

ZBIGNIEW KOBYLIŃSKI

PODSTAWOWE METODY ANALIZY PUNKTOWYCH UKŁADÓW
PRZESTRZENNYCH

Na gruncie ekologii roślin, a następnie także geografii matematycznej, w ciągu ostatnich 30 lat rozwinięto wiele zaawansowanych metod statystycznych pozwalających zidentyfikować przestrzenne struktury analizowanych zdarzeń, a następnie wskazać procesy, które mogły ewentualnie te struktury generować. Również zjawiska badane przez archeologię, takie jak np. rozmieszczenie stanowisk osadniczych w regionie, czy też rozmieszczenie pewnego typu artefaktów na stanowisku archeologicznym, mogą być uznane za realizacje procesów punktowych. Metodami identyfikacji tych właśnie procesów poświęcone będą poniższe uwagi.

METODY STATYSTYCZNE A MOŻLIWOŚCI POZNAWCZE ARCHEOLOGII

Zainteresowanie zastosowaniem statystycznych metod analizy rozkładów punktowych w badaniu różnych poziomów strukturalnej organizacji przestrzennej zjawisk pradziejowych datuje się od lat siedemdziesiątych i wiąże m. in. z nowatorskimi studiami M. F. Dacey (1973), G. A. Johnsona (1972) i R. Whallona (1973; tenże 1974), a następnie z pracami I. Hoddera i jego współpracowników (np. I. Hodder, M. Hassall 1971; I. Hodder, C. Orton 1976; I. Hodder 1977). W Polsce, jak dotąd, opublikowane zostały jedynie przykłady zastosowań metod najprostszych i nie zawsze najbardziej odpowiednich dla danych archeologicznych, w rodzaju testu Clarka-Evansa (J. Kruk 1980, s. 72-79; W. Brzeziński, M. Dulinicz, Z. Kobylński, B. Lichy, W. Moszczyński 1985).

Wydaje się, że spośród różnych sposobów stosowania matematyki w archeologii, właśnie analiza przestrzenna jest szczególnie obiecującym kierunkiem współpracy obu dyscyplin. Równocześnie jednak, z zastosowaniem statystycznych technik analizy przestrzennej w archeologii wiąże się szereg nieprzewidywanych dotąd, a może nawet nieprzewidywalnych trudności dwóch typów: wynikających z istoty źródeł archeologicznych oraz z niedoskonałości samych technik statystycznych.

Przed wszystkim zastosowanie testów statystycznych wymaga dyspono-

wania informacją o pełnym przestrzennym układzie zdarzeń analizowanych, bądź też o reprezentatywnej dla niego (a więc pobranej losowo) próbie znacznej liczebności. Konieczne jest przy tym przyjęcie założenia, że wszystkie zdarzenia obserwowane w próbie zostały wygenerowane przez ten sam proces społeczno-kulturowy. Tymczasem archeolog rzadko tylko może mieć pewność, że analizowany układ nie jest wynikiem mechanicznej agregacji kilku prób otrzymanych z różnych populacji charakteryzujących się odmiennymi rozkładami przestrzennymi zdarzeń. Stąd niezbędne jest, choć nie zawsze możliwe, maksymalnie ściśle oznaczenie charakteru i chronologii wszystkich zdarzeń i analizowanie tych jedynie, które są względnie synchroniczne. Nawet i w tej sytuacji przyjąć trzeba założenie, że w okresie równym odciinkowi czasu, jaki dzieli najmłodsze z analizowanych zdarzeń od najstarszego, mechanizm przestrzenny generujący zdarzenia nie ulegał zmianom, a nie trzeba dodawać, że założenie to jest zazwyczaj niemożliwe do udowodnienia.

