

## Dasymetric mapping for public health planning

Patricia Steiner & Gernot Paulus

School of Geoinformation, Carinthia Tech Institute, Europastrasse 4, A-9524 Villach, Austria

### ABSTRACT

This paper discusses a dasymetric mapping-based approach to estimate the population density as an important decision criteria in public health planning. Dasymetric mapping is a method to disaggregate census data to finer scales by integrating population data and ancillary data providing a more realistic impression about the population distribution than using arbitrary administrative boundaries such as census tracts.

The application of Geographic Information Systems (GIS) for solving public health related questions can analyze various variables that are connected by a common spatial denominator. Such variables can be public health data, demographic data, social data, economic data, as well as environmental data. For representing health related and demographic data, conventionally thematic visualization techniques like choroplethe maps are used. One special thematic representation method is dasymetric mapping. Dasymetric maps integrate settlement patterns and enable the same representation as choropleth maps but on a more detailed spatial reference unit. This paper compares two different approaches for creating dasymetric maps: the "limiting variable method" and the "variable class division method". The objective of the two methods is to investigate the population density delivering a more accurate spatial distribution and therefore a better key variable for decision support concerning public health related planning questions.

*Keywords: dasymetric mapping, population density, population distribution, data disaggregation, public health*

### INTRODUCTION

The usage of GIS as working and visualization tool for spatial information, to answer health related questions has been growing for the last years. The potential of GIS for public health applications is rising, due to the fact that more spatial data like health data, demographic data, ecological and environmental data as well as social data are available and widely used in this application field.

When it comes to demographic data, the problem is, that they are mostly only available for big enumeration units at a coarse level. This makes it impossible to derive more detailed information for smaller planning units (Weichselbaum *et al.* 2005).

When the population density, or other socio-economic data is represented with choroplethe maps, it can happen, that the representation of the size and the shape of the represented area gets more attention than the population itself. Furthermore, the user gets the impression that the population within the area is distributed regularly. The same is true for unpopulated regions (Mennis 2003).

To avoid these problems, the data are disaggregated onto smaller units. This process is known as "dasymetric mapping", and provides the advantage that decisions for planning questions, analysis, queries or other purposes can be done on a more detailed spatial distribution.

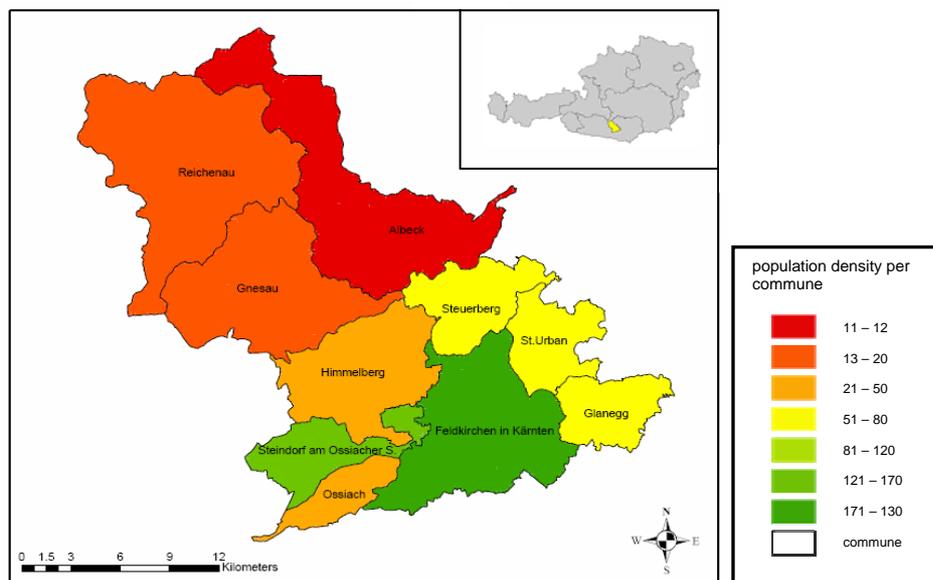
### DASYMETRIC MAPPING APPROACHES

This paper compares two methods for creating dasymetric maps. The first method is the "limiting variable method" which excludes uninhabited areas and calculates the population density only for the inhabited areas. The second method is named "variable class division" and divides inhabited regions

into varying settlement classes, which can range from “low populated” to “high populated”. For every population class the population density is computed separately.

