Fall 2023

Mapping Landcover Change and Population Displacement of Lakshmipur District, Bangladesh due to Riverbank Erosion From 2001-2021: A Geospatial Approach

Israt Jahan

MAPPING LANDCOVER CHANGE AND POPULATION DISPLACEMENT OF LAKSHMIPUR DISTRICT, BANGLADESH DUE TO RIVERBANK EROSION FROM 2001-2021: A GEOSPATIAL APPROACH

by

ISRAT JAHAN

(Under the Direction of Munshi Rahman)

ABSTRACT

Due to the geographical setting, the Lakshmipur district of Bangladesh experiences adverse effects of global climate change that include but are not limited to natural disasters such as tropical cyclones, storm surges, coastal flooding, and riverbank erosion. While riverbank erosion is an implicit impact of climate change, it directly affects human settlements, agricultural activity, and the overall livelihoods of people in this area. Examining the spatiotemporal changes in land cover and population due to riverbank erosion in this region could help us better understand the dynamics of human-environmental relations. This study aimed to classify land cover for every five years from 2001-2021, examine land cover changes in this area from 2001-2021, map populations in Lakshmipur district for every five years from 2001-2021, and estimate population displacement due to riverbank erosion every five years from 2001-2021. Landsat 5 TM 30 m satellite Imagery from 2001-2011 and Landsat 8 OLI 30 m resolution Imagery from 2016-2021 were used to classify landcover and observed landcover changes from 2001-2021. We classified Imagery using the smile random forest classifier in Google Earth Engine and calculated landcover change using the Change Detection Wizard in ArcGIS Pro 3.0. The overall classification accuracies range between 79.04% to 87 %. Our landcover change result suggests that from 2001-2021 fallow land/agricultural land has lost the largest area of land, 341.81 sq km, to homestead forest and waterbody among all the classes. To map populations vector and raster-

based dasymetric mapping approaches were used. The vector and raster-based binary dasymetric mapping and population displacement calculation were carried out in ArcGIS Pro 3.0. The findings suggest that the lowest number of population displacements (1844 people using vector-based approach and 5241 raster-based approach) happened from 2001-2006, and the highest number (86107 people using vector-based approach, 63453 using raster-based approach) were displaced between 2016-2021.

INDEX WORDS: Landcover Change, Population Mapping, Google Earth Engine, Riverbank Erosion, Dasymetric mapping, Population displacement, Lakshmipur District, Bangladesh.

MAPPING LANDCOVER CHANGE AND POPULATION DISPLACEMENT OF LAKSHMIPUR DISTRICT, BANGLADESH DUE TO RIVERBANK EROSION FROM 2001-2021: A GEOSPATIAL APPROACH

by

ISRAT JAHAN

B.URP, Khulna University, Bangladesh, 2019

A Dissertation Submitted to the Graduate Faculty of Georgia Southern University

in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE IN APPLIED GEOGRAPHY

MAPPING LANDCOVER CHANGE AND POPULATION DISPLACEMENT OF LAKSHMIPUR DISTRICT, BANGLADESH DUE TO RIVERBANK EROSION FROM 2001-2021: A GEOSPATIAL APPROACH

by

ISRAT JAHAN

Major Professor:	Munshi Rahman
Committee:	Wei Tu
	Christine M. Hladik

Electronic Version Approved:
December 2023

DEDICATION

To my beloved parents

ACKNOWLEDGMENTS

I am grateful to the Almighty Allah for granting me the patience and strength to complete this thesis successfully.

I am immensely grateful to my primary advisor, Dr. Munshi Rahman, who not only inspired the inception of this project but also encouraged me to use Google Earth Engine, which has been a significant contribution to my work, and I am sincerely thankful for that.

To my remarkable committee members, Dr. Wei Tu and Dr. Christine Hladik, I owe a debt of gratitude. Dr. Wei Tu's rigorous feedback and thought-provoking insights have been instrumental in refining this project. Dr. Christine Hladik's invaluable feedback has also played a crucial role in improving this thesis. I must also extend my appreciation for her unwavering support and guidance throughout my academic journey at Georgia Southern.

Intellectual discussions with Galen Costomiris and Zainab Akinsemoyin were constructive for this thesis. I am thankful that they gave me their valuable time. My appreciation also goes to Syed Sabit Tamim for his assistance with the study area map and for providing associated data.

In times of challenge, my mom and sisters have been sources of mental support. My sister, Nusrat Jahan, has been my mentor and a constant source of inspiration. While I have highlighted some of the key figures who directly contributed to my research and personal growth, some people contributed indirectly to this research. I am thankful for their contributions as well.

TABLE OF CONTENTS

ACKNOWLEDGMENTS 3

TABLE OF CONTENTS 4

LIST OF TABLES 7

LIST OF FIGURES 9

CHAPTER 1 13

INTRODUCTION 13

1.1 Purpose of the Research 13

1.2 Background 14

1.3 Remote Sensing in Population Estimation 19

1.4 Areal Interpolation Techniques 20

1.5 Objectives 24

1.6 Study Area 25

1.7 Significance of the Research 27

1.8 Overview of the research 28

1.9 References 28

1.10 Figures 35

CHAPTER 2 37

METHODOLOGY 37

2.1 Data 37

2.1.1 *Satellite Data for Landcover Classification* 37

2.1.2 *Census Data* 37

2.2 Workflow 39

2.2.1 *Landcover Classification and Change Detection* 39

2.2.2 *Dasymetric Approach based Population Mapping* 41

2.2.2.1 *Vector-Based Dasymetric Population Mapping* 42

2.2.2.2 *Raster-Based Dasymetric Population Mapping* 43

2.2.3 *Volume Preservation and Visual Inspection of Dasymetric Population Maps* 45

2.2.4 *Calculation of Land Loss and Population Displacement due to Riverbank Erosion* ... 46

2.3 References 46

2.4 List Tables and Figures 48

CHAPTER 3 54

RESULT 54

3.1 Image Classification 54

3.1.1 *Random Forest Performance* 54

3.1.2 *Accuracy Assessment* 55

3.2 Change detection 58

3.3 Population Mapping 58

3.3.1 *Dasymetric Mapping Vector and Raster-Based Approach* 58

3.3.2 *Dasymetric Mapping Visual Inspection* 59

3.4 Amount of Land Loss and Population Displacement Due to Riverbank Erosion 59

3.5 References 60

3.6 Tables and Figures 62

CHAPTER 4 89

DISCUSSION 89

4.1 Landcover Classification 89

4.2 Landcover Change 92

4.3 Population Mapping and Estimating Population Displacement Due to Riverbank Erosion. 93

4.4 References 96

CHAPTER 5 98

CONCLUSION 98

5.1 Reference 99

APPENDIX A 100

APPENDIX B 106

APPENDIX C 122

Binary Dasymetric Mapping Vector-Based Approach 122

APPENDIX D 125

Binary Dasymetric Mapping Raster-Based Approach 125

LIST OF TABLES

Table 3.1 Accuracy Assessment of 2001 landcover Classification 67

Table 3.2 Errors of commission, errors of omission of for each landcover class of 2001 67

Table 3.3 Producer's and user's accuracies for each cover class of 2001 landcover Classification 68

Table 3.4 Accuracy Assessment of 2006 landcover Classification (Image acquisition date 02/04/2006). 68

Table 3.5 Errors of commission, errors of omission for each landcover class of 2006 landcover Classification (Image acquisition date02/04/2006). 69

Table 3.6 Producer's and User's accuracy for each landcover class of 2006 landcover Classification (Image acquisition date02/04/2006). 69

Table 3.7 Accuracy Assessment of 2011 landcover Classification (Image acquisition date01/02/2011). 70

Table 3.8 Errors of commission, errors of omission of for each landcover class of 2011 landcover Classification (Image acquisition date01/02/2011) 70

Table 3.9 Producer's and User's accuracy for each landcover class of 2011 landcover Classification (Image acquisition date 01/02/2011) 71

Table 3.10 Accuracy Assessment of 2016 landcover Classification (Image acquisition date 01/15/2016). 71

Table 3.11 Errors of commission, errors of omission of for each landcover class of 2016 landcover Classification (Image acquisition date 01/18/2016). 72

Table 3.12 Producer's and User's accuracy for each landcover class of 2016 landcover Classification (Image acquisition date 01/18/2016). 72

Table 3.13 Accuracy Assessment of 2021 landcover Classification using Landsat 8 OLI (Image acquisition date 17/03/2021 73

Table 3.14 Errors of commission, errors of omission of for each landcover class of 2021landcover Classification using Landsat 8 OLI (Image acquisition date 17/03/2021)..... 73

Table 3.15 Producer's and User's accuracy for each landcover class of 2021 landcover Classification using Landsat 8 OLI (Image acquisition date 17/03/2021). 74

Table 3.16 Landcover change matrix from 2001-2021 using Landsat 5TM and Landsat 8 OLI Imagery (Image acquisition date 01/21/2021 and Acquisition date 17/03/2021)..... 74

Table A.1Satellite data description 100

Table A.2 Number of training points for each landcover class in different years. Training points were drawn inside Google Earth Engine based on our familiarity with the study area. 100

Table A.3Random Forest hyperparameter Settings. 101

Table A.4 Table A4 Spectral Separability analysis of training data for the 2001 image... 101

Table A.5 2006 Spectral separability analysis of training data for the 2006 image.. 101

Table A.6 Spectral separability analysis of training data for the 2011 image.. 102

Table A.7 Spectral separability analysis of training data for the 2016 image.. 102

Table A.8 Spectral separability analysis of training data for the 2021 image.. 103

LIST OF FIGURES

Figure1.1 Study area, Lakshmipur District Bangladesh 35

Figure 1.2 Homestead Forest, Canopies planted at the premises of rural household of Lakshmipur District, Bangladesh..... 36

Figure 2.1 Conceptual diagram of vector-based dasymetric mapping 48

Figure 2.2 Workflow vector-based dasymetric mapping and population displacement..... 49

Figure 2.3 Conceptual diagram of raster based dasymetric mapping. 49

Figure 2.4 Workflow raster based dasymetric mapping and population displacement. 50

Figure 2.5 Choropleth map representing union level population distribution in 2001. The lowest density value was 494 persons/km2 and the highest was 9604 persons/km2. 51

Figure 2.6 Representation of 30 m vector based dasymetric map. The vector representation of 30 m grids could not produce cartographically sound map. 52

Figure 3.1Classified image for year 2001.Overall accuracy 79.04 % 62

Figure 3.2 Classified image for the year 2006. Overall accuracy 81 % 63

Figure 3.3 Classified image for the year 2011. Overall accuracy 87.87 % 64

Figure 3.4 Classified image for year 2016. Overall accuracy 88 % 65

Figure 3.5 Classified image for year 2021. Overall accuracy 83 % 66

Figure 3.6 Dasymetric Population Distribution in 2001, 30 m vector grids were converted to 30 m raster grids..... 75

Figure 3.7 Dasymetric Population Distribution in 2006, 30 m vector grids were converted to 30 m raster grids..... 76

Figure 3.8 Dasymetric Population Distribution in 2011, 30 m vector grids were converted to 30 m raster grids..... 77

Figure 3.9 Dasymetric Population Distribution in 2016, 30 m vector grids were converted to 30 m raster grids..... 78

Figure 3.10 Dasymetric Population Distribution in 2021, 30 m vector grids were converted to 30 m raster grids..... 79

Figure 3.11 Choropleth map representing union-level population distribution in 2001..... 80

Figure 3.12 Dasymetric Population Distribution in 2001 using raster based approach..... 81

Figure 3.13 Dasymetric Population Distribution in 2006 using raster based approach..... 82

Figure 3.14 Dasymetric Population Distribution in 2011 using raster based approach..... 83

Figure 3.15 Dasymetric Population Distribution in 2016 using raster based approach..... 84

Figure 3.16 Dasymetric Population Distribution in 2021 using raster based approach..... 85

Figure A.1 Random forest variable of importance chart for classification 2001 using Landsat 5TM..... 103

Figure A.2 Random forest variable of importance chart for classification 2006 using Landsat 5TM..... 104

Figure A.3 Random forest variable of importance chart for classification 2011 using Landsat 5TM..... 104

Figure A.4 : Random Forest variable of importance chart for classification 2016 using Landsat 8 OLI.. 105

Figure A.5 : Random Forest variable of importance chart for classification 2021 using Landsat 8 OLI. 105

Figure B.1 Shorelines in 2001 and 2006 overlaid with 2001 30 m raster surface of population produced using vector based approach. 106

Figure B.2 Shorelines in 2006 and 2011 overlaid with 2006 30 m raster surface of population produced using vector based approach. 107

Figure B.3 Shorelines in 2011 and 2016 overlaid with 2011 30 m raster surface of population produced using vector based approach. 108

Figure B.4 Shorelines in 2016 and 2021 overlaid with 2021 30 m raster surface of population produced using vector based approach. 109

Figure B.5 Shorelines in 2001 and 2006 overlaid with 2001 30 m raster surface of population produced using raster based approach. 110

Figure B.6 Shorelines in 2006 and 2011 overlaid with 2006 30 m raster surface of population produced using raster based approach. 111

Figure B.7 Shorelines in 2011 and 2016 overlaid with 2011 30 m raster surface of population produced using raster based approach. 112

Figure B.8 Shorelines in 2016 and 2021 overlaid with 2016 30 m raster surface of population produced using raster based approach. 113

Figure B.9 Shorelines in 2001 and 2006 overlaid with 2001 30 m Landcover data..... 114

Figure B.10 Shorelines in 2006 and 2011 overlaid with 2006 30 m Landcover data..... 115

Figure B.11 Shorelines in 2011 and 2016 overlaid with 2011 30 m Landcover data..... 116

Figure B.12 Shorelines in 2016 and 2021 overlaid with 2016 30 m Landcover data..... 117

Figure B.13 Sample code of Landcover classification in 2001 118

Figure B.14 Sample code of Landcover classification in 2001 119

Figure B.15 Sample code of Landcover classification in 2001 120

Figure B.16 Sample code of Landcover classification in 2001 121

CHAPTER 1

INTRODUCTION

1.1 Purpose of the Research

Riverbank erosion impacts both the natural and human environment (Islam et al., 2023; Rahman et al., 2020), causing loss of land, loss of infrastructure, the displacement of people, and significant alterations to local livelihoods (Alam et al., 2017). While previous research works (Islam et al., 2023; Ritu et al., 2023; Mallik et al., 2023) are mainly concentrated on the landcover change due to coastal and riverbank erosion, estimating population displacement along landcover change is important to account for the impacts of riverbank erosion on both the natural and human environment. Therefore, this study focuses on landcover classification, land cover change, and estimates population displacement due to riverbank erosion.

Population data published by the census bureaus are often data that are aggregated at the administrative boundary level and do not match the geographical extent of hazards (Mennis, 2015), limiting the scope of estimating hazard-struck populations. For instance, the number of people being displaced due to an estimated 1km of shoreline shift, as a consequence of riverbank erosion, is difficult to identify using choropleth map as choropleth maps are produced based on census data that distributes population evenly within an administrative boundary. Besides, as census data are aggregated at administrative boundaries, they do not provide actual representations of where people live (Wu & Murray, 2005; Liu, 2003). As a result, using choropleth map produces erroneous estimation. Given that, estimating the number of affected people by natural hazards requires a population distribution map that represents where people

live and can be superimposed with the geographic extent of the hazard to estimate affected populations.

Although several readily available global population datasets, for instance, World Pop (www.worldpop.gov) and Landscan (www.landscan.org), enable small-area population estimation, this research attempts to develop population maps integrating local knowledge with geospatial techniques. Our effort involves producing population maps that are representative of on-ground population distribution. Producing such maps is instrumental in estimating the number of people being relocated due to riverbank erosion. This information is key to disaster planning, preparedness, mitigation, and policy formulation.

1.2 Background

Globally, the coastal regions are vulnerable to climate change-related events such as sea level rise, cyclones, storm surges, saltwater intrusion, and coastal erosion (Nicholls et al., 2007; Islam et al., 2015). Low-lying deltas are likely to be significantly affected by the climate change impacts, especially sea level rise (Nicholls, 2004; Weisse et al., 2014) due to their natural subsidence, reduced sediment deposition caused by construction of artificial dams in the upstream, tectonic plate movement and other geomorphological and anthropogenic stresses (Nicholls, 2007). Bangladesh is a low-lying delta facing toward the Bay of Bengal with more than 700 km of coastline (Haque et al., 2012). Being part of the Bengal Basin, one of the world's largest geo-synclinals (Karim et al., 2008), the country is prone to natural subsidence, experiencing a 2.2 cm yearly subsidence rate (Alam, 1996). Only 10 percent of the country is above 1m mean sea level (Karim et al., 2008). Sea level rise is considered one of the reasons for erosion and shoreline changes in coastal regions (Prasad & Kumar, 2004). Bangladesh lost 1368 sq km from 1973-2016

due to coastal erosion; however, at the same time, 2815.5 sq km has been gained by accretion (Hassan et al., 2017). Besides the coastline, inland coastal areas also experience the adverse effects of climate change that trigger riverbank erosions and saltwater intrusion (Armah et al., 2005).

According to Alam et al. (2017), riverbank erosion in Bangladesh takes away an average of 34 square miles (88 square km) of homestead and farmland annually. The three large rivers called Ganges, Brahmaputra, and Meghna originate in the Himalayas, enter Bangladesh, and merge into the Bay of Bengal through the single outlet of the lower Meghna River (Curtis et al., 2018). During the monsoon flood periods, 138,700 m3/s of water is discharged through the lower Meghna estuary, the highest in the world (Crawford et al., 2019). Analyzing the erosion rate and braiding index, Mahmud et al. (2020) found that from 1988 to 2017, the lower Meghna estuary widened by 49 percent, increasing in width from 8.7km to 13km, which indicates a significant amount of erosion happened over the years.

Lakshmipur district, situated on the eastern bank of the lower Meghna estuary (Figure 1.1), is prone to riverbank erosion due to its location on the Bay of Bengal terminus (Crawford et al., 2020; Paul et al., 2022). Crawford et al. (2021) found that the region experienced 90 meters of net shoreline movement from 1988 to 2018 when the sea level rise rate was 5.7 mm/year (Sarwar, 2013), possibly one of the significant factors that aggravated the erosion event (Crawford et al., 2020). As a consequence of riverbank erosion, the region experienced significant repercussions that include the loss of households, infrastructure, communication systems, population displacement, degradation of agricultural land, and impacts on household livelihoods (Islam et al., 2023).

Several studies (Crawford et al., 2020; Paul et al., 2020; Crawford et al., 2021; Rahman et al., 2021; Paul et al., 2022) have explored the shoreline movement and associated household relocation in the subdistricts of Lakshmipur. Islam et al. (2023) studied shoreline movement landcover change in the Ramgati and Kamalnagar subdistricts of Lakshmipur district. Hasnat et al. (2016) studied the impact of climate change on agriculture and peoples' adaptive strategies in Lakshmipur and Kamalnagar subdistricts. To the best of our knowledge, based on published literature, no studies have quantified the landcover change in the entire Lakshmipur district and estimated the number of population dislocations due to riverbank erosion to date. Therefore, our study aims to fill the research gap by examining land cover change and quantifying population displacement due to riverbank erosion in the Lakshmipur district using a geospatial approach.

Estimating the number of displaced populations helps identify vulnerable coastal communities to address the human-environmental issues that the communities experience. Using traditional field survey methods (e.g., household surveys) to count displaced populations could be time-consuming and costly. Due to the advancement of geospatial science, remote sensing-based population estimation methods have gained popularity for cost efficiency, flexible methodology, and rapid data production capacity (Hossain et al., 2015). Instead of counting the population, researchers have been using remote sensing-derived proxies such as night light, road density, land-use classes, and building footprints to estimate the population (Lo, 1995; Mennis, 2003; Mennis, 2004; Mennis, 2006; Langford, 2006; Fan et al., 2014; Roni & Jia, 2020). For example, previously produced global population datasets that use remote sensing-based proxies include the Landscan global population dataset (landscan.ornl.gov) 1km resolution, Global Rural-Urban Mapping Project (GRUMP) (sedac.ciesin.columbia.edu), 1km resolution, Worldpop (www.worldpop.org) resolution 100m -1km, project. These data sets are great resources; however, they often lack clarity

of method and proper ancillary data source for small area population estimation in the rural areas of developing countries, where spatial census data are unavailable, and indicators of human footprint are different from those of developed countries.

For instance, GRUMP published 1km global gridded population data utilizing nighttime light processed by the National Centers for Environmental Information (www.ngdc.noaa.gov), formerly known as the U.S. National Geophysical Data Center (NGDC), as a proxy of human footprint for urban areas. Using nighttime light data as a proxy for human settlement in both urban and rural areas is problematic, especially in developing countries such as Bangladesh, where rural areas are not luminous during nighttime. Besides, these datasets are in coarse resolution and suitable for global or subcontinental scale analysis (Hoffman et al., 2019). Therefore, population mapping in finer spatiotemporal resolution is crucial to better understand where people live.

Sentinel-2 Multispectral imagery (MSI) and very high resolution 3-5m Planetscope imagery can be effective in mapping rural settlements and disaggregating census populations to those areas. However, the unavailability of freely available historical satellite Imagery limits the opportunity to use sentinel and planet scope Imagery. Using freely available Landsat 30 m imagery settlement mapping is challenging in rural Bangladesh as rural settlements are hard to identify in 30 m Landsat imagery due to dense canopy cover (Islam et al., 2023). In that case, homestead forests, canopies planted at household premises, can be considered as a proxy for human settlement (Figure 1.2) to estimate the population as an alternative to human settlements. Prior studies that mapped populations are primarily in urban (Eicher & Brewer, 2001; Lo, 2003; Mennis, 2003; Mennis, 2004; Mennis, 2009) where satellite imagery encounters less interference to identify human settlement and proxies of human settlements such as night light, impervious surface. Considering rural Bangladesh's geographic and socio-economic context, the homestead forest is a

better representation of human footprints as 30 m resolution Landsat Imagery fails to differentiate settlements due to canopy cover. One study by Hoffman et al. (2019) mapped the location of rural settlements in Rakhine, Myanmar using Landsat 30 m Imagery. The authors used distance to burnt areas as an indicator of human activities in rural areas of Rakhine, Myanmar, to map settlements. They assumed that burning activity was abnormal in deciduous tropical forest ecology; thus, burning activities in Rakhine rural areas indicated human activity, such as burning crop residues. Therefore, in the context of our study area, using homestead forests as an indicator of human settlements can be effective in the absence of any other suitable auxiliary data pertaining to human activities.

Saha (2016) explored the relationship between homestead tree species and farmers' socio-economic condition in Chandpur district, also situated on the eastern bank of the lower Meghna estuary in the north of Lakshmipur. The authors concluded homestead size has a moderate positive relationship with homestead forests. These findings support our assumption that the homestead forests in Lakshmipur represent human settlements as Lakshmipur and Chandpur Districts are situated close to one another and represent similar socio-cultural contexts. However, this assumption has some inherent limitations; for instance, the extent of homestead forests and settlements is not always equal; transferring population in those areas might result in erroneous estimation as areas of homestead forests might be, in some cases, larger than the areas of settlements. However, considering the difficulty of identifying the human settlements from homestead forests in moderate-resolution satellite imagery, we choose homestead forests to indicate settlements in the Lakshmipur district. Therefore, this study attempts to create small area population maps for the Lakshmipur District using the homestead forest as a proxy, which will

eventually help us estimate the number of displaced people every five years from 2001-2021 due to riverbank erosion.

1.3 Remote Sensing in Population Estimation

The use of remote sensing technology and geospatial data in population estimation started in 1950(Lo & Welch, 1977). At that time, the product of average household size based on census data and the number of housing settlements derived from aerial Imagery were used to estimate the population of a particular area (Lo, 1986). Although this method was accurate in estimating the population in smaller areas, this approach was laborious and unsuitable for estimating the population of larger areas due to its dependency on the visual identification of settlements than the classification of aerial Imagery using any raster-based computer software (Wang & Li, 2018).

With the advancement of remote sensing technology and the availability of satellite imagery, researchers started using radiance/reflectance values, extracted through digital image processing techniques, to estimate population using regression models instead of manually counting human settlements and multiplying them with average household size. Iisaka and Hegedus (1982) used the correlation between radiance values and population density using a regression equation between population density and spectral reflectance value to estimate the population density of the Kanto area (including Tokyo Metropolitan). The authors estimated the population density of 500×500 meter meshes using the relationship between population density at census units and radiance values across urban areas. One of the significant difficulties of this process was correlating reflectance/ radiance values from human settlements and urban areas directly to the population density. In 1995, Lo estimated population and dwelling units in 44

tertiary planning units (TPUs) in Kowloon, Hong Kong. This study used a combined method that included spectral radiance of image pixels and pixel counts for residential classes using SPOT multispectral satellite imagery. The author developed four regression equations using the stepwise linear regression method with the following independent variables: average of SPOT bands 1, 2, and 3; average of SPOT band three alone; percentages of pixels classified as high and low-density residential use in each TPU; percentage of pixels classified as high-density residential use in each TPU. By multiplying the estimated population density and dwelling unit density with the area of each TPU, the population in each TPU and dwelling unit was obtained. The study validated its result with actual population count and dwelling units on the ground and that calculating dwelling units was more accurate than calculating population. Besides, pure pixels in the residential land use classes are crucial to estimating the population accurately (Lo, 1995).

1.4 Areal Interpolation Techniques

In contrast to the abovementioned methods, where the population is estimated directly by using parameters estimated from regression analysis of population density and land use classes, in a dasymetric method, the census populations are spatially disaggregated on a finer scale than the parent source, usually census unit, using the relationship between population density and land use classes or other ancillary remote sensing data. Dasymetric mapping is a form of areal interpolation technique to disaggregate census-level populations among known inhabitable areas using ancillary data (Mennis, 2009). The idea of disaggregating population from parent source to target source came from the areal interpolation technique, which uses areal weight derived from area ratio and multiplies with the population density, previously used to disaggregate population data into finer scale without ancillary remote sensing or geospatial data (Mennis, 2003). As an

improvement of the assumptions of the areal weighting, the dasymetric mapping technique gained popularity among geographers as a means of population estimation over the past three decades. Although dasymetric mapping is applicable to estimate any statistical surface data, such as earth surface elevation or air pressure, research works in dasymetric mapping have focused mainly on population, particularly in disaggregating available census population data to make estimates over small areas where census population data are not available. This technique is reliable as the relationship between the variable of interest and the ancillary data provides robustness in estimating low- and high-density areas (Eicher & Brewer, 2001).

Mennis (2009) reviewed traditional and statistical cartographical techniques of dasymetric mapping and explained that the traditional cartographic techniques are binary, three-class, and limiting variables. The most widely used technique, binary dasymetric mapping, uses an area class map as ancillary data classified as habitable and inhabitable areas (e.g., habitable built-up area, not habitable water body) to redistribute the population of the parent map. As a result, the new map shows the nonuniformity of population density within the area of interest that was previously homogeneous for equally distributed populations in habitable and non-habitable areas. Alternatively, land use maps with three or more classes can be used, with different population percentages assigned to different land use categories. For example, considering the total population as 100 percent, in a three-class dasymetric mapping technique, 70 percent of the population is distributed in urban areas, 10 percent near forestlands, and 20 percent in rural areas. When the binary and three-class methods distribute population without creating a threshold for land use classes, the limiting variable method developed by McCleary (1969) is a method of redistributing population into target zones when maintaining the maximum

density threshold for area classes. This method iterates distributing the population among zones according to their threshold value until each zone is below the threshold level.

Researchers have used regression-based approaches to identify the weights of different land use classes in the dasymetric mapping method (Roni et al., 2020; Harvey, 2002). Instead of assigning arbitrary weight to the areal class, regression estimates the relationship between ancillary data and population. The most explored relationships between ancillary data and population distribution are population and urban areas, land use, road density, imperviousness, and other physical, e.g., elevation or socio-economic characteristics. For example, using multispectral satellite Imagery, studies identified a proportional relationship between population growth rate and urban residential areas (Lo & Welch, 1977). Initially, research works (Lo, 1995; Harvey, 2002) used radiance values to identify the built-up residential areas. Later, a volume of research works (Sutton et al., 1997; Prosperie & Eyton, 2000; Lo, 2002) used the Defense Meteorological Satellite Program Operational Linescan System (DMSP-OLS) satellite nighttime Imagery to identify urban residential areas using regression models. The second type of relationship is land-use classes and population distribution. Identifying the varying degree of correlation between population distribution and different land use classes using regression equations, researchers distributed the population based upon the coefficient values into target zones (Weber, 1994; Lo, 2003). Imperviousness is another highly correlated variable with population distribution in urban areas. Sutton et al. (2009) found impervious surfaces (e.g., Roads, buildings, parking lots) as positive correlators of human footprints in urban areas.

The relationship between ancillary data, such as land-use classes, road density, imperviousness, and nighttime light data, are used as exploratory variables to estimate the relationship between population density and those ancillary classes. Then, the parameter of the

regression equation is used as a weight for different regions in the target zones. Although this method is seemingly effective for estimating the population, the problem of a negative parameter in the regression model can indicate a negative population estimation.

For example, when regressed against population density, residential, commercial, and industrial land use classes provide positive parameters, whereas agriculture and forests might provide negative parameters. This problem was first addressed by Yuan et al. (1997). The authors suggest that the solution to this problem is shifting the regression line to the origin and scaling the estimated density to fit with the original choropleth zone counts. In other words, to offset the negative parameter from the regression equation, the absolute value of the negative parameter is added to all the parameters, including the negative parameter. This results in zero parameters for the land use class that had a negative coefficient previously. As a result, a zero population will be assigned to that land use class. However, adding the absolute value of the negative coefficient to the other parameters does not change the relative proportion between them, as all the parameters are being increased by equal value; as a result, the regression equation consists of all positive values, and the ratio between all the parameters stays the same before adjustment (Yuan et al., 1997). Later, Liu et al. (2008) made further improvements to a regression-based analysis by accounting for the spatial dependencies of the regression residual using area-to-point residual kriging. The authors used Root Mean Square Error (RMSEs) and Mean Absolute Error (MAE) values to observe the difference between the estimated and observed population in block level population and observed reduction of error values when opting for area-to-point residual kriging.

Recent studies have worked on improved population estimation methods using numerous geospatial and remote sensing datasets as ancillary data. For example, Stevens et al. (2015) used

semi-automatic dasymetric modeling using land cover types, MODIS-derived net primary productivity (NPP), and nighttime lights as ancillary data. The authors used a random forest algorithm to identify the contribution of these variables to population distribution weighting. Deriving the weights from the random forest technique, they applied them in dasymetric mapping. They used root mean square (RMSE) values to identify the error level in the dasymetric map. Su et al. (2010) developed multilayer multiclass dasymetric mapping to divide the regional population into smaller spatial areal units in the Taipei metropolitan area of Taiwan. The supporting information, such as remote sensing Imagery, land-use zoning, geography, transportation, and facility accessibility, was added in a hierarchical sequence according to how it correlated to the characteristics of population distribution. For instance, the regions were initially divided into building and nonbuilding areas using remotely sensed SPOT satellite Imagery, and then the layer was divided based on land-use zoning. Grippa et al. (2019) used the random forest technique to disaggregate census populations in urban areas. They used the random forest technique with both moderate and high-resolution satellite Imagery to prepare dasymetric population maps. Their findings based on RMSE values of these two approaches reveal that the random forest algorithm with high-resolution satellite imagery accurately redistributes the population compared to moderate-resolution imagery.

1.5 Objectives

The objectives of this study are: 1) to classify landcover for every five years between 2001-2021 and examine twenty years of land cover change in Lakshmipur District from 2001-2021, 2) to map populations in Lakshmipur district for every five years from 2001-2021 in Lakshmipur district, and 3) to estimate the population displacement due to riverbank erosion for every five years from 2001-2021in the Lakshmipur district in Bangladesh. Several research

works have explored the amount of land loss in the eastern and western banks of the estuary due to erosion (Anwar et al., 2021; Crawford et al., 2021; Hussain et al., 2014); however, the estimation of displaced households and population has remained unexplored. Although Crawford et al. (2021) investigated the location of households that disappeared due to erosion, they did not use any robust methodology to estimate the number of displaced populations in this event. The authors roughly estimated the number of displaced populations using Worldpop data (www.worldpop.org), which needs to include publications that describe methods and data used in gridded population estimation for data-scarce environments like Bangladesh.

