

EWA KRZYWICKA-BLUM  
Katedra Geodezji i Fotogrametrii  
Akademii Rolniczej we Wrocławiu

## **Agregacja danych punktowych i pól odniesienia a informacyjne własności map gęstości**

**Zarys treści.** W pracy przeprowadzona została analiza wpływu wielkości pól odniesienia i sposobu konstrukcji skali tematycznej na zmianę informacyjnych własności map przedstawiających zróżnicowanie gęstości danych punktowych. Na podstawie 44 przykładów różnych rozmieszczeń 300 punktów w 100-, a następnie 400-półowym podziale obszaru, dokonano analizy wpływu sposobu agregacji gęstości w klasy oraz wielkości pola odniesienia na zmianę poziomu i zakresu gęstości. Zastosowano metody statystyczne, wyrównanie metodą najmniejszych kwadratów i ocenę różnorodności za pomocą entropii. Ustalono korelację pewnych charakterystyk i wskazano możliwości ich wykorzystania we wstępnych etapach projektowania map.

### **1. Uwagi ogólne**

Współczesne techniki pozyskiwania danych i przechowywania ich w bazach o różnie skodyfikowanym sposobie określenia odniesień przestrzennych każą kartografom oraz innym specjalistom opracowującym mapy zwrócić uwagę na problem wyboru pola podstawowego w dostosowaniu do rozmieszczenia przestrzennego ważnych dla użytkownika elementów. Tak jak osiągnięcia geometrii fraktalnej potwierdziły słuszność stosowanej w tradycyjnej kartografii zasady wnikliwego doboru „skali obserwacyjnej” do „natury zjawisk” (dla przykładu – charakterystyk klimatycznych lub morfologicznych), tak i pewne metody analiz statystycznych mogą być szerzej wykorzystywane we wstępnym, koncepcyjno-studialnym etapie redagowania map. Dotyczy to między innymi map kropkowych.

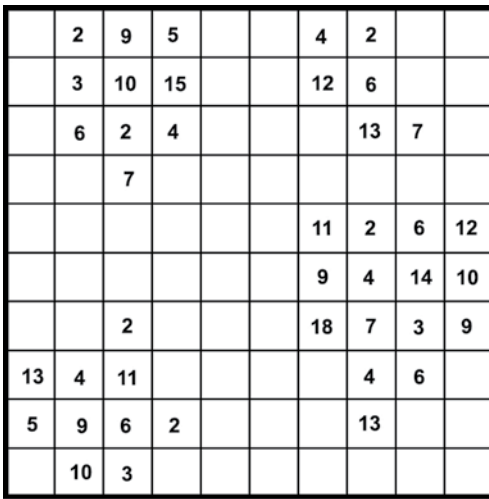
Dane źródłowe, na podstawie których opracowuje się mapy kropkowe, mogą mieć charakter obrazowy lub statystyczny. Jeśli kropka przedstawia grupę takich elementów, jak punkty emisji zanieczyszczeń lub poboru wody, rozdrobnione uprawy, szkoły, szpitale, hotele – dane źródłowe czerpie się z dokumentacji pomiarowej zorga-

nizowanej obrazowo. Stanowią ją mapy topograficzne, szkice geologiczne, zdjęcia termalne, akustyczne, fotograficzne, fotogrametryczne i różne formy rejestracji teledetekcyjnej.

Dane statystyczne, źródłowe dla opracowywanych map kropkowych to zebrane w system tabel liczbowe charakterystyki jakiejś całości, którą może być zarówno dział gospodarki, jak demografia (J. Kurkiewicz 1992). Dane uzyskane na podstawie okresowych spisów urzędowych, deklaracji, ankiet lub metodycznie zorganizowanych badań reprezentacyjnych, udostępnia się, po wstępnym opracowaniu (selekcji, grupowaniu a nawet przetworzeniu), w formie wielkości odnoszących się do jednostek podziału geograficznego, politycznego, administracyjnego lub wyznaniowego.

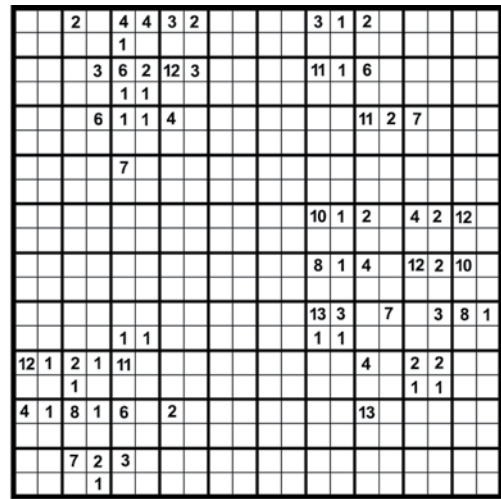
Poprawnie opracowana mapa kropkowa stanowi świetną podstawę formułowania wniosków dotyczących przestrzennej zmienności liczebności wybranej grupy elementów, jednak często tak przygotowany materiał nie odpowiada potrzebom holistycznie traktowanych analiz międzyregionalnych czy międzytematycznych (np. relacji gęstości sieci szpitali i gęstości zaludnienia). W studiach takich wygodne okazuje się dysponowanie mapami opracowanymi metodami kartogramu, kartodiagramu o skali skokowej lub izolinii. Mapy takie charakteryzują się wyraźnym odślonieniem kilku wybranych poziomów zróżnicowań.

