CREATING A GRIDDED POPULATION DATA OF INDIA: A COMPARISON AND EVALUATION STUDY OF DASYMETRIC MAPPING METHODS

Deepika Mann [1], Oinam Bakim Chandra[2]

[1] Student, TERI University

[2] Assistant Professor, Shiv Nadir University

Address: 10, Institutional Area, TERI University, Vasant Kunj, Delhi-110070

Abstract:

Demographic datasets such as population are associated with analytical errors due to arbitrary nature of areal unit partitioning. These problems or errors mainly include MAUP (Modifiable Areal Unit Problem), which is the variation in results that can occur when data from one scale of areal units is aggregated into another spatial scale. Choropleth maps of population density by areal administrative units represent a homogeneously distributed population at that scale (say state-wise) even where places are uninhabited. To account for these problems dasymetric mapping is done which is a method of thematic mapping that uses areal symbols to spatially classify volumetric data. The aim of this study is to generate population distribution raster surface for India at state level by Dasymetric mapping using three different methodological approach proposed by Holloway et al. (1997), Mennis (2003) and tool developed by USGS for Dasymetric Mapping in ArcGIS software.

The resulting raster surface can be used for various applications like geographic studies, socio-economic studies, crime analysis, suitability mapping etc.

About the Author:

Ms Deepika Mann, M.Sc Geoinformatics is a student of Department of Natural Resource Management at TERI University, New-Delhi, India.

E mail ID: deepikamann25@gmail.com ,

Contact No: +91 – 8010666352

Introduction

Most Census generated datasets are aggregated to areal units such as states, districts etc. Demographic datasets such as population are associated with analytical errors due to arbitrary nature of areal unit partitioning. Perhaps the most prominent of these errors is the modifiable areal unit problem (MAUP), defined as a situation in which modifying the boundaries and/or scale of data aggregation significantly affects the results of spatial data analysis (Openshaw 1983). A related problem is the visual description of the demographic datasets. Map of population by areal administrative units represents a homogeneously distributed population even at places that are uninhabited, say rivers, snow cover etc. Dasymetric Mapping provides a potential solution to these problems by generation of a surface-based demographic data representation, in which data are modelled as a continuous field independent of partitioning into arbitrary areal units. Surface-based population representation offers certain advantages over areal unit representation. Population data can be aggregated to nearly any desired areal unit using surface-based representation and hence is not subject to the MAUP and other areal unit-derived problems (Bracken 1993). In addition, since surface representations present a cell as unit of display that is uniform in size across a region, surfaces of population may offer a more accurate cartographic representation of population distribution than do conventional choropleth maps (Langfor and Unwin 1996). Raster GIS provides a platform to develop useful surface based representations of population and other demographic datasets from aggregated census data. Dasymetric mapping is a method of thematic mapping by assessing the relationship between population distribution (from census data) and remote sensing derived land use land cover as ancillary data. Various approaches to dasymetric mapping have been used, the most accepted being the methodology by Jeremy Mennis (2003). This study aims at generating population distribution raster surface for India at state level by Dasymetric using methodology given by Holloway et al. (1997), Jeremy Mennis (2003) and by ArcGIS tool for Dasymetric Mapping. Resulting raster surfaces by these methods are then compared.

Datasets used:

(i) State-wise population data of India for the year 2001 (downloaded from Indiastats.com) is used as the demographic data which is to be distributed to a continuous raster surface. Population data and Land Use Land Cover (LULC) data should be of nearly same time period because LULC data greatly influences the population distribution. Since the most recent LULC data readily available is for 2000, hence census data for 2001 is being used.

(ii) LULC data of India (2000) downloaded from GLC2000 having a resolution of 981.56 m is used. This is recoded to four classes uninhabited (where people do not reside), non-urban (where very less people live), rural (not highly urbanized settlements), urban (includes highly urbanized settlements) being mentioned in Table 1.

