Development of GIS Tool for Dasymetric Mapping

Lwin, K. K., and Murayama, Y.,
Division of Spatial Information Science, Graduate School of Life and Environmental Sciences
University of Tsukuha, Tennodai 1-1-1, Tsukuba Shi, Ibaraki Ken, Japan 305-8572
E-mail: kokolwin@live.com, mura1@sakura.cc.tsukuba.ac.jp

Abstract
Advantages of dasymetric map over traditional choropleth map have been well documented in many cartographic journals. Dasymetric uses ancillary dataset to create smaller geographical unit of population. In fact, the smaller geographical unit of population data is required for effective disaster management, emergence preparedness, retail market competition, health and disease studies, crime analysis and other population data analysis at micro-scale level. In this paper, we discuss new dasymetric mapping technique based on GIS estimated building population which was computed from building footprints, census tract and LIDAR derived Digital Volume Model DVM.

1. Introduction

Many methods to map population distribution have been practiced in geographic information systems (GIS) and remote sensing fields. Common cartographic forms of population mapping are the choropleth map and the dasymetric map. Choropleth maps provide an easy way to visualize how a measurement varies across a geographic area. However, they have limited utility for detailed spatial analysis of population data, especially where the population is concentrated in a relatively small number of villages, towns and cities. Moreover, choropleth maps cannot express statistical variation within the administrative areal units, for example changing population density. One way to avoid this limitation is by transforming the administrative units into smaller and more relevant map units through a process known as dasymetric mapping (Bielecka, 2005). Voss et al. (1999) noticed that most census boundaries do not coincide with the boundaries of geographic features such as land use/land cover, soil type, geological units, and floodplain and watershed boundaries; this is known as "spatial incongruity" and it arises when spatially aggregated data are available for one set of geographic areal units but not the areal units of primary interest. Spatial incongruity presents a major obstacle to the integration of social and natural science data and consequently places limitations on interdisciplinary research efforts. The population data segregation method was firstly utilized and termed dasymetric by the Russian cartographer Tian-Shansky, who developed the multi-sheet population density map of European Russia, scale 1:420 000, published in the 1920s (Preobrazenski, 1954). As populated territory, Tian-Shansky mapped areas within the equidistant of one verst (1067 m) from built-up terrains. The first cartographer who popularized dasymetric mapping was Wright (1936). He set forth a new method of presenting population density based upon the division of a given administrative unit into smaller areas complying with different types of geographical environments. Although dasymetric mapping has been in use since at least the early 1800s, it has never achieved the ubiquity of other types of thematic mapping, and thus the means of producing dasymetric maps have never been standardized and codified in the way other types of thematic mapping techniques have been (Eicher and Brewer, 2001 and Slocum, 1999). Therefore, dasymetric methods remain highly subjective, with inconsistent criteria. The reason for this relative lack of popularity and the paucity of standard methodology surely lies at least partially in the difficulty inherent in constructing dasymetric maps, and until recently, the difficulties in obtaining the necessary data, as well as access to the computer power required to generate them (Maantay et al., 2007). Transferring data from one set of geographic zones or districts to another set of non-coincident zones is often necessary in spatial analysis. For instance, we might have data on the number of people living within a certain census tract but need to estimate the number of people in a smaller area within the tract, or an area that includes only part of that tract and part of other tracts. We may be interested in population or other data at a watershed level and only have population data available at the census enumeration units. Several methods have been developed to generate smaller geographical units of population distribution based on aggregated values with ancillary datasets, commonly known as "Dasymetric Mapping" by using GI Science theory and practice. Maantay et al., (2007) extensively reviews on existing dasymetric methods and techniques. The following are some developed methods and approachcs in dasymetric mapping: areal interpolation, filtered areal weighting (binary method) (Eicher and Brewer, 2001), filtering with land use/land cover data (Sleeter, 2004), and cadastral-based expert dasymetric system (CEDS) (Maantay et al., 2007). Recent ycars, rescareh into micro-spatial analysis has increased due to remote sensing data available at finer spatial resolution with more diverse geo-information sources (IKONOS, QuickBird, LIDAR, etc.) and the availability of fine-scale GIS data with enhanced attribute information (e.g. building footprints with the number of floors, building use type and building name). In this paper, we estimate building population from building footprints and census tracts by integrating LIDAR derived Digital Volume Model DVM.