W pracy przeprowadzono analizę porównawczą 44 różnych rozmieszczeń trzystu punktów w obszarze zawierającym 100 (10×10) pól, które następnie podzielono na 400 (20×20) ćwiartek. Stwarza to warunki porównań rozkładów punktów w układzie zagregowanym i rozdrobnionym. Rozkład statystyczny punktów, a więc liczebność grup przed i po podziale pól, charakteryzuje każdy z 44 układów, które ze względu na 44 osoby projektujące można traktować jako wzajemnie



Liczba		0	2	3	4	5	6	7	9	10	11	12	13	14	15	18
elementów	pól	59	6	3	5	2	5	3	4	3	2	2	3	1	1	1

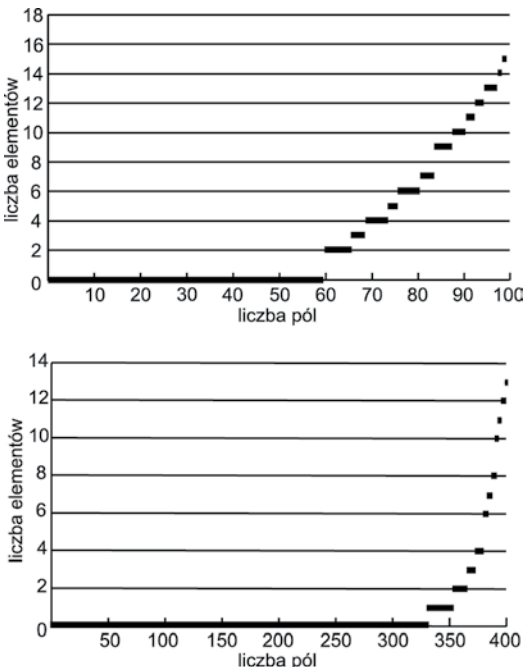
$\bar{x}$  Średnia 3,0  
 $\sigma$  Odchylenie standartowe 4,49  
 $\frac{\sigma}{\bar{x}}$  Wariancja/Średnia 6,71



Liczba		0	1	2	3	4	6	7	8	10	11	12	13
elementów	pól	329	22	13	7	7	4	4	3	2	3	4	2

$\bar{x}$  Średnia 0,75  
 $\sigma$  Odchylenie standartowe 2,25  
 $\frac{\sigma}{\bar{x}}$  Wariancja/Średnia 6,78

Ryc. 1. Rozmieszczenie 300 punktów w 100 i 400 polach odniesienia (przykład 1)  
 Fig. 1. Distribution of 300 dots in 100 and 400 reference areas (example 1)

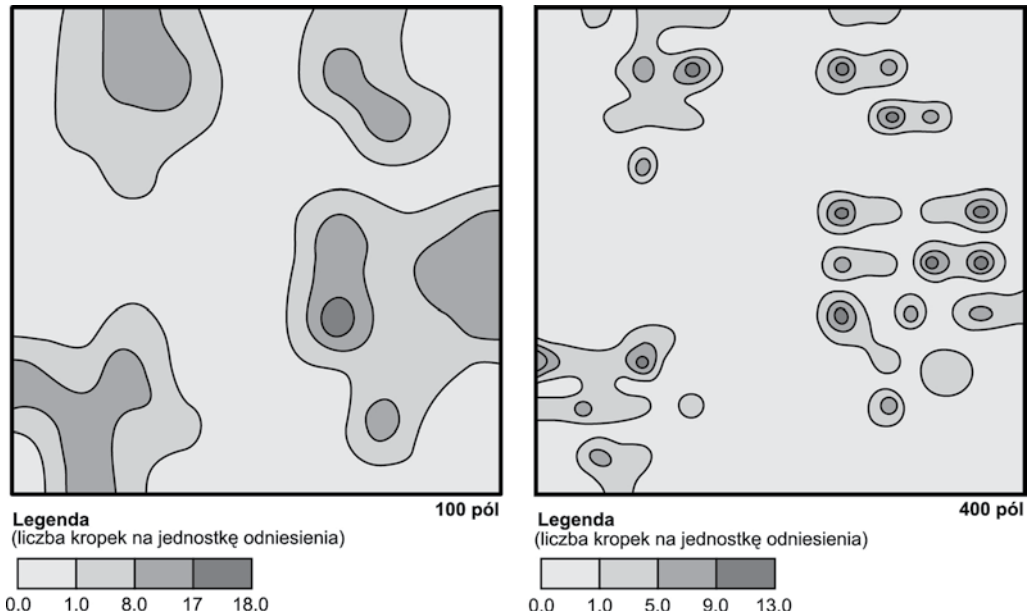


niezależne. Przykładowy rozkład punktów przedstawiono na rycinie 1 w obu podziałach, a poniżej zamieszczono parametry rozkładu gęstości. Wyniki przeprowadzonego doświadczenia, nie rozszcząć sobie praw do szerszych uogólnień, mogą być wskazówką co do możliwych efektów agregacji, a sam sposób analizy – propozycją metodyczną, przydatną we wstępnej fazie projektowania ważnej klasy map tematycznych sporządzanych na podstawie danych punktowych.