Table 1
Table showing different land cover classes

Class Value	Recoded Class Name	Original LULC Classes
1	Uninhabited	Sea, Swamps, Deserts, Snow, etc
2	Non-Urban	Forest, Grasslands, Vegetation
3	Rural	Agriculture, Mud Flats
4	Urban	Settlements

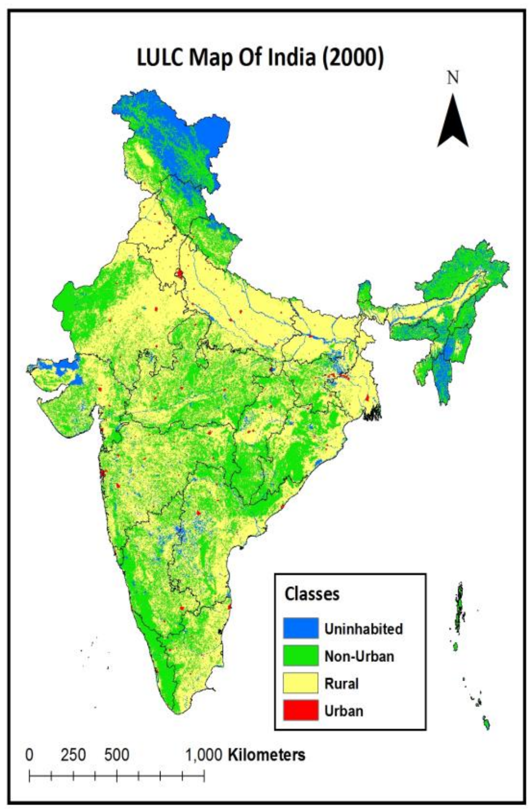


Fig: 1 - LULC map of India for the year 2000

Methodology:

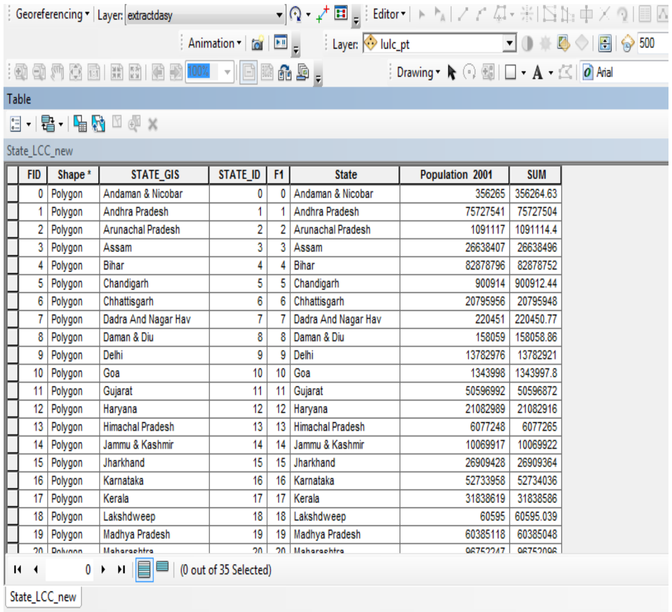
S.No.	Approach	Methodology	Reference
1.	**Mennis (2003)**	The population of each block group is distributed to each cell in the population surface based on two factors: (1) the relative difference in population densities among the three urbanization classes; and (2) the percentage of total area of each block group occupied by each of the three urbanization classes.	Mennis, J.,2003,Generating Surface Models of Population Using Dasymetric Mapping. *The Professional Geographer*, 55(1), 31-42.
2.	**Modified equation of Holloway et al (1997)**	It assigns a predetermined percentage of population to each class, which is same for all the states. Table 1 is used to assign values to different LULC classes for India. Since Urban areas are highly populated than others so this relative density population is being assigned.	Holloway, S., J. Schumacher, and R. Redmond, 1997, People & Place: Dasymetric Mapping Using Arc/Info. *Cartographic Design Using ArcView and Arc/Info*, Missoula: University of Montana, Wildlife Spatial Analysis Lab
3.	**Dasymetric Mapping tool**	The tool developed by USGS follow the original Mennis (2003) method of dasymetric mapping and it is implemented in ArcGIS platform as a tool.	http://geography.wr.usgs.gov /science/dasymetric/data.htm