This study is an attempt to fill this research gap by mapping the population into 30 m grids as well as estimating the number of displaced populations along the riverbank. To achieve that goal, we classified the land cover of Lakshmipur district for every five years between 2001-2021 using landcover classes of every five years from 2001-2021; we mapped populations into 30 m grids and finally estimated the displaced population due to river erosion by overlaying shorelines extracted by Crawford et al. (2021). We also observe changes in landcover classes of Lakshmipur district from 2001-2021 and explore the drivers of landcover change based on existing literature.

1.6 Study area

The study area is Lakshmipur district/zilla, situated on the eastern bank of the lower Meghna estuary,22° 57' 0" N and 90° 50' 0" E (Figure 1.1), with an area of 1455.95 sq km. According to the Bangladesh Bureau of Statistics (BBS), in 2011, the total population of the zilla was 1,729,188. The District/Zilla consists of 5 subdistricts/upazillas called Kamalnagar, Raipur, Ramganj, and Ramgati and is further divided into 62 unions, the smallest administrative unit.

These unions are not defined by certain attributes such as measure of area since their boundaries are determined by political boundaries (BBS, 2011).

Due to its location in the central coastal estuarine zone of Bangladesh, the district experiences the severe consequences of climate change and sea level rise, such as tropical cyclones coupled with frequent storm surges, flood events, salinity intrusion, and riverbank erosion (Murshed et al., 2022). The sea level rise rate at Charchanga station (22°06' 0" N 91°30' 0" E) is 5.7 mm/year (Sarwar, 2013), and an additional 0.2 m of mean sea level rise will inundate the areas adjacent to the lower Meghna estuary (Sarwar et al., 2007; Murshed et al., 2022). The region is frequently flooded by storm surges driven by tropical storms (Dasgupta et al., 2010). Besides, the area regularly experiences high diurnal tides, which worsen in monsoon (Jun-July) with heavy rainfall (Paul et al., 2020). Thus, river erosion intensifies during the monsoon in this region (Rahman et al., 2015; Rashid & Paul, 2014). The 80km stretch of the eastern bank of the lower Meghna estuary experienced an average end point rate (EPR) of -36 m between 1988 and 2018, indicating that erosion predominated over accretion (Crawford et al., 2021). The north and south regions of the 80km stretch experienced erosion over the years, and the two most affected Unions, Ramgati and Kamalnagar, have an annual erosion rate of -88.7m/y (Crawford et al., 2021; Paul et al., 2020).

Erosion has substantially impacted the lives of the residents along the riverbank. According to Crawford et al. (2021), around 720000 people lived within 5 km of the riverbank in 2008. Their result shows that among 407 surveyed households' forty-three percent of surveyed households lost their property to riverbank erosion and relocated elsewhere since 2008. Fifty-three percent of households expressed fears of losing households, and agricultural land to erosion and being relocated.

1.7 Significance of the research

Climate change-induced hazards, cyclones, storm surges, saltwater intrusion, and riverbank erosion impact the coastal region's natural and human environment (Islam et al., 2021). Sea level rise, coastal erosion, saltwater intrusion, and storm surges reduce crop production and create food Insecurity (Huq et al., 2001; Parry et al., 2014). In response to the negative impacts, the inhabitants alter their livelihood strategies, indiscriminately impacting land uses, such as fragmenting agricultural land into waterbodies for shrimp cultivation (Pokrant, 2014). Farmers even changed their farming practices and switched to cultivating saline-tolerant species (Swapan et al., 2011). Besides, due to riverbank erosion, people relocate from the coast to farther inland, which adds extra pressure on agricultural land (Paul et al., 2021). Examining land cover change could help us better understand the implicit impact of climate change on land cover. It could also help agricultural and land use planning in rural areas to reduce environmental degradation.

Natural hazards such as cyclones, storm surges, and floods have different extents based on the event's intensity and impacts varying geographical extent (Hoffman et al., 2019). Choropleth maps evenly distribute the population in spatial units that do not represent where people live and limit the opportunity to estimate affected people by the hazard (Maantay & Maroko, 2009). On the other hand, dasymetric mapping disaggregates populations in small spatial units that are part of human-inhabited areas (Maantay & Maroko, 2017). As a result, Dasymetric maps provide a correct representation of where people live and make estimating hazard-struck people convenient.

Estimating population displacement due to riverbank erosion in Lakshmipur District has several advantages. One of the significant benefits is early planning and policy formulation to reduce the impact of riverbank erosion on affected communities. Others include rehabilitation and

estimating loss of assets. Apart from that, this study also outlines the improvement of methodology, which can guide future studies.

1.8 Overview of the research

This study is organized into five chapters. Chapter One introduces the research objectives, justifies the purpose and importance of this study, reviews the methods applied in the literature, and identifies the suitable method for our study. Chapter two details the data and methods used in this study and tests that have been performed to validate results. Chapter three presents the results, and chapter four includes a discussion to support the findings. Chapter Five concludes with an overview of our findings, limitations, and guidelines for future study.

1.8 References

Alam K, Rahman MH (2014) Women in natural disasters: A case study from southern coastal region of Bangladesh. Int J Disaster Risk Reduction 8: 68-82.

Anwar, M. S., & Rahman, K. (2021). The spatiotemporal shore morphological changes at east Bhola Island in Meghna Estuary of Bangladesh's central coast. *Regional Studies in Marine Science*, *47*, 101937.

Anwar, M. S., & Rahman, K. (2021). The spatiotemporal shore morphological changes at east Bhola Island in Meghna Estuary of Bangladesh's central coast. *Regional Studies in Marine Science*, *47*, 101937.

Armah, A. K., Wiafe, G., & Kpelle, D. G. (2005). Sea-level rise and coastal biodiversity in West Africa: a case study from Ghana. *Climate change and Africa*, 204-217.

Barbier, E.B., Koch, E.W., Silliman, B.R., Hacker, S.D., Wolanski, E., Primavera, J., Granek, E.F., Polasky, S., Aswani, S., Cramer, L.A. and Stoms, D.M., 2008. Coastal ecosystem-based management with nonlinear ecological functions and values. *science*, *319*(5861), pp.321-323.

BBS (2011). Statistical yearbook of Bangladesh, Ministry of Planning, Dhaka, Bangladesh.

Crawford, T. W., Islam, M. S., Rahman, M. K., Paul, B. K., Curtis, S., Miah, M. G., & Islam, M. R. (2020). Coastal erosion and human perceptions of revetment protection in the Lower Meghna Estuary of Bangladesh. *Remote Sensing*, *12*(18), 3108.

Crawford, T. W., Rahman, M. K., Miah, M. G., Islam, M. R., Paul, B. K., Curtis, S., & Islam, M. S. (2021). Coupled adaptive cycles of shoreline change and households in deltaic Bangladesh: analysis of a 30-year shoreline change record and recent population impacts. *Annals of the American Association of Geographers*, *111*(4), 1002-1024.

Curtis, S., Crawford, T., Rahman, M., Paul, B., Miah, M. G., Islam, M. R., & Patel, M. (2018). A hydroclimatological analysis of precipitation in the Ganges–Brahmaputra–Meghna River Basin. *Water*, *10*(10), 1359.

Dasgupta, S., Huq, M., Khan, Z. H., Ahmed, M. M. Z., Mukherjee, N., Khan, M., & Pandey, K. D. (2010). Vulnerability of Bangladesh to cyclones in a changing climate: Potential damages and adaptation cost. *World Bank Policy Research Working Paper*, (5280).

Fan, J., Ma, T., Zhou, C., Zhou, Y., & Xu, T. (2014, August 22). Comparative Estimation of Urban Development in China's Cities Using Socioeconomic and DMSP/OLS Night Light Data. *Remote Sensing*, *6*(8), 7840–7856. https://doi.org/10.3390/rs6087840

Grippa, T., Linard, C., Lennert, M., Georganos, S., Mboga, N., Vanhuysse, S., ... & Wolff, E. (2019). Improving urban population distribution models with very-high resolution satellite information. *Data*, *4*(1), 13.

Haque, U., Hashizume, M., Kolivras, K. N., Overgaard, H. J., Das, B., & Yamamoto, T. (2012). Reduced death rates from cyclones in Bangladesh: what more needs to be done?. Bulletin of the World Health Organization, 90, 150-156.

Harley, M.D., Turner, I.L., Kinsela, M.A., Middleton, J.H., Mumford, P.J., Splinter, K.D., Phillips, M.S., Simmons, J.A., Hanslow, D.J. and Short, A.D., 2017. Extreme coastal erosion enhanced by anomalous extratropical storm wave direction. *Scientific reports*, *7*(1), pp.1-9.

Harvey, J. T. (2002). Estimating census district populations from satellite imagery: Some approaches and limitations. *International journal of remote sensing*, *23*(10), 2071-2095.

Harvey, J. T. (2002). Estimating census district populations from satellite imagery: Some approaches and limitations. *International journal of remote sensing*, *23*(10), 2071-2095.

Hasnat, M. A., Hossain, N., Muhibbullah, M., Sarwar, M. D., & Shormin, T. (2016). Impacts of climate change on agriculture and changing adaptive strategies in the coastal area of Lakshmipur district, Bangladesh. *Current World Environment*, (3), 700.

Hassan, S. T., Syed, M. A., & Mamnun, N. (2017). Estimating erosion and accretion in the coast of Ganges-Brahmaputra-Meghna Delta in Bangladesh. In *6th Int. Conference on Water & Flood Management* (pp. 115-124).

Hoffman-Hall, A., Loboda, T. V., Hall, J. V., Carroll, M. L., & Chen, D. (2019). Mapping remote rural settlements at 30 m spatial resolution using geospatial data-fusion. Remote Sensing of Environment, 233, 111386.

Hossain, M.S., Bujang, J.S., Zakaria, M.H. and Hashim, M., 2015. The application of remote sensing to seagrass ecosystems: an overview and future research prospects. *International Journal of Remote Sensing*, *36*(1), pp.61-114.

Hosseiny, B., Abdi, A. M., & Jamali, S. (2022). Urban land use and land cover classification with interpretable machine learning–A case study using Sentinel-2 and auxiliary data. *Remote Sensing Applications: Society and Environment*, *28*, 100843

Huq, S. (2001). Climate change and Bangladesh. Science, 294(5547), 1617-1617.

Hussain, M. A., Tajima, Y., Gunasekara, K., Rana, S., & Hasan, R. (2014, June). Recent coastline changes at the eastern part of the Meghna Estuary using PALSAR and Landsat Imagery. In *IOP Conference Series: Earth and Environmental Science* (Vol. 20, No. 1, p. 012047). IOP Publishing.

Iisaka, J., & Hegedus, E. (1982). Population estimation from Landsat imagery. *Remote sensing of environment*, *12*(4), 259-272.

Islam, M. A., Hossain, M. S., & Murshed, S. (2015). Assessment of coastal vulnerability due to sea level change at Bhola Island, Bangladesh: using geospatial techniques. Journal of the Indian Society of Remote Sensing, 43, 625-637.

Islam, M. N., van Amstel, A., Islam, M. N., Tamanna, S., van Amstel, A., Noman, M., ... & Ghosh, A. (2021). Climate change impact and comprehensive disaster management approach in Bangladesh: a review. Bangladesh II: climate change impacts, mitigation and adaptation in developing countries, 1-39.

Islam, M. S., Crawford, T. W., & Shao, Y. (2023). Evaluation of predicted loss of different land use and land cover (LULC) due to coastal erosion in Bangladesh. *Frontiers in Environmental Science*, *11*, 479.

Khan, N. I., Elahi, F., & Rana, M. A. R. (2015). A study on the effects of global warming in Bangladesh. *International Journal of Environmental Monitoring and Analysis, 3 (3), 118, 121*.

Langford, M. (2006). Obtaining population estimates in non-census reporting zones: An evaluation of the 3-class dasymetric method. Computers, environment and urban systems, 30(2), 161 180.

Liu, X. (2003). Estimation of the spatial distribution of urban population using high spatial resolution satellite imagery. University of California, Santa Barbara.

Liu, X. H., Kyriakidis, P. C., & Goodchild, M. F. (2008). Population-density estimation using regression and area-to-point residual kriging. *International Journal of geographical information science*, *22*(4), 431-447.

Lo, C. P. (1986). Accuracy of population estimation from medium-scale aerial photography. In *American Congress on Surveying and Mapping and American Society for Photogrammetry and Remote Sensing, Annual Convention, Washington, DC*,(1-10).

Lo, C. P. (2002). Urban indicators of China from radiance-calibrated digital DMSP-OLS nighttime Imagery. *Annals of the Association of American Geographers*, *92*(2), 225-240.

Lo, C. P. (2003). Zone-based estimation of population and housing units from satellite-generated land use/land cover maps. *Remotely sensed cities*, *157*.

Lo, C.P. and Welch, R., (1977). Chinese urban population estimates. *Annals of the Association of American Geographers*, *67*(2),246-253.

Lo, C.P.(1995). Automated population and dwelling unit estimation from high-resolution satellite Imagery: a GIS approach. *Remote Sensing*, *16*(1),17-34.

Lu, D., Weng, Q., & Li, G. (2006). Residential population estimation using remote sensing derived impervious surface approach. *International Journal of Remote Sensing*, *27*(16), 3553 3570.

Maantay, J., & Maroko, A. (2009). Mapping urban risk: Flood hazards, race, & environmental justice in New York. Applied geography, 29(1), 111-124.

Maantay, J., & Maroko, A. (2017). Areal interpolation and dasymetric mapping. The Routledge Handbook of Environmental Justice, 190.

Mahmud, M. I., Mia, A. J., Islam, M. A., Peas, M. H., Farazi, A. H., & Akhter, S. H. (2020). Assessing bank dynamics of the Lower Meghna River in Bangladesh: an integrated GISDSAS approach. *Arabian Journal of Geosciences*, *13*, 1-19.

Mallick, R. H., Bandyopadhyay, J., & Halder, B. (2023). Impact assessment of river bank erosion in the lower part of Mahanadi River using geospatial sciences. Sustainable Horizons, 8, 100075.

Mallick, R. H., Bandyopadhyay, J., & Halder, B. (2023). Impact assessment of river bank erosion in the lower part of Mahanadi River using geospatial sciences. Sustainable Horizons, 8, 100075.

McCleary Jr, G. F. (1969). *The dasymetric method in thematic cartography*. The University of Wisconsin-Madison.

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.

Mennis, J. (2015). Increasing the accuracy of urban population analysis with dasymetric mapping. Cityscape, 17(1), 115-126.

Mennis, J. (2016). Dasymetric Mapping. *International Encyclopedia of Geography: People, the Earth, Environment and Technology: People, the Earth, Environment and Technology*, 1 10.

Mentaschi, L., Vousdoukas, M.I., Pekel, J.F., Voukouvalas, E. and Feyen, L., (2018). Global long term observations of coastal erosion and accretion. Sci. Rep. 8, 12876.

Murshed, S., Griffin, A. L., Islam, M. A., Wang, X. H., & Paull, D. (2022). Assessing multi climate-hazard threat in the coastal region of Bangladesh by combining influential environmental and anthropogenic factors. Progress in Disaster Science, 16, 100261.

Nicholls, R. J. (2004). Coastal flooding and wetland loss in the 21st century: changes under the SRES climate and socio-economic scenarios. Global environmental change, 14(1), 69-86.

Nicholls, R. J., Wong, P. P., Burkett, V., Codignotto, J., Hay, J., McLean, R., & Saito, Y. (2007). Coastal systems and low-lying areas.

Pal, M. and Mather, P.M., 2001, November. Decision tree based classification of remotely sensed data. In *Proceedings of 22nd Asian conference on remote Sensing* (Vol. 5, p. 9).

Parry, M. L., Rosenzweig, C., Iglesias, A., Livermore, M., & Fischer, G. (2004). Effects of climate change on global food production under SRES emissions and socio-economic scenarios. *Global environmental change*, *14*(1), 53-67.

Paul BK (2009) Why relatively fewer people died? The case of Bangladesh's cyclone Sidr. Nat Hazar 50(2): 289-304. 9. Alam K, Rahman MH (2014) Women in natural disasters: A case study from southern coastal region of Bangladesh. Int J Disast Risk Reduc 8: 68-82.

Paul, B. K., Rahman, M. K., Crawford, T., Curtis, S., Miah, M. G., Islam, M. R., & Islam, M. S. (2020). Explaining mobility using the Community Capital Framework and Place Attachment concepts: A case study of riverbank erosion in the Lower Meghna Estuary, Bangladesh. *Applied Geography*, *125*, 102199.

Paul, B. K., Rahman, M. K., Crawford, T., Curtis, S., Miah, M. G., Islam, R., & Islam, M. S. (2021). Coping strategies of people displaced by riverbank erosion in the lower Meghna estuary. Living on the Edge: Char Dwellers in Bangladesh, 227-239.

Paul, B. K., Rahman, M. K., Lu, M., & Crawford, T. W. (2022). Household migration and intentions for future migration in the climate change vulnerable lower Meghna estuary of coastal Bangladesh. *Sustainability*, *14*(8), 4686.

Pokrant, B. (2014). Brackish water shrimp farming and the growth of aquatic monocultures in coastal Bangladesh. Historical perspectives of fisheries exploitation in the Indo-Pacific, 107-132.

Prasad, D. H., & Kumar, N. D. (2014). Coastal erosion studies—a review. International Journal of Geosciences, 2014.

Prosperie, L., & Eyton, R. (2000). The relationship between brightness values from a nighttime satellite image and Texas county population. *Southwestern Geographer*, *4*, 16-29.

Rahman, M. K., Crawford, T. W., Paul, B. K., Sariful Islam, M., Curtis, S., Giashuddin Miah, M., & Rafiqul Islam, M. (2021). Riverbank erosions, coping strategies, and resilience thinking

of the Lower-Meghna River Basin community, Bangladesh. *Climate Vulnerability and Resilience in the Global South: Human Adaptations for Sustainable Futures*, 259-278.

Rahman, M. M., Ghosh, T., Salehin, M., Ghosh, A., Haque, A., Hossain, M. A., ... & Hutton, C. W. (2020). Ganges-Brahmaputra-Meghna delta, Bangladesh and India: a transnational mega-delta. *Deltas in the Anthropocene*, 23-51.

Ritu, S. M., Sarkar, S. K., & Zonaed, H. (2023). Prediction of Padma river bank shifting and its consequences on LULC changes. Ecological Indicators, 156, 111104.

Roni, R., & Jia, P. (2020). An Optimal Population Modeling Approach Using Geographically Weighted Regression Based on High-Resolution Remote Sensing Data: A Case Study in Dhaka City, Bangladesh. *Remote Sensing*, *12*(7), 1184.

Saha, t. K. (2016). *Homestead tree species diversity and it's impact on socioeconomic condition of farmers in chandpur district of bangladesh* (doctoral dissertation, dept. Of agroforestry & environmental science).

Sarwar, G. M., & Khan, M. H. (2007). Sea level rise. A threat to the coast of Bangladesh. *Internationales Asienforum*, *38*(3/4), 375.

Stevens, F. R., Gaughan, A. E., Linard, C., & Tatem, A. J. (2015). Disaggregating census data for population mapping using random forests with remotely-sensed and ancillary data. *PloS one*, *10*(2), e0107042.

Su, M. D., Lin, M. C., Hsieh, H. I., Tsai, B. W., & Lin, C. H. (2010). Multi-layer multi-class dasymetric mapping to estimate population distribution. *Science of the Total Environment*, *408*(20), 4807-4816.

Sutton, P. C., Anderson, S. J., Elvidge, C. D., Tuttle, B. T., & Ghosh, T. (2009). Paving the planet: impervious surface as proxy measure of the human ecological footprint. *Progress in Physical Geography*, *33*(4), 510-527.

Sutton, P., Roberts, D., Elvidge, C., & Meij, H. (1997). A comparison of nighttime satellite imagery. *Photogrammetric Engineering & Remote Sensing*, *63*(11), 1303-1313.

Swapan, M. S. H., & Gavin, M. (2011). A desert in the delta: participatory assessment of changing livelihoods induced by commercial shrimp farming in Southwest Bangladesh. Ocean & Coastal Management, 54(1), 45-54.

Wang, L. & Li, X. (2018) Population Estimation With Remote Sensing. In J. V. Craig & R. L. Smith (Eds.) Comprehensive Remote Sensing. (Volume 1., pp. 59-65). Netherlands: ElsevierScience.https://www.google.com/books/edition/Comprehensive_Remote_Sensing/x9skDwAAQBAJ?hl=en&gbpv=0

WEBER, C. (1994). Per-zone classification of urban land cover for urban population estimation. *In Environmental remote sensing from regional to global scales*, 142-148.

Weisse, R., Bellafiore, D., Menéndez, M., Méndez, F., Nicholls, R. J., Umgiesser, G., & Willems, P. (2014). Changing extreme sea levels along European coasts. *Coastal engineering*, *87*, 4-14.

Wu, C., & Murray, A. T. (2005). A cokriging method for estimating population density in urban areas. Computers, Environment and Urban Systems, 29(5), 558-579.

Yuan, Y., Smith, R. M., & Limp, W. F. (1997). Remodeling census population with spatial information from Landsat TM imagery. *Computers, Environment and Urban Systems*, *21*(3-4), 245-258.

1.10 Figures

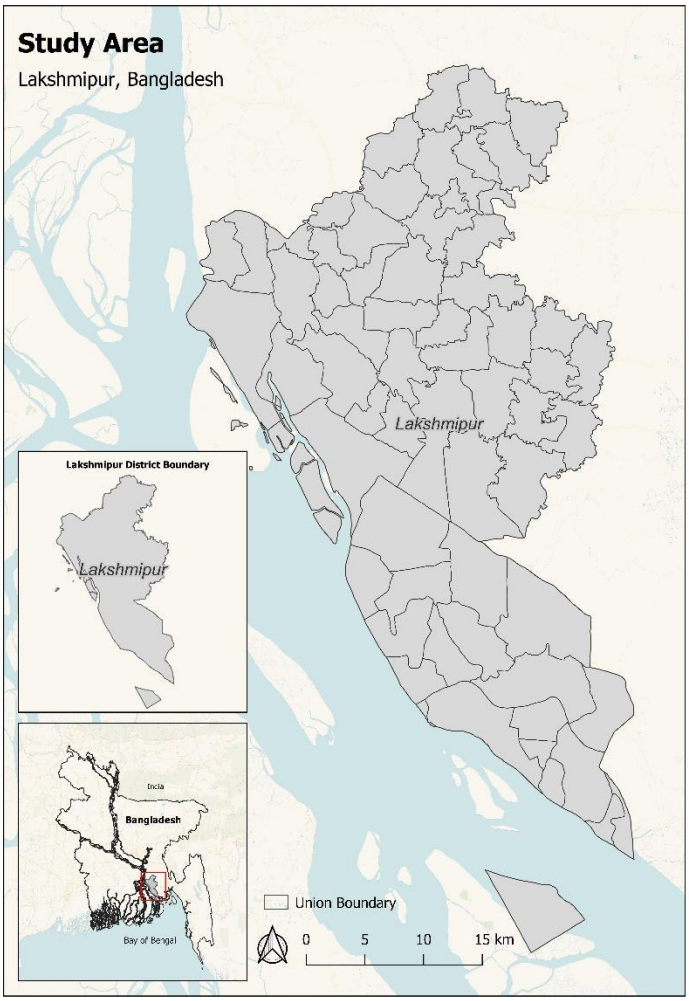


Figure 1.1 Study area, Lakshmipur District Bangladesh

Figure1.2 Homestead Forest, Canopies planted at the premises of rural households of Lakshmipur District, Bangladesh. Homestead forests have been used as proxies of human settlements. This image is a snapshot of Google earth image and does not have any spectral bands associated with it.

CHAPTER 2

METHODOLOGY

We classified landcover for every five years from 2001-2021 using satellite imagery. To examine twenty years of landcover change from 2001-2021 we subtracted the 2001 and 2021 post-classification Imagery to calculate the change from 2001-2021. To map the population, we used a dasymetric mapping approach to downscale the census data, National Population and Housing Census, Union statistics 2001, 2011, and 2021, and inter-census years 2006 and 2016 data. The following sections outline data and methods we used to produce a population map of Lakshmipur district that later aided in estimating people displacement by river erosion.

2.1 Data

2.1.1 *Satellite Data for Landcover Classification*

Landcover data are integral to dasymetric mapping (Zandbergen et al., 2010; Mennis, 2009). Several studies (Sleeter, 2004; Baynes, 2022) used national landcover data for population disaggregation in the conterminous United States of America (USA). Due to non-existence of landcover data set in our study area, we produced landcover classifications for our study period (2001, 2006, 2011, 2016, 2021) using moderate spatial resolution (30 m) Landsat 5 Thematic Mapper (TM) data and Landsat 8 Operational Land Imager (OLI) data (See Appendix A, Table A.1).

2.1.2 *Census Data*

Bangladesh Bureau of Statistics provides union-level tabular population data every ten years (BBS, 2011). We used union-level tabular census data, the lowest administrative unit of

published population data for our analysis unit, for 2001 and 2011, and unpublished population data, official data provided by BBS upon request, for 2021. For inter-census years, the fifth year between two consecutive census years, 2006 and 2016, we interpolated the population, calculating the growth rate of each union. We used the following equation in the BBS 2011 report to calculate growth rates and population for each union.

$$P = P_o * (1+r)^t \quad \text{(Eq 2.1; BBS, 2011)}$$

$$\text{So, Growth rate } r = (P/P_0)^{1/t} - 1 \quad \text{(Eq. 2.2)}$$

Where, P = Target year population

r = Growth rate

P_o = Base year Population

t = Time between the base year and target year

We calculated growth rates between two consecutive census years and used those growth rates to calculate inter-census year populations. For instance, to calculate populations in 2006, we calculated the growth rate between 2001 and 2011 using equation 2.2. Next, we calculated the population in 2006 using 2001's population as the base year population and the growth rate between 2001 and 2011 as input in equation 2.1. Similarly, we calculated the population for the year 2016. As our census data is in tabular form, we created union-level spatial data, downloaded from humanitarian data exchange (The Humanitarian Data Exchange, n.d.), by joining tabular data with union feature classes (Figure 2.2) using ArcGIS pro 3.0 (www.esri.com).

All our spatial data were reprojected to the Bangladesh Transverse Mercator (BTM) projection system before all analyses. Table 2.1 represents the census population (BBS 2001,

2011, 2021) and inter-census year population data (interpolated using equations 1 and 2). We only considered the residential population count from the census data. Besides, there was missing data in some unions in the upper right corner of the district due to the discrepancy between population data and union boundary. We filled the gaps in those data using the areal interpolation method in Geostatistics Wizard of ArcGIS pro 3.0 (www.esri.com).

2.2 Workflow

2.2.1 *Landcover Classification and Change Detection*

Image classification was performed in Google Earth Engine (Google Earth Engine), a cloud-based geospatial analysis and data visualization platform (Gorelick et al., 2017). With petabytes of freely available satellite Imagery, it provides access to Google's hosted cloud computing environment, reducing dependency on user computers (Mutanga et al., 2019; Zhao et al., 2021). Its Application Programming Interface (API) provides the flexibility of Java and Python scripting with libraries available for geospatial analysis (Zhao et al., 2021).

We used Landsat 5 Thematic Mapper (TM) collection 2 Tier 1 top of atmosphere reflection Imagery (TOA) for the years 2001-2011 and Landsat 8 Operational Land Imager (OLI) collection 2 Tier 1 top of atmosphere reflection Imagery (TOA) Imagery for the years 2016-2021(Table A1, Appendix A) from GEE's publicly available online data catalog. Collection 2 tier 1 has improved the level of geometric correction and lower RMSEs as these datasets have been corrected with improved algorithms, ground control points, and elevation datasets (Lubke et al., 2021). The image search was limited between January and February to obtain cloud-free Imagery (See Table A1, Appendix A) of our study area. Based on our

familiarity with the study area, we drew our samples for four classes: homestead forest, fallow land, other vegetation, and waterbody.

We used the random forest classification algorithm in Google Earth Engine. A random forest algorithm is an ensemble of decision trees that uses bootstrapping and aggregating/Bagging. Bootstrap sampling ensures the use of random subsets of variables and data with replacement when generating trees and then aggregates the results based on the majority vote of all trees (Breiman, 1996). It works well in classifying unknown data since it uses multiple decision trees and introduces randomness in the sampling process (Breiman, 1996).

The random forest classifier inside Google Earth Engine allows manual hyperparameter setting. Although hyperparameters set with software defaults yield optimum results, tweaking parameters can often increase the model performance (G & Sumathi, 2020; Kilany et al., 2021). As tuning hyperparameters and changing the ratio of training and validation data can improve the overall performance of the classifier (Probst et al., 2019), we manually tuned the number of variables per split, bag fraction, and the number of trees and changed the proportion of training and testing samples (Table A3, Appendix A). We determined our best classification result (Table A3, Appendix A) considering the combinations that provided lower out-of-bag error, which is the number of accurately predicted elements from out-of-bag samples and higher validation accuracy. The classification algorithm also provided us with variables of importance based on the Gini impurity index that calculates the probability of misclassification of a randomly selected element in a node, and provides us with which bands/variables are important/effectively split data into different classes. It analyzes how the reduction in impurity is obtained, using a particular band to split the data. The higher the reduction in impurity, the more important the

band is. We calculated the relative reduction in impurity for a band by summing up the weighted impurity reductions across all nodes. The higher the sum, the more important the band is considered for classification.

For each study year, we reserved 100 ground reference points for external validation. Those points were generated using the stratified random sampling technique in ArcGIS pro3.0. As ground reference Imagery, we used Google Earth Imagery (spatial resolution ranges 0.3m Maxar-30 m Landsat) for 2006, 2011, 2016, and 2021 and a 10m resolution SPOT image for 2001. Finally, confusion matrices were generated using reserved validation data with overall accuracy, error of commission, error of omission, and user's and producer's accuracy.

For change detection, we used ArcGIS Pro 3.0 (www.esri.com). We calculated thematic change from 2001 to 2021 using Landsat 5 TM image for 2001 and Landsat 8 OLI image for 2021. The change detection provided from class to class table that has been used to create a change matrix and calculate net gain and loss of landcover classes.

2.2.2 *Dasymetric Approach-based Population Mapping*

As we only used homestead forest as a proxy of human settlement binary dasymetric mapping, where the land cover class was reclassified as habitable/non-habitable, is appropriate. We explored both vector and raster-based dasymetric approaches to map population and calculate displacement. We explored vector and raster-based approaches to disaggregate union-level population data and calculated displacement. Appendix C and D contain the tools and techniques used for vector and raster-based dasymetric mapping.

2.2.2.1 *Vector-Based Dasymetric Population Mapping*

The vector-based dasymetric approach is based on basic areal interpolation adopted from Goodchild and Lam (1980). Figure 2.1 explains the conceptual framework of vector-based dasymetric mapping. The authors applied the following equation (equation 2.3) using polygon overlay in GIS to transfer population count from the census tracts to planning districts of London, Ontario, without any ancillary data. In our case, we used homestead forest as ancillary data and transferred the population from union to homestead forest polygons and then to 30 m grids. This means that at first, our source zones were unions and target zones were homestead forests, and to transfer the population in 30 m grids, homestead forest polygons were source zones, and 30 m grids were target zones. As we used a vector-based approach, we initially converted 30 m raster landcover maps into vectors, and to transfer population in 30 m grid homestead forests, we created 30 m fishnets. Figure 2.2 explains the workflow diagram of vector-based dasymetric mapping, and Appendix C contains the tools and detailed processes used to implement the equation in GIS.

Population in each target zone t, $Pt = \sum Ats / As * Ps$ ……(Eq 2.3)

Where Ats is the intersection area between the source and target zone.

As is the Area of the source zone.

Ps is the Population in the source zone.

We started by converting our landcover maps into polygons and extracting our homestead forest class from the landcover map. Intersect tool was used to join union-level population data

with our habitable layer homestead forest. To assign populations from source zones/Union level to the target zones/ homestead forest polygons from the source zones/unions, we applied equation 2.3.

