Advances in Social Science, Education and Humanities Research, volume 79
1st International Conference on Geography and Education (ICGE 2016)

Using OpenStreetMap Data for Population Distribution Model

Nurrokhmah Rizqihandari
Department of Geography
Faculty of Mathematics and Natural Science, University of Indonesia
Depok, Indonesia
Corresponding email: n.rizqihandari@ui.ac.id

Satria Indratmoko
Center for Applied Geography Research
Faculty of Mathematics and Natural Science, University of Indonesia
Depok, Indonesia

***Abstract*—Understanding urban population dynamic with its spatial structure becomes problems because of the rapid changes in a city. The detailed scale of continuum data is needed to get an exact picture of this phenomenon. Census as a method to discover the population distribution detail was constrained by its own time limit of the data. The development of GIS technology able to bridge these problems. Based on participative mapping via OpenStreetMap, population surface models can be created into detail scale. Depok City development which is characterized by high rates of population growth and the expansion of residential areas and areas that woke up quickly, make this city a perfect place for the application of this method.**

Keywords—OpenStreetMap, population, urban.

I. INTRODUCTION

The high rate of urbanization is a main problem in most major cities in developing countries. This happens because of two things: the rural-urban migration caused by the appeal of the main city (primate city) and expansion of the urban area is massive because the land needs to be settlements and centers, Jakarta and the area surrounding it (Bodetabek) also with this issue.

Statistical data mention that in the period 1971-2000 in-migration rate in Jakarta likely to increase. There are 1.8 million people went to Jakarta in 1971 and it's doubled in 2000. But after 2000, the trend had been changed. Migrant no longer make Jakarta as a migrant destination, but other areas around Jakarta. If in the 1971-1980, the Jakarta population growth rate about 5.4% per year, but in the 2000-2010 it has dropped to 1.42% per year. In contrast, during 2005-2010, the rate of in-migration rate in Bodetabek is rising between 2 - 5,5% (BPS, 2010). Those statistical data are in line with the growth of Jakarta's urban area, which even exceeded its jurisdiction (over bounded) as shown in Figure 1 below.

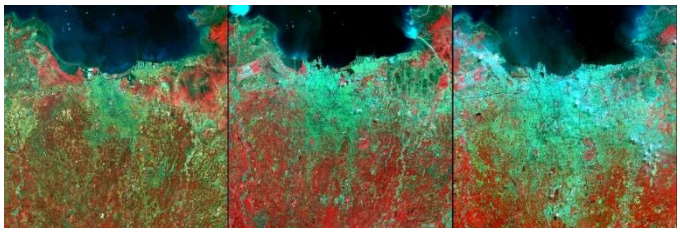

Fig. 1. Urban Growth Area in Jabodetabek 1976, 1989 and 2004

The declining of in-migration in Jakarta and the expansion of urban areas out of Jakarta show that migrants are no longer move to Jakarta, but to Jakarta's hinterland. The Increasing of migration flows is reflected in the 2010 Population Census data which the recent migrant ratio in Jakarta below its hinterland.

TABLE I. RECENT MIGRANT RATIO IN JABODETABEK, 2010.

District/City	Recent Migration	Ratio on 100 People
DKI Jakarta	756,778	8.62
Tangerang	136,407	5.35
Tangerang City	176,886	10.88
South Tangerang City	193,974	16.54
Bogor	243,574	5.70
Bekasi	266,081	11.33
Bogor City	53,864	6.25
Bekasi City	269,629	12.73
Depok City	198,423	12.66

Source: BPS, 2010

Referring to The Table 1, although the percentage of recent migration in Depok not as much as other cities in Bodetabek, but statistical data in Depok show the symptoms of the rising recent migration rates. Population data indicate that the population growth rate in Depok is 5.7% per year, with 70% is contributed by in-migration rate (BPS Depok, 2013). The

population density in the city of Depok was doubled just in 3 years, from 4.739 people/km2 in 2000 and increased to 9.796 people/km2 in 2013. These rapid changes cause an increase in residential areas in Depok.