**Data**

Input data for the data disaggregation are grid cell data as well as municipality and district administrative boundaries. The methods are applied for the project area of the district Feldkirchen in Carinthia, located in Southern Austria, which consists of ten municipalities and has an overall population of 30273 inhabitants. The micro grid data represent a special new enumeration unit of the Statistic Austria, which are built up by polygons with a size of 250 x 250 meters. The data includes information about the number of accommodations, the number of capital residences as well as the number of accommodations that are no capital residences. Figure 1 gives an impression of a classical representation of population density based on municipality level.



**Figure 1:** Representation of the population density for every municipality within the district Feldkirchen.

**"Limiting variable method"**

The "limiting variable method" was created by Wright (1936) and is used for generating dasymetric maps when the appearance of the population can be excluded from the outset. This can for instance be valid for a glacier or a highly forested area, for which the value zero is assigned. In the case of the grid cell data all uninhabited grid cells are assigned a value of 0. For calculating the population density of the inhabited areas, following formula (1) is used (Slocum *et al.* 2004):

$$D_N = \frac{D - (D_M \times A_M)}{(1 - A_M)} \tag{1}$$

- $D_N$  = density of region  $N$
- $D$  = overall density of the region
- $D_M$  = density of sub region  $M$
- $A_M$  = fractional area of region  $M$

For this calculation the information about the number of accommodations is relevant. The limiting variable in this case is the number of grid cells without accommodations. The computation is implemented for all ten communities separately within the district Feldkirchen. The population density of the inhabited regions is multiplied with the portion of the inhabited area of the grid cells on the total area of the municipality and subtracted from the population density of the whole community. The result is divided by 1 minus the portion of the inhabited area of the grid cells on the total area of the community. This value is represented with a value between 0 and 1. For the demonstration of the calculation, the municipality Albeck is considered:

$$D_N = \frac{11 - (0 \times 0.115)}{(1 - 0.115)}$$

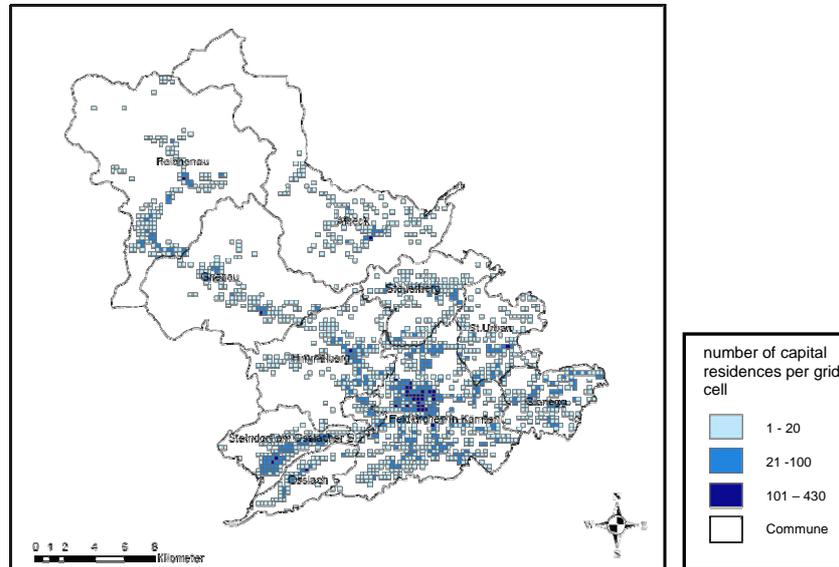
The population density for the whole municipality Albeck mounts up to 11 inhabitants per km<sup>2</sup>. The population density outside of the grid cells is zero. The portion of the populated area on the total area of the commune is 0.115. This results to a new population density of 12.43 inhabitants per km<sup>2</sup>, which stands for an increase of 1 inhabitant per km<sup>2</sup>.

#### **"Variable class division method"**

The second method is the "variable class division method" created by Mennis and is based on the method of Langfort and Unwin (1994) (Mennis 2002).

The "variable class division method" calculates the population density within different settlement classes where the classes reach from high populated to low populated. The basis of the population data is an enumeration unit. As ancillary information, land use data is used. The population density is calculated for the classes of the land use data and consists of four formulas (Mennis 2003).

