for MASDAP administers: National Spatial Data Center, Department of Surveys, (13.973456 S, 33.7878122 E), Lilongwe, Malawi.
- Extent of the shapefiles:
  - xMin,yMin 32.6789,-17.1353 : xMax,yMax 35.9242,-9.36718
- 2008 population attribute description:
  - Contains the following attributes (summary statistics also provided):

Attribute Name	Label	Average	Median	Standard Deviation
Objectid	Objectid	16.88	16.50	9.64
District	District			
Total_Popu	Total_Popu	446583.31	368323.00	283812.60
Popdensity	Popdensity	428.17	163.14	746.52

  - The primary key is objectid and the join field is Districts. There are no missing districts and no duplicates.
- 1998 population attribute description:
  - Contains the following attributes (summary statistics also provided):

Attribute Name	Label	Average	Median	Standard Deviation
Area	Area	13524026.39	5140119.00	306417542.99
Perimeter	Perimeter	13208.25	11057.16	20006.40
Eacode	Eacode	24459340.28	30101015.00	6839567.64
Households	Households	246.16	241.00	90.42
Total	Total	1075.32	1036.00	412.18
Male	Male	526.85	501.00	214.02
Female	Female	548.48	534.00	203.37

Attribute Name	Label	Average	Median	Standard Deviation
Ta	Ta			
Ta_Code	Ta_Code	24263.64	20909.00	7185.34
District	District			
Dist_Code	Dist_Code	242.55	209.00	71.86
Code	Code	1.00	1.00	0.02
Feature	Feature	195699.56	23.00	2154783.65
Ea_Code2	Ea_Code2	24263640.71	2.0909E7	7185342.63
Density	Density	242.40	161.00	292.23
Area_Km	Area_Km	13.51	5.00	306.42

  - o The primary key is eacode, the join field could be District, but I use a spatial join for joining the 2008 and 1998 data. The enumeration areas of the 1998 census are completely nested within the districts for the 2008 census.
- I use 2008 districts as the enumeration areas for this study, and aggregated 1998 census data accordingly.
- There were no spatial errors in the 2008 data and I used centroids of 1998 data.
- The shapefiles were in the WGS 84 spatial reference system. I developed a custom projection named Malawi Transverse Mercator defined by the following proj4 string: '+proj=tmerc +lat_0=-13.24498 +lon_0=34.295 +k=0.999738 +x_0=800000 +y_0=800000 +datum=WGS84 +units=m +no_defs'
- Note: The shape files themselves did not specify which year the population data was from.
  - o For the districts shape file, I found the total population to be 14.3 million. This was the closest to the Malawi 2008 population reported by the Malawi Population and Housing Census 2012 data sheet (14.8 million) and CityPopulation.de (13 million).
  - o For the enumeration areas shape file, the total population was found to be 9.9 million, which matched with the CityPopulation.de population estimate (9.9 million).

Validation

- As described above, I validated the total population data with other sources.
- Checked the statoids website to find that no boundaries had been restructured from 1998 to 2008, there were only a few districts that had split from each other.
- The borders of both the shape files aligned with each other.

Cleaning Population Data

* Custom projection, fixing geometry, intersection, calculating area, pop density, and pop change.

* 2008 population data referred to as census_dist and 1998 population data referred to as eas_bnd. Eas_bnd consists of smaller enumeration data than the 2008 data, but these visually seem to be generally nested within 2008 enumeration areas. I use a spatial join to aggregate population data for 1998 by 2008 enumeration units.

```
SELECT district, srid(geom), AsText(geom)

FROM pop2008_aoi


* Generate centroid of 1998 enumeration areas and spatial join to 2008
enumeration areas

* just to visualize:

SELECT id, AsText(PointOnSurface(geom)) AS pointText, district FROM pop1998


CREATE TABLE pop1998_pos AS

SELECT pop1998_total.id AS id, PointOnSurface(pop1998_total.geom) AS geom,

pop1998_total.district AS district, pop1998_total.ta AS ta,

pop1998_total.area as areaCheck, pop1998_total.total AS total_popu

FROM pop1998_total

JOIN pop2008_aoi

ON Within(PointOnSurface(pop1998_total.geom), pop2008_aoi.geom)


SELECT RecoverGeometryColumn ('pop1998_pos', 'geom', 4326, 'POINT')


*Then vector -> geoprocessing -> intersection. So now it's a table.
pop1998_new2 is the result of the intersect between pop2008_aoi and
pop1998_pos


*Aggregate population by district

select DISTRICT, sum(total_popu_2)

from pop1998_new2

group by DISTRICT


* There were some duplicated entries in mzuzu, checked others but they were
fine

select sum(mzuzu_popu), sum(areacheck)

from mzuzu_unique
```

```
* Join 1998 and 2008

create table mwPop2 as

select pop2008_aoi.id as id, geom, pop2008_aoi.district as district,

pop2008_aoi.total_popu as pop2008, pop2008_aoi.popdensity as
density2008_check,

tojoin_1998.pop1998 as pop1998

FROM pop2008_aoi

JOIN tojoin_1998 on

pop2008_aoi.district LIKE tojoin_1998.DISTRICT


SELECT RecoverGeometryColumn ('mwPop2', 'geom', 4326, 'MULTIPOLYGON')


* Now custom projection and area calculation

insert into spatial_ref_sys values(110004, 'gg328',  110004, 'Malawi
Transverse Mercator', '+proj=tmerc +lat_0=-13.24498 +lon_0=34.295 +k=0.999738
+x_0=800000 +y_0=800000 +datum=WGS84 +units=m +no_defs', 'Undefined')


* copy over my table

create table mwPopProj as

select *

from mwPop2


SELECT RecoverGeometryColumn ('mwPopProj', 'geom', 4326, 'MULTIPOLYGON')


* srid is in the system

select *

from spatial_ref_sys

where srid = 110004


* Set the projection

select discardgeometrycolumn('mwPopProj', 'geom')

update mwPopProj set geom = transform(geom, 110004)

select recovergeometrycolumn('mwPopProj', 'geom', 110004, 'MULTIPOLYGON')


* calculate ellipsoidal and planar area
```

```
alter table mwPopProj add column areaEllip real
update mwPopProj set areaEllip = Area(geom, 1)/1000000

alter table mwPopProj add column areaPlanar real
update mwPopProj set areaPlanar = Area(geom)/1000000

* Calculate values
alter table mwPopChange add column density2008 float
update mwPopChange set density2008 = pop2008/areaPlanar

alter table mwPopChange add column density1998 float
update mwPopChange set density1998 = pop1998/areaPlanar

alter table mwPopChange add column popChange float
update mwPopChange set popChange = 10.0 * (pop2008-pop1998)/pop1998
```

- Tabular summary of population statistics:

District	Total Population 1998	Total Population 2008	Area	Population Density 1998	Population Density 2008	Population Change (%)
Chitipa	126799	194707	4249.23	29.84	45.84	5.36
Karonga	195962	297694	3410.98	57.45	87.28	5.19
Kasungu	479300	707862	8063.93	59.44	87.82	4.77
Likoma	8074	10426	20.26	398.58	514.41	2.91
Mzimba	697974	795708	10624.64	65.69	74.91	1.40
Mzuzu City	86980	168928	140.28	620.06	1204.24	9.42
Nkhata Bay	164761	236978	4170.83	39.50	56.82	4.38
Nkhotakota	229460	334856	4322.57	53.08	77.46	4.59
Rumphi	128360	187137	4651.99	27.59	40.23	4.58

Figure 1: Major cities in Malawi

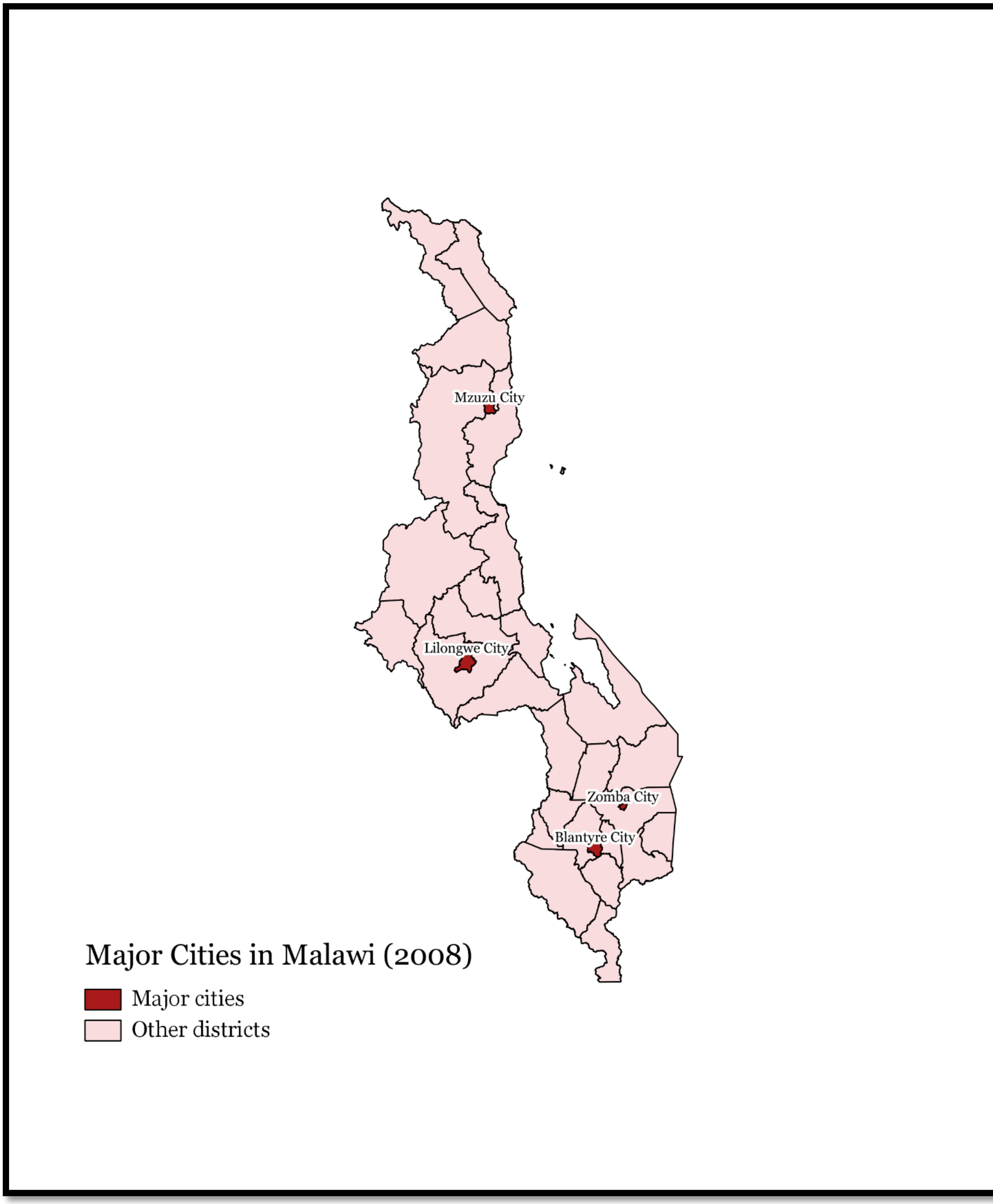


Figure 2: Population Density in Districts in Area of Interest (1998)