Table 2
Table showing relative densities for each land cover class

LandCover Code	Descriptions	Relative Density (R_A)
4	Urban	60
3	Rural	30
2	Non-Urban	10
1	Uninhabited	0

Conclusion:

1) **Dasymetric Mapping using methodology given by Mennis (2003):** Fig: 3 shows the resulting raster surface for population distribution .This proves to be the best method since it gives the population distribution in accordance with the LULC Map. Also, it does not over or underestimates the population. Very less or zero value is assigned to hilly regions of Uttaranchal, Himachal Pradesh, Jammu and Kashmir, forested parts of Meghalaya, Arunachal Pradesh and deserted parts of Gujarat and Rajasthan which are classified as uninhabited of non-urban. Major cities like Delhi, Mumbai, Jaipur are assigned highest population. Also Uttar-Pradesh, Bihar, West Bengal, Assam are given highest population which is also true in reality. Rural and urban parts of other states are assigned moderate population. Sum of the cell values for each state in the final dasymetric map (see Fig: 2) matches well with the actual population of these states. This justifies the accuracy of the method.



FID	Shape *	STATE_GIS	STATE_ID	F1	State	Population 2001	SUM
0	Polygon	Andaman & Nicobar	0	0	Andaman & Nicobar	356265	356264.63
1	Polygon	Andhra Pradesh	1	1	Andhra Pradesh	75727541	75727504
2	Polygon	Arunachal Pradesh	2	2	Arunachal Pradesh	1091117	1091114.4
3	Polygon	Assam	3	3	Assam	26638407	26638496
4	Polygon	Bihar	4	4	Bihar	82878796	82878752
5	Polygon	Chandigarh	5	5	Chandigarh	900914	900912.44
6	Polygon	Chhattisgarh	6	6	Chhattisgarh	20795956	20795948
7	Polygon	Dadra And Nagar Hav	7	7	Dadra And Nagar Hav	220451	220450.77
8	Polygon	Daman & Diu	8	8	Daman & Diu	158059	158058.86
9	Polygon	Delhi	9	9	Delhi	13782976	13782921
10	Polygon	Goa	10	10	Goa	1343998	1343997.8
11	Polygon	Gujarat	11	11	Gujarat	50596992	50596872
12	Polygon	Haryana	12	12	Haryana	21082989	21082916
13	Polygon	Himachal Pradesh	13	13	Himachal Pradesh	6077248	6077265
14	Polygon	Jammu & Kashmir	14	14	Jammu & Kashmir	10069917	10069922
15	Polygon	Jharkhand	15	15	Jharkhand	26909428	26909364
16	Polygon	Karnataka	16	16	Karnataka	52733958	52734036
17	Polygon	Kerala	17	17	Kerala	31838619	31838586
18	Polygon	Lakshdweep	18	18	Lakshdweep	60595	60595.039
19	Polygon	Madhya Pradesh	19	19	Madhya Pradesh	60385118	60385048

(0 out of 35 Selected)

Fig: 2 – Attribute Table showing comparison of population from census data and sum of the population distributed to cells of each state by Dasymetric Mapping using Mennis (2003) method.

