Exploring the Potential of Remote Sensing integrated with Census Data in Electoral Redistricting through BUA extraction: A Hybrid Approach

Abstract

Periodic electoral redistricting is vital in a representative system of government. It uses updated census data to ensure fair voting power, eliminating population biases among districts. Consistent census frequency is a costly endeavor, often unattainable for developing countries. Consequently, it opens the door to gerrymandering, where electoral districts are strategically arranged to benefit specific candidates. This study proposes a cost-effective alternative for estimating updated census data by leveraging previously available census data and establishing its correlation with built-up areas through medium resolution remote sensing datasets. Firstly, an integrated approach of object-based and pixel-based image analysis, using Sentinel data, was employed to extract Built-Up Areas (BUA). This BUA extraction served as the foundational footprint for extracting BUA in the census year using Landsat data. The proportional relationship of BUAs for respective years was integrated with census data to estimate the population of the year under consideration. The findings demonstrated the effectiveness of integrating census and remote sensing data, capturing 97% of population growth and its spatial growth pattern. The developed relationship was utilized to perform a pixel-level disaggregation of the population which proved instrumental in electoral redistricting up to an accuracy of -1.7 to 4.1%. Additionally, this paper examined the impact of different hierarchical levels of aggregated census data. The analysis revealed a variation of 3.06% at the district level, ranging from -0.2.21% to 13.64% at the sub-district level (Tehsils), and a wider range of -21% to 41% at the local district level (Union Councils). By utilizing pixel-level population data, multiple parcels can be generated in accordance with the current redistricting guidelines. This approach also enables efficient and quick review of existing electoral boundaries in a GIS environment.

Keywords: Built-up area extraction, Gerry Gerrymandering, Electoral Redistricting

1. Introduction

In representative democracies, electoral redistricting ensures equal voting strength among individuals. It uses updated census data to ensure fair voting power. Frequent and consistent census-taking poses a significant financial burden in terms of cost, time and human effort, particularly for developing nations. Failure to incorporate updated census in electoral redistricting can result in population biases across districts, leading to different election outcomes despite identical voting patterns. As a result, gerrymandering becomes a potential concern, as electoral districts can be deliberately manipulated to favor particular candidates. Contrary to UN Guidelines [1] i.e. 10 years, time span between last censuses in Pakistan is 19 years (Figure 1) whereas a more consistent census frequency can be observed in India & USA. Consequently, three

(3) Elections (2002, 2008 & 2013) were conducted in Pakistan, based on 1998 census [2-4] which could have potentially created disparity amongst districts. A need thus arises to explore readily available alternatives for creating updated census data with an acceptable degree of accuracy for the purpose of redistricting. Geospatial and Remote Sensing (RS) technologies offer such an alternative by employing

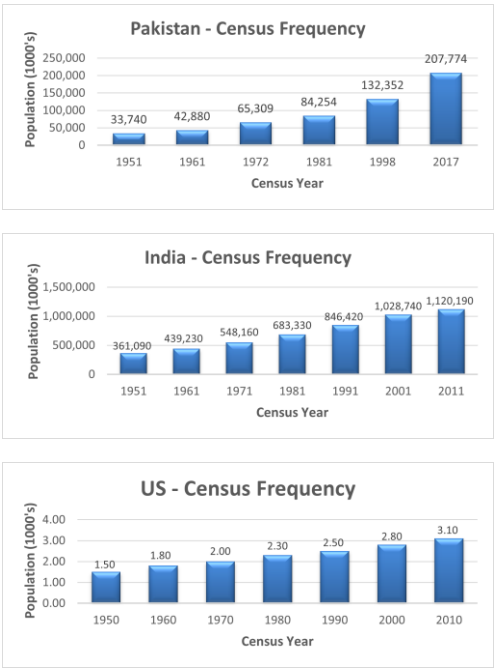


Figure 1: Census timeline comparison among Pakistan, India & USA representing inconsistent census frequency in Pakistan (source: official portals/archives of respective countries)

the allometric growth model to estimate population increase through its relationship with built-up areas (BUA) and the available census data. The allometric growth concept relies on scaling relations and has been

extensively applied in urban and regional analysis. Typically, allometric analyses focus on the proportional relationship between two elements within a geographical system [5]. Having estimated the population increase using the developed relationship, disaggregation of the estimated population can be performed as per the Minimum Mapping Aerial Unit of the available census data. Geographic Information System (GIS) can be utilized as a tool for not only mapping the housing /population growth but also assessing its spatial distribution. In line with the prevailing districting guidelines of the country (parity of population being the basic objective), electoral redistricting can be performed using the BUA pixel level disaggregation of the estimated population. This approach also provides a method for quick assessment and remedy, if felt appropriate, of the objections raised by political stakeholders.

This study focusses on three prongs towards achieving the end goal of electoral redistricting. First, BUA extraction using RS dataset. Second, disaggregation of census data down to individual pixels of BUA. Third, electoral redistricting in a GIS environment.

The building rooftop area is an essential indicator of human activity, and urban planning [6, 7]. In this context, satellite remote sensing has been the prominent measure for urban mapping of our earth [8] Compared to the traditional survey-based methods [9, 10] remote sensing could observe large areas at a potentially low cost, thus allowing tracking of the building dynamic of developing regions [11]. Advent of increased spatial resolution datasets opened new avenues of approaches for extracting BUA in an object-based image analysis paradigm. Further, multiple efforts have been made by researchers to disaggregate the census data down to individual buildings using the latest census data whereas this study mainly focusses on using the previous census data (in case of non-availability of updated census). Another study utilized remote sensing night light data and Points of Interest (POIs) data for the purpose of population mapping. The primary objective was to disaggregate existing census data in order to generate a population map for China, achieving a spatial resolution of 100 meters [12].

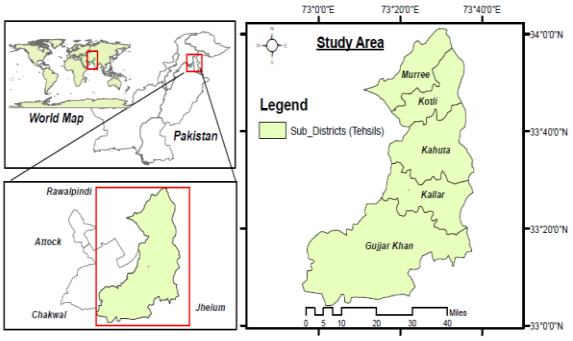


Figure-1-2. Map Highlighting the study area

Geospatial technology was first developed almost 60 years ago, and entered academic classrooms a couple of decades later, just as politicians were beginning to take advantage of computer programming to create electoral districts[13]. New technologies have created a powerful paradigm shift in political campaigning [14] and the same is true for election management. The advent of geospatial technology has enabled cartographers to create finely drawn electoral maps for particular purposes; whoever has the information about what data the maps contain and how to draw the maps has the power (Snoot, 2019) [15]. Based on limited exposure in educational institutions, we know that the majority of people in the United States are not familiar with geospatial technology, which would be the primary method of drawing electoral districts in [16]. [17] used GIS approach for Delimitation of Local Government Electoral Areas, Scotland using a semi-automated approach. Location-allocation technique was used for partitioning of regions to produce maximum compactness. [18] redefined Greek electoral districts through the application of a region building algorithm displayed in GIS environment. [19] analyzed the basic criterion that countries of the world take into account when dividing electoral districts, the most important of which are: equality of population in different districts (52.8% of the countries in the world).

