GEODESY AND CARTOGRAPHY
Vol. 65, No 2, 2016, pp. 149-162

DOI: 10.1515/geocart-2016-0011

Redistribution population data across a regular spatial grid according to buildings characteristics

Beata Calka, Elzbieta Bielecka*, Katarzyna Zdunkiewicz

Military University of Technology
Faculty of Civil Engineering and Geodesy
Institute of Geodesy
2 Gen. S. Kaliskiego St. 01-476 Warsaw
e-mail: beata.calka@wat.edu.pl; elzbieta.bielecka@wat.edu.pl
zdunkiewicz.ka@gmail.com;

*Corresponding author: Elzbieta Bielecka

Received: 15 August 2016 / Accepted: 10 October 2016

Abstract: Population data are generally provided by state census organisations at the predefined census enumeration units. However, these datasets very are often required at user-defined spatial units that differ from the census output levels. A number of population estimation techniques have been developed to address these problems. This article is one of those attempts aimed at improving county level population estimates by using spatial disaggregation models with support of buildings characteristic, derived from national topographic database, and average area of a flat. The experimental gridded population surface was created for Opatów county, sparsely populated rural region located in Central Poland. The method relies on geolocation of population counts in buildings, taking into account the building volume and structural building type and then aggregation the people total in 1 km quadrilateral grid. The overall quality of population distribution surface expressed by the mean of RMSE equals 9 persons, and the MAE equals 0.01. We also discovered that nearly 20% of total county area is unpopulated and 80% of people lived on 33% of the county territory.

Keywords: population data, dasymetric modeling, spatial grid, choropleth map, topographic data

1. Introduction

Research on modeling spatial distribution of population dates back to the early twentieth century. Tian-Shansky, Russian cartographer who elaborated the dasymetric population density map of European Russia, is considered as the precursor in this field (Bielecka, 2005). However, Wright (1936) was the first who set forth this new method of portrayal population density based upon the division of a given administrative unit into smaller areas complying with different types of geographical environment. Wright derived population density values subjectively, the subjectivity is yet considered as the main disadvantage of the method. Recent advances in GIS and remote sensing, as well as increased availability of digital datasets both vector and raster, have revitalized the idea of mapping population density in units that correspond well with real population distribution. Based on the literature Wu et al. (2005) classified the methods into two main categories: areal interpolation and statistical modelling (as cited in Murshed, 2009). Areal interpolation utilizes census data as input and apply different disaggregation techniques, with ancillary or without ancillary information, to obtain more robust population (Mennis, 2003; 2009). The additional data usually produces more accurate results than those without auxiliary information, assuming that ancillary information reflects the spatial distribution of the variables being mapped. The more often land cover/land use data derived from satellite imageries (Eicher and Brewer, 2001; Mennis, 2003; Ciołkosz and Bielecka, 2005; Gallego, 2010; Azar et al., 2013; Langford, 2013), topographic data (Su et. Al., 2010; Wu et al., 2008), street network (Zandbergen and Ignizio, 2010) or even cadastral data (Maantay et al., 2007) are used as ancillary information. However, recently buildings extracted from LIDAR data (Sridharan and Qiu, 2013) or imperviousness layer (Wu et al., 2005; Azar et al., 2010; Wu and Murray, 2005) are the source of additional information for the purposes of dasymetric population modelling. Mennis and Hultgren (2006) introduced the "intelligent" dasymetric mapping (IDM) *technique that supports a variety of methods for characterizing the relationship between the ancillary data and underlying statistical population surface*. The advantage of IDM method is it returns a set of statistics that summarize the quality of the resulting dasymetric map.

Statistical modelling infers relations between population and other variables (e.g. land use/land cover, dwellings, urban zones, image pixel characteristics) to estimate population of any given unit. These methods could also be used for prediction population distribution in a given time. Stevens at el. (2015) outline that the combination of widely avaialble remotely and geospatial data to contributed the modelled dasymetric weights and the use of statistical model Random Forest can predict population density at approximately 100 m spatial resolution. This was also confirmed by Anderson et al. (2014).

Moreover, each of these methods could preserve population totals in a census enumeration units, which is called after Tobler et al. (1995) pycnophylactic. Table 1 gives an overview of methods of disaggregation of population data described in the literature.