To transfer the Population from Homestead forests to 30 m grids, we followed a similar process by creating 30 m grids for our study area. To translate our Population, count in 30 m grids from each homestead forest polygon; we first intersected the 30 m grids with homestead forest polygons that contain the Population assigned to each polygon. This process brought the 30 m grid layer that contains homestead forest polygons inside each of those 30 m grids. To transfer the population in those homestead forest polygons that fell inside each of the 30 m grids, we divided the areas of those homestead forests by the areas of those 30 m grids. Then, we multiplied the ratio by the population in those homestead forest polygons within 30 m grids.

The cartographic representation of vector-based dasymetric maps was unsuitable to represent (Figure 2.6) the Population in each 30 m grid. To achieve a better cartographical representation of 30 m vector-based dasymetric maps, we converted the 30 m grid cells to 30 m raster.

2.2.2.2 *Raster-Based Dasymetric Population Mapping*

The concept of dasymetric mapping in raster-based dasymetric mapping is almost identical to the vector-based method (See Figure 2.3). The only exception is that the vector-based population data is converted into raster, and the ancillary landcover data is preserved in raster format. The process followed the equation below that was transformed by Holloway et al. (1996) to calculate per pixel population. Likewise, in the vector-based approach, we also calculated the areal weight and multiplied the areal weight of the homestead forest by the population of each

union. We assumed each homestead forest pixel has a relative density of 100 and multiplies that with the product of population and areal weight. Finally, we divided it with the total number of homestead forest cells to obtain per-pixel populations. The tools and techniques we used to implement this method have been described in Appendix D.

Population Per Pixel = $((R_H) * N*E)/(A_H/30*30)$ …….(Eq. 2.4)

Or, = $(R_H) * N*30*30*E/A_H$…… (Eq. 2.5)

Where,

R_H = the relative density of each homestead forest cell

$A_H/30*30$ = Total number of homestead forest cells in each union

N = Population in a Union

E = Areal weight of land use classes in each union

$E = P_H*100$……(Eq. 2.6)

P_H is the area proportion of homestead forest in a union calculated using the following equation,

$P_H = A_H/ A_H+A_F+A_O+A_W$ …….(Eq. 2.7)

A_H = Area of homestead forest, A_F= Area of Fallow land, A_O = Area of Water body A_W = Area of Waterbody

We divided the total homestead forest area by 30x30 to derive per pixel population to obtain the number of homestead forest cells.

We converted our vector-based union level population data to 30 m raster since we assigned Population in each 30 m pixels of homestead forests. As we intended to assign all our population to the class homestead forest, we reclassified the landcover data to provide homestead

forest class 100 percent relative density (See Appendix D for detailed processes). Next, the Area of each landcover class in each union and the proportion of homestead forest in each union was calculated using equations 2.7 and 2.6. Multiplying the areal proportion of homestead forests with 100, we obtained the area weight of homestead forests in each union (See Equation 2.5). We finally used equation 2.4 to calculate the population in each 30 m pixel.

2.2.3 *Volume Preservation and Visual Inspection of Dasymetric Population Maps*

We tested the accuracy of the dasymetric maps by observing the volume preservation. As described by Waldo Tobler (1979), volume preservation is preserving the total number of populations when it is transferred into new target zones. When populations are transferred from higher spatial units to lower spatial units, there are chances that populations might increase or decrease compared to their source-level population due to the calculation process (Tobler, 1979). To assess the volume preservation, we calculated the total sum of the union-level and dasymetric-level populations.

Dasymetric mapping techniques can distribute population density heterogeneously, with uneven population distribution across its parent source; however, the scope can be limited depending on ancillary data limitations. Wood (2006) stressed that dasymetric maps might represent homogeneous population density, even population distribution, within parent zones due to the lack of variation in land cover data. For instance, in urban areas, disaggregating census tract level populations into finer resolutions using ancillary land-use maps categorized as high, moderate, and low-density classes of residential land uses will provide a more realistic and heterogeneous distribution than using one broad class of residential land-use. Dasymetric mapping provides a more realistic and heterogeneous representation of population distribution

than a census unit-level choropleth map (Mennis, 2003). To understand how well our dasymetric maps represent the heterogeneity of population distribution compared to the choropleth map, we visually inspected the difference between the choropleth map and dasymetric maps.

2.2.4 *Calculation of Land Loss and Population Displacement due to Riverbank Erosion*

To estimate the amount of land loss and population displacement, we overlaid our dasymetric level population data with the shorelines of each study year. We used shoreline extracted by Crawford et al. (2021) in our study. The authors achieved an overall 25-30 m accuracy for the shorelines using Landsat Imagery. We took the shorelines and superimposed them on landcover data to auto-digitize shorelines on the landcover data. The purpose of doing this is to get accurate shorelines. Next, we extracted eroded areas every five years. Overlaying the areas with disaggregated population maps, we calculated the population displacement in each five years from 2001 to 2021.

2.3 References

BBS (2001). Statistical yearbook of Bangladesh, Ministry of Planning, Dhaka, Bangladesh.

BBS (2011). Statistical yearbook of Bangladesh, Ministry of Planning, Dhaka, Bangladesh.

Goodchild, M. F. (1992). Geographical data modeling. *Computers & Geosciences*, *18*(4), 401-408.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment.

Goodchild, M. F., & Lam, N. S. N. (1980). Areal interpolation: A variant of the traditional spatial problem. *Geo-processing*, *1*(3), 297-312.

Jitt-Aer, K., Wall, G., Jones, D., & Teeuw, R. (2022). Use of GIS and dasymetric mapping for estimating tsunami-affected population to facilitate humanitarian relief logistics: a case study from Phuket, Thailand. *Natural Hazards*, *113*(1), 185-211.

Mutanga, Onisimo, and Lalit Kumar. "Google earth engine applications." *Remote sensing* 11, no. 5(2019): 591.

Holloway, S. R., Schumacher, J., & Redmond, R. L. (1999). People and place: Dasymetric mapping using Arc/Info. *GIS solutions in natural resource management*, 283-291.

Zhao, Qiang, Le Yu, Xuecao Li, Dailiang Peng, Yongguang Zhang, and Peng Gong. "Progress and trends in the application of Google Earth and Google Earth Engine." *Remote Sensing* 13, no. 18 (2021): 3778.

Breiman, Leo. "Random forests." *Machine learning* 45 (2001): 5-32.

Probst, P., Wright, M. N., & Boulesteix, A. L. (2019). Hyperparameters and tuning strategies for random forest. *Wiley Interdisciplinary Reviews: data mining and knowledge discovery*, *9*(3), e1301.

The Humanitarian Data Exchange. (n.d.). Retrieved August, 2022, from https://data.humdata.org/

Mennis, Jeremy. "Dasymetric mapping for estimating population in small areas." *Geography Compass* 3, no. 2 (2009): 727-745.

Eicher, Cory L., and Cynthia A. Brewer. "Dasymetric mapping and areal interpolation: Implementation and evaluation." *Cartography and Geographic Information Science* 28, no. 2 (2001): 125-138.

Immitzer, M., Vuolo, F., & Atzberger, C. (2016). First experience with Sentinel-2 data for crop and tree species classifications in central Europe. *Remote sensing*, *8*(3), 166.

Tobler, Waldo R. "Smooth pycnophylactic interpolation for geographical regions." *Journal of theAmerican Statistical Association* 74, no. 367 (1979): 519-530.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*.

Liu, X. H., Phaedon C. Kyriakidis, and Michael F. Goodchild. "Population-density estimation using regression and area-to-point residual kriging." *International Journal of geographical information science* 22, no. 4 (2008): 431-447.

Wu, Changshan, and Alan T. Murray. "A cokriging method for estimating population density in urban areas." *Computers, Environment and Urban Systems* 29, no. 5 (2005): 558-579.

Zandbergen, P. A., & Ignizio, D. A. (2010). Comparison of dasymetric mapping techniques for small-area population estimates. *Cartography and Geographic Information Science*, *37*(3), 199-214.

Sleeter, R. (2004, November). Dasymetric mapping techniques for the San Francisco Bay region California. In *Urban and regional information systems association annual conference proceedings* (pp. 7-10).

Sumathi, B. (2020). Grid search tuning of hyperparameters in random forest classifier for customer feedback sentiment prediction. International Journal of Advanced Computer Science and Applications, 11(9).

Baynes, J., Neale, A., & Hultgren, T. (2022). Improving intelligent dasymetric mapping population density estimates at 30 m resolution for the conterminous United States by excluding uninhabited areas. *Earth system science data*, *14*(6), 2833-2849.

Kilany, M., Zhang, C., & Li, W. (2021). Optimization of urban land cover classification using an improved Elephant Herding Optimization algorithm and random forest classifier. *International Journal of Remote Sensing*, *42*(15), 5741-5763.

Sleeter, R., & Wood, N. (2006). Estimating daytime and nighttime population density for coastal communities in Oregon. In *44th Urban and Regional Information Systems Association Annual Conference, British Columbia* (pp. 1-15).

Lubke, M., Rengarajan, R., & Choate, M. (2021, August). Preliminary assessment of the geometric improvements to the Landsat Collection-2 archive. In *Earth Observing Systems XXVI* (Vol. 11829, pp. 125-137).

2.4 List Tables and Figures

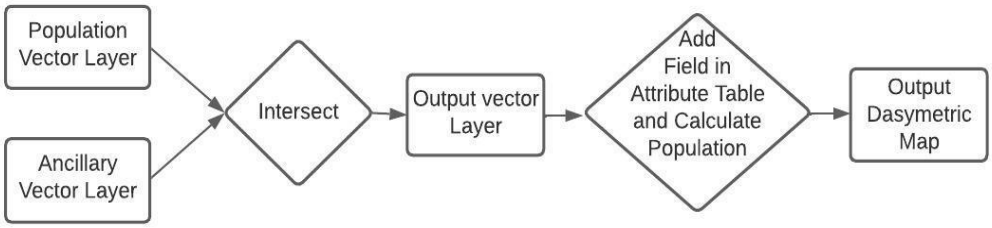


Figure 2.1 Conceptual diagram of vector-based dasymetric mapping

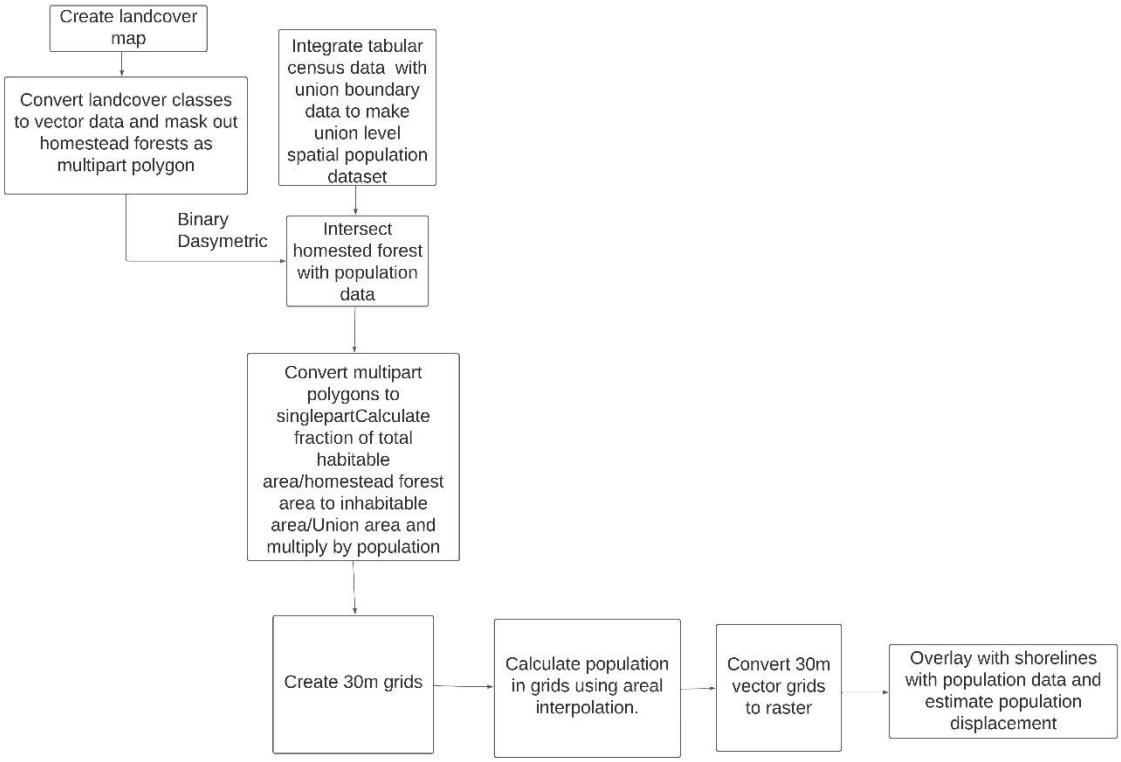


Figure 2.2 Workflow vector-based dasymetric mapping and population displacement.

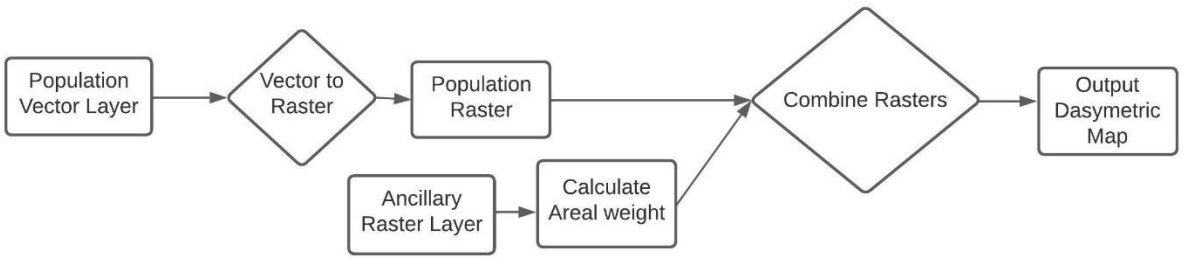


Figure 2.3 Conceptual diagram of raster-based dasymetric mapping.

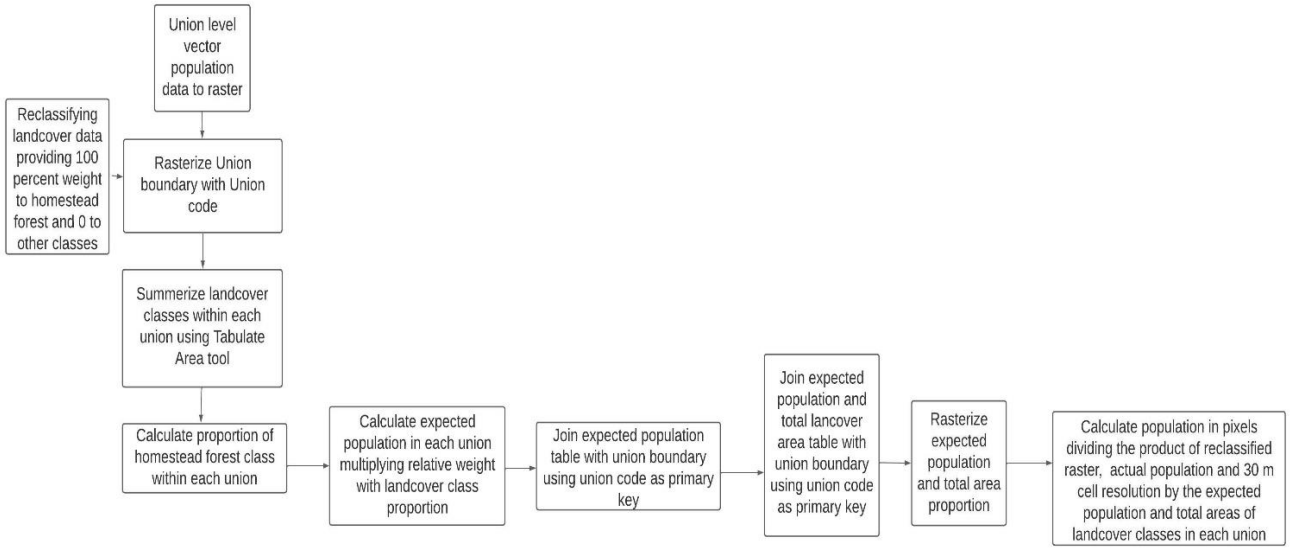


Figure 2.4 Workflow raster-based dasymetric mapping and population displacement.

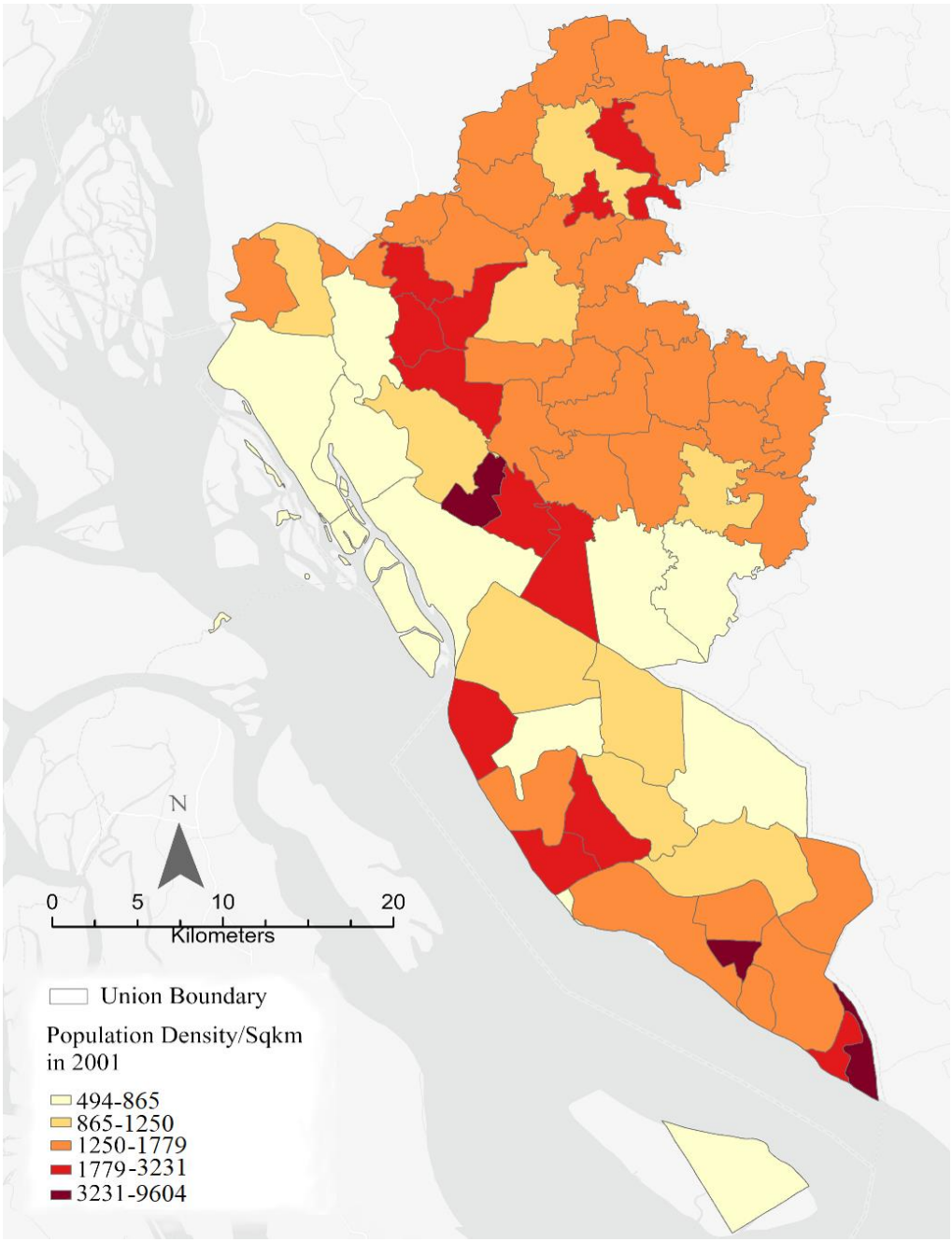


Figure 2.5 Choropleth map representing union-level population distribution in 2001. The lowest density value was 494 persons/km2 and the highest was 9604 persons/km2.

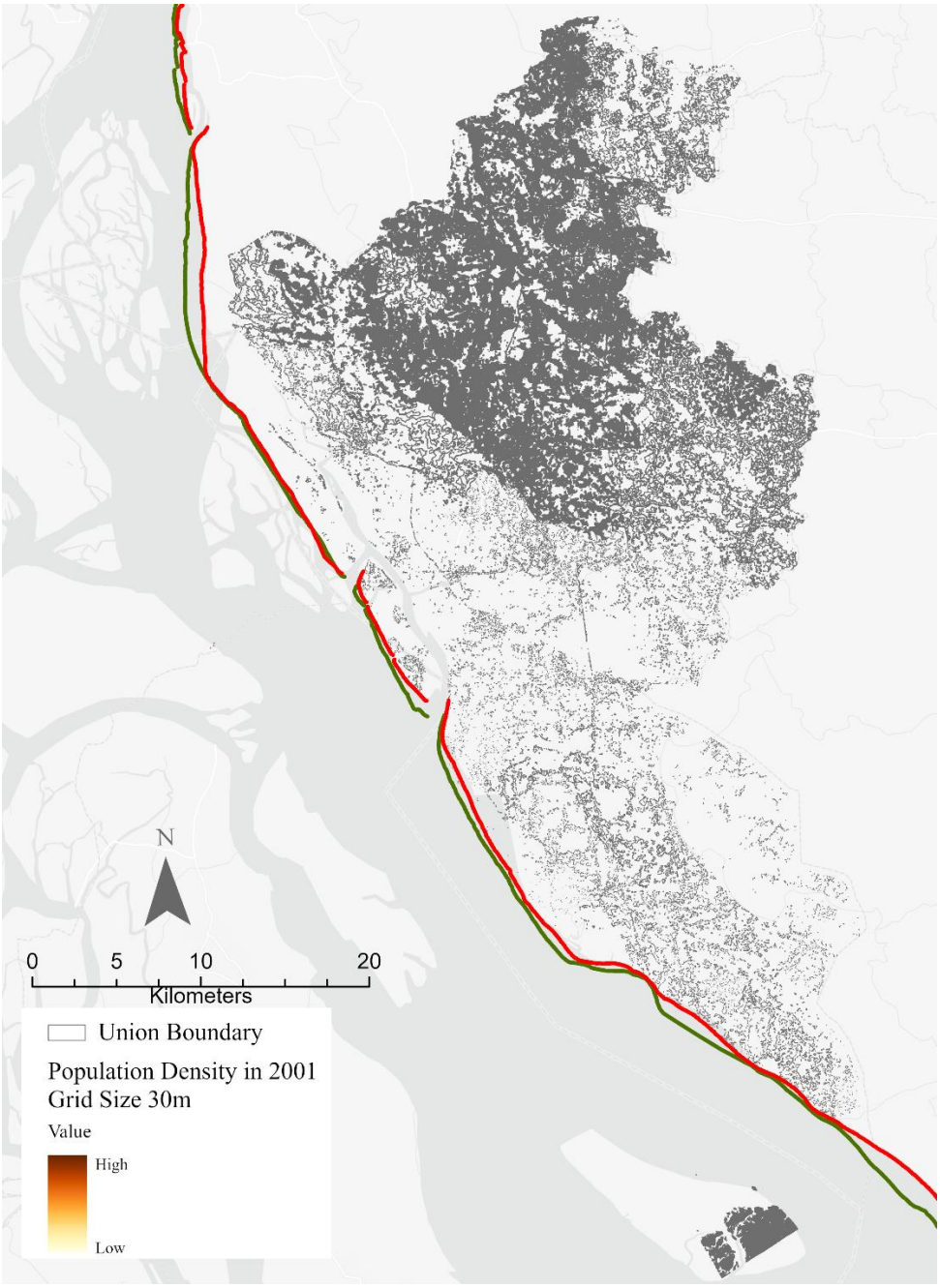


Figure 2.6 Representation of vector-based dasymetric map with lowest 30 m mapping unit. The vector representation of 30 m grids could not produce a cartographically sound map.

Table 2.1 Census year and mid-census year population, Lakshmipur District, Bangladesh

Year	Union Level Total Population
2001	1,706,858
2006	1,643,838
2011	1,672,552
2016	2,115,626
2021	2,239,718

CHAPTER 3

RESULTS

3.1 Image Classification

3.1.1 *Random Forest Performance*

As stated earlier in Chapter 3, the smile random forest classifier in the Google Earth Engine allows manual tuning of the data split, number of trees, variables per split, bag fraction, and maximum nodes. This section describes the combinations of training and testing data, number of trees, bag fraction, and variable per split used for different years to achieve optimum results from a random forest classifier. Classification parameters were tested on the 2001 Landsat TM 5 image, image acquisition date 01/21/2001. We tested several proportions of training and testing samples, adjusted the number of variables per split, number of trees, and bag fraction, and observed the model performance. We found a 70/30 split of training and testing data with 200 numbers of trees, two variables per split, and 0.4 bag fraction as the best combination that provided the lowest out-of-bag error of 0.27. The most essential variable in this classification was Band 5 Near-Infrared (1.55 - 1.75 µm), and the least essential band was B2(Green 0.52 - 0.60 µm).

Utilizing a 70/30 split of training and testing data with the default of 200 trees and 0.5 bag fraction for the 2006 Landsat TM 5 image (image acquisition date 02/04/2006), we obtained the lowest out-of-bag error of 0.06. The most crucial feature in the classification was Band 4 Near-Infrared (0.76 - 0.90 µm), and the least important band was B2(Green 0.52 - 0.60 µm).

For the 2011 Landsat TM 5 image (acquisition date 01/02/2011), 75/,25 splits of training and testing data with the default of 200 trees and a bag fraction of 0.6 resulted in the lowest out-

of-bag error of 0.01. Band 7 Mid-Infrared (2.08 - 2.35 µm) was the classification's most significant variable, and the least important band was B2(Green 0.52 - 0.60 µm).

For the 2016 Landsat 8 OLI image (acquisition date: 01/15/2016), a 75/25 split of training and testing data produced the optimum result. Changing hyperparameters of the random forest classifier did not have any effect on out-of-bag error. With default settings, we obtained a 0.021 out-of-bag error. Band 5 (Near Infrared 0.85-0.88 µm) was the most important band and the least important band B1 (Coastal aerosol 0.43-0.45 µm).

When, in some cases, tuning hyperparameters improved the results of our previous classifications, for 2021's Landsat 8 OLI image (Acquisition date 17/02/2021), default hyperparameters with a 70/30 split came out as the best combination. The default setting yielded a 0.14 out-of-bag error. The most significant band was band 5 (Near Infrared 0.85-0.88 µm). The least important band was B1 and B2(Coastal aerosol 0.43-0.45 µm, blue 0.45-0.51µm).

3.1.2 *Accuracy Assessment*

We produced our confusion matrices using ground reference data (see Chapter 3). The classification of the 2001 image, based upon the Landsat TM 5 image (image acquisition date 01/21/2021), provides an overall accuracy of 79.04 % (Table 3.1). The landcover map (Figure 3.1) shows four landcover classes: homestead forest, fallow land, other vegetation, and water body, with individual class accuracy ranging from 93.33 percent for homestead forest to 57.14 percent for the waterbody (Table 3.3). Other vegetation was the most overclassified class (commission error 44.44 percent), mostly confused with fallow land and homestead forest. The classification worked well to classify our most important class, homestead forest, with an accuracy of 93.33 percent, 6.66 percent omission error, and 20 percent commission error while

contributing pixels to other vegetation type classes and received from fallow land and water bodies. It also has the highest producer, 93.3 percent, and user accuracy of 80 percent compared to other classes (Table 3.3).

Based on the Landsat 5TM image, acquisition date 02/04/2006, the classification of the 2006 image yields an overall accuracy of 81 percent (Table 3.4). Four landcover classes—homestead forest, fallow land, other vegetation, and water body—are represented on the landcover map (Figure 3.2), with the accuracy of each class varying from 92.68 percent for the homestead forest to 44.44 percent for the other vegetation (Table 3.6). Likewise, the previous classification of other vegetation types was the most frequently misclassified (commission error: 54.54%) class, confused with homestead forests and water bodies. The classification provided satisfactory results, classifying our most important class, homestead forests, with the highest producer and user accuracy of 92.68 percent, 82.60 percent, 7% omission error, and 17.39 % commission error while supplying pixels to the class other vegetation and receiving them from them fallow land, other vegetation, and water body.

The overall accuracy for the 2011 image classification, using Landsat 5 TM image (image acquisition date 01/02/2011), is 87.87 % (Table 3.7). The landcover map depicts four landcover classes (Figure 3.3): homestead forest, fallow land, other vegetation, and water body. The accuracy of each class ranges from 91.66 percent for the homestead forest to 66.66 percent for the water body (Table 3.9). The class of other vegetation types is the most overclassified commission error of 50% (Table 3.8), and it is frequently mistaken for fallow land and homestead forest. While providing pixels to other vegetation type classes and receiving them from fallow land and water bodies, the classifier successfully classified our most significant

class, homestead forest, with an accuracy of 91.66 percent, 8.33 percent omission error, and 6.33 percent commission error.

We have achieved 82% overall accuracy (Table 3.10) for 2016 using Landsat 8 OLI image (acquisition date: 01/15/2016). The landcover map shows four types of landcover (Figure 3.4): water bodies, fallow land, homestead forests, and other vegetation. The accuracy of each class varies, as shown in (Table 3.12), from 87% for the homestead forest to 77% percent for the water body. The most overclassified class is the other vegetation, with a commission error of 33% (Table 3.11), which has been mistaken for homestead forests. Random forest classifiers classified our most important class, homestead forest, with an accuracy of 87%, an omission error of 12%, and a commission error of 10 % while providing pixels to other vegetation types and receiving them from fallow land, other vegetation, and water bodies as well.

Landcover classification for the year 2021 using Landsat 8 OLI (Image acquisition date 17/03/2021) the overall accuracy is 83% (Table 3.13). The landcover map (Figure 3.5) shows four landcover classes: homestead forest, fallow land, other vegetation, and water body, with individual class accuracy ranging from 91.83 percent for homestead forest to 56.25 percent for the other vegetation (Table 3.15). Other vegetation was the most overclassified class (commission error 43.75 percent), mostly confused with fallow land and homestead forest. The classification worked well to classify our most important class, homestead forest, with a user accuracy of 93.75 percent, 8.1 percent omission error, and 6.25 percent commission error while contributing pixels to waterbody and fallow land classes and received from fallow land and other vegetation.

3.2 Change detection

We observed 21 years of change from 2001-2021. The overall accuracy of the 2001 classification (Landsat 5 TM, image acquisition date 01/21/2001) was 79.04%, and the 2021(Landsat 8 OLI, image acquisition date 17/03/2021) classification was 83.0 %. Fallow land experienced a substantial net loss of 341.81 sq. km, mainly due to its conversion into homestead forests and water bodies. Specifically, 276.65 sq. km of fallow land became homestead forests, and 132 sq. km were transformed into water bodies. This class of land lost 420 sq. km and gained only 78 sq. km, resulting in a net loss of 341 sq. km. Homestead Forest had the highest net gain of 220 sq. km, most of which was received from fallow land. The class lost 76.91 sq. km to the water body. The second highest net gain happened to waterbody 162.93 sq. km, with the highest gain from fallow land 132.69 sq. km and homestead forest 76 sq. km.

3.3 Population Mapping

3.3.1 *Dasymetric Mapping Vector and Raster-Based Approach*

Table 3.17 exhibits union level and dasymetric population count using a vector-based approach. The count was compared to understand if any over or underestimation happened during the disaggregation process. It should be noted that the results have been derived from a vector-based approach; however, the dasymetric maps (Figure 3.6- 3.10) have been presented in a raster version for better cartographic representation. Our result shows (Table 3.17) that volume has not been preserved in the year 2001. The total population was 1,706 858 in 2001, and the estimated population was 1,666,537. We lost 40,321 people in disaggregating the population from the union level to the 30 m dasymetric level population.

Table 3.18 represents the union level and dasymetric population count using a raster-based approach. The total population and estimated population counts are not equal in 2001 and 2016, indicating the violation of the volume preservation assumption. In 2001 and 2016, the total union-level population was 1706858 and 2,115,626, whereas the estimated dasymetric level population was 1,728,554 and 2,119,186. In the process of disaggregation, we overestimated 21696 and 3560 people.