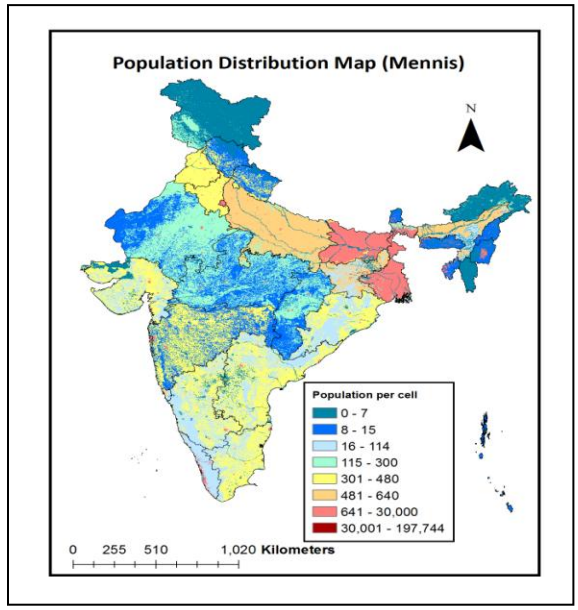
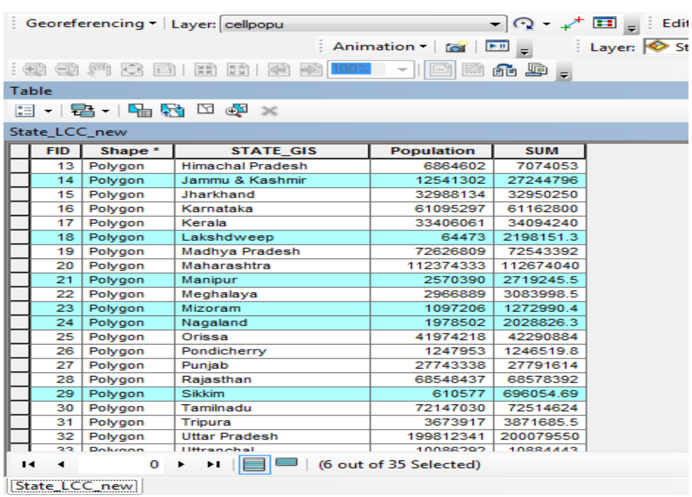


Fig: 3 – Map showing population distribution generated by using methodology given by Mennis (2003)

2) **Dasymetric Mapping using modified equation given by Holloway et al (1997):** Fig: 5 shows the resulting raster surface for population distribution .This method overestimates the population for most of the states. Jammu and Kashmir should be assigned no data values but is moderately populated in the map. Western part i.e. the deserted parts of Rajasthan is given nearly moderate population but should be amongst least populated areas. Population for Kerala is also overestimated. Population for eastern parts of India, namely, Arunachal Pradesh, Nagaland, Manipur, Mizoram, Tripura, Meghalaya and Sikkim is overestimated. These should have been assigned less population since these majorly comprises of forested areas. The population value for highly populated states and cities like Punjab, Haryana, Bihar, Uttar Pradesh, Delhi, Mumbai, etc is rightly assigned. Exact figures can be seen in the table provided below in Fig: 4.



FID	Shape *	STATE_GIS	Population	SUM
13	Polygon	Himachal Pradesh	6864602	7074053
14	Polygon	Jammu & Kashmir	12541302	27244796
15	Polygon	Jharkhand	32988134	32950250
16	Polygon	Karnataka	61095297	61162800
17	Polygon	Kerala	33406061	34094240
18	Polygon	Lakshdweep	64473	2198151.3
19	Polygon	Madhya Pradesh	72626809	72543392
20	Polygon	Maharashtra	112374333	112674040
21	Polygon	Manipur	2570390	2719245.5
22	Polygon	Meghalaya	2966889	3083998.5
23	Polygon	Mizoram	1097206	1272990.4
24	Polygon	Nagaland	1978502	2028826.3
25	Polygon	Orissa	41974218	42290884
26	Polygon	Pondicherry	1247953	1246519.8
27	Polygon	Punjab	27743338	27791614
28	Polygon	Rajasthan	68548437	68578392
29	Polygon	Sikkim	610577	696054.69
30	Polygon	Tamilnadu	72147030	72514624
31	Polygon	Tripura	3673917	3871685.5
32	Polygon	Uttar Pradesh	199812341	200079550