2. Development of GIS Tool

2.1 Methodology

Lwin and Murayama (2009) introduee two building population estimation methods: (1) Areametric (whieh does not require information on the number of building floors); and (2) Volumetric (which does require information on the number of floors). For improved accuracy, the two methods allow filtering by other categories such as: minimum footprint area and building use types, e.g. commereial, industrial, educational, and other building use types that are not occupied by residents. Figure 1 shows the abstract idea of estimations and the calculation is demonstrated by the following mathematical expressions:

Areametric Method:

$$BP_i = \left(\frac{CP}{\sum_{k=1}^{n} BA_k}\right) BA_i$$

Equation 1

Using building footprint surface area

Volumetric Method:

$$BP_i = \left(\frac{CP}{\sum_{k=1}^{n} BA_k . BF_k}\right) BA_i . BF_i$$

Equation 2

Using number of floors information:

The two algorithms have been tested with actual building population data acquired from City Office for study purpose by applying various filtering footprint sizes.

The beset result was achieved in volumetric method, filtered by 20 m^2 footprints size. Moreover, advances in remote sensing data acquisition technologies such as LiDAR can be used for the extraction of building footprints, building height (Digital Hcight Modcl (DHM)) and building volumc (Digital Volume Model (DVM)). Equations (3) and (4) can be used for LiDAR data:

$$BP_i = \left(\frac{CP}{\sum_{k=1}^{n} BA_k . BH_k}\right) BA_i . BH_i$$

Equation 3

Using average building height

$$BP_i = \left(\frac{CP}{\sum_{k=1}^{n} BV_k}\right) BV_i$$

Equation 4

Using total building volume

where:

BPi	Population of building i
CP	Census tract population
BAi	Footprint area of building i
BFi	Number of floors of building i
BHi	Average height of building i (from LiDAR data)
BVi	Total volume of building i (from LiDAR data)
i, k	Summation indices
n	Number of buildings that meet user-defined criteria and fall inside the CP polygon

2.2 A GIS Tool

We implemented a standalone GIS tool named as PopShapeGIS using the Visual Basic programming language and TatukGIS DK (Development Kit). Figure 2 shows the program flowchart of PopShapeGIS tool. Under this tool, users can define the minimum ignored footprint size such as for portieos, garbage boxes and other unpopulated areas. They ean also apply filtering by attributc field(s) such as building-use type and other attribute information. Three additional approachcs are available under the Volumetric method, namely Usc Number of Floors, Use Average Building Height and Use Total Building Volume (See in Figure 3). After processing, the estimated building population attribute field, "EST_POP", appears in a new ESRI Shape file. A map viewer is also provided for viewing the processed results by performing common GIS functions such as add map layer, zoom in, zoom out, get attribute information, label by attribute field and change map layer properties.

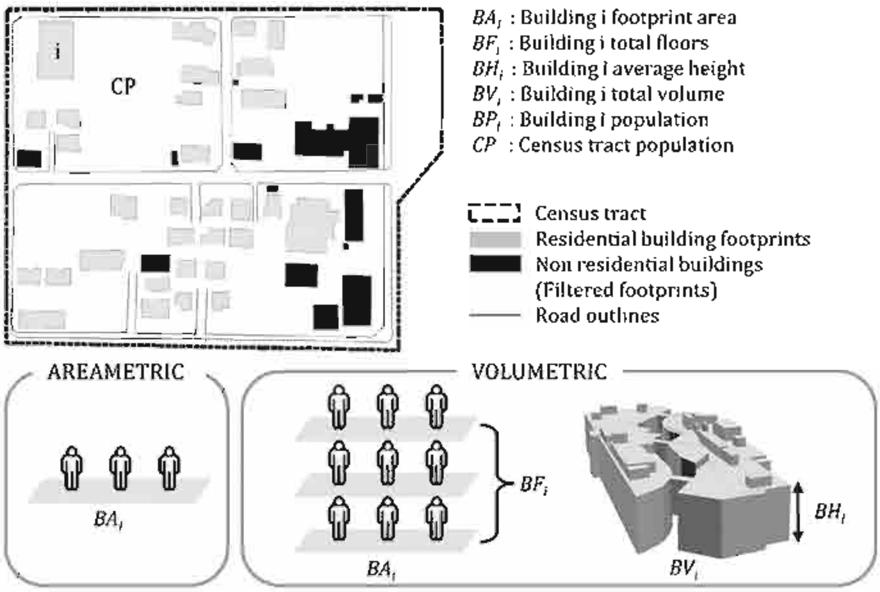


Figure 1: Graphical illustration of equations (modified from Lwin and Murayama, 2009)