2. Materials and Methods

Study Area: The focused study area is Eastern half of Rawalpindi District of the Punjab Province in Pakistan. Geographically it is located in Potohar Plateau; 33°43'N 73°04'E in the north of the country. Its total area is 3291 Sq. km which is divided into five (5) sub-districts (Tehsils): Murree, Kotli Sattian, Kahuta, Kallar and Gujjar Khan (Figure 2). Topography varies from hilly to plain areas North to South.

Satellite Data: Landsat TM images of April 14, 2018 have been used for BUA extraction 1998 (census year). Near infrared (NIR) and short-wave infrared (SWIR) bands (originally 30 m) have been resampled to 10 m spatial resolution for the purpose of developing the relationship with Sentinel data. All the images used in this study were cloud free barring one image of Landsat data which had only 1 % cloud cover. Sentinel 2 images of April 16, 2017 have been used for BUA extraction 2017 (year under consideration).

Sentinel 2 provides four 10 m visible bands i.e., RGB and the NIR bands, six 20 m SWIR and red-edge bands, and three 60 m bands [20]. In this study, 10 m bands (i.e., RGB and NIR) have been utilized, since introducing bands with coarser resolution potentially brings degradation in the performance of BUA extraction. Selected product was Level-1C Top of the Atmosphere (TOA) reflectance product, which has been conducted with systematic radiometric calibration, geometric and terrain correction by ESA [21]

Ancillary data included Provincial Census Report, (PCR) 1998 for development of relationship with BUA 1998, PCR 2017 for validation of estimated population, Election Act 2017 Guidelines for electoral redistricting and Electoral District profiles 2017 for validation / analysis of redistricting carried out in this study. All the ancillary data incorporated in this study has been obtained from official gazette notifications of federal and provincial governments.

Methodology

Cardinals: Methodology followed for this study involved four (4) cardinal stages. (Figure 2a)

- First, BUA extraction 2017 using integrated approach of object-based image analysis (OBIA) and pixel-based image analysis.
- Second, BUA extraction 1998 using Normalized Difference Built-up Index (NDBI), applied over 2017 BUA as base BUA footprint.
- Third, development of BUA relationship with 1998 census data, estimation of population 2017 using the developed relationship at various aggregate levels of census data & pixel level disintegration of population.
- Fourth, redistricting as per the electoral redistricting guidelines.

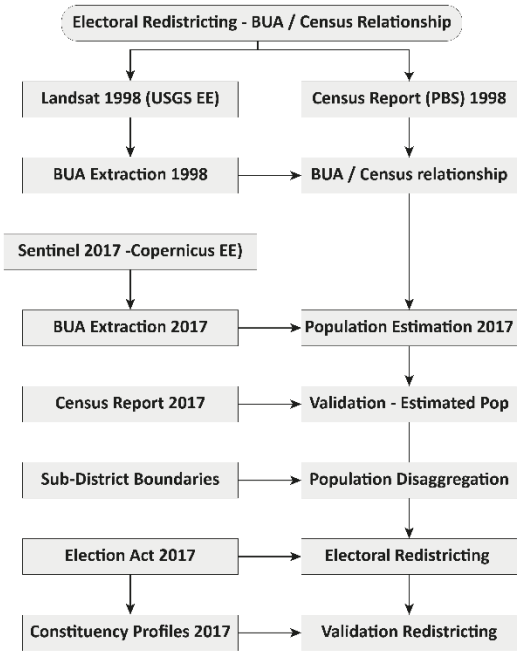


Figure 2-1. Methodological Framework - Redistricting

BUA Extraction 2017

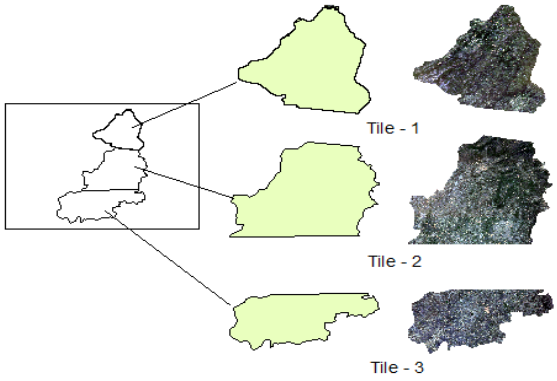


Figure 3: Downscaling of the study area into three (3) tiles

Tiling of Study Area: Keeping in view the heterogeneity in the land cover (hilly, semi hilly, plain), Digital number (DN) values of the data and processing cost (3291 sq km), the study area was divided into three (3) tiles (Figure 3). Each tile was considered as independent entity and results were combined. In this paper, BUA extraction for only Tile 1 will be covered.

Pixel Based: For pixel-based analysis two (2) classes i.e., BUA and Non-BUA were considered. Normalized difference vegetation index (NDVI) was applied and it

was observed that most of buildings were correctly identified with suitable NDVI threshold, however, some of the non-building areas, such as roads, barren areas, water and even some sparse vegetation, were also being identified as BUA. (Figure 4).

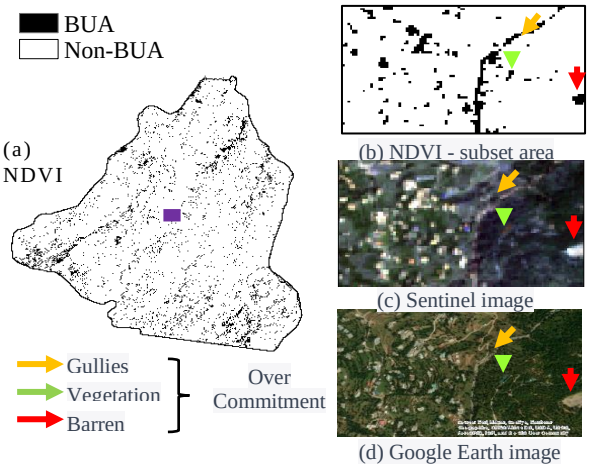


Figure 4: Pixel-based analysis of identified classes

Consequently, OBIA was carried out to extract the objects which are at least partially encompassing BUAs.

