



METODYKA

Bogdan Brzeziecki

Katedra Hodowli Lasu
Wydziału Leśnego SGGW-AR
ul. Rakowiecka 26/30
02-528 Warszawa

Wykorzystanie programów z pakietu Wildiego i Orlókiego do analizy danych fitosocjologicznych

The use of Wildi—Orlóci package
in analysis of phytosociological data

1. Wstęp

Możliwością zastosowania metod statystycznych w analizie danych fitosocjologicznych zainteresowano się stosunkowo wcześniej (Kulczyński 1927, Motyka 1947, Matuszkiewicz 1948, Matuszkiewicz i Polakowska 1955), dążąc do większego obiektywizmu w badaniach nad zbiorowiskami roślinnymi. Duża pracochłonność znanych wówczas algorytmów (np. diagramu Czekanowskiego i dendrytu wrocławskiego), przy braku odpowiednich środków obliczeniowych, uniemożliwiała jednak szersze wykorzystanie tych metod w praktyce badawczej. Między innymi z tych właśnie względów fitosocjologia pozostawała przez dłuższy czas, nie tylko zresztą w Polsce, nauką opierającą się w większym stopniu na doświadczeniu i intuicji badaczy niż na sformalizowanych analizach matematyczno-statystycznych. Znaczenie tych ostatnich mogło wzrosnąć dopiero w wyniku szybszego rozwoju elektronicznej techniki obliczeniowej i poprawy parametrów użytkowych komputerów (wielkości pamięci i szybkości przeprowadzania operacji).

Datującemu się od początku lat sześćdziesiątych rozwojowi komputeryzacji towarzyszy szybki wzrost zastosowań metod statystycznych, głównie metod wielu zmiennych (multivariate methods) w fitosocjologii. Wykorzystanie tych technik na szerszą skalę ułatwia fakt, że autorzy nowych zastosowań bardzo często publikują tzw. listingi programów komputerowych, opracowanych specjalnie na potrzeby fitosocjologów. Przykładowo już od 1975 r. są dostępne programy, opracowane w BASIC-u przez Orlókiego (1975), poświęcone różnym metodom ordynacji (porządkowania) i klasyfikacji numerycznej fitocenoz. Do znanych i często wykorzystywanych programów należy także seria programów ekologicznych Uniwersytetu w Cornell, np. programy DECORANA (Hill 1979a) i TWINSPAN (Hill 1979b).

Wśród programów, które ukazały się dotychczas, na wyjątkową uwa-

gę zasługuje także, i to z kilku powodów, pakiet, którego autorami są Wildi i Orłóci (1983). Poszczególne programy wchodzące w skład tego pakietu stanowią elementy zintegrowanego systemu. Można je wykorzystywać w różnych konfiguracjach i połączeniach w zależności od założonych celów analiz. W skład systemu weszły tylko te metody, które oparły się próbie czasu i rzeczywiście są najczęściej wykorzystywane przy obróbce danych fitosocjologicznych. Rdzeniem systemu są programy opracowane dla różnych metod wielu zmiennych (np. analiza głównych składowych, analiza kanoniczna, wzajemne uśrednianie, analiza skupień), tym niemniej pakiet zawiera także programy, które mogą znacznie ułatwić opracowywanie surowych materiałów fitosocjologicznych metodami „klasycznymi”. Z tych względów może on, jak się wydaje, zainteresować każdego fitosocjologa.

Poniżej przedstawiono najważniejsze kierunki wykorzystania metod numerycznych w analizie danych fitosocjologicznych i na tym tle omówiono poszczególne programy z pakietu Wildiego i Orłóci. W pracy wykorzystano opracowanie Wildiego i Orłóci (1983) oraz własne doświadczenia, uzyskane w trakcie stażu naukowego w Szwajcarskim Instytucie Badawczym Leśnictwa w Birmensdorfie.

2. Funkcje metod numerycznych

2.1. Wprowadzenie

Surowe dane fitosocjologiczne, uzyskane w wyniku prac terenowych, muszą być opracowane w postaci umożliwiającej bezpośrednią analizę i interpretację. Ze względu na dużą objętość typowych danych, główny problem przy ich przetwarzaniu sprowadza się najczęściej do odpowiedniej redukcji, nie powodującej nadmiernych strat informacji. Problem ten usiłuje się rozwiązać bądź przez zmniejszenie liczby obiektów (zdjęć) i wyróżnienie grup zdjęć o podobnym składzie florystycznym (metody klasyfikacji), bądź też przez redukcję liczby cech (gatunków) i zastąpienie ich czynnikami środowiska (metody ordynacyjne). Stosowane do tych celów algorytmy stanowią zwykle końcowe etapy opracowywania danych i wymagają na ogół przeprowadzenia szeregu operacji wstępnych, takich jak transformacja danych, ustalenie rang gatunków, obliczenie współczynników podobieństwa itp. Pozwala to na wyróżnienie w ramach szeroko pojętej analizy statystycznej danych fitosocjologicznych pewnych typowych problemów, przy rozwiązywaniu których mogą znaleźć zastosowanie metody numeryczne. Najważniejsze z tych problemów przedstawiono w ogólnym zarysie niżej, wykorzystując przy tym proste przykłady w celu zilustrowania podstawowych zasad i idei. Bardziej szczegółowe informacje dotyczące omawianej tu problematyki można znaleźć m.in. w opracowaniach Mueller-Domboisa i Ellenberga (1974), Pielou (1977), Orłóci (1978), Greena (1979) oraz Gaucha (1981).

2.2. Transformacja

Przed przystąpieniem do właściwych obliczeń „surowe” dane fitosocjologiczne zwykle muszą być w odpowiedni sposób przekształcone (transformowane). Do najbardziej typowych transformacji należą:

— Zastąpienie kodów alfanumerycznych ze skali Brauna-Blanqueta wartościami liczbowymi. Jedno z możliwych rozwiązań zostało zaproponowane przez K u h n a (1983; również van der M a a r e l 1979). Zgodnie z propozycją tego autora symbole Brauna-Blanqueta zastępuje się odpowiednimi wartościami liczbowymi:

— — 0; r — 30; + — 50; 1 — 58; 2 — 65; 3 — 71; 4 — 74; 5 — 75.

— Transformacja typu signum, która polega na zastąpieniu wszystkich wartości różnych od zera — wartością 1, dzięki czemu uzyskuje się dane typu jest albo nie ma (1/0). Przykładowy wektor danych

[1 23 0 5 17]

po transformacji signum składa się z następującego ciągu liczb:

[1 1 0 1 1]

— Przydzielanie wag. Elementy danych są mnożone lub dzielone przez pewną wielkość, np. sumę składowych wektorów danych. Przykładowo wektor

[1 2 0 1 1]

dla którego suma składowych wynosi 5, po podzieleniu każdej wartości przez tę sumę, zamienia się na wektor

[1/5 2/5 0 1/5 1/5]

dla którego suma składowych wynosi 1.