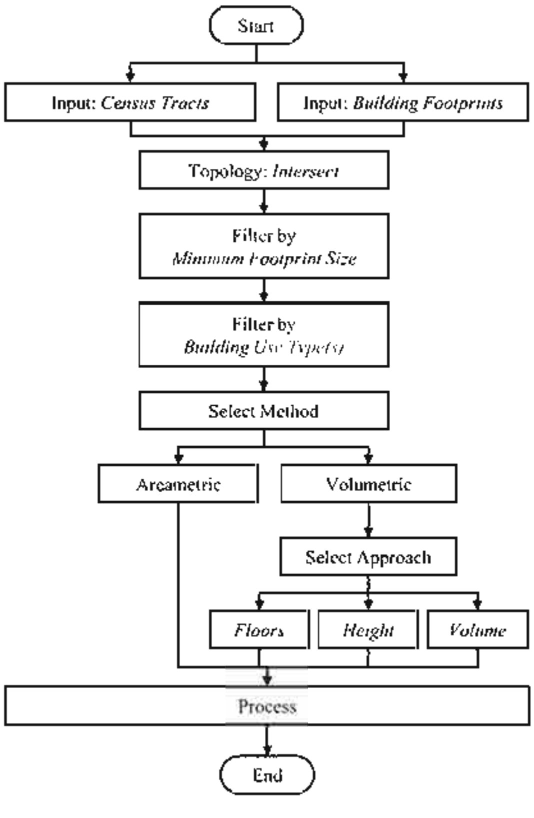


Figure 2: PopShapeGIS program flowchart (Lwin and Murayama, 2009)

More details about this tool and program can be downloaded from the following: URL. http://giswin.geo.tsukuba.ac.jp/sis/en/software.html

2.3 Data Requirements

We need two GIS dataset, census tracts and building footprints dataset. Areametric method requires only building footprints and volumetric method requires either number of floor or building height or building volume. Areametric method is suitable for rural area and volumetric method is suitable for urban area. Building use type is also required for estimation of residential building population. Nowadays, building footprints data with number of floors, tenant information and other attribute information can be purchased from commercial GIS data vendors. In addition, by utilizing modern remote sensing data acquisition system like LIDAR, we can measure building height (known as Digital Height Model DHM) and building volume (known as Digital Volume Model DVM) that can be used in volumetric method. LIDAR techniques have been studied and utilized since the early 1960s, but appear to have become more prominent in the past few years. LIDAR has found applications in a wide variety of fields of study, including atmospheric science, bathymetric data collection, law enforcement, telecommunications, and even steel production (Maune et al., 2000). Advantages of using LIDAR for terrain and urban applications include the following: LIDAR allows rapid generation of a large-scale DTM (digital terrain model); LIDAR is daylight independent, is relatively weather independent, and is extremely precise. In addition, because LIDAR operates at much shorter wavelengths, it has higher accuracy and resolution than microwave radar (Jelalian, 1992).

