

MODELOWANIE POWIERZCHNI STATYSTYCZNEJ PRZEDSTAWIAJĄCEJ GĘSTOŚĆ ZALUDNIENIA W POLSCE PRZY POMOCY METODY DAZYMetrycznej

MODELING STATISTICAL SURFACE OF POPULATION DENSITY IN POLAND USING DASYMETRIC MAPPING

Elżbieta Bielecka, Anna Kuczyk, Ewa Witkowska

Instytut Geodezji i Kartografii

Słowa kluczowe: metoda dazymetryczna, kartogram, interpolacja powierzchniowa, dane statystyczne, gęstość zaludnienia

Keywords: dasymetric method, choropleth map, areal interpolation, census data, population density

Wprowadzenie

Dokładność i poprawność map gęstości zaludnienia, jak każdych map wskaźnikowych, zależy od właściwego doboru pól odniesienia, sposobu transformacji danych oraz zastosowanej metody prezentacji. Najczęściej gęstość zaludnienia jest przedstawiana kartogramicznie, metodą izolinii lub metodą dazymetryczną. Do połowy lat dziewięćdziesiątych metoda dazymetryczna była stosowana niezwykle rzadko. Z badań przeprowadzonych przez Golenia i Ostrowskiego (1994, 1996) wynika, że zarówno w Polsce jak i na świecie kartogramy dazymetryczne były wypierane przez inne metody prezentacji danych ilościowych. W dobie komputeryzacji kartografii i jej synergii z GIS sytuacja ta uległa zmianie.

Metoda dazymetryczna (grec. *dasy* – gęsty, mocny; *metron* – miara) należy do prezentacji kartogramicznych przedstawiania danych statystycznych za pomocą znaków powierzchniowych. Tak jak w kartogramie właściwym zjawisko jest ujmowane relatywnie, dane są pogrupowane w przedziały klasowe i odniesione do pól odniesienia. Podstawowa różnica pomiędzy kartogramem dazymetrycznym a kartogramem właściwym wynika ze sposobu przyjęcia pól odniesienia. W kartogramie właściwym stanowią je arbitralnie określone jednostki, w których zbierano dane (obwody spisowe) lub do których są odnoszone dane (jednostki podziału administracyjnego), podczas gdy w kartogramie dazymetrycznym układ pól odniesienia wynika ze zmienności zjawiska. Według Ratajskiego (1989) metoda dazymetryczna może być stosowana jako kompromis na rzecz bardziej geograficznego przedstawiania powierzchni statystycznej.

Brak jednoznacznych zasad (subiektywizm) delimitacji jednostek odniesienia oraz nieustalone zasady przeliczania danych z jednostek wejściowych do jednostek dazymetrycznych, a także duży nakład pracy związany z opracowaniem map dazymetrycznych uważane są za największe wady tej metody prezentacji kartograficznej i znacznie ograniczają jej szerokie stosowanie.

Metodyka opracowywania map dazymetrycznych

W piśmiennictwie kartograficznym podawane są sposoby wykonania kartogramu dazymetrycznego polegające na transformacji jednego rodzaju mapy na inny. Jako mapę wyjściową wykorzystuje się zwykle mapę kropkową lub kartogram prosty. Ratajski (1989) dodatkowo wymienia mapy przedstawiające dane w sieci zmiennogęstej nieregularnej, które podlegają podobnej transformacji jak mapy kropkowe. Transformacja jednego rodzaju mapy na inny prowadzi zawsze do straty informacji, ponieważ każde przetworzenie oddala obraz zjawiska od danych wyjściowych. Dlatego też dziś, w dobie dominacji technik kartografii komputerowej i GIS, niezwykle rzadko opracowuje się kartogram dazymetryczny poprzez transformację mapy kropkowej, sieci zmiennogęstej, czy też kartogramu prostego. Zazwyczaj wykorzystuje się do tego celu systemy informacji geograficznej oraz bazy danych geograficznych pozwalające na wyznaczenie jednostek odniesienia oraz oszacowanie, w ich granicach, wartości analizowanego zjawiska.