W praktyce badań fitosocjologicznych potrzeba wstępnej transformacji danych zachodzi bardzo często. Stosowanie skali Brauna-Blanqueta wymaga zamiany symboli alfanumerycznych (r, +, 1, 2 itd.) na wartości liczbowe. Transformacja typu signum stosowana jest wówczas, gdy z jakiegoś względu istotniejszy jest sam fakt występowania gatunków niż ich ilościowość. (Niestety, nie są możliwe transformacje odwrotne, tzn. zamiana danych typu 0/1 na dane ilościowe!) Używanie różnych jednostek, np. przy pomiarach cech glebowych, także narzuca konieczność transformacji zapewniającej porównywalność cech, co najczęściej osiąga się poprzez odpowiednią standaryzację wektorów danych (jednostkowa długość wektora, pomiar w jednostkach odchylenia standardowego itp.). Szerzej o problemie transformacji danych i wpływie rodzaju transformacji na wyniki dalszych analiz numerycznych traktuje m.in. opracowanie O r l ó c i e g o (1978).

2.3. Analiza podobieństwa

W celu określenia ogólnego podobieństwa albo niepodobieństwa obiektów opisywanych przez dużą liczbę cech zaproponowano szereg różnych współczynników. Aby określić pełną strukturę podobieństwa dla zbioru składającego się z n obiektów (np. zdjęć), należy obliczyć $n(n-1)/2$ współczynników podobieństwa (similarity coefficients) lub taką samą liczbę komplementarnych współczynników niepodobieństwa (dissimilarity coefficients). Współczynnik niepodobieństwa można zdefiniować w prosty sposób, np. jako sumę absolutnych różnic między elementami dwóch wektorów danych:

$$\begin{array}{ll} \text{zdjęcie 1} & [1 \ 3 \ 0 \ 3] \\ \text{zdjęcie 2} & [2 \ 0 \ 0 \ 1] \\ \text{różnice } |d| & 1 \ 3 \ 0 \ 2, \quad \Sigma |d| = 6. \end{array}$$

W tym przypadku liczba 6 jest miarą niepodobieństwa zdjęć 1 i 2. Im większe jest niepodobieństwo rozważanych zdjęć, tym większa jest wartość $\Sigma |d|$ i odwrotnie. Podobnie współczynnik podobieństwa można określić przykładowo jako sumę iloczynów elementów dwóch wektorów danych:

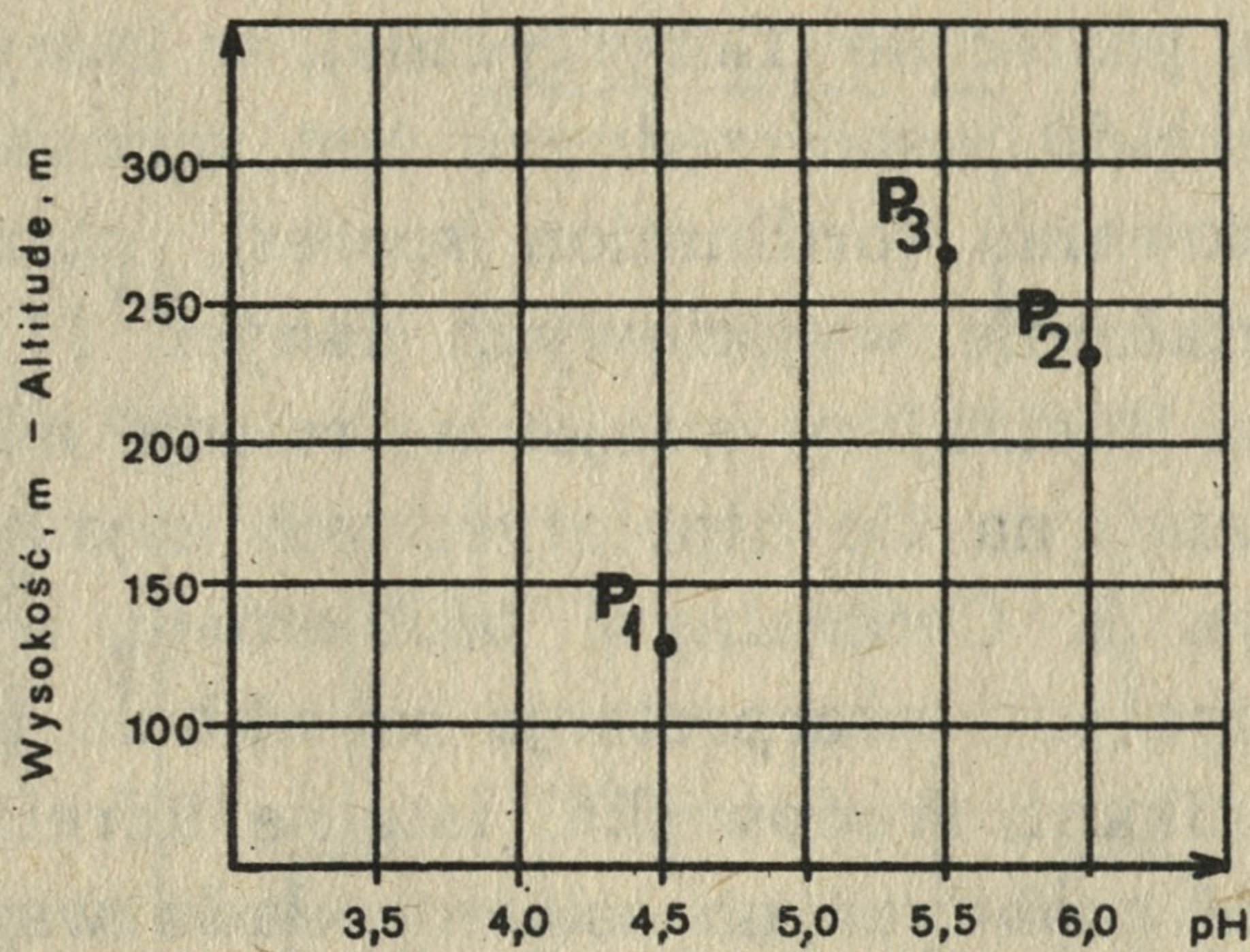
$$\begin{array}{ll} \text{zdjęcie 1} & [1 \ 3 \ 0 \ 3] \\ \text{zdjęcie 2} & [2 \ 0 \ 0 \ 1] \\ \text{iloczyny } s & 2 \ 0 \ 0 \ 3, \quad \Sigma s = 5. \end{array}$$

Dla tego przykładu liczba 5 jest miarą wzajemnego podobieństwa zdjęć 1 i 2. Oczywiście obie miary są bardzo niedoskonałe (nie mają ustalonej górnej granicy oraz zależą od liczby gatunków) i zostały podane tu tylko w celu ilustracji problemu. W praktyce wykorzystuje się bardziej złożone miary, określane często w taki sposób, żeby wartości współczynników zmieniały się w pewnych ustalonych granicach, np. od 0 do 100%, jak to jest w często stosowanym w Polsce wzorze Jaccarda ($P = 2c / (a + b) \times 100\%$, gdzie a i b oznaczają liczby gatunków w porównywanych zdjęciach, natomiast c jest to liczba gatunków wspólnych).