OBIA: For segmentation of OBIA, number of iterations were performed and selected parameters were: Scale - 80, Algorithm - Multiresolution, Shape and Compactness 0.5 each (Figure 5a).

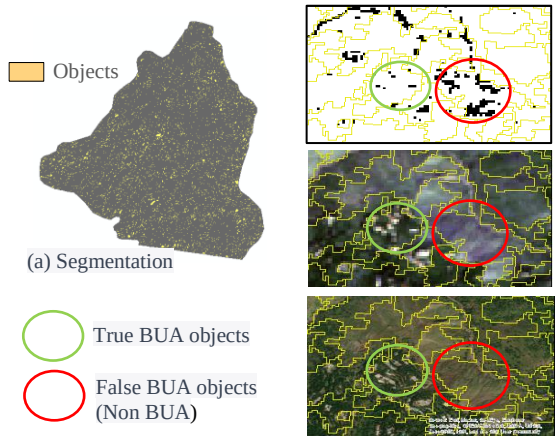


Figure 5: Segmentation analysis for BUA objects

Ruleset Preparation: True BUA objects (Figure 5) were selected using sampling technique wherein more than 600 samples were taken with selected features: NDVI, Brightness, RGB & NIR values, compactness & GLCM Homogeneity (Gray-Level Co-Occurrence Matrix) which is a feature used to quantify texture that assesses the proximity of element distribution within individual objects. Minimum and maximum (range) values of sampled object features were worked and ruleset was prepared which was applied over entire

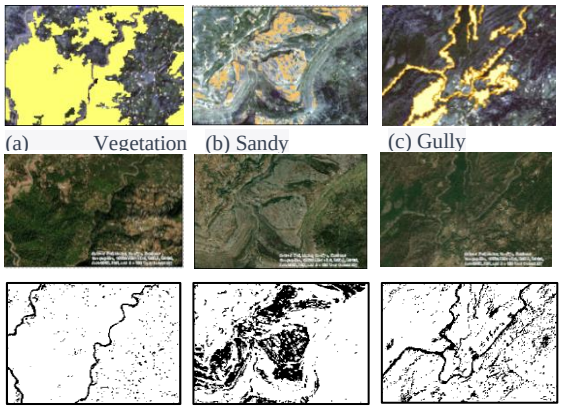


Figure 6: Refinement of Initial

population using Multiple "AND" query operators.

Refinement was performed for general brightness filter. Further refinement of rule set was performed with separate sampling of vegetation, sandy and gully areas (Figure 6) and initial rules set was symmetrically differentiated thereby only refining ruleset & not changing it. Visual representation of the BUA extraction can be visualized in Figure 6. Lastly, the results were masked with road network [22] to exclude roads, tracks, and pathways etc. which create striking spectral confusion with BUA. And also to cater for linear urban sprawl due to which BUA objects astride roads include a significant portion of roads as part of BUA objects during segmentation process. Finally,

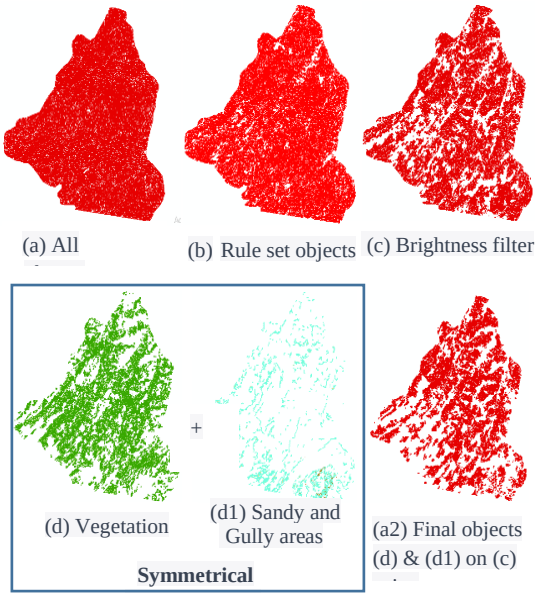


Figure 6: Visual representation of extracted BUA

refined extracted BUA objects were overlain with pixel-based results. (Figure 7)

Accuracy Assessment: The choice of accuracy assessment method has a marked impact on the way how results are evaluated. Most common way to express the accuracy of classified images/maps is by a statement of the percentage of the map area that has been correctly classified when compared with reference data or "ground truth." [23]. However, considering, the requirement and scope of the study, the 'total BUA' method was employed. Total number of pixels classified as BUA were accumulatively considered instead of evaluating the accuracy of individual pixels or objects. It is pertinent to mention here that individual pixels and area level object-based assessment methods would have

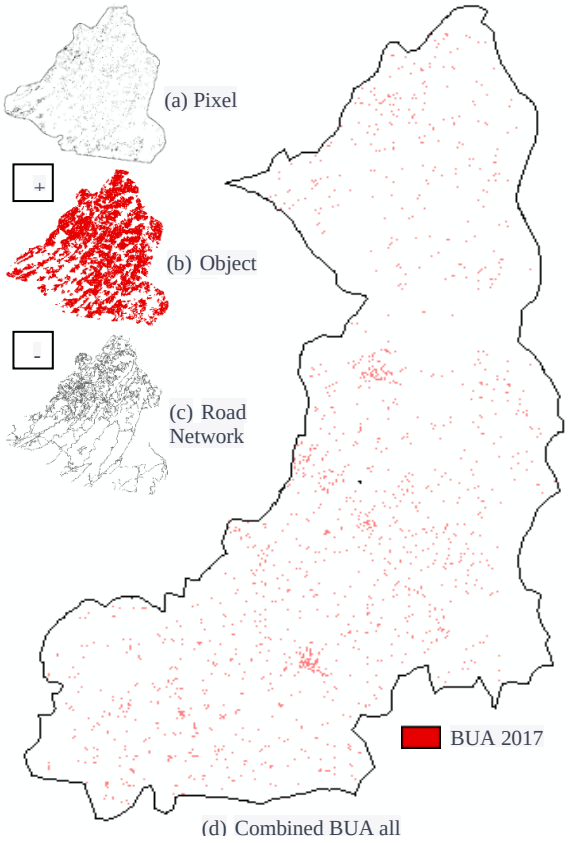
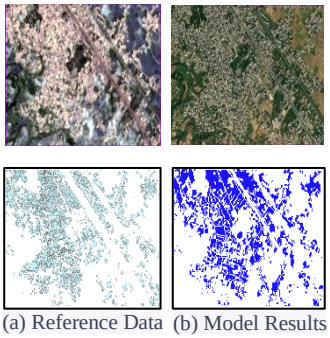