Do opracowania map dazymetrycznych Robinson (1988) proponuje wykorzystanie dwóch rodzajów zmiennych: zmiennych ograniczających i zmiennych powiązań. Zmienne ograniczające ustalają absolutną granicę wartości przedstawianego zjawiska mogącego pojawić się na danym obszarze. Za zmienne powiązań przyjmuje się te zjawiska geograficzne, które wykazują powiązania przestrzenne ze zjawiskiem przedstawionym na mapie i nie były wykorzystane jako zmienne ograniczające. W przypadku map gęstości zaludnienia najczęściej za zmienną ograniczającą przyjmuje się użytkowanie ziemi, natomiast za zmienną powiązań rzeźbę terenu, gleby, warunki fizyczno-geograficzne lub inne. Wybór zmiennych powiązań i sposób ich wykorzystania należą do kluczowych zagadnień rozwiązywanych podczas opracowania map dazymetrycznych.

Kolejnym nie mniej istotnym zagadnieniem jest opracowanie metody przeliczania danych statystycznych, odniesionych do podziału administracyjnego (powszechnie udostępniane i publikowane są tylko dane w przekrojach gminnym, powiatowym lub wojewódzkim) do nowo-wyznaczonych jednostek odzwierciedlających natężenie prezentowanego zjawiska. W trakcie przeliczeń musi być zachowany warunek zgodności względem bazowych danych statystycznych gwarantujący poprawność każdej mapy statystycznej.

Następnym ważnym problemem jest ustalenie liczby przedziałów klasowych i ich granic oraz wybór skali barwnej. Warunkiem prawidłowego podziału danych na klasy jest minimalizowanie różnic wartości w obrębie klas przy jednoczesnym maksymalizowaniu różnic między klasami. Spełnienie tego warunku oznacza, że różnice między powierzchnią statystyczną zbudowaną na podstawie danych wyjściowych i mapą opartą na pogrupowanych danych będą możliwie najmniejsze, a granice klas widoczne na mapie będą zgodne z rzeczywistym rozkładem zjawiska. Należy mieć na uwadze również, że w metodzie dazymetrycznej, należącej do metod ilościowych, ciemniejszy wizualnie odcień lub bardziej intensywna barwa powinien być przypisany wartościom większym, tak aby podkreślić hierarchię i porządek klas.

Metoda badań i uzyskane wyniki

Celem prezentowanych badań było sformalizowanie opracowywania dazymetrycznych map ludnościowych przez przyjęcie określonych wartości gęstości zaludnienia dla poszczególnych typów pokrycia terenu, otrzymywanych w wyniku klasyfikacji treści obrazów satelitarnych. Analizując pokrycie terenu wyodrębniono obszary możliwie jednorodne pod względem gęstości zaludnienia, które nazwano jednostkami dazymetrycznymi. Jednostkom dazymetrycznym przypisano taką liczbę ludności, aby po zsumowaniu przypisanych wartości liczba ludności zgadzała się w obrębie jednostki administracyjnej z danymi statystycznymi.

Informacja o pokryciu terenu pochodziła z bazy danych CORINE Land Cover 2000 (CLC2000), dane ludnościowe z publikacji GUS (2000).

W artykule przedstawiono trzy odmiany dazymetrycznej metody modelowania rozmieszczenia ludności: metodę binarną, powierzchniowo-wagową metodę agregacji oraz powierzchniowo-wagową metodę korelacji. Badania eksperymentalne zostały wykonane dla województwa mazowieckiego.