2.4. Porządkowanie (ordynacja)

Głównym celem porządkowania jest ustalenie kolejności zdjęć lub gatunków na osiach układu współrzędnych i interpretacja zaobserwowanych trendów w kategoriach czynników środowiska. Wyniki prezentowane są z reguły w postaci dwuwymiarowych diagramów, na których punkty wyobrażają powierzchnie próbne (zdjęcia fitosocjologiczne) lub gatunki. Można wyróżnić dwa główne rodzaje porządkowania:

— Porządkowanie bezpośrednie (środowiskowe), w którym współrzędne zdjęć wyznaczane są bezpośrednio na podstawie pomierzonych



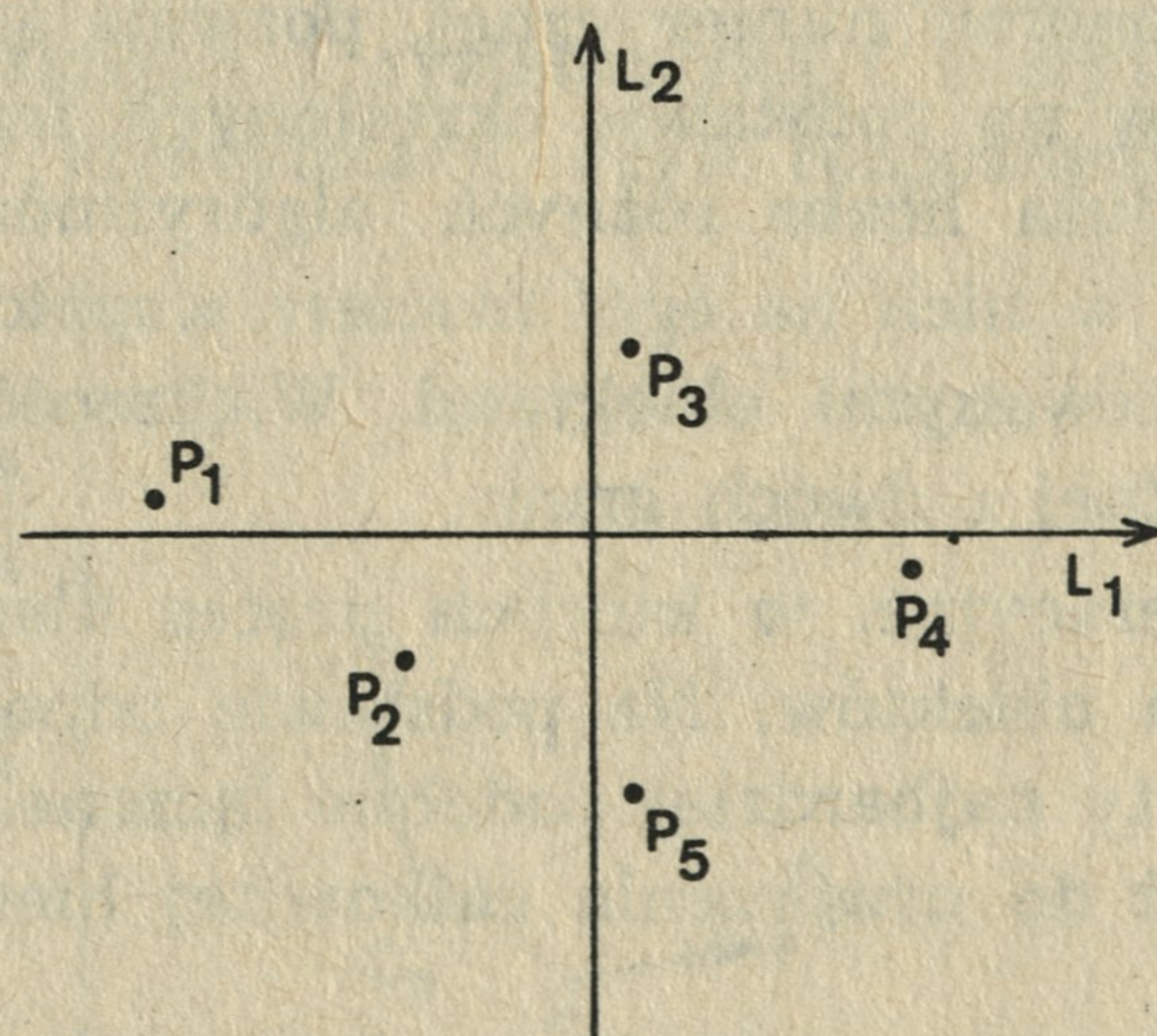
Rys. 1. Porządkowanie bezpośrednie; przykładowy diagram ilustrujący łączny wpływ dwóch zmiennych (pH i wysokość npm) na wzajemne podobieństwo zdjęć fitosocjologicznych (wg Wildiego i Orłóckiego 1983)

Direct ordination; diagram illustrating the joint effect of two variables (pH and altitude a.s.l.) on mutual resemblance of phytosociological relevés (after Wildi and Orłóci 1983)

wartości czynników środowiskowych. Na przykład na podstawie danych dotyczących zmienności dwóch czynników ekologicznych:

pH	4,5	6,0	5,5
wysokość, m	130	230	270

można sporządzić dwuwymiarowy diagram (rys. 1), który pozwala rozpatrywać łączny wpływ dwóch czynników ekologicznych (w tym przypadku pH i wysokości npm) na podobieństwa między zdjęciami.



Rys. 2. Porządkowanie pośrednie; uszeregowanie zdjęć w układzie o osiach odpowiadających czynnikom środowiska o dominującym wpływie na roślinność. P_1 — P_5 — numery zdjęć, L_1 i L_2 — hipotetyczne czynniki środowiska (wg Brzezieckiego 1987)

Indirect ordination; arrangement of relevés in a system with axes corresponding to environmental factors having a dominant influence on vegetation, P_1 — P_5 — numbers of relevés, L_1 and L_2 — hypothetical environmental factors (after Brzeziecki 1987)

— Porządkowanie pośrednie (florystyczne), w którym położenie obiektów względem osi układu współrzędnych jest wyznaczane na podstawie wskaźników porządkowania (ordination scores), otrzymywanych przez odpowiednie przekształcenie wyjściowych danych florystycznych. Przykładowy diagram ilustrujący wzajemne relacje między powierzchniami próbnymi (zdjęciami) na tle hipotetycznych czynników siedliskowych przedstawiono na rys. 2. Centralnym problemem w metodach porządkowania florystycznego jest interpretacja osi układu współrzędnych i powiązanie ich z czynnikami środowiska. Istotne korelacje między współrzędnymi powierzchni próbnych na osiach układu współrzędnych i czynnikami ekologicznymi są podstawą do wnioskowania o tych czynnikach, które są w największym stopniu odpowiedzialne za przestrzenne zróżnicowanie roślinności. Główne założenia modelu porządkowania florystycznego, zasady otrzymywania wskaźników porządkowania dla zdjęć fitosocjologicznych oraz problemy związane z interpretacją ekologiczną wyników porządkowania omówiono szczegółowo w innej pracy (Brzeziecki 1987).

