

Z Katedry Geografii Ekonomicznej Wydziału Biologii i Nauk o Ziemi UMCS
Kierownik: prof. dr Jan Ernst

Ryszard JEDUT

**Problemy porządkowania i podziału jednostek terytorialnych
przy użyciu zespołu metod taksonomicznych**

Проблемы упорядочивания и деления территориальных единиц при помощи
таксономических методов

Problems of Arranging and Subdividing Territorial Units by the Use of a Group
of Taxonomic Methods

Spośród szeregu problemów klasyfikacji (regionalizacji) dotychczas nie rozwiązanych porządkowanie badanych jednostek i ich podział na grupy jest tym zagadnieniem, od rozwiązania którego w poważnym stopniu zależy ostatecznie powodzenie innych etapów klasyfikacji czy regionalizacji. Problem ten zostanie rozpatrzony w niniejszej pracy głównie pod kątem widzenia regionalizacji rolnictwa, ale wiele omawianych zagadnień znacznie wykracza poza problemy regionalizacji rolniczej.

Jest rzeczą oczywistą, że najważniejszym elementem procesu regionalizacji jest sprawa koncepcji, ale także bezsporny jest fakt, że cel może być osiągnięty tylko przy pomocy odpowiedniej metody, umożliwiającej zrealizowanie tej koncepcji. W tym sensie priorytet dany rozważaniom na temat metody wiąże się bezpośrednio z istotą zagadnienia i ma umożliwić rozwiązanie problemu regionalizacji.

Przed ostatecznym zdecydowaniem, którą z możliwych metod przyjmą za podstawową, wypróbowano szereg z nich na materiałach wykorzystanych w niniejszej pracy. Spośród wypróbowanych metod pozwalających na uwzględnienie kilku cech jednocześnie — metody rang M. G. Kendella, metody względnej dominanty J. Ernsta, metody korelacji wielorakiej A. Wankego i kilku innych — stosunkowo niezłe wyniki, po pewnych przekształceniach (11), dała metoda względnej do-

minanty („względnego uprzywilejowania”). Niemniej nie udało się przy pomocy tej metody w pełni rozwiązać problemu granic. Trudności wystąpiły szczególnie w obszarach przejściowych, wykazujących zmiany o charakterze strefowym. Z tego względu metodę tę w dalszych rozważaniach potraktowano jako rozpoznawczą i pomocniczą, a główny ciężar rozwiązania problemu postawiono przed pewną grupą metod taksonomicznych, którą ogólnie możemy wiązać z metodą różnic J. Czekanowskiego.

Problemy metodologiczne procesu regionalizacji w oparciu o metody taksonomiczne sprowadzić można do kilku zagadnień realizowanych na różnych etapach postępowania: 1) wyboru cech regionalizacyjnych, 2) standaryzacji i ważenia cech, 3) obliczania różnicy między badanymi jednostkami, 4) porządkowania i podziału na grupy (regiony). Postępowanie badającego w każdym z tych etapów może być różne i niezależne od decyzji podejmowanych na innych etapach.

Zajęcie się w niniejszej pracy tylko ostatnim etapem — porządkowania i podziału na grupy (regiony) wiąże się z dużymi trudnościami, na jakie napotykamy w tym stadium regionalizacji. Oczywiście możliwość przejścia do porządkowania i podziału na grupy wymaga uprzedniego wykonania wcześniejszych etapów regionalizacji.

Podstawę porządkowania stanowi tablica odległości (różnic)¹, przedstawiająca porównanie każdej badanej jednostki z każdą inną (tab. 1). Tablica różnic została obliczona w oparciu o zespół sześciu głównych upraw i umożliwia wydzielenie regionów produkcji roślinnej (polowej) w województwie lubelskim dla r. 1965.

TAKSONOMICZNE METODY GRUPOWANIA

Różnorodność technik stosowanych do wydzielenia regionów wynika z jednej strony ze zróżnicowania ich historycznego rozwoju, z drugiej — z faktu, że żadna ze stosowanych metod nie jest doskonała i może spełniać swoje zadanie tylko w określonym zakresie.

Omówienie na tym miejscu kilku z nich (diagramu Czekanowskiego, dendrytu wrocławskiego i „drzewa połączeń”) wynika nie tyle z chęci konfrontowania wyników każdej z wymienionych metod,

¹ Pojęcia odległości i różnic mają w metodach taksonomicznych ścisłe znaczenie wynikające z odpowiedniego sposobu obliczania podobieństwa między badanymi jednostkami. W tym jednak przypadku pojęcie odległości będzie używane jako synonim ogólnego pojęcia odległości taksonomicznej. Natomiast pojęcie różnic będzie odnosiło się do różnic przeciętnych obliczanych według odpowiedniego wzoru.

co z powodu ich wzajemnego uzupełniania się. Nieco szczegółowsze omówienie podyktowane jest także tym, że z wyjątkiem dendrytu pozostałe metody zostaną zastosowane w formie zmodyfikowanej, różnej od pierwotnego wzoru.

DIAGRAM CZEKANOWSKIEGO²

Diagram Czekanowskiego opracowuje się na podstawie macierzy odległości (tablicy różnic). Zasada sporządzania diagramu polega na zastąpieniu liczb — tablicy różnic — przez odpowiednie sygnatury. Zwykle odległości (różnice) najmniejsze zastępuje się najbardziej intensywnym zaczerzeniem, a odległości coraz to większe — w odpowiednio dobranych przedziałach klasowych — coraz mniej intensywnym zaczerzeniem. Następnie porządkuje się diagram poprzez przestawianie poszczególnych kolumn i odpowiadających im wierszy tak, aby pola zaczerznięte znalazły się blisko przekątnej diagramu. Na podstawie pól zaczerzniętych, które dążą do skupienia się w postaci kwadratów przy przekątnej, przeprowadza się podział badanych jednostek na grupy (regiony). W oparciu o graficzny obraz stopni podobieństwa dokonuje się analizy powiązań między poszczególnymi jednostkami.

I chociaż sama zasada opracowania diagramu wydaje się być prosta, to jednak istnieją dotychczas pewne problemy nie rozwiązane w sposób zadowalający. Jednym z problemów, który już przed przystąpieniem do porządkowania należy rozstrzygnąć, jest ustalenie przedziałów klasowych (ich ilości, rozpiętości) dla poszczególnych stopni podobieństwa. W dotychczasowej praktyce albo stosowano przedziały konwencjonalne (np. Ernst, Fierich), albo też podejmowano próby opracowania określonych zasad ustalania klas podobieństw (np. Gługniiewicz, Steczkowski).

Zagadnienie jest tym ważniejsze, że jeszcze przed uporządkowaniem diagramu wartości liczbowe grupowane są w pewne klasy. Na przykład gdy przyjmujemy przedziały 0—5, 5—10, itd., to w praktyce porządkowania diagramu i jego podziału na części większość decyzji podejmujemy w oparciu o wartości graniczne (5, 10, ...), a głównie pierwszą. Uznajemy tym samym niejako automatycznie, że wartości 0,1 i 4,9 są

² Zasadę porządkowania przy pomocy diagramu podał J. Czekanowski w r. 1909 w pracy „Zur Differenzialdiagnose der Neanderthalgruppe” *Korrespondenz — Blatt d.D.G.f. Anthr. u. Urgesch* XI. Jahrg. nr 6/7 Braunschweig 1909. Diagram Czekanowskiego doczekał się tak obfitych zastosowań w kraju i za granicą nie tylko w antropologii, ale także w wielu innych dziedzinach, że wymienianie na tym miejscu pozycji innych niż bezpośrednio związanych z tematem pracy wymagałoby sporządzenia wielostronicowej bibliografii. Podstawowe założenia porządkowania diagramu nie zostały jak dotychczas w niczym zmienione.

bliższe siebie niż wartości 4,9 i 5,1, ponieważ wszystkie jednostki w granicach pierwszego przedziału stanowią podobieństwo pierwszego stopnia, natomiast między przedziałami — podobieństwo drugiego stopnia itd.

Ponieważ zagadnienie przedziałów klasowych w każdym przypadku jest elementem obciążającym metodę, dlatego podjęta zostanie tutaj próba nieco innego rozwiązania. Mianowicie zamiast porządkować diagram na podstawie przyjętych klas podobieństw, spróbujemy uporządkować go w oparciu o konkretne wartości liczbowe.

Wymaga to jednak nieco precyzyjniejszego sformułowania najlepszego uporządkowania. Zgodnie z definicją pierwowzoru za najlepsze uporządkowanie należałoby uznać takie, które posiada najwyższą koncentrację odległości (różnic) minimalnych wokół przekątnej, a jednocześnie najwyższą koncentrację odległości maksymalnych w strefie najbardziej oddalonej od przekątnej, tak aby każde przedstawienie nie realizujące warunku najlepszego uporządkowania zmniejszało wskaźnik koncentracji, a każde prawidłowe przedstawienie zwiększało go. Byłoby to jednocześnie uporządkowanie jednakowo właściwe z punktu widzenia wszystkich badanych jednostek, a nie tylko tej grupy, dla której przedziały klasowe zostały poprawnie ustalone.

Przykład konkretnego uporządkowania według wymienionych zasad stanowi tab. a ryc. 1. Oceny poprawności uporządkowania dokonujemy poprzez sumowanie różnic w kolejnych szeregach równoległych do przekątnej według następującej reguły:

$$\begin{array}{r} r_{2,1} + r_{3,2} + r_{4,3} + \dots + r_{19,18} = r_{1+1,1} = S_1 \\ r_{3,1} + r_{4,2} + r_{5,3} + \dots + r_{19,17} = r_{1+2,1} = S_2 \\ \dots \\ r_{19,1} = r_{1+18,1} = S_{18} \end{array}$$

Wzór ogólny ma postać:

$$S_j = \sum_{i=1}^{n-j} r_{i+j,i} \quad \text{dla } j = 1, 2, 3, \dots, n-1$$

gdzie: S_j — kolejna suma (j — ta suma)
 n — oznacza ilość wierszy macierzy.

Obliczenia dla interesującego nas przykładu przedstawiają się następująco:

$$S_1 = 2,48 + 2,53 + 1,25 + \dots + 2,42 = 44,67$$

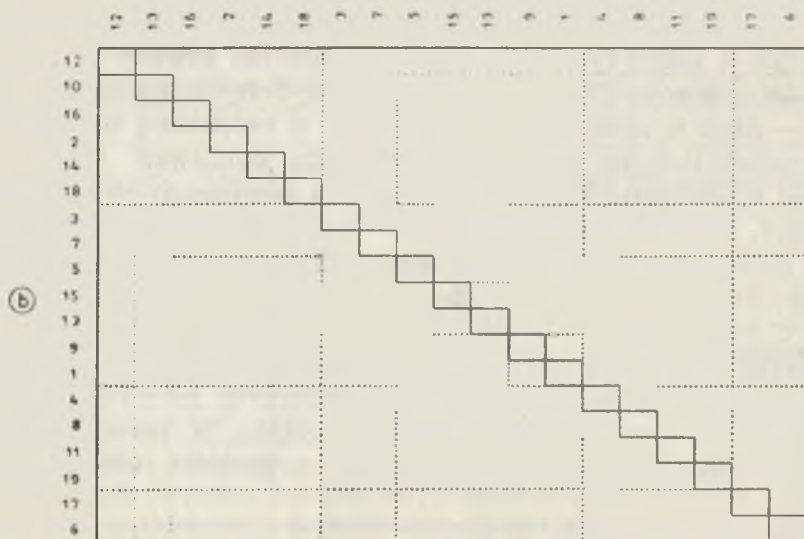
$$S_2 = 2,41 + 2,22 + 1,34 + \dots + 3,47 = 54,67$$

$$S_{18} = 15,03$$

Ponieważ sumy pojedynczych szeregów mogą wykazywać nieznaczące różnice (+, -) przy różnych wariantach uporządkowania, ostatecznej oceny „lepszości” uporządkowania dokonujemy na podstawie szeregu

(a)

	12	10	7	14	18	3	7	5	15	13	9	1	4	8	11	19	17	6	
12		25	24	25	37	42	71	83	91	82	86	96	100	124	118	114	118	141	150
10	25		25	22	26	26	54	64	66	57	61	76	81	99	93	89	93	119	126
16	24	25		13	13	21	47	59	81	72	76	86	90	114	108	104	108	132	140
2	25	22	13		12	17	50	58	85	76	80	89	94	117	111	108	112	135	144
14	37	26	13	12		11	39	49	73	65	71	78	83	106	100	96	100	124	133
18	42	26	21	17	11		34	42	69	61	64	74	78	102	96	92	96	120	128
3	71	54	47	50	39	34		30	43	49	48	52	44	67	61	57	61	86	94
7	83	64	59	58	49	42	30		61	47	43	31	36	59	56	63	65	95	93
5	91	66	81	85	73	69	43	61		30	35	52	47	58	34	39	43	66	61
15	82	57	72	76	65	61	49	47	30		12	36	44	59	45	52	56	79	80
13	86	61	76	80	71	64	48	43	35	12		25	36	52	42	47	51	80	78
9	96	76	86	89	78	74	52	31	52	36	25		12	38	41	51	47	74	72
1	100	81	90	94	83	78	66	36	47	44	36	12		32	37	49	45	69	68
4	124	99	114	117	106	102	67	59	58	59	52	38	32		28	30	27	39	41
8	118	93	108	111	100	96	61	56	34	45	42	41	37	28		26	17	41	37
11	114	89	104	108	96	92	57	63	39	52	47	51	49	30	26		16	36	37
19	118	93	108	112	100	96	61	65	43	56	51	47	45	27	17	16		29	35
17	141	119	132	135	124	120	86	95	66	79	80	74	69	39	41	38	29		24
6	150	126	140	144	133	128	94	93	61	80	78	72	68	41	37	37	35	24	



Ryc. 1. Przykład zmodyfikowanego (rachunkowego) sposobu porządkowania tablicy różnic; a—tablica różnic wstępnie uporządkowana, b—gradienty w tablicy różnic

An example of modified (calculated) method of arranging tables of differences; a—table of differences preliminary set in order, b—gradients in the table of differences

kumulacyjnego³ z sum kolejnych szeregów. Za poprawniejsze uznajemy to uporządkowanie, którego suma kolejnych wartości szeregu kumulacyjnego jest mniejsza.

Jeżeli przez P oznaczymy wartości szeregu kumulacyjnego, to:

$$P_1 = S_1$$

$$P_2 = S_1 + S_2$$

$$P_3 = S_1 + S_2 + S_3$$

$$P_{n-2} = S_1 + S_2 + S_3 + \dots + S_{n-2}$$

$$P_{n-1} = S_1 + S_2 + S_3 + \dots + S_{n-2} + S_{n-1}$$

$$\text{natomiast } \sum_{j=1}^{n-1} P_j = (n-1)S_1 + (n-2)S_2 + (n-3)S_3 + \dots + 2S_{n-2} + S_{n-1}$$

a wobec tego wzór ogólny przyjmuje następującą postać:

$$P = \sum_{j=1}^{n-1} P_j = \sum_{j=1}^{n-1} (n-j) S_j$$

Na podstawie obliczeń według wyżej przyjętej zasady bardzo łatwo możemy ustalić, który z dwu wariantów daje lepsze uporządkowanie: pierwszy, jak w tab. a ryc. 1, gdzie interesujące nas powiaty występują w kolejności..., Janów, Chełm, Opole, Kraśnik, Bełżyce, Bychawa..., czy też drugi — gdzie te same powiaty ułożone są w zmienionej kolejności..., Janów, Kraśnik, Bełżyce, Opole, Puławy, Chełm, Bychawa.

Zgodnie z ustalonym wzorem dla wariantu pierwszego otrzymujemy $P' = 11\,586,72$, dla wariantu drugiego $P'' = 11\,622,21$. Rozstrzygnięcie, które z dwu uporządkowań jest lepsze, następuje w myśl uprzednio ustalonej reguły. Ponieważ $P' < P''$, wobec tego wariant pierwszy uznajemy za lepszy⁴.