In this case, micro grid data is used instead of land use data. The "variable class division method" does not only differentiate between "inhabited" or "not inhabited" areas but also between "high", "middle" and "low" inhabited areas. There is a fourth class too, but this class only includes non inhabited polygons that have a population density value of zero. For that reason the fourth class is excluded for all calculations. This results to a smaller overall area of the test area. The division of the inhabited classes is carried out on the number of capital residences of the grid cell data. It is not possible to apply the number of accommodations for this circumstance, because the number of accommodations include all permanent residences as well as all secondary residences. This would falsify the result, because some inhabitants would be counted twice. The lowest settlement class goes from one capital residence to 20 capital residences. The class "middle" ranges from 21 capital residences to 100 capital residences. The "high" inhabited class includes 101 to 430 capital residences (as illustrated in figure 2).



**Figure 2:** Representation of the three settlement classes and their number of capital residences per grid cell for the district Feldkirchen.

The “variable class division method” is applied by using following formulas (2)-(5):

*Population density fraction*

The population density fraction (2) stands for the portion of the population density on the whole population density for every settlement class.

$$d_{uc} = \frac{p_{uc}}{(p_{hc} + p_{lc} + p_{nc})} \quad (2)$$

$d_{uc}$  = population density fraction of urbanization class  $u$  in county  $c$

$p_{uc}$  = population density of urbanization class  $u$  in county  $c$

$p_{hc}$  = population density of urbanization class  $h$  (high) in county  $c$

$p_{lc}$  = population density of urbanization class  $l$  (low) in county  $c$

$p_{nc}$  = population density of urbanization class  $n$  (non urban) in county  $c$

The population density fraction of the several settlement classes for the district Feldkirchen is calculated as follows:

$$pdf_{uFK} = \frac{PD_{uFK}}{PD_{IFK} + PD_{mFK} + PD_{hFK}} \quad pdf_{IFK} = \frac{PD_{IFK}}{total\_density}$$

$$pdf_{IFK} = \frac{113.39752}{3456.2408} = 0.0328 = 3.3\%$$

The population density fraction (*pdf*) for the population class low (*l*) of the district Feldkirchen (*FK*) is calculated by dividing the population density of the low populated class through the sum of the population density of all three settlement classes. The calculation is done similarly for the classes “middle” and “high”. The population density for every urbanization class and the total density of the district Feldkirchen is represented in table 1.

**Table 1:** Population density fraction for the district Feldkirchen.

District	Settlement class	Residents	Area (km <sup>2</sup> )	Population density (R/km <sup>2</sup> )	Total density	Pdf
Feldkirchen	l = low	10305	90.875	113.39752	3456.2408	3.3
Feldkirchen	m = middle	15308	24.0625	636.17662	3456.2408	18.4
Feldkirchen	h = high	6090	2.25	2706.6667	3456.2408	78.3

#### Area ratio

The next step is to calculate the area ratio (3). It is necessary to declare, how high the appearance of one class, on the whole area of the region is. It is assumed that all three settlement classes are distributed steady, so 0.33 is assigned as the percent value. This is done to ease the calculation.

$$a_{ub} = \frac{\left( \frac{n_{ub}}{n_b} \right)}{0.33} \quad (3)$$

$a_{ub}$  = area ratio of urbanization class *u* in block group *b*

$n_{ub}$  = number of grid cells of urbanization class *u* in block group *b*

$n_b$  = number of grid cells in block group *b*

The number of grid cells of one settlement class (for instance class middle) within a community (for instance Feldkirchen which is the capital city of the district Feldkirchen) is divided by the number of grid cells which build up the community Feldkirchen. The result is divided by the value 0.33. Following area ratios are derived for the community Feldkirchen:

$$A_{lF} = 2$$

$$A_{mF} = 0.9$$

$$A_{hF} = 0.1$$

The total number of the three area ratios is three and represents the number of classes.

#### Total fraction

The total fraction (4) calculates the entire fraction of each settlement class, within one block group, inside one region.

$$f_{ubc} = \frac{(d_{uc} \times a_{ub})}{[(d_{hc} \times a_{hb}) + (d_{lc} \times a_{lb}) + (d_{nc} \times a_{nb})]} \quad (4)$$

$f_{ubc}$  = total fraction of urbanization class *u* in block group *b* and in county *c*

$d_{uc}$  = population density fraction of urbanization class *u* in county *c*

$a_{ub}$  = area ratio of urbanization class *u* in block group *b*

$d_{hc}$  = population density fraction of urbanization class  $h$  in county  $c$   
 $a_{hb}$  = area ratio of urbanization class  $h$  in block group  $b$   
 $d_l$  = population density fraction of urbanization class  $l$  in county  $c$   
 $a_{lb}$  = area ratio of urbanization class  $l$  in block group  $b$   
 $d_{nc}$  = population density fraction of urbanization class  $n$  in county  $c$   
 $a_{nb}$  = area ratio of urbanization class  $n$  in block group  $b$