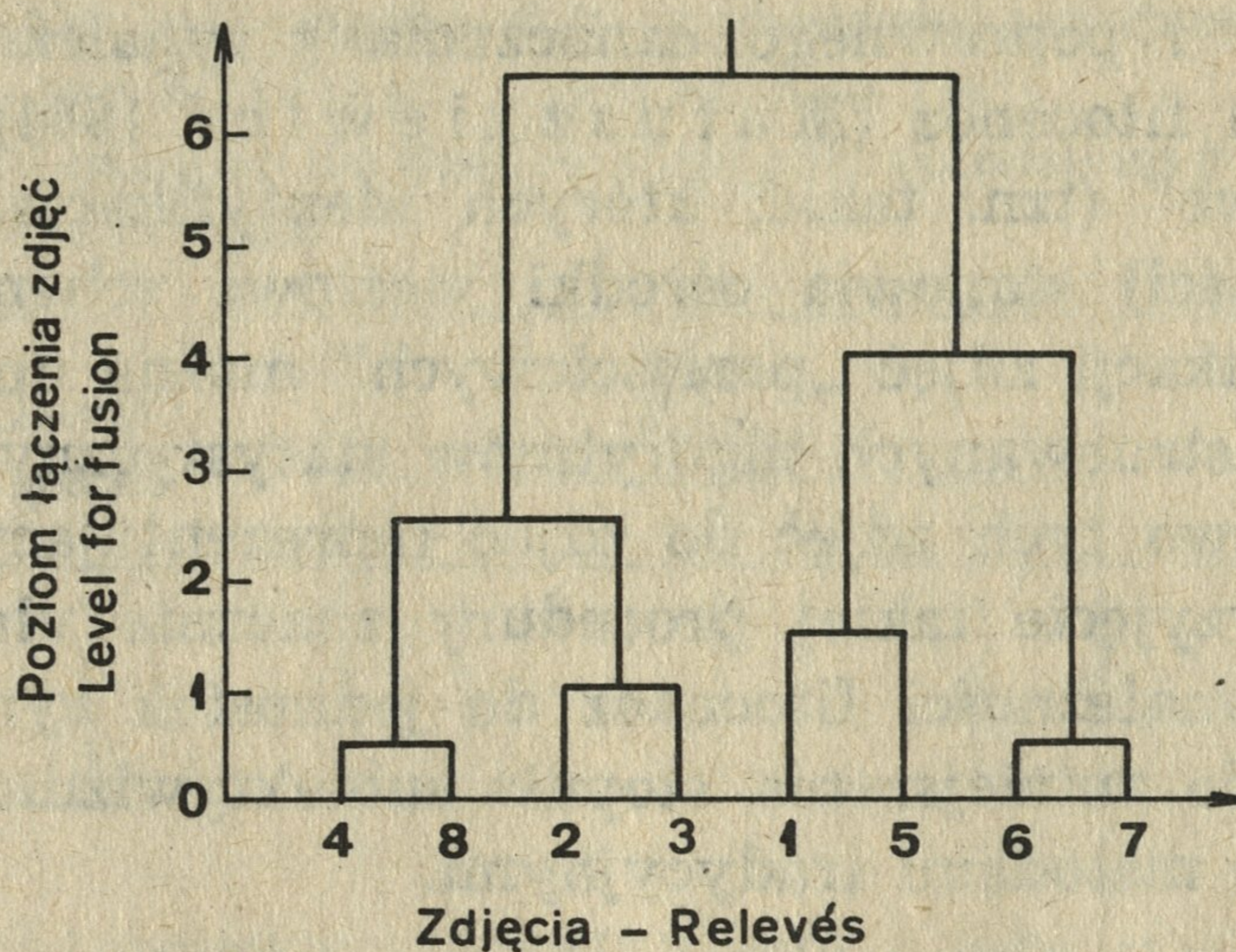
2.5. Klasyfikacja numeryczna (hierarchiczna)

Klasyfikacją w ścisłym logicznym znaczeniu jest taki podział zbioru elementów na podzbiory (klasy, grupy), który prowadzi do powstania systemu wyczerpującego, rozłącznego i konsekwentnie opartego na jednolitych kryteriach wynikających z istotnych cech elementów danego zbioru (Matuszkiewicz 1984). Powyższe warunki spełniają metody współczesnej syntaksonomii numerycznej, pozwalające na dokonanie podziału zbioru obiektów na podstawie określonych formalnych kryteriów. Istnieje stosunkowo duża liczba różnych algorytmów klasyfikacyjnych; punktem wyjścia jest w nich na ogół macierz współczynników podobieństwa między klasyfikowanymi obiektami. Większość tych algorytmów można zaliczyć do jednej z dwóch grup:

— Metody aglomeracyjne, w których proces klasyfikacji rozpoczyna się od poszczególnych obiektów. Na podstawie odpowiednio zdefiniowanych kryteriów obiekty najbardziej podobne łączone są stopniowo w coraz większe grupy, aż do utworzenia całkowitej hierarchii klasyfikacyjnej (rys. 3).

— Metody podziałowe, które prowadzą proces klasyfikacyjny w odwrotnym kierunku, tzn. najpierw dzielą cały zbiór obiektów na dwa podzbiory, następnie każdy z nich na dwa dalsze itd. aż do poziomu pojedynczych obiektów. W rezultacie powstaje dendrogram analogiczny do tego, który przedstawiono na rys. 3.

Jak podkreśla Matuszkiewicz (1984), fitocenozy są niepowtarzalnymi układami syntetycznymi, dlatego podział oparty na zupełnie