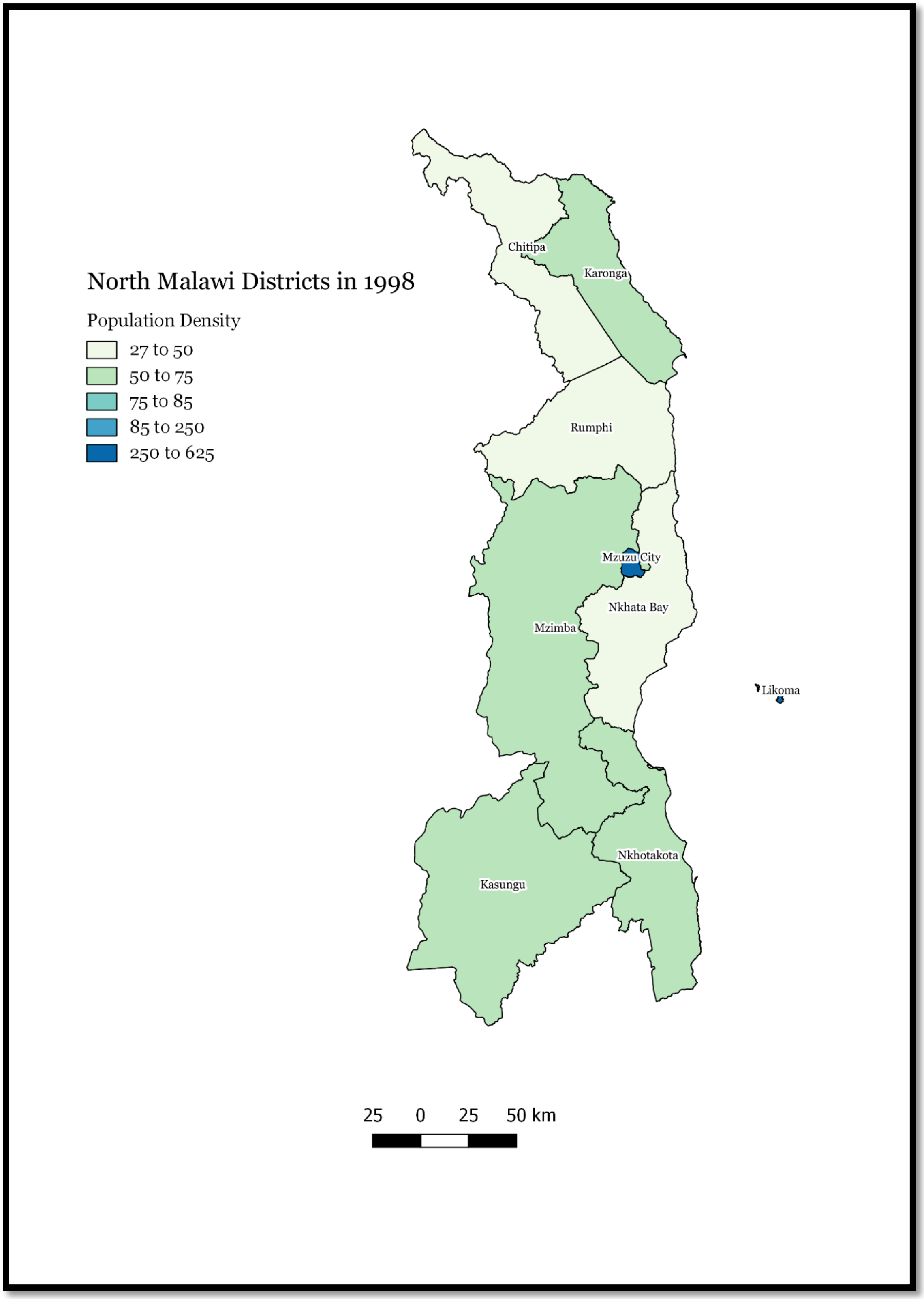


Figure 3: Population Density in Districts in Area of Interest (2008)

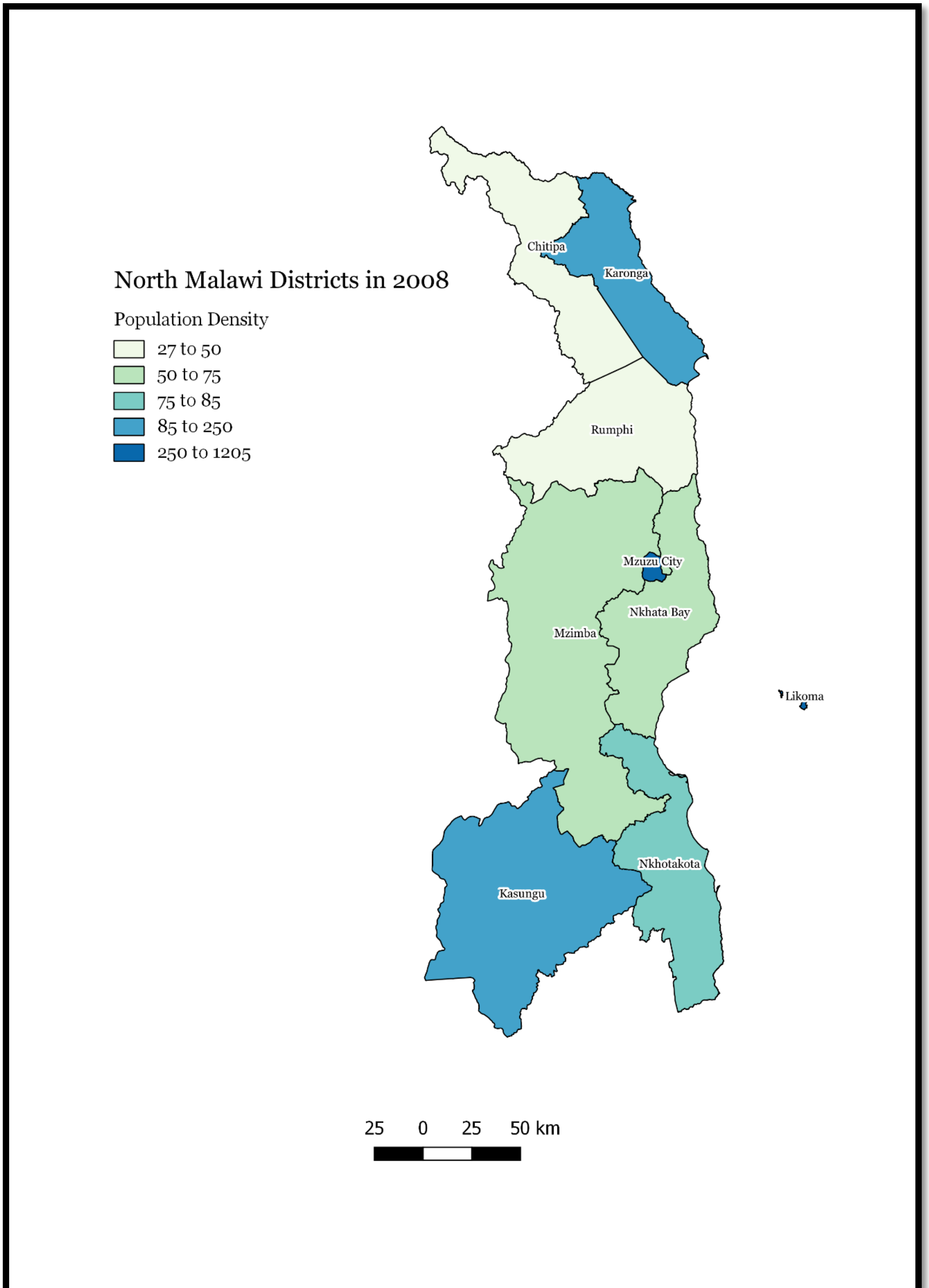
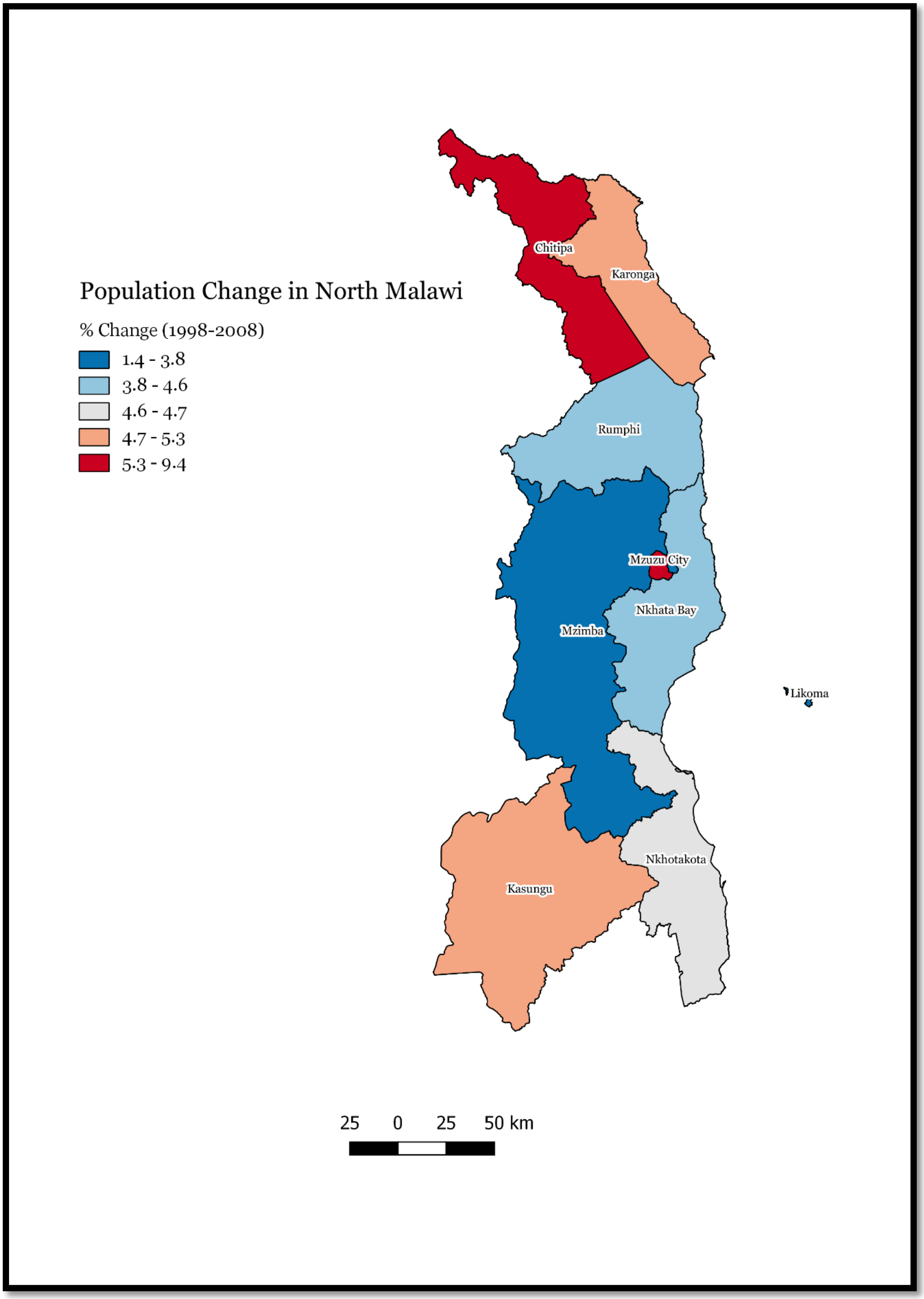