The total fraction of each settlement class inside the municipality Feldkirchen and within the district Feldkirchen is determined as follows:

$$f_{uFFK} = \frac{(pdf_{uFK} \times a_{uF})}{[(pdf_{iFK} \times a_{iF}) + (pdf_{mFK} \times a_{mF}) + (pdf_{hFK} \times a_{hF})]}$$

To determine the total fraction of the settlement class low, the population density fraction of the settlement class low is multiplied by the area ratio of this settlement class. The result is divided by the sum of the population density fraction of every settlement class multiplied with the according area ratio.

$$\begin{aligned}
 f_{iFFK} &= 0.2 \\
 f_{mFFK} &= 0.5 \\
 f_{hFFK} &= 0.3
 \end{aligned}$$

The entire sum of all fractions is one.

#### Total population

The last step is to calculate the population within the grid cells of every settlement class (5):

$$pop_{ubc} = \frac{(f_{ubc} \times pop_b)}{n_{ub}} \quad (5)$$

$pop_{ubc}$  = population assigned to one grid cell of urbanization class  $u$  in block group  $b$  and in county  $c$

$f_{ubc}$  = total fraction of urbanization class  $u$  in block group  $b$  and in county  $c$

$pop_b$  = population of block group  $b$

$n_{ub}$  = number of grid cells of urbanization class  $u$  in block group  $b$

The total fraction of an urbanization class within one community is multiplied with the amount of the population within that community and divided by the number of grid cells of the specific settlement class. For the commune Feldkirchen following calculation is generated.

$$pop_{uFFK} = \frac{(f_{uFFK} \times pop_F)}{n_{uF}}$$

$$\begin{aligned}
 pop_{iFFK} &= 7.8 \\
 pop_{mFFK} &= 46.5 \\
 pop_{hFFK} &= 162.3
 \end{aligned}$$

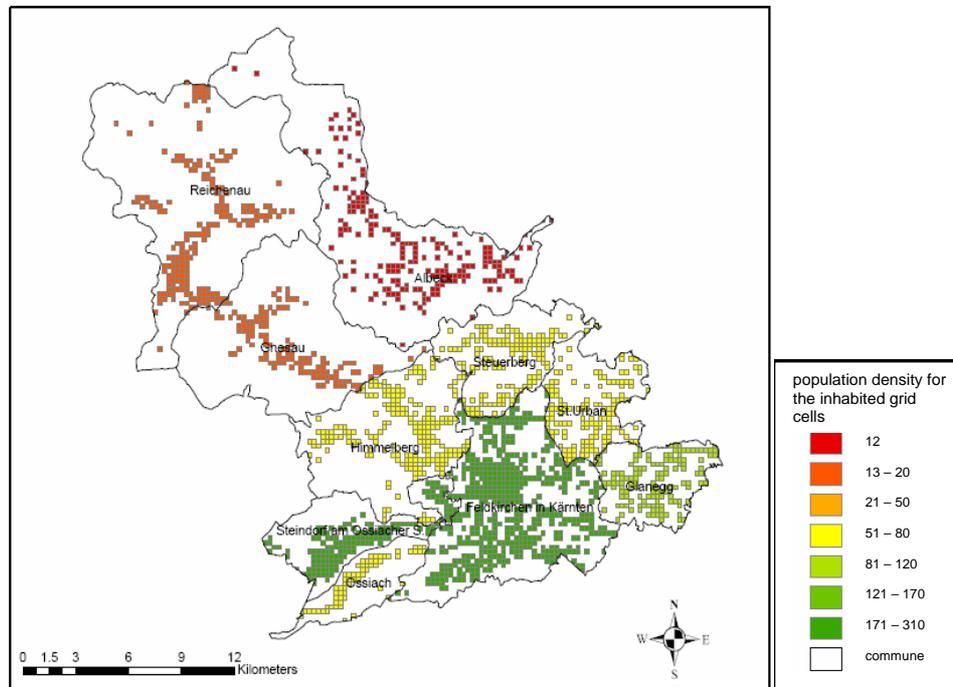
**RESULTS**

The results of the “limiting variable method” are represented in Table 2. The table includes the population density for the whole commune as well as the new population density for the inhabited grid cells.

**Table 2:** Population density for the whole commune and population density within the inhabited regions of the district Feldkirchen.