W zasadzie nigdy nie możemy także mieć pewności, że badany układ nie jest zredukowany w stosunku do przeszłej rzeczywistości. Procesy depozycyjne, podepozycyjne i odkrywcze, jak to pokazał np. F. W. Hamond (1980, s. 195-198 ryc. 2-8; por. także J. Kruk 1980, s. 15 ryc. 3), eliminować mogą wiele zdarzeń z zakresu obserwowalności archeologicznej. Problem ten jest bardzo poważny, bowiem, jak wykazano, losowe eliminacje elementów układu punktowego mogą doprowadzić do zaniku jakiegokolwiek wykrywalnej prawidłowości w tym układzie, jeśli tylko zmniejszenie gęstości zdarzeń jest znaczące (por. M. F. Dacey, T. Tung 1962; J. G. Skellam 1952; P. Greig-Smith 1964, s. 217; L. J. King 1969, s. 103-107; w literaturze polskiej problem ten sygnalizował M. Dulicz 1983).

Wreszcie, analizowany rozkład może być nie tylko liczniejszy lub mniej liczny od rzeczywistego, ale także zdeformowany przestrzennie, przede wszystkim w wyniku działania czynników podepozycyjnych. W miarę „bezpiecznie” badać można przestrzenne rozmieszczenie obiektów trwale związanych z podłożem (jak np. grodziska), ich układy bowiem uznać można za względnie niezmiennie wobec transformacji podepozycyjnych. Większe ograniczenia nałożone być muszą na statystyczną analizę rozmieszczenia pojedynczych zabytków ruchomych na stanowisku, ponieważ badania etnoarcheologiczne (np. D. P. Gifford 1980) i eksperymentalne (np. C. L. Redman, P. J. Watson 1970, s. 280; D. Roper 1977, s. 372-375; N. Broadbent 1979, s. 9-10) wskazują na możliwość znacznych przesunięć artefaktów w wyniku późniejszej działalności zwierząt czy uprawy roli.

Również same metody statystyczne analizy przestrzennej nie są dotąd instrumentem doskonałym i są współcześnie nadal dyskutowane i udoskonalane na gruncie matematyki (por. np. krytyczne ich omówienie przez C. Ortona 1982). Nie ma jak dotąd powszechnej zgody matematyków co do mocy informacyjnej poszczególnych metod (por. np. sprzeczne opinie W. G. S. Hines i R. J. O. Hines 1979 oraz C. Ortona 1982 na temat mocy informacyjnej tzw. statystyki Eberhardta). Generalnie rzecz ujmując, techniki

bardziej informatywne wymagają zastosowania elektronicznej techniki obliczeniowej, nie tylko ze względu na dużą liczbę niezbędnych obliczeń, ale także ze względu na konieczność symulacyjnego testowania wyników. Równocześnie metody o mniejszej mocy informacyjnej, które mogą z łatwością być zastosowane przez archeologa, mają znaczenie jako instrument wstępnej przynajmniej analizy danych. Trzeba jednak pamiętać, że wszystkie metody ilościowe redukują potencjalnie zawartą w źródłach informację, a przeto wyniki ich zastosowania muszą być traktowane z największą ostrożnością.

METODY ANALIZY ROZKŁADU JEDNEJ ZMIENNEJ

Mówiąc najogólniej, metody, o które tu chodzi, polegają na statystycznej ocenie zgodności przestrzennego empirycznego rozkładu zdarzeń z różnego typu rozkładami losowymi (wyliczenie typów stochastycznych dyskretnych procesów przestrzennych znajduje się np. w pracach S. Golachowskiego, B. Kostrubca i A. Zagożdżona 1974, s. 103-104 oraz C. Ortona 1982, s. 5). W szczególności większość testów jako hipotezę zerową przyjmuje punktowy jednorodny rozkład Poissona (zob. np. J. W. Prochorow, J. A. Rozanow 1972, s. 49), czyli rozkład losowy, w którym prawdopodobieństwo lokalizacji zdarzenia w dowolnym punkcie jest funkcją jedynie gęstości zdarzeń na danym obszarze, niezależne jest natomiast od lokalizacji innych zdarzeń. Zgodnie z tym rozkładem prawdopodobieństwo pojawienia się x zdarzeń w dowolnym kwadracie, na jakie podzielona jest płaszczyzna, wynosi:

$$P(x) = e^{-\lambda} \cdot \frac{\lambda^x}{x!}, \quad (1)$$

gdzie: $x = 0, 1, 2, \dots$; e — stała Eulera; λ — średnia arytmetyczna liczby zdarzeń we wszystkich kwadratach.