3.3.2 *Dasymetric Mapping Visual Inspection*

Our visual inspection demonstrates that the variation captured by the dasymetric mapping is slightly improved over a choropleth population density map. Figure 3.6 -3.10 and Figure 3.12-3.16, show that dasymetric mapping population density only represents a few variations in most unions compared to the choropleth map (Fig 3.11). However, unions near the riverbank show some variation in population distribution within union boundaries.

3.4 Amount of Land Loss and Population Displacement Due to Riverbank Erosion

Table 3.19 represents the amount of land loss to river erosion every five years from 2001-2021. Our results indicate that land loss was highest between 2011-2016, 30.3 square kilometers, and lowest between 2001-2006, 4.93 square kilometers. Table 3.20 shows the estimated population displacement due to river erosion using vector-based dasymetric population maps. The highest number of people, 86107, were relocated between 2016-2021, and the lowest number of people, 1844, were relocated between 2001-2006. On average, 9582 people were displaced in the past 21 years due to river erosion. The years of the highest and lowest percentage of population displacement were consistent with the years of counted highest and lowest number of population displacement in the vector-based method. The lowest percentage of

population displacement happened from 2001-2006, which was 0.1 percent of the total population in 2001. The highest percentage of population displacement was from 2016-2021, which was 4.07 percent of the total population in 2016. Table 3.21 shows the population displacement calculation due to river erosion using a raster-based approach. The maximum number of people 63453 was displaced between 2016-2021, whereas the minimum number of people displaced 5241, was between 2001-2006. The highest percentage of displacements, 3.7 percent of the total population, happened between 2011-2016, although the highest count shows that the maximum number of displacements occurred between 2016-2021. This is because the percentage is a relative measure and relative to the total population; the highest number of displacements happened from 2011-2016, in raster-based calculation.

3.5 References

BBS (2001). Statistical yearbook of Bangladesh, Ministry of Planning, Dhaka, Bangladesh.

BBS (2011). Statistical yearbook of Bangladesh, Ministry of Planning, Dhaka, Bangladesh.

Crawford, T. W., Rahman, M. K., Miah, M. G., Islam, M. R., Paul, B. K., Curtis, S., & Islam, M. S. (2021). Coupled adaptive cycles of shoreline change and households in deltaic Bangladesh: analysis of a 30-year shoreline change record and recent population impacts. *Annals of the American Association of Geographers*, *111*(4), 1002-1024.

Fisher, P. F., & Langford, M. (1996). Modeling sensitivity to accuracy in classified imagery: A study of areal interpolation by dasymetric mapping. *The Professional Geographer*, *48*(3), 299-309.

Holt, J. B., Lo, C. P., & Hodler, T. W. (2004). Dasymetric estimation of population density and areal interpolation of census data. *Cartography and Geographic Information Science*, *31*(2), 103-121.

Hoque, M. Z., Cui, S., Islam, I., Xu, L., & Tang, J. (2020). Future impact of land use/land cover changes on ecosystem services in the lower meghna river estuary, Bangladesh. *Sustainability*, *12*(5), 2112.

Hasnat, M. A., Hossain, N., Muhibbullah, M., Sarwar, M. D., & Shormin, T. (2016). Impacts of climate change on agriculture and changing adaptive strategies in the coastal area of Lakshmipur district, Bangladesh. *Current World Environment*, (3), 700.

Islam, M. S., Crawford, T. W., & Shao, Y. (2023). Evaluation of predicted loss of different land use and land cover (LULC) due to coastal erosion in Bangladesh. *Frontiers in Environmental Science*, *11*, 479.

Jin, Y., Liu, X., Chen, Y., & Liang, X. (2018). Land-cover mapping using Random Forest classification and incorporating NDVI time-series and texture: A case study of central Shandong. *International journal of remote sensing*, *39*(23), 8703-8723.

Langford, M. (2006). Obtaining population estimates in non-census reporting zones: An evaluation of the 3-class dasymetric method. *Computers, environment and urban systems*, *30*(2), 161-180.

Langford, M. (2007). Rapid facilitation of dasymetric-based population interpolation by means of raster pixel maps. *Computers, Environment and Urban Systems*, *31*(1), 19-32.

Maantay, J. A., Maroko, A. R., & Herrmann, C. (2007). Mapping population distribution in the urban environment: The cadastral-based expert dasymetric system (CEDS). *Cartography and Geographic Information Science*, *34*(2), 77-102.

Mennis, Jeremy. "Dasymetric mapping for estimating population in small areas." *Geography Compass* 3, no. 2 (2009): 727-745.

Magidi, J., Nhamo, L., Mpandeli, S., & Mabhaudhi, T. (2021). Application of the random forest classifier to map irrigated areas using google earth engine. *Remote Sensing*, *13*(5), 876.

Stahler, G. J., Mazzella, S., Mennis, J., Chakravorty, S., Rengert, G., & Spiga, R. (2007). The effect of individual, program, and neighborhood variables on continuity of treatment among dually diagnosed individuals. *Drug and alcohol dependence*, *87*(1), 54-62.

Tobler, W. R. (1979). Smooth pycnophylactic interpolation for geographical regions. *Journal of the American Statistical Association*, *74*(367), 519-530.

Wang, D., Wan, B., Qiu, P., Su, Y., Guo, Q., Wang, R., ... & Wu, X. (2018). Evaluating theperformance of Sentinel-2, Landsat 8 and Pléiades-1 in mapping mangrove extent and species. *Remote Sensing*, *10*(9), 1468.

Wu, C., & Murray, A. T. (2005). A cokriging method for estimating population density in urban areas. *Computers, Environment and Urban Systems*, *29*(5), 558-579.

3.6 Tables and Figures

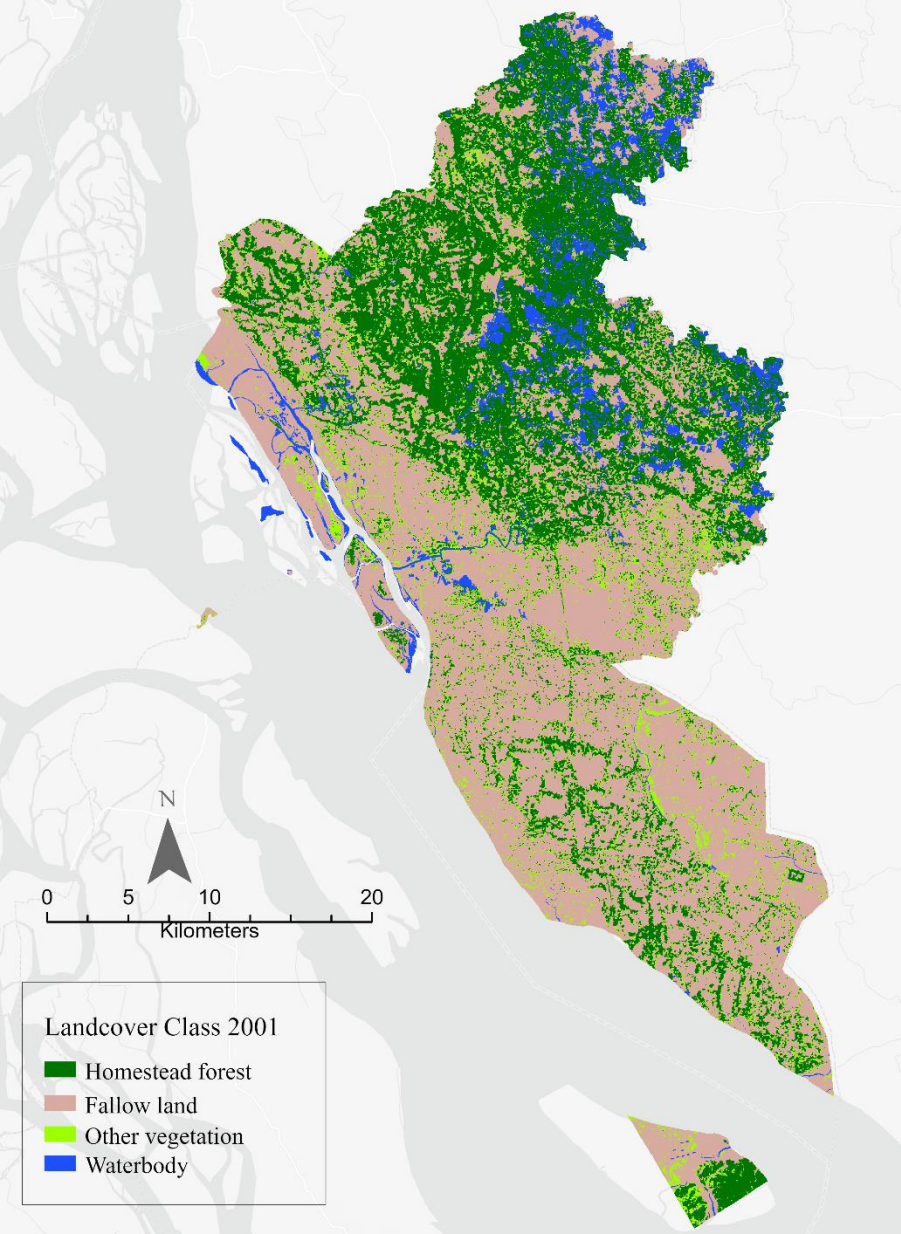


Figure 3.1Classified image for year 2001.Overall accuracy 79.04 %

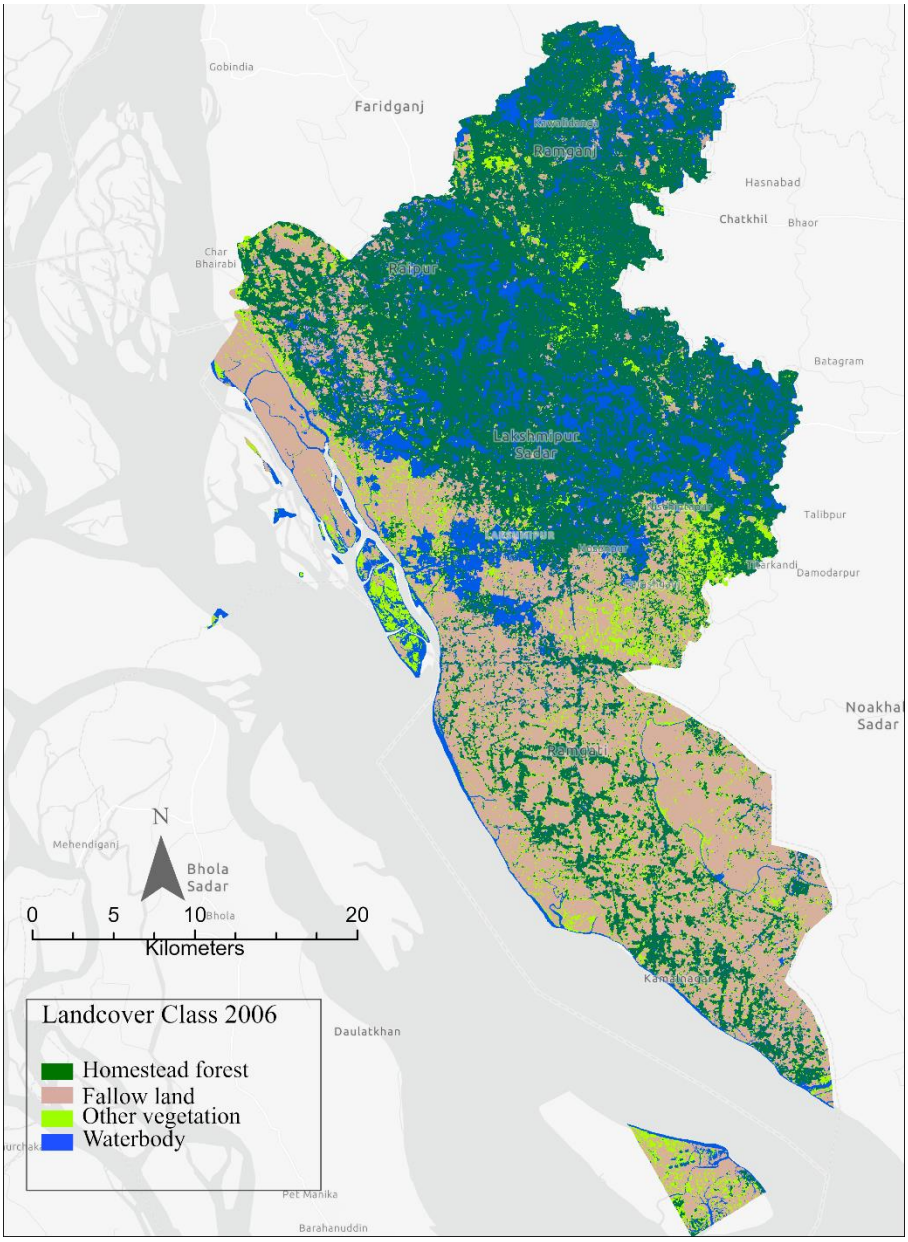


Figure 3.2 Classified image for the year 2006. Overall accuracy 81 %

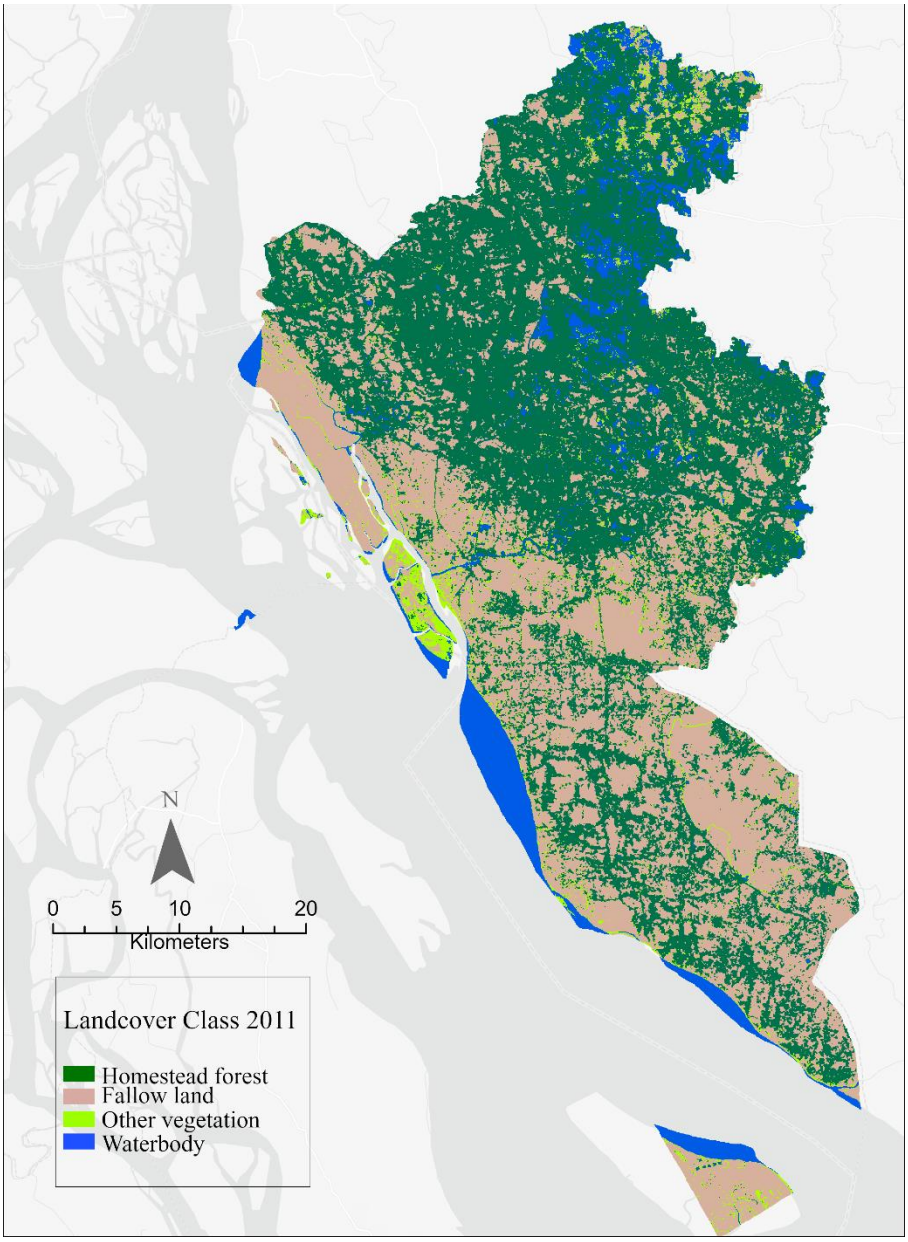


Figure 3.3 Classified image for the year 2011. Overall accuracy 87.87 %

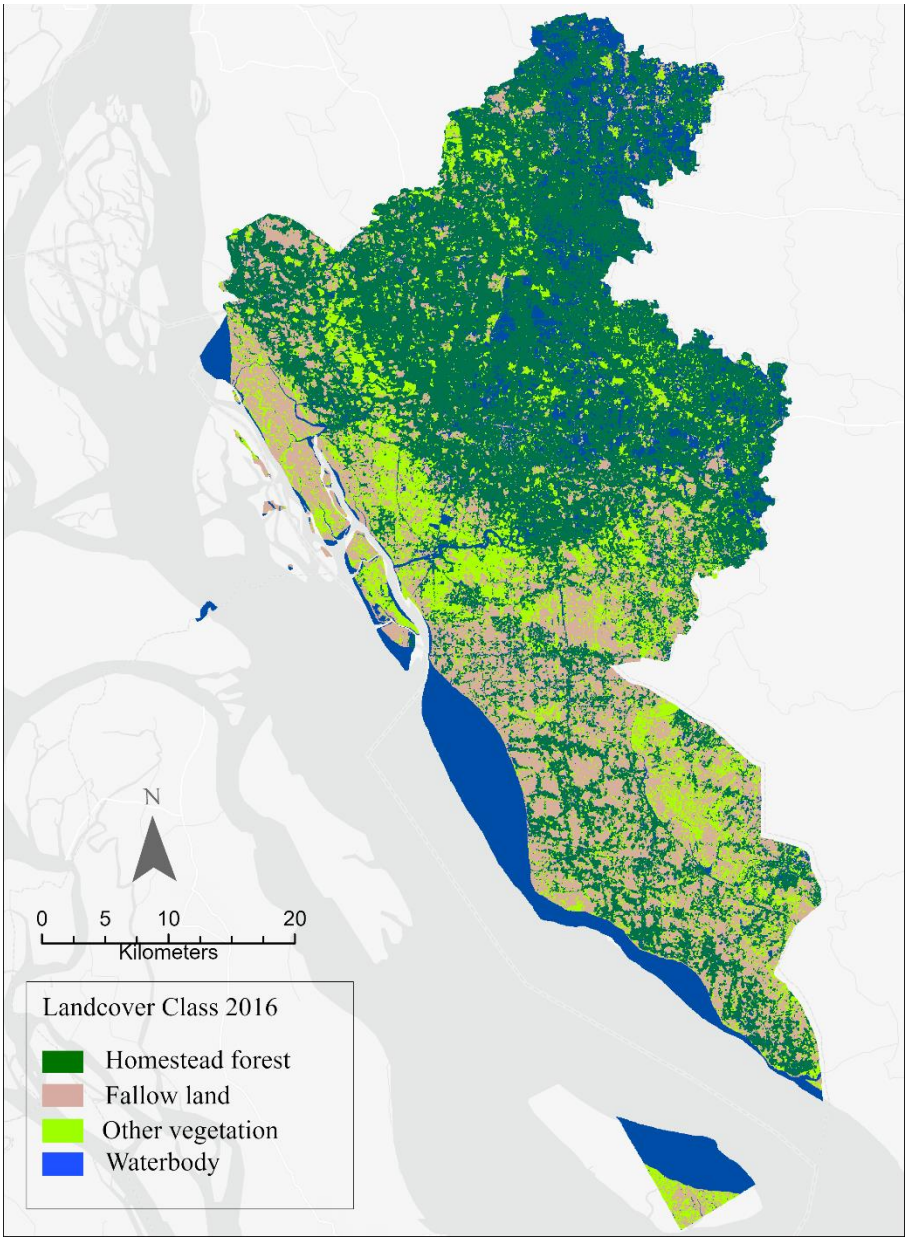


Figure 3.4 Classified image for year 2016. Overall accuracy 88 %

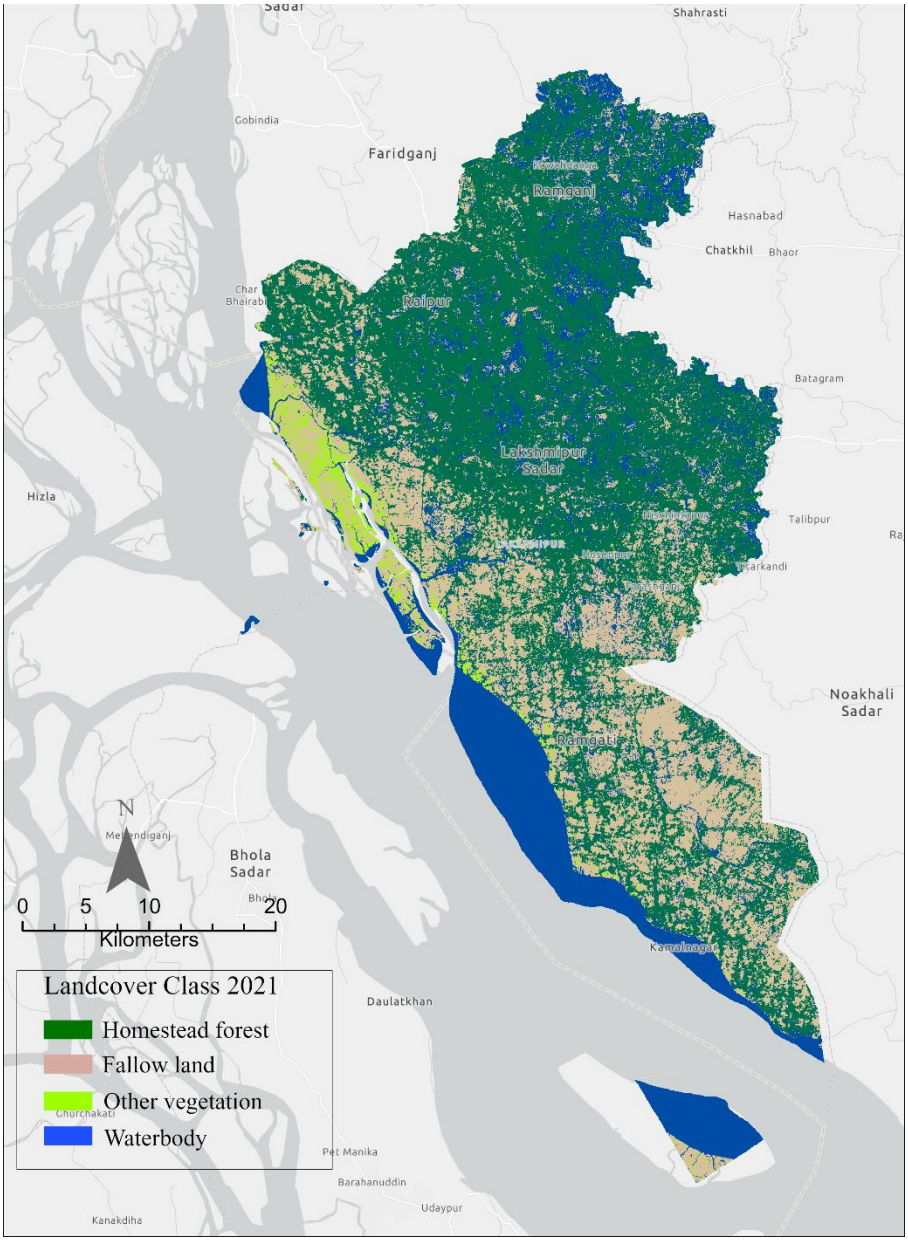


Figure 3.5 Classified image for year 2021. Overall accuracy 83 %

Table 3.1 Accuracy Assessment of 2001 landcover Classification (Image acquisition date 01/21/2001). The columns represent ground truth points (GTP) reference, rows represent classification from the satellite image. Overall accuracy 79.04%

	GTP			
Classified	Homestead forest	Fallow land	Vegetation	Waterbody
Homestead forest	28	5	0	2
Fallow land	0	41	0	0
Vegetation	2	5	10	1
Waterbody	0	1	0	4

Table 3.2 Errors of commission, errors of omission for each landcover class of 2001 landcover Classification (Image acquisition date 01/21/2001)

	Error of Omission(percent)	Error of Commission (percent)
Homestead forest	6.66	20
Fallow land	21.15	0
Vegetation	0	44.44
Waterbody	42.85	20

Table 3.3 Producer's and user's accuracies for each cover class of 2001 landcover Classification (Image acquisition date 01/21/2001).

	Producer's accuracy (Percent)	User's accuracy (Percent)
Homestead forest	93.33	80
Fallow land	78.84	100
Vegetation	100	55.55
Waterbody	57.14	80

Table 3.4 Accuracy Assessment of 2006 landcover Classification (Image acquisition date 02/04/2006). The columns represent ground truth points (GTP)/reference data (acquired using Google Earth's historical Imagery), and the row represents classification from satellite Imagery. Overall accuracy 81%

	GTP			
Classified	Homestead forest	Fallow land	Vegetation	Waterbody
Homestead forest	38	3	4	1
Fallow land	0	28	0	0
Vegetation	3	1	5	2
Waterbody	0	5	0	10

Table 3.5 Errors of commission, errors of omission for each landcover class of 2006 landcover Classification (Image acquisition date02/04/2006).

	Error of Omission(percent)	Error of Comission(percent)
Homestead forest	7.31	17.39
Fallow land	10.81	0
Vegetation	55.55	63.63
Waterbody	23.07	33.33

Table 3.6 Producer's and User's accuracy for each landcover class of 2006 landcover Classification (Image acquisition date02/04/2006).

	Producer's accuracy	User's accuracy
Homestead forest	92.68	82.60
Fallow land	89.18	100
Vegetation	44.44	36.36
Waterbody	76.92	66.66

Table 3.7 Accuracy Assessment of 2011 land cover Classification (Image acquisition date 01/02/2011). The columns represent ground truth points (GTP)/reference data (acquired using Google Earth's historical Imagery), row represents classification from satellite images. Overall accuracy 87.87%

	GTP			
Classified	Homestead forest	Fallow land	Vegetation	Waterbody
Homestead forest	44	0	3	0
Fallow land	0	34	0	1
Vegetation	3	4	7	0
Waterbody	1	0	0	2

Table 3.8 Errors of commission, errors of omission for each landcover class of 2011 landcover Classification (Image acquisition date01/02/2011)

	Error of Omission(percent)	Error of Commission (percent)
Homestead forest	8.33	6.38
Fallow land	10.52	2.85
Vegetation	30	50
Waterbody	33.33	33.33

Table 3.9 Producer's and User's accuracy for each landcover class of 2011 landcover Classification (Image acquisition date 01/02/2011)

	Producer's accuracy	User's accuracy
Homestead forest	91.66	93.61
Fallow land	89.47	97.14
Vegetation	70	50
Waterbody	66.66	66.66

Table 3.10 Accuracy Assessment of 2016 land cover Classification (Image acquisition date 01/15/2016). The columns represent ground truth points (GTP)/reference data (acquired using Google Earth's historical Imagery), and the rows represent classification from satellite images. Overall accuracy 82%

	GTP			
Classified	Homestead forest	Fallow land	Other vegetation	Waterbody
Homesteadforest	34	2	1	1
Fallow land	1	24	0	2
Other vegetation	3	1	10	1
Waterbody	1	1	1	14

Table 3.11 Errors of commission, errors of omission for each landcover class of 2016 landcover Classification (Image acquisition date 01/18/2016).

	Error of Omission	Error of Commission
Homestead forest	12.82	10.52
Fallow land	14.2	11.11
Other vegetation	16.66	33.33
Waterbody	22.22	17.64

Table 3.12 Producer's and User's accuracy for each landcover class of 2016 landcover Classification (Image acquisition date 01/18/2016).

	Producer Accuracy	User Accuracy
Homestead forest	87.17	89.47
Fallow land	85.71	88.88
Other vegetation	83.33	66.66
Waterbody	77.77	82.35

Table 3.13 Accuracy Assessment of 2021 landcover Classification using Landsat 8 OLI (Image acquisition date 17/03/2021). The columns represent ground truth points (GTP)/reference data (acquired using Google Earth's historical Imagery), and rows represent classification from satellite images. Overall accuracy 83.00%

	GTP				
Classified	Homesteadforest	Fallow land	Other vegetation	Waterbody	Overall accuracy
Homesteadforest	45	2	1	0	0.83
Fallow land	2	20	1	1	
Other vegetation	0	7	9	0	
Waterbody	2	1	0	9	

Table 3.14 Errors of commission, errors of omission for each landcover class of 2021landcover Classification using Landsat 8 OLI (Image acquisition date 17/03/2021).

	Error of Omission(percent)	Error of Commission(percent)
Homesteadforest	8.16	6.25
Fallow land	33.33	16.66
Other vegetation	18.18	43.75
Waterbody	10	25

Table 3.15 Producer's and User's accuracy for each landcover class of 2021 landcover Classification using Landsat 8 OLI (Image acquisition date 17/03/2021).

	Producer's Accuracy	User's Accuracy
Homesteadforest	91.83	93.75
Fallow land	66.66	83.33
Other vegetation	81.81	56.25
Waterbody	90	75

Table 3.16 Landcover change matrix from 2001-2021 using Landsat 5TM and Landsat 8 OLI Imagery (Image acquisition date 01/21/2021 and Acquisition date 17/03/2021). Each column represents the initial area of each class, while each row represents the final area of each class. Shaded cells indicate the areas that did not change. Positive signs indicate a net gain, whereas negative signs indicate a net loss.

	Homesteadforest	Fallow land	Other vegetation	Waterbody	Total Loss
Homesteadforest	270.59	50.18	0.36	76.91	127.46
Fallow land	276.65	291.75	11.45	132.69	420.80
Other vegetation	33.32	14.20	1.57	8.19	55.72
Waterbody	38.23	14.60	2.03	10.74	54.86
TotalGain	348.20	78.98	13.85	217.80	
NETLOSS/GAIN	220.74	-341.81	-41.86	162.93	

Figures Dasymetric Mapping

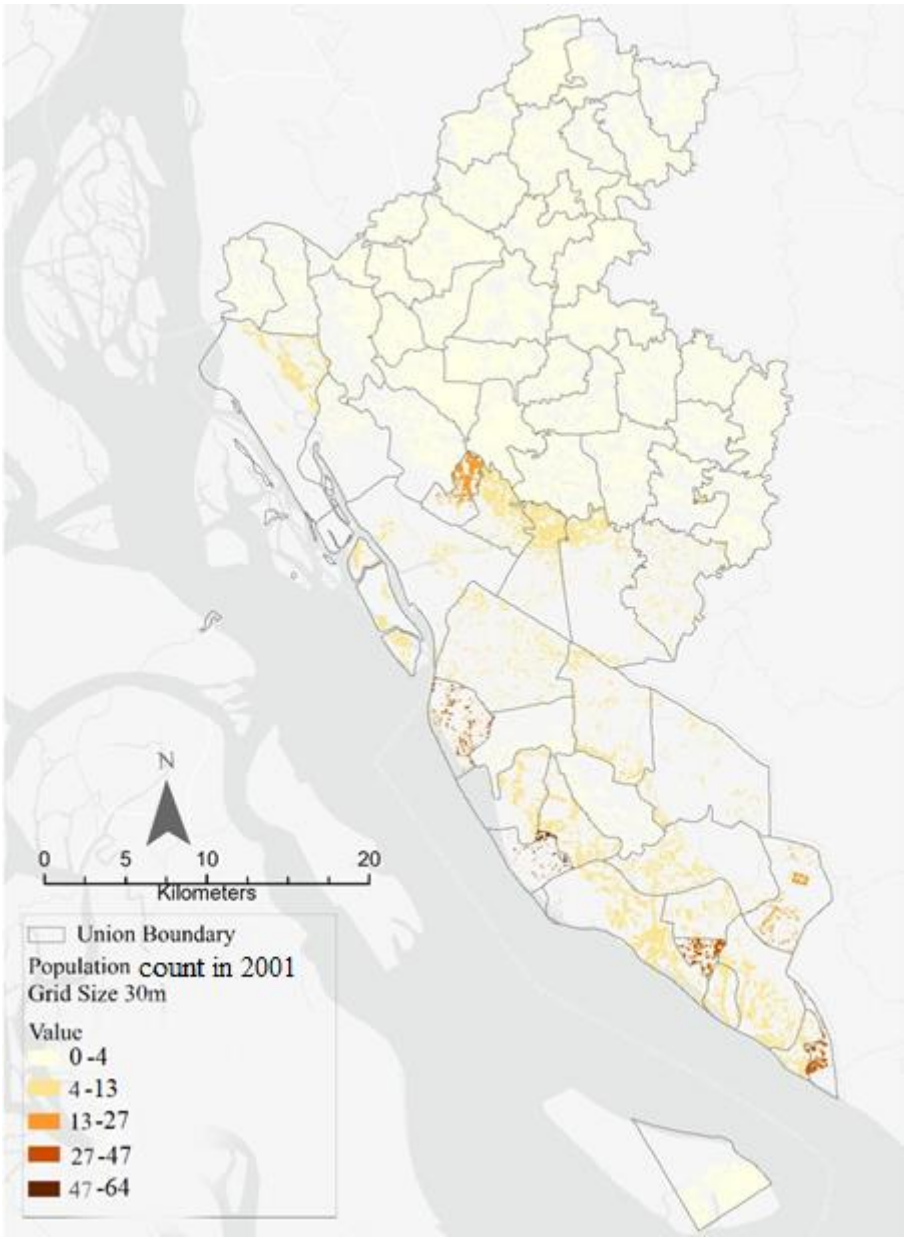


Figure 3.6 Dasymetric Population Distribution In 2001, 30 m vector grids were converted to 30 m raster grids. The 30 m raster surface has been symbolized using the Natural Jenks classification method. Original population densities have decimal places that have been rounded up in the legend to avoid representation of fractional population values. The raster surface represents population count (number of persons in each 30 m grid).