(6 out of 35 Selected)

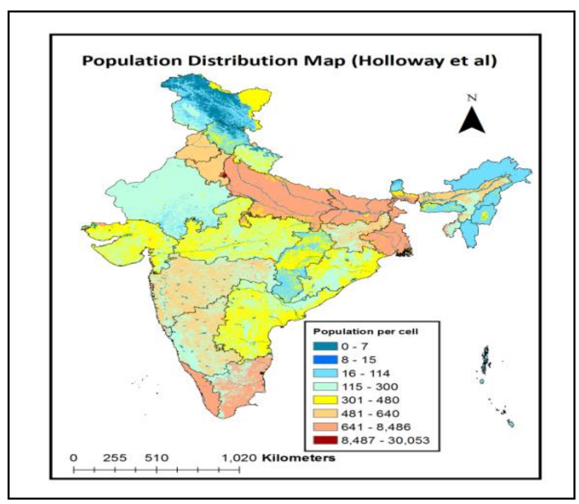
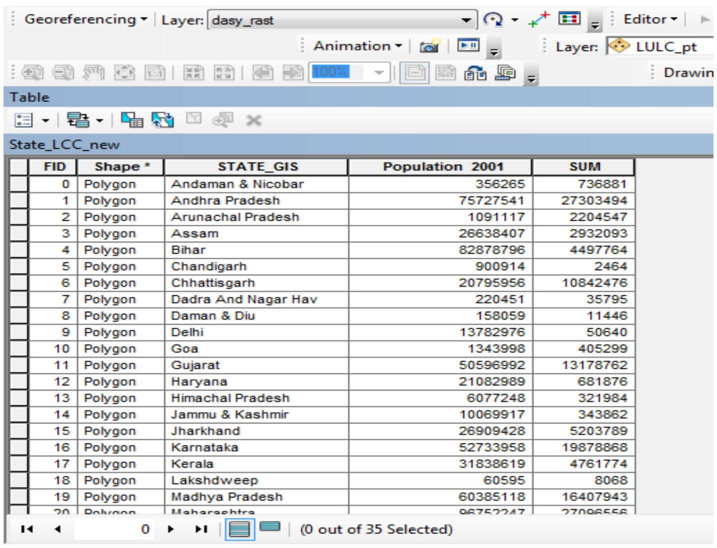


Fig: 4 – Attribute Table showing comparison of population from census data and sum of the population distributed to cells of each state by Dasymetric Mapping using Holloway et al (1997) method.

Fig: 5 – Map showing population distribution generated by using methodology given by Holloway et al (1997)

3) **Dasymetric Mapping using tool developed for ArcGIS:** Fig: 7 shows the resulting raster surface for population distribution. It overestimates the population distribution for some regions and underestimates for some other regions. It overestimates the population for Jammu and Kashmir where most of the area is uninhabited and a fairly good population is assigned to it. It underestimates the population for highly populated areas like Uttar Pradesh, Bihar, West Bengal, Delhi etc and overestimates the population for less populated areas(mainly forested) like Sikkim, Manipur, Tripura, Arunachal Pradesh. Exact figures can be seen in the table provided below in Fig: 6.