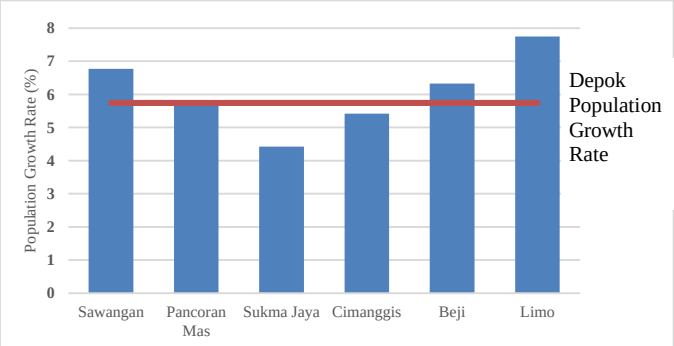

Fig. 2. Sub-District and Depok Population Growth Rate, 2000 – 2013.

Population statistical database just showing static data of population and avoiding the dynamic changes in demographic situation in Depok City. Furthermore, statistical database just shown in administrative unit. This administrative boundary make population data is shown as a discrete data, whereas population distribution spread continuously along settlement area.

Geographic Information System technology continues to evolve to address this problem some research done by applying surface models population as was done by Martin (1998) in Southampton, Eincher and Brewer (2001) in several cities in the UK, or Wang and Meng (1999) in Shenyang and Beijing China. Mennis's (2003) research in Philadelphia, USA used 100m image resolution to obtain settlement area. He has succeeded in visualizing the distribution of the population near to its reality.

Based on some previous research, population surface model can be applied to show the dynamically demographic situation in Depok City. OpenStreetMap (OSM), one of crowdsourcing apps, can be used to replacing the detail image that used in Mennis's research. By using OSM database can answer the need of continuously population distribution database. From Figure 2 Sub-District Pancoran Mas and Sub-District Beji were choose to generate a population surface model that showed population distribution as a continuum data.

II. METHOD

A. Data Collection

The data collection consists of two types of data, ie data spatial and tabular data. Building parcel data as a spatial data in detail scale was developed from crowdsource mapping via Open Street Map (OSM). A project was created to build building parcel data of Beji Area and it involved 10 students who digitized whole building plots in the scale of 1:7000. Beside OSM database, land use data was needed to define the build-up area as a basis data to determining Depok city center. Meanwhile, tabular data was obtained from Depok's Statistical Office.

The utilization of public domain database, such as OSM, has many weaknesses. Misidentification building function or different building's shapes, usually happened during digitation. To eliminate this type of error, a survey was conducted to re-checked building function and add some points of interest to improve the quality of data in database OSM. Data verification results are then uploaded back to the OSM website http://www.openstreetmap.org.

B. Raster Population Data

Raster data creation was conducted as a basis for determining the structure of population distribution. Raster data was developed by 100mx100m area per unit grid.

The initial stage before determining the distribution of the population was estimating the number of people per house building that had been digitized associated with vast proportions throughout the building located in a village.

So apply the following formula:

$$ParcelPop_{ij} = \frac{SubPop_j \times PFLArea_i}{SubPFLArea_j}$$

Where:

$ParcelPop_{ij}$ = population in building i, kelurahan j
$SubPop_j$ = population in Kelurahan j
$PFLArea_i$ = building area i
$SubPFLArea_j$ = building area in Kelurahan j

By overlapped building parcel map and grid map the population density in each grid was created. This process was done by the sum of the population within each building in the grid. Residents of the building work was interrupted by the grid (see Figure 3), calculated based on the proportion of its range. The population is then converted into density by dividing the amount to the width of one grid (10,000 m2)