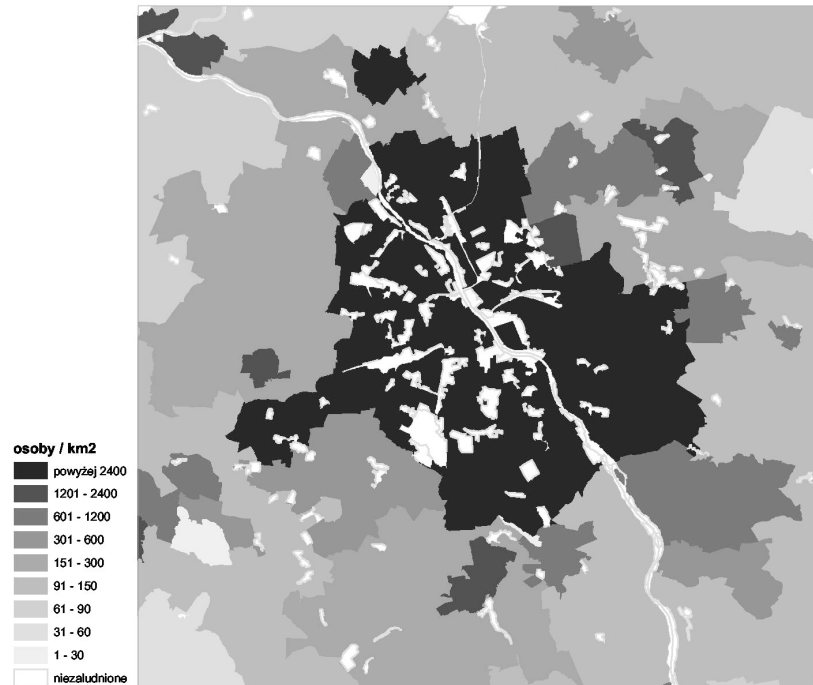
Metoda binarna

Metoda binarna należy do najprostszej odmiany metody dazymetrycznej. Bazując na danych o pokryciu terenu powierzchnię gminy dzieli się na dwie klasy zamieszkaną i niezamieszkaną, przypisując liczbę ludności tylko do części zamieszkaną. Pierwszym etapem opracowania map jest więc wybór tych klas pokrycia terenu, które są niezamieszkałe. Z trzydziestu jeden klas pokrycia terenu, wyznaczonych w ramach projektu CLC2000 (Bielecka i Ciołkosz 2005), aż osiemnaście to obszary niezaludnione. Należą do nich m.in.: rzeki, jeziora, bagna, torfowiska, plaże, odkryte skały, roślinność naturalna, miejskie tereny zielone, tereny sportowe, drogi i koleje, porty, lotniska, kopalnie odkrywkowe, budowy. Dazymetryczna mapa gęstości zaludnienia opracowana metodą binarną została przedstawiona na rysunku 1.

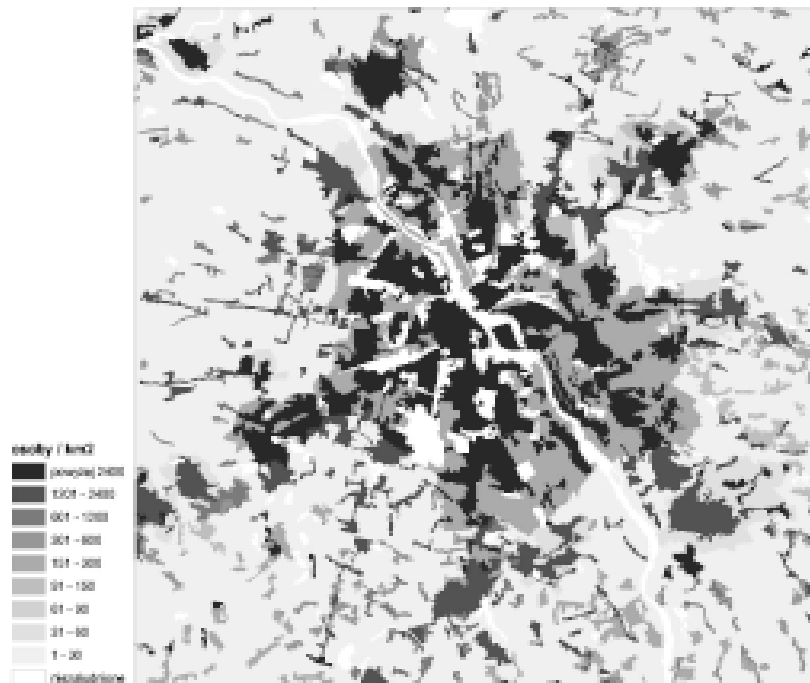
Prostota i zachowanie wolumenu danych w jednostce administracyjnej (gminie) należą do największych zalet tej odmiany metody dazymetrycznej. Subiektywizm metody jest w tym przypadku znacznie ograniczony, kartograf podejmuje tylko jedną decyzję dotyczącą wyboru obszarów niezamieszkałych.

Powierzchniowo-wagowa metoda agregacji

W powierzchniowo-wagowej metodzie agregacji dane o pokryciu terenu, traktowane jako zmienna ograniczająca, zostały zagregowane do czterech typów: (1) zabudowy zwartej, (2) zabudowy rozproszonej, (3) terenów słabo zaludnionych i (4) terenów niezamieszkałych. Jako drugą zmienną ograniczającą przyjęto, arbitralnie określony, procent ludności zamieszkującej cztery wymienione typy pokrycia terenu. Dodatkowo za zmienną powiązań przyjęto, określany przez GUS, typ gminy (miejska, miejsko-wiejska, wiejska). I tak założono, że 70% ludności w gminach miejskich i miejsko-wiejskich mieszka na terenie zabudowy zwartej, 28% – zabudowy rozproszonej i 2% na gruntach rolnych, terenach zajmowanych przez roślinność naturalną oraz w lasach przyjętych za tereny słabo zaludnione. W gminach wiejskich, gdzie nie ma zabudowy zwartej 98% ludności zostało przypisane do zabudowa rozproszonej, pozostałe 2% – do obszarów słabo zaludnionych. Dazymetryczna mapa gęstości zaludnienia opracowana powierzchniowo-wagową metodą agregacji metodą została przedstawiona na rysunku 2.