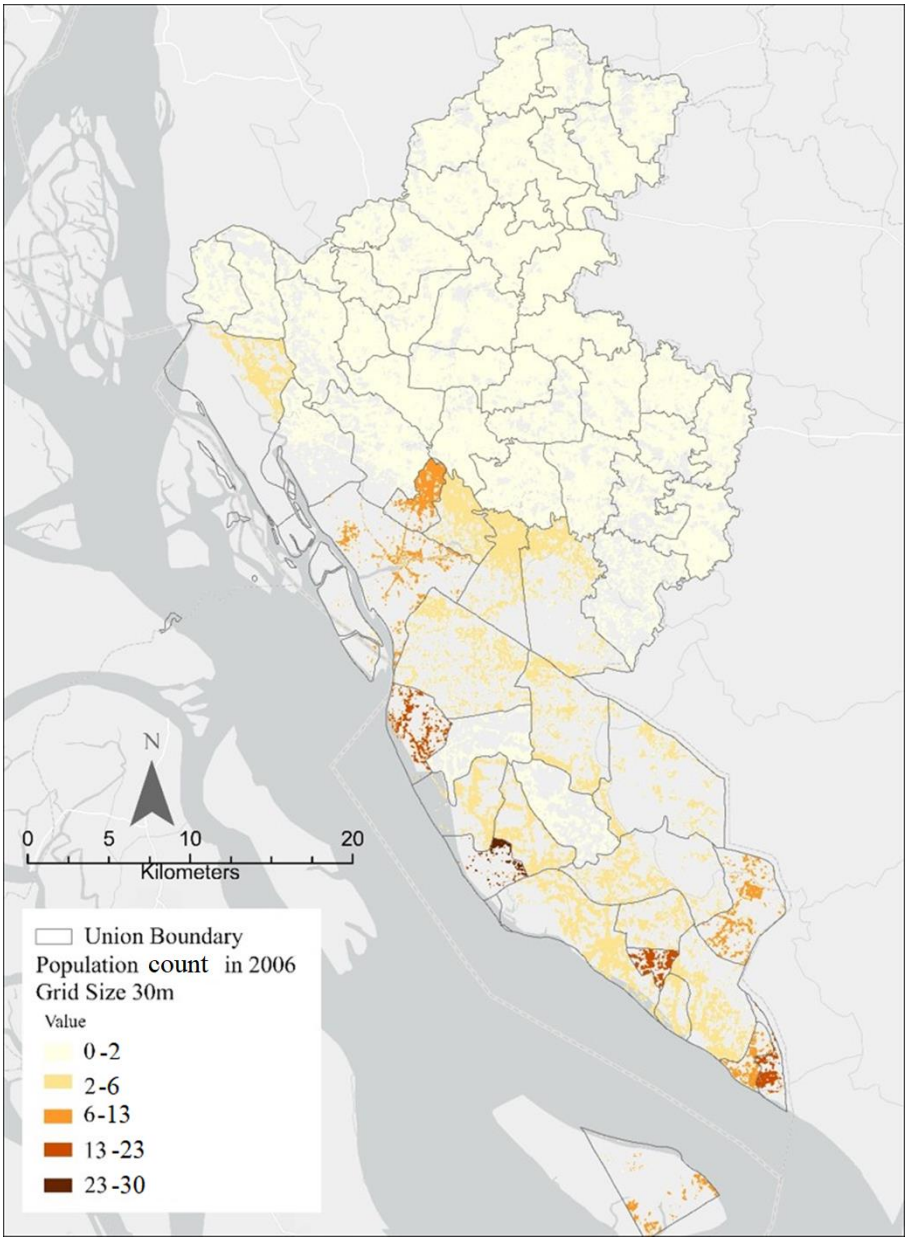


Figure 3.7 Dasymetric Population Distribution in 2006, 30 m vector grids were converted to 30 m raster grids. The 30 m raster surface has been symbolized using the Natural Jenks classification method. Original population densities have decimal places that have been rounded up in the legend to avoid representation of fractional population values. The raster surface population count (number of persons in each 30 m grid).

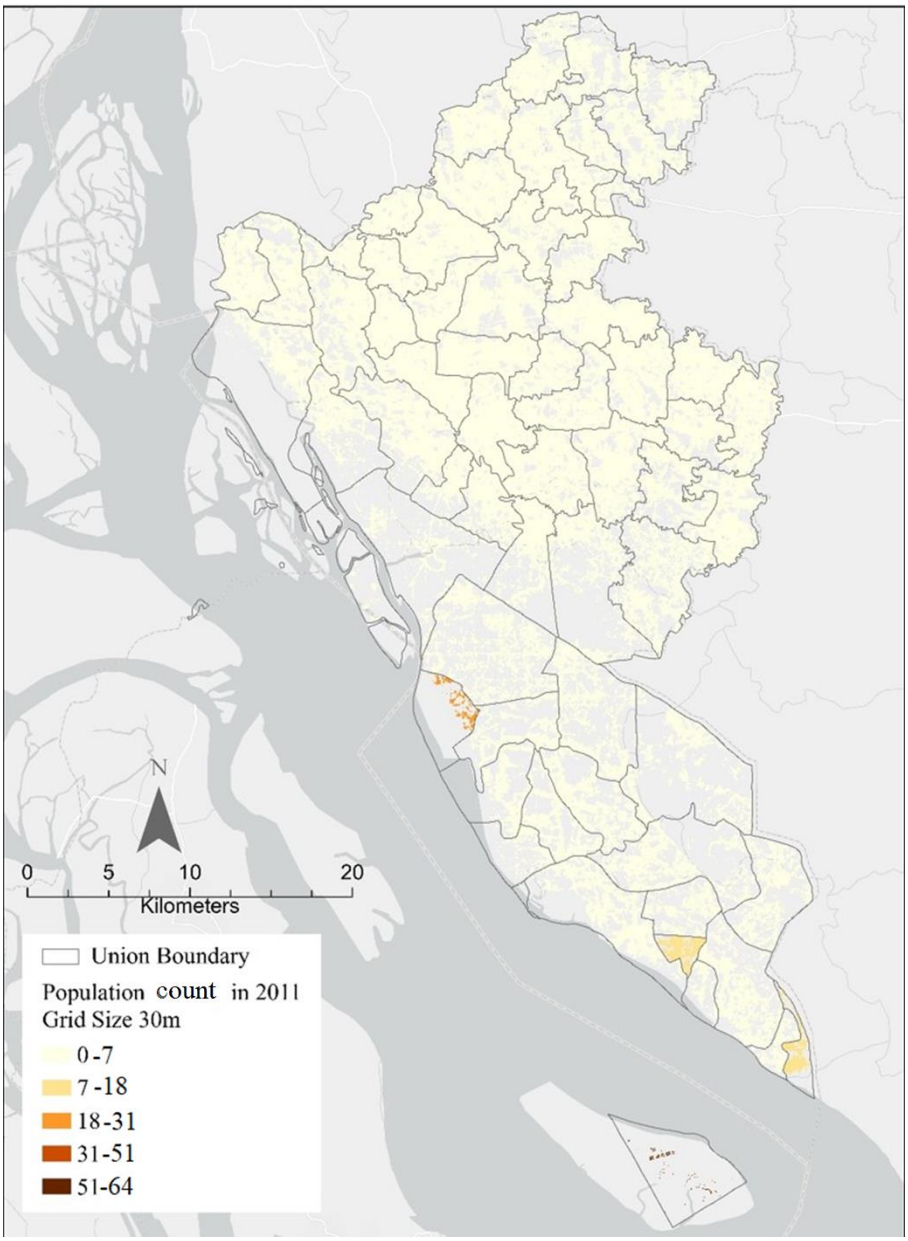


Figure 3.8 Dasymetric Population Distribution in 2011, 30 m vector grids were converted to 30 m raster grids. The 30 m raster surface has been symbolized using the Natural Jenks classification method. Original population densities have decimal places that have been rounded up in the legend to avoid representation of fractional population values. The 30 m raster surface has been symbolized using the Natural Jenks classification method. The raster surface represents population count (number of persons in each 30 m grid).

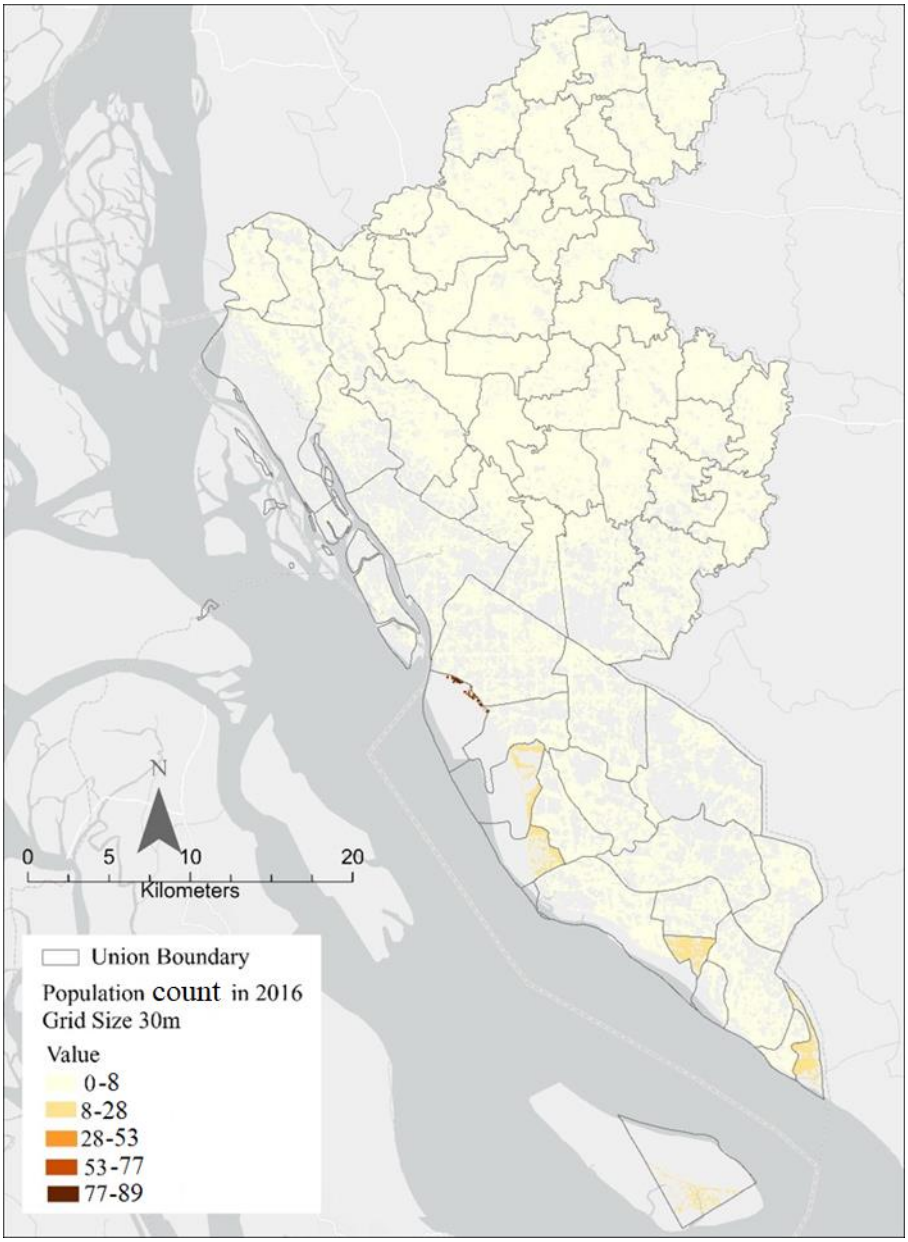


Figure 3.9 Dasymetric Population Distribution in 2016, 30 m vector grids were converted to 30 m raster grids. The 30 m raster surface has been symbolized using the Natural Jenks classification method. Original population densities have decimal places that have been rounded up in the legend to avoid representation of fractional population values. The raster surface represents population count (number of persons in each 30 m grid).

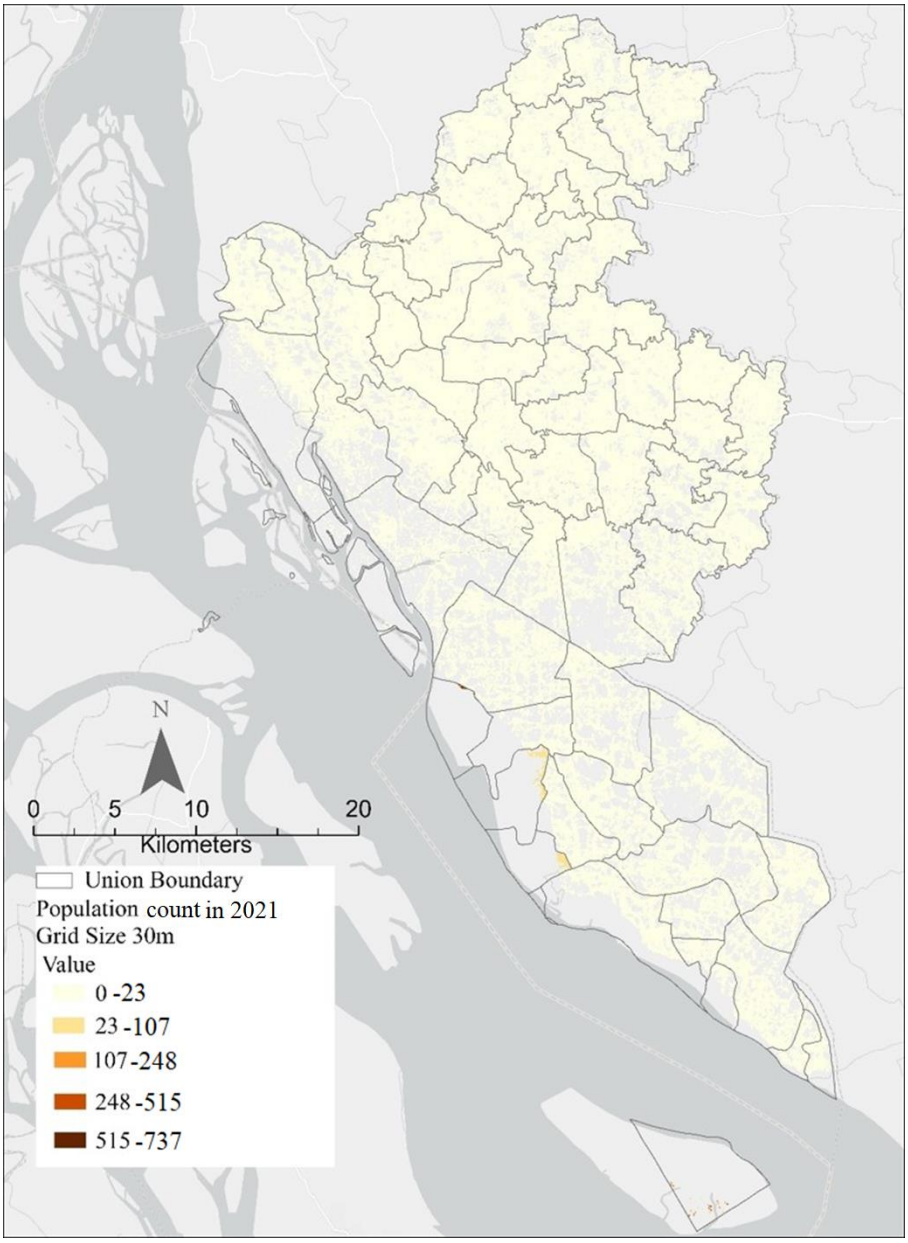


Figure 3.10 Dasymetric Population Distribution in 2021, 30 m vector grids were converted to 30 m raster grids. The 30 m raster surface has been symbolized using the Natural Jenks classification method. Original population densities have decimal places that have been rounded up in the legend to avoid representation of fractional population values. The raster surface represents population count (number of persons in each 30 m grid).

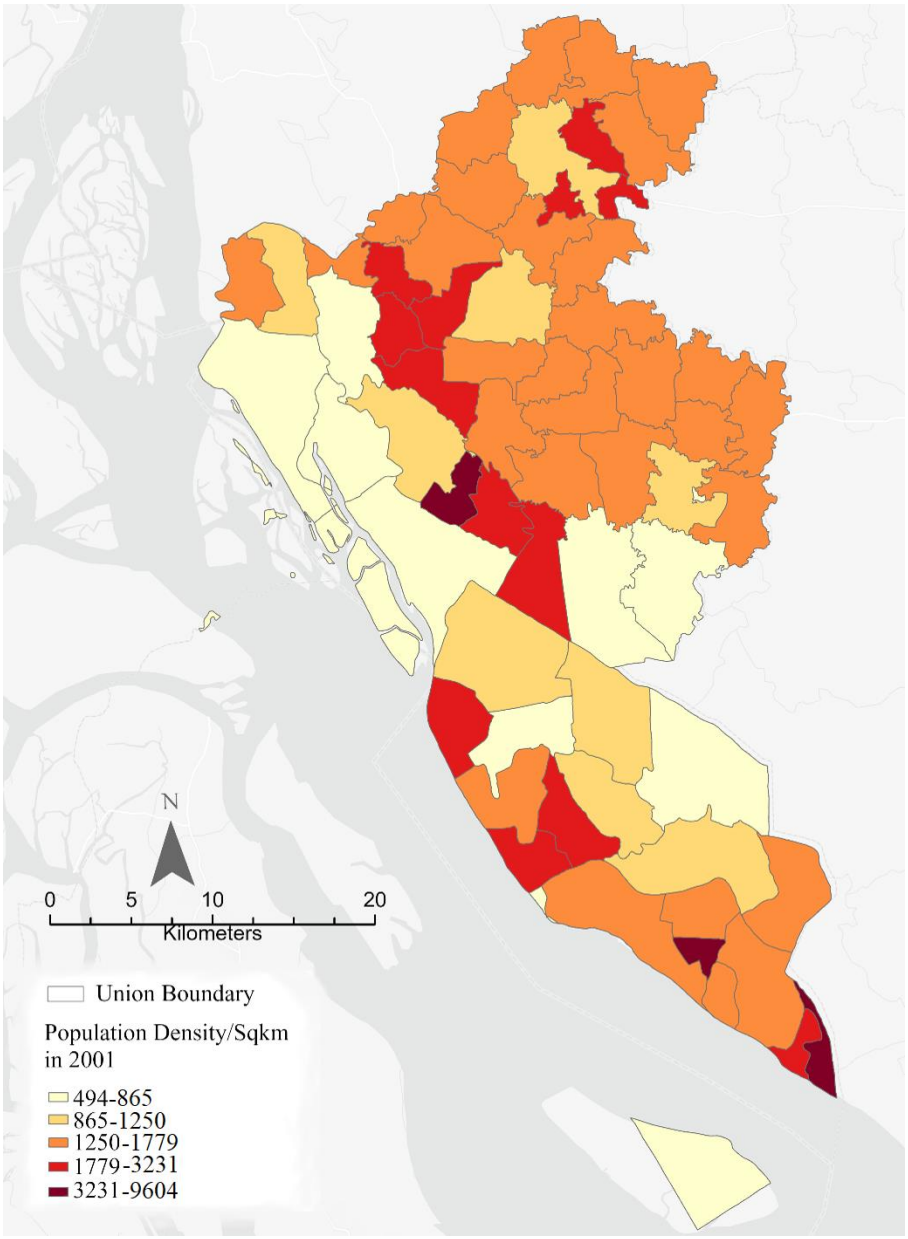


Figure 3.11 Choropleth map representing union-level population distribution in 2001. Legend for density has been categorized from low to high for simplification purpose and does not represent original density values. The lowest density value was 494 persons/$km^{2,}$ and the highest was 9604 persons/$km^{2.}$

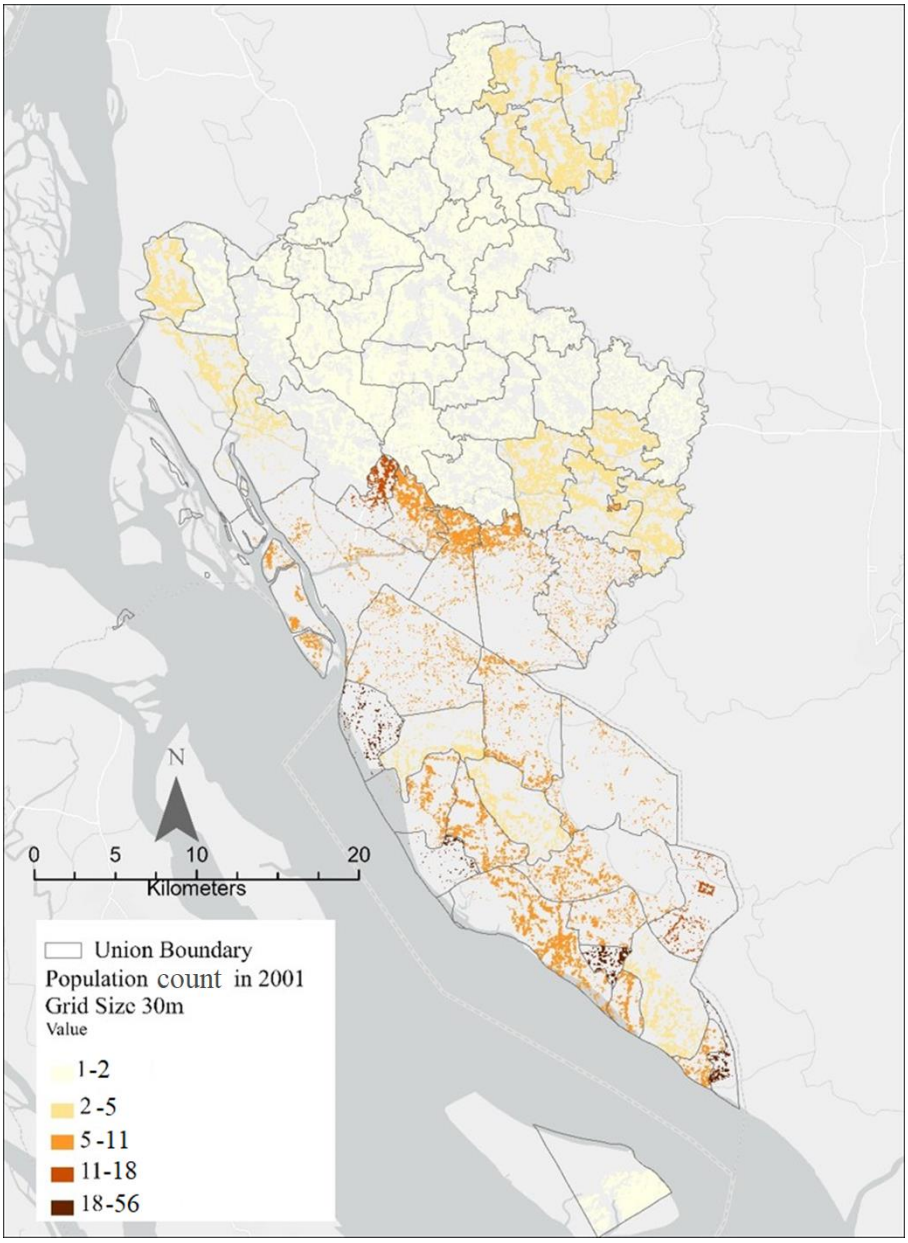


Figure 3.12 Dasymetric Population Distribution in 2001 using a raster-based approach. The 30 m raster surface has been symbolized using the Natural Jenks classification method. Original population densities have decimal places that have been rounded up in the legend to avoid representation of fractional population values. The raster surface represents population count (number of persons in each 30 m grid).

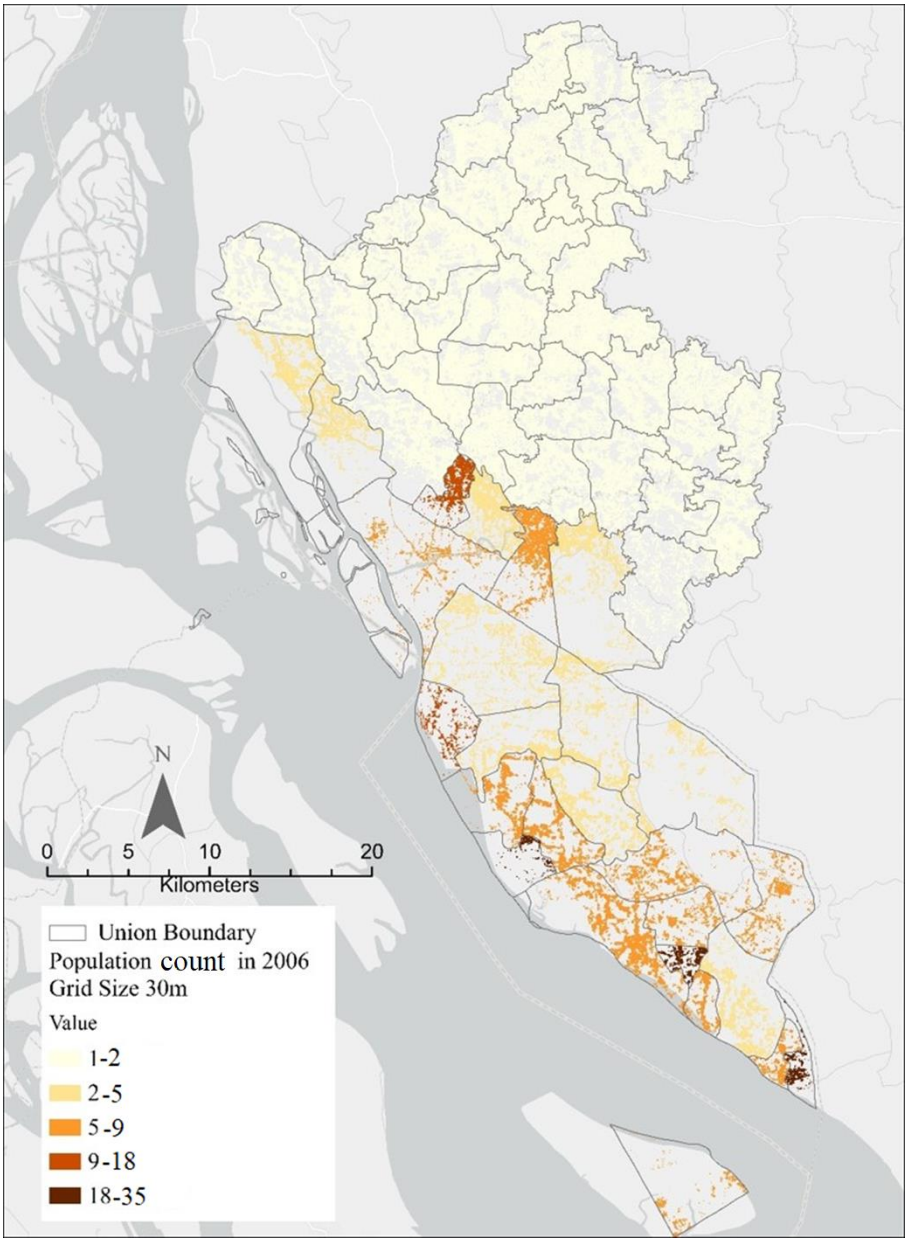


Figure 3.13 Dasymetric Population Distribution in 2006 using the raster-based approach. The 30 m raster surface has been symbolized using the Natural Jenks classification method. Original population densities have decimal places that have been rounded up in the legend to avoid representation of fractional population values. The raster surfac represents population count (number of persons in each 30 m grid) and population density.

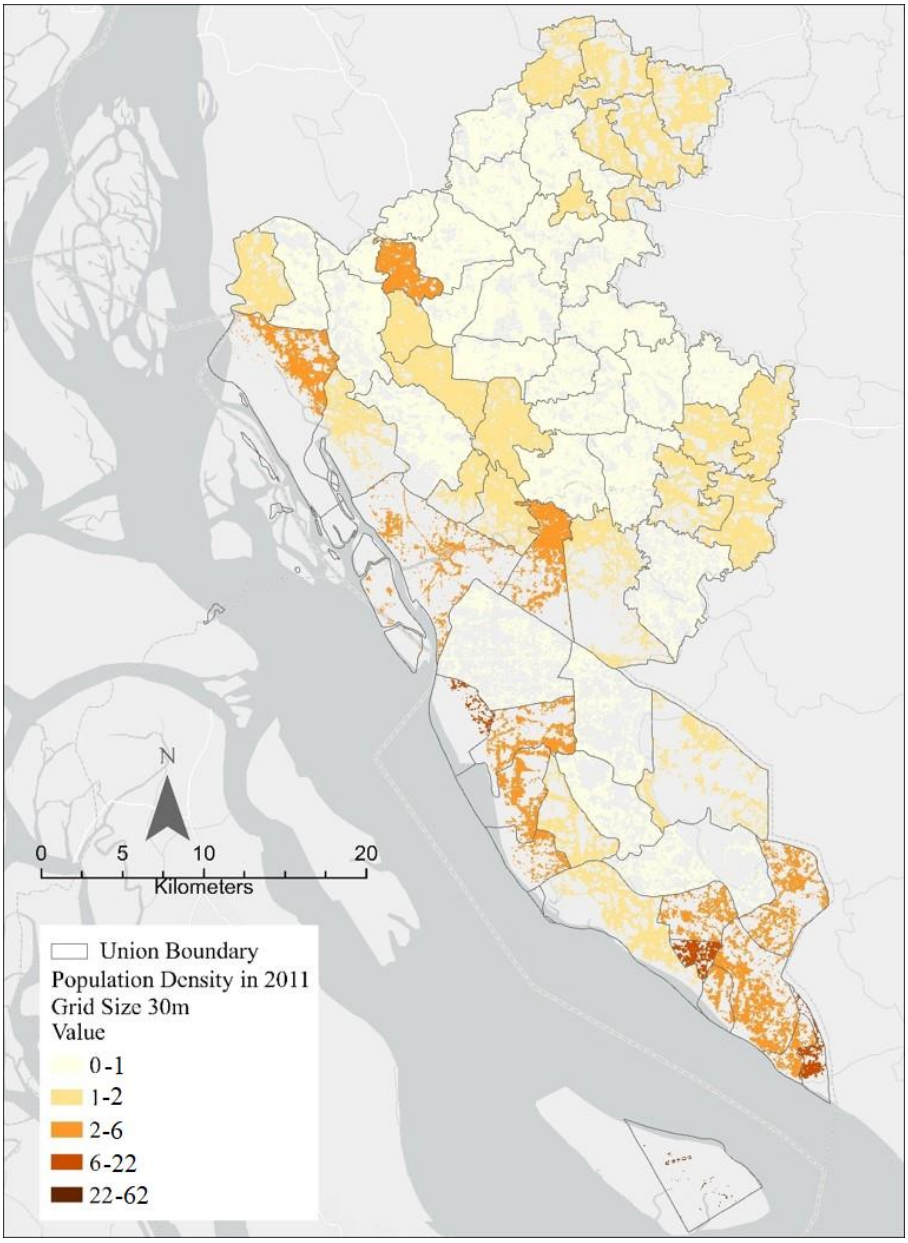


Figure 3.14 Dasymetric Population Distribution in 2011 using a raster-based approach. The 30 m raster surface has been symbolized using the Natural Jenks classification method. Original population densities have decimal places that have been rounded up in the legend to avoid representation of fractional population values. The raster surface represents population count (number of persons in each 30 m grid).