Figure 7: Combined BUA pixel results of all three tiles obtained by integrating pixel based & object-based extraction with masked road network

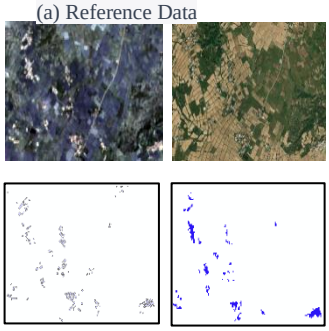
yielded lesser accuracy with a varying degree of branching or mis factor. However, in case of total covered area of BUA, error of commission would add to the total covered area. Two (2) locations each for all the tiles, representing compacted and isolated settlements, were selected for automatic processing of BUA extraction using a commercial software (Figure 8a). Upon visual inspection it was analyzed that returned results needed further refinement since the accuracy assessment can only be as accurate as the quality of reference data itself. As can be seen in Figure 8b that model results largely conform to the reference data qualitatively and shows an over commitment of around 6 % quantitatively. (Figure 8c1 &c2). Although the samples taken for accuracy assessment represent only a fraction of the entire image, however, it gives a fair idea of the model performance since the individual samples have been taken for all three (3) tiles separately. Segregation of samples between compacted and isolated settlements also provide an insight for the variation of detection for the settlement patterns across the study area.



(a) Reference Data (b) Model Results

Description	Reference Data	Model Results
Total Area (metres square)	2213267	2213267
Total Covered BUA (metres square)	547115	599600
Accuracy (%)	100	94.58
Remarks	Reference Data	5.41 % over commitment

(c1) Quantitative Assessment Compacted Settlements



(a) Reference Data

Description	Reference Data	Model Results
Total Area (metres square)	2172782	2172782
Total Covered BUA (metres square)	34314	36700
Accuracy (%)	100	93.49
Remarks	-	6.51 % over commitment

(c1) Quantitative Assessment Isolated Settlements

Figure 8: Methodology, qualitative and quantitative accuracy assessment of compacted & isolated settlements

Graphical representation of the methodology adopted for BUA extraction through integrated approach is summarized in Figure 9.

BUA Extraction 1998

The methodology utilized for BUA 1998 builds upon the results of BUA 2017, where an integrated approach with 10 m resolution data was employed using extensive processing techniques to exclude non-BUA areas. The extracted BUA results

from 2017 were transformed into polygons, serving as the fundamental BUA footprint for 1998. It is reasonable to assume, as is generally the case, and particularly in developing countries, that any new construction predominantly occurs in non-BUA areas

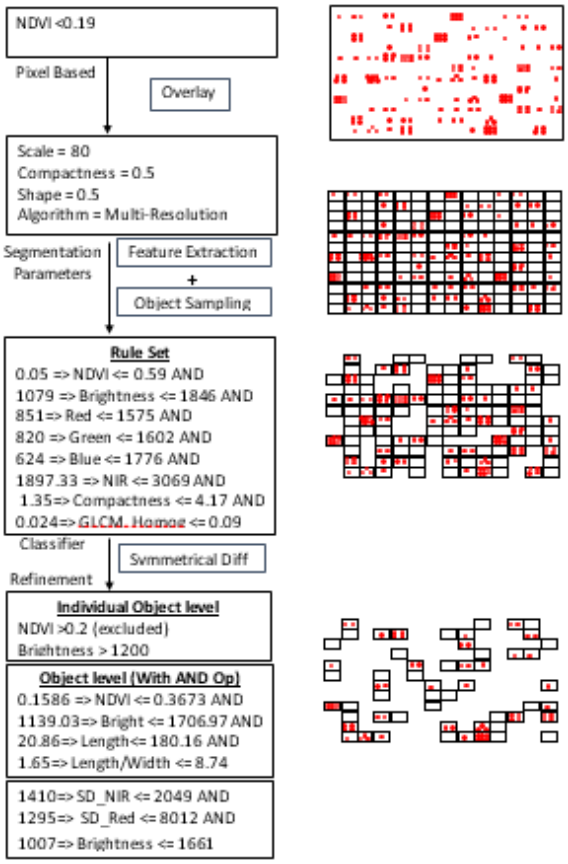


Figure 9: Methodological framework for preparation and refinement of ruleset for BUA extraction 2017

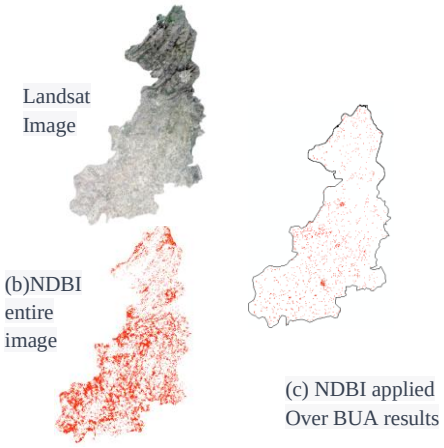
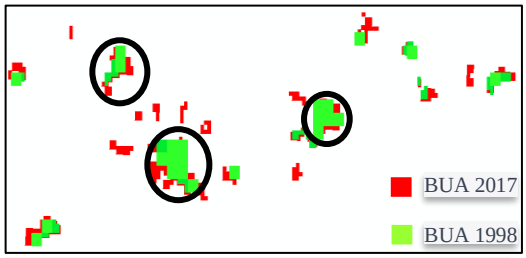


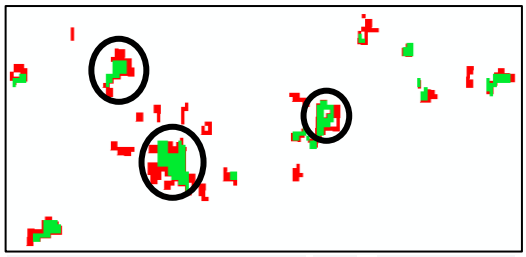
Figure 10: BUA 2018 obtained by applying Landsat NDBI over BUA 2017.

areas and rarely at locations where buildings previously existed. Conversely, this suggests that there would be minimal non-BUA areas in 2017 where actual buildings existed in 1998. Leveraging this hypothesis, extraction processing was focused solely on the BUA 2017 results. Thus, suitable threshold of NDBI, using NIR & SWIR bands, was employed only over the results of BUA 2017.