W oparciu o tak uporządkowaną macierz odległości (tablicy różnic) można bezpośrednio przystąpić do podziału badanych jednostek na grupy, wykorzystując do tego celu gradienty różnic. W prezentowanym przykładzie (ryc. 1 b) zaznaczono gradienty o wartości różnicy ponad

³ Można również obliczyć, ile punktów wypada przeciętnie na jedno pole dla kolejnych szeregów, albo też wykreślić krzywą koncentracji i obliczyć wskaźnik koncentracji. Poprzestajemy jednak tylko na obliczaniu sum z szeregów kumulacyjnych, ponieważ są one wystarczającym kryterium poprawności uporządkowania, pozwalającego na wydzielenie istniejących grup (regionów).

⁴ Nie mamy tu jednak zupełnej pewności, że jest to rzeczywiście najlepsze z możliwych uporządkowań. Znalazienie takiego optymalnego uporządkowania wymaga zastosowania maszyny cyfrowej. Aby jednak praca maszyny była ekonomiczna, konieczne jest znalezienie algorytmu pozwalającego na szybkie porządkowanie.

2,00, które naszym zdaniem zupełnie poprawnie i przekonywująco dzielią badane jednostki na grupy.

Wreszcie wykreślenie diagramu (ryc. 4 a) posłużyło do interpretacji stopni podobieństw pomiędzy jednostkami różnych regionów, jak również wewnątrz samych regionów.

DENDRYT WROCŁAWSKI⁵

Metoda dendrytowa, podobnie jak metoda Czekanowskiego, wychodzi z definicji typu jako miejsca skupienia. Celem jej jest z jednej strony porządkowanie jednostek ze względu na określony zespół cech, z drugiej — podział badanej zbiorowości na pewną liczbę typów.

Dendryt jest graficznym obrazem przedstawiającym na płaszczyźnie najmniejsze odległości między badanymi jednostkami, rozumianymi jako punkty wielowymiarowej przestrzeni, których współrzędnymi są ich cechy. Według definicji (6, s. 195) „przez dendryt rozumiemy tu łamaną, która może się rozgałęziać, lecz nie może zawierać łamanych zamkniętych, i taką, że każde dwa punkty zbioru Z są przez nią połączone”.

Ponieważ jednak takich połączeń może być wiele, wprowadza się dodatkowo pojęcie tzw. n a j k r ó t s z e g o u p o r z ą d k o w a n i a (tak aby suma długości boków dendrytu była najmniejsza), polegającego na tym, że poszczególne jednostki łączy się według najmniejszych odległości.

W celu opracowania dendrytu z poszczególnych kolumn macierzy odległości (tablicy różnic) wyszukujemy wartości minimalne i podajemy numer wiersza, w którym wartość minimalna wystąpiła (z wyjątkiem wartości zerowej). Wartość minimalna wskazuje na odległość dwu jed-

⁵ Podstawy teoretyczne dendrytu wrocławskiego znaleźć można w pracach 6, 18, 19. Jak podaje J. Perkal, metoda dendrytowa znalazła zastosowanie w różnych dziedzinach nauk filologicznych, przyrodniczych i technicznych.

Przykładowo można wymienić kilka interesujących prac z dziedziny botaniki i zootechniki. W botanice dendryt stosowali między innymi: Matuszkiewicz W., Borowik M.: Materiały do systematyki fitosocjologicznej lasów łęgowych w Polsce, *Acta Societatis Botanicorum Poloniae* 25 (1956), z. 2; Faliński J. B.: Zastosowanie taksonomii wrocławskiej do fitosocjologii, *Acta Societatis Botanicorum Poloniae* 27 (1958), z. 1; Kowal T.: Zasady i przykłady systematyki roślin metodą dendrytową, *Prace Wrocławskiego Towarzystwa Naukowego*, seria B, nr 117, Wrocław 1965.

Natomiast w zootechnice dendryt stosowali: Janowski T.: Próba zastosowania taksonomii wrocławskiej, *Medycyna Weterynaryjna*, nr 3, 1955; Nowicki B.: Charakterystyka bydła nizinnego czarno-białego w regionie Wałbrzyskim, *Zeszyty Naukowe WSR we Wrocławiu*, Zoot. II, Wrocław 1956; Borowski B.: Próby zastosowania taksonomii wrocławskiej do badań biometrycznych nad kozami, *Zeszyty Naukowe WSR we Wrocławiu*, Zootechnika I, Wrocław 1956.

nostek, a numer wiersza pokazuje numer jednostki, z jaką należy połączyć jednostkę odpowiadającą danej kolumnie.

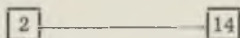
I tak w kolumnie pierwszej (tab. 1) najmniejsza wartość (1,20) znajduje się w wierszu dziewiątym, co oznacza, że jednostki 1 i 9 należy połączyć.



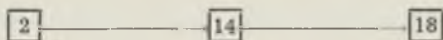
Następnie szukamy najmniejszej odległości w kolumnie jednostki dziewiątej. Okazuje się, że jednostka dziewiąta wykazuje najmniejszą odległość (1,20) z jednostką pierwszą. Wobec tego nie kreślimy już nowego połączenia, a zaznaczamy tylko, że tworzą one wzajemną parę.



Podobnie postępujemy z następnymi jednostkami w kolejnych kolumnach. Druga jednostka wykazuje najmniejszą odległość z czternastą. Jednostki 2 i 14 łączymy.



Z kolei jednostka czternasta wykazuje najmniejszą odległość z osiemnastą, wobec czego jednostki 14 i 18 łączymy, otrzymując połączenie trzech jednostek.

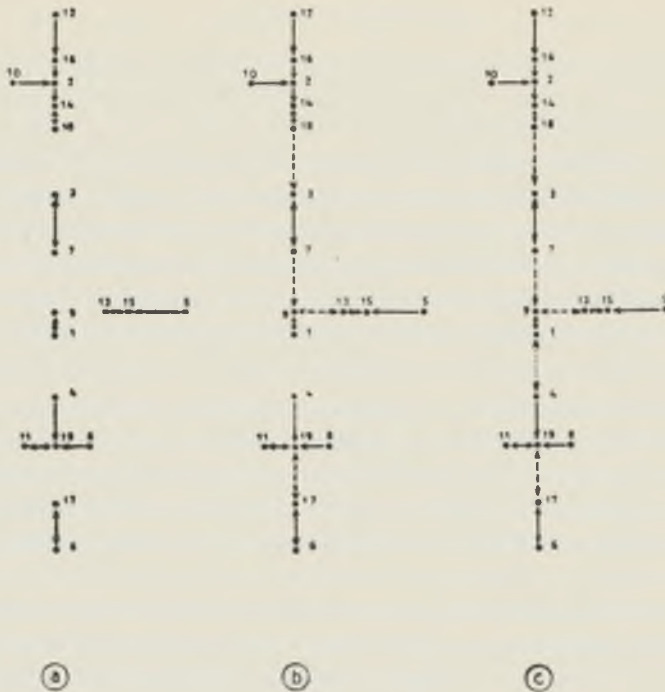


Proces ten kontynuujemy dalej i prowadzimy dotąd, aż wszystkie kolumny (jednostki) zostaną uwzględnione. Połączenia te utworzą pewne grupy (dendryty). Tak na przykład jednostka dziesiąta zostanie połączona z jednostką drugą, a jednostka dwunasta połączy się z szesnastą, ta zaś z kolei uzyska połączenie z drugą, tworząc w ten sposób grupę składającą się z jednostek: 18, 14, 2, 12 i 10 (ryc. 2a).

Jak się okazuje, po wyczerpaniu możliwych połączeń nie zdołaliśmy utworzyć jednej grupy. W tym przypadku musimy dalej kontynuować proces łączenia, szukając najbliższej odległości między poszczególnymi grupami (z pominięciem odległości pomiędzy jednostkami danej grupy). Na przykład: nr 1 łączy z nr 4 odległość 3,16; nr 9 łączy z nr 13 odległość 2,54. Oznacza to, że grupa jednostek (1, 9) połączy się z inną grupą poprzez jednostkę nr 13 przy pomocy odległości 2,54 (ryc. 2b).

Jeżeli w drugim etapie grupowania wszystkie dendryty nie połączą się, proces grupowania kontynuujemy etapami dalej (identycznie jak w etapie drugim), aż do uzyskania jednej grupy — pełnego dendrytu.

Przy interpretacji dendrytu odróżnienie połączeń różnych stopni nie jest istotne — jest nam potrzebne tylko do konstrukcji dendrytu. Podobnie nie ma znaczenia sposób ułożenia dendrytu. Natomiast istotna jest kolejność ułożonych jednostek i odległości między nimi. Odcinki



Ryc. 2. Etapy porządkowania dendrytowego; a — etap pierwszy, b — etap drugi, c — etap trzeci

Stages of dendritic arrangement; a — first stage, b — second stage, c — third stage

najkrótsze łączą bowiem jednostki najbardziej podobne, natomiast odcinki najdłuższe łączą jednostki (grupy jednostek) różniące się między sobą. Aby uzyskać odpowiednie grupy jednostek podobnych, w praktyce postępujemy w ten sposób, że odrzucamy kolejno odcinki najdłuższe.

W prezentowanym przykładzie (ryc. 2c) po odrzuceniu odcinka najdłuższego oddzieli się od dendrytu grupa jednostek (są to powiaty: 18-Włodawa, 14-Parczew, 2-Biała Podlaska, 16-Radzyń Podlaski, 12-Łuków, 10-Lubartów), które — w tym samym zespole cech — również przy zastosowaniu innych metod stanowią dosyć zwartą grupę. W drugim i w następnych etapach będą oddzielały się grupy powiatów lub pojedyncze powiaty: na przykład powiat chełmski po odrzuceniu czwartego z kolei odcinka.

Oczywiście istnieje uzasadniona granica dalszego dzielenia dendrytu. Pewne wskazówki mające na celu otrzymanie optymalnej ilości grup podają autorzy metody. Wydaje się jednak, że decydujące znaczenie mieć tu powinna analiza merytoryczna wyodrębnionych grup.

DRZEWO POŁĄCZEŃ⁶

Zależnie od ostrości kryterium przynależności do grupy stawiane są zasadniczo trzy definicje grupy, które realizowane są przy pomocy nieco odmiennych metod klasyfikacji.

Pierwsza z nich mówi, że „jednostka która ma być zaliczona do grupy musi być bliższa jakiemuś członkowi tej grupy aniżeli jakiemukolwiek członkowi innej grupy lub pozostałym jednostkom”. W myśl tej definicji wszystkie jednostki zostają zgrupowane ze swoimi najbliższymi sąsiadami (podobnie jak to było w przypadku dendrytu) bez względu na dzielące je różnice. Zbyt łagodne kryterium przynależności powoduje, że w grupie mogą znaleźć się jednostki bardzo zróżnicowane.

Przeciwieństwem pierwszej jest inna definicja, mówiąca, że „jednostka, która ma być zaliczona do grupy, powinna być bliższa wszystkim członkom tej grupy niż jakiemukolwiek członkowi innej grupy lub pozostałym jednostkom”. Jest to z kolei zbyt ostre kryterium przynależności, powodujące, że bardzo wielu jednostek nie można włączyć do żadnej z grup.

W tym świetle najsluszniejsza wydaje się być trzecia — kompromisowa — definicja mówiąca, że „jednostka która ma być zaliczona do grupy powinna być bliższa do jakiegoś punktu odniesienia dla danej grupy aniżeli do jakiegokolwiek punktu odniesienia innej grupy lub pozostałych jednostek”. Bliższego sprecyzowania wymaga tylko sam punkt odniesienia.

Spośród proponowanych za McQuitty'm punktów odniesienia (odległość minimalna, odległość maksymalna i centroid) za najpoprawniejszy w naszym przypadku uznać należy centroid⁷. Może on być obliczany jako średnia odległość między wszystkimi jednostkami grupy a daną jednostką (lub wszystkimi jednostkami danej grupy) albo też między punktem środkowym grupy a jednostką (lub punktem środkowym innej grupy). Przy posługiwaniu się centroidem odległość użyta między jednostką i grupą jest punktem środkowym danej grupy, a w przypadku dwu grup odległość obliczana jest między dwoma centroidami.

Postępując zgodnie z przedstawioną definicją i przyjętym punktem odniesienia, w pierwszym etapie klasyfikacji na ogół nie osiągamy

⁶ Zagadnieniom metodycznym związanym z problemem klasyfikacji przy pomocy drzewa połączeń, poza pozycjami wymienionymi w spisie literatury (1, 12, 16), szereg prac zamieszczonych w „Educational and Psychological Measurement” poświęcił McQuitty.

⁷ Należy pamiętać, że każda jednostka terytorialna jest właśnie takim centroidem dla poszczególnych jej części, a różni się od obecnie tworzonych centroidów tym, że nie posiada określonego stopnia jednorodności. Stąd przyjęcie centroidu jako punktu odniesienia stanowi jak najbardziej naturalny sposób postępowania.

pożądaney liczby grup, szczególnie gdy badane jednostki tworzą dużą liczbę par wzajemnych. Dlatego dalsze grupowanie uzyskanych grup i jednostek po pierwszym etapie jest konieczne. Grupowanie to osiągamy przez zbudowanie nowej macierzy między jednostkami i grupami pozostałymi po pierwszym etapie — w następnych etapach.

Mamy tu do wyboru dwie drogi postępowania: albo odrywamy wszystkie wzajemne pary jednostek od macierzy w jednym etapie, albo też łączymy tylko najbardziej podobną parę i obliczamy nową macierz odległości (tablicę różnic). I chociaż jest to powolniejszy proces łączenia jednostek w grupy, w niniejszej pracy został przyjęty ten drugi sposób postępowania.

Grupując jednostki zgodnie z przyjętymi założeniami, zachowujemy w poszczególnych etapach maksymalną jednolitość grup. W pierwszym etapie poprzez połączenie pary najbardziej podobnych jednostek i zastąpienie jej centroidem (średnią odległością) sprowadzamy macierz odległości z postaci $n \times n$ do postaci $(n-1) \times (n-1)$. W drugim etapie łączymy następne dwie najbardziej podobne jednostki lub jednostkę i grupę jednostek (centroid), a po połączeniu otrzymujemy macierz $(n-2) \times (n-2)$. Proces połączeń prowadzimy dalej, aż otrzymamy macierz 2×2 . W ostatnim etapie zawsze otrzymujemy jedną grupę.

Przykład praktycznego obliczenia zostanie przedstawiony tylko dla części jednostek z prezentowanej już macierzy odległości (tablicy różnic). Odległości (różnice) między wybranymi jednostkami przedstawia tablica:

	A	B	C	D	E	F	G
A	0	5,78	2,84	9,88	2,98	10,26	10,15
B	5,78	0	3,37	6,60	3,89	7,34	6,88
C	2,84	3,37	0	9,27	2,57	10,01	9,55
D	9,88	6,60	9,27	0	8,90	2,62	2,59
E	2,98	3,89	2,57	8,90	0	9,64	9,18
F	10,62	7,34	10,01	2,62	9,64	0	1,10
G	10,15	6,88	9,55	2,59	9,18	1,10	0
	(C)	(C)	(E)	(G)	(C)	(G)	(F)
	2,84	3,37	2,57	2,59	2,57	1,10	1,10

gdzie: 1) litery oznaczają wybrane powiaty; A — Bychawa, B — Chełm, C — Kraśnostaw, D — Lubartów, E — Lublin, F — Parczew, G — Włodawa; 2) wartości liczbowe pod tablicą dotyczą najmniejszych odległości między jednostką danej kolumny a inną jednostką, której symbol podano w nawiasie.

W pierwszym etapie łączymy dwie najbardziej podobne jednostki F i G (1,10) w jedną grupę. Obliczamy dla niej średnie wartości z kolumn i wierszy F i G. Zastępujemy wartości kolumn i wierszy F, G przez przeciętną wartość F—G, przedstawiającą odległości między tą parą jednostek a wszystkimi pozostałymi (centroid). Otrzymujemy nową macierz odległości 6×6 .