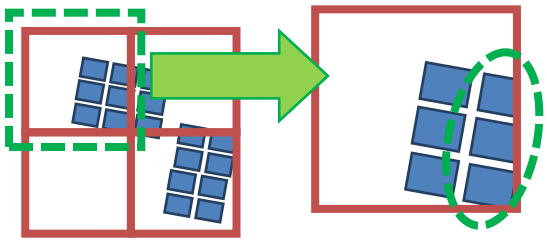

Fig. 3. An important concern in the building that is cut by the unit of analysis

Apply the following formula:

$$CellPop_i = \sum_{j=i}^{n} ParcelPop_j \times \%PInCell_{ij}$$

Where:

$CellPop_i$ = population in grid i
$ParcelPop_{ij}$ = population in building parcel j was cut by grid i.
$\%PInCell_{ij}$ = Parcel j percentage in grid i.

The next step is to change the vector data densities at each grid into raster data to create surface models using the population of polygon to raster tools in ArcGIS

III. FINDING AND DISCUSSION

A. *Cluster of Central Services Area*

Based on land use classification of Depok Land use Map, there are several land use that can be defined as non-settlement build-up area. Those land use are industrial area, strategic area, trade and services area, office area, mix-use area, transportation facilities and universities, next will be stated as central services area.

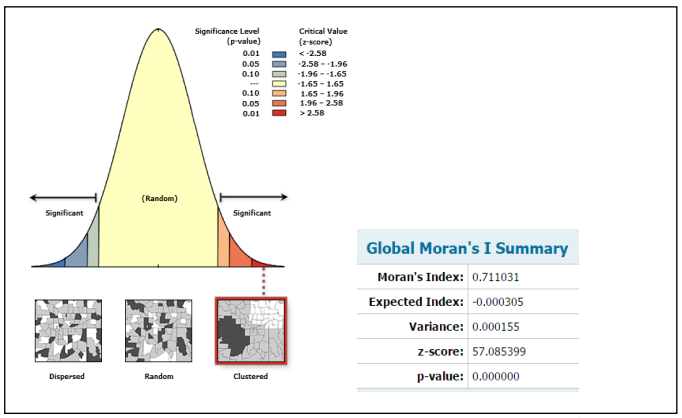

Fig. 3. Moran I Output on central services area in 100mx100m grid framework, Pancoran Mas and Beji Sub-District. Data Processing, 2012.

By using Cluster and Outlier Analysis (Anselin Local Morans I) on central services area in 100mx100m grid framework, clustered pattern had been generated (Moran I = 0,711). Those clusterisation occurs in following the main road, in the Mid and South part of Margonda Raya Street and along with Sawangan Raya Street.

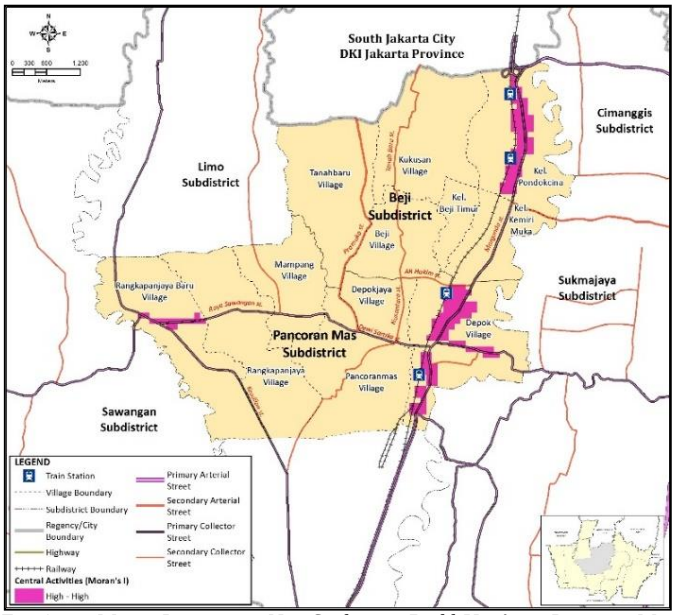

Fig. 4. Moran I output on Non-Settlement Build-Up Area. Pancoran Mas and Beji Sub-District. 2014.