Rys. 1. Dazymetryczna mapa gęstości zaludnienia Warszawy i okolic; metoda binarna



Rys.2. Dazymetryczna mapa gęstości zaludnienia Warszawy i okolic; powierzchniowo-wagowa metoda agregacji

Powierzchnia statystyczna przedstawiająca gęstość zaludnienia utworzona z wykorzystaniem tej metody umożliwia bardziej zbliżone do rzeczywistego przestrzenne przedstawienie rozmieszczenia ludności niż w metodzie binarnej. Tak jak metoda binarna, zachowuje wolumen danych w jednostkach administracyjnych najniższego szczebla. Obarczona jest jednak większym subiektywizmem związanym z agregacją klas pokrycia terenu i przyjęciem procentowego podziału liczby ludności w zależności od typu gminy i rodzaju pokrycia terenu.

Powierzchniowo-wagowa metoda korelacji

Powierzchniowo-wagowa metoda korelacji została opracowana jako modyfikacja metody zastosowanej przez Gallego i Peedella (2001) do sporządzenia mapy gęstości zaludnienia włoskiej prowincji Arezzo. Metoda bazuje na założeniu, że stosunek gęstości zaludnienia dwu typów pokrycia terenu jest stały dla każdej gminy w województwie. Transformacja danych ludnościowych odniesionych do gmin do jednostek dazymetrycznych odbywa się zgodnie ze wzorem:

$$D_c = U_c \frac{P}{\sum_c A_c U_c}$$

gdzie:

D_c – gęstość zaludnienia typu c pokrycia terenu w województwie,

U_c – współczynnik wagowy (różnicujący gęstość zaludnienia pomiędzy różnymi typami pokrycia terenu i rodzajami gmin),

P – liczba ludności w województwie,

A_c – powierzchnia zajmowana przez typ c pokrycia terenu w województwie

Również w tej metodzie zmienną ograniczającą było pokrycie terenu, zagregowane tym razem do sześciu typów: (1) zabudowanych terenów miejskich, (2) terenów przemysłowych i handlowych, (3) terenów rolniczych, (4) osadnictwa rozproszonego, (5) lasów, roślinności naturalnej, (6) terenów niezamieszkałych (m.in. wody, bagna, torfowiska, plaże, nagie skały). Drugą zmienną powiązań był rodzaj gminy określany na podstawie: typu gminy GUS, liczby ludności, gęstość zaludnienia i charakter zabudowy w gminie. Wyróżniono osiem rodzajów gmin: I – miasta o liczbie ludności powyżej 100 tys., II – pozostałe gminy miejskie, III – część miejska gmin miejsko-wiejskich o gęstości zaludnienia powyżej 500 os/km², IV – część miejska gmin miejsko-wiejskich o gęstości zaludnienia poniżej 500 os/km², V – gminy wiejskie z zabudową o charakterze miejskim, VI – gminy wiejskie bez zabudowy o charakterze miejskim, VII – część wiejska gmin miejsko-wiejskich z zabudową o charakterze miejskim oraz VIII – część wiejska gmin miejsko-wiejskich bez zabudowy o charakterze miejskim.

W celu obliczenia wagi przypisanej do odpowiedniego typu pokrycia terenu w obrębie gmin danego rodzaju obliczono współczynniki korelacji ρ_{cr} pomiędzy stosunkiem liczby ludności przypisanej do każdej gminy P_m^0 a liczbą ludności podaną przez GUS P_m oraz stosunkiem powierzchni zajmowanej przez poszczególny typ pokrycia terenu A_{cm} do całkowitej powierzchni gminy A_m (Bielecka 2005).

$$\rho_{cr} = \text{corr} \left(\frac{P_m^0}{P_m}, \frac{A_{cm}}{A_m} \right)$$

Współczynniki wagowe U'_{cr} , różnicujące gęstość zaludnienia pomiędzy różnymi typami pokrycia terenu i rodzajami gmin, obliczane są iteracyjnie, aż do momentu, gdy różnica δ_r , pomiędzy liczbą ludności przypisaną do każdej gminy P_m^0 a wartością podaną przez GUS P_m' ustabilizuje się.