	F—G	A	B	C	D	E
F—G	0	10,38	7,11	9,78	2,60	9,41
A	10,38	0	5,78	2,84	9,88	2,98
B	7,11	5,78	0	3,37	6,60	3,89
C	9,78	2,84	3,37	0	9,27	2,57
D	2,60	9,88	6,60	9,27	0	8,90
E	9,41	2,98	3,89	2,57	8,90	0
	(D)	(C)	(C)	(E)	(F—G)	(C)
	2,60	2,84	3,37	2,57	2,60	2,57

W etapie drugim postępujemy identycznie, z tym tylko, że jako najbardziej podobne (2,57) połączone zostają jednostki C i E, a macierz odległości zmniejsza się do 5×5 .

	C—E	F—G	A	B	D
C—E	0	9,59	2,91	3,63	9,08
F—G	9,59	0	10,38	7,11	2,60
A	2,91	10,38	0	5,78	9,88
B	3,63	7,11	5,78	0	6,60
D	9,08	2,60	9,88	6,60	0
	(A)	(D)	(C—E)	(C—E)	(F—G)
	2,91	2,60	2,91	3,63	2,60

W etapie trzecim dochodzi do interesującego połączenia jednostki D z grupą jednostek — centroidem⁸ F—G (2,60) i dalszego zmniejszenia macierzy do 4×4 .

⁸ Pojawia się tu zagadnienie wag. Mianowicie błędem byłoby traktowanie centroidu F—G i jednostki D jako równorzędnych składowych i obliczanie z nich średniej, jak to było w poprzednich przykładach. Musimy tu obliczyć wartości nowego centroidu przy pomocy średniej arytmetycznej ważonej $x = \frac{\sum xn}{\sum n}$, tak jak

	F—G—D	C—E	A	B
F—G—D	0	9,42	10,21	6,94
C—E	9,42	0	2,91	3,63
A	10,21	2,91	0	5,78
B	6,94	3,63	5,78	0
	(E)	(A)	(C—E)	(C—E)
	6,94	2,91	2,91	3,63

W następnych etapach kolejno połączone zostają centroid C—E z jednostką A, następnie nowy centroid C—E—A z jednostką B, tak że w ostateczności tworzą się dwie grupy jednostek F—G—D i C—E—A—B.

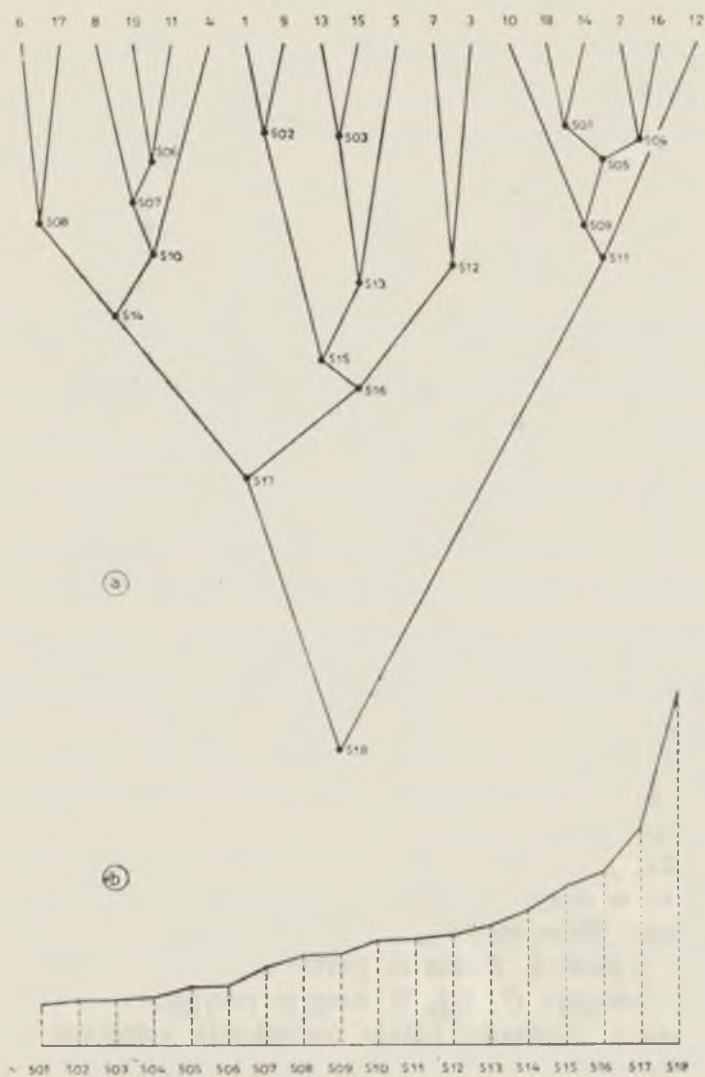
Na podstawie tego typu obliczeń dla wszystkich jednostek tab. 1 zostało sporządzone „drzewo połączeń” z zachowaniem poszczególnych etapów połączeń i odległości, za pomocą których były łączone jednostki i grupy jednostek (ryc. 3). W oparciu o sporządzone drzewo połączeń z łatwością możemy podzielić badane jednostki na grupy, zaczynając ten podział w kierunku odwrotnym (od końca) do kierunku łączenia badanych jednostek w grupy.

Jeżeli zechcemy podzielić badane jednostki na dwie grupy, to cofniemy się do macierzy 2×2 , a na rysunku drzewa połączeń odrzucimy odcinki łączące punkty 511 i 517 w etapie osiemnastym. Jeżeli zaś chcemy badane jednostki podzielić na trzy grupy, to cofniemy się do macierzy 3×3 , a w drzewie połączeń odrzucimy odcinki łączące punkty 514 i 516 w etapie siedemnastym itd.

Pytanie tylko, jak daleko należy prowadzić ten proces podziału na części. Zależy to w dużym stopniu od pożądanego uogólnienia, a także od momentu, przy jakim etapie generalizacji pojawiają się zwarte terytorialnie grupy jednostek. Znane są próby unifikowania oceny dopuszczalnej granicy podziału (1, 23). W naszym przypadku zostały w tym celu wykorzystane gradienty różnic (odległości) kolejnych połączeń. Na podstawie kwadratów tych różnic został obliczony procent uogólnienia dla poszczególnych etapów połączeń, podobnie jak to czynili Berry, Tarrant i inni. W oparciu o te dane została wykreślona krzywa koncentracji.

Operując interpretacją gradientu różnic, wskaźnikiem procentu uogólnienia i wykresem krzywej koncentracji, możemy stwierdzić, że podział badanych jednostek na trzy grupy jest jak najbardziej uzasadniony.

gdybyśmy ją obliczali z poszczególnych wartości F, D, G. Rozwiązanie to dotyczy wyłącznie prezentowanego przykładu i jest dopuszczalne tylko w przypadku, gdy mamy do czynienia z jednostkami o zbliżonej wielkości.



Ryc. 3. Schemat porządkowania powiatów przy pomocy „drzewa połączeń”; a — drzewo połączeń, b — kolejne różnice
 Pattern of arrangement of counties by means of "linkage tree"; a — linkage tree, b — successive differences

Natomiast mniej pewne jest wydzielenie dalszych grup — maksymalnie do etapu trzynastego.

WSTĘPNA OCENA PRZYDATNOŚCI ZASTOSOWANYCH METOD

Prezentowane metody (diagram, dendryt i drzewo połączeń) mają pewne cechy wspólne i wiele różnic. Wszystkie obciążone są jedną wspólną „właściwością”, tkwiącą w samej tablicy różnic (odległości), — w tym sensie, w jakim możliwe jest przedstawienie wszystkich różnic między cechami w postaci jednej wspólnej różnicy (odległości). Chodzi o to, że taka sama wspólna różnica może być wynikiem różnic z odmiennych kombinacji cech. Z tego też powodu niekiedy celowe będzie nie poprzestawanie na różnicach wspólnych (przeciętnych), a sięganie do materiału wyjściowego (wartości poszczególnych cech).

Zarówno diagram, jak i dendryt traktują wszystkie jednostki od początku do końca indywidualnie. Ale diagram porządkuje poszczególne jednostki liniowo (rzutuje punkty wielowymiarowej przestrzeni na prostą), natomiast dendryt — inaczej (rzutuje punkty n wymiarowej przestrzeni na płaszczyznę). W tym punkcie widoczna jest przewaga uporządkowania dendrytowego.

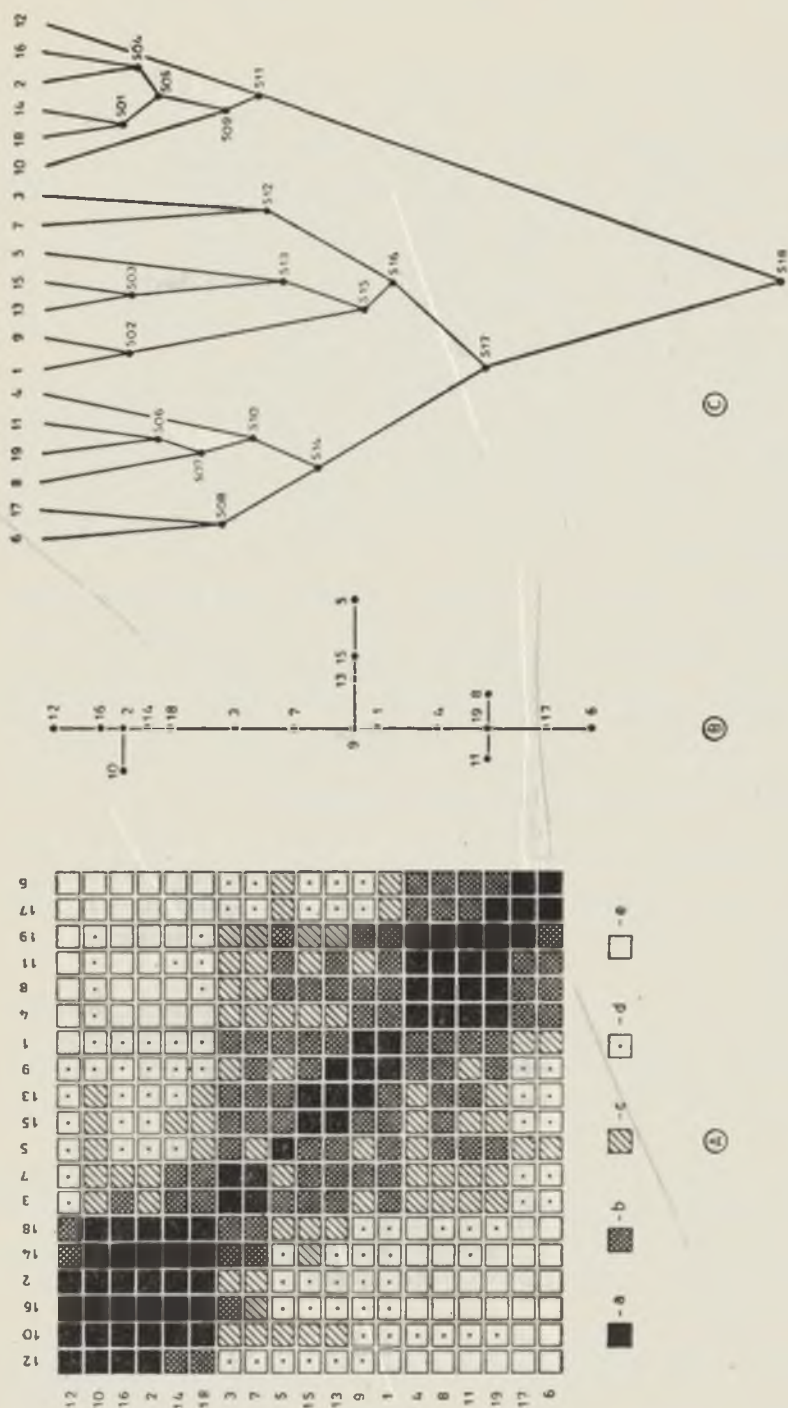
Diagram jednak odtwarza odległości (różnice) między jedną jednostką a wszystkimi pozostałymi — dzięki czemu stwarza lepsze warunki do interpretacji powiązań — natomiast dendryt uwzględnia tylko najmniejsze odległości pomiędzy jednostkami. Jeżeli zatem te najmniejsze odległości nie różnią się w sposób istotny od pozostałych (a tak jest bardzo często), uporządkowanie dendrytowe nie może stanowić dostatecznie pewnej podstawy grupowania badanych jednostek.

Podobnie jak w przypadku dendrytu, również przy konstrukcji drzewa połączeń uwzględnia się tylko najbliższe odległości (połączenia). Ale dla dendrytu tych najbliższych połączeń może być tylko $n - 1$, natomiast dla drzewa połączeń minimum $2n - 1$, a więc przynajmniej dwukrotnie więcej.

Jednak najbardziej istotna różnica pomiędzy dendrytem a drzewem połączeń polega na tym, że podczas gdy dendryt łączy poszczególne jednostki pojedynczo (każdą z osobna), to drzewo połączeń dla jednostek najbardziej podobnych tworzy centroidy (średnie), do których, jak do jakiegoś ośrodka krystalizacyjnego, przyciągane są jednostki podobne.

Obrazowo działanie metody drzewa połączeń można by przyrównać do łączenia jednostek na zasadzie terasowania (zrównywania jednostek najbardziej podobnych przez tworzenie centroidów — średnich), natomiast działanie metody dendrytowej — do łączenia jednostek (punktów) na zasadzie serpentyny.

Ocena przydatności poszczególnych metod może być różna, zależnie od charakteru badanej zbiorowości. I tak w stosunku do prezentowanego przykładu regionów 6 głównych upraw (ryc. 4) wszystkie trzy



Ryc. 4. Uporządkowanie powiatów według: A — diagramu Czekanowskiego (gdzie odpowiednie sygnatury oznaczają różnice: a) 0—3, b) 3—5, c) 5—7, d) 7—10, e) powyżej 10); B — dendrytu wrocławskiego, C — drzewa połączeń Arranging of counties after: A — Czekanowski's diagram (where suitable the symbols indicate the differences: a) 0—3, b) 3—5, c) 5—7, d) 7—10, e) more than 10); B — the Wrocław dendrite; C — linkage tree

metody w sposób wystarczający spełniają swoje zadanie. Otrzymany podział jest klarowny i uzasadniony analizą poszczególnych cech. Zasadniczo wszystkie trzy metody ujawniają podział na trzy podstawowe grupy. Grupę pierwszą stanowią powiaty: 12-lukowski, 10-lubartowski, 16-radzyński, 2-bialski, 14-parczewski i 18-włodawski. W grupie drugiej znajdują się powiaty: 6-hrubieszowski, 17-tomaszowski, 19-zamojski, 11-lubelski, 8-krasnostawski i 4-bychawski. W grupie tej powiaty hrubieszowski i tomaszowski stanowią wyraźną podgrupę, natomiast pozostałe powiaty wykazują więcej powiązań z grupą trzecią. Grupa trzecia, do której można zaliczyć powiaty: 1-bełżycki, 9-kraśnicki, 13-opolski, 15-puławski, 5-chełmski, 7-janowski i 3-biłgorajski, jest wewnętrznie najbardziej zróżnicowana. Część z tych powiatów (biłgorajski, janowski) wykazuje silne powiązania z grupą pierwszą, inne (bełżycki, kraśnicki, chełmski) z grupą drugą. Tak więc na podstawie wszystkich metod, zależnie od pożądanego stopnia generalizacji, powiaty mogą być podzielone na dwie, trzy lub pięć oddzielnych grup.

Sprawa komplikuje się jednak bardziej, gdy mamy do czynienia ze zbiorowością bardziej liczną, posiadającą poza określonymi typami także jednostki o charakterze przejściowym. Taki typ zbiorowości niewątpliwie reprezentuje przykład regionalizacji województwa w oparciu o gromady.