Figure 4: Simple Yearly Population Change in Area of Interest (1998-2008)



Annex B: Environmental Change Analysis

Metadata

- Data obtained from earthdata: https://earthdata.nasa.gov/, Landsat 5 imagery.
- 2 granules for 1998, date acquired: 07/08/1998:
    - LT05_L1TP_169068_19980708_20161223_01_T1
        - Cloud cover: 1
        - Image quality: 9
        - Sun azimuth: 45.45363692
        - Sun elevation: 41.44229409
        - Earth sun distance: 1.0166856
    - LT05_L1TP_169067_19980708_20161223_01_T1.
        - Cloud cover: 3
        - Image quality: 9
        - Sun azimuth: 46.21962361
        - Sun elevation: 42.60934645
        - Earth sun distance: 1.0166856
- 2 granules for 2008, date acquired: 07/03/2008:
    - LT05_L1TP_169068_20080703_20161030_01_T1
        - Cloud cover: 0
        - Image quality: 9
        - Sun azimuth: 42.96052993
        - Sun elevation: 42.82922655
        - Earth sun distance: 1.0167006
    - LT05_L1TP_169067_20080703_20161030_01_T1
        - Cloud cover: 1
        - Image quality: 9
        - Sun azimuth: 43.71674284
        - Sun elevation: 44.02659822
        - Earth sun distance: 1.0167006
- Contact: National Aeronautics and Space Administration (NASA) and U.S. Geological Survey (USGS)
- Source URL:https://search.earthdata.nasa.gov/
- Required attribution: USGS/NASA Landsat Program
- Temporal resolution: March 1984-January 2013
- Cell size: 30m reflective and thermal
- Method of creation: Radiometrically calibrated and orthorectified using ground control points and digital elevation model (DEM) data to correct for relief displacement. These are the highest quality Level-1 products suitable for pixel-level time series analysis.
- Bands used:
    - Blue: Band 1, 0.45-0.52 micrometers, 30m
    - Green: Band 2, 0.52-0.60 micrometers, 30m

  - Red: Band 3, 0.63-0.69 micrometers, 30m
  - NIR: Band 4, 0.76-0.90 micrometers, 30m
  - SWIR1: Band 5, 1.55-1.75 micrometers, 30m
  - SWIR2: Band 7, 2.08-2.35 micrometers, 30m
- File format: GeoTiff
- Coordinate reference system: UTM-36, WGS84

<u>Documentation</u>

Step 1: Obtain Remote Sensing Data

- Save a kml file of the area of interest and load it in Earthdata to look for images (spatial filter).
- Download appropriate granules and geotiff products.
- Downloaded two granules for each year in the same path.

Step 2: Correct and Clean Images

- Conducted this in Orfeo.
- Concatenated blue, green, red, NIR, SWIR1, SWIR2 in that order.
- Optically calibrated using metadata, calibration level: top of atmosphere
- Set no data values to 0.
- Mosaic two granules from the same time period to obtain one raster for each year.
- Create data mask using band math expression <im1b1 && im1b2 && im1b3 && im1b4 && im1b5 && im2b1 && im2b2 && im2b3 && im2b4 && im2b5>.
- Apply this as no data mask to trim images across years.

Step 3: Unsupervised Classification

- Conducted this in SAGA.
- Calculated enhanced NDVI layer for each year.
- Used k-means clustering for grids on cleaned rasters, added NDVI layer.
- Reclassify grid to group values that are similar, end categories: forest, agriculture, clouds, water, urban/barren, other.
- Reduced noise from classification filter by using majority filter.
- Converted SAGA grid to polygon to use in QGIS.

Step 4: Create Training Areas for Supervised Classification

- Loaded polygon of area into Google Earth, ran majority filter more in order to make this polygon manageable.
- Made polygons for training classes.

Step 5: Supervised Classification

- Used clean image of the year and training polygons for supervised classification in SAGA. Initial classes: Urban, Forest, Industry, Agriculture, Grassland, Sparse agriculture, Open water.
- Reclassified grid to combine any classes.

Step 6: Assess Accuracy of Supervised and Unsupervised

- Resampled supervised and unsupervised grid to 300m and run majority filter to create polygons of classes in SAGA.
- Opened as shapefile in QGIS and took a fixed sample of random points inside each polygon.
- Exported these points to google earth for ground truthing.
- Compared with satellite data for that particular time period.
- Used complete layer of ground truthed points and turned into a grid in SAGA for cross-tabulation.
- Created a cross-tabulation of unsupervised grid and ground truthed points.
- Reclassified supervised classification grid to match unsupervised classes. Cross-tabulated supervised classification with truthed points.
- Compared performance of supervised and unsupervised and decided to use unsupervised.

Step 7: Detect Land Change

- Cross tabulated grid for 1998 and 2008, use 1998 image as first grid.
- Compared land change patterns.