Table 1. Methods of disaggregation population data

Interpolation without ancillary data	Dasymetric interpolation	Statistical modelling
Point-based Methods	Binary method	Correlation with urban areas, land use, buildings, dwellings units, image pixel characteristics, socioeconomic data
Kernel-based interpolation	Three class method	Linear regression
Area-based Methods	Areal weight method	Non-linear regression e.g. Random-Forest
	Intelligent dasymetric mapping	Geographically wight regression
		Kriging, co-kriging
		Spatial autocorrelation
		Expectation-maximization algorithm (EM)

Source: modyfied after Murshed (2009)

Lately, spatio-temporal population modelling techniques have been introduced. They can distinguish population during day and night (Smith et al., 2015) or even hourly population distribution using mobile phone footprints (Deville et al., 2014), Call Detail Records (CDRs) and Person Trip Survey (PTS) data from users of mobile devices (Horanont and Shibasaki, 2010; Horanont et al., 2015).

This paper demonstrates the use of information about buildings location and their characteristic derived from Topographic Object Database, the nationwide topographic database at the scale 1:10,000, to elaborate gridded population surface for the rural areas. The novelty of this approach lays in use of information about number of storey of the building, building footprints and average area of a flat, given by Central Statistical Office (hereafter referred as CSO, in Polish: GUS), as independent variables for estimating the population count in residential buildings. The experimental population surface was created for Opatowski county, sparsely populated, rural region, located in the central part of Poland. The paper is structured as follows: the next section (section 2) describes the main methodological assumptions. Section 3 gives the overview of data used and character of the study area. Results and discussion are provided in section 4, and final remark – in section 5.

2. Adopted methodology

We assume that population and population density are based on functional definition of population, thus means on place of residence, referred also as night time population. Moreover, the elaborated approach tends to preserve the total volume of population at the commune level, the lowest level of administration units in Poland (corresponds to the EUROSTAT NUTS5) as well as, the smallest enumeration area for which statistical population data are commonly available. The method relies on geolocation of population count in buildings, taking into account the building volume and structural building type and then aggregation the people totals in 1 km quadrilateral grid cells. The correctness of people estimation is verified at the commune level by summing population attributed to grid cells that overlap the area of commune.

As the approach lay on population distribution on buildings level the main research problem comes down to estimate population count for a multi apartment (and multi-storey) residential building *Pop(b)* based on equation (1):

$$Pop(b) = n(s) \cdot (A(b) / Av(f) \cdot \mathrm{p(f)_i}) \tag{1}$$

where:
n(s) – number of storey in a residential building,
A(b) – area of footprint of a building,
Av(f) – the average area of flat, given by the Central Statistical Office
$p(f)_i$ – average number of people in one flat in a *i* commune.

Hence, the number of people in a 1km rectangular grid cell *Pop(g)* is calculated as:

$$Pop(g) = \sum_1^j Pop(b) \tag{2}$$

where: *j* – the number of buildings which centroids are completely within the grid cell.

And the estimated population totals in a commune *Pop(c)* is computed according to the formula:

$$Pop(c) = \sum_1^m Pop(b) \tag{3}$$

where:
m – number of buildings in a commune,
Pop(b) – population count for a multi-apartment residential building.

The coefficient *p(f)* is constant within a communes and could differ between communes. Its values are adjusted during an iterative computation of population totals in a commune. The process stops when the difference between estimated population counts and the number of people given for Opatów county by CSO is less than 1 person and for every commune does not exceed 0.2% of population totals. For verification of a number of flats in the residential buildings StreetView photos from Google Maps were used.

The research has been conducted according to the three-stage work flow:

Stage I – people estimation at the building level.

(1) Selection residential buildings using ‘select by attribute’ query.
(2) Computing number of people in a building using equation (1).
(3) Compute number of people in a commune according to equation (3) – first iteration.

(4) Control of computed population totals in a commune and in the Opatów county.
(5) Adjust the $p(f)_i$ coefficient.

The steps from (3) to (5) are repeated during iteration until the population totals are preserved.

Stage II – population redistribution over 1km rectangular grid
(6) Creating the 1 km spatial grid.
(7) Assigning buildings centroid to the grid cells by 'spatial join' relation.

Stage III – visualization of gridded population surface.
(8) Choropleth map creation.

Results of modeling, the raster population surface, is visualized by the means of choropleth map, with regular 1 km quadrilateral grid in the Lambert Azimutal Eaqual-Area projection coordinate reference system. The spatial resolution and the coordinate reference system were adjusted to the requirement of INSPIRE Directive (INSPIRE, 2010). For visualization purpose population total in grid cells were grouped into 6 classes natural breaks method according to rules made by Medyńska-Gulij (2014) and de Smith et al. (2015).