B. *Population Surface Model*

Most population distribution map have been shown based on administrative boundary. Administrative boundary makes population data shown as discrete data. Meanwhile, in the reality, population distribute continually. Applying surface model method on population distribution can show the continuity of population density and can represent data close to original condition.

Several previous studies such as Martin (1998) in Southampton, Eincher and Brewer (2001) in several cities in the UK, or Wang and Meng (1999) in Shenyang and Beijing China obtained the finding that the utilization of the unit of analysis grids capable of further illustrate population distribution. By utilizing the building parcel data and data on the number of residents per municipality, the population density distribution can be obtained as described in previous method.

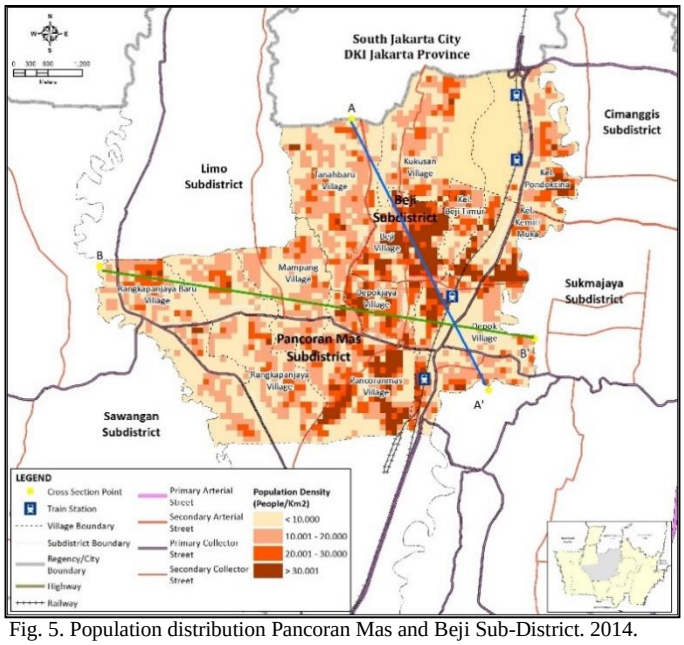

Fig. 5. Population distribution Pancoran Mas and Beji Sub-District. 2014.

The NW-SE Cross-section (A-A') show that the population is concentrated in Beji Village, the center of Beji District Beji, and in Depokjaya Village, the first settlement in Depok. Both of them relatively close to Margonda Raya, the main street of the city of Depok. More towards the outskirts of those two area, the density continually declines. The E-W cross sections (B-B') draw the dense population have association with Village Office.

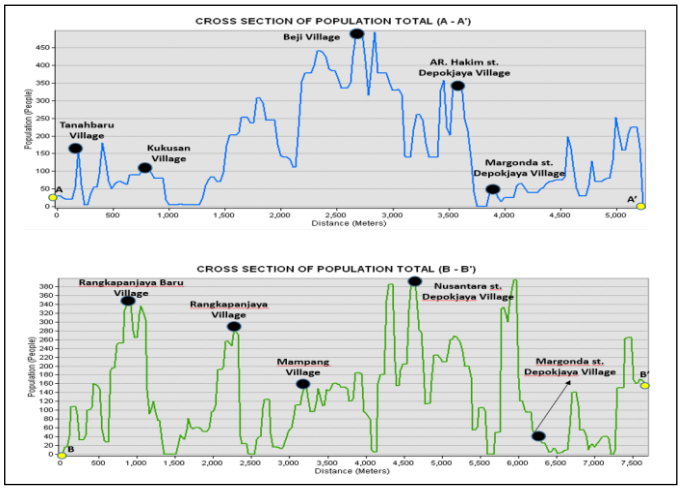

Fig. 6. Cross-Sectional Data That Show Pancoran Mas and Beji Sub-District Population Density. 2014

C. Relationship between Population Density and Cluster of Economic Activities

Figure 6 shows that every peak of dense population has association with settlement area and closed to village office. Most settlement area in Depok has own center of economic activities. Both of those graphics showed there are relationship between population density and human activities. Scatterplot graphic and mathematical exponential model can be used to prove their relationship.