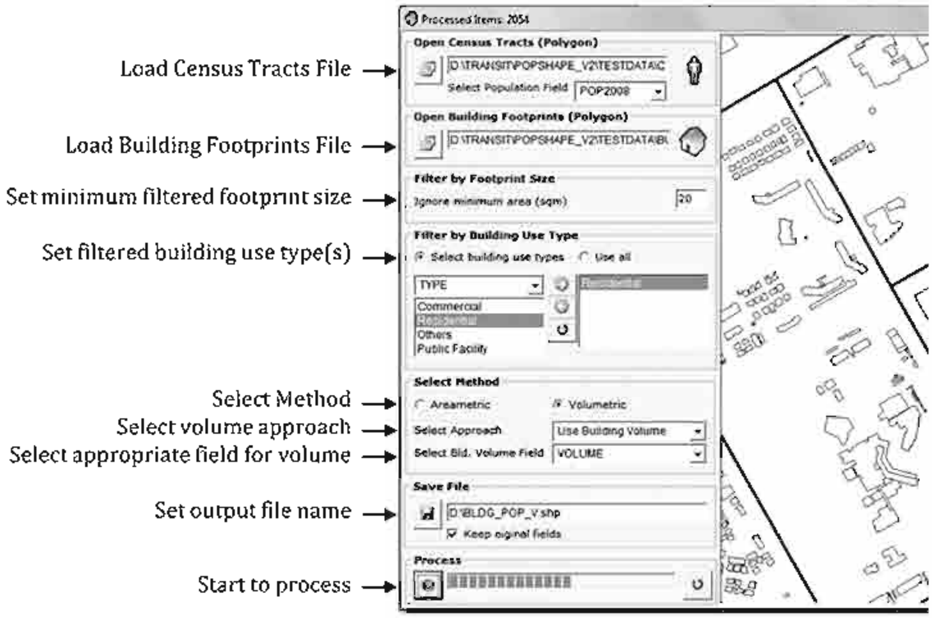


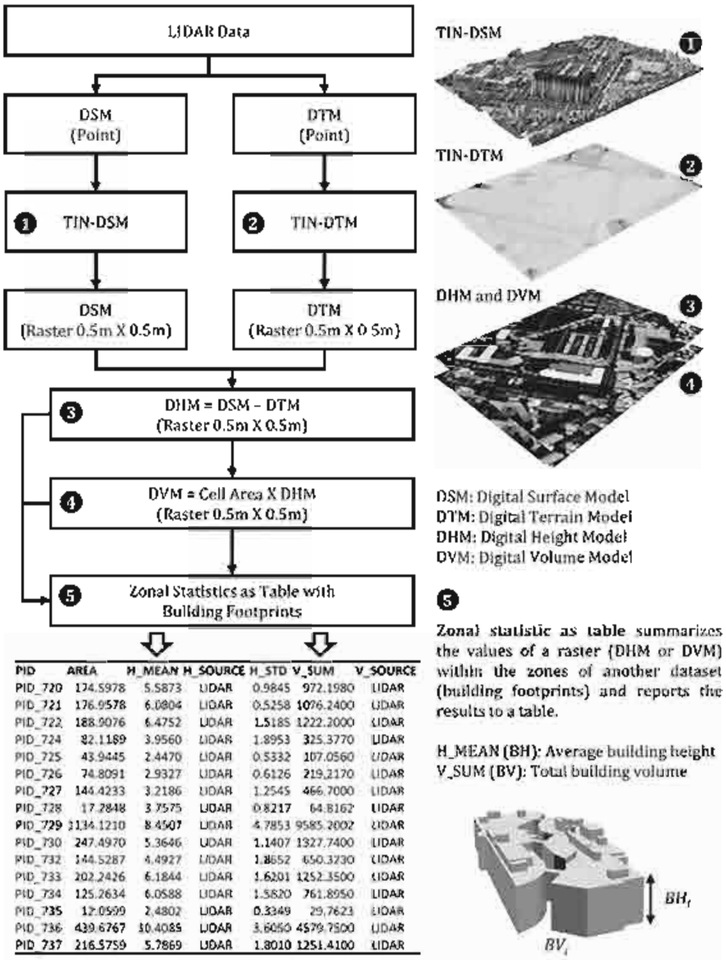
Figure 3: PopShapeGIS Graphical User Interface (modified from Lwin and Murayama, 2009)

2.4 Data Acquisitions

During the past two decades many researchers in Photogrammetry, Remote Sensing and computer vision communities have been trying to study and develop the automatic or semi-automatic approaches for building extraction and reconstruction (Gruen et al., 1997 and Mayer, 1999). Several approaches have been presented for building extraction from the laser altimeter data. Maas and Vosselman (1999) extracted buildings from original laser altimeter point data. Sahar and Krupnik (1999) developed a semiautomatic building extraction approach, which buildings were detected interactively and 3D building outlines were extracted using shadow analysis and stereoscopic processing. However, the aerial photographs are typically very complex and contain a large number of objects in the scene. The automatic building extraction from aerial photograph has proven to be quite difficult. Those approaches are far from being useful in practice for images of different characteristics and complex contents (Mayer, 1999). There are alternate ways to acquire building footprints data. In some countries, building footprints data can be purchased from commercial map vendors. For example, in Japan, building footprints and other fine scale GIS dataset can be purchased from Zenrin Map vendor which product is known as ZmapTOWNII. Moreover, availability of commercial high spatial resolution satellite data such as QuickBird (0.61m at Nadir) or orthorectified aerial images can be used as base map for onscreen building footprints digitizing.