BUA extraction for 2017 was based on Sentinel (10 m) whereas same was based on Landsat (30 m) for 1998. Overlaying both datasets yielded a variable pattern results wherein some of the smaller BUAs of 2017 (based on Sentinel) having eight (8) or lesser number of pixels, were extracted as one pixel of Landsat causing a great deal of over commitment (Figure 11a). Similarly, some of the smaller BUAs of 2017 were skipped altogether during extraction processing. Thus, Landsat data was resampled to 10 m resolution (Figure 11b) so that both data sets are on a same comparable scale.



(a) BUA 1998 at 30 m extracted as per BUA 2017 on 10 m



(a) BUA results 1998 resampled to 10 m Spatial resolution

Figure 11: Comparison of variation in results with inconsistent (a) & consistent (b) resolution of data sets

A better visualization of the comparison of the spatial pattern of extracted results on a larger scale is shown in Figure 12g. As can be observed, the BUA 2017 (shown in red) is being radially increased when compared to BUA 1998 (shown in green). Total BUA 1998 & 2017 came out to be 45.27 & 64.45 respectively in totality which has been used for development of BUA / population relationship with the census data.

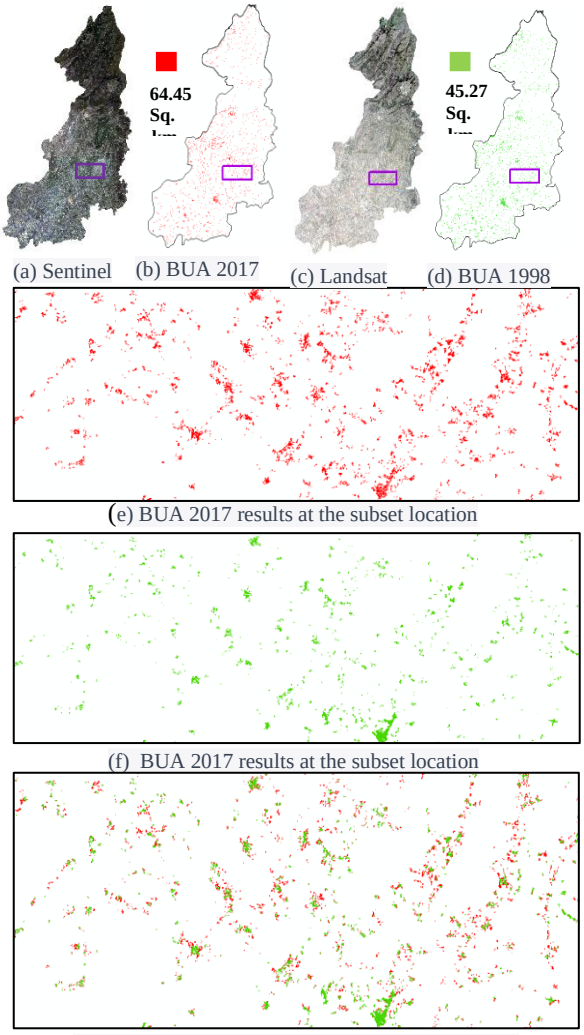


Figure 12: Comparison of BUA 1998 & 2017 results on a scale of 1:100,000

BUA / Census relationship

District Level: Based on the PCR1998, the study area had a population of 1,065,159 individuals, covering an extracted BUA of 45.27 Sq. km. By employing allometric growth analysis on the proportional BUA of 2017 (64.45 Sq. km), the estimated population for 2017 was determined to be 1,503,953 persons, while the PCR reported a population of 1,468,181 for the same year. Upon comparing the estimated and PCR populations for 2017, it becomes evident that the study overestimated the population by approximately 3% when considering the entire study area as the minimum mapping aerial unit (MMAU).

Tehsil (Sub District) Level: The PCRs provide data at both the District and Tehsil (Sub-District) level. To capture the spatial variation across the study area (see Figure 13), individual tehsils have been employed as the MMAU and compared their respective BUAs with the census data. Table 1 clearly shows that the estimated population variations among the different tehsils are more pronounced, highlighting the importance of considering this level of aggregation in our analysis.

Table 1: Tehsil level comparison of estimated pop with census 2017

Tehsil	BUA 1998 (sqkm)	Pop 1998 Census X 10,000	BUA 2017 (sqkm)	Pop Estimate 2017 X 10,000	Pop Census 2017 X 10,000	Variation (%)
Murree	4.67	17.6	6.02	22.7	23.3	-2.21
Kotli	2.57	8.2	3.51	11.3	11.9	-5.12
Kahuta	5.9	15.3	9.62	25.08	22.07	13.64
Kallar	7.46	15.8	11.56	24.5	21.7	12.88
GK	24.87	49.4	33.58	66.7	67.8	-1.63
	45.47	1,06.5	64.31	1,50.3	1,46.8	

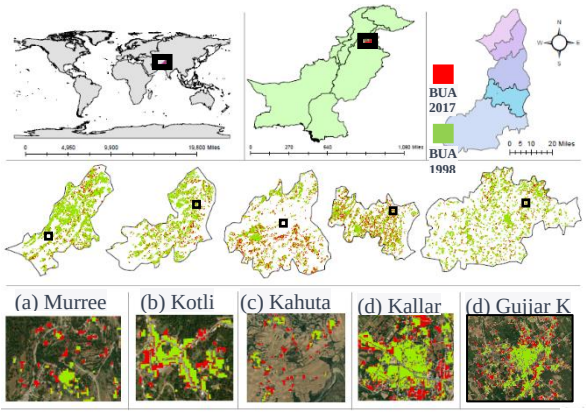


Figure 13: Tehsil level BUA 1998 / 2017 with sample results at larger scales

Pixel (10m) Disintegration of Population: To achieve a more granular understanding, aggregated population was disintegrated down to individual BUA pixels by considering the respective tehsils as MMAU. This level is the smallest at which population census data is officially published. For each tehsil, population / BUA pixel was calculated by dividing the estimated population with total number of BUA pixels covered under that specific tehsil. This process allowed to

obtain a finer resolution of population distribution within each tehsil and gain valuable insights into the spatial patterns of population density.

Table 2: Statistical results - pixel disintegration of population

Tehsil	BUA Pixels 2017	Estimated Population 2017	Persons / per pixel	Weighted as per covered BUA
Murree	60256	227,862	3.78	0.35
Kotli	35171	113,192	3.22	0.18
Kahuta	96281	250,851	2.61	0.39
Kallar	115693	245,022	2.12	0.38
G.K	335835	667,023	1.99	1.04
	643236	1,503,950		2.34

Table 2 highlights a distinctive pattern of results when distributing the population to each pixel of Built-Up Area (BUA) using census statistics for respective tehsils. The overall weighted average population per pixel (10 m) is 2.34 persons, but individual tehsils exhibit variations, ranging from 1.99 to 3.78 persons per pixel.