State_LCC_new

FID	Shape *	STATE_GIS	Population 2001	SUM
0	Polygon	Andaman & Nicobar	356265	736881
1	Polygon	Andhra Pradesh	75727541	27303494
2	Polygon	Arunachal Pradesh	1091117	2204547
3	Polygon	Assam	26638407	2932093
4	Polygon	Bihar	82878796	4497764
5	Polygon	Chandigarh	900914	2464
6	Polygon	Chhattisgarh	20795956	10842476
7	Polygon	Dadra And Nagar Hav	220451	35795
8	Polygon	Daman & Diu	158059	11446
9	Polygon	Delhi	13782976	50640
10	Polygon	Goa	1343998	405299
11	Polygon	Gujarat	50596992	13178762
12	Polygon	Haryana	21082989	681876
13	Polygon	Himachal Pradesh	6077248	321984
14	Polygon	Jammu & Kashmir	10069917	343862
15	Polygon	Jharkhand	26909428	5203789
16	Polygon	Karnataka	52733958	19878868
17	Polygon	Kerala	31838619	4761774
18	Polygon	Lakshdweep	60595	8068
19	Polygon	Madhya Pradesh	60385118	16407943

(0 out of 35 Selected)

Fig: 6 – Attribute Table showing comparison of population from census data and sum of the population distributed to cells of each state by Dasymetric Mapping using ArcGIS tool .

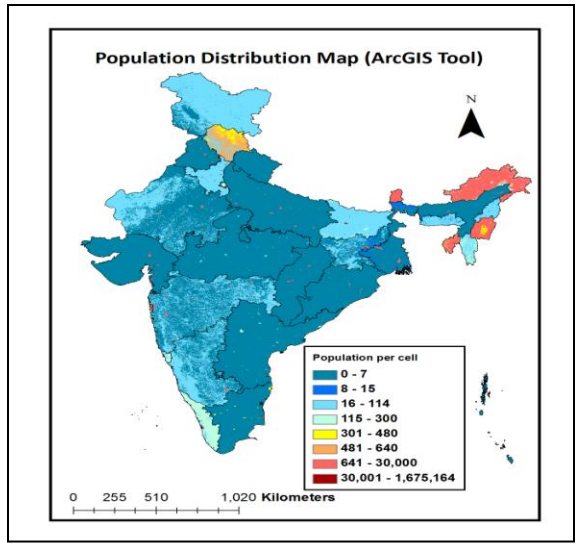


Fig: 7 – Map showing population distribution generated by using ArcGIS tool

On comparing the three methods the methodology given by Mennis (2003) proves to be promising since it takes into account both the population densities per urbanization class and also the proportion of area covered by each urbanization class. On the other hand Holloway et al assigns a predetermined proportion of population for each class for all the states which is not very realistic because every state will have different proportion of population residing in different urbanization classes. Also, the subjectivity and accuracy of this percentage assignment can be questioned. The tool does not provide satisfactory results for most of the states, in fact it gives an incorrect estimation for most of the highly populated and least populated areas.

Dasymetric Mapping proves to be promising solution for problems like MAUP (modifiable areal unit problem). For any Raster based GIS analysis, when there is a need to incorporate socio-economic data, using such map produced by dasymetric mapping might produce more accurate result rather than simply converting the vector based feature data to raster data. Studies like Rain Water Harvesting suitable sites, crime mapping, disaster planning, population modelling or any other studies which incorporate demographic or socio-economic data as parameters utilise these maps, which, if not generated properly can greatly affect the final results.

References:

1. Holloway, S., J. Schumacher, and R. Redmond, 1997, People & Place: Dasymetric Mapping Using Arc/Info. *Cartographic Design Using ArcView and Arc/Info*, Missoula: University of Montana, Wildlife Spatial Analysis Lab.
2. Mennis, J., 2003, Generating Surface Models of Population Using Dasymetric Mapping. *The Professional Geographer*, 55(1), 31-42.
3. Unwin, D.J., 1996, GIS, Spatial Analysis and Spatial Statistics. *Progress in Human Geography*, 20(4), 540-441.

4. Openshaw, Stan. 1983. The modifiable areal unit problem. *Concepts and Techniques in Modern Geography*, vol. 38. Norwich: Geobooks.
5. Eicher, Cory L., and Cynthia A. Brewer, 2001, Dasymetric mapping and areal interpolation: Implementation and evaluation. *Cartography and Geographic Information Science* 28 (2), 125–38.