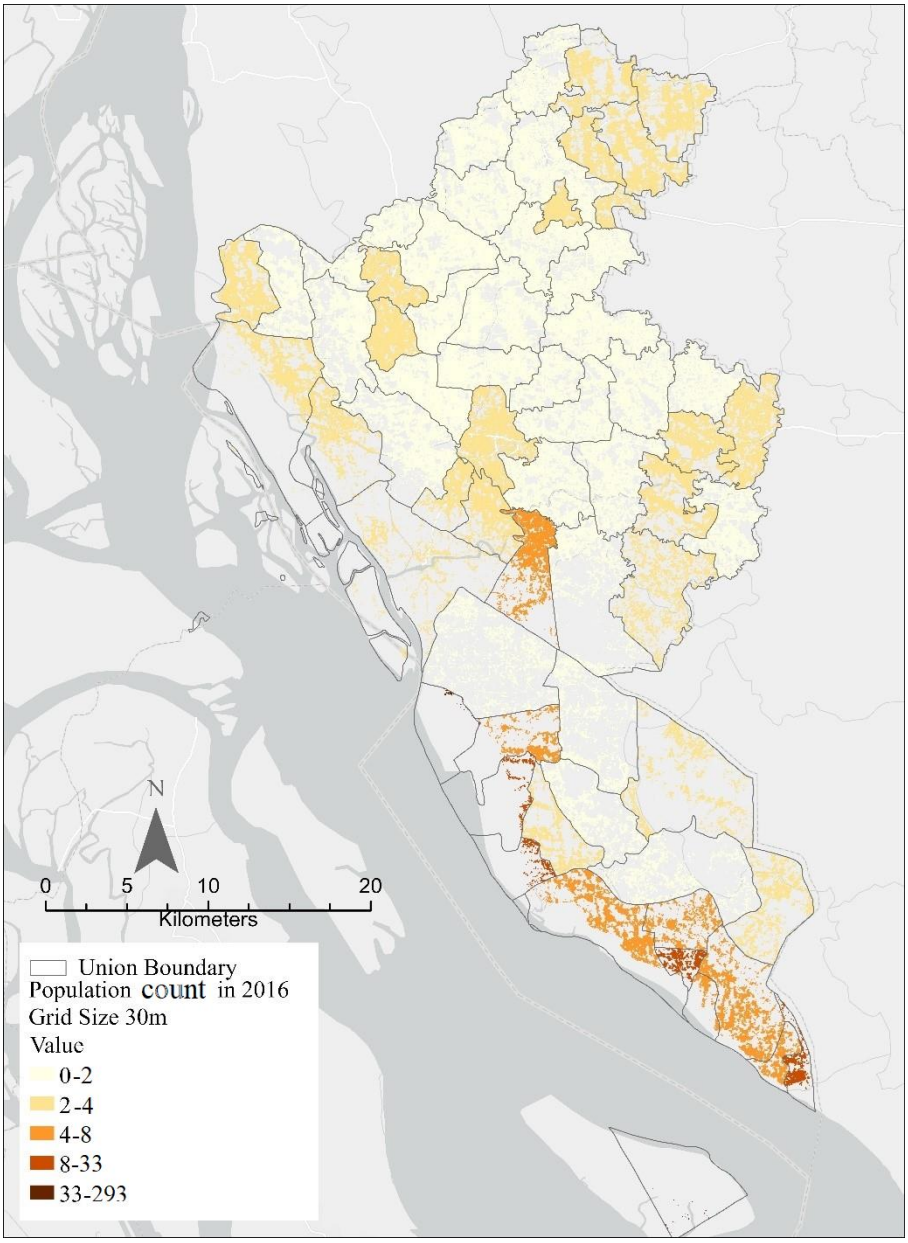


Figure 3.15 Dasymetric Population Distribution in 2016 using raster-based approach. The 30 m raster surface has been symbolized using the Natural Jenks classification method. Original population densities have decimal places that have been rounded up in the legend to avoid representation of fractional population values. The raster surface represents population count (number of persons in each 30 m grid).

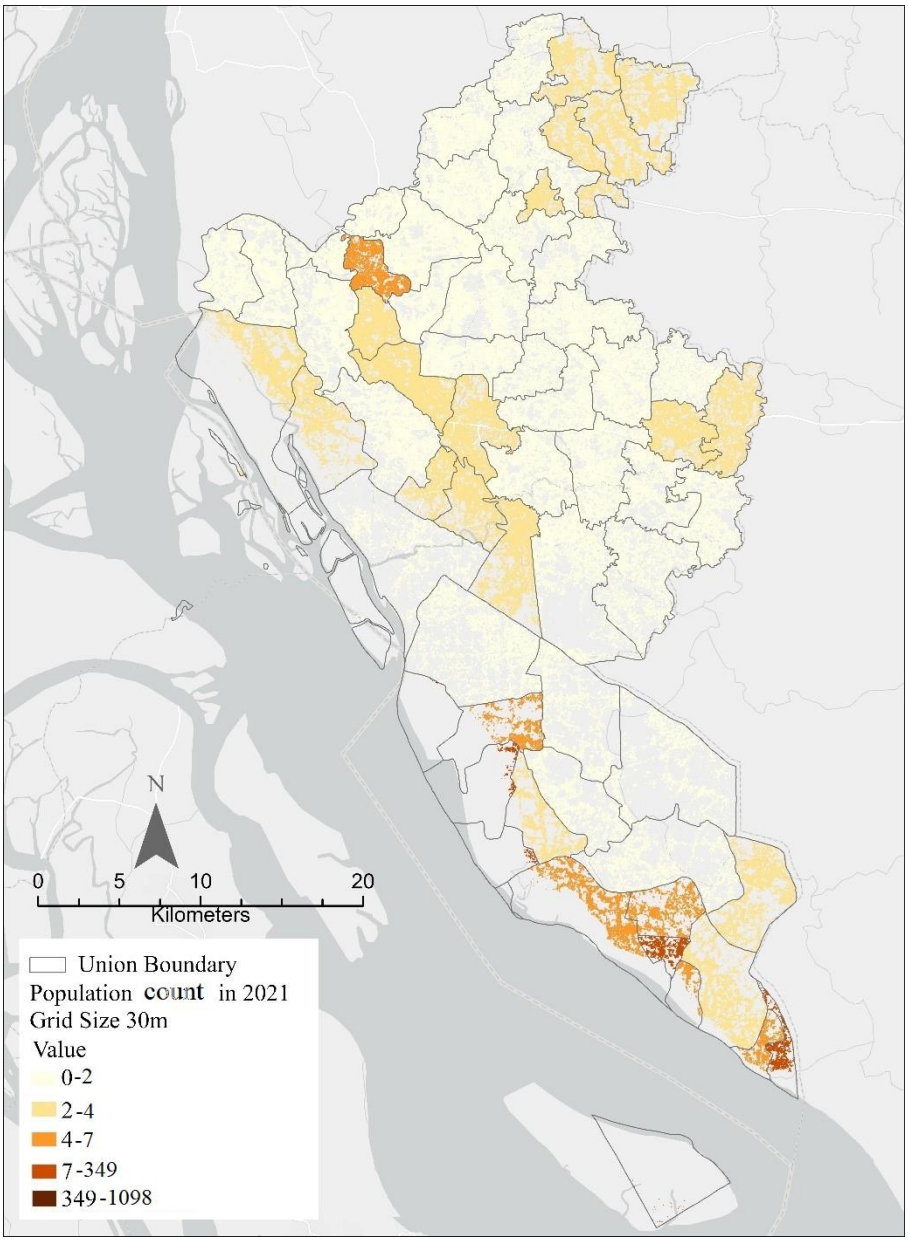


Figure 3.16 Dasymetric Population Distribution in 2021 using the raster-based approach. The 30 m raster surface has been symbolized using the Natural Jenks classification method. Original population densities have decimal places that have been rounded up in the legend to avoid representation of fractional population values. The raster surface represents population count (number of persons in each 30 m grid).

Table 3.17 Union level and Dasymetric level population count using the vector-based dasymetric approach. Equal values indicate volume preservation.

Year	Union Level Total Population	Dasymetric Level Total Population
2001	1,706,858	1,666,537
2006	1,643,838	1,643,838
2011	1,672,552	1,672,552
2016	2,115,626	2,115,626
2021	2,239,718	2,239,718

Table 3.18 Union level and Dasymetric level population count using the raster-based dasymetric approach. Equal values indicate volume preservation.

Year	Union Level Total Population	Dasymetric Level Total Population
2001	1,706,858	1,728,554
2006	1,643,838	1,643,838
2011	1,672,552	1,672,552
2016	2,115,626	2,119,186
2021	2,239,718	2,239,718

Table 3.19 Amount of land loss and homestead forest loss to river erosion every five years from 2001-2021

Year	Amount of Land loss in Sqkm	Amount of Homestead Forest loss in Sqkm
2001-2006	4.93	0.36
2006-2011	19.57	2.33
2011-2016	30.3	9.06
2016-2021	16.53	5.64
Max	**30.3**	**9.06**
Min	**4.93**	**0.36**
Average	**3.39**	**0.82**

Table 3.20 Estimated Population Displacement Due to RiverBank Erosion Vector-based Binary Dasymetric Approach

Year	Number of Displaced People	Maximum	Minimum	Total Population	Population Displacement (in percent)
2001-2006	1844		**1844**	1,706,858	0.10
2006-2011	51703			1,643,838	3.14
2011-2016	61579			1,672,552	3.68
2016-2021	86107	**86107**		2,115,626	4.07
2001-2021 Average	**9582**				

Table 3.21 Estimated Population Displacement Due to RiverBank Erosion Raster-based Binary Dasymetric Approach

Year	Number of Displaced People	Maximum	Minimum	Total Population	Population Displacement (in percent)
2001-2006	5241		**5241**	1,706,858	0.30
2006-2011	51703			1,643,838	3.14
2011-2016	62478			1,672,552	3.73
2016-2021	63453	**63453**		2,115,626	2.99
2001-2021 Average	**8708**				

CHAPTER 4

DISCUSSION

4.1 Landcover Classification

Our overall classification accuracies range from 79.04% to 87 % (See Chapter 3 Table 3.1-3.15). Our most important class, homestead forest, which needs to be accurately classified for dasymetric mapping, had the highest user and producer accuracy (90% to 93.75%, See Chapter 3, Table 3.1-3.15) throughout the five different study years 2001, 2006, 2011, 2016, and 2021. Although we observed overestimation in areas where homestead forest is a dominant class compared to fallow land and water bodies, we accepted those results due to low error of commission, ranging from 5% to 20% (Chapter 3, Table 3.1-3.15). Among all the classes, the other vegetation class was the most frequently misclassified. The presence of crop residues on some fallow lands in our study area and the presence of waterbodies with aquatic plants created confusion with other vegetation and, over the years, produced frequent high errors of commission for that class. Although the image search was deliberately restricted to the dry season, our study area's complex farming system led to misclassification, resulting from heterogeneity within the classes. For instance, fallow lands had high intra-class variability due to inconsistent soil moisture in irrigated and non-irrigated lands, which led to confusion with water bodies. The reflectance of other vegetation classes was sometimes identical to fallow land and homestead forest due to the presence of remains of harvested crops and winter crops. These ambiguities are due to large intra-class variability, consistent with other studies (Jin et al., 2018; Magidi et al., 2021), in complex agricultural environments where within-class heterogeneity exists. Despite these misclassifications, we are satisfied with the classification of the homestead

forest that we will be using for dasymetric mapping for having high producer and user accuracies.

The most significant variables across the classifications are near-infrared and mid-infrared bands (Appendix A, Fig: A1-A5), consistent with the findings of Abdi, A. M. (2020) and Phan et al. (2020). Their study found NIR as an important variable after topographic variables and argued that the importance is attributable to NIR's capability of discerning vegetation classes in scenes. Our study area is primarily rural, and our classes are predominantly vegetation classes, such as homestead forest, fallow land, and other vegetation, which reinforces the importance of near-infrared and mid-infrared in classifying scenes that mostly contain vegetation classes. NIR reflectance is influenced by canopy leaf structure and chlorophyll concentration (Gausman, 1985; Ray et al., 1993). In our study area, homestead forests are mainly canopies planted at rural residences, substantiating the importance of the NIR band.

Numerous landcover classification literature (Hurskainen et al., 2019; Khalmurzayeva, 2019; Hosseiny et al., 2022) have used auxiliary variables such as elevation, textural properties, and spectral indices to improve land use/landcover classification. Dobrinić et al. (2021) demonstrated that incorporating spectral indices such as normalized difference vegetation index (NDVI) and soil-adjusted vegetation index (SAVI) helps separate forests from other land cover. The author argued that SAVI is especially important in sparsely vegetated areas where a considerable amount of reflection is coming from background soil. In the case of our study site, both NDVI and SAVI might be useful as we have both dense canopy cover and sparsely vegetated areas. Phan et al. (2020) found elevation to be an important variable in separating other vegetation and grassland, other vegetation, and settlements. Similar to their study area, our study site includes settlements and other vegetation classes; in the future, adding elevation as an

auxiliary variable might provide better results. Our current classifications misclassified water bodies and fallow lands due to the presence of irrigated land in our study area. Incorporating the SAVI index or downscaled soil moisture active passive (SMAP) data might help separate waterbodies and irrigated land.

Drawing training points and determining classes based on spectral separability is crucial to obtaining satisfactory classification results. Our training points of defined classes have values of more than 1.8(Appendix A, Table A4- A8), indicating that the classes are spectrally separable. We combined grassland and dispersed medium and small plants, which were not homestead forests, as other vegetation classes for having class separability less than 1.8. As a result, other vegetation, being a broad class, had mixed spectral signature with homestead forest and showed low-class separability. As stated earlier, prior studies (Hurskainen et al., 2019; Khalmurzayeva, 2019; Hosseiny et al., 2022) improved classification using elevation and spectral indices; we assume that for our study site, including those auxiliary variables might improve class separability leading to more detailed classification. Besides, using higher spectral resolution data than Landsat might also help improve the land use classification. For instance, Sentinel 2-MSI has narrow spectral bands, red edge, NIR, and SWIR, capable of carrying out species-level vegetation classification (Chaves et al., 2020). In the future, using such fine spectral resolution imagery might help separate that broad class into more detailed classes.

Our ground reference data points was drawn from SPOT Satellite image and Google Earth Imagery (See Chapter 3). Google Earth Imagery might often be misleading as Google Earth provides low-resolution historical Imagery for our study area. In the future, access to higher spatial-resolution historical imagery will help construct more accurate confusion matrices.

We used 100 ground reference points. Increasing the number of ground reference points and collecting them using field surveys might increase classification accuracy.

4.2 Landcover Change

In terms of landcover change, the major change took place from 2001-2021 in the fallow landcover class. A total of 341.81 sq. km of fallow land was transformed into homestead forests and water bodies. Islam et al. (2023) conducted a land use land cover change analysis from 1988-2020 in Kamalnagar and Ramgati subdistrict and found that agricultural lands are decreasing due to river erosion. A study conducted in 2014-15 by Hasnat et al. (2016) surveyed 120 households in those subdistricts and found that 39 percent of people think that the primary reason behind agricultural land degradation is riverbank erosion. These studies were conducted in only two subdistricts of Lakshmipur district and do not represent the scenario of the Lakshmipur district. Our findings also support that a significant amount of fallow land, 132 sqkm of fallow land/agricultural land (we defined both agricultural land and fallow land as fallow land), was transformed into waterbodies and went into the river from 2001-2021. However, at the same time, a massive amount of 276.65 sq km of fallow land/agricultural land was also transformed into homestead forests. In our study, the homestead forest is a proxy of human settlement. As a result, we assume that increasing population and population displacement due to riverbank erosion could be related to the decline of agricultural land in the Lakshmipur district. Crawford et al. (2021) found that people who lose their households to riverbank erosion mostly migrate to nearby places and build settlements, which aligns with our assumption that the people being displaced are also contributing to the transformation of agricultural land to human settlements. Future studies should incorporate household surveys and in-depth interviews to reveal the dynamics of land use change and population displacement due to riverbank erosion.

4.3 Population Mapping and Estimating Population Displacement Due to Riverbank Erosion.

The dasymetric maps (Fig 3.6 -3.10 and Fig 3.12-3.16) are the better representation of population density over the choropleth map (Fig 3.11) for disaggregating populations in known habitable areas. Our vector-based dasymetric maps displayed heterogeneous population distribution within the unions, and raster-based calculation provided heterogeneous distribution among the unions. The reason behind this is that the vector-based method provides flexibility to calculate the areal proportion of each habitable unit within each union, which creates different areal proportions of habitable areas within the unions. Therefore, when we multiplied the areal proportion with the Union-level populations, it gave heterogeneous density to some extent. On the other hand, the raster-based approach provides the opportunity to calculate the total proportion of total habitable areas instead of individual habitable areas within a union. Since our raster classification maps contained pixels from only a single class of habitable area, which was homestead forests, the raster calculation assigned populations to only that class, which resulted in homogeneous distributions. Heterogeneity in the raster approach would be possible in the presence of multiple habitable classes within the unions. This observation is aligned with Wood's (2006) argument that dasymetric maps might occasionally represent homogeneous population density due to the limited variability in land cover data.

Although a vector-based approach was provided within union heterogeneity to some extent, it also exhibited homogeneity in some areas. This is primarily attributable to the use of 30 m resolution Landsat imagery. Using Landsat imagery, we could not classify settlements or other land use classes than homestead forests that indicate human presence in the upper right portion of the district (See Chapter 3, Fig 3.1-3.5). Therefore, some unions, especially unions in the upper portion of the maps (See Chapter 3, Fig 3.1-3.5), that are entirely covered by homestead

forests only contain the class homestead forest, leaving no space for uninhabited areas. As a result, the areal interpolation equations contain equal areas of homestead forests as source and target zones in equation 2.3 (See Chapter 2), resulting in equal nominators and denominators in equation 2.7. This produces a homogeneous population density across the unions like a choropleth map (See Chapter 3, Fig 3.6-3.10). For this reason, the upper portion of the vector-based maps shows a higher density of population distributions compared to the raster-based maps.

The second reason is the data aggregation in the final step of the vector-based method (See Appendix C, Step 8). The data aggregation created some smoothing effects. However, some unions, especially unions that are situated in proximity to the riverbank, still show heterogeneity in distribution. Due to the aggregation process, the entire vector-based process might be called areal interpolation instead of dasymetric mapping (Eicher & Brewer, 2001). However, there is evidence in the existing literature (Eicher & Brewer, 2001) that supports the method to be called a dasymetric method for using ancillary data.

Our estimated number of displaced population ranges from 1844 (0.10 percent of the total population) to 86107(4.07 percent of the total population) from 2001 to 2021 using a vector-based approach (See Chapter 3, Table 3.20) and 5241(0.3 percent of the total population) to 63453(2.9 percent of the total population) using a raster-based approach (See Chapter 3, Table 3.21). We found an increasing trend of population displacement over the years and assume that this is related to the increased population and high riverbank erosion rate. According to BBS (2011) and our population projection, over 21 years, the total 532,860 people increased in the Lakshmipur district, which also increased the displaced population. Crawford et al. (2021) found that between 2008 and 2018, the area experienced extreme erosion with an EPR (endpoint rate)

of 300m per year, supporting our findings of high population displacement. Their study also conducted a household survey creating stratified random sampling, and referring to BBS (2011), they mentioned that in their surveyed area, which is a small portion of our study area, around 40,000 people were living. Our estimate suggests that from 2011-2016, using a raster-based approach, 62,478 people have been displaced due to riverbank erosion, which is a reasonable estimate considering the extent of our study area.

The amount of land loss and homestead forest loss is highest from 2011-2016, 30.3 sq. km (See Chapter 3, Table 3.19) and 9.6 sq. km; however, the population displacement due to riverbank erosion is highest from 2016-2021. The reason might be attributed to the increased population in 2016 in this area (Table 3.17-3.18). The number of population displacements cannot be verified due to the nonexistence of ground validation data. Besides, as our study calculates population displacement every five years, comparing the displaced population with the yearly population would not give an accurate assumption of yearly displacement. A finer-level temporal analysis is needed to compare yearly displacement with total population change.

Although our Dasymetric maps offer an actual representation of where people live, there are uncertainties associated with them. The sources of these uncertainties are uncertainty in population data and uncertainty in the relationship between ancillary data and population distribution (Nagle et al., 2014). Although these uncertainties cannot be completely controlled, they can be measured. Nagle et al. (2014) highlight the Penalized Maximum Entropy Dasymetric Model (P-MEDM) method to address the uncertainties in dasymetric mapping. Although we do not delve into the implications of those uncertainties, we acknowledge their existence. Exploring the uncertainties in our dasymetric maps using the method from Nagle et al. (2014) is a potential future direction of this research.

4.4 References

BBS (2011). Statistical yearbook of Bangladesh, Ministry of Planning, Dhaka, Bangladesh.

Crawford, T. W., Rahman, M. K., Miah, M. G., Islam, M. R., Paul, B. K., Curtis, S., & Islam, M. S. (2021). Coupled adaptive cycles of shoreline change and households in deltaic Bangladesh: analysis of a 30-year shoreline change record and recent population impacts. *Annals of the American Association of Geographers*, *111*(4), 1002-1024.

Dobrinić, D., Gašparović, M., & Medak, D. (2021). Sentinel-1 and 2 time-series for vegetation mapping using random forest classification: A case study of Northern Croatia. *Remote Sensing*, *13*(12), 2321.

ED Chaves, M., CA Picoli, M., & D. Sanches, I. (2020). Recent applications of Landsat 8/OLI and Sentinel-2/MSI for land use and land cover mapping: A systematic review. *Remote Sensing*, *12*(18), 3062.

Eicher, C. L., & Brewer, C. A. (2001). Dasymetric mapping and areal interpolation: Implementation and evaluation. *Cartography and Geographic Information Science*, *28*(2), 125-138.

Eklundh, L., Harrie, L., & Kuusk, A. (2001). Investigating relationships between Landsat ETM+ sensor data and leaf area index in a boreal conifer forest. Remote sensing of Environment, 78(3), 239-251.

Gausman, H. W. (1985). Plant Leaf Optical Parameters in Visible and Near-Infrared Light, Graduate Studies No. 29 (Lubbock, Texas).

Hasnat, M. A., Hossain, N., Muhibbullah, M., Sarwar, M. D., & Shormin, T. (2016). Impacts of climate change on agriculture and changing adaptive strategies in the coastal area of Lakshmipur district, Bangladesh. *Current World Environment*, (3), 700.

Hosseiny, B., Abdi, A. M., & Jamali, S. (2022). Urban land use and land cover classification with interpretable machine learning–A case study using Sentinel-2 and auxiliary data. *Remote Sensing Applications: Society and Environment*, *28*, 100843.

Hurskainen, P., Adhikari, H., Siljander, M., Pellikka, P. K. E., & Hemp, A. (2019). Auxiliary datasets improve accuracy of object-based land use/land cover classification in heterogeneous savanna landscapes. *Remote sensing of environment*, *233*, 111354.

Immitzer, M., Vuolo, F., & Atzberger, C. (2016). First experience with Sentinel-2 data for crop and tree species classifications in central Europe. *Remote sensing*, *8*(3), 166.

Islam, M. S., Crawford, T. W., & Shao, Y. (2023). Evaluation of predicted loss of different land use and land cover (LULC) due to coastal erosion in Bangladesh. *Frontiers in Environmental Science*, *11*, 479.

Jin, Y., Liu, X., Chen, Y., & Liang, X. (2018). Land-cover mapping using Random Forest classification and incorporating NDVI time-series and texture: A case study of central Shandong. *International journal of remote sensing*, *39*(23), 8703-8723.

Khalmurzayeva, G. (2019). Land Cover Classification using Random Forest Technique and DEM Auxiliary Data. *International Journal of Geoinformatics*.

Magidi, J., Nhamo, L., Mpandeli, S., & Mabhaudhi, T. (2021). Application of the random forest classifier to map irrigated areas using google earth engine. *Remote Sensing*, *13*(5), 876.

Nagle, N. N., Buttenfield, B. P., Leyk, S., & Spielman, S. (2014). Dasymetric modeling and uncertainty. *Annals of the Association of American Geographers*, *104*(1), 80-95.

Paul, B. K., Rahman, M. K., Crawford, T., Curtis, S., Miah, M. G., Islam, M. R., & Islam, M. S. (2020). Explaining mobility using the Community Capital Framework and Place Attachment concepts: A case study of riverbank erosion in the Lower Meghna Estuary, Bangladesh. *Applied Geography*, *125*, 1021

Phan, T. N., Kuch, V., & Lehnert, L. W. (2020). Land cover classification using Google Earth Engine and random forest classifier—The role of image composition. Remote Sensing, 12(15), 2411.

Ray, T. W., Murray, B. C., Chehbouni, A., & Njoku, E. (1993, October). The red edge in arid region vegetation: 340-1060 nm spectra. In *Summaries of the 4th Annual JPL Airborne Geoscience Workshop. Volume 1: AVIRIS Workshop*.

Sleeter, R., & Wood, N. (2006). Estimating daytime and nighttime population density for coastal communities in Oregon. In 44th Urban and Regional Information Systems Association Annual Conference, British Columbia (pp. 1-15).

Wang, D., Wan, B., Qiu, P., Su, Y., Guo, Q., Wang, R., & Wu, X. (2018). Evaluating theperformance of Sentinel-2, Landsat 8 and Pléiades-1 in mapping mangrove extent and species. *Remote Sensing*, *10*(9), 14

CHAPTER 5

CONCLUSION

We aimed to examine land cover change and population displacement due to riverbank erosion in the Lakshmipur district, Bangladesh. The landcover change result indicates that a significant amount of fallow land, 276.65 sq km, was transformed into homestead forests, and 132 sq km of fallow land was transformed into waterbodies, indicating changes due to river erosion and population displacement. Regarding population displacement, we found an increasing trend of population displacement over the years and assume that this is related to the increased population and high riverbank erosion rate. Besides, we observed population underestimation in one year using a vector-based dasymetric mapping approach and overestimation in two years using a raster-based dasymetric mapping approach. The errors in dasymetric mapping might result from several sources. An error might be inherited from an error in ancillary landcover data (Mennis, 2003; Mennis, 2009), as our ancillary landcover data is not 100 % accurate. Besides, our landcover class does not have many classes that can be used as a proxy for human footprint. The limitations also include difficulty defining high, moderate, and low-density residential areas using homestead forest as a proxy. Landcover data created from high-resolution satellite imagery and availability of additional ancillary data, e.g., land parcel level population data, could help define high, moderate, and low-resolution residential and improve overall estimation, producing a map with more variance in population density.

This study will help replicate the process of dasymetric mapping in similar geographical settings. The population maps represent population distribution in the entire district, benefitting natural resource management, facility planning and distribution, disaster management and planning, and enhancing community resilience. In addition, the most recent population map can

be used in emergency evacuations in case of any potential natural disaster event that may disrupt people and their livelihoods on the coast of Lakshmipur and beyond as appropriate.

5.1 References

Mennis, Jeremy. "Dasymetric mapping for estimating population in small areas." *Geogr Compass* 3, no. 2 (2009): 727-745.

Mennis, J. (2003). Generating surface models of population using dasymetric mapping. *The Professional Geographer*, *55*(1), 31-42.

APPENDIX A

Table A.1Satellite data description

Year	Image ID	Acquisition date	Correction level	Cloud Coverage
2001	LANDSAT/LT05/C02/T1_TOA/LT05_137044_20010121	01/21/2001	TOA	0%
2006	LANDSAT/LT05/C02/T1_TOA/LT05_137044_20060204	02/04/2006	TOA	0%
2011	LANDSAT/LT05/C02/T1_TOA/LT05_137044_20110101	01/02/2011	TOA	0%
2016	LANDSAT/LC08/C02/T1_TOA/LC08_137044_20160115	01/15/2016	TOA	0%
2021	LANDSAT/LC08/C02/T1_TOA/LC08_137044_20210317	03/17/2021	TOA	0%

Table A.2 Number of training points for each landcover class in different years. Training points were drawn inside Google Earth Engine based on our familiarity with the study area.

Year	Homestead forest	Fallow land	Other vegetation	Waterbody
2001	508	238	351	185
2006	285	263	85	300
2011	739	996	78	235
2016	703	812	611	354
2021	508	231	351	158

Table A.3Random Forest hyperparameter Settings.

Year	Training and Testing Data Split	Variable per split	Bag Fraction
2001	70/30	2	0.4
2006	70/30	2	0.5
2011	75/25	2	0.6
2016	75/25	2	0.5
2021	70/30	3	0.5

Table A.4 Table A4 Spectral Separability analysis of training data for the 2001 image. Values range from 0 to 2. Higher values indicate high spectral separability between classes. As other vegetation is a broad class low class separability has been observed between homestead forest and other vegetation.

	Homestead forest	Fallow land	Other vegetation	Waterbody
Homestead forest				
Fallow land	1.8			
Other vegetation	1.4	1.9		
Waterbody	1.9	1.9	1.9	

Table A.5 2006 Spectral separability analysis of training data for the 2006 image. Values range from 0 to 2. Higher values indicate high spectral separability between classes.

	Homestead forest	Fallow land	Other vegetation	Waterbody
Homestead forest				
Fallow land	1.9			
Other vegetation	1.78	1.94		
Waterbody	1.9	1.83	1.98	

Table A.6 Spectral separability analysis of training data for the 2011 image. Values range from 0 to 2. Higher values indicate high spectral separability between classes.

	Homestead forest	Fallow land	Other vegetation	Waterbody
Homestead forest				
Fallow land	1.9			
Other vegetation	1.64	1.94		
Waterbody	1.99	1.99	1.98	

Table A.7 Spectral separability analysis of training data for the 2016 image. Values range from 0 to 2. Higher values indicate high spectral separability between classes.

	Homestead forest	Fallow land	Other vegetation	Waterbody
Homestead forest				
Fallow land	1.99			
Other vegetation	1.98	1.99		
Waterbody	1.99	1.99	1.99	

Table A.8 Spectral separability analysis of training data for the 2021 image. Values range from 0 to 2. Higher values indicate high spectral separability between classes.

	Homestead forest	Fallow land	Other vegetation	Waterbody
Homestead forest				
Fallow land	1.99			
Other vegetation	1.87	1.88		
Waterbody	1.99	2.00	1.99	

Figure A1-A5 Random Forest variable of Importance. Smile random forest classifier has an importance function. Google Earth engine calculates variables of importance using Gini impurity. Figures A to E are the five graphs for 2001-2021

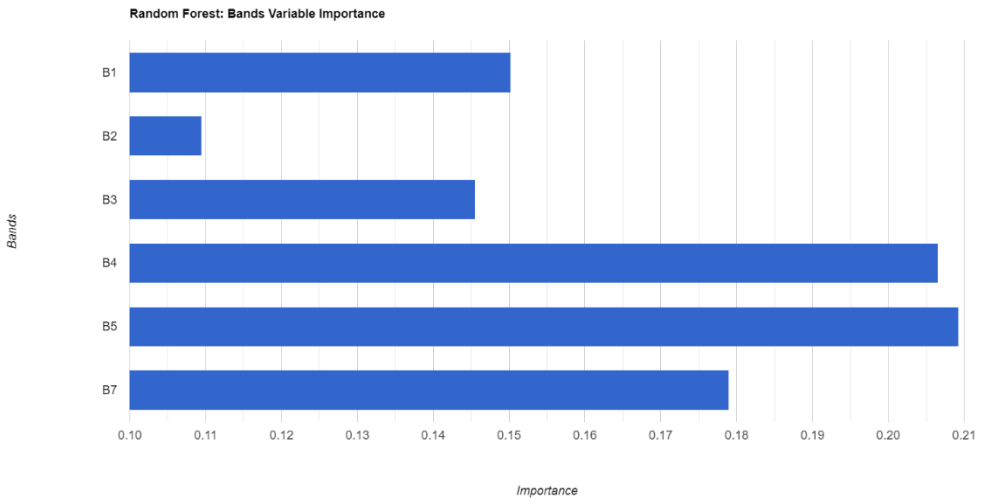


Figure A.1 Random forest variable of importance chart for classification 2001 using Landsat 5TM. The most important band is B5 (Near-infrared 1.55 - 1.75 µm), and the least important band is B2(Green 0.52 - 0.60 µm).