W eksperymencie zastosowano podział obszaru na równe co do powierzchni pola odniesienia, co jest powszechnie dziś stosowanym sposobem kodowania danych w systemach informacji przestrzennej. Rozpatrywanym punktom można przypisać różne stosowane w kartografii znaczenia, a więc zarówno 15 sklepów, 5 domów czy stacji benzynowych, jak i 100 osób niewidomych lub 20 000 sztuk owiec (R. Kucharski 2001, *Oxford Regional Economic Atlas 1967*).

Ryc. 2. Skumulowany wykres rozkładu liczby punktów w podziale obszaru P na 100 i 400 pól odniesienia (por. ryc. 1)

Fig. 2. Cumulated chart of the distribution of a number of dots in the division of P area into 100 and 400 reference areas



Ryc. 3. Mapy izarytmiczne przy podziałach obszaru na 100 i 400 pól odniesienia (por. ryc. 1)  
Fig. 3. Isarithmic maps in the case of the division of the area into 100 and 400 reference areas

## 2. Charakterystyka zbioru

Wstępny etap opracowania obejmuje dla każdego z 88 przykładów rozmieszczenia 300 punktów (44 w 100 polach i 44 w 400 – powstałych przez ćwiartkowy podział pól pierwotnych) ustalenie zakresu zmienności gęstości  $x$  określonej jako liczby elementów w polu<sup>1</sup> – a następnie średniej gęstości  $\bar{x}$  i odchylenia standardowego  $\sigma^2$

znajdącej się w polu  $x$  jako charakterystyki typu rozkładu statystycznego.

Aby zbadać wpływ wielkości pól podstawowych i procesu agregacji gęstości w klasy na stopień uogólnienia informacji wyjściowej i wynikające stąd pragmatyczne własności map, dokonano zgrupowania wartości w klasy, w każdym przykładzie z osobna, przed i po podziale pól podstawowych. Przyjęto zasadę łączenia wartości mniej różniących się, a rozdzielania – bardziej. Na wykresie skumulowanych częstości wystąpień (poszczególnych wartości gęstości) granice ustalano w miejscach „rozerwań” ukła-

du punktów, a następnie dzielono zbyt długie, przebiegające bez rozerwań części wykresu w punktach załamania łamanej aproksymującej krzywą, czyli w miejscach zmiany tempa przyrostu zmiennej. Przykład ustalania liczby klas przedstawiono na rycinie 2, a na rycinie 3 – odpowiadające podziałom mapy izopletowe. Jak można było przewidywać, różna wielkość pola podstawowego (1/100 i 1/400 część całości obszaru), mająca wpływ na rozkład statystyczny gęstości, a w konsekwencji na różne wartości granic przedziałów klasowych, generuje odmienny obraz kartograficzny, stanowiący informację o charakterze rozmieszczenia danych.

Przyjmując liczbę klas w podziale 100- i 400-polowym całego obszaru jako pewien „wskaźnik” charakteryzujący typ rozkładu statystycznego odpowiadającego rozmieszczeniu punktów, podzielono 44 przykłady na 7 grup zaliczając do grupy „3–3” – 14 przykładów, do „4–4” – 12 i kolejno do: „4–3” – 7, „3–4” – 4, „5–4” – 4, „5–3” – 2 i „5–5” – 1. Przykład na rycinie 1 zaliczono do grupy „4–4”, co oznacza, że liczba klas w przypadku obu podziałów nie uległa zmianie i wynosi 4. W grupach określono charakteryzujące je zakresy wartości wymienionych uprzednio wskaźników przed i po podziale ćwiartkowym pól podstawowych. Dla uzyskanych po zgrupowaniu w klasy liczebności (w poszczególnych klasach

<sup>1</sup> Modelowy charakter rozważań uzasadnia traktowanie pola podstawowego jako umownej jednostki powierzchni, a więc i wprowadzenie nazwy „gęstość”, zgodnie z terminologią statystyczną i odniesieniem do potocznego odbioru liczby elementów w polu jako gęstości.

gęstości) obliczono wartość charakteryzujących stopień różnorodności: entropii ( $H_1$ ) i entropii względnej ( $h_1$ ) według wzorów:

$$H_1(N) = - \sum_{i=1}^N w_i \lg_2 w_i$$

$$h_1 = \frac{H_1}{\lg_2 \frac{1}{N}}$$

gdzie  $N$  oznacza liczbę klas,  $w_i$  – częstość wystąpienia elementu klasy  $i$ , a więc udział pól charakteryzujących się gęstością „ $i$ ” w całym zbiorze pól wypełniających rozważany obszar  $P$ .