Rozkład punktów w danym obszarze zgodny z rozkładem Poissona może być określony jako „przypadkowy”. Jeśli analizowany jest rozkład stanowisk osadniczych, to wówczas wnioskować można, że naturalna przestrzeń geograficzna, w której ramach odbywał się proces osadniczy, była postrzegana jako jednorodna. Ta jednorodność może być oczywiście pozorna i wynikać z nałożenia się działania różnych reguł lokalizacyjnych, związanych z pozytywnym wartościowaniem różnych, równomiernie rozmieszczonych zasobów środowiska.

Innego typu procesy stochastyczne generują przestrzenne rozkłady punktowe o charakterze skupionym. Procesy te, rozpoczynając się od jednorodnego procesu Poissona, w każdym zdarzeniu generują pewną liczbę potomnych zdarzeń w otoczeniu. Rozkład skupień ma wówczas charakter rozkładu Poissona, natomiast liczba zdarzeń w każdym skupieniu może mieć np. rozkład logarytmiczny (powstaje wówczas rozkład dwumianowy negatywny), bądź również rozkład Poissona (co daje w efekcie tzw. rozkład Neymana typu A; zob. J. Neyman 1939). Procesy prowadzące do powstania rozkładów

skupionych zwane są procesami epidemii prawdziwej, a ich stwierdzenie w odniesieniu do rozmieszczenia stanowisk archeologicznych w regionie sugerować może rozprzestrzenianie się osadnictwa jako proces generujący (por. J. J. Wood 1971, s. 69).

Trzeci typ procesów stochastycznych to procesy Poissona o zmiennej intensywności (złożone procesy Poissona; zob. J. W. Prochorow, J. A. Rozanow 1972, s. 51), tzn. takie, w których parametr λ we wzorze (1) zmienia się w zależności od miejsca na płaszczyźnie. Działanie takiego procesu, zwanego procesem epidemii pozornej, daje w efekcie rozkłady podobne do rozkładów skupionych. Równocześnie jednak interpretacja archeologiczna musi być w tym wypadku zupełnie inna, sugerując subiektywną niejednorodność naturalnej przestrzeni geograficznej. Stąd też próby rozróżnienia tych rozkładów, jakkolwiek skomplikowane matematycznie (por. analizę rozmieszczenia stanowisk neolitycznych na małopolskich wyżynach lessowych, dokonaną przez I. Hoddera i C. Ortona 1976, s. 89-98; także I. Hodder 1977, s. 244-254), są niezwykle istotne.

Czwarty wreszcie typ procesu przestrzennego, który należy tu wziąć pod uwagę, to procesy generujące rozkłady o regularnym rozmieszczeniu punktów na płaszczyźnie. W rozkładach tego typu każde ze zdarzeń otoczone jest okręgiem o promieniu proporcjonalnym do „wagi” zdarzenia, a okręgi sąsiadujących zdarzeń nie mogą się przecinać (por. B. D. Ripley 1977, s. 176). Najbardziej regularny rozkład, oparty na sieci trójkątów równobocznych, zwany jest w geografii osadnictwa siecią Christallera. Do rozkładów tych należy również tzw. regularny rozkład Poissona, opisany przez M. F. Daceyą (1964); w odniesieniu do danych osadniczych odpowiadający procesom lokalizacyjnym związanym ze współzawodnictwem w eksploatacji środowiska naturalnego.

Pierwszym krokiem w statystycznej analizie przestrzennej rozkładu zdarzeń powinna być, zdaniem wielu badaczy (np. R. Whallona 1973, czy M. F. Daceyą 1973; choć por. opinię odmienną: H. J. Hietala, D. E. Stevens 1977, s. 539), ocena jego zgodności z rozkładem przypadkowym (jednorodnym Poissona) w celu wykrycia ewentualnej znaczącej odmienności od tego rozkładu, co pozwoliłoby mówić o obecności wzoru rozmieszczenia zdarzeń.