3. Study area and data used

3.1. Study area characteristic

Opatów, the rural, sparsely populated county, is located in the central part of Poland, in the Świętokrzyskie voivodship (Figure 1). The census data are provided for 8 communes, out of which 6 are the rural and 2 urban-rural (including small cities: Opatów and Ożarów). The commune type (rural or urban-rural) has indisputable impact on the character of built-up areas. In rural areas, single-family housing prevails, while in cities are mostly multi-storey, multi-family buildings. The centres of Opatów and Ożarów are made of a mixture of urban fabric of varied density and height, commercial, communication and public areas as well as green urban areas. Opatów County is one of the least populated counties in the voivodship. For several years the population in the county has been steadily decreasing, in 2014 was 54,248 and was by 3.2% lower than in 2013, a similar decline was recorded in the years 2012-2013. This is due to negative natural growth and negative migration balance. According to the 2014 official statistical data (GUS, 2015) the population density of Opatów county is 60 people per square kilometer, this is 48 people lower than average population of the voivodship, and more than twice time than national average (123 person per sq. km). Characteristic of population distribution in the study area does not vary considerably, the dense populated is the city of Opatów (711 person per sq. kilometer) and city of Ożarów (601 person per sq. km), the lowest population is observed in Tarłów (33 person per square kilometer). This is illustrated in Figure 1 (see also Table 1).

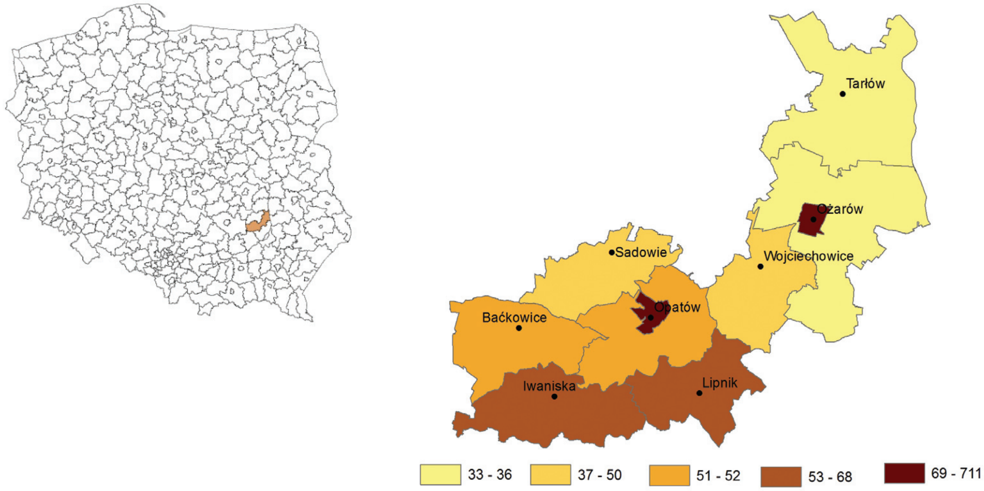

Fig. 1. Opatów county location and population density

The average indicators of living conditions in the analysed area are somewhat above Polish average. According to CSO data for 2015 (GUS, 2015a) the average flat has an area of 74 m^2 (national average: 73.40 m^2) and is home to 2.90 people (national average: 2.98).

3.2. Data used

Population totals aggregated to the communes, county and voivodships are delivered by the CSO in the form of reports, xls tables, graphs and maps. The population is a population balance as of 31st December 2014 made on the basis of Population and Housing Census 2011 then recalculated according to the territorial division of the country as of 1st January 2015 (GUS, 2015b).

Information on buildings' location, number of storey, as well as the type of the building are derived from Topographic Objects Data Base (referred thereafter as BDOT10k), the national topographic data cover the whole country, at the level of details corresponds to civilian topographic maps 1:10,000. Each building, represented geometrically by polygon, is characterised by several thematic attributes and locational accuracy expressed by RMSE equals 1.3 m (Bielecka 2015). The BDOT10k data are integrated with administrative units boundaries, at the commune level, originally available from National Border Register. Based on several previous research (Bielecka et al., 2014; Bielecka, 2015) we assume that it is a reliable data source for creating population distribution surface. For Opatowski county the BDOT10k data are up-to-date to the year 2013. Therefore, the temporal validity of population distribution is the January 2013. Consequently, population totals from national census and administrative units boundaries were taken for the year 2013.

The study exploits only residential buildings, for which the values of '*kod*' attribute are equals: *BUBD01* (single family building), *BUBD02* (two-families building), and *BUBD03* (multi- apartment building with 3 and more flats). The detail description of the application schema of topographic objects, including data on buildings, can be found in Regulation of Ministry of Internal Affairs and Administration (Regulation, 2011) and textbook edited by Olszewski and Gotlib (2013).

4. Results and discussion

4.1. Population estimation

The total number of residential buildings in the study area equals 16,592, out of which 222 are multi-apartments (multi-family) and 16,370 are single-family. With 75% of multi-family buildings located in two cities Opatów and Ożarow (Table 1, Figure 2).