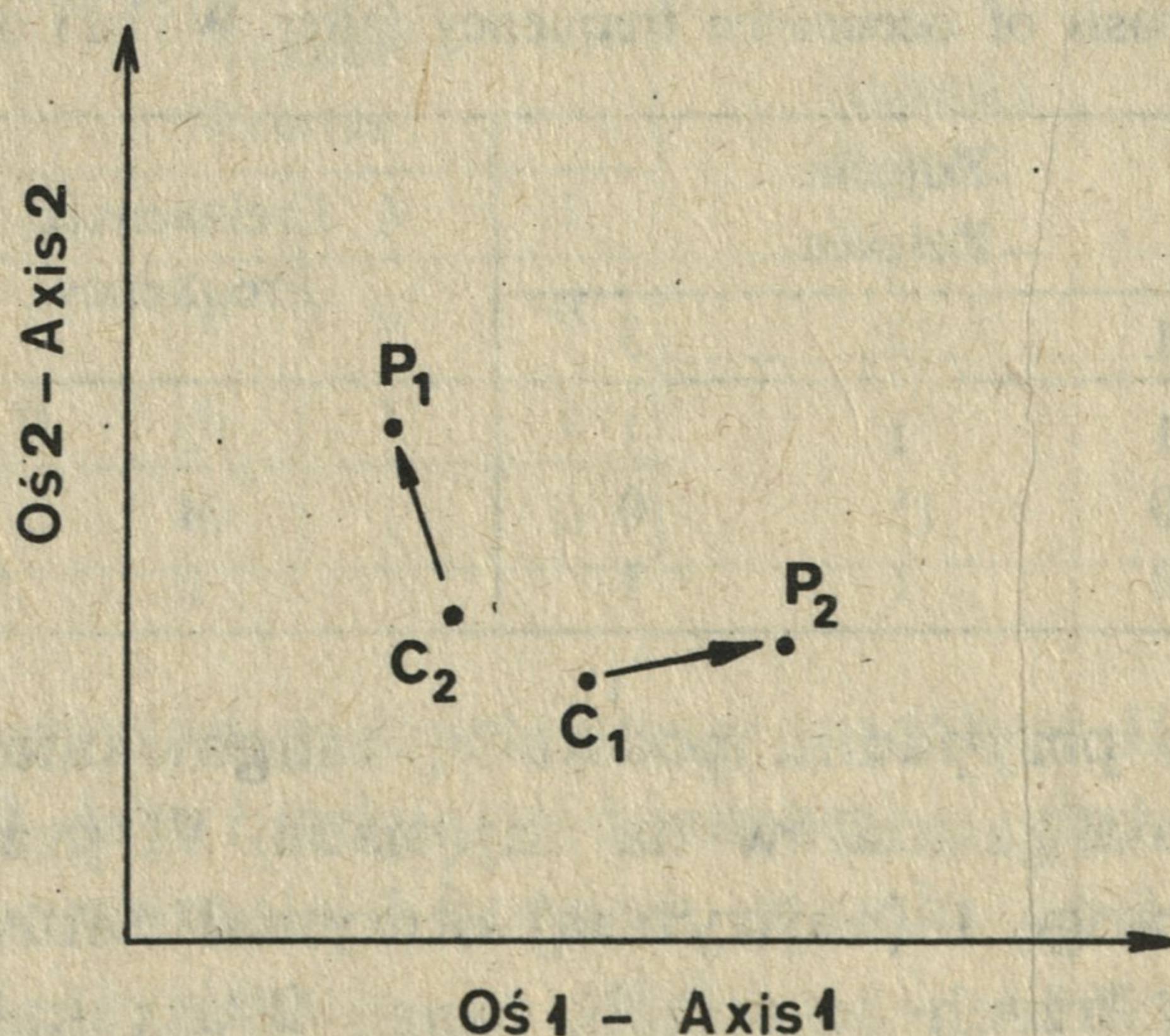
Rys. 3. Hierarchia klasyfikacyjna zdjęć fitosocjologicznych uzyskana za pomocą numerycznych metod klasyfikacyjnych; przykładowy dendrogram

Hierarchical classification of phytosociological relevés obtained by means of numerical classification methods; an exemplary dendrogram

sztywnych formalnych kryteriach nie zawsze jest podziałem naturalnym i najlepszym. Mimo to klasyfikacja numeryczna zdjęć z pewnością może stanowić istotną pomoc przy wyróżnianiu jednostek syntaksonomicznych różnego rzędu. Przykłady wykorzystania klasyfikacji numerycznej do takich celów zawierają m. in. prace Dzwonki (1977a, 1977b, 1984)

2.6. Identyfikacja

Na obecnym etapie rozwoju badań fitosocjologicznych maleje znaczenie problemów związanych z systematyką i klasyfikacją fitocenoz, natomiast wzrasta rola identyfikacji, czyli rozpoznania zbiorowisk roślin-



Rys. 4. Identyfikacja przynależności syntaksonomicznej zdjęć „przejściowych” (c_1 i c_2). Punkty P_1 i P_2 — zdjęcia „typowe” (wg Wildiego i Orłóci 1983)
Establishing the syntaxonomical status for “transitional” relevés (c_1 and c_2). Points P_1 and P_2 — “typical” relevés (after Wildi and Orłóci 1983)

nych jako typów i poprawnego oznaczenia przynależności syntaksonomicznej badanych fitocenoz (Matuszkiewicz 1984). Przy założeniu, że zdjęcia „typowe” (tzn. takie, których identyfikacja nie przedstawia większych trudności) stanowią ośrodki węzłowe schematu syntaksonomicznego, identyfikacji zdjęć „przejściowych” można dokonać za pomocą odpowiednio skonstruowanych algorytmów statystycznych, analizujących ogólne podobieństwo tych zdjęć do zdjęć reprezentujących jednostki typowe (rys. 4). Przyjęcie takiej procedury znacznie ułatwia praktyczną identyfikację przynależności fitocenoz do jednostek syntaksonomicznych i przyczynia się do zmniejszenia stopnia subiektywizmu, nieuniknionego przy identyfikacji metodami tradycyjnymi.

2.7. Rangowanie

Rangowanie należy do tych operacji, które poprzedzają analizę danych za pomocą bardziej skomplikowanych i czasochłonnych algorytmów klasyfikacyjnych i ordynacyjnych. Celem rangowania jest uszeregowanie gatunków pod względem ich udziału w całkowitej zmienności występującej w analizowanym zbiorze danych. Gatunek, którego udział jest największy, otrzymuje rangę najwyższą, gatunek o kolejnym udziale rangę o jeden mniejszą itd. aż do całkowitego wyczerpania listy gatunków. Z praktyki wiadomo, że gatunki o najniższych udziałach mają znikomy wpływ na wyniki porządkowania i klasyfikacji zdjęć metodami numerycznymi, dlatego często można z nich zrezygnować, uzyskując w zamian przyspieszenie obliczeń oraz oszczędność czasu i pamięci komputera.

Tab. I. Ustalanie rang gatunków na podstawie częstotliwości występowania (wg Wildiego i Orlóciego 1983)

Species ranking on the basis of occurrence frequency (after Wildi and Orlóci 1983)

Gatunki Species	Zdjęcia Relevés			Frekwencja Frequency	Ranga Rank
	1	2	3		
1	1	1	1	3	1
2	0	1	0	1	3
3	0	1	1	2	2

W najprostszym przypadku podstawą rangowania może być częstotliwość występowania gatunków na zdjęciach. W przedstawionym w tabeli I przykładzie rangę 1 (najwyższą) otrzymał gatunek 1, który występuje we wszystkich trzech zdjęciach, rangę 2 uzyskał gatunek 3 (obecność w dwóch zdjęciach) i rangę 3 gatunek 2, występujący tylko w jednym zdjęciu. Inny, ulepszony sposób rangowania uwzględnia nie tylko częstotliwość, ale także korelacje między gatunkami. W przykładzie przedstawionym w tab. II gatunek 2 otrzymał rangę 3 (najniższą), ponie-

Tab. II. Ustalenie rang gatunków na podstawie korelacji międzygatunkowych (wg Wildiego i Orłóciego 1983)

Species ranking on the basis of interspecific correlations (after Wildi and Orłóci 1983)

Gatunki Species	Zdjęcia Relevés			Ranga Rank
	1	2	3	
1	1	1	0	1
2	1	1	0	3
3	0	1	1	2

waż jest silnie skorelowany z gatunkiem 1 (oba występują na tych samych zdjęciach), gatunek 3 otrzymał rangę 2, ponieważ uwzględnia mniej korelacji niż gatunek 1, który otrzymuje rangę 1 (najwyższą).