<b>Commune</b>	<b>population density for whole community</b>	<b>New population density within inhabited regions</b>
Albeck	11	12
Feldkirchen	181	310
Gnesau	16	17
Glanegg	80	119
Himmelberg	42	54
Ossiach	43	53
Reichenau	18	20
St.Urban	53	77
Steindorf	121	173
Steuerberg	52	70

When the population density of the whole municipality is low, the new calculated population density for the inhabited regions within every municipality is low, too. For example, the population density of Albeck is 11 residents per km<sup>2</sup>. The new population density adds up to 12 residents per km<sup>2</sup>. This is a change of one person. The population density for the commune Feldkirchen is 181 persons per km<sup>2</sup>. The new calculated population density amounts to 310 residents per km<sup>2</sup>. This is a change of 129 inhabitants per km<sup>2</sup>. Figure 3 represents the new population density inside the inhabited areas using the dasymetric map approach. A low population density is represented with a red hue; a high population density is symbolized by greenish colors. The commune Albeck has the lowest population density. Feldkirchen, Steindorf am Ossiacher See und Glanegg have the highest population density within the populated grid cells.



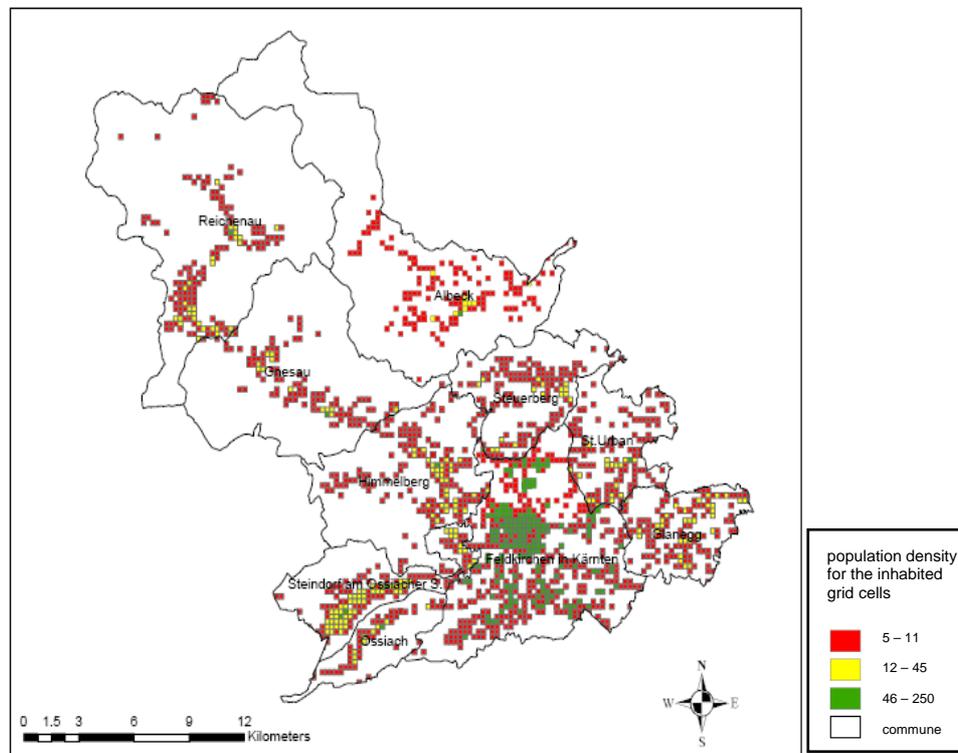
**Figure 3:** Population density within the populated grid cells inside the district Feldkirchen.

The results of the “variable class division calculation” are shown in Table 3, which includes the total population within the three settlement classes for the whole district Feldkirchen.

**Table 3:** Total population within every settlement class inside the district Feldkirchen.

Municipality	Pop <sub>l</sub>	Pop <sub>m</sub>	Pop <sub>h</sub>
Albeck	6.2	26.8	120.6
Feldkirchen	7.8	46.5	162.3
Glanegg	10.7	25.3	0
Gnesau	6.5	34.6	138.3
Himmelberg	5.8	38.2	244.4
Ossiach	6	34.8	156.6
Reichenau	6.8	38.3	83.7
St. Urban	6.2	30.5	160.1
Steindorf am Ossiacher See	7.6	42.7	121.9
Steuerberg	7.7	39.4	0

The population density of the settlement class "high" in Glanegg and Steuerberg is down to zero because in these communities no grid cells exist with more than 100 permanent residences. The resulting values in table 3 are represented in form of a dasymetric map as illustrated in figure 4. The value for the class “low” ranges from 5 to 11 inhabitants per grid square. The class “middle” goes up from 12 to 45 inhabitants and the highest class ranges from 46 to 245 inhabitants. Figure 4 shows the population density of the three settlement classes for the district Feldkirchen. The population density of the lowest class is represented by a red hue, the class “middle” is symbolized by the color yellow and green represents the settlement class “high”.