$$U'_{cr} = U_c \left(1 - \frac{P_{cr} \times \delta_r}{2 \times P_r} \right)$$

Satysfakcjonujące wartości współczynników uzyskuje się po około czterdziestu iteracjach. Ostateczne wartości współczynników wagowych dla poszczególnych typów pokrycia terenu i rodzajów gmin województwa mazowieckiego zostały zestawione w tabeli 1. Mapa, pokazująca gęstość zaludnienia okolic Warszawy oszacowaną powierzchniowo-wagową metodą korelacji, została pokazana na rysunku 3. Mapa ta przedstawia obraz najbardziej zróżnicowany i najwierniej oddający charakter zjawiska. Dobrze widoczne jest, że gęstość zaludnienia zmienia się wraz z typem pokrycia terenu. Wadą tej metody, poza żmudnymi i czasochłonnymi obliczeniami, jest nie zachowanie sformułowanego przez Toblera (1979) warunku zachowania objętości danych w źródłowych jednostkach.

Ponieważ powierzchniowo-wagowa metoda korelacji zachowuje objętość danych tylko w obrębie województw, w celu oszacowania dokładności wolumetrycznej obliczono, dla każdej gminy, różnicę pomiędzy rzeczywistą liczbą ludności (dane statystyczne) a liczbą ludności przypisaną w wyniku procedury dazymetrycznej. Analiza różnic pokazała, że w gminach miejsko-wiejskich położonych na obrzeżach Warszawy (powiaty: wołomiński, pruszkowski, warszawski-zachodni) liczba ludności została niedoszacowana, natomiast w gminach uznawanych za rekreacyjne (powiaty: białobrzeski, sokołowski, wyszkowski) liczba ludności jest przeszacowana. Błąd względny szacowania liczby ludności w gminach województwa mazowieckiego został zilustrowany na rysunku 4. Lepsze dopasowanie współczynników wagowych można by prawdopodobnie osiągnąć wprowadzając dodatkową zmienną powiązań np. funkcje gminy.

Tabela 1. Wartości współczynników wagowych dla poszczególnych typów pokrycia terenu i rodzajów gmin w województwie warszawskim

Typy pokrycia terenu		Zabudowane tereny miejskie	Tereny przemysłowe i handlowe	Tereny rolnicze	Osadnictwo rozproszone	Lasy i roślinność naturalna
I	Miasta powyżej 100 tys.	20,3057	4,506	0,2538	6,4827	0,0749
II	Pozostałe gminy miejskie	21,2693	4,3446	0,6254	6,3867	0,0354
III	Część miejska gmin miejsko-wiejskich o gęstości zaludnienia >500 os/km ²	14,3739	2,8035	0,099	9,4669	0,4771
IV	Część miejska gmin miejsko-wiejskich o gęstości zaludnienia <500 os/km ²	10,2616	1,2296	0,5841	4,3265	0,2058
V	Gminy wiejskie z zabudową miejską	11,7465	0,7767	0,42	2,8618	0,278
VI	Gminy wiejskie bez zabudowy miejskiej			0,4943	2,6439	0,0654
VII	Część wiejska gmin miejsko-wiejskich z zabudową miejską	13,9312	0,975	0,5827	5,6549	0,2637
VIII	Część wiejska gmin miejsko-wiejskich bez zabudowy miejskiej			0,4474	1,5188	0,1606

Zakończenie

Kartowanie dazymetryczne i metoda dazymetryczna stały się terminami wieloznacznymi. W piśmiennictwie kartograficznym metoda dazymetryczna jest zaliczana do przedstawień kartogramicznych, w literaturze z dziedziny GIS stanowi formę interpolacji powierzchniowej danych statystycznych lub ich transformacji pomiędzy różnymi jednostkami odniesień przestrzennych (Langford i in. 1991, Langford i Unwin 1994, Eicher i Brewer 2001, Mennis 2002, Mennis 2003).

Prezentowane w artykule metody modelowania powierzchni statystycznej przedstawiającej gęstość zaludnienia są metodami interpolacji danych odniesionych do powierzchni, forma prezentacji wyników ma postać kartogramu. Można więc powiedzieć, że łączą przedstawione wyżej podejścia do metody dazymetrycznej.