Step 8: Summarize Change by Enumeration Area

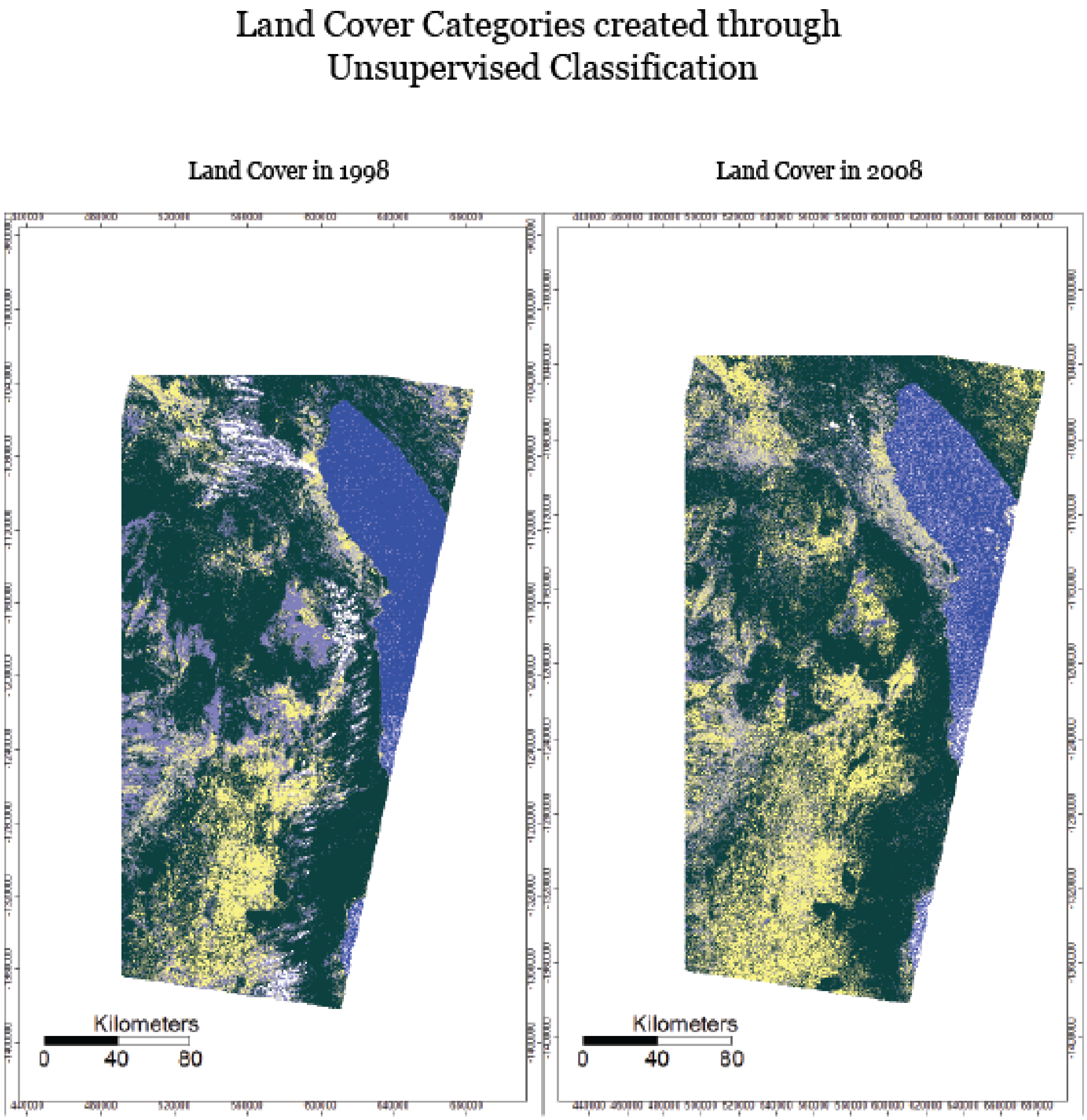
- Can do this is SAGA.
- Loaded shapefile of enumeration areas (2008 districts) into saga and converted into a grid.
- Cross-tabulated raster districts with land cover change. This is a more efficient way of doing a spatial join.
- Also reclassified land cover change to 3 classes: remain in agriculture, new agriculture, loss of agriculture to show agricultural land use patterns. Can also aggregate this by enumeration area by cross-tabulating.

Step 9: Clean agricultural land cover data in QGIS

```
*Add agricultural data

create table agrLand as

select *

FROM mwPopProj

JOIN agrTable2008 on

mwPopProj.id LIKE agrTable2008.id
```

```
alter table agrLand3 add column agr2008 float

update agrLand3 set agr2008 = 900.0 *(remainAgr + newAgr)/1000000


alter table agrLand3 add column agr1998 float

update agrLand3 set agr1998 = 900.0 *(remainAgr + lossAgr)/1000000


alter table agrLand3 add column agrPerc2008 float

update agrLand3 set agrPerc2008 = 100.0 * agr2008/areaPlanar


alter table agrLand3 add column agrPerc1998 float

update agrLand3 set agrPerc1998 = 100.0 * agr1998/areaPlanar


alter table agrLand3 add column agrChange float

update agrLand3 set agrChange = 10.0 * (agr2008 - agr1998)/ ifnull(agr1998,
1)


alter table agrLand3 add column agrDensity2008 float

update agrLand3 set agrDensity2008 = pop2008 / agr2008


alter table agrLand3 add column agrDensity1998 float

update agrLand3 set agrDensity1998 = pop1998 / agr1998


alter table agrLand3 add column agrDensityChange float

update agrLand3 set agrDensityChange = 10.0 * (agrDensity2008 -
agrDensity1998)/ ifnull(agrDensity1998, 1)
```

Figure 5: Land Cover Classification of Area of Interest (1998)



COLOR	NAME
	Agriculture
	Clouds
	Forest
	Other
	Urban/barren
	Water

Table 1: Accuracy Assessment of 2008 Unsupervised Classification

Found on ground →

Found on map ↓

Name	Agriculture	Forest	Urban/barren	Water	Other	SumUser	AccUser
Agriculture	6	0	3	2	0	11	54.54546
Forest	1	11	2	0	0	14	78.57143
Urban/barren	0	0	7	0	0	7	100
Water	0	0	1	9	1	11	81.81818
Other	0	0	0	0	8	8	100
SumProd	7	11	13	11	9	**Kappa**	**0.755513**
AccProd	85.714286	100	53.846154	81.818182	88.888889	**Overall Accuracy**	**0.803922**

Table 2: Accuracy Assessment of 2008 Supervised Classification

Found on ground →

Found on map ↓

Name	Agriculture	Forest	Urban/barren	Water	Other	SumUser	AccUser
Agriculture	5	0	4	7	0	16	31.25
Forest	1	10	2	0	0	13	76.92308
Urban/barren	0	0	0	4	0	4	0
Water	0	0	0	0	0	0	
Other	0	0	0	0	1	1	100
SumProd	6	10	6	11	1	**Kappa**	**0.323757**
AccProd	83.333333	100	0	0	100	**Overall Accuracy**	**0.470588**