W tym przypadku posługiwanie się diagramem — zakładając nawet możliwość poprawnego uporządkowania przy pomocy maszyny cyfrowej, w myśl poprzednio przedstawionych założeń — mogłoby napotkać na trudności związane z podziałem tego diagramu na części. Nie wydaje się zresztą, aby posługiwanie się diagramem dla zbyt dużej liczby jednostek było celowe także z punktu widzenia możliwości interpretacji diagramu. Zastosowanie dendrytu jako metody niewątpliwie najprostszej nie dało niestety dostatecznie przekonujących wyników (ryc. 8). Okazuje się bowiem, że dendryt nie jest tak precyzyjnym narzędziem badawczym, jak to sądzili twórcy tej metody (6), a otrzymane uporządkowanie bywa często, niestety, dosyć przypadkowe. W tej sytuacji za jedyne możliwe do przyjęcia rozwiązanie uznano zastosowanie „drzewa połączeń”. Nie oznacza to jednak, aby drzewo połączeń w każdej sytuacji bezbłędnie rozwiązywało problem grupowania. Przykłady niedoskonałości tej metody były na tyle przekonujące, aby dojść do wniosku, że tylko przy jednoczesnym wykorzystaniu także pozostałych metod można będzie postawiony problem rozwiązać.

PROPONOWANA TECHNIKA OBLICZEŃ I KONSTRUKCJI
DRZEWA POŁĄCZEŃ

Zagadnienie porządkowania jest dosyć proste, gdy idzie o małą liczbę jednostek. Ale w tym przypadku chodzi o znalezienie rozwiązania dla czterystu siedemnastu jednostek. Z powodów, które już zostały omówione, należało zrezygnować z porządkowania przy pomocy diagramu Czekanowskiego, a również dendryt wrocławski okazał się niewystarczającą podstawą klasyfikacji. Pozostawało właściwie do dyspozycji tylko drzewo połączeń. Ale i tutaj wyłoniły się trudności natury technicznej. Mianowicie symetryczna macierz odległości dla 417 jednostek posiada 174 306 liczb, które należałoby jednocześnie umieścić w pamięci maszyny cyfrowej. Jeżeli nawet weźmiemy pod uwagę fakt, że można by umieścić tylko połowę macierzy, to i tak mielibyśmy 87 153 liczby. Natomiast maszyna, z której mogłem korzystać, „Odra 1013”, posiada pamięć zaledwie 8 192 słów maszynowych. W tej sytuacji byłem zmuszo-

Tab. 1. Tablica przeciętnych różnic,
Table of mean differences in six

		1	2	3	4	5	6	7
Bełżyce	1	0	9,37	4,37	3,16	4,65	6,76	3,55
Biała Podl.	2	9,37	0	5,02	11,73	8,46	14,41	5,83
Biłgoraj	3	4,37	5,02	0	6,72	4,33	9,39	2,98
Bychawa	4	3,16	11,73	6,72	0	5,78	4,14	5,92
Chełm	5	4,65	8,46	4,33	5,78	0	6,10	6,13
Hrubieszów	6	6,76	14,41	9,39	4,14	6,10	0	9,26
Janów Lub.	7	3,55	5,83	2,98	5,92	6,13	9,26	0
Krasnystaw	8	3,65	11,13	6,11	2,84	3,37	3,70	5,56
Kraśnik	9	1,20	8,93	5,16	3,77	5,18	7,23	3,12
Lubartów	10	8,05	2,22	5,35	9,88	6,60	12,55	6,36
Lublin	11	4,90	10,76	5,74	2,98	3,89	3,65	6,26
Łuków	12	9,99	2,51	7,07	12,35	9,08	15,03	8,31
Opole Lub.	13	3,57	7,98	4,84	5,18	3,48	7,82	4,25
Parczew	14	8,25	1,22	3,90	10,62	7,34	13,29	4,87
Puławy	15	4,41	7,55	4,91	5,88	3,04	7,99	4,70
Radzyń Podl.	16	9,00	1,29	4,70	11,37	8,09	14,04	5,93
Tomaszów Lub.	17	6,95	13,52	8,55	3,93	6,65	2,42	9,45
Włodawa	18	7,79	1,68	3,44	10,15	6,88	12,83	4,24
Zamość	19	4,48	11,15	6,13	2,68	4,28	3,48	6,52
	W a g a	3,6	8,8	6,1	4,0	9,4	10,7	3,2

ny opracować taki sposób obliczeń, który nie uwzględniałby konieczności jednoczesnego wprowadzania całej macierzy do pamięci maszyny.

Możliwość takiego rozwiązania ma ogólniejsze znaczenie, ponieważ pozwala na zastosowanie metody bez krępowania się liczbą badanych jednostek. Pozwala to, najogólniej biorąc, na posługiwanie się w badaniach geograficznych jednostkami o najdogodniejszej wielkości. Możliwości poprzednio prezentowanej metody obliczeń dotyczyły raczej tylko podziałów pierwszego rzędu, a więc podziału Polski w oparciu o województwa, podziału województwa w oparciu o powiaty itd. Natomiast w tym przypadku chodzi o zastosowanie metody obliczeń umożliwiającej dokonywanie podziałów w oparciu o jednostki drugiego rzędu. Na przykład podziału województwa w oparciu o gromady, Polski w oparciu o powiaty itd. A zatem liczba jednostek przeciętnie wzrasta z kilkunastu lub kilkudziesięciu do kilkuset. W związku z tym wzrasta skala trudności, wyłaniają się przy tym inne problemy i to zarówno natury

6 głównych upraw, 1965 (powiaty)

basic crops, for 1965 (by counties)

8	9	10	11	12	13	14	15	16	17	18	19
3,65	1,20	8,05	4,90	9,99	3,57	8,25	4,41	9,00	6,95	7,79	4,48
11,13	8,93	2,22	10,76	2,51	7,98	1,22	7,55	1,29	13,52	1,68	11,15
6,11	5,16	5,35	5,74	7,07	4,84	3,90	4,91	4,70	8,55	3,44	6,13
2,84	3,77	9,88	2,98	12,35	5,18	10,62	5,88	11,37	3,93	10,15	2,68
3,37	5,18	6,60	3,89	9,08	3,48	7,34	3,04	8,09	6,65	6,88	4,28
3,70	7,23	12,55	3,65	15,03	7,82	13,29	7,99	14,04	2,42	12,83	3,48
5,56	3,12	6,36	6,26	8,31	4,25	4,87	4,70	5,93	9,45	4,24	6,52
0	4,14	9,27	2,58	11,75	4,21	10,01	4,51	10,76	4,11	9,55	1,74
4,14	0	7,64	5,11	9,58	2,54	7,82	3,56	8,57	7,41	7,35	4,71
9,27	7,64	0	8,90	2,48	6,12	2,62	5,69	2,53	11,93	2,59	9,29
2,58	5,11	8,90	0	11,38	4,73	9,64	5,22	10,39	3,64	9,18	1,58
11,75	9,58	2,48	11,38	0	8,60	3,67	8,17	2,41	14,14	4,15	11,77
4,21	2,54	6,12	4,73	8,60	0	7,13	1,23	7,63	8,00	6,40	5,11
10,01	7,82	2,62	9,64	3,67	7,13	0	6,50	1,34	12,40	1,10	10,03
4,51	3,56	5,69	5,22	8,17	1,23	6,50	0	7,18	7,94	6,12	5,61
10,76	8,57	2,53	10,39	2,41	7,63	1,34	7,18	0	13,15	2,07	10,78
4,11	7,41	11,93	3,64	14,14	8,00	12,40	7,94	13,15	0	11,97	2,92
9,55	7,35	2,59	9,18	4,15	6,40	1,10	6,12	2,07	11,97	0	9,57
1,74	4,71	9,29	1,58	11,77	5,11	10,03	5,61	10,78	2,92	9,57	0
9,5	6,7	6,6	7,8	8,7	3,7	3,9	5,6	7,3	8,8	5,0	8,1

Tab. 2. Przykład macierzy z etapu /5/ 17: a—w
 Example of a matrix from the stage /5/ 17: a—

a)	1	2	3	4	5	6	7	8	9
1		9,52							
2	9,52	0	9,52	9,52	9,52	9,52	9,52	9,52	9,52
3		9,52							
4		9,52							
5		9,52							
6		9,52							
7		9,52							
8		9,52							
9		9,52							
10	9,52	0	9,52	9,52	9,52	9,52	9,52	9,52	9,52
11		9,52							
12	9,52	0	9,52	9,52	9,52	9,52	9,52	9,52	9,52
13		9,52							
14	9,52	0	9,52	9,52	9,52	9,52	9,52	9,52	9,52
15		9,52							
16	9,52	0	9,52	9,52	9,52	9,52	9,52	9,52	9,52
17		9,52							
18	9,52	0	9,52	9,52	9,52	9,52	9,52	9,52	9,52
19		9,52							

b)	511	517
511	0	9,52
517	9,52	0

technicznej jak i merytorycznej, których przy operowaniu małą liczbą jednostek nie było.

Przykład praktycznego zastosowania proponowanego sposobu obliczeń zostanie przedstawiony na tym samym materiale, na którym były prezentowane poprzednie rozwiązania.

Spróbujmy bliżej przyjrzeć się macierzy odległości odnoszącej się do przedostatniego etapu połączeń /5/17 w ujęciu centroidalnym (tab. 2b). Pokazuje ona odległości pomiędzy dwoma centroidami 511 i 517. Jak wiemy, jest to odległość między środkami obu grup (centroidami) albo

układzie jednostkowym, b — w układzie centroidalnym
in unit pattern, b — in centroidal pattern

10	11	12	13	14	15	16	17	18	19
9,52		9,52		9,52		9,52		9,52	
0	9,52	0	9,52	0	9,52	0	9,52	0	9,52
9,52		9,52		9,52		9,52		9,52	
9,52		9,52		9,52		9,52		9,52	
9,52		9,52		9,52		9,52		9,52	
9,52		9,52		9,52		9,52		9,52	
9,52		9,52		9,52		9,52		9,52	
9,52		9,52		9,52		9,52		9,52	
0	9,52	0	9,52	0	9,52	0	9,52	0	9,52
9,52		9,52		9,52		9,52		9,52	
0	9,52	0	9,52	0	9,52	0	9,52	0	9,52
9,52		9,52		9,52		9,52		9,52	
0	9,52	0	9,52	0	9,52	0	9,52	0	9,52
9,52		9,52		9,52		9,52		9,52	
0	9,52	0	9,52	0	9,52	0	9,52	0	9,52
9,52		9,52		9,52		9,52		9,52	
0	9,52	0	9,52	0	9,52	0	9,52	0	9,52
9,52		9,52		9,52		9,52		9,52	
0	9,52	0	9,52	0	9,52	0	9,52	0	9,52
9,52		9,52		9,52		9,52		9,52	

przeciętna odległość pomiędzy wszystkimi jednostkami centroidu 511 i 517. Oznacza to, że wszystkie jednostki wchodzące w skład obu centroidów możemy traktować tak, jak gdyby miały one tę samą przeciętną odległość (9,52). A wobec tego również macierz odległości 2×2 , powstała w etapie /5/17, można przedstawić w ten sposób, ażeby wszystkie jednostki centroidu były w odpowiedni sposób uwidocznione (tab. 2a).

Cechą charakterystyczną tablicy jest zachowana symetryczność rozmieszczenia odległości jednostek danego centroidu. Na specjalną uwagę zasługuje fakt, że wystarczy znać odległości w jednej kolumnie (np. 15) danego etapu (a więc odległość środka centroidu od środka lub jednostek innego centroidu), aby można było umieścić je we wszystkich innych kolumnach odpowiadających jednostkom danego centroidu. Identycznie jest w przypadku wierszy. Utożsamienie identyczności odległości środka centroidu i przeciętnej odległości jego składowych mieć będzie decydujące znaczenie w dalszym postępowaniu.

Tab. 3. Tablica drzewa połączeń dla
 Table of linkage tree for six basic

	501	502	503	504	505	506	507	508
1	7,99	0	3,62	8,98	8,55	4,79	4,46	7,08
2	1,48	9,08	7,72	0	0	10,19	10,26	13,24
3	3,64	4,88	4,88	4,87	4,44	5,94	6,00	9,01
4	10,36	3,56	5,60	11,15	11,14	2,83	2,83	4,05
5	7,08	4,99	3,22	8,29	7,86	4,09	3,82	6,34
6	13,03	7,07	7,92	14,24	13,81	3,56	3,61	0
7	4,52	3,27	4,51	5,88	5,39	6,39	6,08	9,35
8	9,75	3,97	4,39	10,96	10,53	2,15	0	3,58
9	7,55	0	3,62	8,98	8,55	4,79	4,46	7,08
10	2,60	7,78	5,86	2,36	2,45	9,10	9,16	12,27
11	9,38	5,03	5,03	10,59	10,16	0	0	3,58
12	3,94	9,73	8,34	2,46	2,99	11,58	11,64	14,63
13	6,72	2,90	0	7,60	7,21	5,17	4,87	7,93
14	0	7,74	6,49	1,57	0	10,19	10,26	13,24
15	6,29	3,86	0	7,60	7,21	5,17	4,87	7,93
16	1,75	8,72	7,36	0	0	10,19	10,26	13,24
17	12,16	7,25	7,96	13,35	12,95	3,27	3,59	0
18	0	7,74	6,49	1,57	0	10,19	10,26	13,24
19	9,77	4,63	5,42	10,98	10,55	0	0	3,58
Waga	8,9	10,3	9,3	16,1	25,0	15,9	25,4	19,5

Należy przy tym zauważyć, że na poszczególnych etapach połączeń nie zmieniają się wszystkie odległości, ale tylko ich część. Mianowicie, zamiast dwu wierszy i dwu kolumn powstaje jeden nowy wiersz i jedna nowa kolumna jako centroid z dwu poprzednich. Wszystkie pozostałe wartości z poprzedniego etapu połączeń pozostają bez zmian. Wystarczy wobec tego zapisać tylko nowe odległości na każdym etapie połączeń, bez każdorazowego powtarzania tych wszystkich odległości, które nie zmieniły się od poprzedniego etapu. Ponieważ w każdej macierzy, na każdym etapie połączeń, jak już zauważyliśmy, symetria macierzy zostaje zachowana, tzn. odpowiedniemu wierszowi odpowiada identyczna kolumna, wystarczy tylko zachować wartości kolumn albo wierszy w każdym etapie połączeń, aby można było na tej podstawie uzyskać wszystkie interesujące nas odległości.

Na tej identyfikacji środka centroidu i jednostek centroidu oraz

6 głównych upraw (powiaty — 1965)
crops, for 1965 (by counties)

509	510	511	512	513	514	515	516	517	518
8,41	4,32	8,70	4,27	4,27	5,42	0	0	0	0
0	10,13	0	5,26	6,89	11,56	7,45	6,92	9,52	0
4,63	6,10	5,15	0	4,97	7,31	4,65	0	0	0
10,87	0	10,65	6,14	4,72	0	5,43	5,84	0	0
7,60	4,09	7,92	4,95	0	5,76	0	0	0	0
13,13	3,66	13,45	9,10	7,15	0	5,43	5,84	0	0
5,59	6,06	6,18	0	4,97	7,31	4,65	0	0	0
10,15	0	10,65	6,14	4,72	0	5,43	5,84	0	0
8,41	4,32	8,70	4,27	4,27	5,42	0	0	0	0
0	10,13	0	5,26	6,89	11,56	7,45	6,92	9,52	0
10,15	0	10,65	6,14	4,72	0	5,43	5,84	0	0
2,88	11,74	0	5,26	6,89	11,56	7,45	6,92	9,52	0
6,94	4,96	7,25	4,74	0	5,76	0	0	0	0
0	10,13	0	5,26	6,89	11,56	7,45	6,92	9,52	0
6,94	4,96	7,25	4,74	0	5,76	0	0	0	0
0	10,13	0	5,26	6,89	11,56	7,45	6,92	9,52	0
13,13	3,66	13,45	9,10	7,15	0	5,43	5,84	0	0
0	10,13	0	5,26	6,89	11,56	7,45	6,92	9,52	0
10,15	0	10,65	6,14	4,72	0	5,43	5,84	0	0
31,6	29,4	40,3	9,3	18,7	48,9	29,0	38,2	87,1	127,4

wiersza i kolumny oparty jest tok dalszego postępowania. Polega ono na znalezieniu centroidu z dwu kolumn wybieranych na zasadzie największego podobieństwa (najmniejszej różnicy) przez obliczenie odległości pomiędzy nowym centroidem a wszystkimi pozostałymi jednostkami lub centroidami.