Dla przykładu rozmieszczenia przedstawionego na ryc. 1:  $H_1(4) = 1.13$ ,  $h_1 = 0.72$  przy podziale na 100 pól, a  $H_1'(4) = 0.89$ ,  $h_1' = 0.44$  – przy podziale na 400 pól.

Aby ustrzec się przypadkowości podziału zakresu zmienności gęstości na klasy przeprowadzono dla każdego z 44 rozmieszczeń, podział zgodny z pierwszym co do liczby, ale z założeniem ich równej rozpiętości i obliczono odpowiadające mu wartości  $H$  i  $h$ , nadając im, dla odróżnienia, indeksy „2”. Podział na klasy o równej rozpiętości ma charakter obiektywny, nie zapewnia jednak dostosowania do zasad agre-

gacji, stwarza bowiem możliwość wystąpienia klas pustych oraz zaliczenia do różnych klas elementów mniej się wzajemnie różniących, a do tych samych – różniących się bardziej. W przykładzie przedstawionym na rycinie 1:  $H_2(4) = 1.21$ ,  $h_2 = 0.61$ ,  $H_2'(4) = 0.49$  i  $h_2' = 0.24$ .

W zbiorczej tabelcy 1 zestawiono wyniki obliczeń dla grup. Oprócz wymienionych wskaźników, w ostatniej kolumnie zamieszczono wartości procentowego zmniejszenia części obszaru, w której występują elementy punktowe, a więc – przyrostu obszaru pól pustych przy podziale pól odniesienia i bardziej szczegółowej lokalizacji wystąpienia elementów punktowych. Jest oczywiste, że tylko w przypadku, gdy każde z wypełnionych punktami pól dzieli się na ćwiartki również zawierające punkty, nie następuje redukcja obszaru wystąpienia cechy przedstawianej na mapie. W rozważanych przykładach brak jest takiego przypadku, natomiast zakres pomniejszenia się obszaru wystąpienia punktów wyniósł od 14 do 54 procent, w wielu przykładach przekraczając poziom 40 procent. W przykładzie przedstawionym na rycinie 1, z 82-procentowej części obszaru występowania elementów punktowych, po rozdrobnieniu następuje zmniejszenie do 53%, a więc poziom redukcji DP wynosi 29%.

Grupowe relatywne wskaźniki, najczęściej obliczane jako wartości średnie ilorazu charakterystyk przed i po podziale ćwiartkowym pól odniesienia,

Tabela 1. Zakresy zmienności wskaźników

$N_{100}$	$(X_{100})_{\max}$	$(\sigma^2/X)_{100}$	$H_1$	$h_1$	$H_2$	$h_2$	$\Delta P(\%)$
$N_{400}$	$(X_{400})_{\max}$	$(\sigma^2/X)_{400}$	$H_1'$	$h_1'$	$H_2'$	$h_2'$	
3	8 - 19	0.87 - 5.29	0.52 - 1.52	0.44 - 0.96	0.78 - 1.53	0.50 - 0.97	20 - 45
3	4 - 13	0.95 - 4.54	0.33 - 1.48	0.21 - 0.93	0.27 - 0.98	0.17 - 0.62	
4	10 - 22	0.79 - 7.86	0.71 - 1.90	0.36 - 0.94	0.79 - 1.48	0.39 - 0.74	17 - 45
4	8 - 16	0.86 - 6.78	0.33 - 1.65	0.17 - 0.82	0.32 - 0.83	0.16 - 0.41	
4	11 - 17	1.24 - 6.17	1.00 - 1.78	0.50 - 0.89	1.03 - 1.46	0.29 - 0.73	14 - 54
3	5 - 14	0.87 - 4.04	0.32 - 1.15	0.20 - 0.72	0.32 - 0.85	0.20 - 0.53	
3	8 - 12	1.45 - 4.23	1.06 - 1.56	0.66 - 0.98	0.67 - 1.56	0.42 - 0.98	37 - 44
4	8 - 10	2.36 - 4.05	0.45 - 1.44	0.22 - 0.72	0.37 - 0.89	0.18 - 0.44	
5	17 - 22	7.33 - 11.11	1.32 - 1.72	0.57 - 0.74	1.14 - 2.60	0.49 - 0.60	14 - 25
4	12 - 19	5.05 - 7.74	0.58 - 0.90	0.29 - 0.52	0.29 - 0.36	0.15 - 0.18	
5	22 - 23	4.49 - 10.90	0.50 - 1.39	0.22 - 0.60	0.65 - 0.84	0.28 - 0.36	11 - 48
3	12 - 17	2.50 - 7.92	0.28 - 0.39	0.17 - 0.27	0.16 - 0.25	0.10 - 0.16	
5	0.21	11.30	1.33	0.57	1.06	0.46	15
5	0.19	8.49	0.33	0.14	0.33	0.14	