*Figure 4:* Population density for the three settlement classes within the district Feldkirchen.

## CONCLUSION

Dasymetric mapping is a method to disaggregate data on a smaller level by integrating ancillary data. By comparing the "limiting variable method" and the "variable-class-division method" the following conclusions, advantages and disadvantages can be shown:

The "limiting variable method" calculates the population density for the populated areas. An advantage is that non-populated areas are totally excluded. Those areas represent the limited variable, which enforces a fast and simple calculation of the population density. A second advantage is that for the calculation only a low portion of information is required, videlicet the old population density and the fraction of the populated area. The "limiting variable method" determines the population density within the populated regions, but it doesn't divide those regions into sub regions.

The "variable-class-division method" enables the calculation of multiple settlement classes. The choice of the number of classes is optional to the user. The calculation gets more complex the more settlement classes are integrated. The calculation itself is more time-consuming than adopting the "limiting variable method". An advantage is that the non-populated areas are also excluded. The application can be accomplished by using diverse data instead of population data or land use data.

The decision, which of the two methods is more appropriate depends on the focus under investigation as well as on the desired data accuracy. In regions that are densely populated, for instances in large cities, it is obvious to apply the "variable class division method" because of the

advantage of a varying class number. The more densely a region is populated the more interesting is it to use a larger number of settlement classes to get a more accurate representation.

The resulting maps for the project area provide concise information about where the population is settling as well as how the population is distributed within the inhabited areas. For this reason dasymetric mapping delivers a basis for further, more accurate analysis and planning tasks, especially in the public health sector, where exact knowledge of the spatial distribution of socio-demographic data distribution are important decision criteria. Unfortunately, many of these data are often only available for coarse spatial administrative units and their acquisition is more expensive the spatially finer the socio-demographic data acquisition has to be performed. Dasymetric mapping circumvents these facts by disaggregating coarser data to finer scales. This approach can for instance be used for planning a nursing home for elder people in individual city districts by using the age distribution at total city level and the number of inhabitants for the single city districts. Using dasymetric mapping the age distribution is disaggregated from the city level to the district level resulting in a more detailed spatial information about the age distribution within the city area and therefore providing better a basis for planning such a nursing home.

#### **ACKNOWLEDGEMENTS**

This diploma thesis represents a collaboration project between the School of Geoinformation and the School of Health Care Planning at Carinthia Tech Institute. The discussions and contribution of Herbert Janig and Holger Penz are gratefully acknowledged.

#### **BIBLIOGRAPHY**

- Eicher, C. and Brewer, C. 2001, Dasymetric mapping and areal interpolation: implementation and evaluation, *Cartography and Geographic Information Science*, Vol. 28, pp. 125-138.
- Langford, M. and Unwin, D., 1994, Generating and mapping population density surface within a geographical information system, *The Cartographic Journal*, Vol. 31, pp. 21-26.
- Mennis, J., 2002 Using Geographic Information Systems to Create and Analyze Statistical Surfaces of Population and Risk for Environmental Justice Analysis, *Social Science Quarterly*, Vol. 83, No. 1, pp. 281 – 297.
- Mennis, J., 2003, Generating Surface Models of Population Using Dasymetric Mapping, *The Professional Geographer*, Vol. 55, No.1, pp. 31 – 42.
- Slocum, T., McMaster, R., Kessler, F. and Howard, H., 2004, Dot and Dasymetric Mapping, in: *Thematic Cartography and Geographic Visualization*, editor: Clarke, K.C., pp. 328 – 340, Pearson/Prentice Hall.
- Weichselbaum, J., Petrini-Monteferrri, F., Papathoma, M., Wagner, W. and Hackner, N., 2005, Sharpening census information in GIS to meet real-world conditions: the case for Earth Observation, *WIT Transactions on Ecology and the Environment*, Vol. 84, pp. 143 – 154.
- Wright, J.K, 1936, A Method of mapping densities of population with Cape Cod as an example, *Geographical Review*, Vol. 26, pp:103-110.