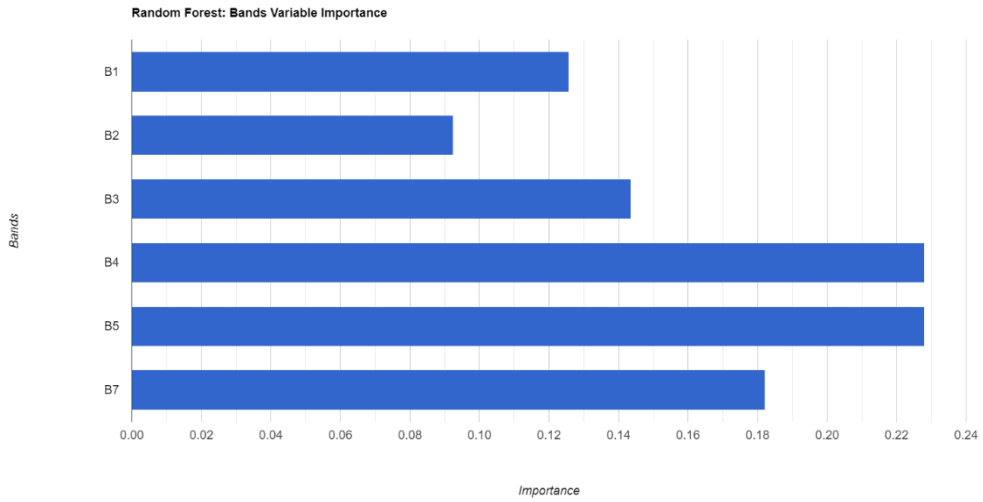


Figure A.2 Random forest variable of importance chart for classification 2006 using Landsat 5TM. The most important bands are B4 (Near-Infrared 0.76 - 0.90 µm) and B5 (Near-infrared 1.55 - 1.75 µm); the least important band is B2(Green 0.52 - 0.60 µm).

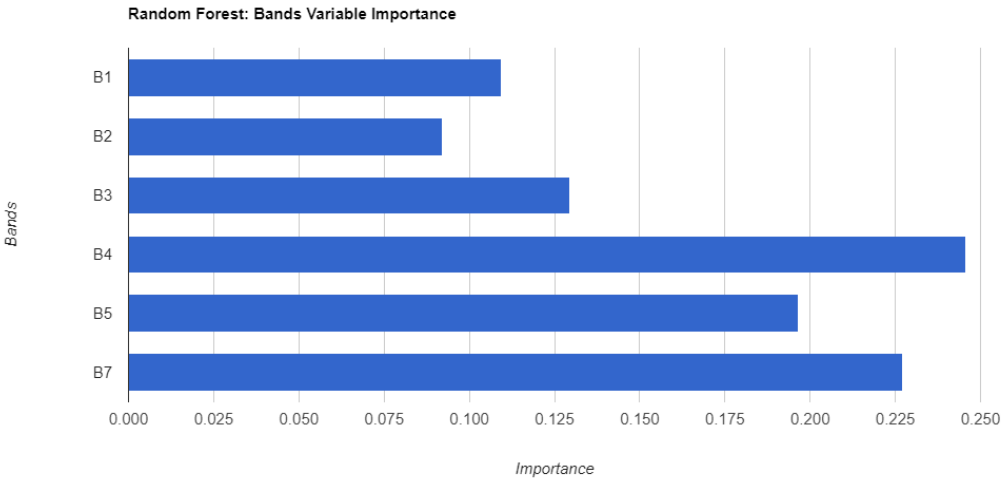


Figure A.3 Random forest variable of importance chart for classification 2011 using Landsat 5TM. The most important band is B5 (Near-infrared 55 - 1.75 µm) and the least important band is B2(Green 0.52 - 0.60 µm).

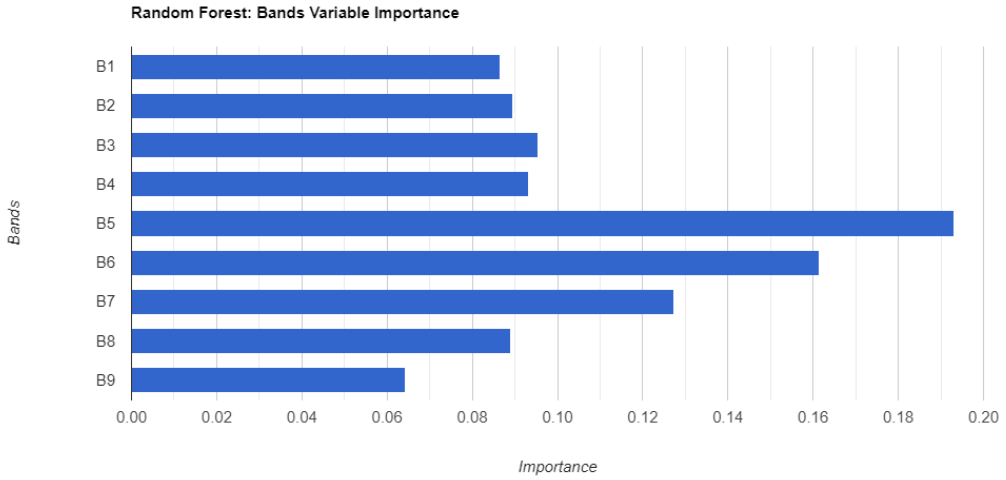


Figure A.4: Random Forest variable of importance chart for classification 2016 using Landsat 8 OLI. The most important band is B5 (Near Infrared 0.5-0.88 µm) and the least important band is B1 (coastal aerosol 0.43-0.45 µm).

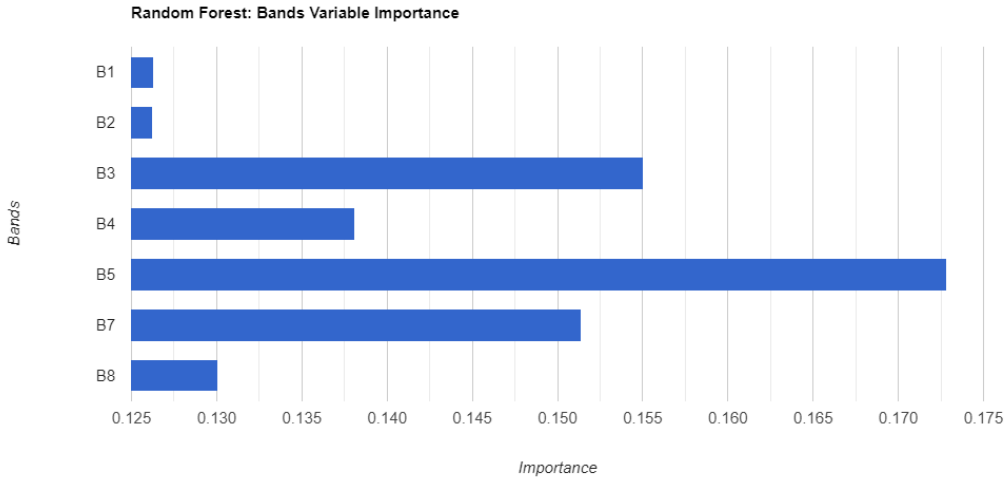


Figure A.5: Random Forest variable of importance chart for classification 2021 using Landsat 8 OLI. The most important band is B5 (Near Infrared 0.85-0.88 µm) and the least important band is B1 and B2 (Coastal aerosol 0.43-0.45 µm and Blue 0.45-0.51).

APPENDIX B

Figures Overlaying Dasymetric Maps with Shorelines.

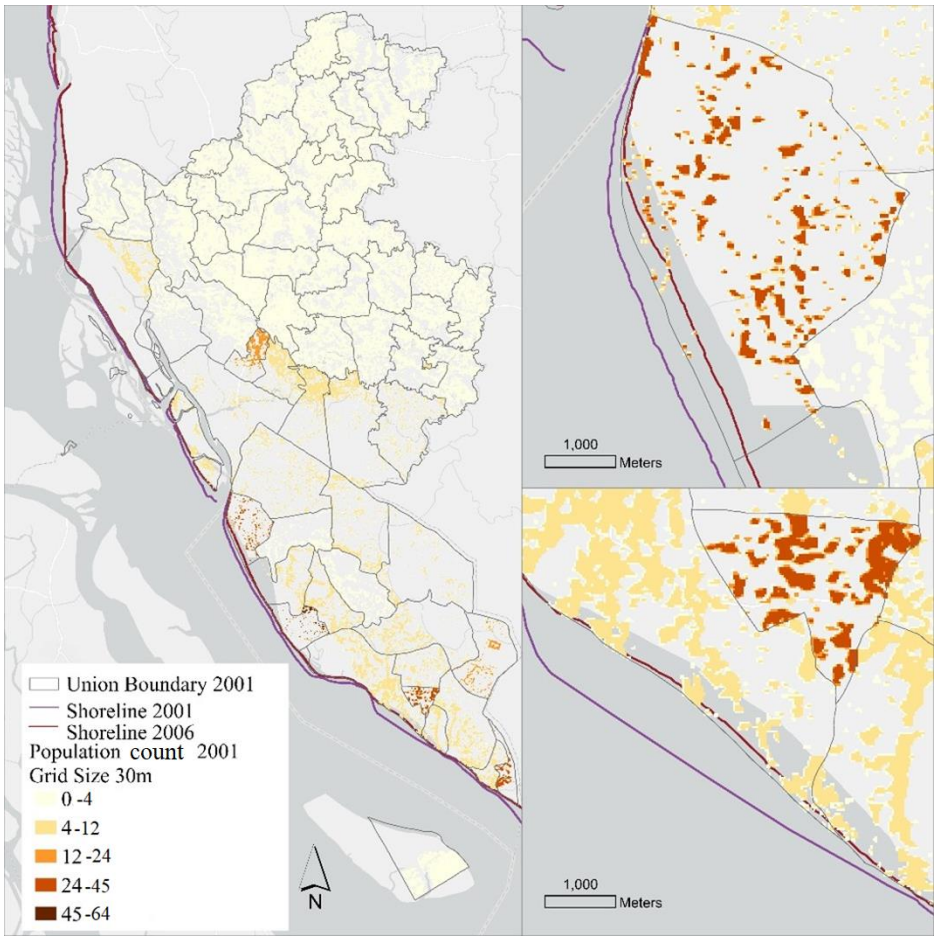


Figure B.1 Shorelines in 2001 and 2006 overlaid with the 2001 30 m raster surface of the population produced using vector-based approach.

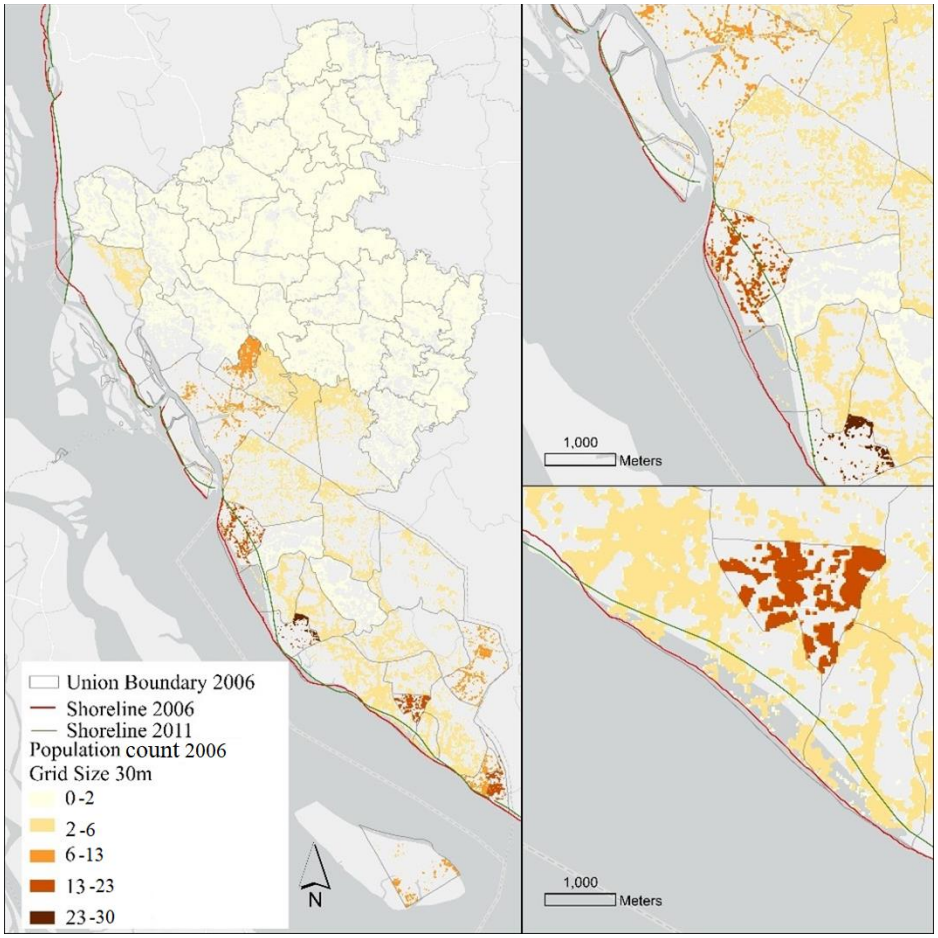


Figure B.2 Shorelines in 2006 and 2011 overlaid with 2006 30 m raster surface of population produced using vector-based approach.

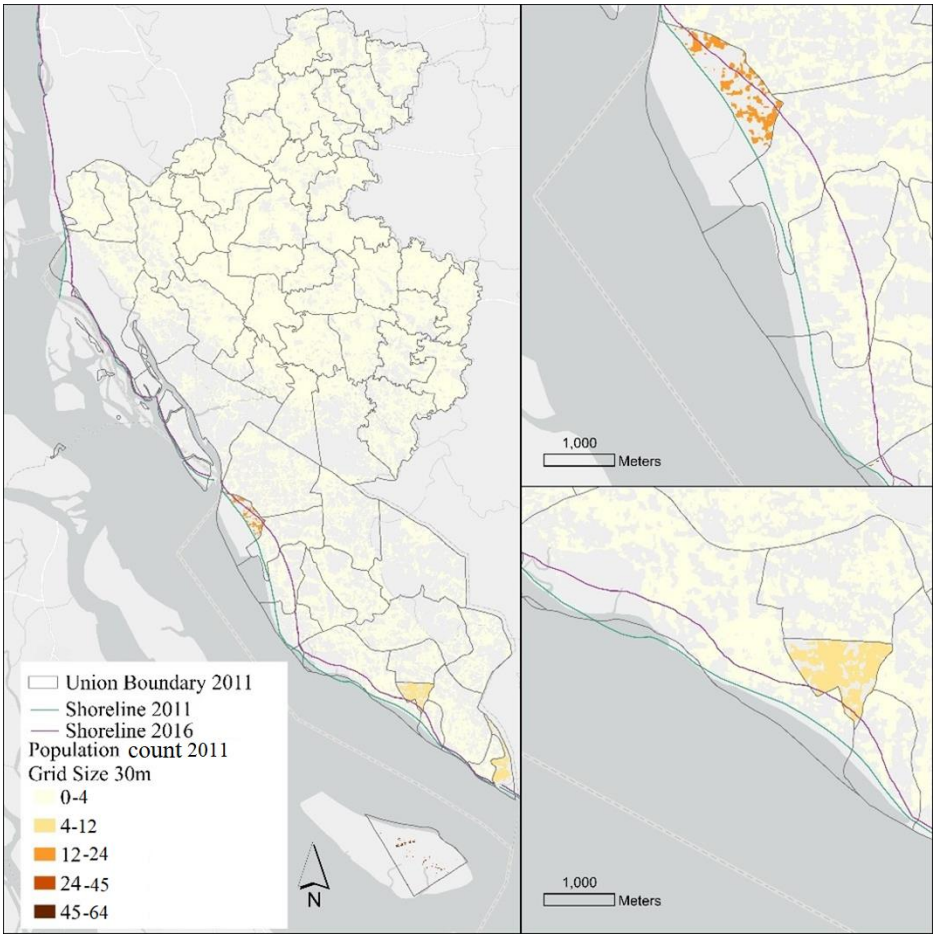


Figure B.3 Shorelines in 2011 and 2016 overlaid with the 2011 30 m raster surface of the population produced using a vector-based approach.

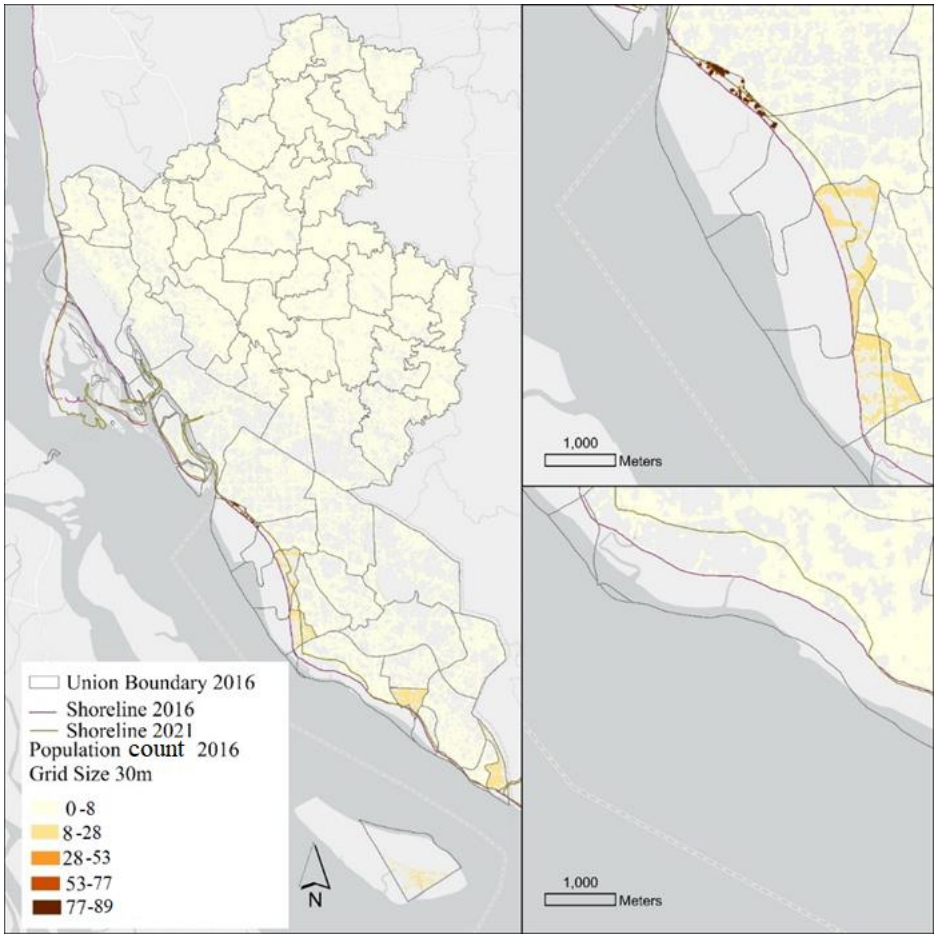


Figure B.4 Shorelines in 2016 and 2021 overlaid with 2021 30 m raster surface of population produced using vector-based approach.

Raster Based Approach

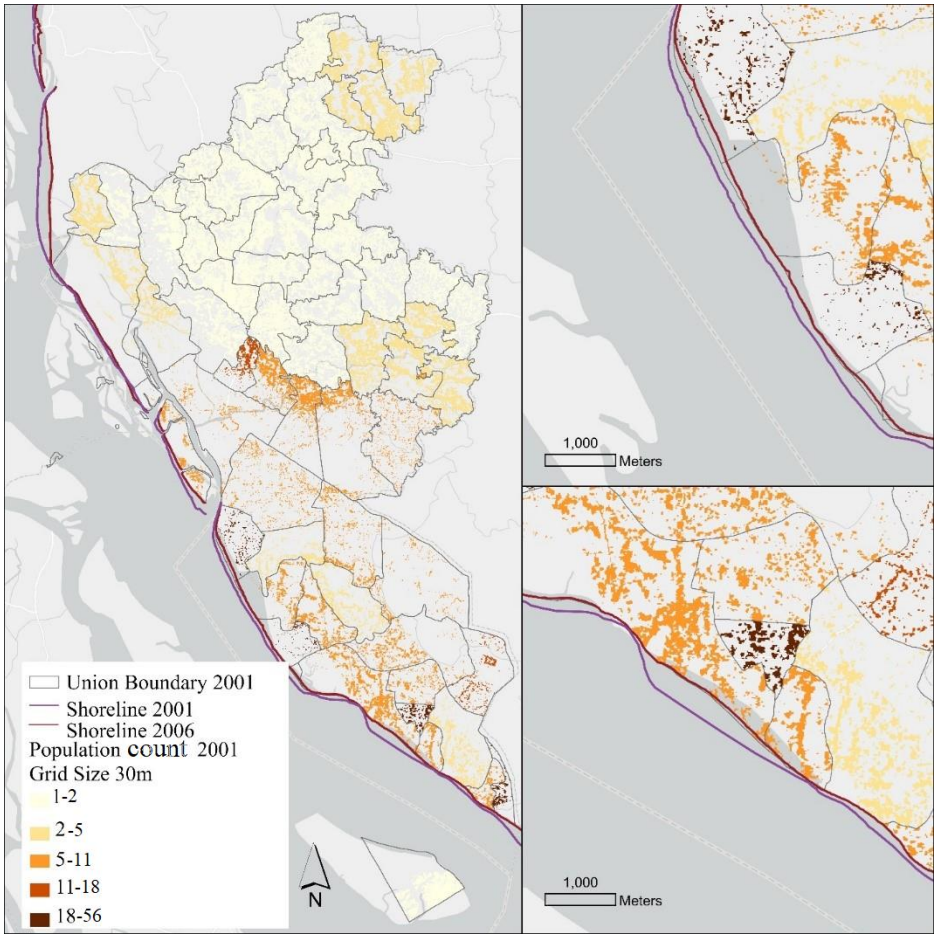


Figure B.5 Shorelines in 2001 and 2006 overlaid with the 2001 30 m raster surface of the population produced using raster raster-based approach.

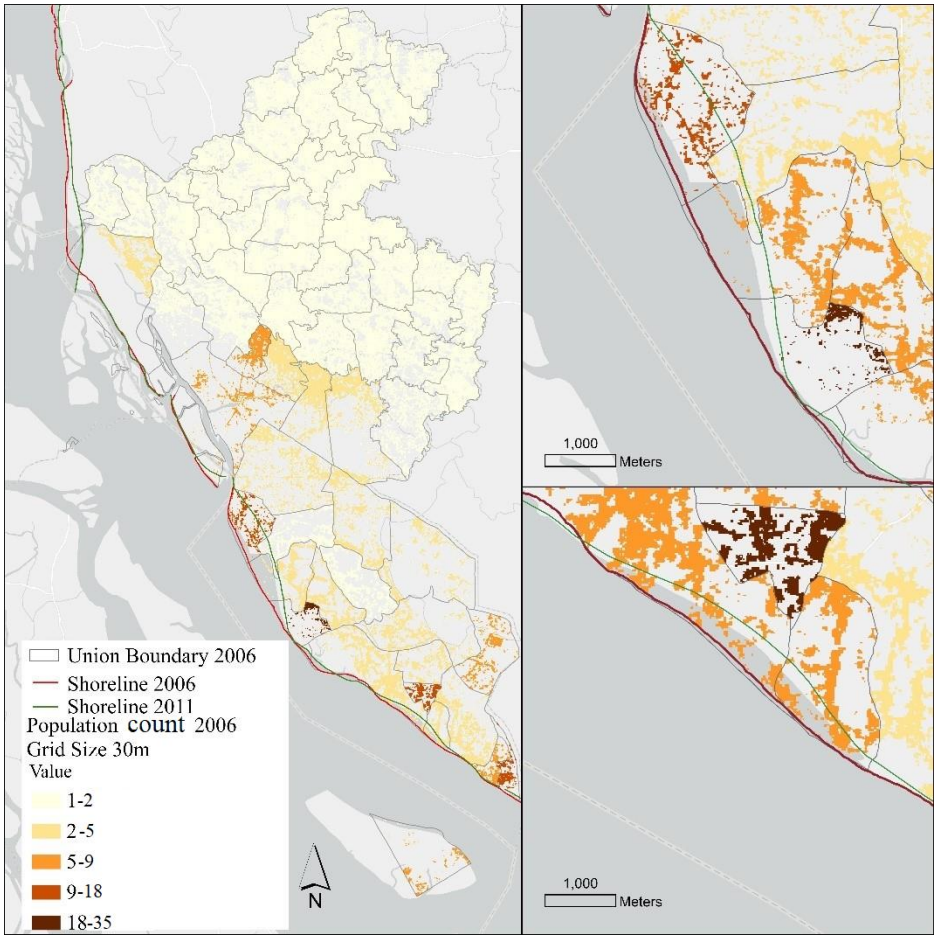


Figure B.6 Shorelines in 2006 and 2011 overlaid with the 2006 30 m raster surface of the population produced using raster raster-based approach.

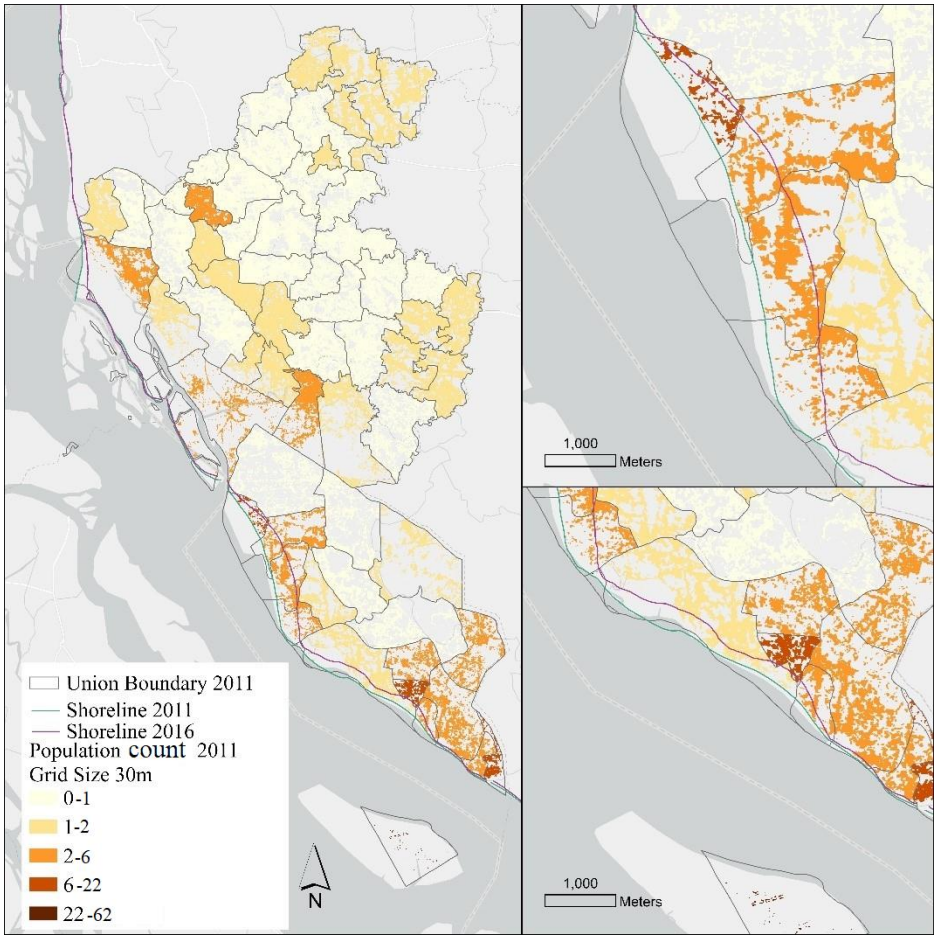


Figure B.7 Shorelines in 2011 and 2016 overlaid with 2011 30 m raster surface of population produced using raster-based approach.

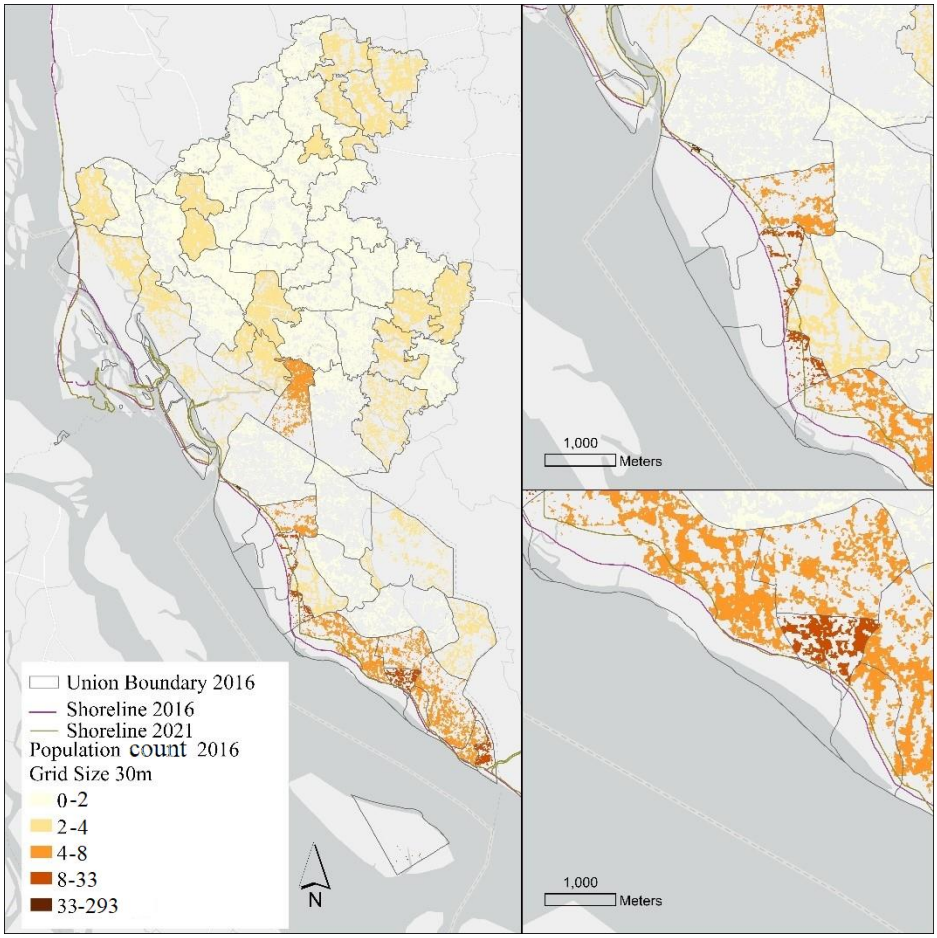


Figure B.8 Shorelines in 2016 and 2021 overlaid with the 2016 30 m raster surface of the population produced using a raster-based approach.

Figures Overlaying Landcover Maps with Shorelines.