Matematyka dysponuje czterema klasami metod badania procesów generujących rozkłady punktowe (B. D. Ripley 1977, s. 172; także C. Orton 1982, s. 5): 1 – metodami opartymi na obliczaniu i analizie częstości zdarzeń w kwadratach sieci nałożonej na badany obszar; 2 – metodami opartymi na analizie odległości zdarzenia od najbliższego sąsiedniego zdarzenia; 3 – metodami „drugiego rzędu”, za pomocą których analizuje się wszystkie odległości między zdarzeniami na badanym obszarze; oraz 4 – metodami oceniającymi prawdopodobieństwo znalezienia się punktu (reprezentującego zdarzenie) w niewielkim obszarze testowym umieszczanym losowo na badanym terytorium. Spośród nich najbardziej dostępne do samodzielnego zastosowania przez archeologa, choć równocześnie stosunkowo „najsłab-

sze" informacyjnie – są dwie pierwsze klasy – metody „kwadratowe” i metody „najbliższego sąsiada”, i one to będą omawiane w dalszym ciągu niniejszego artykułu.

METODY „KWADRATOWE”

Pośród metod opartych na konstrukcji kwadratowych sieci na badanej płaszczyźnie i zliczaniu częstości zdarzeń w poszczególnych jednostkach tych sieci, najprostsze jest obliczanie tzw. wskaźnika dyspersji I , zgodnie ze wzorem:

$$I = n \cdot \frac{V}{\bar{x}} \quad (2)$$

gdzie: n – liczba kwadratów; \bar{x} – średnia arytmetyczna liczby zdarzeń w kwadratach sieci; V – wariancja liczby tych zdarzeń. Test zgodności rozkładu empirycznego z rozkładem Poissona polega na porównaniu wartości I z wartością krytyczną testu chi-kwadrat przy $n-1$ stopniach swobody na odpowiednim poziomie istotności (I. Hodder, C. Orton 1976, s. 34).

Niewątpliwą zaletą tej metody jest prostota procedury statystycznej, jak również fakt, że nie wymaga ona dysponowania informacjami o dokładnej lokalizacji zdarzenia. Stąd metoda ta może zostać zastosowana w sytuacji, gdy dane zbierane były w terenie w obrębie równych pod względem powierzchni jednostek (wykopów). Równocześnie jednak metoda oceny wskaźnika dyspersji ma wielkie wady (zob. np. I. Hodder, C. Orton 1976, s. 38; H. J. Hietala, D. E. Stevens 1977, s. 539-540; C. Orton 1982, s. 15). Jedną z nich jest niesprecyzowanie hipotezy alternatywnej. Nie wiadomo, czy brak zgodności z rozkładem Poissona oznacza rozkład skupiony czy regularny, choć zazwyczaj za hipotezę alternatywną uważane jest skupianie się zdarzeń w skali pojedynczego kwadratu sieci. Znacznie poważniejszą wadą metody jest zależność uzyskanego wyniku od arbitralnie dobieranej skali podziału badanego obszaru na kwadraty. Ten sam wzór rozmieszczenia może wykazywać odmienne wartości wskaźnika dyspersji, w zależności od wielkości kwadratów. Wreszcie, obliczając wskaźnik dyspersji całkowicie ignoruje się informację o pozycji poszczególnych kwadratów sieci w stosunku do siebie, przeto traconą jest możliwość zauważenia skupień w skali większej niż pojedynczy kwadrat.

Dla przezwyciężenia tych trudności zaproponowano metodę analizy wariancji (ANOVA) rozmieszczenia punktów w zależności od skali podziału obszaru, w celu znalezienia skali, dla której rozmieszczenie to znacząco różni się od rozkładu przypadkowego (P. Greig-Smith 1964, s. 86-93; R. Mead 1974). W archeologii na zalety poznawcze metody ANOVA wskazywał przede wszystkim R. Whallon (1973). Warto zauważyć, że analogiczną procedurę zastosowali w Polsce już w 1957 r. S. Kucharczyk i H. Szczepanowski (1957) współpracujący z W. Hołubowiczem w czasie badań na Ślęży. Symulu-

jąc różnej wielkości siatki wykopów na przebadanej części stanowiska, porównywali uzyskane rozkłady fragmentów ceramiki kultury łużyckiej z rozkładem Poissona i obliczali wartości wskaźników Steinhausa i Perkala, uzyskując w efekcie znaczące wskaźniki skupienia w skali kwadratów o boku 2 m.