3. Dasymetric Mapping

We have produced Tsukuba City dasymetric map using building footprints (ZMapTOWNII product) and census tracts population from city office. LIDAR data provided by PASCO Corp. was utilized for measurement of building heights and volume in order to use in volumetric method. Additional iTownpage (Internet town page) from Nippon Telegraph & Telephone Corp. (NTT) which includes address of business centers, government organizations, public facility centers and other business activities information was converted into point features in ESRI Shape format using geocoding and address matching software which geocoding accuracy is building level. This iTownpage data was used to separate residential, non-residential and mixed building use types. Figure 4 shows the building height and volume extraction from LIDAR point cloud data. Figure 5 shows the separation of residential vs. non-residential building footprints using NTT iTownpage data. Original form of this data is Comma Separated Value CSV format and converted into ESRI Point feature using geo-coding software. This point feature layer is intersects with building footprints layer and then assigned to non-residential buildings. We also assigned the building footprints which height is less than 2m and area is less than 20m^2 such as bicycle stand roofs, garbage boxes, porticos, etc. to non-residential.



PID	AREA	H_MEAN	H_SOURCE	H_STD	V_SUM	V_SOURCE
PID_720	174.5978	5.5873	LIDAR	0.9845	972.1980	LIDAR
PID_721	176.9578	6.0804	LIDAR	0.5258	1076.2400	LIDAR
PID_722	188.9076	6.4752	LIDAR	1.5185	1222.2000	LIDAR
PID_724	82.1189	3.9560	LIDAR	1.8953	325.3770	LIDAR
PID_725	43.9445	2.4470	LIDAR	0.5332	107.0560	LIDAR
PID_726	74.8091	2.9327	LIDAR	0.6126	219.2170	LIDAR
PID_727	144.4233	3.2186	LIDAR	1.2545	466.7000	LIDAR
PID_728	17.2848	3.7575	LIDAR	0.8217	64.8162	LIDAR
PID_729	1134.1210	8.4507	LIDAR	4.7853	9585.2002	LIDAR
PID_730	247.4970	5.3646	LIDAR	1.1407	1327.7400	LIDAR
PID_732	144.5287	4.4927	LIDAR	1.8652	650.3230	LIDAR
PID_733	202.2426	6.1844	LIDAR	1.6201	1252.3500	LIDAR
PID_734	125.2634	6.0588	LIDAR	1.5820	761.8950	LIDAR
PID_735	12.0599	2.4802	LIDAR	0.3349	29.7623	LIDAR
PID_736	439.6767	10.4085	LIDAR	3.6050	4579.7500	LIDAR
PID_737	216.5759	5.7869	LIDAR	1.8010	1251.4100	LIDAR

Figure 4: The work process of building height and volume extraction from LIDAR point cloud data

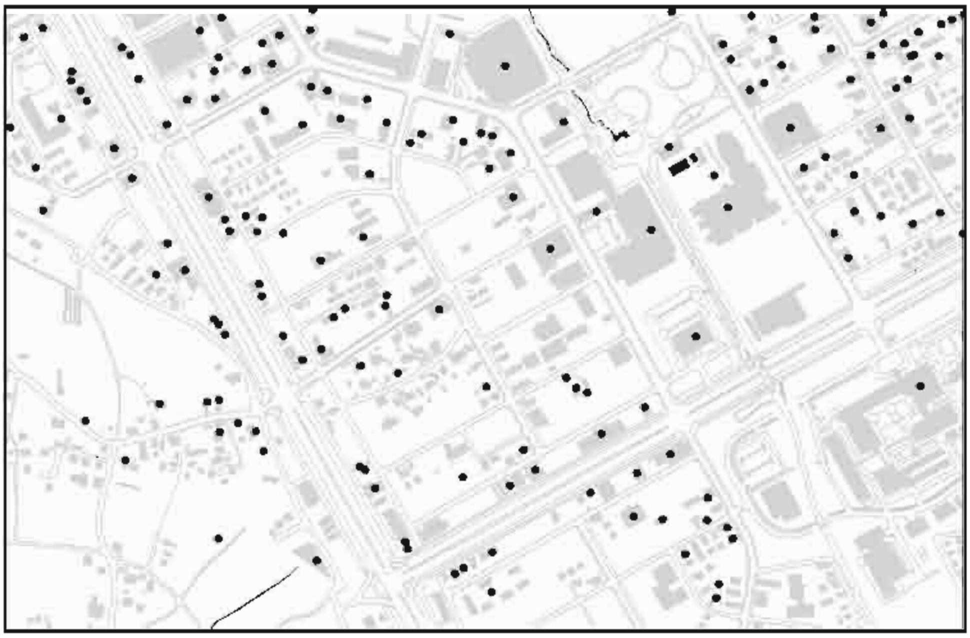


Figure 5: NTT iTownpage point layer overlays on building footprints polygon layer to separate residential vs. non-residential buildings