Etap pierwszy. Obliczenia rozpoczynamy od znalezienia absolutnego minimum w macierzy odległości (tab. 1), oznaczającego dwie najbardziej podobne jednostki w całej badanej zbiorowości. Wartość minimalna (1,10) znajduje się w osiemnastym wierszu kolumny czternastej i w czternastym wierszu kolumny osiemnastej. Oznacza to, że w pierwszym etapie należy utworzyć centroid z wartości kolumny czternastej i osiemnastej. W czternastym i osiemnastym wierszu nowo utworzonej kolumny 501⁹ wpisujemy wartość minimalną (1,10) lub wartości ze-

⁹ Oznaczenie nowych kolumn jest dowolne. Przy małej liczbie można

rowe.¹⁰ Pozostałe wartości kolumny czternastej mnożymy przez jej wagę (3,9)¹¹, podobnie pozostałe wartości kolumny osiemnastej mnożymy przez wagę tej kolumny (5,0). Otrzymane w ten sposób iloczyny sumujemy w odpowiednich wierszach i dzielimy je przez sumę wag łączonych kolumn, a wyniki wpisujemy do odpowiednich wierszy nowej kolumny 501 (tab. 3), natomiast kolumny czternastą i osiemnastą wykreślamy z tab. 1.

E t a p d r u g i. Spośród osiemnastu kolumn pozostałych po etapie pierwszym znajdujemy najmniejszą odległość (1,20). Łączy ona kolumny pierwszą i dziewiątą. W wierszach pierwszym i dziewiątym nowej kolumny 502 wpisujemy zera. Dla wierszy czternastego i osiemnastego kolumny 502 obliczamy wspólną średnią. Otrzymujemy ją przez pomnożenie wartości kolumny pierwszej w obu wymienionych wierszach przez wagę (3,6) tej kolumny, następnie mnożymy wartości tych wierszy w kolumnie dziewiątej przez wagę kolumny, otrzymane iloczyny sumujemy, wynik dzielimy przez sumę wag, a otrzymaną wartość wpisujemy w wierszach czternastym i osiemnastym kolumny 502. Dla pozostałych wierszy odległości centroidu obliczamy identycznie jak w etapie pierwszym. Wartości kolumny pierwszej mnożymy przez jej wagę, następnie wartości kolumny dziewiątej mnożymy również przez wagę tej kolumny. Otrzymane wartości sumujemy w wierszach, wyniki dzielimy przez łączną wagę obu kolumn, a otrzymane wartości wpisujemy pod odpowiednimi wierszami kolumny 502.

Chodzi więc w obu etapach o obliczanie wartości centroidu według wzoru na średnią arytmetyczną ważoną:

$$x = \frac{\sum xw}{\sum w}$$

gdzie: x — odległości pomiędzy jednostkami i centroidami,
 w — wagi jednostek lub centroidów.

oznaczać je symbolami literowymi. Można też oznaczać je cyframi, np. od 101 dla liczebności nie przekraczającej stu jednostek. W naszym przypadku oznacza się je od 501, ponieważ największa liczebność wynosi 417 jednostek, a nie może być symboli powtarzających się. Na przykład 505 jest symbolem nowej kolumny, a numer 505 — 500 = 5 — symbolem etapu połączeń.

¹⁰ Nie ma to wprawdzie istotnego znaczenia, ale w dalszym postępowaniu będziemy wpisywali wartości zerowe. Bardziej podkreślają one istotę zagadnienia, gdyż różnice pomiędzy jednostkami tego samego centroidu (w myśl przedstawionego założenia) wynoszą zero.

¹¹ Ze względu na istotne różnice w wielkości poszczególnych jednostek podstawowych przyjęto zasadę, że każda jednostka otrzyma indywidualną wagę, proporcjonalną do jej wielkości. W tym konkretnym przykładzie wagi poszczególnych jednostek terytorialnych są proporcjonalne do wielkości powierzchni zasianej w tych jednostkach.

Identycznie obliczamy odległości w etapie trzecim i następnych. Różnice polegają tylko na tym, że wzrasta liczba średnich ważonych obliczanych z centroidów. Zmienia się więc tylko organizacja obliczeń, natomiast ich istota pozostaje ta sama.

Na przykład w etapie piątym obliczamy średnią z dwu centroidów: 501 — który składa się z jednostek czternastej i osiemnastej, oraz 504 — który powstał z jednostek drugiej i szesnastej. Zatem w wierszach 14, 18, 2 i 16 nowo tworzonej kolumny 505 wpisujemy wartości zerowe. Dla grup jednostek składających się na centroidy 1—4 i dla pojedynczych jednostek obliczamy odległości jak w etapie drugim (z tym tylko, że używamy wag odpowiadających kolumnom 501 i 504). Podobnie w etapie szóstym dla wartości kolumny jedenastej i dziewiętnastej z wierszy 2, 14, 16 i 18 obliczamy średnią odległość i wpisujemy ją do odpowiednich wierszy kolumny 506. Po zakończeniu obliczeń wszystkich odległości kolumny 506 kolumnę jedenastą i dziewiętnastą wykreślamy. W ten sposób po każdym etapie obliczeń tabela połączeń zmniejsza się o jedną kolumnę (jedna kolumna przybywa, ale dwie ubywają), natomiast ilość wierszy pozostaje bez zmian.

Na podstawie sporządzonej tablicy połączeń (tab. 3) możemy odtworzyć tabelę symetryczną dla dowolnego etapu połączeń. Dla etapu dwunastego np. symetryczna tabela sporządzona na podstawie tabeli połączeń przedstawia się następująco:

	5	502	503	508	510	511	512
5	0						
502	4,99	0					
503	3,22	3,62	0				
508	6,34	7,08	7,93	0			
510	4,09	4,32	4,96	3,66	0		
511	7,92	8,70	7,25	13,45	10,65	0	
512	4,95	4,27	4,74	9,10	6,14	5,26	0

Możliwość ta może być z powodzeniem wykorzystana w przypadku, gdy chcemy na pewnym etapie połączeń przerwać kontynuowanie tego procesu i przejść na inną metodę, albo też gdy chcemy kontynuować proces łączenia, ale niezależnie od niego na pewnym etapie pragniemy zainteresować się powiązaniem nie tylko pierwszego stopnia, ale również dalszych stopni.

Poza tym gdybyśmy chcieli drukować całość wyników łączenia kolumn przy pomocy metody obliczeń poprzednio prezentowanej i obecnie

przedstawionej, to w tym przypadku zyskujemy dużą oszczędność miejsca. Dla wydrukowania całości wyników w poprzednim ujęciu potrzeba było siedemnastu tabel o wymiarach od 18×18 do 2×2 wierszy i kolumn, a obecnie otrzymamy tylko jedną tabelę o 19 wierszach i 18 kolumnach.

Jeżeli nie mamy żadnych dodatkowych celów na uwadze, a jedynie konstrukcję drzewa połączeń, wystarcza nam tylko znajomość symboli jednostek, które ulegają połączeniu w poszczególnych etapach, oraz znajomość odległości minimalnej decydującej o danym połączeniu. W ten właśnie sposób zostały sporządzone skrócone tabele, które posłużyły do konstrukcji drzewa połączeń dla gromad (ryc. 9).

ZAGADNIENIE PORZĄDKOWANIA I GRUPOWANIA GROMAD

Porządkowanie gromad

Podobnie jak w przypadku powiatów, tak i tutaj tablica przeciętnych różnic stanowiła podstawę porządkowania, a następnie podziału badanych jednostek na grupy (regiony).

Początkowo zastosowano porządkowanie dendrytowe (ryc. 8). Z powodów, które już omówiono wcześniej, uporządkowanie dendrytowe nie dawało dostatecznie pewnych przesłanek do pogrupowania gromad i wydzielenia na tej podstawie regionów. Z tego też względu zrezygnowano z dendrytu jako metody głównej, a wykorzystano go jedynie jako metodę uzupełniającą przy analizie drzewa połączeń i jako metodę pomocniczą dla wstępnego porządkowania diagramu Czekanowskiego.

Podstawową część grupowania przeprowadzono w oparciu o metodę „drzewa połączeń” (ryc. 9). W przeciwieństwie do dendrytu i diagramu Czekanowskiego drzewo połączeń jednocześnie z porządkowaniem dokonuje grupowania gromad i w ten sposób przygotowuje stopniowo drogę do podziału na regiony.

Na podstawie otrzymanego grupowania opracowano mapy (ryc. 5, 6) przedstawiające różne stopnie generalizacji, wynikające z przyjętego dopuszczalnego zróżnicowania zespołów gromad. Opracowane mapy są wynikiem tylko generalizacji rachunkowej, wynikającej z charakteru metody, bez jakiegokolwiek generalizacji geograficznej. Stąd nawet przy najwyższym stopniu generalizacji nie wszystkie zespoły gromad (regiony) tworzą zupełnie zwarte terytorialnie obszary.

Zróżnicowanie w poszczególnych przekrojach klasyfikacji nie przekracza górnej granicy przyjętej dla odpowiedniego stopnia generalizacji. Na podstawie analizy gradientu w przeciętnych różnicach centroi-

dów ustalono następujące wartości graniczne: dla pierwszego stopnia generalizacji 3,5, dla trzeciego 5,5 przeciętnej różnicy. Dolna granica nie została tak wyraźnie określona, gdyż uzależniona jest z jednej strony od tworzenia się wyraźnych zespołów gromad podobnych, z drugiej — od momentu pojawiania się zwartych terytorialnie obszarów z zespołów tych gromad.

Przyjęte zasady, szczególnie w etapie pierwszym, nie zawsze były precyzyjnie przestrzegane. Niekiedy zachodziły przypadki, że do dwu wyraźnie określonych zespołów dopiero po ich zgrupowaniu przyłączyły się pojedyncze gromady lub niewielkie grupy gromad. Wcześniejsze wydzielenie grup powodowało konieczność przydziału tych gromad do jednej z nich. Decyzję w dwu przypadkach (odnośnie zespołów gromad: Bełzec, Łabunie i m. Janów, m. Biłgoraj) uzależniono od wyniku analizy różnic przeciętnych pomiędzy gromadami i od wartości zespołu cech typologicznych tych gromad, a w jednym przypadku (m. Puławy, m. Chełm) także od położenia geograficznego badanych jednostek. Tak więc jednostki te przydzielono do odpowiednich grup wcześniej, niżby to wynikało z drzewa połączeń. W innych przypadkach jednorodnie zespoły gromad zwartych terytorialnie mogły być wydzielone znacznie wcześniej. Na przykład w zespole 8 gromady: Czajki, Orłów Murowany, Izbica, Tarzymiechy stanowią wyraźnie odrębną grupę, jednorodną od względem zespołu cech typologicznych i zwartą terytorialnie. Mimo to wyodrębniono tylko jeden zespół gromad, chcąc zachować odpowiedni stopień generalizacji.

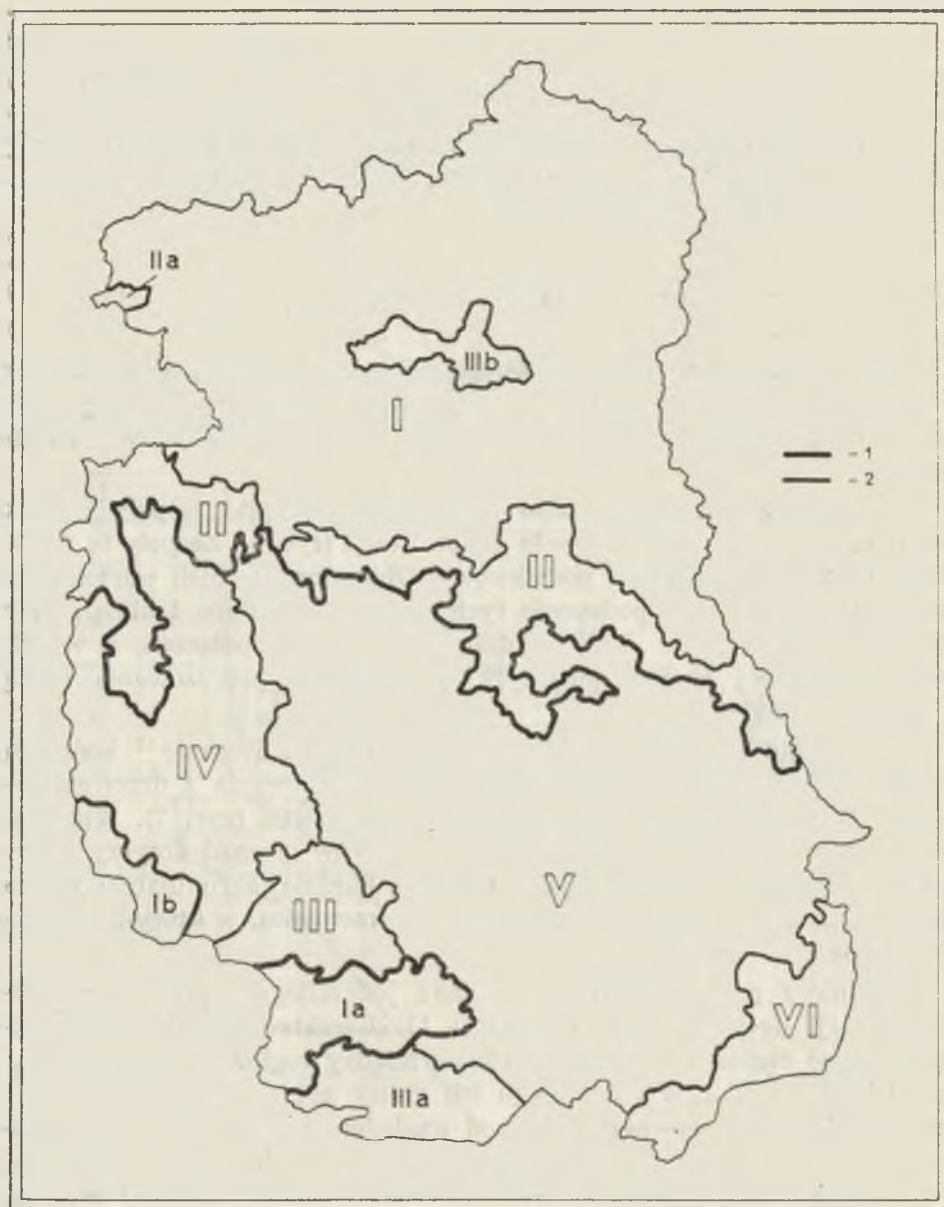
Odrębne zagadnienie stanowi sprawa kilku jednostek bardzo zróżnicowanych, które w pierwszym, drugim, a nawet trzecim etapie generalizacji nie uzyskały połączenia z innymi jednostkami. Są to jednostki ostatecznie zgrupowane w dwa odrębne zespoły: 1) Lubycza Królewska, Machnów i m. Kraśnik Fabryczny oraz 36) Gołęb, m. Terespol, m. Opole. Szczególnie dołączenie Kraśnika Fabrycznego do zespołu pierwszego oraz połączenie jednostek w zespole trzydziestym szóstym budzi pewne wątpliwości.