Tablica 2. Wskaźniki relatywne

Grupa	n	$\bar{S}(x)$	$S(z)$	$\bar{S}(\sigma^2/x)$	$\bar{S}(h_1)$	$\bar{S}(h_2)$	$\Delta P$
3 - 3	14	1.8	1.2	1.4	1.2	2.0	37.1
4 - 4	12	1.4	1.5	1.3	1.65	2.4	28.75
4 - 3	7	1.9	0.6	1.9	1.8	2.0	28.1
3 - 4	4	1.2	2.0	1.0	2.0	2.4	39.5
5 - 4	4	1.2	0.8	1.5	1.6	3.2	18.75
5 - 3	2	1.6	0.2	1.6	1.85	2.7	29.5
5 - 5	1	1.1	( )	(1.3)	(4.0)	(3.2)	(15)
wartości średnie:	1.46	1.05	1.43	1.43	2.0	2.6	28.04
odchylenie stand.:	0.29	0.59	0.26	0.26	0.84	0.47	8.20

przedstawiono w tablicy 2. W tablicy tej n oznacza liczbę rozkładów w danej grupie, a z – rozrzut zakresu gęstości. Dla przykładu w grupie „3–3”, w której zakresem rozkładów: 8, 8, 8, 10, 11, 12, 12, 12, 14, 15, 15, 17 i 19 dla podziału 100-polowego odpowiadały zredukowane zakresy odpowiednio: 4, 4, 5, 7, 6, 7, 7, 8, 8, 8, 7, 11, 5 i 13 dla podziału 400-polowego, współczynniki redukcji zakresu wynosiły: 2, 2, 1.6, 1.4, 1.8, 1.7, 1.7, 1.5, 1.5, 1.8, 2.1, 1.4, 3.4, 1.5. Średnia wartość w grupie „3–3” współczynnika redukcji zakresu gęstości  $\bar{S}(x) = 1.8$  (por. tab. 2), natomiast  $S(z)$  czyli stosunek całego rozrzutu gęstości:  $z_{100}$  przed i  $z_{400}$  – po rozdrobnieniu podziału obszaru kształtuje się (por. tab. 1) jako iloraz wartości  $19 - 8 = 11$  (przed) i  $13 - 4 = 9$  (po podziale) czyli wynosi 1.2. Nie jest to wartość średnia. Pozostałe wskaźniki:  $\bar{S}(\sigma^2/x)$ ,  $\bar{S}(h_1)$ ,  $\bar{S}(h_2)$ ,  $\Delta P$ , związane z oceną typu rozkładu i różnorodności oraz zmianą redukcji zakresu występowania cechy, są wartościami średnimi w grupie.

**3. Uzyskane wyniki**

47 procent wszystkich przykładów stanowią rozkłady gęstości, które zostały podzielone na 3 klasy. Jest to 18 rozkładów należących do dwóch grup w podziale obszaru P na 100 pól i 23 rozkłady z trzech grup w podziale obszaru na 400 pól. 44% rozkładów podzielono na 4 klasy gęstości. Jest to 19 z dwóch grup w podziale 100-polowym i 20 z trzech grup w podziale rozdrobnionym.

Pozostałe 9% zbioru przykładowych rozkładów podzielono na 5 klas: 7 należących do 3 grup w podziale 100-polowym i 1 – w 400-polowym.

Średnia wartość współczynnika  $\frac{\sigma^2}{x}$  dla wszystkich rozkładów z podziałem gęstości na 3 klasy wynosi 2.83, dla rozkładów z podziałem na 4 klasy – 4.42, a na 5 – 8.89. Świadczy to o „naturalnym” rozkładzie funkcji

$$N = N \left( \frac{\sigma^2}{x} \right)$$

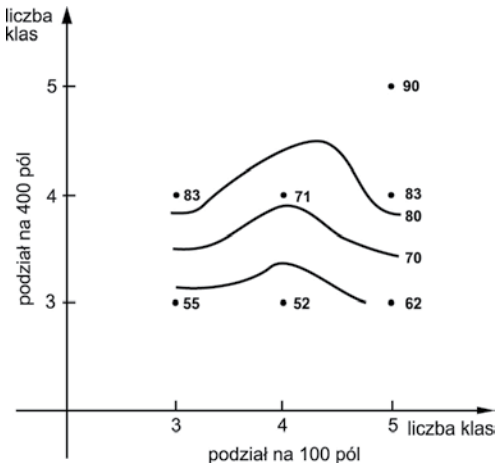
Przy czterokrotnym zmniejszeniu wielkości pola odniesienia odwrotność średniego współczynnika