Dysponując takimi samymi danymi, jak wymagane w metodzie analizy wariancji, a więc częstościami zdarzeń w kwadratach sieci, można stestować zgodność rozkładu przestrzennego z rozkładem Poissona lub dowolnym innym teoretycznym rozkładem losowym także i w inny, prostszy sposób. Zgodność przebiegu dystrybuanty rozkładu empirycznego może być bezpośrednio porównana z dystrybuantą rozkładu teoretycznego za pomocą testu chi-kwadrat lub testu Kołmogorowa (A. Góralski 1974, s. 228). Prawdopodobieństwa rozkładu Poissona znaleźć można w każdym podręczniku statystyki (np. C. Platt 1977, tablica V), natomiast prawdopodobieństwa w bardziej skomplikowanych rozkładach obliczyć trzeba samodzielnie (odpowiednie wzory podaje np. J. J. Wood 1971, s. 68-70) lub odszukać w rozproszonej literaturze fachowej. Niestety, opisana procedura nie eliminuje najpoważniejszego z ograniczeń metod „kwadratowych”, a mianowicie silnego uzależnienia od wielkości jednostek sieci.

Te same wady charakteryzują wszystkie inne opracowane dotąd procedury oceny rozkładu przestrzennego oparte na obliczaniu częstości w kwadratach: obliczanie entropii formułą Miedwiedkowa (M. Chilczuk 1975, s. 33-34; S. Golachowski, B. Kostrubiec, A. Zagożdżon 1974, s. 146-151) bądź też przy zastosowaniu modelu grawitacyjnego (S. Golachowski, B. Kostrubiec, A. Zagożdżon 1974, s. 72-78); obliczanie wskaźnika koncentracji C (B. Kostrubiec 1969, s. 229-251), czy wreszcie obliczanie wskaźnika habilitacji Lexisa-Steinhausa (S. Golachowski, B. Kostrubiec, A. Zagożdżon 1974, s. 107-112).

METODY „NAJBLIŻSZEGO SĄSIADA”

Generalnie „mocniejsze”, bo uniezależnione od arbitralności podziału badanego obszaru na kwadraty, są metody, za pomocą których bada się odległości pomiędzy punktami reprezentującymi zdarzenia. Znane są trzy rodzaje tych metod: 1 — analiza odległości od zdarzeń do najbliższych sąsiednich zdarzeń; 2 — analiza odległości od punktów rozmieszczonych losowo na badanym obszarze do najbliższych sąsiednich zdarzeń; i wreszcie 3 — analiza obejmująca oba wymienione rodzaje odległości (wykaz znanych metod odległościowych podaje C. Orton 1982, s. 6-7 tabela I).

Pośród metod pierwszej grupy (za pomocą których tworzy się dendryt najkrótszych odległości między zdarzeniami) najbardziej znana jest, i najczęściej stosowana na gruncie archeologii, metoda obliczania tzw. statystyki Clarka-Evansa (P. J. Clark, F. C. Evans 1954). Autorzy ci pokazali, że dla losowego jednorodnego rozkładu Poissona średnia najbliższa odległość do

sąsiedniego zdarzenia jest zależna wyłącznie od gęstości zdarzeń na danym obszarze i wynosi:

$$R_E = \frac{1}{2 \sqrt{\frac{N}{A}}}, \quad (3)$$

gdzie: N — całkowita liczba zdarzeń; A — pole powierzchni obszaru badanego. Przypadkowość rozkładu empirycznego jest zatem wskazana przez stosunek średniej najbliższej odległości obserwowanej (R_0) do średniej najbliższej odległości teoretycznej:

$$R = \frac{R_0}{R_E} \quad (4)$$

Dla rozkładu losowego stosunek ten wynosić powinien 1, podczas gdy dla rozkładu skupionego przyjmuje wartości mniejsze niż 1, osiągając wartość 0 w przypadku skrajnym, gdy wszystkie zdarzenia zlokalizowane są w jednym punkcie, a dla rozkładu regularnego R jest większe od 1, osiągając 2,1491 w ekstremalnym przypadku rozkładu punktów w regularnej christallerowskiej siatce trójkątów równobocznych.