2.8. Ocena wyników klasyfikacji

Miarą „sukcesu” klasyfikacji jest stopień wzajemnej odpowiedniości wyróżnionych grup zdjęć i gatunków (stopień strukturyzacji tabeli fitosocjologicznej). Odpowiednie metody statystyczne, takie jak analiza kanoniczna, umożliwiają ilościowe porównywanie „dobroci” różnych klasyfikacji (np. klasyfikacji a priori i klasyfikacji numerycznej). Zasadę tego

Tab. III. Analiza stopnia strukturyzacji tabeli fitosocjologicznej (wg Wildiego i Orłóciego 1983, zmodyfikowane)

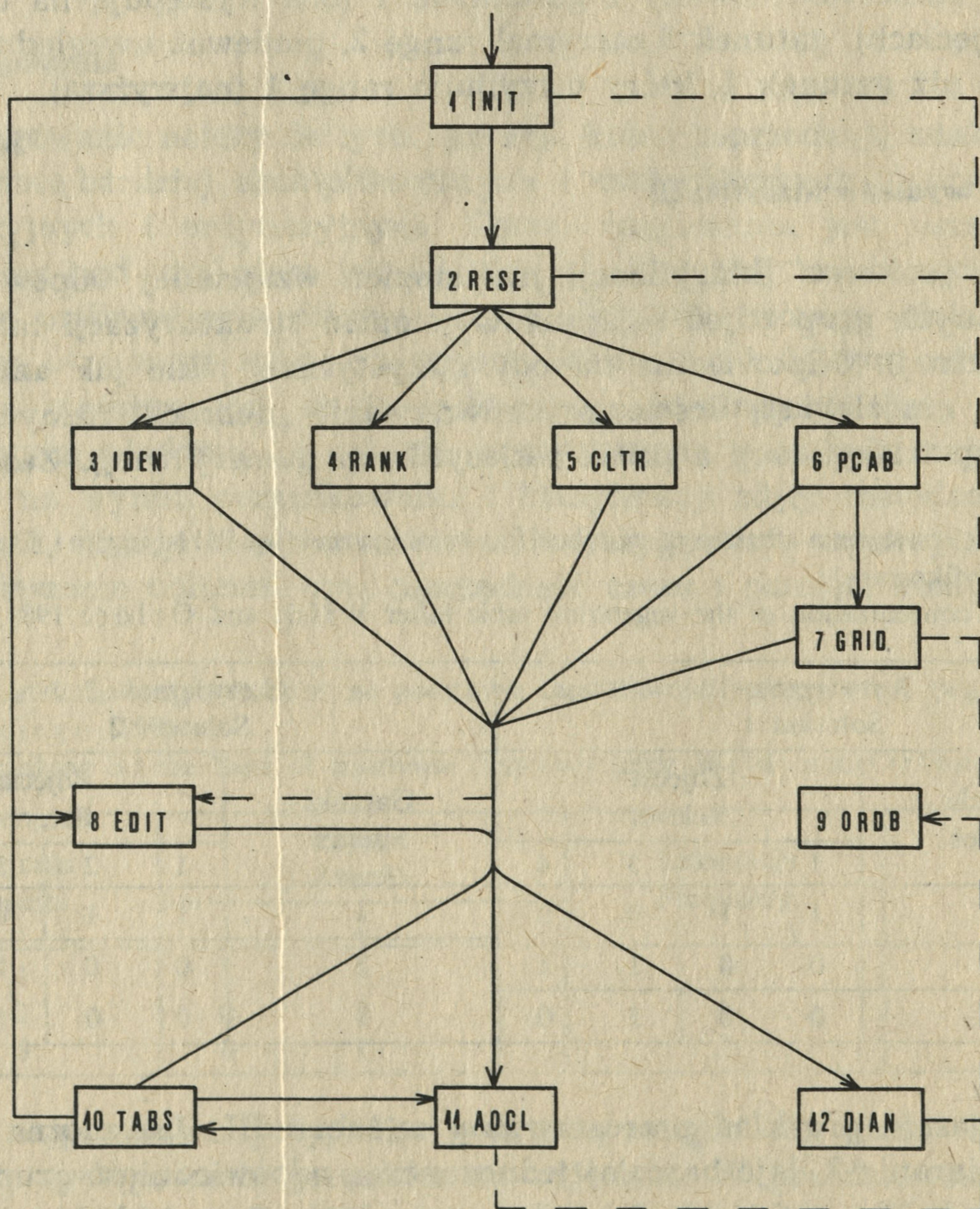
Analysis of concentration of the vegetation table (after Wildi and Orłóci 1983, modified)

Rozwiązanie 1 Solution 1					Rozwiązanie 2 Solution 2				
Gatunki Species	Zdjęcia Relevés				Gatunki Species	Zdjęcia Relevés			
	1	2	3	4		1	2	3	4
1	1	1	0	0	1	1	1	0	0
2	0	0	1	1	2	0	0	1	1
3	0	0	1	0	3	0	0	1	0

testu wyjaśnia przykład przedstawiony w tabeli III. Jak można zauważyć, rozwiązanie 2 daje bardziej jednoznaczną odpowiedniość grup gatunków i zdjęć. W rozwiązaniu 1 bowiem gatunki 1 i 2 znalazły się w jednej grupie, mimo że występują w różnych grupach zdjęć! Natomiast w rozwiązaniu 2 gatunek 1 jest w oczywisty sposób związany z pierwszą grupą zdjęć o numerach 1 i 2, podczas gdy gatunki 2 i 3, stanowiące wspólną grupę, występują tylko w drugiej grupie zdjęć, oznaczonych jako 3 i 4. Z tych względów rozwiązanie 2 należy uznać za poprawniejsze.

3. Charakterystyka programów z pakietu Wildiego i Orłóci

Większość przedstawionych wyżej funkcji można zrealizować za pomocą 12 programów komputerowych wchodzących w skład zintegrowanego pakietu, którego autorami są Wildi i Orłóci (1983). Programy zostały napisane w języku FORTRAN 77 i zainstalowane pod systemem operacyjnym UNIX na komputerze Cyber CDC, znajdującym się w Politechnice Federalnej w Zurychu. Kopie programów oraz informacje dotyczące szczegółów związanych z implementacją programów na innych komputerach można uzyskać od dra O. Wildiego ze Szwajcarskiego Instytutu Badawczego Leśnictwa w Birmensdorfie.



Rys. 5. Ogólna struktura pakietu programów Wildiego i Orłóci. Kierunki logicznych powiązań między programami zaznaczono za pomocą strzałek (wg Wildiego i Orłóci 1983)

General structure of Wildi-Orlóci package. Logical connections between programmes are indicated by arrows (after Wildi and Orłóci 1983)

Na schemacie ilustrującym ogólną strukturę pakietu przedstawiono za pomocą strzałek kierunki logicznych powiązań między programami (rys. 5). Współpraca między nimi polega na tym, że wyniki obliczeń uzyskiwane w jednym programie są w postaci oddzielnych zbiorów wykorzystywane w innych programach. Na przykład strzałka łącząca program RESE z programem PCAB (rys. 5) wskazuje, że ten ostatni wymaga na „wejściu” danych obliczonych wcześniej w programie RESE.