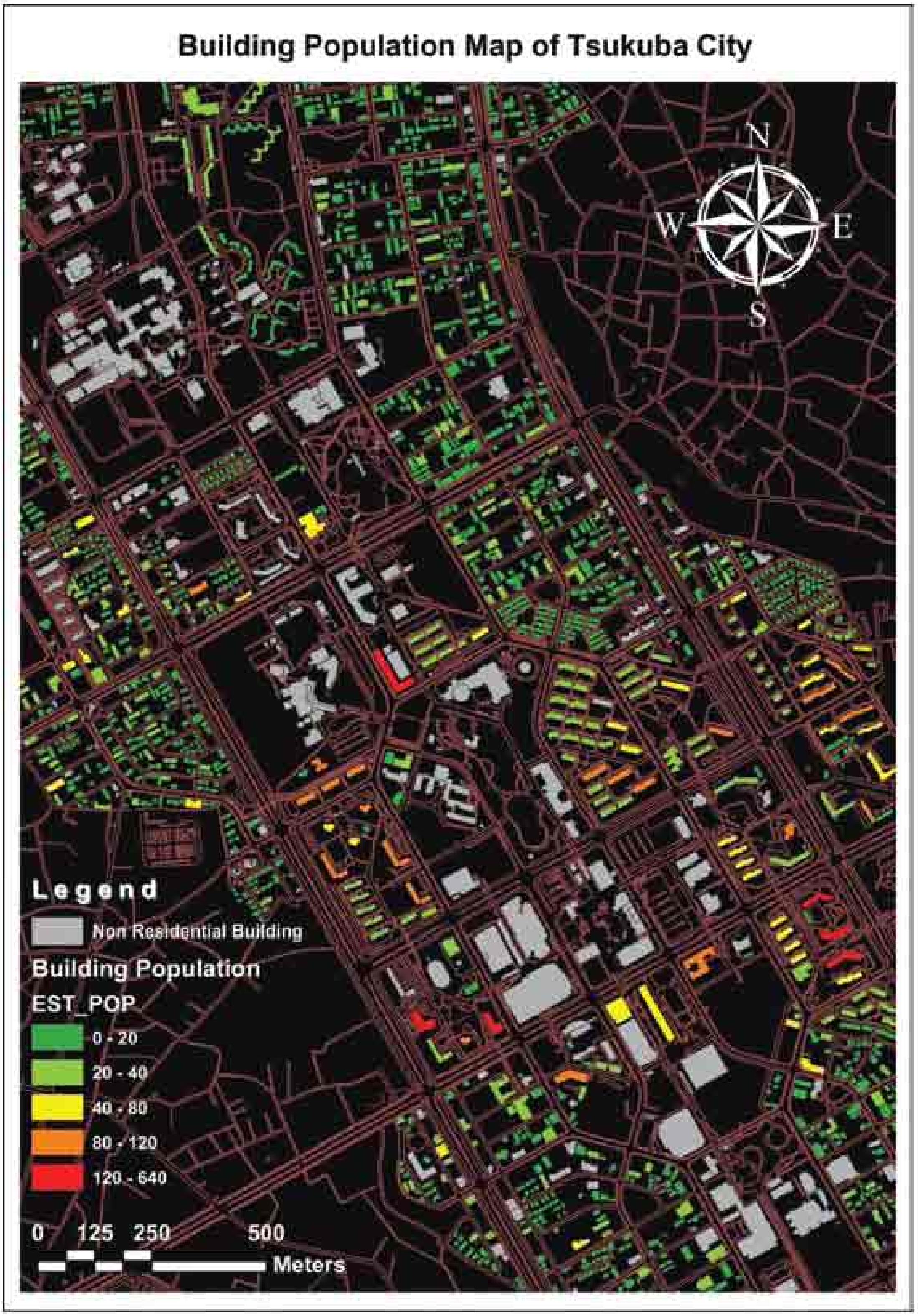


Figure 6: Dasymetric mapping of Tsukuba City Central Area based on GIS estimated building population