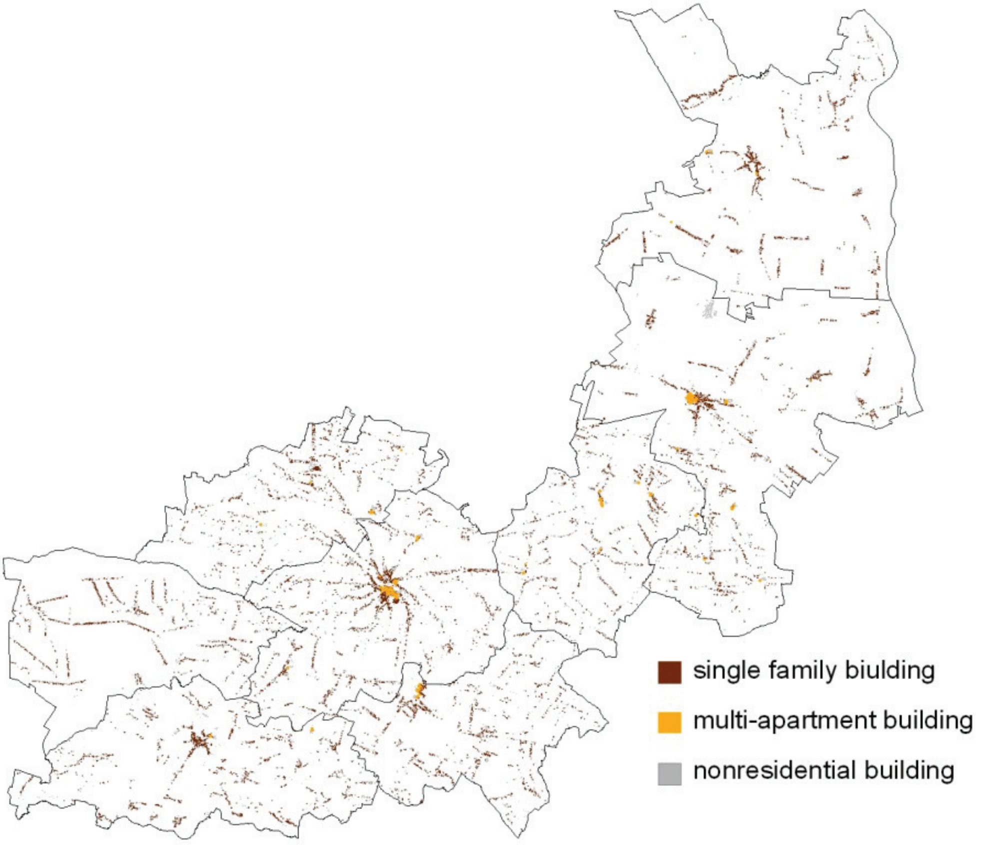

Fig. 2. Buildings in Opatów county

Table 1. Population within study area

Commune	Type of commune	Single family buildings	Multi-apartment buildings	Population total by CSO*	Population density*
Opatów – town – rural area	urban-rural town rural area	3023 1174 1849	105 95 10	12102 6657 5445	107 711 52
Ożarow – town – rural area	urban-rural town rural area	2889 536 2353	60 47 13	11067 4682 6385	60 601 36
Baćkowice	rural	1565	0	4984	52
Iwaniska	rural	2113	3	6869	65
Lipnik	rural	1666	17	5502	68
Sadowie	rural	1395	11	4096	50
Tarłów	rural	2303	7	5407	33
Wojciechowice	rural	1416	22	4221	49
Opatów county	rural	16370	222	54248	60

**source: data published by CSO (GUS, 2015b)*

The number of people in each building was estimated using equation (1), starting with the value of average number of people in one flat *p(f)* equals for any commune and takes value 2.90. (GUS,2015a) The average area of a flat in the analysed region (Świętokrzystkie voivodship) *Av(f)* is 74.00 m^2. The results after the first iteration was unsatisfactory, the population total in Opatów county was overestimates approximately of 3.5%, and any correlation with the type of commune has been observed. Population counts were overestimated in 2 urban-rural communes as well as in 3 rural communes. For the remaining 3 rural communes the number of people was underestimated, from 2.3% to 10.7%. The highest overestimation (25.3%) was denoted in Tarlow and the lowest (2.3) in Sadowie.

The adjusting *p(f)* coefficient was done iteratively with the step 0.01. The number of iteration varies from 7 for Sadowie to 59 for Tarłów. The final results of population estimation in a commune based on number of flats and average number of people living in 1 flat are presented in Table 2.

The root mean square error (RMSE) and the mean absolute error (MAE), which are well known accuracy measure in the field of estimation (Bielecka and Bober, 2014; Langford, 2013) equals 9 and 0.01 respectively. Moreover, the coefficient of determination $R^2 = 0.96$ means that population totals in communes are in 96% explained by the number of flats.