Pomijając celowość łączenia w grupy jednostek zbyt zróżnicowanych, wydaje się konieczne zwrócenie uwagi na dodatkowy fakt rzucający pewne światło na samo zagadnienie doboru cech regionalizacyjnych. Większość z tych jednostek, co do których mamy wątpliwości odnośnie ich poprawnego zakwalifikowania, wykazuje mianowicie niskie wartości udziału procentowego w strukturze zasiewów. Co więcej dla niektórych z tych jednostek w zespole cech typologicznych brakuje cechy najważniejszej, np. dla m. Terespolu i m. Kraśnika Fabrycznego — uprawy warzyw. Okazuje się więc, jak ważny jest postulat poprawnego i pełnego



Ryc. 5. Zespoły gromad — 6 głównych upraw — r. 1965. Numeracja zespołów zgodna z numeracją ryc. 7

Groups of villages with regard to six basic crops, for 1965. The numbering of the group conforms to that shown in Fig. 7



Ryc. 6. Granice regionów produkcji roślinnej (polowej) w r. 1965; 1—granice pierwszego rzędu, 2—granice drugiego rzędu
Boundaries of regions with regard to agricultural field produce in 1965; 1—first rank boundaries, 2—second rank boundaries

określenia przez zespół cech typologicznych w s z y s t k i c h badanych jednostek.

Należy ponadto podkreślić, że metoda drzewa połączeń posiada określone „właściwości” wynikające z samej jej istoty: daje wyniki pewniejsze przy grupowaniu jednostek podobnych, natomiast ze wzrostem różnicowania jednostek grupowanych, przy małych różnicach między grupami, niezawodność metody maleje. Fakt ten ma istotne znaczenie dla dalszego toku postępowania.

W y d z i e l a n i e r e g i o n ó w

Brak pewności co do poprawnego połączenia gromad w pierwszym etapie klasyfikacji, szczególnie na wyższym stopniu generalizacji, spowodował powtórzenie tego procesu przy zmienionej podstawie odniesienia.

Za podstawę drugiego etapu grupowania przyjęto zespoły gromad wyróżnione w pierwszym stopniu generalizacji (ryc. 5). Zespoły te potraktowano tak jak jednostki podstawowe. Obliczono dla nich wartości poszczególnych cech. Na podstawie tych wartości obliczono tablicę różnic. Następnie opracowano dendryt, diagram i drzewo połączeń, a wyniki przeprowadzonej analizy posłużyły dla dalszego precyzowania liczby i granic regionów.

W analizie tej najbardziej wymowne jest porównanie¹² końcowej części drzewa połączeń z pierwszego etapu grupowania z drzewem połączeń sporządzonym w drugim etapie grupowania (ryc. 7). Widoczna odmienność zakwalifikowania pewnych zespołów gromad dotyczy zespołów o charakterze przejściowym i wymaga bardziej skrupulatnego przeanalizowania tych zespołów, szczególnie od momentu, w którym powstają istotne rozbieżności.

Zgodnie z przyjętą definicją regionu jednolitego o poprawności regionalizacji decydują dwa momenty: 1) charakter różnicowania wewnętrznego regionu, 2) istotność różnic między regionami. Zatem decyzje o liczbie regionów i przebiegu ich granic muszą uwzględniać badanie różnicowania wewnętrznego pod względem zespołu cech typologicznych oraz ocenę stopnia istotności różnic tych cech między regionami.

Problemy weryfikacji liczbowej przebiegu granic regionów oraz za-

¹² W drugim etapie grupowania z zespołów gromad 29, 33, 34 wyróżniono dodatkowo oddzielny zespół gromad (31), zwarty terytorialnie i położony w innej części województwa niż pozostałe podobne zespoły. Wskutek tego na ryc. 7 b mamy o jeden zespół więcej niż na ryc. 7 a. Różnica ta nie ma istotnego wpływu na końcowy wynik grupowania.



Ryc. 7. Rozbieżność w grupowaniu przy przyjęciu za podstawę gromad lub zespołów gromad; A —końcowy fragment drzewa połączeń z pierwszego etapu grupowania (gromady), B —drzewo połączeń dla drugiego etapu grupowania (zespoły gromad)

Discrepancies in grouping when basing it on villages or on groups of villages; A —final fragment of linkage tree from first stage of grouping (villages), B — linkage tree for second stage of grouping (groups of villages)

sadności tych granic zostaną rozstrzygnięte przy pomocy miar zmienności ($\hat{\sigma}_x$ i V_x) oraz testu „t” Studenta.

Jeżeli mówimy o jednorodności wewnętrznej regionu, oznacza to, że poszczególne regiony mają wykazywać maksymalnie zbliżony poziom cech typologicznych oraz podobną strukturę wartości tych cech, a różnice między jednostkami zaliczonymi do określonego typu (regionu) powinny być możliwie najmniejsze. Z tego powodu również wartości dyspersji ($\hat{\sigma}_x$) obliczone dla poszczególnych regionów powinny być znacznie mniejsze od wartości dyspersji dla całego badanego obszaru, ale także powinny charakteryzować się stosunkowo wyrównanym stopniem zmienności w regionach i to nie tyle dla poszczególnych cech, co średniej zmienności całego zespołu cech typologicznych. Do takiego rozwiązania dążymy w trakcie całego procesu porządkowania (grupowania). Niemniej jednak pozostają pewne przypadki wątpliwe, których rozstrzygnięcie może nastąpić na drodze zbadania zmienności.

O poprawnym zakwalifikowaniu poszczególnych jednostek czy zespołów jednostek możemy mówić dopiero wtedy, jeżeli przesunięcie jakiegokolwiek jednostki (zespołu jednostek) „z jednego regionu do drugiego nie spowoduje: 1) jednoczesnego zmniejszenia współczynników zmienności dla obu regionów, między którymi zachodzi przesunięcie [...]; 2) większego spadku współczynników zmienności w jednym regionie, aniżeli wynosi wzrost współczynników zmienności w drugim regionie” (8, s. 116).

Naturalnie badanie zmienności nie może mieć charakteru permanentnego, a może dotyczyć tylko przypadków wątpliwych dla ustalonego na innej drodze stopnia generalizacji i związanej z nim liczby regionów. Zresztą i wtedy z powodu trudności rozstrzygnięcia, która z miar zmienności $\hat{\sigma}_x$ czy V_x jest bezwzględnie lepsza, nie zawsze możemy w sposób jednoznaczny przesądzać o absolutnej wyższości takiego a nie innego przebiegu granic regionów.

Wracając do przykładu zademonstrowanego na ryc. 7, sędzę, że należy przede wszystkim omówić sprawę przynależności zespołów gromad 8, 9, 22 oraz problem przebiegu granicy między regionami IV i V, uzyskanymi w trzecim stopniu generalizacji.

Wyjaśnienie pierwszego zagadnienia tkwi niewątpliwie w genezie zespołu 22. W obu przypadkach na ryc. 7A i 7B zespoły gromad 8 i 9 łączą się ze sobą wzajemnie i ta sprawa nie wymaga wyjaśnień. Natomiast zespół 22 raz uzyskuje połączenie z zespołami 23, 24, innym razem z zespołami 8, 9. Połączenie zespołu 22 z jedną lub drugą grupą powoduje zmianę przebiegu granicy między regionami z powodu peryferyjnego położenia jednej z trzech jednostek tego zespołu. Z tego

względem rozstrzygnięcie, które z uporządkowań jest lepsze, ma także znaczenie dla ustalenia przebiegu granicy między głównymi regionami.

Powstanie zespołu gromad 22 nastąpiło na drodze połączenia jednostek: Aleksandrów i m. Tomaszów L. w kroku 709 (czyt. 209) w pierwszym etapie grupowania, a następnie przyłączenia do nich jeszcze gromady Ruskie Piaski w kroku 824 (czyt. 324). Należy przy tym wspomnieć, że o ile jednostki m. Tomaszów L. i Aleksandrów tworzyły parę wzajemną, to gromada Ruskie Piaski wykazywała największe podobieństwo z gromadą Żabia Wola, lecz na skutek wcześniejszego połączenia tej jednostki w inny zespół została dołączona do zespołu 709. Mimo to zespół gromad 22 należy oceniać jako względnie jednorodny.

Przejściowy charakter zespołu 22 (824) i wyjątkowo niekorzystny dla niego cykl połączeń spowodował, że zespół ten, wykazujący podobieństwo pierwszego stopnia z zespołem 688 (ryc. 9) składającym się z jednostek: Bondyż, Krasnobród, Suchowola i Majdan Sopocki, oraz podobieństwo drugiego stopnia z zespołem 813 składającym się z sąsiednich gromad: Józefów i Zwierzyniec, nie został z nimi połączony ze względu na wcześniejsze połączenie tych zespołów z innymi jednostkami w zespół 864 (9). Należy podkreślić, że wymienione zespoły także — zwykle na drugim miejscu — wykazywały podobieństwo z zespołem 824 (22).

Jeżeli do tego uwzględnimy analizę rozmieszczenia geograficznego jednostek wchodzących w skład zespołu 22, to przyłączenie tego zespołu do zespołów 8 i 9 jak na ryc. 7B wydaje się jak najbardziej uzasadnione. Do podobnych wniosków prowadzi analiza diagramu Czekanowskiego, sporządzonego dla drugiego etapu grupowania, aczkolwiek niejednorodność tej grupy na diagramie jest już widoczna. Szczególnie widoczne jest bliższe podobieństwo zespołu 9 z zespołem 15, a także bardziej widoczny przejściowy charakter zespołu 22. Niemniej wymienione zespoły stanowią bardzo podobne typy z nieznacznymi odchyleniami in + w zespole 22 dla owsa, a w zespole 8 dla jęczmienia i buraków cukrowych.

O ile jednak połączenie zespołów gromad 8, 9 i 22 w większy zespół 517 jak na ryc. 7B nie budzi zastrzeżeń, to dalsze grupowanie przedstawione na tej rycinie — aczkolwiek słuszne z punktu widzenia mechanizmu metody — może stwarzać uzasadnione wątpliwości. Z tego względu proces grupowania został przerwany na pewnym etapie generalizacji (na ryc. 7B zaznaczono linią przerywaną) i dla utworzonych wcześniej grup sporządzono „wstępnie uporządkowaną” tablicę różnic przeciętnych między utworzonymi zespołami gromad (tab. 4). Wgląd w tablicę różnic daje podstawę do traktowania zespołu 517 łącznie z zespołami 523, 519 i 520 jako całości, która w trzecim stopniu generalizacji stanowi jeden zwarty region.

Tab. 4. Tablica różnic dla etapu III wstępnie uporządkowana (dla rozstrzygnięcia przynależności zespołu 517)
 Table of differences for stage III. Preliminary arrangement (for deciding the appartenance of group 517)

	2	1	523	519	520	517	518	524	36	508	25	522
2	0	4,81	9,56	5,54	7,51	8,55	12,58	11,24	12,68	11,59	13,91	17,07
1	4,81	0	7,55	6,58	7,37	8,03	11,53	10,65	11,92	10,31	12,21	15,40
523	9,56	7,55	0	3,98	4,86	4,60	9,12	7,98	9,70	8,23	10,82	13,64
519	5,54	6,58	3,98	0	4,48	4,22	7,28	6,62	8,14	5,63	9,07	12,54
520	7,51	7,37	4,86	4,48	0	4,99	7,22	4,78	7,50	5,63	8,50	11,53
517	8,55	8,03	4,60	4,22	4,99	0	4,60	5,28	7,74	6,19	5,74	8,56
518	12,58	11,53	9,12	7,28	7,22	4,60	0	4,49	6,20	6,22	6,39	6,05
524	11,24	10,65	7,98	6,62	4,78	5,28	4,49	0	4,71	4,83	7,87	8,79
36	12,68	11,92	9,70	8,14	7,50	7,74	6,20	4,71	0	8,23	9,55	4,54
508	11,59	10,31	8,23	5,63	5,63	6,19	6,22	4,83	8,23	0	4,01	7,58
25	13,91	12,21	10,82	9,07	8,50	5,74	6,39	7,87	9,55	4,01	0	5,42
522	17,07	15,40	13,64	12,54	11,53	8,56	6,05	8,79	4,54	7,58	5,42	0

Takie rozwiązanie potwierdza także poniższe zestawienie miar zmienności.

Wariant	Region	Liczba jednostek w regionie	σ_x	V_x
1	III	26	18,91	344,46
	V	125	21,97	250,40
2	III	23	18,48	333,10
	V	128	21,09	246,90

Przesunięcie zespołu 22 z regionu trzeciego do regionu piątego w wariacie drugim obniża wskaźniki zmienności w obu regionach, co jest zgodne z naczelnym warunkiem poprawnej regionalizacji.

Na marginesie tej analizy można pokusić się o sformułowanie pewnego wniosku o charakterze ogólniejszym. Mianowicie rozwiązanie to potwierdza w sposób wyjątkowo przekonujący konieczność stosowania wielu środków (metod) równocześnie dla rozwiązania określonego zadania.

Trudniejszy problem stanowi ustalenie granicy pomiędzy regionami IV i V. Rozbieżności dotyczą zakwalifikowania zespołów gromad 16, 17 i 18. Według klasyfikacji na podstawie drzewa połączeń z pierwszego etapu grupowania zespoły te w całości zostały zakwalifikowane do regionu IV, natomiast według klasyfikacji na podstawie drzewa połączeń w drugim etapie grupowania — w całości do regionu V. Istnieje jeszcze inna możliwość, uwidoczniła częściowo na ryc. 7B, a potwierdzona w pełni opracowaniem drugiego etapu grupowania przy nieco odmiennych wagach dla poszczególnych cech. W wyniku takiego grupowania następuje rozbicie grupy zespołów na dwie części: zespoły 17 i 18 zostają zakwalifikowane do regionu IV, a zespół 16 — do regionu V (ryc. 6). Podobną tendencję wykazuje uporządkowanie zespołów gromad w diagramie Czekanowskiego, sporządzonym dla drugiego etapu grupowania.

Rozstrzygnięcie, który z trzech wariantów najbardziej odpowiada definicji regionu jednolitego, powinno nastąpić na drodze zbadania zmienności wewnątrz regionów. Analiza danych tab. 5, sporządzonej w tym celu, nie daje jednak zupełnie jednoznacznej odpowiedzi. Gdybyśmy bowiem bez żadnych zastrzeżeń za kryterium poprawności przyjęli sumę zmienności w regionach poszczególnych wariantów, to przy przyjęciu za kryterium δ_x wariant pierwszy byłby najlepszy, a przy kryterium V_x — wariant trzeci.

Analiza zmienności poszczególnych cech nie upoważnia do tak zdecydowanej diagnozy. Wybór wariantu trzeciego przy V_x wynika tylko z obniżenia zmienności mieszanek zbożowych w tym wariacie. Pamiętać jednak należy, że mieszanki zbożowe wprowadzono jedynie jako

Tab. 5. Miary zmienności (δ_x i V_x) dla cech typologicznych w różnych wariantach regionu IV i V
 Scales of variability (δ_x and V_x) for typological features in different alternatives of regions IV and V

Wariant	Region	Liczba jednostek w regionie	Zyto	Pszensica	Owies	Jęczmień	Ziemiaki	Buraki cukrowe	Mieszanki zbożowe	Suma
1	IV	39	δ_x 3,87 V_x 17,57	2,51 19,36	3,46 35,90	2,31 43,57	1,70 9,89	2,05 52,49	2,61 24,37	18,51 213,15
	V	168	δ_x 4,51 V_x 23,41	3,30 17,22	2,75 44,88	2,55 42,40	2,22 14,86	5,08 60,52	3,60 70,76	24,01 271,05
2	IV	60	δ_x 4,71 V_x 24,19	3,44 32,51	2,98 30,52	2,40 44,63	2,08 12,75	2,23 51,55	2,46 22,66	20,30 218,81
	V	147	δ_x 4,48 V_x 22,67	2,66 13,34	5,53 41,53	2,55 41,80	2,24 15,03	5,19 58,55	2,93 68,28	22,58 261,23
3	IV	79	δ_x 4,62 V_x 24,93	4,90 38,54	2,96 32,85	2,22 41,92	2,09 13,27	2,07 46,83	2,34 22,68	21,20 220,42
	V	125	δ_x 4,41 V_x 21,85	2,64 13,19	2,54 42,51	2,67 42,80	2,28 15,21	5,30 56,32	2,13 58,52	21,97 250,40

δ_x — dyspersja (dispersion).