Table 3: Comparison between 2008 Supervised and Unsupervised Clusters

Supervised classification →

Unsupervised classification ↓

Name	Agriculture	Forest	Urban/barren	Water	Other	SumUser	AccUser
Agriculture	13380855	39859	8396	0	23	13429133	99.6405
Forest	71768	24666856	867	0	225	24739716	99.70549
Urban/barren	0	0	0	0	0	0	
Water	856504	12138	852438	0	3057	1724137	0
Other	0	0	0	0	0	0	
SumProd	14309127	24718853	861701	0	3305		
AccProd	93.512728	99.789646	0		0		

Table 4: Land Cover Changes from 1998 to 2008 (sq km)

Name	Agriculture	Forest	Urban/barren	Water	Other	SumUser	AccUser
Agriculture	5814.31	346.75	127.46	345.27	3.74	6637.53	87.60
Forest	2833.93	20568.68	1223.84	317.07	21.58	24965.10	82.39
Urban/barren	3210.64	3335.56	716.56	287.16	20.04	7569.95	9.47
Water	1153.41	466.71	563.45	849.84	3.22	3036.62	27.99
Other	14.26	20.86	6.87	4.64	4795.05	4841.68	99.04
SumProd	13026.56	24738.56	2638.18	1803.96	4843.63		
AccProd	44.63	83.14	27.16	47.11	99.00		

Table 5: Land Cover Change Summary by Districts (1998-2008)

District	Area under agriculture in 2008 (sq km)	Area under agriculture in 1998 (sq km)	% Change in Agricultural Land (1998 to 2008)	% of Land under Agricultural in 2008	% of Land under Agriculture in 1998
Chitipa	1267	499	15.39	29.82	11.74
Karonga	278	263	0.57	8.15	7.71
Nkhata Bay	37	23	6.09	0.89	0.55
Rumphi	1031	392	16.30	22.16	8.43
Mzimba	5313	3574	4.87	50.01	33.64
Kasungu	29	16	8.13	0.36	0.20
Nkhotakota	38	14	17.14	0.88	0.32
Mzuzu City	49	19	15.79	34.93	13.54

Figure 6: Agricultural Land Cover Changes (1998-2008)

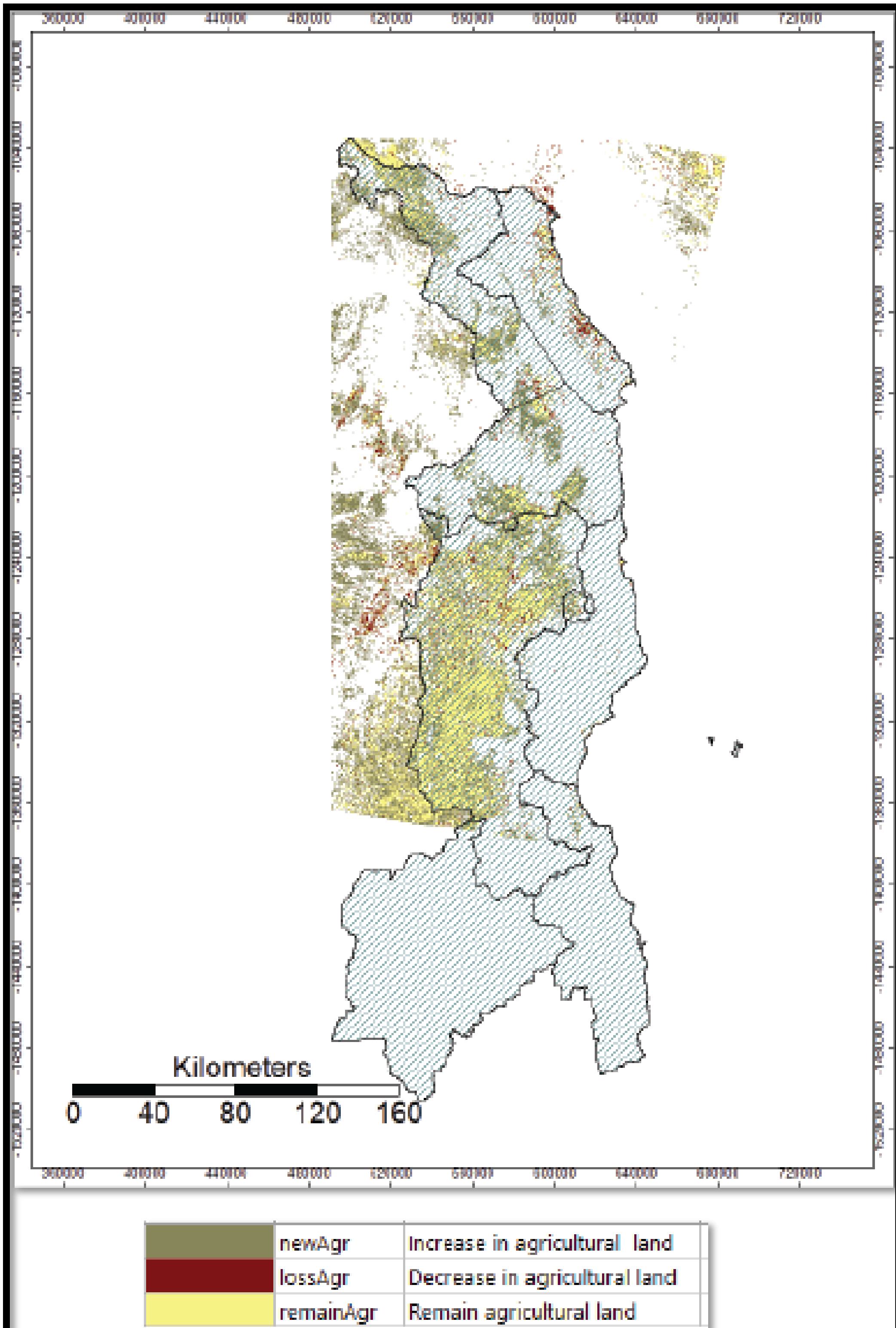


Figure 7: Agricultural Land Cover Change per District (1998-2008)