Table 2. Population estimation using ancillary data on building characteristics

Commune	Estimated number of flats (based on BDOT10k)	Number of iteration	Adjusted *p(f)*	Estimated population	Population total by CSO*	Difference in population totals (%)
Opatów	4582	20	2.70	12092	12102	-10 (0.08%)
Ożarow	4104	17	2.73	11086	11067	19 (0.17%)
Baćkowice	1565	33	3.23	4989	4984	5 (0.10%)
Iwaniska	2135	35	3.25	6870	6869	1 (0.01%)
Lipnik	1846	8	2.98	5501	5502	-1 (0.02%)
Sadowie	1466	7	2.83	4090	4096	-6 (0.15)
Tarłów	2369	59	2.31	5397	5407	-10 (0.18)
Wojciechowice	1567	15	2.75	4224	4221	3 (0.07)
Opatów county	19634		2.795	54249	54248	1 (0.002%)

**source: data published by CSO (GUS, 2015b)*

4.2. Population redistribution and visualisation

The population totals in grid cells are presented in the form of choropleth map, using 6 groups and natural breaks method, excluding unpopulated area (Fig. 3). The natural breaks method is a kind of classification scheme that minimize within-class variance and maximize between-class differences. One drawback of this approach is that it is a data-specific classifications and not useful for comparing multiple maps built from different underlying information. This classification is based on the Jenks Natural Breaks algorithm (see *Univariate classification schemes* in *Geospatial Analysis—A Comprehensive Guide*, 3rd edition; 2016; de Smith, Longley, Goodchild). This allows to portrayal areas totally unpopulated, which comprises nearly 20% (175.8 sq. km) of analysed territory. They are mainly located in the north-east part of the Opatów county, within two communes: Tarłów where dominate forest (69.8%) and Ożarów with industry and intensive agriculture. The most densely populated areas are located in the centre of cities Opatów (3802 in 1 sq. km) and Ożarów (2667 in 1 sq. km).

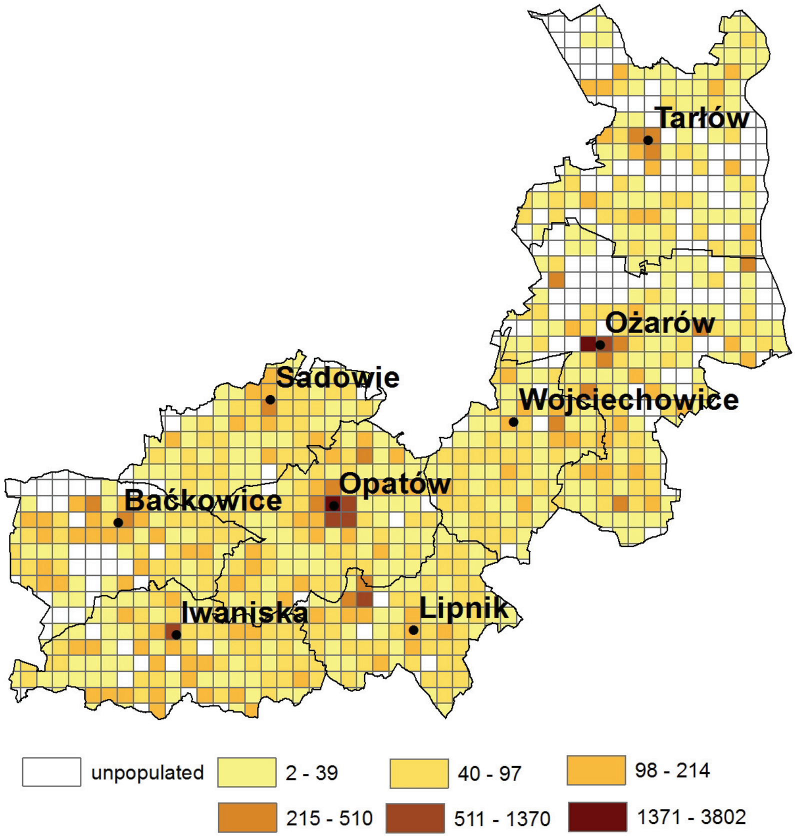

Fig. 3. Opatów county, choropleth map of population distribution, natural break method

More than 80% of total population is concentrated on the area comprising 33% of the county territory. So the allocation of people distribution is closed to the Pareto principle states that, for many events, roughly 80% of the effects come from 20% of the causes (Hazewinkel, 2001).