V_x — współczynnik zmienności (coefficient of variability).

cechę pomocniczą i gdybyśmy obliczali zmienność dla wartości ważonych, miałyby one mniejsze znaczenie. Podobnie jak w przypadku wariantu pierwszego przy δ_x , mniejsza zmienność wynika przede wszystkim z mniejszej dyspersji żyta. Pamiętając o dużej różnicy (in +) między dyspersją a wagą dla żyta, znowu mamy wątpliwości co do większej poprawności tego wariantu.

Biorąc pod uwagę stosunkowo wyrównany przebieg wskaźników zmienności w wariacie drugim oraz fakt, że przebieg granicy tego wariantu potwierdzany jest innymi metodami, skłonni jesteśmy przyjąć ten wariant, zdając sobie sprawę z pewnej subiektywności wyboru.

Rozstrzygnięcie, które jednostki lub zespoły jednostek przydzielić do odpowiedniego regionu, może nastąpić na drodze zbadania zróżnicowania wewnętrznego regionów. Ale trzeba brać pod uwagę fakt, że z reguły regiony niższego rzędu będą wykazywały mniejsze zróżnicowanie, natomiast regiony wyższego rzędu będą charakteryzowały się większym zróżnicowaniem wewnętrznym. Dlatego rozstrzygnięcie o przebiegu granicy między regionami może mieć miejsce dla ustalonego już — na innej drodze — typu regionu.

Zagadnienie hierarchii regionów, ich liczby i granic wymaga potwierdzenia z punktu widzenia istotności różnic między regionami. Dla oceny stopnia zróżnicowania między regionami został zastosowany test istotności „t” i „c”¹³.

¹³ L. Zobler proponuje stosowanie w badaniach geograficznych testu chi kwadrat (26). Podobnie test chi kwadrat stosuje się w szeregu innych opracowań przestrzennych. Zastosowanie w tym przypadku testu „t” i „c” wiąże się z chęcią bardziej analitycznego badania istotności z punktu widzenia poszczególnych cech typologicznych.

Zgodnie z zaleceniem W. Okta by (17), przy spełnieniu założenia o równości wariancji (test istotności F) zastosowano tu wzór w postaci:

$$t = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 1} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

gdzie: \bar{y}_1 i \bar{y}_2 — średnie w pierwszej i drugiej grupie

S_1^2 i S_2^2 — wariancje pierwszej i drugiej grupy

n_1 i n_2 — liczebności w pierwszej i drugiej grupie

natomiast przy niespełnianiu założenia o równości wariancji korzystano z przybliżonego testu Cochrona i Coxa w postaci:

$$C^0 = \sqrt{\frac{\bar{y}_1 - \bar{y}_2}{Z_1 + Z_2}}$$

$$\text{gdzie: } Z_1 = \frac{n_1 S_1^2}{n_1(n_1 - 1)} \quad \text{i} \quad Z_2 = \frac{n_2 S_2^2}{n_2(n_2 - 4)}$$

Obliczane wartości graniczne istotności, przy 5% ryzyku błędu, wahały się w granicach od 2.0 do 2.4. Można zatem zgodnie z powszechną praktyką przyjąć, że różnice dla „t” i „c” > 3 są istotne, a różnice dla „t” i „c” < 3 nieistotne.

Tab. 6. Testy istotności „t” i „c” pomiędzy przeciętnymi wartościami cech typologicznych według regionów

Tests of significance of "t" and "c" for mean values of typological features, by regions

Regiony	II	III	IV	V	VI	
I	1	c = 7,180	t = 16,937	t = 29,163	t = 39,168	t = 19,881
	2	c = 6,400	t = 11,067	c = 10,031	t = 44,523	c = 5,738
	3	c = 15,960	c = 6,281	c = 6,645	c = 17,451	c = 19,055
	4	c = 7,544	t = 6,578	c = 13,600	c = 21,325	c = 5,895
	5	t = 5,510	t = 8,365	c = 19,605	c = 25,024	t = 12,709
	6	c = 5,472	c = 4,675	c = 14,528	c = 19,764	c = 4,367
	7	c = 9,474	c = 3,541	c = 30,123	c = 14,027	t = 2,277
II	1		t = 5,745	c = 9,942	c = 10,926	t = 9,758
	2		t = 3,452	t = 1,440	c = 16,654	c = 4,113
	3		t = 19,808	c = 7,514	c = 1,136	t = 4,006
	4		c = 3,365	t = 1,598	t = 2,804	t = 2,768
	5		c = 3,796	c = 6,697	c = 9,036	t = 10,486
	6		c = 3,321	c = 1,941	c = 7,835	c = 2,954
	7		t = 3,746	c = 7,777	t = 2,781	c = 5,664
III	1			t = 2,844	t = 3,402	t = 7,563
	2			t = 2,491	t = 11,839	c = 2,926
	3			t = 9,921	t = 18,081	c = 21,677
	4			c = 6,310	c = 3,807	c = 4,222
	5			c = 1,237	c = 3,081	t = 8,199
	6			c = 8,727	c = 15,429	c = 3,888
	7			t = 12,708	t = 2,430	c = 1,402
IV	1				t = 0,204	t = 6,076
	2				c = 18,711	c = 3,756
	3				t = 8,025	c = 11,443
	4				t = 1,255	t = 2,359
	5				t = 3,974	t = 12,376
	6				c = 8,243	c = 2,488
	7				t = 15,009	t = 10,743
V	1					t = 6,500
	2					c = 0,230
	3					c = 6,804
	4					t = 1,863
	5					t = 10,689
	6					t = 0,795
	7					c = 4,765

Cechy: 1 — żyto, 2 — pszenica, 3 — owies, 4 — jęczmień, 5 — ziemniaki, 6 — buraki cukrowe, 7 — mieszanki zbożowe.

Features: 1 — rye, 2 — wheat, 3 — oat, 4 — barley, 5 — potatoes, 6 — sugar beets, 7 — cereal mixtures.

Nie będzie tu omawiane szczegółowo zagadnienie istotności na poszczególnych etapach generalizacji. W każdym razie najniższe jednostki regionalne poczęto wyróżniać, jeżeli zespoły gromad różniły się w sposób istotny chociaż pod względem jednej cechy, natomiast najwyższe jednostki regionalne (tab. 6) — gdy istotne różnice wystąpiły przynajmniej dla większości (tzn. dla czterech spośród siedmiu) cech typologicznych. Dane tab. 6 wykazują, że nawet między regionami o charakterze przejściowym (II i IV) występujące różnice przynajmniej dla czterech cech mają charakter różnic istotnych. Pozwala to sądzić, że różnice pomiędzy wydzielonymi regionami są dostatecznie znamienne i nie ma potrzeby prowadzenia dalszej generalizacji.

WNIOSKI OGÓLNE

Jak już zaznaczono na wstępie, zagadnienie porządkowania i podziału badanych jednostek terytorialnych stanowi końcowy fragment szerszego problemu metodologicznego procesu klasyfikacji czy regionalizacji, obejmującego także inne zagadnienia dotyczące wyboru cech regionalizacyjnych, standaryzacji i ważenia cech, metody obliczania różnic między badanymi jednostkami. Zagadnieniom tym poświęcona jest inna praca przygotowana do druku.

Zawężenie rozważań tylko do zagadnień porządkowania i podziału stwarza, rzecz jasna, ograniczone możliwości wnioskowania. Niemniej pozwala na wyciągnięcie kilku wniosków rzutujących na całość metodologii klasyfikacji (regionalizacji) przy zastosowaniu zespołu metod taksonomicznych.

Konieczność stosowania wielu metod dla rozwiązania różnych etapów regionalizacji, jak i jednoczesnego stosowania różnych metod w szczególnie trudnych momentach klasyfikacji jest ważnym elementem lepszej oceny i konfrontacji otrzymanych wyników. Konieczność ta wynika z ograniczonych możliwości i względnej tylko niezawodności wszystkich stosowanych metod badawczych. Chodzi przecież o to, aby otrzymane wyniki nie były tylko rezultatem zastosowanej metody, a stosowane metody prowadziły do obiektywnego uchwycenia badanego zjawiska.

Problem obiektywności wyników był podstawowym założeniem podjętego tematu. Realizacja tego założenia wymagała odpowiedniego ustalenia stosowanych technik porządkowania: diagramu Czekanowskiego, dendrytu wrocławskiego i drzewa połączeń, określenia ich roli w tym porządkowaniu oraz uściślenia niektórych metod.

Możliwość ścisłego ustalenia przynależności jednostek do odpowiedniego regionu pozwoliła na jednoznaczne rozgraniczenie badanego obsza-

ru. Zadanie to wymagało zastosowania elektronicznej techniki obliczeniowej. Ułatwiło to nie tylko postawienie zadań wymagających dotychczas bardzo żmudnych i pracochłonnych obliczeń, ale umożliwiło także rozwiązanie problemu porządkowania dużej liczby jednostek, którego nie można byłoby opracować tradycyjnymi środkami.

Problem dokonywania klasyfikacji praktycznie nieograniczonej liczby jednostek ma tym większe znaczenie, że w badaniach geograficznych liczba obserwacji (jednostek) wzrasta w miarę podejmowania coraz szczegółowszych badań. Zachęcać to powinno do podejmowania badań nie tylko nad rozwiązywaniem konkretnych problemów, ale także nad dalszym doskonaleniem metod taksonomicznych.

Dotychczasowe doświadczenia wynikające z porównania wyników regionalizacji przeprowadzonej dla różnych okresów czasowych potwierdzają walory zastosowanej metody, która pozwala na stosunkowo precyzyjne uchwycenie zmian między różnymi stanami czasowymi. Mimo że szereg zagadnień klasyfikacji i regionalizacji wymaga jeszcze wielu prób i udoskonaleń, to jednak wyniki wykonanej pracy upoważniają do optymistycznego spojrzenia na tę kwestię.

LITERATURA

1. Berry B. J. L.: A Method for Deriving Multi-factor Uniform Regions. *Przegl. Geogr.*, 33, z. 2, 1961, ss. 263—282.
2. Czekanowski J.: Zarys metod statystycznych w zastosowaniu do antropologii. *Prace Towarzystwa Naukowego Warszawskiego*, nr 5, 1913.
3. Ernst J.: Regiony geograficzno-rolnicze Polski (Les régions agricoles en Pologne). *Czas. Geogr.*, 10, z. 4, 1932, ss. 143—168.
4. Ernst J.: Niektóre metody określania regionów geograficzno-rolniczych (Some Methods for the Determination of Agricultural Regions). *Ann. Univ. Mariae Curie-Skłodowska, sectio B*, vol. XXI (1966), 1, Lublin 1968.
5. Fierich J.: Próba zastosowania metod taksonomicznych do rejonizacji systemów rolniczych w województwie krakowskim. *Myśl Gospodarcza*, 1, Kraków 1957, ss. 73—100.
6. Florek K., Łukasiewicz J., Steinhaus H., Zubrzycki S.: Taksonomia wrocławska. *Przegl. Antropol.*, 17, 1951, ss. 193—211.
7. Główniewicz Z.: Rejony podaży kontraktowanej trzody chlewnej w województwie poznańskim. *Ruch Prawniczy i Ekonomiczny*, 22, z. 2, 1960.
8. Główniewicz Z.: Taksonomiczna metoda różnic przeciętnych jako metoda delimitacji rejonów gospodarczych. *Biblioteka Wiadomości Statystycznych*, t. 9, Warszawa 1969, ss. 97—126.
9. Grigg D.: The Logic of Regional Systems. *Annals of the Association of American Geographers*, 55, 1965, ss. 465—491.
10. Isard W.: *Metody analizy regionalnej*. Warszawa 1965. (Methods of Regional Analysis, An Introduction to Regional Science. New York 1960).
11. Jedut R.: Próba regionalizacji głównych form użytkowania ziemi w woj. lubelskim metodą „względnej uprzywilejowania” (Essai de division en regions des formes principales d'utilisation de la terre dans woivodie de Lublin

- par la méthode de „favorisation relative”). *Ann. Univ. Mariae Curie-Skłodowska, sectio B.* vol. XIX (1964), 10, Lublin 1966.
12. Johnston R. J.: Choice in Classification: the Subjectivity of Objective Methods. *Annals of the Association of American Geographers*, 58, 1968, ss. 578—589.
 13. Kwiecień W.: Metodyka ustalania i weryfikacji rejonów produkcji rolniczej (*Méthodes de délimitation et de vérification des régions de production agricole lá l'exemple de la voïvodie de Rzeszów*). *Ann. Univ. Mariae Curie-Skłodowska, sectio H*, vol. I (1967), 10, Lublin 1967.
 14. Lewiński S.: Taxonomic Methods in Regional Studies. *Geographia Polonica*, 15, 1968, ss. 189—198.
 15. Liczkowski J.: Badanie intensywności rolnictwa w ujęciu przestrzennym (na przykładzie województwa poznańskiego). *FWRiL*, Warszawa 1964.
 16. McQuitty L. L.: Single and Multiple Classification by Reciprocal Paris and Rank Order Types. *Educational and Psychological Measurement*, 26, 1966, s. 253—265.
 17. Oktaba W.: Elementy statystyki matematycznej i metodyka doświadczalnictwa. PWN, Warszawa 1966.
 18. Perkal J.: Taksonomia wrocławska. *Przegl. Antropol.*, 19, 1953, ss. 82—96.
 19. Perkal J.: *Matematyka dla przyrodników i rolników. T. I i II*, PWN, Warszawa 1958, 1963.
 20. Rodoman B. B.: Mathematical Aspects of the Formalization of Regional Geographical Characteristics. *Geographia Polonica*, 15, 1968, ss. 37—57.
 21. Steczkowski J.: Zasady i metody rejonizacji produkcji rolniczej. *PWRiL*, Warszawa 1966.
 22. Stone R.: A Comparison of the Economic Structure of Regions Based on the Concept of Distance. *Journal of Regional Science*, vol. 2, No 2, 1960, ss. 1—20.
 23. Tarrant J. R.: A Note Concerning the Definition of Groups of Settlements for a Central Place Hierarchy. *Economic Geography*, 44, 1968, ss. 144—151.
 24. Wanke A.: Metoda badań częstości występowania zespołu cech czyli metoda stochastycznej korelacji wielorakiej. *Przegl. Antropol.*, 19, 1963, ss. 106—147.
 25. Wysocki Z.: Zagadnienie taksonomii geograficznej (*The Problem of Geographical Taxonomy*). *Przegl. Geogr.*, 37, z. 2, 1965, ss. 313—339.
 26. Zabler L.: Decision Making in Regional Construction. *Annals of the Association of American Geographers*, 48, 1958, ss. 140—148.

РЕЗЮМЕ

Представленная в работе тема — это часть большой методологической проблемы процесса классификации и районирования. Этот процесс содержит также другие проблемы: выбор черт районирования, стандартизация и взвешивание черт, метод подсчета разниц между исследуемыми единицами. Все эти этапы важны и влияют на окончательный результат, однако упорядочивание и деление исследованных единиц и до настоящего времени — это самый трудный этап данного процесса (особенно при большом числе основных единиц отнесения).

Попытка решения этого процесса является очень существенной для прогресса исследований объективизации всего процесса районирования.

База упорядочивания (в данных примерах) — таблица средних разниц, представляющая сравнение каждой исследованной единицы с каждой другой (табл. 1). Таблица разниц была подсчитана на основе комплекса 6 главных культур и дает возможность выделить регионы растительной продукции (полевой) в Люблинском воеводстве в 1965 г.