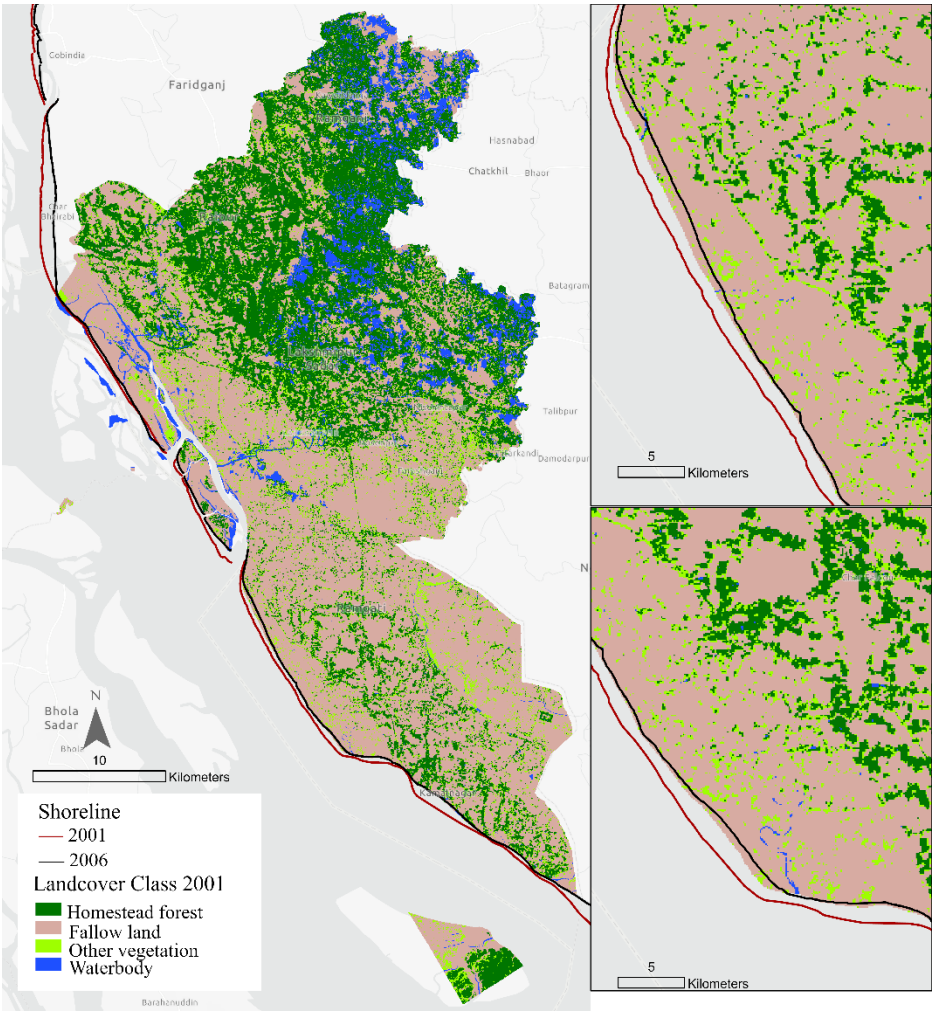


Figure B.9 Shorelines in 2001 and 2006 overlaid with 2001 30 m Landcover data.

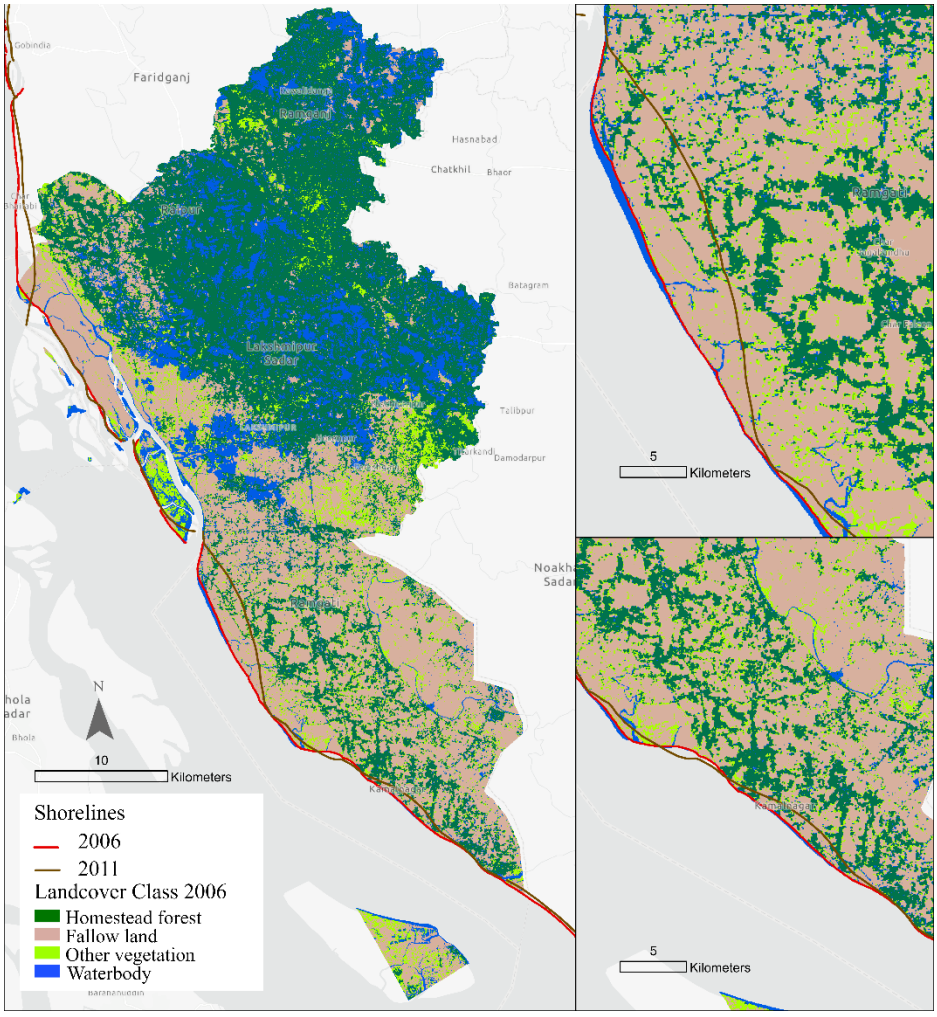


Figure B.10 Shorelines in 2006 and 2011 overlaid with 2006 30 m Landcover data.

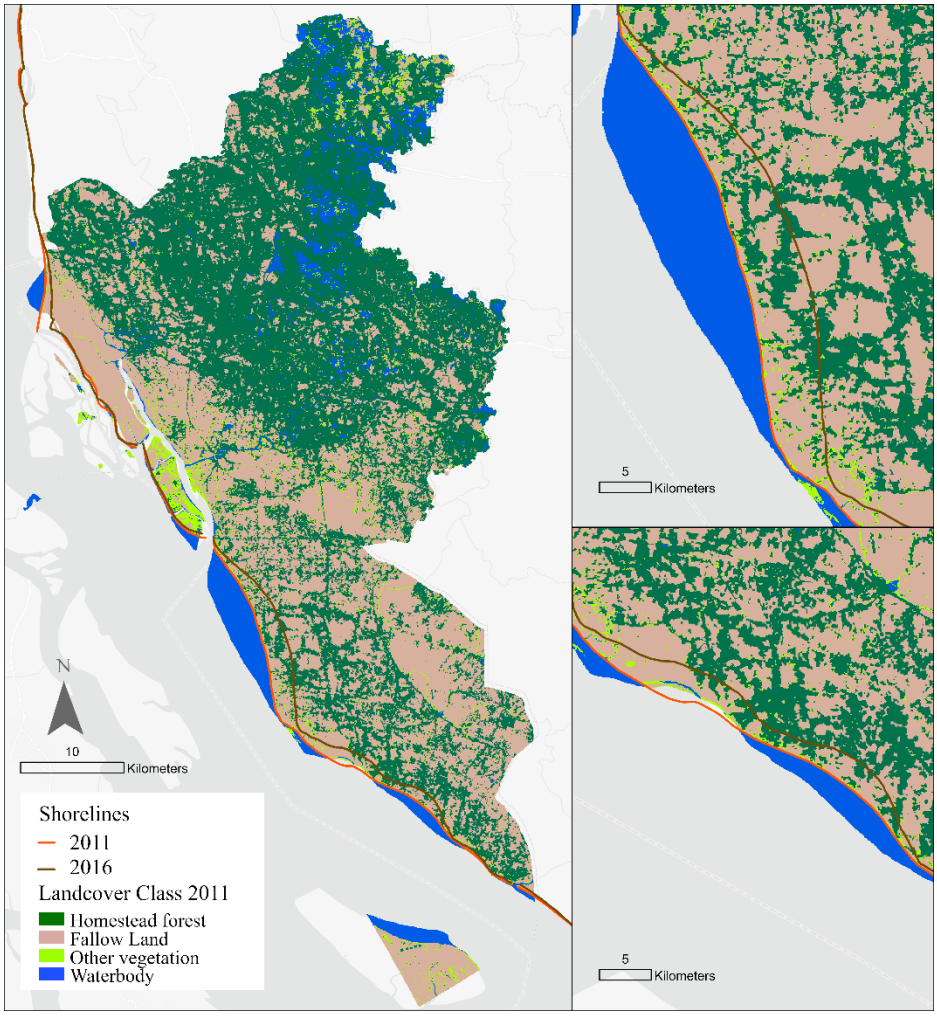


Figure B.11 Shorelines in 2011 and 2016 overlaid with 2011 30 m Landcover data.

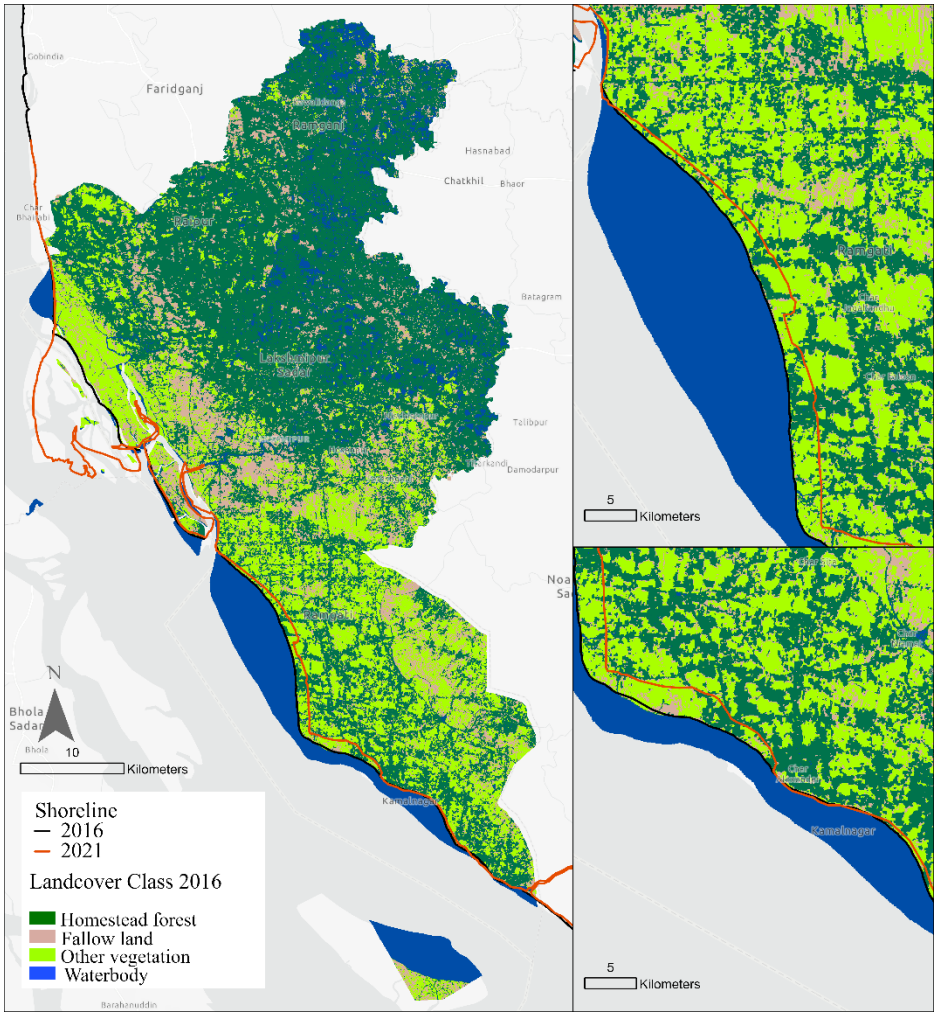
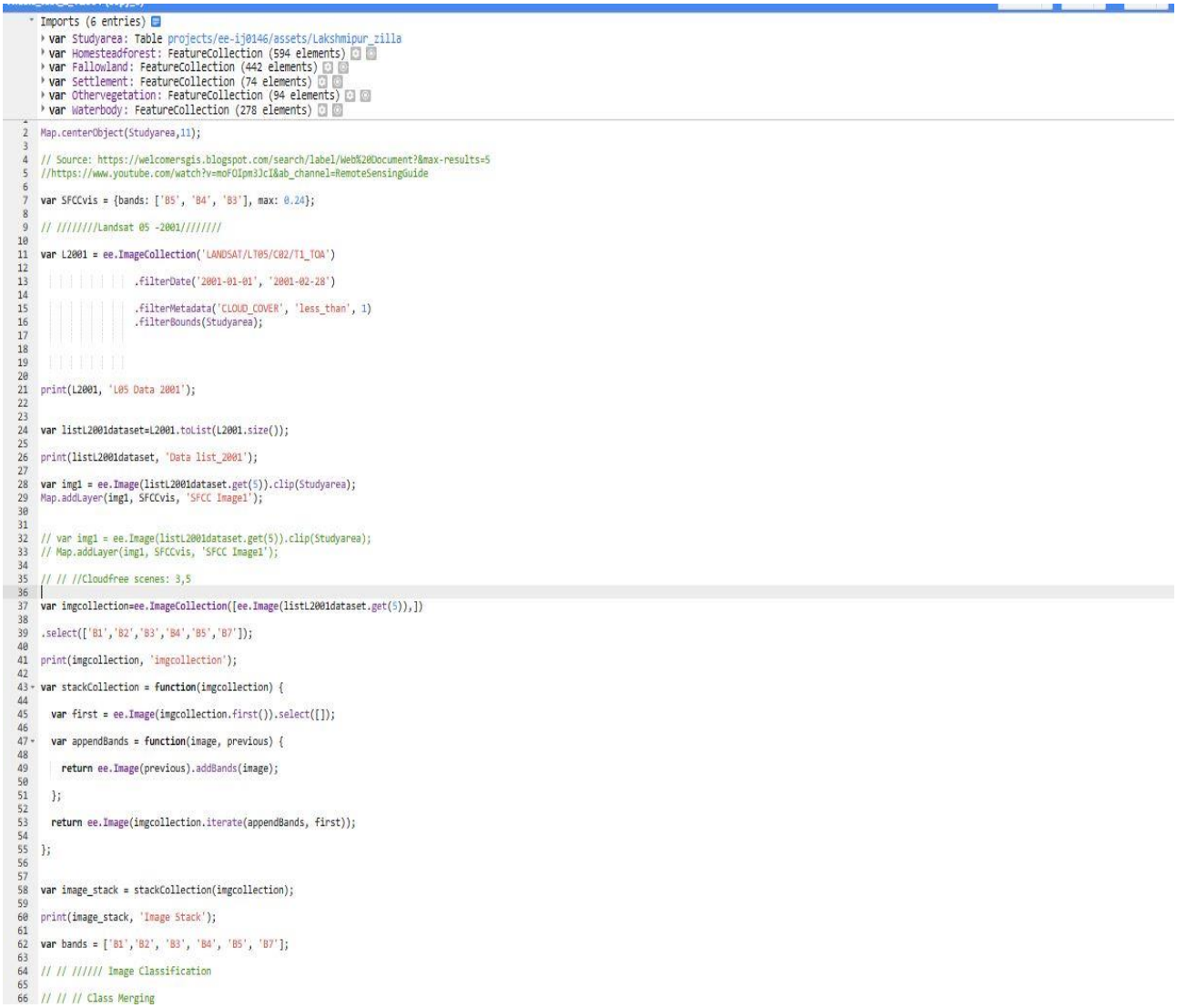


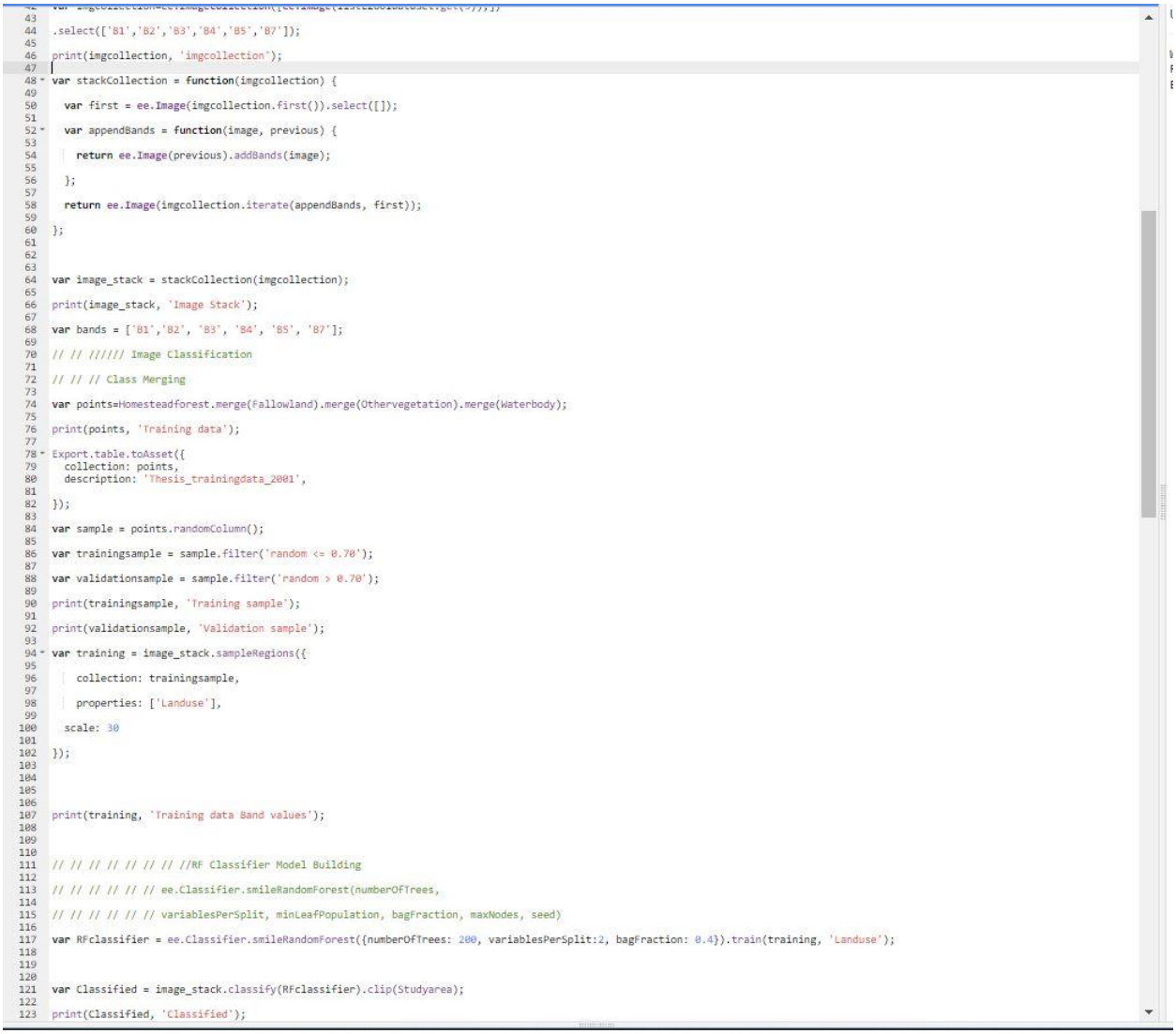
Figure B.12 Shorelines in 2016 and 2021 overlaid with 2016 30 m Landcover data.

Fig B13-B16 Sample code of Landcover classification in 2001



```
Imports (6 entries)
  var Studyarea: Table projects/ee-ij0146/assets/Lakshmipur_zilla
  var Homesteadforest: FeatureCollection (594 elements)
  var Fallowland: FeatureCollection (442 elements)
  var Settlement: FeatureCollection (74 elements)
  var Othervegetation: FeatureCollection (94 elements)
  var Waterbody: FeatureCollection (278 elements)

Map.centerObject(Studyarea,11);

// Source: https://welcomersgis.blogspot.com/search/label/Web%20Document?&max-results=5
//https://www.youtube.com/watch?v=moFOIpm3JcI&ab_channel=RemoteSensingGuide

var SFCCvis = {bands: ['B5', 'B4', 'B3'], max: 0.24};

// ////////Landsat 05 -2001///////

var L2001 = ee.ImageCollection('LANDSAT/LT05/C02/T1_TOA')

                  .filterDate('2001-01-01', '2001-02-28')

                  .filterMetadata('CLOUD_COVER', 'less_than', 1)
                  .filterBounds(Studyarea);


print(L2001, 'L05 Data 2001');


var listL2001dataset=L2001.toList(L2001.size());

print(listL2001dataset, 'Data list_2001');

var img1 = ee.Image(listL2001dataset.get(5)).clip(Studyarea);
Map.addLayer(img1, SFCCvis, 'SFCC Image1');


// var img1 = ee.Image(listL2001dataset.get(5)).clip(Studyarea);
// Map.addLayer(img1, SFCCvis, 'SFCC Image1');

// // //Cloudfree scenes: 3,5

var imgcollection=ee.ImageCollection([ee.Image(listL2001dataset.get(5)),])

.select(['B1','B2','B3','B4','B5','B7']);

print(imgcollection, 'imgcollection');

var stackCollection = function(imgcollection) {

  var first = ee.Image(imgcollection.first()).select([]);

  var appendBands = function(image, previous) {

    return ee.Image(previous).addBands(image);

  };

  return ee.Image(imgcollection.iterate(appendBands, first));

};


var image_stack = stackCollection(imgcollection);

print(image_stack, 'Image Stack');

var bands = ['B1','B2', 'B3', 'B4', 'B5', 'B7'];

// // ////// Image Classification

// // // Class Merging
```

Figure B.13 Sample code of Landcover classification in 2001



```
.select(['B1','B2','B3','B4','B5','B7']);

print(imgcollection, 'imgcollection');

var stackCollection = function(imgcollection) {

  var first = ee.Image(imgcollection.first()).select([]);

  var appendBands = function(image, previous) {

    return ee.Image(previous).addBands(image);

  };

  return ee.Image(imgcollection.iterate(appendBands, first));

};


var image_stack = stackCollection(imgcollection);

print(image_stack, 'Image Stack');

var bands = ['B1','B2', 'B3', 'B4', 'B5', 'B7'];

// // ////// Image Classification

// // // Class Merging

var points=Homesteadforest.merge(Fallowland).merge(Othervegetation).merge(Waterbody);

print(points, 'Training data');

Export.table.toAsset({
  collection: points,
  description: 'Thesis_trainingdata_2001',

});

var sample = points.randomColumn();

var trainingsample = sample.filter('random <= 0.70');

var validationsample = sample.filter('random > 0.70');

print(trainingsample, 'Training sample');

print(validationsample, 'Validation sample');

var training = image_stack.sampleRegions({

    collection: trainingsample,

    properties: ['Landuse'],

  scale: 30

});



print(training, 'Training data Band values');


// // // // // // // //RF Classifier Model Building

// // // // // // ee.Classifier.smileRandomForest(numberOfTrees,

// // // // // // variablesPerSplit, minLeafPopulation, bagFraction, maxNodes, seed)

var RFclassifier = ee.Classifier.smileRandomForest({numberOfTrees: 200, variablesPerSplit:2, bagFraction: 0.4}).train(training, 'Landuse');


var Classified = image_stack.classify(RFclassifier).clip(Studyarea);

print(Classified, 'Classified');
```

Figure B.14 Sample code of Landcover classification in 2001

```
var Classified = image_stack.classify(RFclassifier).clip(Studyarea);

print(Classified, 'Classified');


var Palette =
[  '#008000', //Homesteadforest

  '#FFFF00', //Fallowland
  '#00FF00', //Othervegetation
  '#0000FF', //Waterbody

];


Map.addLayer(Classified,{palette: Palette, min: 1, max: 5},'Classified map');

// // // //// Accuracy Assessment

// // // // Get a confusion matrix and overall accuracy for the training sample.

var trainAccuracy = RFclassifier.confusionMatrix();

print('Training error matrix', trainAccuracy);

print('Training overall accuracy', trainAccuracy.accuracy());


// // // // Get a confusion matrix and overall accuracy for the validation sample.

validation = validation.classify(RFclassifier);

var validationAccuracy = validation.errorMatrix('Landuse', 'classification');

print('Validation error matrix', validationAccuracy);

print('Validation accuracy', validationAccuracy.accuracy());

// ///// Variable Importance

var explain = RFclassifier.explain();

print(explain, 'Explain');


// // // //Variable Importance of RF Classifier

var variable_importance = ee.Feature(null, ee.Dictionary(explain).get('importance'));

// // // // // Chart of Variable Importance of RF Classifier

var chartTitle = 'Random Forest: Bands Variable Importance';

var chart =

    ui.Chart.feature.byProperty(variable_importance)

      .setChartType('BarChart')

      .setOptions({

        title: chartTitle,

        legend: {position: 'none'},

        hAxis: {title: 'Importance'},

        vAxis: {title: 'Bands'}

      });

// // // // // Chart: Location and Plot

chart.style().set({

  position: 'bottom-left',

  width: '400px',
```

Figure B.15 Sample code of Landcover classification in 2001

```
  });

// // // // // Chart: Location and Plot

chart.style().set({

  position: 'bottom-left',

  width: '400px',

  height: '400px'

});

Map.add(chart);

// // // ////// Export

Export.image.toDrive({

  image: Classified,

  description: 'Lakshmipur_2001_RF_Classified_70prcent_test3_30m',

  region: Studyarea,

  scale: 30,
  crs: 'EPSG:32645',

  fileFormat: 'GeoTIFF',

  maxPixels: 1e9,

});
```

Figure B.16 Sample code of Landcover classification in 2001

APPENDIX C

Binary Dasymetric Mapping Vector-Based Approach

Population in each target zone t, $P_t = \sum At_s / A_s * P_s$

Where At_s is the intersection area between the source and target zone.

A_s is the area of the source zone.

P_s is the population in the source zone.

1. Reclassifying landcover data
2. Converting them from raster landcover data to vector data using raster to polygon. We converted the raster into multipart polygons.
3. We masked out fallow land, other vegetation, and water bodies and kept only homestead forest class. By doing so we obtained only the homestead forest layer polygon.
4. Then intersected the union population layer with homesteadforest polygon layer, and all the attributes were stored in the homesteadforest layer. Besides each homestead forest polygons contained population union populations. Homestead forests in each union act as multipart polygons which are connected and have total shape area in the attribute table. We add a new field in the attribute table called source area and use field calculator to copy value from field shape area. We used Sourcearea=!Shapearea! command in the field calculator. The reason behind this is in the next step we are going to have a single shape area for every homestead forest polygon where we will assign population. To reduce confusion, we created a new field called source area for the total area of homestead forest polygon in a union.

5. We converted the multipart homestead forest polygons into single part polygons and now each polygon is disconnected and has ‘shape area’s of their own. Besides every disconnected homestead forest polygon has its associated union population assigned to them. To calculate how many people live in each single polygon we will do areal interpolation. Which means we will calculate the areal proportion of each individual polygon in a union relative to total area of homestead forest polygons in each union. Then we will multiply them by the population count of their respective unions.
6. Areal interpolation

   We used the following equation to assign population in each individual homestead forest polygon.

   AIT = (!Shapearea!/!Sourcearea!)*Total number of populations in a union
7. We then again create a new field called Source Area 2. Next we calculate the field using field calculator and with expression Sourcearea2= !Shapearea! as the shape area of each polygons is going to work as source area in our next step.
8. Next steps involve transferring the population to the 30 m grid. To do so we created 30 m fishnets. We clipped the fishnets with the union boundary.
9. We intersected the 30 m grids with the homesteadforest layer where we previously disaggregated the union population data on every homestead forest polygon. After the intersection we now have the attributes of homestead forest layer. We now have shape area of each 30 m grids, and we have sourcearea2 from previous step and population in the homestead forest polygons. Now we again do areal interpolation using the following equation.

   AIT2 = (!Shapearea!/!Sourcearea2!)*AIT

10. We then convert our 30 m polygons to 30 m raster.

APPENDIX D

Binary Dasymetric Mapping Raster Based Approach

Equation Population Per Pixel = ((R_H) * N*E)/(A_H/30*30)

Or , = (R_H) * N*30*30*E/A_H

R_H = the relative density of a cell

(A_H/30*30) = Number of homestead forest cells in a union

N = Actual population in a Union

E = Areal weight

$$E = P_H*100+P_F*0+P_o*0+P_w*0$$

P_H is the area proportion of homestead forest in a union calculated using the following equation,

$$P_H = A_H/ A_H+A_F+A_O+A_W$$

A_H = Area of homestead forest, A_F= Area of Fallowland, A_O = Area of Water body A_W = Area of Waterbody

A = Total Area in union, derived from

$$A= A_H+A_F+A_O+A_W$$

1. **Create Union Population Raster**

As the entire method is raster based and we assign the population into 30 m homestead forest pixels our only vector based data needs to be converted to raster to perform raster based analysis.

The first step involves converting the vector based union population data to 30 m raster. We used polygon to raster tool, choose input feature class as **Union_2001,** value field as population 2001.We determined 30 m cell size. For the year 2001 we name the output data **Pop_2001.**

2. **Reclassifying Landcover data**

As we assigned all our population to the homestead forest class, we provided 100 of relative density weight to homestead forest using reclassifying tool. The rest of the classes received 0 weight. We name this output data set as **relative_density** raster.

3. **Calculating Expected population in each Union**

To calculate the areal weight in each union we will follow the following equations.

Areal weight, $E = P_H * 100$

Proportion of Homestead forests in a Union, $P_H = A_H / A$

Total area in a Union, $A = A_H + A_F + A_O + A_W$

Step 1: Total Area Calculation

To calculate total Area in each union we need to know the total areas of each landcover classes in each union.

To do that we first rasterize the Union boundary data based on their union code. Then use tabulate area tool to calculate total area of each landcover class within each union.

Rasterize union boundary.

We used polygon to raster tool.

For value field we set the ADMN_4 field as value which contains unique union codes of each union. We named this dataset as **ADMN4_2001**

After rasterizing the union boundary, we used Tabulate area tool to obtain areas of each landcover class.

Inside Tabulate Area tool, we provided following parameters Input zone data: ADMN4_2001, Zone field: ADMN_4, which is our unique union code, Input raster: relative_density, Output table **Lctab_2001.**

Lctab_2001 table contains areas of homesteadforest, fallowland, othervegetation, Waterbody. We created a new field called **Total** to calculate total area of those classes in each union boundary. We use the following equation.

Total area in a Union, $A = A_H + A_F + A_O + A_W$

Step 2: Proportion of Homestead Forest calculation

To calculate the proportion of the homestead forest in each union we created a new field called P_H. We used $P_H = A_H / A$ to calculate the total areas of homestead forests in each union.

Step 3: Areal weight E

We created a new field called **E** and multiplied the relative density weight of homestead forest 100 with the proportion of homestead forest in each union using the following equation**.**

$$E = P_H * 100$$

4 **Creating Homestead Forest Area and Expected Population Raster**

Next to obtain the E and A_H in the equation 3.4 we created raster data from field E and Total in table Table **Lctab_2001**.

Step 1: Joining **Lctab_2001** with **Union_2001**

We used **ADMN_4** field as primary key to join **LCtab_2001** with **Union_2001** feature class.

Step 2: Rasterizing $\mathbf{A_H}$ and **E** field

As now the A_H and E field joined with our Union boundary feature class, we used polygon to raster tool and created $\mathbf{A_H}$ and **E** rasters using Total and E fields as value fields respectively. For the parameter cell size, we chose 30 m.

5 Masking out only Homestead Forest class from relative_density raster.

As our final step involves assigning population to only homestead forest class, we masked out only homestead forest from **relative_density** raster to decrease processing time. We used reclassified tools and for homestedforest class we provided 100 in the value fields and 'NODATA' for all other classes. We renamed the new relative density raster as **relative_density_2**. The new raster contains only pixels of homestead forest.

6. Calculating Population Per Cell

To calculate population in each homestead forest cell we used a rater calculator and used the following expression. The expression is based on equation 3.4

"relative_density_2" * "POP_2001_Raster" *30 *30 * "E" / ("A_H")

The output raster was named as Dasy_pop_2001