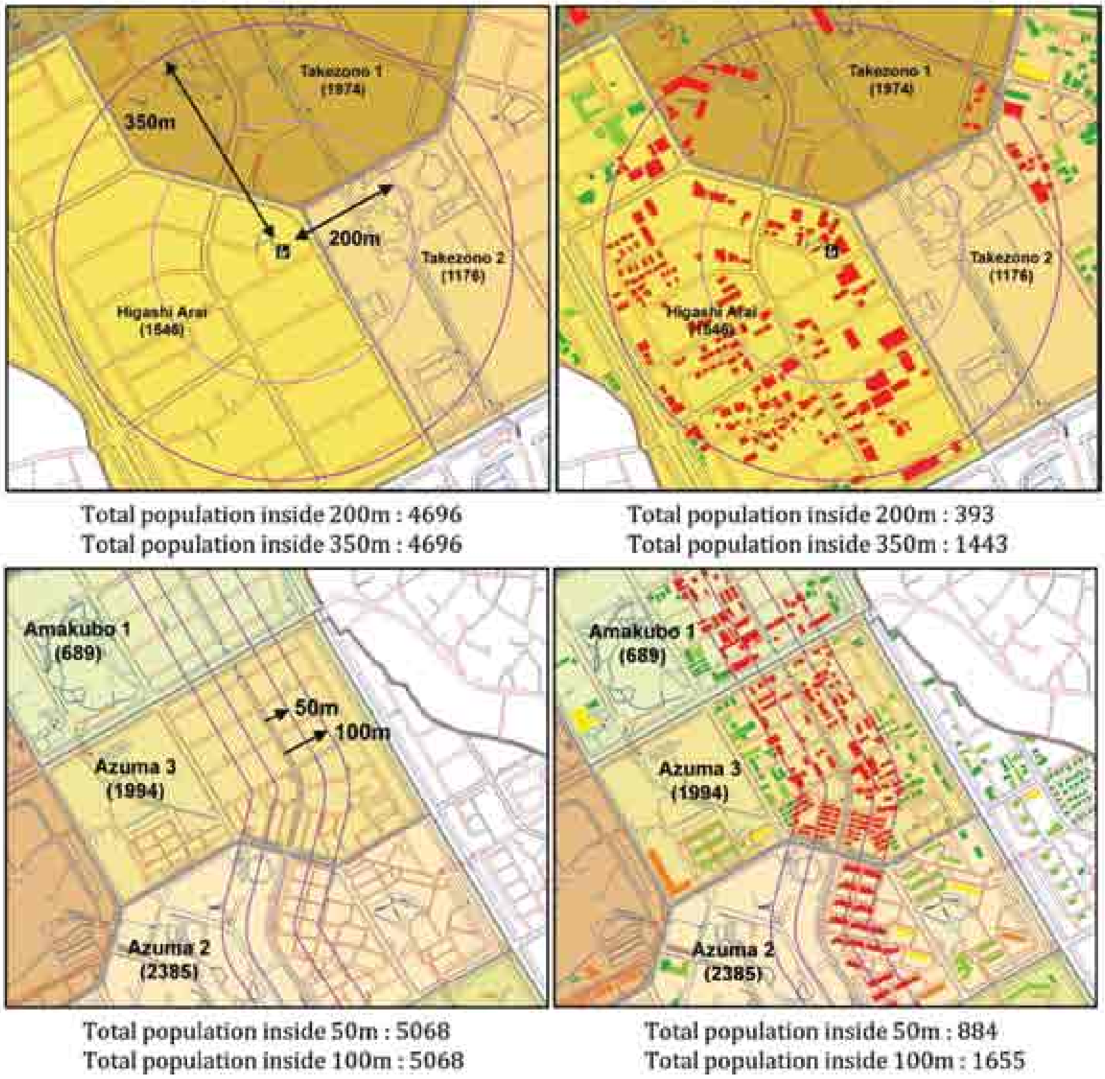


Figure 7: Comparison of various GIS analyses between chrolopleth and dasymetric map example in point and line buffering

Figure 6 shows the Tsukuba City dasymetric mapping based on estimated residential building population. In this map, we computed residential building population from two volumetric approaches, one from building floor information and one form building volume applied to residential building footprints. We averaged two results and created this dasymetric map. Building population attribute field can be used as a weighted factor for many GIS analytical functions such as determining a population weighted mean center point rather than using building mean center. We have developed a web-based interactive micro-spatial population analysis functions based on building population with other ancillary dataset such as public facility and transportation network. This web site can be reached at following URL: http://land.geo.tsukuba.ac.jp/microspa/welcome.aspx. Figure 7 shows the population results of various GIS analyses (i.e. point and line buffering) between chrolopleth and dasymetric map. Spatial analysis based on building population data is key benefit for disaster management teams in order to prepare humanitarian assistance when disaster occurred. They need specific quantitative amount of population with certain geographical unit such as 500m away from coastal lines in the case of tsunami strike or 5Km distance from earthquake's epicenter.