$$\frac{1}{x}$$

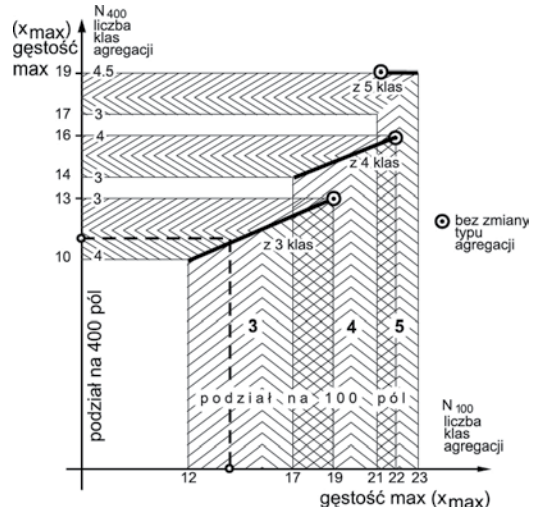
redukcji zakresu gęstości  $\bar{S}x$  (tab. 2) zmienia się od wartości 0.55 dla grupy „3 – 3” poprzez 0.71 dla grupy „4–4” i „4–3” do 0.90 dla grupy „5–5”. Odpowiednie wartości obniżenia zakresu gęstości przy podziale pól odniesienia. Wyinterpolowane linie, wskazujące ogólne tendencje zmian, przedstawiono na rycinie 4, a zmianę poziomu gęstości maksymalnej na rycinie 5.

Dla przykładu przy podziale na 100 pól (górne wiersze w grupach w tablicy 1) trzyklasowe podziały gęstości charakteryzują grupy „3–3” i „3–4”, stąd na rycinie 5 przedział maksymalnych wartości gęstości oznaczono jako 12 – 19. Analogicznie dla podziału czteroklasowego (grupy: „4–4” i „4–3”) 17 – 22 i dla pięcioklasowego (grupy „5–4”, „5–3” i „5–5”) 21 – 23. Z ryciny 5 można odczytać, że przedział gęstości 12 – 19, po podziale pól odniesienia, redukuje się do poziomu gęstości 10 – 13, a więc dla rozkładu z maksymalnym poziomem gęstości 14 można spodziewać się jego obniżenia do poziomu 11. Niejednoznaczność przedziału 17 – 19 jest po-

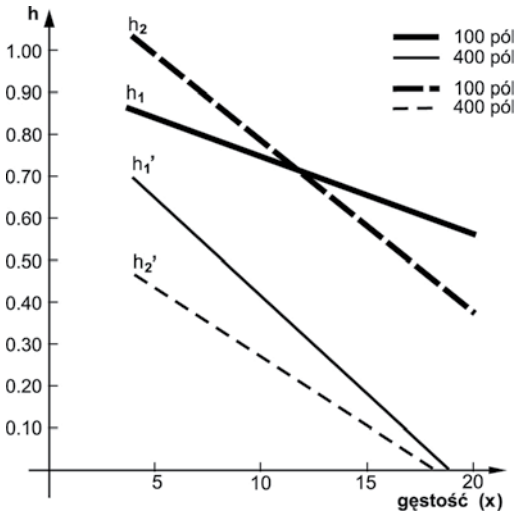




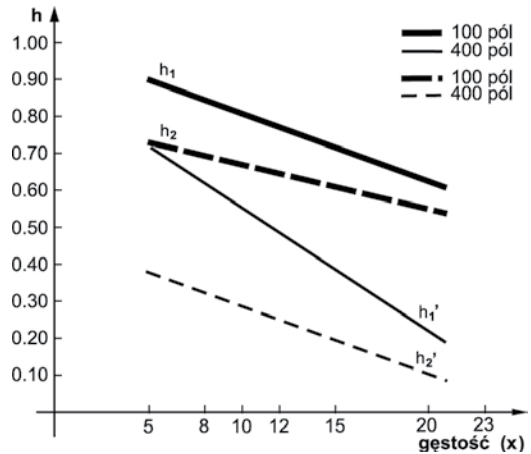
Ryc. 4. Redukcja zakresu gęstości przy czterokrotnym zmniejszeniu wielkości pól odniesienia (w % wartości przed rozdrobnieniem)  
 Fig. 4. Reduction of density range in the case of reference areas being diminished 4 times (% of the initial value)



Ryc. 5. Zmiana poziomu gęstości maksymalnej pól podstawowych po rozdrobnieniu  
 Fig. 5. Change of the level of maximum density of primary areas after subdivision



Ryc. 6. Entropia względna jako miara zróżnicowania zagregowania do trzech klas gęstości w 100- i 400-półowych podziałach obszaru ( $h_1, h_1'$  – w podziale wg rozkładu;  $h_2, h_2'$  – w podziale równowartościowym)  
 Fig. 6. Relative entropy as the measure of diversity of aggregation into three classes of density in 100- and 400-field divisions ( $h_1, h_1'$  – divided according to distribution;  $h_2, h_2'$  – divided according to equal value)



Ryc. 7. Entropia względna jako miara zróżnicowania zagregowania do czterech klas gęstości w 100- i 400-półowych podziałach obszaru ( $h_1, h_1'$  – w podziale wg rozkładu;  $h_2, h_2'$  – w podziale na klasy o równej rozpiętości)  
 Fig. 7. Relative entropy as the measure of diversity of aggregation into four classes of density in 100- and 400-field divisions ( $h_1, h_1'$  – divided according to distribution;  $h_2, h_2'$  – divided according to equal value)