Murree and Kallar stand out with higher population density per pixel, possibly due to the hilly terrain that limits the construction of spacious houses. Census data also supports this observation (see Figure 14) where the average household size in Murree and Kotli Tehsil is comparatively smaller (4.93 and 5.62) compared to other tehsils such as GK (6.10), Kahuta (6.03), and Kallar (5.83).

ADMN - UNIT	AVERAGE HOUSEHOLD SIZE
GUJAR KHAN TEHSIL	6.10
RURAL	6.07
URBAN	6.28
KAHUTA TEHSIL	5.83
RURAL	5.84
URBAN	5.79
KALLAR SAYADDAN TEHSIL	6.03
RURAL	5.98
URBAN	6.17
KOTLI SATTIAN TEHSIL	4.93
RURAL	4.97
URBAN	4.75
MURREE TEHSIL	5.62
RURAL	5.63
URBAN	5.62

Figure 14: Average household size of individual tehsils as published in PCR 2017

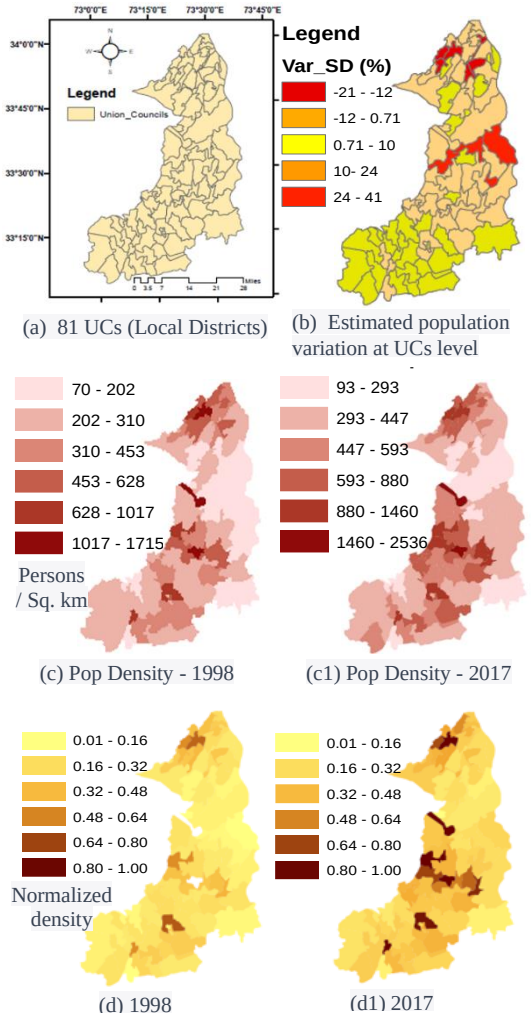
Notably, Gujjar Khan, being the largest city among all tehsils, has the largest average household size, resulting in a smaller population per pixel (1.986). This can be attributed to the presence of non-residential or commercial buildings in the area, such as warehouses, marriage halls, and other commercial activities. The flat topography of Gujjar Khan also allows for the construction of spacious houses, further contributing to this variation in population density. The influence of non-residential structures and varying topographies must be carefully considered when analyzing results obtained through this method. This study relies on the major assumption of linearity in the increase of non-residential settlements relative to residential buildings. The analysis considers proportional increases in the BUA for respective years. This assumption assumes that the growth of non-residential settlements follows a consistent linear pattern in proportion to the growth of residential buildings.

Union Councils (Local District) Level: The study area comprises 81 union councils (UCs). While census data does not provide results at this level, an alternative approach utilizing pixel-level disintegration was employed to create population density maps for each UC. The number of pixels falling within each UC was identified and then multiplied by the population per pixel of the respective tehsil covering that UC. This process enabled the derivation of population density (persons/Sq. km) for each UC by dividing the population (census 1998 & estimated 2017) with the total covered area of the respective tehsils. To facilitate better visualization of population density, both datasets were normalized, ensuring similar representation characteristics. This normalization allowed for the visualization of changes in density with consistency between the two datasets. Additionally, to evaluate the estimated population at the UC level, the standard deviation was calculated (27.2%). Variation in individual tehsils beyond 27.2% was mapped to assess deviations from the expected population change.

Figure 15 c & c1 illustrates the population density distribution for 1998, ranging from 70 to 1715 persons/Sq. km, while in 2017, it increased to 93 to 2536 persons/Sq. km across all UCs. The normalized distribution reveals a more coherent pattern of temporal population increase, particularly noticeable in urban centers (Figure 15 d & d1). As previously discussed, the overall overestimation of 3% across the entire study area varies across different regions, with

some areas showing a greater degree of variation compared to others. This phenomenon is further illustrated in Figure 4.3 (a), where the estimated population variation ranges from -21% to 41%. This variation serves as a critical indicator for qualitatively assessing the performance of the methodology employed in this study.

The variation at UC level remains within acceptable limits (-12% to +24%) across most of the area (72 UCs), with some UCs displaying an increase of +24% to 41% (12 UCs). Such larger variations could be attributed to the use of tehsil-level aggregated census data instead of UCs, which was not available.



(a) 81 UCs (Local Districts) (b) Estimated population variation at UCs level

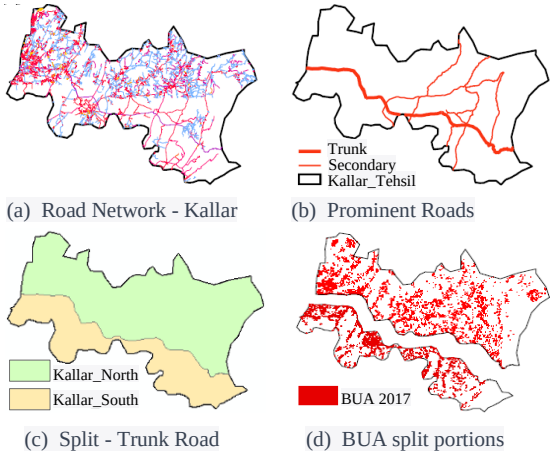
(c) Pop Density - 1998 (c1) Pop Density - 2017

(d) 1998 (d1) 2017

Figure 15: UCs population density maps for 1998 & 2017 derived from tehsil level per pixel population

Alternatively, it may indicate genuine spatial variation or potential shortcomings in the methodology itself. In any case, if the variation is attributed to the extraction BUA, these nine UCs can be analyzed and refined separately. However, this study primarily focuses on redistricting National Assembly (NA) electoral districts thus minor variations of individual UCs will automatically be smoothed out. Therefore, the spatial pattern of variation is largely dependent on the level of aggregation. A lower level of aggregation yields more refined results and enables a more accurate capture of the spatial pattern. A more suitable approach would have been to analyze the results by distinguishing between urban and rural areas. However, the inconsistent delineation of urban boundaries between 1998 and 2017 made it unfeasible to establish a relationship with separate distinctions of urban and rural areas.