4.3. Discussion

The crucial impact on the quality of population redistribution, especially in a fine scale, have ancillary data sources. This issue is still open, however, many findings concerning data quality are commonly available in the literature. Data that are derived by National Mapping Agencies or other governmental institutions are generally more reliable than those from VGI. However, they are generally available for a payment. On the other hand VGI data are of different quality and completeness. Researches that employ VGI data, mainly OSM, for population distribution clearly claim the semantic inadequacy and completeness of information on buildings (Siva et al., 2013; Bakillah et al., 2014). The completeness of buildings in OSM data for Poland approximates to 20% (Marczak, 2015), and is higher in urban area than rural. Semantic correspondence of building characteristic in BDOT10k, Polish national topographic data, and the OSM data reaches 70% (Nowak Da Costa, 2016a; 2016b). Therefore, the OSM is valuable data source be used to model the distribution of population in urban areas, particularly large cities.

We used official data, that are accurate and credible but for some areas are not up-to-date. Regulation of Ministry of Internal Affairs and Administration of BDOT10k (Regulation, 2011) says that data should be updated on a regular basis, but in practice for the study area data about buildings were dated to the end of 2012. An alternative, updated on a regular basis, data source is Lands and Buildings Register, but for the analyzed area these information is available only in analog form.

Unfortunately, it was not possible to verify the number of flats which has been estimated. Nonetheless, the number of flats in Opatów County is available from national census (carried out in 2011) its value (16,061) is lower than number of residential buildings (16,592) in National Topographic Database (BDOT10k). This big divergence in numbers of flats could not be explained by number of buildings (or flats) completed in 2012-2014.

Analysis of population distribution at the building level is the subject of many studies (Bakillah et.al, 2014; Sridharan and Qiu, 2013; Wu et al., 2008). Disadvantages of this type of research are time-consuming and high cost of obtaining information about the buildings, in particular building structural type, volume and height. We overcame mention above drawbacks, but the weak point of the elaborated method is to calculate the number of flats in a multi-apartment building based on the average size of the flat in the voivodship, given by the CSO. The preliminary analysis of real estate market in Opatów county shows that this figure is far too high and should not exceed 66 m^2. In further studies, the average size of the flat could be determined based on the analysis of local real estate markets at the commune level. This certainly contributes significantly to increase the accuracy of estimation the population totals in buildings.

Still, insufficiently explored issue is the variable accuracy of estimation of population distribution within a single set. This simply means that the accuracy of the assignment of the population counts to a reference units should differ significantly. To investigate the accuracy at the level of mapping unit reliable data on the number of people living (or staying) in the buildings are required. For coarse scale the national gridded population could be used, for fine scale the appropriate data could be derived from mobile phones.

5. Conclusions

The elaborated approach is spatial disaggregation. It is based on the assumption that population totals, aggregated to a commune (handled as the global parameter), can be redistributed within the commune by means of local parameters, namely building characteristics. Other words it is a multi-variable dasymetric modeling, where building type (e.g.: multi-residential, single flat, detached) is treated as limiting variable, and building volume, expressed by number of storey as well as area of building footprint – as connection variable. The overall quality of elaborated gridded population distribution surface at the scale 1 km^2 expressed by the mean of RMSE equals 9 persons, and the MAE equals 0.01.

The appropriateness and quality of the ancillary data used influence the accuracy and level of detail of population distribution techniques and algorithms. One of the advantages with the approach put forward in this paper, is that the data set (building data layer from Topographic Objects Data Base) used has the capacity to provide information about the characteristics of the population distribution in a fine scale. Thus, the elaborated population surface could be used in risk exposure and risk evacuation or mitigation plans.

Acknowledgment

The research was conducted under statutory research No 933/2016 at the Institute of Geodesy, Faculty of Civil Engineering and Geodesy, Military University of Technology. People redistribution was calculated by Katarzyna Zdunkiewicz within the MSc thesis "Population modelling using Topographic Objects Database" at the Faculty of Civil Engineering and Geodesy, Military University of Technology, under the supervision of prof. Elzbieta Bielecka.

References

Anderson, W., Guikema, S., Zaitchik, B and Pan, W. (2014). Methods for Estimating Population Density in Data-Limited Areas: Evaluating Regression and Tree-Based Models in Peru. *PLoS ONE* 9(7): e100037, doi:10.1371/journal.pone.0100037.

Azar, D., Graesser, J., Engstrom, R., Comenetz, J., Leddy JR, R.M., Schechtman, N.G. and Andrews, T. (2010). Spatial refinement of census population distribution using remotely sensed estimates of impervious surfaces in Haiti. *International Journal of Remote Sensing*, 31 (21): 5635–5655, doi:10.1080/01431161.2010.496799.