Tablica 3. Współczynniki korelacji wskaźników relatywnych w grupach

$r_{ij}$	$\bar{S}(x)$	$S(z)$	$\bar{S}(\sigma^2/x)$	$\bar{S}(h_1)$	$\bar{S}(h_2)$	$\bar{\Delta}P$
$\bar{S}(x)$	1	- 0.48	0.30	- 0.58	- 0.80	0.45
$S(z)$	- 0.48	1	- 0.84	0.37	- 0.23	- 0.24
$\bar{S}(\sigma^2/x)$	0.30	- 0.84	1	- 0.22	- 0.20	- 0.26
$\bar{S}(h_1)$	- 0.58	0.37	- 0.22	1	0.59	- 0.63
$\bar{S}(h_2)$	- 0.80	- 0.23	- 0.20	0.59	1	- 0.78
$\bar{\Delta}P$	0.45	- 0.24	- 0.26	- 0.63	- 0.78	1

Przykłady z podziałem gęstości na cztery klasy (por. tab. 2) stanowią 39-elementową grupę rozkładów: 19 w podziale na 100 pól (12 w grupie „4–4” i 7 – w grupie „4–3”) i 20 w podziale pola P na 400 jednostek (12 w grupie „4–4” i po 4 w grupach „3–4” i „5–4”). Równania prostych regresji mają postać:  $h_1 = 0.984 - 0.018x$

$$h_2 = 0.788 - 0.012x$$

dla podziału bazowego (100-polowego) i dla rozdrobionego:  $h_1' = 0.905 - 0.034x$

$$h_2' = 0.469 - 0.018x$$

Odpowiednie zależności gęstości od sposobu agregacji i wielkości pól odniesienia przedstawiono na rycinach 6 i 7, dla agregacji 3-klasowej (ryc. 6) i dla agregacji 4-klasowej (ryc. 7).

Zaledwie siedmioelementowy zbiór rozkładów z 5-klasową agregacją (6 w podziale 100-polowym i 1 w 400-polowym) nie może stanowić podstawy do wnioskowania o charakterze zależności w tej grupie rozkładów między zakresem gęstości a zróżnicowaniem.

#### 4. Wnioski ogólne

Pewną ogólną orientację w charakterze relacji między wskaźnikami stosowanymi przy ocenie skutków sposobu agregacji w klasy zbioru gęstości oraz różnej wielkości pól odniesienia można uzyskać z zestawienia, na podstawie danych zebranych w tablicy 2, odpowiednich współczynników korelacji  $r_{ij}$  między relatywnymi średnimi wartościami wskaźników w grupach. Wyniki obliczeń przedstawiono w tablicy 3.

Najwyraźniejszą zależność można dostrzec między relatywnymi współczynnikami zmiennych

$$\frac{\sigma^2}{x}$$

z oraz  $x$ . Oznacza to, że im większa jest, przy zmianie wielkości pól odniesienia, stabilność

$$\frac{\sigma^2}{x}$$

charakteryzującej typ rozkładu wielkości  $x$ , tym większy jest rozrzut zakresów gęstości. Ważniejszą dla praktyki modelowania kartograficznego informacją jest wynik:

$$r(\bar{S}x, \bar{S}h_2) = -0.8$$

oznaczający, że stosunek  $h_2/h_2'$  entropii względnej w podziale na klasy o równej rozpiętości przed i po podziale ćwiartkowym pól odniesienia można traktować jako wskaźnik stopnia redukcji zakresu gęstości. (Ujemny znak jest naturalny, wystarczy bowiem obliczyć

$$r\left(\frac{1}{\bar{S}x}, \bar{S}h_2\right) = 0.8$$

aby uzyskać informację znakowo zgodną z ukierunkowaniem zmian).

Jeszcze jeden związek ma dla modelowania kartograficznego charakter znaczący, mianowicie wynik:

$$r(\bar{\Delta}P, \bar{S}h_2) = -0.8$$

oznacza, że im większa jest zmiana zróżnicowania gęstości, tym mniejszy jest przyrost obszaru, w którym nie występują żadne elementy punktowe, a więc większa stabilizacja zasięgu występowania przedstawianej na mapie cechy. Wielkość  $\bar{S}(h_2)$  można zatem traktować jako wskaźnik poprawności doboru wielkości pól odniesień, co ma w procesie generalizacji podstawowe znaczenie. Praktycznie – obliczenie stosunku  $h_2/h_2'$  dla danego rozkładu powinno zawsze być poprzedzone ustaleniem liczby klas agregacji na podstawie analizy typu rozkładu statystycznego gęstości. Dla przykładu przedstawionego na rycinie 8 na podstawie analizy rozkładu ustalono grupę



agregacyjną „4–3”. Odpowiednia wielkość  $h_2/h_2'$  wynosi 3, a więc duża jest zmiana zróżnicowania. Odpowiada jej mała, bo zaledwie 14-procentowa zmiana zakresu występowania elementów punktowych, a więc na tyle duża stabilizacja zasięgu, że rozdrobnienie nie ma istotnego znaczenia dla samego studium zasięgu, stanowiąc jedynie uszczegółowienie informacji ilościowej.