Electoral Redistricting: Countries worldwide consider various criterion for electoral districts: parity of population, geographic compactness, administrative units, homogeneity etc. A tradeoff is inevitable among these factors since achieving 100% parity may not be practically feasible.



(a) Road Network - Kallar (b) Prominent Roads

(c) Split - Trunk Road (d) BUA split portions

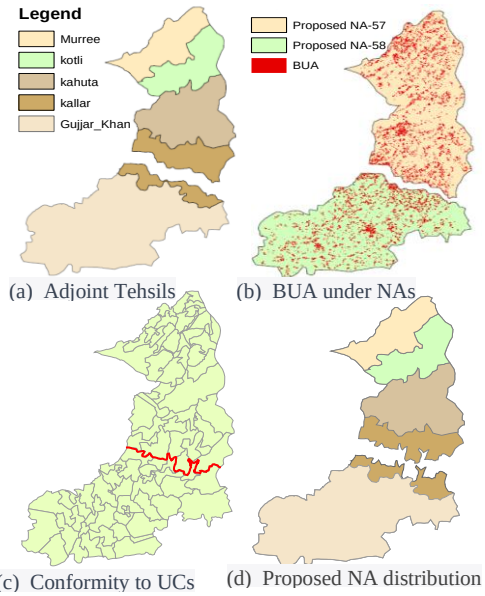
Figure 16: Illustration for splitting of Kallar Tehsil

The population of Gujjar Khan Tehsil is corresponding to the combined population of Murre, Kotli, and Kahuta tehsils (Table 2). Therefore, dividing Kallar tehsil along a prominent linear feature can effectively distribute the population of the study area into two equivalent parts. OSM road network (Figure 16a)) consists of various classes of roads. The trunk road (16b), a prominent linear feature with an East-West orientation was selected as a reference to split Kallar (16c). Resultantly, Kallar north was joined with Murree, Kotli, and Kahuta, while Kallar south with G.K (Figure 17a) in conformity with UCs boundaries (17c) forming the proposed distribution of electoral

districts (NA 57 & NA 58). Population of these electoral districts was calculated by multiplying the population / pixel of the respective tehsil with the BUA pixels falling under that tehsil. The cumulative results for both tehsils indicate only a 2.44% population disparity between the proposed electoral districts (Table 3).

Table 3: Population parity – Proposed NAs

NA	Tehsil	BUA pixels 2017	Pop/ pixel	Est pop 2017	Est Pop NA	Diff
NA - 57	Murree	60256	3.78	227,768	740,755	2.44 %
	Kotli	35171	3.21	112,899		
	ahuta	96281	2.6	250,331		
	Kallar North	70975	2.11	149,757		
		263879		740,755		
NA-58	Kallar South	44718	2.11	94,355	759,307	
	Gk	335835	1.98	664,953		



(a) Adjoint Tehsils (b) BUA under NAs

(c) Conformity to UCs (d) Proposed NA distribution

Figure 17: Illustration for proposed Electoral Districts

Validation of Proposed Redistricting:

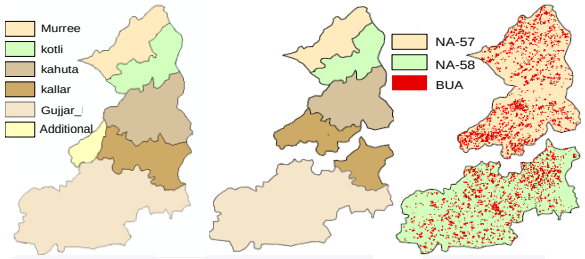
The official NA boundaries of 2017 do not align with tehsil boundaries (Figure 18b). However, for validation purposes, the additional area (18a) was included in the processing. On the other hand, for tehsil level analysis, this area was excluded as it does not cover an entire tehsil, rendering the census data for this area unusable. According to the official NA

Table 4: Population comparison – proposed and official NAs

NA	Tehsil	BUA pixels 2017	Pop/ pixel	Est pop 2017	Official Pop NA	Diff
NA - 57	Murree	60256	3.78	227,768	790,632	4.1 % (+)
	Kotli	35171	3.21	112,899		
	ahuta	96281	2.6	250,331		
	Kallar West	69089	2.11	135,228		
	Addl area	40124	2.11	78,791		
				824,588		
NA-58	Kallar East	51604	2.11	108,884	776,540	1.7% (-)
	Gk	330724	1.98	655,053		
				763,718		

boundaries, Kallar tehsil has been divided into Kallar East and West (18b), which differs from the proposed electoral districts (Kallar North and South). For validation, population/pixel weights were incorporated based on the spatial distribution of official NA boundaries, and the estimated population was then compared with the official population of the respective NAs (Table 4). The analysis indicates that the methodology followed in this study successfully delineated NA 57 with a disparity of +4.1% and NA 58 with a disparity of -1.7%.

Results and Discussion: This study utilized two different spatial resolution RS datasets, with resolutions of 10m and 30m, to extract BUA for the years 2017 and 1998. An integrated approach combining pixel and object-based methods was applied to extract BUA for the year 2017 using the higher resolution dataset. This higher-resolution BUA map served as the base for extracting BUA in 1998 using the indices method on a moderate resolution dataset. This approach helped to mitigate intra-urban heterogeneity and spectral confusion between



(a) Tehsils with additional area (b) Adjoint Tehsils (c) NAs 2017

Figure 18: Formation of official NA boundaries 2017

different landcover types on a moderate resolution dataset. Although the samples taken for accuracy assessment represent only a fraction of the entire image yet it gives an insight into the qualitative assessment of extracted BUA. Due to the extensive coverage of the study area, errors of commission and omission were expected to be present. However, cumulatively, the errors of commission would contribute to the total BUA. It is pertinent to mention here that individual pixels and area level object-based assessment methods would have yielded lesser accuracy with a high degree of branching and mis factor. However, the scope and scale of study allows the adaptation of aforesaid assessment approach.

The total BUA for 1998 and 2017 was found to be 45.27 and 64.45, respectively. These figures were then used to establish a relationship between BUA and various levels of 1998 census data, based on the respective MMAUs. At the higher aggregation level (district), there was an overestimation of 3% in the population due to data smoothing. At lower aggregation levels (Sub districts), the population estimation varied from -2.21% to 13.64% indicating a variable pattern of spatial variation. At the lowest aggregation level (UCs), a wider range of variation (-21% to 41%) was observed in some parts, which may be linked to the use of tehsil-level aggregated census data instead of UCs level. The study focused primarily on redistricting NA electoral districts, so minor variations in individual UCs would naturally be smoothed out.