Azar D, Engstrom R, Graesser J, Comenetz J (2013) Generation of fine-scale population layers using multi-resolution satellite imagery and geospatial data. *Remote Sens Environ,* 130:219–232, doi: doi:10.1016/j.rse.2012.11.022

Bakillah, M.,. Liang, S, Mobasheri, A., Arsanjani J.J. and Zipf A. (2014). Fine-resolution population mapping using OpenStreetMap points-of-interest, *International Journal of Geographical Information Science*, vol. 28(9):1940-1963, doi: 10.1080/13658816.2014.909045.

Bielecka, E. (2015). Geographical data sets fitness of use evaluation. *Geodetski vestnik* Vol. 59 (2015), No. 2:335–348, DOI: 10.15292/geodetski-vestnik.2015.02.335-348.

Bielecka, E. (2005). A dasymetric population density map of Poland, Proceedings of the 22nd International Cartographic Conference, 9–15 July 2005, A Coruña, Spain (CD).

Bielecka, E. and Bober, A. (2013). Reliability analysis of interpolation methods in travel time maps-the case of Warsaw. *Geodetski vestnik*, 57.2 (Jun 2013), 299–312.

Bielecka, E., Leszczynska, M. and Hall, P. (2014). User perspective on geospatial data quality. Case study of the Polish Topographic Database. The 9th International Conference "ENVIRONMENTAL ENGINEERING" 22–23 May 2014, Vilnius, Lithuania, selected papers, eISSN 2029-7092 / eISBN 978-609-457-640-9. Available online at http://enviro.vgtu.lt, doi: http://dx.doi.org/10.3846/enviro.2014.193.

Ciołkosz, A. and Bielecka, E. (2005). Land cover in Poland. CORINE Land Cover databases. *Biblioteka Monitoringu* Środowiska, Warsaw, pp. 76.

de Smith, M.J. Goodchild, M.F. and Longley, P.A. (2015). *Geospatial Analysis. A Comprehensive Guide to Principles, Techniques and Software Tools*. Fifth Edition, Issue version: 1 (2015).

de Smith, M., Longley, P. and P. Goodchild (2016). *Geospatial Analysis book online* – web version, http://www.spatialanalysisonline.com/

Deville, P., Linard, C., Martin, S., Gilbert, M., Stevens, F.R., Gaughan, A.E., Blondel, V.D. and Tatem, A.J. (2014). Dynamic population mapping using mobile phone data. *Proc. Natl. Acad. Sci.* 2014, *111*:15888–15893.

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

Gallego, J. (2010). A population density grid of the European Union. *Population and Environment,* 31:460–473, doi: 10.1007/s11111-010-0108-y.

GUS. (2015a). *Gospodarka mieszkaniowa w 2014 r. Informacje i opracowania statystyczne*. Główny Urząd Statystyczny, Warszawa, avaiable on-line: http://stat.gov.pl/files/gfx/portalinformacyjny/pl/defaultaktualnosci/5492/7/10/1/gospodarka_mieszkaniowa_2014.pdf

GUS. (2015b). *Powierzchnia i ludność w przekroju terytorialnym w 2015 r*. [Area and population in the territorial profile in 2015], Główny Urząd Statystyczny, Warszawa

Hazewinkel, M. ed. (2001), "Pareto distribution", *Encyclopedia of Mathematics*, Springer, ISBN 978-1-55608-010-4

Horanont, T. and Shibasaki, R. (2010). Estimate ambient population density: discovering the current flow of the city. Available on line: https://www.academia.edu/2004297/ESTIMATE_AMBIENT_POPULATION_DENSITY_DISCOVERING_THE_CURRENT_FLOW_OF_THE_CITY

Horanont, T., Phithakkitnukoon, S. and Shibasaki, R. (2015). Sensing Urban Density Using Mobile Phone GPS Locations: A Case Study of Odaiba Area, Japan. *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, vol. 144:146–155, 10.1007/978-3-319-15392-6_15

INSPIRE. (2010). D2.8.I.2 INSPIRE Specification on Geographical Grid Systems – Guidelines. http://inspire.ec.europa.eu/documents/Data_Specifications/INSPIRE_Specification_GGS_v3.0.1.pdf

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, doi:10.1559/152304007781002190.

Marczak, S. (2015). Assessment of society involvement in creation of spatial data in Poland on the example of OpenStreetMap project [in Polish: Ocena zaangażowania społeczeństwa w tworzenie danych przestrzennych w Polsce na przykładzie projektu OpenStreetMap]. *Annals of Gematics*, t. XIII, 3(69), 239–253.

Medyńska-Gulij, B. (2015). *Kartografia. Zasady i zastosowania geowizualizacji*, PWN Warszawa, ISBN: 9788301183288.