4. Conclusion

The estimated or quantitative mapping of building population is essential for micro-spatial analysis especially in terms of emergency management. Effective disaster preparedness requires quantitative spatial distribution patterns of population in order to position emergency response centers and prepare food and shelter in the event of disaster. Building population data is also required for improved accuracy in cost estimation of food and shelter for emergency preparedness and other humanitarian assistance. City and urban planners need to know how many local residents will benefit from newly constructed public facilities such as bus centers, railway stations and hospitals. Hydrologists require an cstimate on the number of people living on a floodplain. Potential business owners can define their business location and perform consumer analysis. Quantitative building population data can be used as a weighted factor in spatial statistical analysis such as for determining population mean eenter and standard distance. This is important for decision making related to population such as in selecting a voting site or construction a new public facility. Specific sub-population distribution for urban areas in order to develop an improved "denominator," which would enable the calculation of more correct rates in GIS analyses involving public health, crime, and urban environmental planning. Additionally, the knowledge of accurate population distribution can be extremely valuable in the sphere of urban planning. The understanding of the locational characteristics of target populations would allow for more equitable rcsource allocation in areas such as community infrastructure development, provision of open space and recreational opportunities, transportation access, and necessary environmental facilities (Maantay et al., 2007). In this dasymetrie mapping, we considered population data as a 3 dimensional space and developed a GIS tool for generating smaller geographical unit of population (i.e., building population) based on building footprints and census tracts. Although LIDAR data is expensive and require additional computer resources to handle, still we can use building surface area for rural area and building floor information for urban area. This study approach is suitable for miero-scale population analysis such as market competition analysis, public health, traffic noise impact studies, public facility planning and disaster management.

References

Bielecka, E., 2005, A Dasymetric Population Density Map of Poland, in: *Proceedings of the International Cartographic Conference, July*, A Coruna, Spain, 9-15.

Eicher, C. L., and Brewer, C. A., 2001, Dasymetric Mapping and Areal Interpolation: Implementation and Evaluation, *Cartography and Geographic Information Science*. 28(2):125-138.

Gruen, A., K_bler, O., and Agouris, P., 1997, *Automatic Extraction of Man-Made Objects from Aerial and Space Images (II)*, Birkhauser Verlag, Basel-Boston- Berlin.

Jelalian, A., 1992, *Laser Radar Systems*, Artech House, Boston.

Lwin, K. K., and Murayama, Y., 2009, A GIS Approach to Estimation of Building Population for Micro-Spatial Analysis, *Transactions in GIS*, 13(4):401-414.

Maantay, J. A., Maroko, A. R., and 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.

Maas, H. G., and Vosselman, G., 1999, Two Algorithms for Extracting Building Models from Raw Laser Altimetry Data, *ISPRS Journal of Photogrammetry and Remote Sensing*. 54(2-3):153-163.

Maune, D., Daniel, C. and Damron, J., 2000, LIDAR and IFSAR: Pitfalls and opportunities for our future, in: *Proceedings of the American Society for Photogrammetry and Remote Sensing Conference, Washington, DC.*

Mayer, H., 1999, Automatic Object Extraction from Aerial Imagery-A Survcy Focusing on Buildings, *Computer Vision and Image Understanding*. 74(2):138-149.

Preobrazenski, A., 1954, Dorewolucjonnyje I Sovietskije Karty Razmieszczenija Nasielenia, *Woprosy Geografii Kartographia*. 34:134-149.

Sahar, L., and Krupnik, A., 1999, Semiautomatic Extraction of Building Outlines from Large-Seale Aerial Images, *Photogrammetric Engineering and Remote Sensing*. 65(4):459-466.

Sleeter, R., 2004, Dasymetric Mapping Techniques for the San Francisco Bay Region, California, in: *Proceedings of the Urban and Regional Information Systems Association Annual Conference*, Reno, Neveda.

Slocum, T. A., 1999, *Thematic Cartography and Visualization*, Prentice Hall, New Jersey.

Voss, P. R., David, D. L., and Roger, B. H., 1999. When Census Geography Doesn't Work: Using Ancillary Information to Improve the Spatial Interpolation of Demographic Data, Center for Demography and Ecology, University of Wisconsin, Madison.

Wright, J. K., 1936, A Method of Mapping Densities of Population: With Cape Cod as an example, *Geographical Review*. 26(1):103-110.