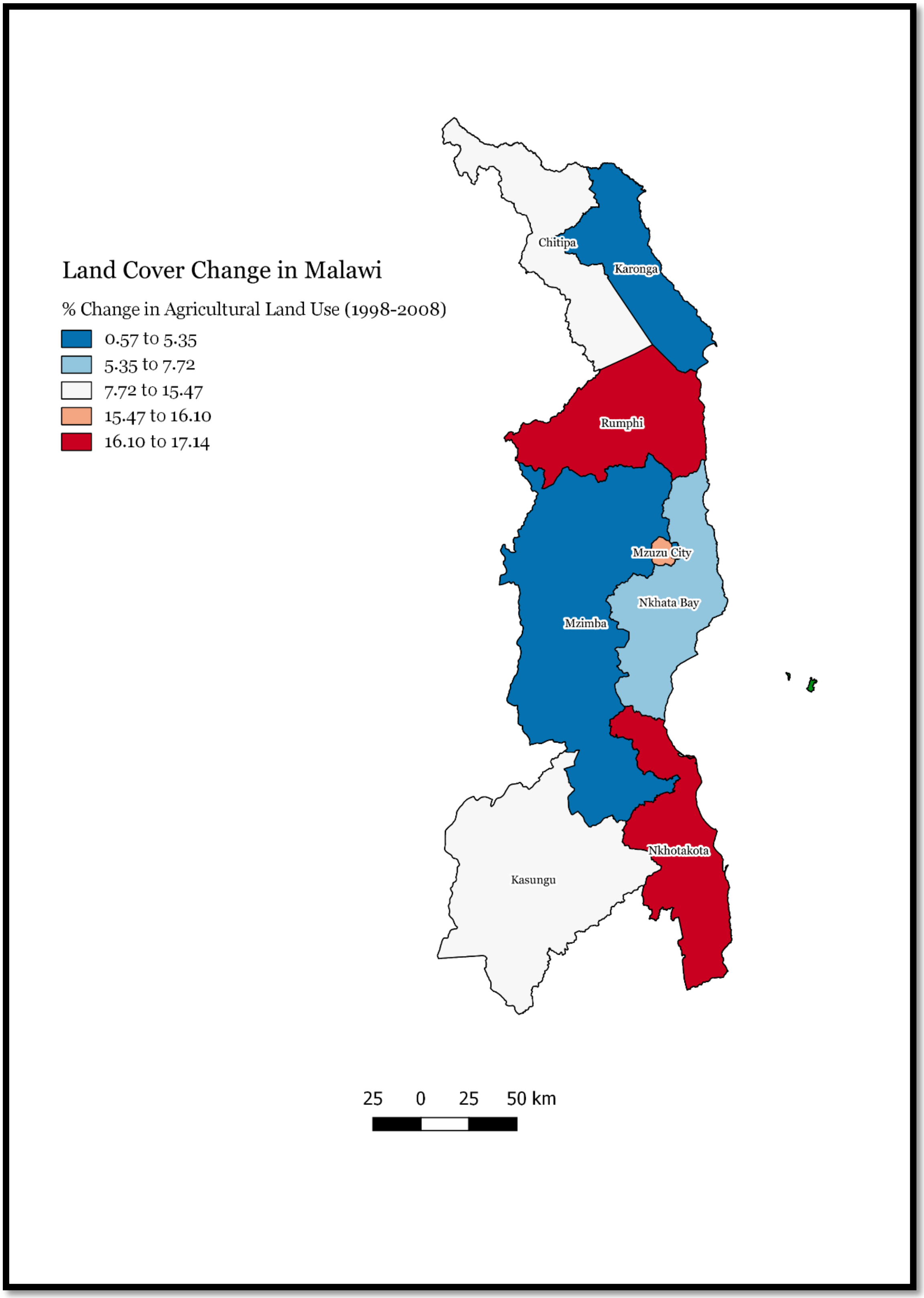
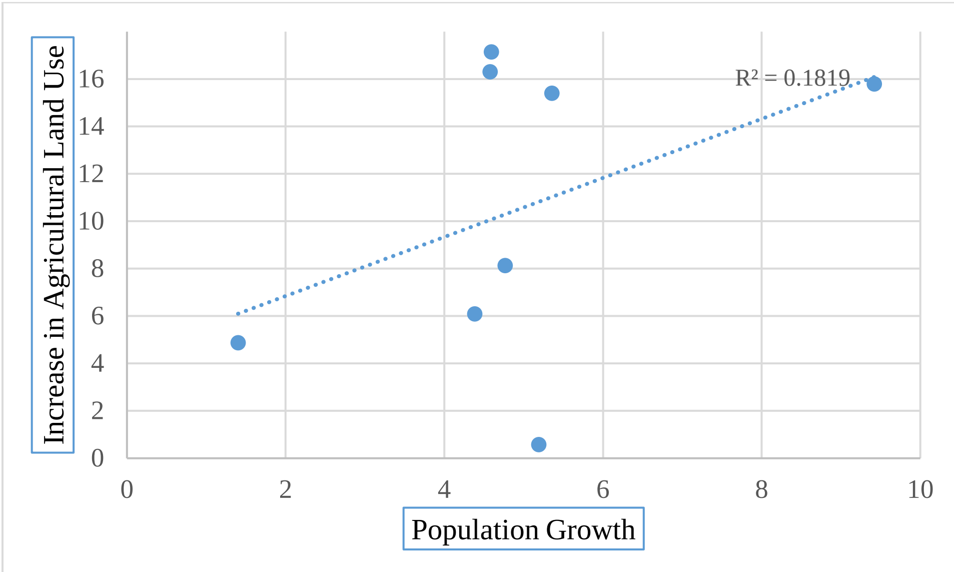


Figure 8: Relationship between Population growth and Agricultural Land Use in Malawi's Northern Region Districts (1998-2008)



Annex C: Malnutrition Outcomes

Metadata

- Two polygon shape files with location of each cluster.
  - Extent: xMin,yMin 32.9054,-17.1212 : xMax,yMax 35.8312,-9.57217
  - Spatial reference system: +proj=longlat +datum=WGS84 +no_defs
  - 560 clusters in all of Malawi for 2000, 848 for 2010.
- Two survey datasets, each row is an individual response:
  - 2000: MWHR41FL.DTA
  - 2010: MWHR61FL.DTA
- All data sources found on DHS website after applying for access. https://dhsprogram.com/.

Documentation (For conducting the aggregation explained in methodology)

In STATA:

```
* Aggregate for 2000

Use
"W:\geog0328\fp\dhs_data\MW_2000_DHS_05162018_1540_98784\MWHR41DT\MWHR41FL.DT
A", clear


drop if abs(hc5_1) > 600 | abs(hc8_1) > 600 | abs(hc11_1) > 600 | hc5_1 == .
| hc8_1 ==. | hc11_1==.

gen ha_sd = hc5_1/100  * This is std deviations from height for age z score

gen wa_sd = hc8_1/100 * Same for weight for age

gen wh_sd = hc11_1/100 * Same for weight for height


gen stunting_perc = (ha_sd < -2) /* long term nutritional health*/

gen wasting_perc = (wa_sd < -2) /* current nutritional health*/

gen underwt_perc = (wh_sd < -2) /*avg*/


collapse (mean) ha_sd wa_sd wh_sd stunting_perc wasting_perc underwt_perc,
by(hv001)

* Save as csv to load in QGIS

* Aggregate for 2010

use
"W:\geog0328\fp\dhs_data\MW_2010_DHS_05162018_1542_98784\MWHR61DT\MWHR61FL.DT
A", clear
```

```
drop if abs(hc5_01) > 600 | abs(hc8_01) > 600 | abs(hc11_01) > 600 | hc5_01
== . | hc8_01 ==. | hc11_01==.

gen ha_sd = hc5_01/100

gen wa_sd = hc8_01/100

gen wh_sd = hc11_01/100


gen stunting_perc = (ha_sd < -2) /* long term nutritional health*/

gen wasting_perc = (wa_sd < -2) /* current nutritional health*/

gen underwt_perc = (wh_sd < -2) /*avg*/

collapse (mean) ha_sd wa_sd wh_sd stunting_perc wasting_perc underwt_perc,
by(hv001)

*save as csv to load in QGIS
```