Modelowanie powierzchni statystycznej przedstawiającej rozmieszczenie ludności w województwie mazowieckim przy pomocy metody dazymetrycznej pozwoliło na uzyskanie kompromisu między prostotą przekazu kartograficznego a zróżnicowaniem gęstości zaludnienia oraz umożliwiło rozwiązanie jednego z najistotniejszych problemów kartografii statystycznej – zwiększenie dokładności przekazu informacji ilościowych.

Literatura

- Bielecka E., 2005: A Dasymetric population density map of Poland. Materiały XXII Międzynarodowego Kongresu Kartograficznego, 11-16 lipca A Coruna, Hiszpania.
- Bielecka E., Ciołkosz A. 2005: *CORINE Land Cover*. Biblioteka Monitoringu Środowiska.
- Eicher C. L., and C. A. Brewer., 2001: Dasymetric mapping and areal interpolation: implementation and evaluation. *Cartography and Geographic Information Science* 28 (2): 125-138.
- Gallego J., Peedell S., 2001: Using CORINE Land cover to map population density. Towards agri-environmental indicators. EEA Topic report 6/2001:94-105.
- Goleń J., Ostrowski W., 1994: Metoda dazymetryczna – rys historyczny. *Polski Przegląd Kartograficzny* 26(1):3-16.
- Goleń J., Ostrowski W., 1996: Z problematyki dazymetrycznych map zaludnienia. *Polski Przegląd Kartograficzny* 28(2):79-85.
- GUS, 2005: Powierzchnia i ludność w przekroju terytorialnym w 2000 r. Informacje i Opracowania Statystyczne, Warszawa.
- Langford M., Maguire D.J., Unwin D.J. 1991: The areal interpolation problem: estimating population using remote sensing in a GIS framework. [W:] *Handling Geographical Information*, Essex UK, Longman Scientific & technical, pp: 55-77.
- Langford M., Unwin D. J., 1994: Generating and mapping population density surface within a geographical information system. *The Cartographic Journal* 31:21-26.
- Mennis J., 2002: Using geographic information systems to create and analyse statistical surface of population and risk for environmental justice analysis. *Social Science Quarterly*, 83(1).
- Mennis J., 2003: Generating surface models of population using dasymetric mapping. *The Professional Geographer* 55(1):31-42.
- Ratajski L., 1989: *Metodyka kartografii społeczno-gospodarczej*. Wyd.2. PPWK, Warszawa.
- Robinson A., Sale R., Morrison J., 1988: *Podstawy kartografii*. PWN, Warszawa.
- Tobler W., 1979: Smooth pycnophylactic interpolation for geographic regions. *Journal of the American Statistical Association* 74: 519-530.

Summary

Population mapping, in general, has two purposes: firstly, to cartographically portray the extent and density of population across an area of interest, and secondly, to derive a quantitative estimation of population density for use in subsequent spatial analytical modeling tasks. A cartographic portrayal of population traditionally has the form of a choropleth map. This kind of mapping is very simple, but despite its simplicity, choropleth maps have limited utility for detailed spatial analysis of population data, especially where human populations are concentrated in relatively small numbers of villages, towns and cities. One way to avoid this limitation is to transform the administrative units into smaller and more relevant map units through the process known as dasymetric mapping. The dasymetric technique maps a quantitative variable according to boundaries derived from the character of the data distribution. It is a form of an areal interpolation that uses ancillary data to transform population data from one set of spatial units to another.

This paper demonstrates the use of satellite derived ancillary land cover data to map population densities using dasymetric mapping. The three dasymetric methods presented, revealed the inter-regional variation in population density more realistically, in particular, among urban and rural areas. The methods were tested for Mazovia Region.

The binary method, the simplest, is easy to implement in GIS and gives a better view of population distribution over a given area than conventional choropleth maps (fig. 1). The only drawback is the delimitation of uninhabited areas.

The areal weighting aggregation method uses land cover data as limiting variable and a typology of communes as a correlation variable. We a priori assign the percentage of people attributed to each land cover type and groups of communes. Subjectivity of these decisions is considered the drawback of the method. The results are detailed enough and portray population density very realistic (fig.2).

The areal weighting correlation method presents a new way of calculating weighting coefficients. This is based on the method proposed by Gallego and Peedell (2001), but is adapted to Polish conditions by grouping land cover classes, stratifying communes and computing coefficients (tab.1). This dasymetric mapping method is based on the assumption that the ratio between the population density of two land cover categories is the same for any given commune. This method, on contrary to previous ones, does not preserve pycnophylactic property of statistical data, so it is necessary to evaluate the results. Relative errors were computed for evaluating the modified areal weighting method. The analyse of a comparison between attributed population and population data known from statistical measurements indicate that the population value attributed to most of the communes is approximately in agreement with the statistical data (fig.4). The coefficients seem to be too high for a few urban communes and too low for some rural areas. The areal weighting correlation method provides realistic view of population distribution in the Mazovia Region (fig 3).