Jak zaznaczono we wstępie, uzyskane wyniki nie mają pretensji do zbyt szerokich uogólnień,

#### Literatura

Kucharski R., 2001, *Mapy tematyczne. Siła analityczna MapInfo*. „IMAGIS Forum” nr 5, Warszawa.  
Kurkiewicz J., 1992, *Podstawowe metody analizy demograficznej*. Warszawa: PWN.

wskazują jednak na empirycznie potwierdzoną efektywność metod, które mogą być wykorzystane w etapie wstępnych analiz procesów agregacji danych w klasy i doboru wielkości pól odniesienia.

*United States & Canada. Oxford Regional Economic Atlas*. 1967, London: Oxford Univ. Press.

Recenzował  
prof. dr hab. Andrzej Makowski

### Aggregation of dot elements and reference space units and informational characteristics of density maps

#### Summary

The process of transformation of dot elements into specific thematic maps, which present the distribution of a particular feature using a choropleth or isopleth method depends on a proper size of reference space units and the criteria of aggregation of data into classes.

In the paper 44 spatial distributions of 300 points have been analyzed according to the division of  $P$  area into 100 squares, which were later subdivided into 400. Two methods of grouping into density classes (the number of dots in reference area) have been applied. The first follows the natural breaks and the second bases on the classes of equal range.

All 88 density patterns (before and after subdivision) have been aggregated into classes. 7 groups have been established according to the number of classes: „3–3”, „4–4”, „5–5”, „4–3”, „3–4”, „5–4” and „5–3”. Statistic characteristics have been determined within each group. Relative entropy has been used as an especially important measure of diversity (of density in a group). Comparison of indicator levels before and after the subdivision of each area has been used to analyze the relation of area size and the reduction of the density range. Through the application of the method of the least squares for the two most numerous groups of aggregation (to 3 and, separately, to 4 classes of density) it was possible to establish a type of linear dependence

between the density range and the level of diversity.

In the case of aggregation into three classes it has been proved, that the condition of the same level of diversity is:

$$x_{\max} \{100\} = 3 \times x_{\max} \{400\}$$

For instance, when before the division of reference areas the diversity range of density was 0–18 and after the division 0–6, then (see fig. 6):

$$h_1(18) = h_1'(6) = 0.60$$

Finally, some interesting correlation indicators between average characteristics have been determined. Interdependence between the value of the indicator of relative change of density range and the compensation of the range of dot elements due to the subdivision of reference areas is a significant result of the analysis.

Although presented conclusions can only be treated as empirically obtained results, the number and variety of examples allows to accept their usefulness for generalization procedures necessary in initial preparation of important types of thematic maps.

Translated by M. Horodyski

## Агрегация точечных данных и территориальных единиц и информационные свойства карт плотности

### Резюме

Процесс преобразования множества точечных данных в специальные тематические карты, представляющие размещение некоторой действительной черты с применением картограммного или изоплетного метода зависит от правильного подбора величин территориальной единицы и от принятых критериев агрегации данных в классы.

В работе анализом охвачены были 44 пространственных распределений 300 точек (пунктов) применительно к делению территории  $P$  сначала на 100, а потом на 400 базовых квадратных полей, образованных путём четвертичного деления. Были применены два метода группирования в классы плотности (число точек в поле), а именно, приспособленный к размещению плотности и в классы с равным диапазоном.

Была произведена агрегация в классы всех 88 примерных размещений плотности (до и после деления полей) и определён ряд статистик в каждой из 7 выделенных согласно числу классов: „3–3”, „4–4”, „5–5”, и „4–3”, „3–4”, „5–4”, „5–3” групп. Особенно важной мерой дифференциации (всего набора плотности в группе) признано относительную энтропию. Сравнение уровня показателей до и после деления территориальных единиц на четверти использовано для анализа влияния величины территориальной единицы на редукцию охвата (диапазона) плотности. Применяя метод наименьших

квадратов для двух наиболее многочисленных групп агрегации (до 3 и отдельно до 4 классов плотности), был установлен вид линейной зависимости диапазона плотности и степени дифференциации. При агрегации плотности до трёх классов доказано, что выполнено условие уравнивания дифференциации:

$$x_{\max} \{100\} = 3 \times x_{\max} \{400\}$$

Для примера, когда перед делением территориальных единиц диапазон изменчивости плотности равнялся 0–18, а после деления 0–6, тогда (срав. рис. 6):

$$h_1(18) = h_1'(6) = 0.60$$

В итоге вычислены величины ряда коэффициентов корреляции между средними величинами характеристик. Значимым результатом анализа может быть признана зависимость между величиной показателя относительного изменения диапазона плотности и компенсацией территории выступления точечных элементов в результате внутреннего деления территориальных единиц.

Несмотря на то, что представленные в работе итоги могут считаться единственно результатами, полученными эмпирически, число и дифференциация примеров даёт возможность признать их пригодными для генерализирующих процедур, необходимых в предварительных работах по проектированию важных видов тематических карт.

*Перевод Р. Толстикова*