In QGIS Sqlite:

```
create table geoNutr2000 as

SELECT *

FROM nutrition_2000

JOIN points2000

ON nutrition_2000.hv001 like points2000.dhsclust


SELECT RecoverGeometryColumn ('geoNutr2000', 'geom', 4326, 'MULTIPOINT')


create table geoNutr2010 as

SELECT *

FROM nutrition2010

JOIN points2010

ON nutrition2010.hv001 like points2010.dhsclust


SELECT RecoverGeometryColumn ('geoNutr2010', 'geom', 4326, 'MULTIPOINT')


* Now joining by district

create table distNutrJoin2000 as

SELECT geoNutr2000.ha_sd as ha_sd, geoNutr2000.wa_sd as wa_sd,
geoNutr2000.wh_sd as wh_sd,
```

```
geoNutr2000.stunting_perc as stunting_perc, geoNutr2000.wasting_perc as
wasting_perc, geoNutr2000.underwt_perc as underwt_perc,

geoNutr2000.geom as geom, geoNutr2000.id as id, geoNutr2000.dhsclust as
dhsclust, popChange_aoi.district as district

FROM geoNutr2000

JOIN popChange_aoi

ON Within(geoNutr2000.geom, popChange_aoi.geom)


create table nutrition2000_summary as

select avg(ha_sd), avg(wa_sd), avg(wh_sd), avg(stunting_perc),
avg(wasting_perc), avg(underwt_perc),

district

from distNutrJoin2000

group by district

*Repeat same for 2010.
```

Table 6: Relationship between Population Growth, Agricultural Land Use, and Nutritional Status Change

District	Population Change (%)	% Increase in Agricultural Area	% Reduction in Stunting Prevalence	% Reduction in Wasting Prevalence	% Reduction in Underweight Prevalence
Kasungu	4.77	8.13	7.08	36.15	50.18
Mzimba	1.40	4.87	3.44	28.95	70.26
Nkhata Bay	4.38	6.09	-5.15	-20.37	59.78
Nkhotakota	4.59	17.14	29.87	52.98	77.23
Rumphi	4.58	16.30	43.24	76.71	100.00

Figure 9: Relationship between Stunting Prevalence along children and Agricultural Land Use in Malawi's Northern Region Districts (1998-2008)

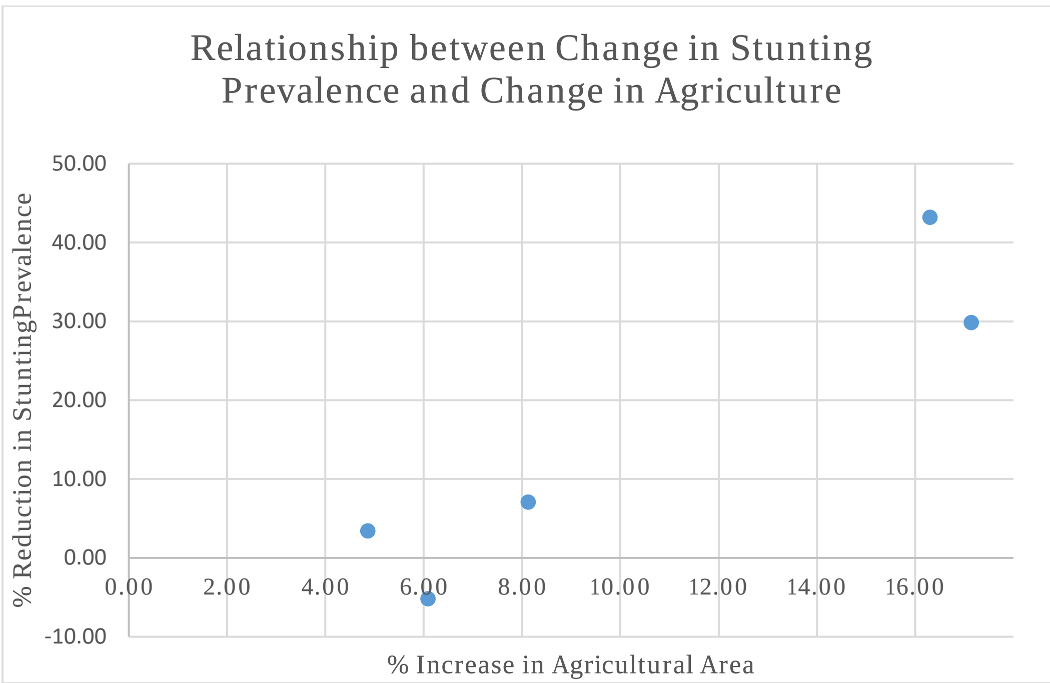


Figure 10: Relationship between Wasting Prevalence among children and Agricultural Land Use in Malawi's Northern Region Districts (1998-2008)

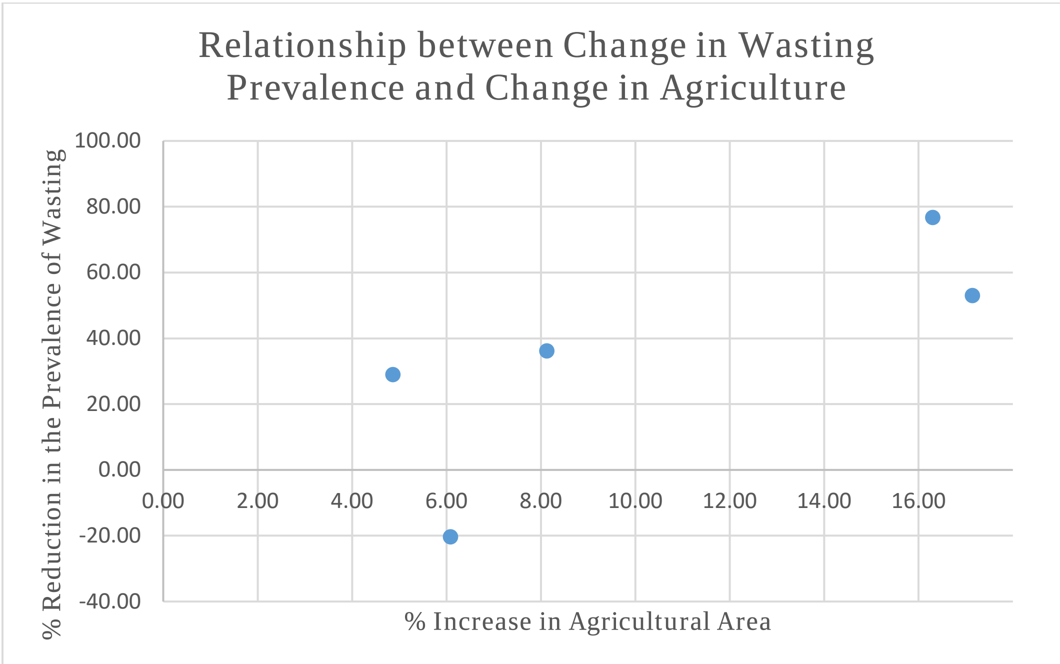


Figure 11: Relationship between Prevalence of Underweight Children and Agricultural Land Use in Malawi's Northern Region Districts (1998-2008)