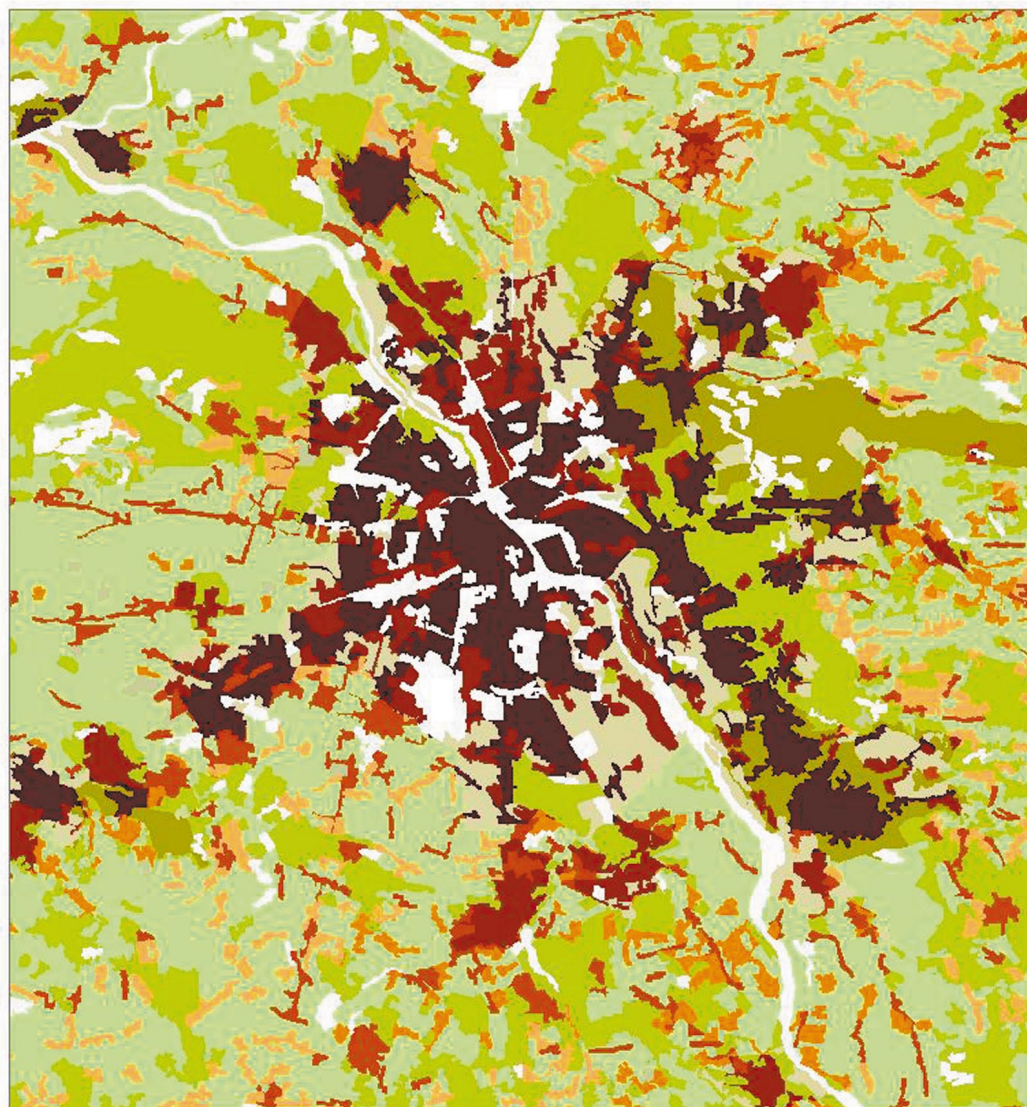
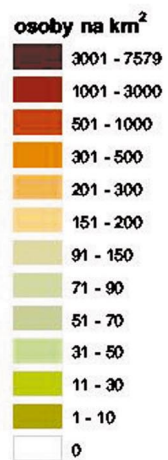
The dasymetric population map was visualized according to choropleth map rules in such a way that it focuses on the relationship between settlement and the natural environment. The information generated from a dasymetric population density map could provide useful assistance to district administrations, especially those responsible for regional or city development and land management.