Medyńska-Gulj, B. (2014). Cartographic sign as a core of multimedia map prepared by non-cartographers in free map services. *Geodesy and Cartography*. Vol. 63:55–64, doi: 10.2478/geocart-2014-0004.

Mennis, J. and Hultgren, T. (2006). Intelligent Dasymetric Mapping and Its Application to Areal Interpolation. *Cartography and Geographic Information Science*, 33:179–194.

Mennis, J. (2003). Generating Surface Models of Population Using Dasymetric Mapping. *The Professional Geographer.*, 55:1–42.

Mennis, J. (2009). Dasymetric mapping for estimating population in small areas. *Geography Compass,* 3:727–745.

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

Murshed, S. M. (2009). Disaggregation of regional population data for residential hot water demand assessment, Proceedings of 24th ICC/ICA cartographic Conference, Santiago de Chile, Santiago de Chile, Chile, 15–21 November 2009.

Nowak Da Costa, J. (2016a). Novel tool to examine data completeness based on comparative studof VGI data and official building datasets. *Geodetskij vestnik*, 60(3): 495–508. DOI: 10.15292/geodetski-vestnik.2016.03.495-508)

Nowak Da Costa, J. (2016b). Towards building data semantic similarity analysis: OpenStreetMap and the Polish Database of Topographic Objects. 2016 Baltic Geodetic Congress (BGC Geomatics), Gdansk, 2016, pp. 269–275. DOI: 10.1109/BGC.Geomatics.2016.55

Langford, M. (2013). An Evaluation of Small Area Population Estimation Techniques Using Open Access Ancillary Data. *Geographical Analysis*, 45:324–344.

Olszewski, R. and Gotlib. D. (red.). (2013). *Rola bazy danych obiektów topograficznych w tworzeniu infrastruktury informacji przestrzennej w Polsce*. Joint publication, GUGiK, ISBN 978-83-254-1975-2.

Regulation. (2011). Regulation of Ministry of Internal Affairs and Administration of 11 November 2011 as regards *topographic objects database and database of general topographic objects as well as standard cartographic products*. Dz. U. 2011 No. 279, entry 1642, Poland, 2011.

Silva, M. and Pereira, S. (2014). Assessment of physical vulnerability and potential losses of buildings due to shallow slides, *NatHazards*, 72(2), 1029–1050, doi 10.1007/s11069-014-1052-4, 2014.

Sridharan, H. and Qiu, F. (2013). A Spatially Disaggregated Areal Interpolation Model Using Light Detection and Ranging-Derived Building Volumes. *Geographical Analysis*, 45:238–258

Smith, A., Newing, A., Quinn, N., Martin, D.; Cockings, S. and Neal, J. (2015). Assessing the Impact of Seasonal Population Fluctuation on Regional Flood Risk Management. *ISPRS Int. J. Geo-Inf.* 2015, 4:1118–1141, DOI:10.3390/ijgi4031118.

Stevens, F.R., Gaughan A., E. and Tatem, A. J. (2015). Disaggregation census data for population mapping using Random Forest with remotely-sensed and ancillary data. *PLoS One*, 2015, 10(2), e0107042.

Su, M.D., Lin, M.C., Hsieh, H.I., Tsai, B.W. and Lin, C.H. (2010). Multi-layer multi-class dasymetric mapping to estimate population distribution. *Science of the Total Environment*, 408 (20):4807–4816, doi:10.1016/j.scitotenv.2010.06.032

Tobler, W. R., Deichmann, U, Gottsgen J. and Maloy, K. (1995). The global demography project. National Center for Geographic Information and Analysis (NCGIA), University of California; Santa Barbara, California: 1995. Technical Report TR-6-95.

Wright, J K. (1936). A Method of Mapping Densities of Population: With Cape Cod as an Example. *Geographical Review*, 26:104–15.

Wu, S.-S., Qiu, X. and Wang, L. (2005). Population Estimation Methods in GIS and Remote Sensing : a review. *GIScience and Remote sensing,* 42 (1):58–74; doi: 10.2747/1548-1603.42.1.80.

Wu, S-S., Wang, L., and Qiu, X. (2008). Incorporating GIS building data and census housing statistics for subblock-level population estimation. *The Professional Geographer*, 60 (1):121–135, doi: 10.1080/00330120701724251

Wu, C. and Murray, A.T. (2005). A cokriging method for estimating population density in urban areas. *Computers, Environment and Urban Systems*, 29 (5), 558–579, doi:10.1016/j.compenvurbsys.2005.01.006

Zandbergen, P. and Ignizio, D. (2010). Comparison of dasymetric mapping techniques for small-area population estimates. *Cartography and Geographic Information Science*, 37 (3):199–214, doi:10.1559/152304010792194985.