Pixel disaggregation of population using the estimated population of respective tehsils indicate a variation of 3.78 to 1.99 persons / pixel (North to South). Murree and Kallar stood out with higher population density per pixel, possibly due to the hilly terrain limiting the construction of spacious houses. Census data supported this observation, as the average household size in Murree and Kotli Tehsil was comparatively smaller than in other tehsils. Gujjar Khan, being the largest city among all tehsils, had the largest average household size, resulting in a smaller population per pixel. This was attributed to the presence of non-residential or commercial buildings, which included warehouses, marriage halls, and other commercial activities. The flat topography of Gujjar Khan also allowed for the construction of spacious houses, further contributing to the variation in population density. The influence of non-residential structures and varying topographies must be considered when interpreting the results obtained through this method.

The study made a major assumption of linearity in the increase of non-residential settlements relative to residential buildings, with proportional increases in BUA for respective years. This assumption relied on the growth of non-residential settlements following a consistent linear pattern in proportion to the growth of residential buildings.

For electoral district creation, population parity was the primary criterion. The population of the proposed electoral districts showed a disparity of only 2.44%, well below the permissible limit of 10%. Validation results against official NA boundaries indicated an underestimation of 4.1% for NA-57 and an overestimation of 1.7% for NA-58 based on the pixel-level disaggregation of population. These minor variations follow a similar pattern of tehsil level variation in the estimated population which formed the basis for pixel level disaggregation of population.

Conclusions and Recommendations

This study highlights the effectiveness of an integrated approach, involving Remote Sensing and Census data to extract Built-Up Areas (BUA). This approach adeptly captured 97% of population growth and its spatial nuances, offering a cost-effective means to update census data. The developed relationship between BUAs and census data demonstrated robust accuracy, with a pixel-level disaggregation exhibiting precision ranging from -1.7% to 4.1% during electoral redistricting. Spatial analyses revealed substantial variations in population density across hierarchical levels, emphasizing the need for fine-scale considerations. The study also illuminated the impact of non-residential structures and topography on population distribution. Electoral redistricting, guided by our population estimates, adhered to the primary criterion of population parity, with minor variations falling within acceptable limits. Furthermore, the generated pixel-level population data showcased its potential for efficient GIS-based reviews of electoral boundaries. This study not only contributes valuable insights into population estimation methodologies but also lays the groundwork for future research, suggesting avenues for refining intra-urban heterogeneity handling, addressing errors, and exploring the dynamics of non-residential settlement growth.

References

1. Affairs, D.o.E.a.S. and S. Division, *Principles and Recommendations for Population and Housing Censuses*, D.o.E.a.S. Affairs and S. Division, Editors. 2017.
2. Statistics, P.B.o. *Population Census*. 2017; Available from: https://www.pbs.gov.pk/content/population-census.
3. Bureau, U.S.C. *Census Bureau Data*. 2020; Available from: https://data.census.gov/.
4. India, C.o. *Census Tables*. 2011; Available from: https://censusindia.gov.in/census.website/node/325.
5. Chen, Y., *Multi-Scaling Allometric Analysis for Urban and Regional*

Development. 2016.
6. Huang, W., H. Tang, and P. Xu, *OEC-RNN: Object-Oriented Delineation of Rooftops With Edges and Corners Using the Recurrent Neural Network From the Aerial Images.* IEEE Transactions on Geoscience and Remote Sensing, 2022. **60**: p. 1-12.
7. Nadal, A., et al., *Urban planning and agriculture. Methodology for assessing rooftop greenhouse potential of non-residential areas using airborne sensors.* Sci Total Environ, 2017. **601-602**: p. 493-507.
8. Zhu, X.X., et al., *The urban morphology on our planet - Global perspectives from space.* Remote Sens Environ, 2022. **269**: p. 112794.
9. Kuthanazhi, V., et al., *Estimating Mumbai's Rooftop PV Potential through Mobilization of IEEE Student Community.* IEEE Transactions on Geoscience and Remote Sensing, 2016.
10. Jones, L. and P. Hobbs, *The Application of Terrestrial LiDAR for Geohazard Mapping, Monitoring and Modelling in the British Geological Survey.* Remote Sensing, 2021. **13**(3).
11. Liu, Z., et al., *China Building Rooftop Area: the first multi-annual (2016–2021) and high-resolution (2.5 m) building rooftop area dataset in China derived with super-resolution segmentation from Sentinel-2 imagery.* Earth System Science Data, 2023. **15**(8): p. 3547-3572.
12. Ye, T., et al., *Improved population mapping for China using remotely sensed and points-of-interest data within a random forests*

model. Sci Total Environ, 2019. **658**: p. 936-946.
13. Brace, K.W., *A Personal Prospective on the Use of Technology in Redistricting over the past Thirty Years.* The Brookings Institute Conference on Congressional Redistricting,

April 16, 2004, Washington, D.C, 2004.
14. Spiller, L.D. and J. Bergner, *American Elections and the Competition to Govern.* American Society for Competitiveness, 2010. **8(2)**.
15. Smoot, R. *A look into the past, present and future of partisan gerrymandering in N.C.* 2019; Available from: https://www.dailytarheel.com/article/2019/01/gerrymandering-history-future-0128.
16. Leadbeater, R. *Open Redistricting and Rebuilding Trust*. 2017; Available from: https://www.govloop.com/community/blog/open-redistricting-rebuilding-trust/.
17. ROBERTSON and I.M. L., *The Delimitation of Local Government Electoral Areas in Scotland: A Semi-Automated Approach.* Journal of the Operational Research Society, 1982.
18. PHOTIS, Y.N., *REDEFINITION OF THE GREEK ELECTORAL DISTRICTS THROUGH THE*

APPLICATION OF A REGION-BUILDING ALGORITHM European Journal of Geography, 2012. **3(2)**: p. 72-83.
19. Salameh, M.T.B., *Electoral districts in Jordan*, in *Routledge Handbook on Elections in the Middle East and North Africa*. 2023. p. 54-65.
20. Huang, X., et al., *Mapping 10 m global impervious surface area (GISA-10m) using multi-source geospatial data.* Earth System Science Data, 2022. **14**(8): p. 3649-3672.
21. Tools, S.U.H.a.E., *Sentinel-2 User Handbook*. 2015: European Space Agency. 60.
22. OpenStreetMap, *Planet dump retrieved from https://planet.osm.org*. 2023, OpenStreetMap contributors.
23. Story, M., *Accuracy Assessment: A User's Perspective.* American Society for Photogrammetry and Remote Sensing, 1986.