Проблема упорядочивания и деления исследованных единиц решена при помощи трех основных методов: диаграммы Чекановского, Вроцлавского дендрита и „связывающего дерева“. Применение этих методов было вызвано ограниченными возможностями и относительной безошибочностью всех примененных методов исследования, Оно требовало их соответствующего подбора для разных этапов, определения их роли в этом упорядочивании и уточнения некоторых методов.

Так, например, главный недостаток метода Чекановского — неточное определение (когда неизвестно, есть ли упорядочивание диаграммы действительно наилучшим), т.е. необъективность, что было ликвидировано добавочным уточнением принципов упорядочивания диаграммы, которое дало возможность установить наилучшее упорядочивание однозначным способом.

За наилучшее упорядочивание принято такое, которое имеет самую высокую концентрацию минимальных расстояний (разниц) около диагонали и самую высокую концентрацию максимальных разниц в наиболее отдаленной от диагонали зоне. Соответственно этому предположению из двух упорядочиваний то есть правильным, в котором сумма очередных значений кумулятивного ряда есть меньшая (стр. 182).

Также введение весов „связывающего дерева“ соответственно величине основных единиц отнесения и соответственно предложенный способ подсчета (без необходимости одновременного применения целой матрицы разниц для цифровой вычислительной машины) даст возможность правильно решить проблему группировки.

Наконец, точное установление принадлежащих единиц к соответствующему региону при помощи исследований изменчивости внутри их и оценка степени значимости разниц между регионами дали бы возможность проведения однозначного деления исследованной территории.

Каждый метод применяли с учетом его свойств. Наравне с одновременным применением трех методов их применяли и отдельно, исключительно в последней фазе группировки. „Связывающее дерево“ было, например, особенно полезным на первом этапе группировки,

т.е. при объединении наиболее подобных громад. Дендрит был очень полезным в том смысле, что упростил предварительное упорядочивание диаграммы Чекановского. Зато диаграмма есть незаменима при анализе связей между исследованными единицами, а также между выделенными региональными единицами.

Основная цель работы — получение наиболее объективных результатов. Для этого была использована электронная техника. Она дала возможность решить те проблемы, которые нельзя было разрешить при помощи традиционных методов.

Возможность проведения классификации практически неограниченного числа единиц имеет большое значение потому, что при геологических исследованиях число наблюдений (единиц) увеличивается по мере выполнения более подробных исследований. Это должно, по мнению автора, послужить стимулом для разрешения не только конкретных проблем, а также для дальнейшего усовершенствования таксономических методов.

ОБЪЯСНЕНИЕ ТАБЛИЦ И РИСУНКОВ

Табл. 1. Таблица средних разниц, 6 главных культур, 1965 г. (повяты).

Табл. 2. Пример матрицы для 5/17 этапа: а) в единичной системе, б) центроидальной системе.

Табл. 3. Таблица „связывающего дерева” для 6 главных культур, 1965 г. (повяты).

Табл. 4. Таблица разниц для III этапа. Предварительное упорядочивание (для определения принадлежности комплекса 517).

Табл. 5. Величины вариантности ($\hat{\sigma}_x$ и V_x) для типологических черт в разных вариантах регионов IV и V.

Табл. 6. Критерий значимости „t” и „с” между средними значениями типологических черт по регионам.

Рис. 1. Пример модифицированного (численного) способа упорядочивания таблицы разниц: а) предварительно упорядоченная таблица разниц, б) градиенты в таблице разниц.

Рис. 2. Этапы дендритического упорядочивания: а) первый этап, б) второй этап, с) третий этап.

Рис. 3. Схема упорядочивания повятов при помощи „связывающего дерева”: а) „связывающее дерево”, б) очередные разницы.

Рис. 4. Упорядочивание повятов по: А — диаграмме Чекановского (где соответствующие сигнатуры обозначают разницы: а) 0—3, б) 3—5, с) 5—7, d) 7—10, е) больше 10); В — Вроцлавскому дендриту, С — „связывающему дереву”.

Рис. 5. Комплекс громад, 6 главных культур, 1965 г., Номера комплексов соответствуют номерам на рис. 7.

Рис. 6. Границы регионов растительной продукции (полевой), 1965 г.: 1) границы первого ряда, 2) границы второго ряда.

Рис. 7. Разности в группировке при условии, что за основу приняты громады или комплексы громад. А — последний фрагмент „связывающего дерева” первого

этапа группировки (громады). В — „связывающее дерево” второго этапа группировки (комплексы громад).

Рис. 8. Дендрит для громад, 6 главных культур, 1965 г.

Рис. 9. „Связывающее дерево” для громад, 6 главных культур, 1965 г.

S U M M A R Y

The subject presented by the author is a part of a wider methodological problem of the process of classification and regionalization, comprising also a variety of further problems dealing with the selection of regionalizing features, standardization and weighing of these features, and the method of calculating differences between the investigated units. All these stages are important and to some definite degree affect the final result; but so far the problem of how to arrange and subdivide the examined units has been the most difficult stage in this process (especially when a large number of basic units of reference are involved). Hence dealing with this stage of regionalization, and attempting to reach a solution represents an essential element in the progress of research towards making the whole process of regionalization unprejudiced.

In the example cited the author used, as basis for his arranging a table of mean differences showing a comparison of every investigated unit with each other unit (Table 1). This table of differences has been calculated on the basis of a set of six principal crops, and based on this classification, different regions of agricultural produce in the Lublin Voivodeship can be distinguished for the year 1965.

The author solved the problem of setting in proper order and of subdividing the investigated units by the adoption of three basic methods: Czekanowski's diagram, the so-called Wrocław dendrite, and linkage tree. The fact that three different techniques of arranging have been applied, results from the limited applicability and the merely relative infallibility of all research methods commonly in use, and demands their suitable adaptation, i.e. the definition of the part they are to play in this process of arranging and, moreover requires specifying some of the methods more accurately.

Thus, for instance, the principal objection to Czekanowski's method, asserting that it lacks an accurate definition (since it fails to indicate whether the arrangement of the diagram is indeed the best possible) and therefore should be considered biased, has been refuted by the additional clear definition of what arranging the diagram is aiming at, so that now the most favourable arrangement can be indisputably defined.

The author considers best that manner of arranging which brings about the highest concentration of least distances (i.e. of differences) around the diagonal and, at the same time, the highest concentration of greatest differences in the zone farthest away from the diagonal. In the aspect of this presupposition, of the two ways of arranging more correct is the one for which the sum of successive values in a cumulative series is smaller (for the relevant formula see p. 182).

In a similar way the introduction of weighing into the linkage tree, conforming to the values of the basic units of reference, as well as the method of calculating suggested by the author, which eliminates the necessity of introducing at the same time the whole matrix of differences into the memory of the computer, opened the way for the correct solution of the problem of grouping.

Finally, the close determination of the appurtenance of units to their respective regions by studies of the variability occurring within regions, and the estimate of the degree of importance of differences between regions made it possible to uncontestedly demarcate an investigated area.

Taking into account the specific properties of particular methods the author applied methods conforming to these properties. And, apart from adopting all three methods simultaneously, especially in the final stage of grouping, he also used them individually for special purposes. The linkage tree, for instance, proved particularly useful in the first stage of grouping when it came to combine villages which resemble each other the most. The dendrite method was helpful as a means facilitating the preliminary proper arrangement of Czekanowski's diagram. On the other hand, all the same the diagram is and remains the irreplaceable means for analyzing interlinkings between investigated units as well as between regional units already distinguished.

The author's aim in preparing this paper was to obtain the results as unbiased as possible. He managed to put this intention into effect by applying the computer technique of calculating. This indeed not only facilitated time-consuming calculations, but also made it possible to solve such problems as were not to be solved by traditional methods.

This possibility of accomplishing the classification of a practically unlimited number of units is the more important, since in geographic research the number of observations (or units) increases at the rate the given research is going into greater detail. In the author's opinion this should stimulate the inclination to undertake investigations not only for solving definite problems but also to further improve our taxonomic methods.

Papier druk. sat. III kl. 80 g

Format B5 (70×100)

Stron druku: 46+wkł.

Annales UMCS, Lublin 1970

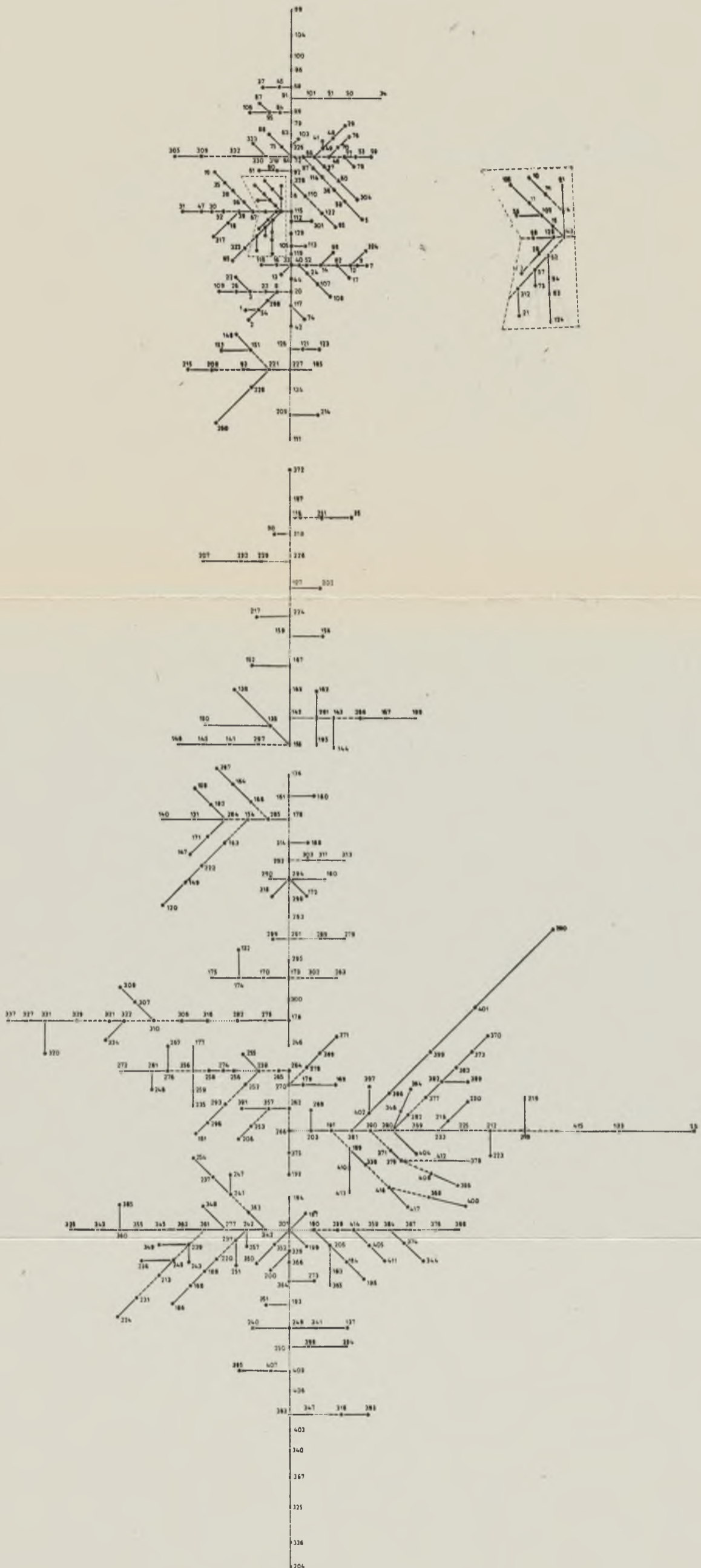
Drukarnia Uniwersytecka w Lublinie

Zam. nr 275 z dn. 5 V 1971

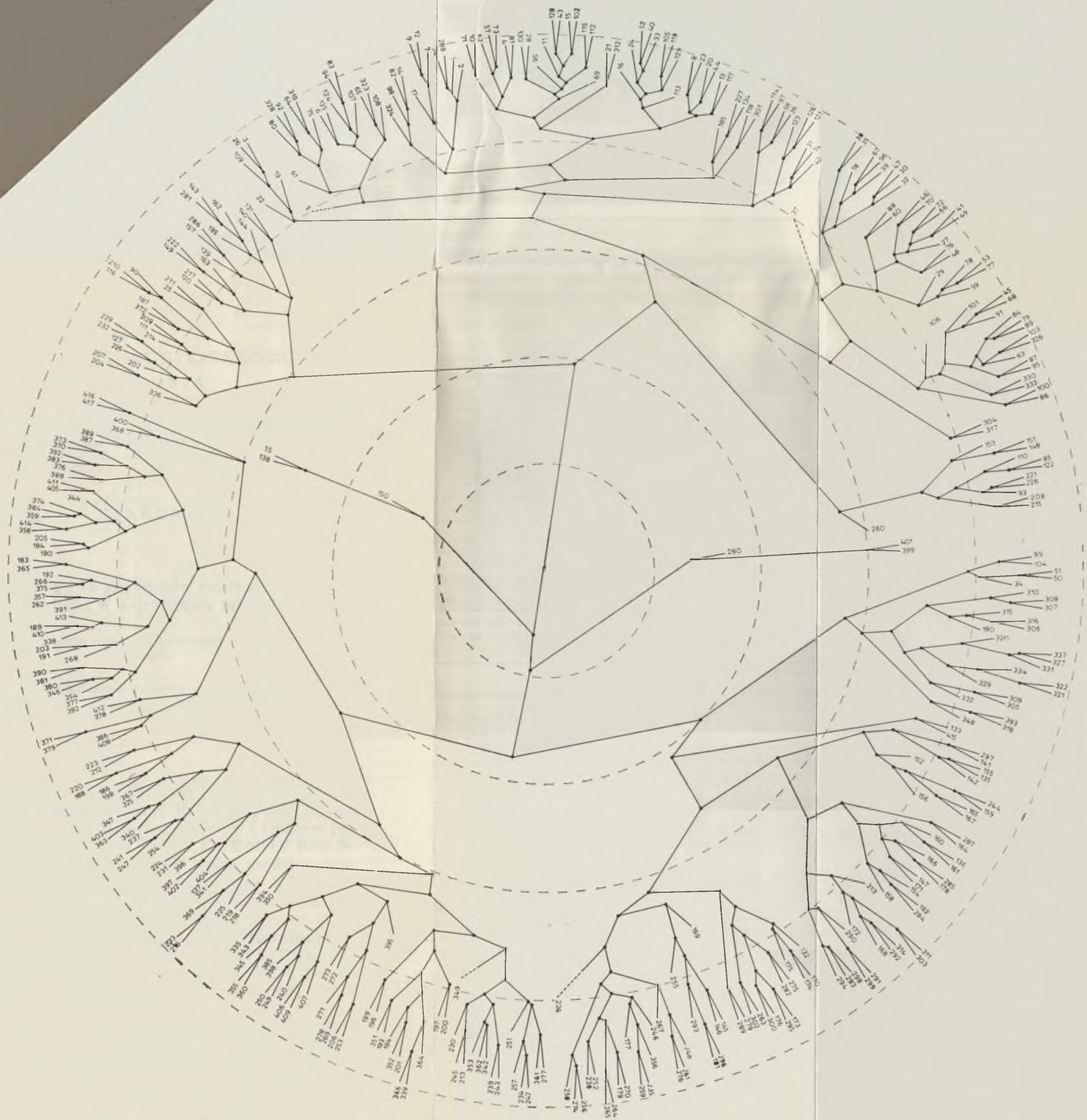
Nakład 900+125 egz. B-4

Maszynopis otrzymano w maju 1971

Druk ukończ.: styczeń 1972



Ryc. 8. Dendryt dla gromad — 6 głównych upraw — r. 1965
 Dendrite for villages with regard to six basic crops, for 1965



Ryc. 9. Drzewo połączeń dla gromad—6 głównych upraw—r. 1965
 Linkage tree for villages, with regard to six basic crops, for 1965