Każdy program może wykonywać z reguły kilka funkcji i każdy oferuje użytkownikowi możliwość wyboru różnych dodatkowych opcji. Krótki opis działania poszczególnych programów przedstawia się następująco.

Program INIT wczytuje dane, oblicza częstość występowania gatunków w zdjęciach i liczbę gatunków w każdym zdjęciu, dokonuje wstępnej transformacji danych (np. zamiana symboli Brauna-Blanqueta na wartości numeryczne), oblicza średnie arytmetyczne dla zdjęć i dla gatunków, drukuje (po wyborze odpowiedniej opcji) macierz danych. W programie znajduje się ponadto 6 różnych opcji, pozwalających na dalszą transformację danych.

W programie RESE następuje obliczanie kwadratowej (symetrycznej) macierzy współczynników podobieństwa, w zależności od wybranej opcji, dla gatunków lub zdjęć. Aktualna wersja programu pozwala na wybór jednej z 7 miar podobieństwa (m.in. odległość euklidesową, kowariancję, współczynnik Ochiai i współczynnik van der Maarela). Obliczona macierz może być na żądanie użytkownika wydrukowana.

Program IDEN umożliwia przydzielanie nowych zdjęć (lub gatunków) do istniejących, zdefiniowanych w innych programach, grup klasyfikacyjnych. Od użytkownika wymaga się m.in. określenia „progu identyfikacyjnego”, czyli minimalnej wartości podobieństwa, pozwalającej na zaliczenie danego obiektu do jednej z istniejących grup.

Program RANK przydziela rangi i ustala kolejność gatunków lub zdjęć pod względem procentowego udziału w całkowitej wariancji zawartej w danych. Gatunek (lub zdjęcie) o największym udziale otrzymuje rangę najwyższą, kolejny rangę o jeden mniejszą itd. Istnieje możliwość określenia minimalnego dopuszczalnego udziału, a tym samym wcześniejszego zakończenia procesu rangowania.

Program CLTR konstruuje klasyfikację hierarchiczną obiektów za pomocą jednego z trzech algorytmów: tzw. metoda pojedynczych połączeń (single linkage), kompletnych połączeń (complete linkage) i minimalnej wariancji (minimum variance). Wyniki drukowane są w postaci dendrogramów, a użytkownik określa m.in. szerokość dendrogramu oraz przypuszczalną liczbę grup.

Program PCAB wykonuje obliczenia związane z porządkowaniem zdjęć i gatunków. Opcje dostępne w programie pozwalają na wybór między analizą głównych składowych i metodą wzajemnego uśredniania.

Program PCAB korzysta z macierzy współczynników podobieństwa (dla zdjęć lub gatunków) wygenerowanych wcześniej w programie RESE. Standardowo drukowane są wskaźniki dla zdjęć i gatunków na pierwszych sześciu osiach wielowymiarowego porządkowania.

Z programem PCAB ściśle współpracuje program GRID który pozwala na uzyskanie dodatkowej klasyfikacji zdjęć lub gatunków, przy której wykorzystuje się współrzędne obiektów na pierwszych czterech osiach porządkowania. Użytkownik określa m.in. liczbę „komórek” klasyfikacyjnych na poszczególnych osiach. Dokładniejszy opis użytego algorytmu można znaleźć w pracy Wildiego (1979).

Do często wykorzystywanych programów należy program EDIT, za pomocą którego można dokonywać dowolnych manipulacji na wyjściowym zbiorze danych, np. redukować liczbę gatunków lub zdjęć, dodawać nowe podzbiory danych, ustalać podział zdjęć i gatunków na grupy, określać a priori ich kolejność w tabeli fitosocjologicznej itp.

Program ORDB stanowi typowy program prezentacyjny (wynikowy), za pomocą którego można szybko otrzymać graficzne diagramy rozrzutu punktów, przedstawiających ułożenie zdjęć lub gatunków w układzie osi porządkowania. Od użytkownika wymaga się m.in. określenia wielkości diagramu oraz odpowiedniej pary osi porządkowania (np. 1 i 2, 1 i 3, 2 i 3 itp). Podobny charakter ma także program TABS, wykorzystywany przy wydruku gotowych, uporządkowanych tabel fitosocjologicznych (na podstawie wyników uzyskanych w innych programach). W programie uwzględniono także możliwość zamieszczania dodatkowych informacji w „główce” tabeli fitosocjologicznej, takich jak wielkość powierzchni zdjęcia, data wykonania, miejscowość itp.

Program AOCL służy do analizy stopnia strukturyzacji uporządkowanych tabel oraz do porównywania efektywności różnych klasyfikacji. W programie DIAN wykorzystano algorytm analizy dyskryminacyjnej, umożliwiający ocenę „siły dyskryminacyjnej” poszczególnych czynników środowiskowych dla wyróżnionych grup lub typów syntaksonomicznych. Daje to podstawy do ekologicznej interpretacji wyników analiz florystycznych.

W zależności od charakteru analizowanych danych oraz założonych celów, należy odpowiednio zaplanować przebieg obliczeń i ustalić, w jakiej kolejności i w jakiej konfiguracji programy powinny być połączone. Z reguły przy analizie konkretnego problemu wykorzystywana jest tylko część programów. Przykładowo, jeżeli założymy, że celem jest przeprowadzenie apriorycznej klasyfikacji zdjęć, to wówczas należałoby wykorzystać następującą konfigurację programów:

INIT — EDIT — INIT — EDIT — TABS — AOCL — ORDB.

W przedstawionej konfiguracji dwukrotnie wywołane są programy INIT i EDIT. Wynika to stąd, że program INIT standardowo rozpoczyna wszy-

stkie obliczenia (wczytuje dane, kreuje zbiory potrzebne w dalszych programach itd.). Natomiast program EDIT za pierwszym razem jest wykorzystywany w celu modyfikacji wejściowego zbioru danych (np. do usunięcia zbędnych zdjęć lub gatunków). Po uporządkowaniu danych wejściowych i ponownym opracowaniu ich przez program INIT, program EDIT jest wywoływany powtórnie, tym razem już w celu podziału całego analizowanego materiału na grupy (typy zdjęć). Opcje programu pozwalają na zdefiniowanie dowolnej liczby grup dla zdjęć i dla gatunków oraz na ustalenie dowolnej kolejności zdjęć i gatunków w wyróżnionych grupach. Po przeprowadzeniu podziału można wykorzystać program prezentacyjny TABS do wydruku uporządkowanych tabel fitosocjologicznych. Jeżeli w tym momencie dojdziemy do wniosku, że w uporządkowanej tabeli należy jeszcze coś poprawić, wówczas można powrócić jeszcze raz do programu EDIT. Wydruk tabeli za pomocą programu TABS mógłby w zasadzie zakończyć ten krok, ale istnieje dodatkowa możliwość sprawdzenia „jakości” klasyfikacji za pomocą programu AOCL. Wyniki analizy kanonicznej uzyskane w tym programie przesyłane są do programu ORDB, który pozwala na graficzne przedstawienie środków grup zdjęć i gatunków względem osi kanonicznych i względem siebie, a tym samym także na wizualną ocenę stopnia strukturyzacji tabeli fitosocjologicznej.