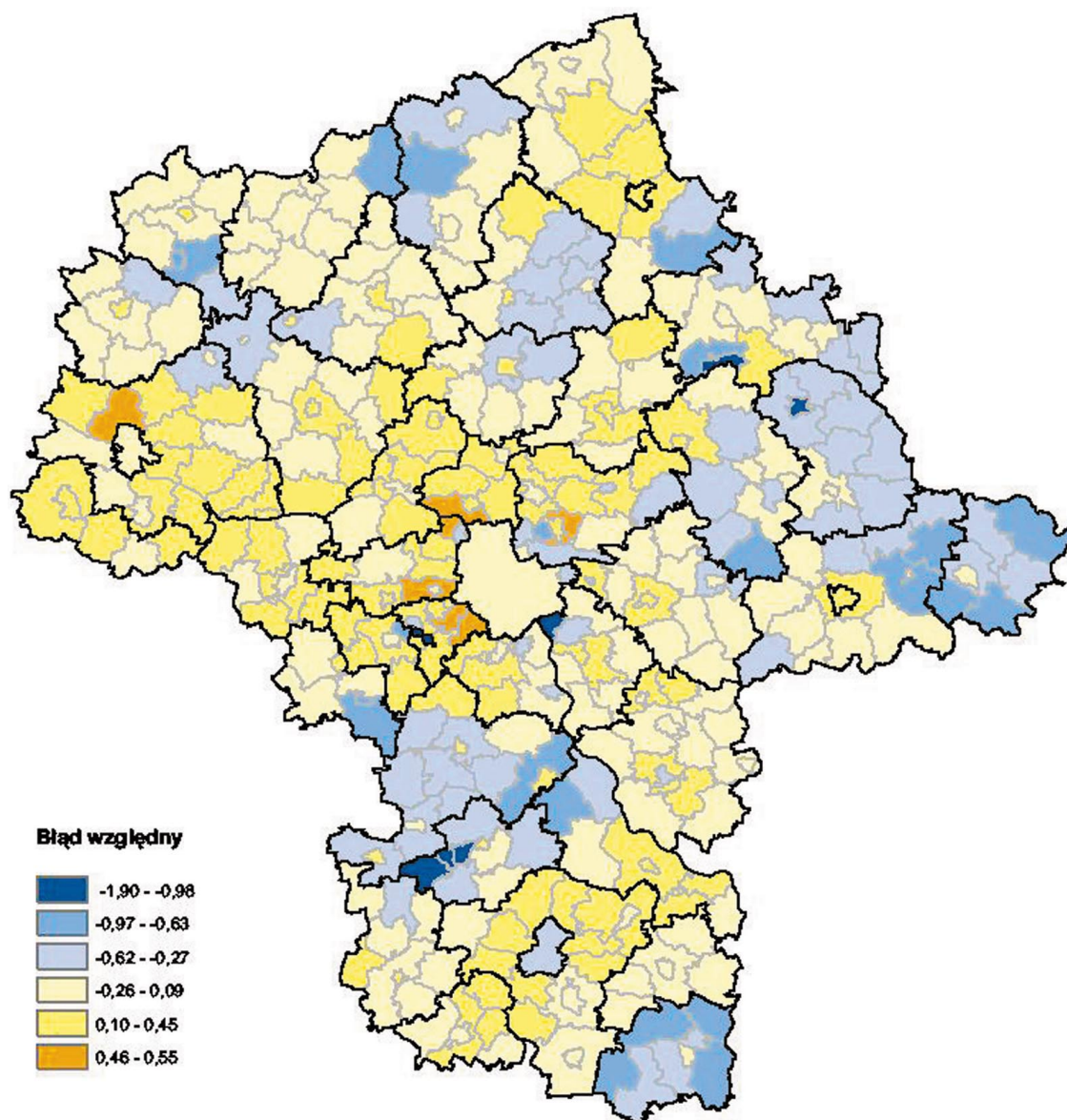
dr hab. inż. Elżbieta Bielecka
elzbieta.bielecka@igik.edu.pl

mgr Anna Kuczyk
anna.kuczyk@igik.edu.pl

mgr Ewa Witkowska
ewa.witkowska@igik.edu.pl

Rys. 3. Dazymetryczna mapa gęstości zaludnienia Warszawy i okolic; powierzchniowo-wagowa metoda korelacji





Rys. 4. Błąd względny szacowania liczby ludności w gminach województwa mazowieckiego