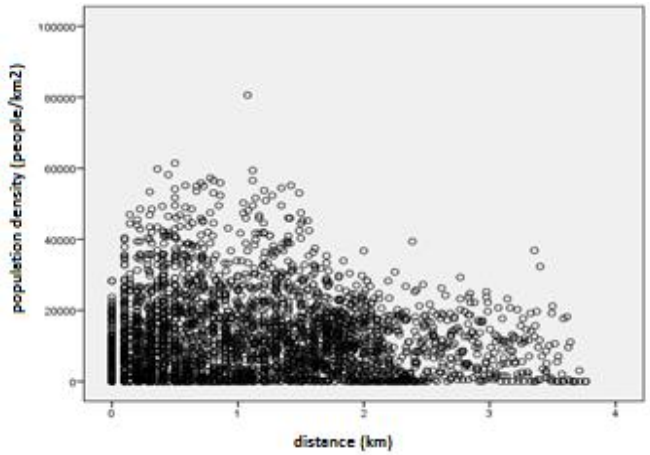

Fig. 7.Scatterplot graphic between population density and distance from city center

Scatterplot graphic between population density and distance from city center presented that dense population centrally distributed closer to economic activities, and vice versa. Statistical result showed the significances model with mathematical exponential model (R2 = 0,355; α = 99%), same result as Cuthbert and Anderson's research (2002). This exponential model followed the concentric design.

TABLE I. EXPONENTIAL OUTPUT BETWEEN EXPONENTIAL POPULATION DENSITY AND EXPONENTIAL DISTANCE

Model Statistics

Model	Number of Predictors	Model Fit statistics	Ljung-Box Q (18)			Number of Outliers
		R-squared	Statistics	DF	Sig.	
Population Density (People/km2)-Model_1	0	.345	31.502	15	.008	0

Source: Data Processing, 2012

IV. Conclusions and Suggestions

OpenStreetMap database, one of crowdsourcing mapping application, with statistical population database can be used to produce population surface model. Population surface model will show population data at detail scale. Applied population surface model in Depok Area showed that dense population centrally distributed closer to economic activities, and vice versa. This is in line with the exponential model between population density on the distance from the city center, which indicates a concentric pattern in the distribution of the population.

References

[1] Anselin L., 1995, "Local indicators of spatial association - LISA" Geographical Analysis 27 (2) 93-115.

[2] BPS, 2010. Depok Dalam Angka 2009. Depok.

[3] BPS, 2013. Depok Dalam Angka 2012. Depok.

[4] Cuthbert A L, and Anderson W P, 2002, "Using spatial statistics to examine the pattern of urban land development in Halifax-Dartmouth" The Professional Geographer 54(4) 521-532.

[5] Eicher C. L., and Brewer C A, 2001, "Dissymmetric mapping and areal interpolation: Implementation and evaluation" Cartography and Geographic Information Science 28 (2) 125-138.

[6] Martin D., 1998, "Automatic neighborhood identification from population surfaces" Computers, Environment and Urban Systems 22 (2) 107-120.

[7] Mennis J., 2003, "Generating surface models of population using dissymmetric mapping" The Professional Geographer 55(1) 31-42

[8] Wang F., and Meng Y. C., 1999, "Analyzing urban population change patterns in Shenyang, China 1982-90: density function and spatial association approaches" Geographic Information Sciences 5(2) 121-130.