4. Uwagi końcowe

Profesor J. Motyka, będący gorącym zwolennikiem wykorzystania metody Czekanowskiego w fitosocjologii, do uporządkowania tą metodą 280 zdjęć z Podola potrzebował 2 lat, w czasie których poświęcał każdego dnia 3—4 godz. na obliczenia i układanie zdjęć (M o t y k a 1947). Charakterystyczne jest przy tym, że w stosunku do uzyskanych wyników nie uważał tego czasu za długi. Obecnie opracowanie tego materiału podobną, a może nawet lepszą metodą za pomocą średniej klasy komputera jest możliwe dosłownie w ciągu kilku minut (jeżeli nie liczyć czasu potrzebnego na przygotowanie maszynowych nośników danych). Przykład ten ilustruje ogromny postęp, jaki dokonał się w ostatnim czasie w zakresie technicznych i statystycznych możliwości przetwarzania danych fitosocjologicznych.

Gotowe programy, a w jeszcze większym stopniu banki programów w rodzaju przedstawionego tu pakietu Wildiego i Orłóciego znakomicie przyspieszają proces opracowywania danych i odciążają fitosocjologów od konieczności czysto mechanicznej i podatnej na błędy pracy obliczeniowej. Specjalna konstrukcja programów, cechująca się dużą elastycznością i dopuszczająca wykorzystanie wielu opcji, umożliwia użytkownikom

sprawdzanie różnych hipotez i przeprowadzanie eksperymentów metodycznych, których wyniki uzyskiwane są niemal natychmiast.

Wszystko to sprawia, że rozwój fitosocjologii w kierunku „numerycznym” jest właściwie przesądzony. Należy tylko mieć nadzieję, że jego dalszą konsekwencją będzie także postęp w zakresie teorii zbiorowisk roślinnych. Najlepsze bowiem programy i komputery nie zwolnią swych użytkowników od konieczności koncepcyjnego myślenia oraz znajomości głównych założeń wykorzystywanych metod i ich ekologicznej interpretacji. Jak podkreśla K u h n (1983), w im większym stopniu korzysta się ze środków technicznych i numerycznych, tym większego znaczenia nabiera szeroka znajomość problemu, wprawa i doświadczenie.

Piśmiennictwo

- Brzeziecki B. 1987 — Analiza związków między roślinnością i środowiskiem za pomocą modelu porządkowania florystycznego — *Wiad. ekol.* 33: 391—405.
- Dzwonko Z. 1977a — The use of numerical classification in phytosociology — *Fragm. Florist. Geobot.* 23: 327—343.
- Dzwonko Z. 1977b — A numerical classification of the forest communities of the Słonne Góry Mountains (The Polish Eastern Carpathians) — *Fragm. Florist. Geobot.* 23: 345—353.
- Dzwonko Z. 1984 — Klasyfikacja numeryczna zbiorowisk leśnych polskich Karpat — *Fragm. Florist. Geobot.* 30: 93—167.
- Gauch H. G. 1981 — *Multivariate analysis in community ecology* — Cambridge University Press, Cambridge, ss. 298.
- Green R. H. 1979 — *Sampling design and statistical methods for environmental biologists* — Wiley, New York, ss. 257.
- Hill M. O. 1979a — DECORANA — a FORTRAN program for detrended correspondence analysis and reciprocal averaging — *Ecology and Systematics*, Cornell University, Ithaca, New York 14850, USA, ss. 52.
- Hill M. O. 1979b — TWINSPLAN — a FORTRAN program for arranging multivariate data in an ordered two-way table by classification of the individuals and attributes — *Ecology and Systematics*, Cornell University, Ithaca, New York 14850, USA, ss. 90.
- Kuhn N. 1983 — VEGTAB, ein Computer-Program als Hilfe zur tabellarischen Vegetationsgliederung — *Tuexenia*, 3: 499—522.
- Kulczyński S. 1927 — Pflanzenassoziationen der Pieninen — *Extrait Bull. Acad. pol. Sci. Ser. B, Biol.* 57—203.
- Maarel E., van der 1979 — Transformation of cover-abundance values in phytosociology and its effects on community similarity — *Vegetatio*, 39: 97—144.
- Matuszkiewicz W. 1948 — Roślinność lasów okolic Lwowa — *Ann. Univ. Mariae Curie-Skłodowska, Sect. C, Biol.* 3: 119—194.
- Matuszkiewicz W. 1984 — *Przewodnik do oznaczania zbiorowisk roślinnych Polski* — PWN, Warszawa, ss. 298.
- Matuszkiewicz W., Polakowska M. 1955 — Materiały do fitosocjologicznej systematyki borów mieszanych w Polsce — *Acta Soc. Bot. Pol.* 24: 421—458.
- Motyka J. 1947 — O zadaniach i metodach badań geobotanicznych — *Ann. UMCS, Sec. C, Suppl. I*, 1—168.

- Mueller-Dombois D., Ellenberg H. 1974 — Aims and methods of vegetation ecology — Wiley, New York, ss. 547.
- Orlóci L. 1975 — Multivariate analysis in vegetation research — Dr. W. Junk B. V. Publishers, The Hague, ss. 276.
- Orlóci L. 1978 — Multivariate analysis in vegetation research — Dr. W. Junk B. V. Publishers, The Hague, ss. 451.
- Pielou E. C. 1977 — Mathematical ecology — Wiley-Interscience, New York, ss. 385.
- Wildi O. 1979 — GRID — a space density analysis for recognition of nodes in vegetation samples — *Vegetatio*, 41: 95—100.
- Wildi O., Orlóci L. 1983 — Management and multivariate analysis of vegetation data — *EAFV*, Ber. 215, ss. 68.

Summary

Fast development of computer technique since the early nineteen-sixties has created new possibilities for the use of various statistical methods (including the so-called multivariate methods) in analysis and processing of phytosociological data. The most important trends in using these methods for phytosociological purposes are: floristic data transformation, resemblance analysis of relevés and species direct multidimensional ordination (Fig. 1) and indirect one (Fig. 2), hierarchical classification (Fig. 3), establishing the syntaxonomical status for relevés (Fig. 4), species ranking (Tables I, II) and analysis of concentration of vegetation tables (Table III).

Special computer programmes allow to use statistical methods in phytosociology, among which Wildi-Orlóci (1983) package is one of the best. The package contains 12 integrated programmes (Fig. 5) fulfilling the majority of the above mentioned functions. Special construction of programmes allowing, among other things, to choose between various computational options and to plan different courses of analyses depending on the character of data and purpose of research, explains their usefulness as a versatile tool for phytosociologists.

(wpłynęło: 6 V 1988 r.)