Możliwe jest również sprawdzenie, jakie jest prawdopodobieństwo odrzucenia hipotezy zerowej przy danej wartości R . Wartość statystyki C :

$$\frac{(R_0 - R_E) \sqrt{N \cdot \frac{N}{A}}}{0,26136}, \quad (5)$$

może być porównana z rozkładem normalnym na różnych poziomach istotności, jeśli N jest większe niż 100, lub (co trudniejsze do wykonania) z rozkładem Pearsona typu III, jeśli N jest mniejsze niż 100 (R. Whallon 1974).

Alternatywna statystyka zaproponowana została przez J. G. Skellama (1952) i P. G. Moore'a (1954):

$$2\pi \cdot \frac{N}{A} \cdot \sum r^2, \quad (6)$$

gdzie: r — odległość do „najbliższego sąsiada”. Zaletą jej jest łatwość testowania poprzez porównanie z rozkładem chi-kwadrat przy $2N$ stopniach swobody (test dwustronny). Dla oceny wartości uzyskanej przy dużych N można użyć tzw. wzoru Romanowskiego (M. Krzysztofiak, A. Luszczewicz 1979, s. 166):

$$\frac{|\text{chi-kwadrat} - s|}{\sqrt{2s}}, \quad (7)$$

gdzie: s — liczba stopni swobody (w tym przypadku równa $2N$). Do

odrzućcenia hipotezy zerowej na poziomie istotności 0,05, wartość uzyskana ze wzoru (6) powinna być większa niż 3. Innym sposobem oceny istotności wartości chi-kwadrat jest porównanie wartości wyrażenia:

$$\sqrt{2\text{chi-kwadrat}} - \sqrt{2s-1}, \quad (8)$$

z tabelą wielkości obszaru pod krzywą normalną. Wartość uzyskana ze wzoru (6) do odrzućcenia hipotezy zerowej na poziomie istotności 0,05 powinna być większa niż 1,96 lub mniejsza niż $-1,96$ (R. Whallon 1974, s. 20).

Wielu autorów zwraca uwagę na liczne niedoskonałości testu Clarka-Evansa w jego oryginalnym sformułowaniu (wady te są również własnością statystyki Skellama-Moore'a). Przede wszystkim wielkości użyte do opisu rozkładu populacji zdarzeń w badanym obszarze: gęstość i średnia odległość do najbliższego sąsiada, nie charakteryzują obszaru jednoznacznie (poważnym utrudnieniem jest również sama konieczność znajomości gęstości zdarzeń). Wszelkie rozciągnięcia grafu najbliższych odległości bez rozrywania nie zmieniają średniej najbliższej odległości, co oznacza, że istnieje nieskończenie wiele układów przestrzennych o jednakowej wartości testu, mimo wybitnie różniących się konfiguracji (B. Kostrubiec 1969, s. 236-237; por. też M. Dulicz 1983, s. 306 ryc. 3). Na wyniki testu ma również silny wpływ wielkość badanego obszaru i jego umieszczenie w stosunku do całkowitego obszaru zajmowanego przez populację zdarzeń (por. D. Pinder, I. Shimada, D. Gregory 1979, s. 434-435, zwłaszcza ryc. 2). Najbardziej zgodne z wymogami testu byłoby umieszczenie obszaru badanego w obrębie centrum układu zdarzeń. Tymczasem, w przypadku archeologii, obszary badawcze są z reguły określane szeroko, w celu uzyskania informacji o pełnym układzie analizowanych zdarzeń. Powoduje to, że odległości do „najbliższego sąsiada” nie stanowią zbioru obserwacji niezależnych (P. J. Diggle 1979, s. 87). Tym samym wartości testu dla kompletnych układów nie mogą być uważane za dokładne, jakkolwiek symulacyjnie wykazano, że błąd obliczeń jest względnie niewielki (P. J. Diggle 1976, s. 247).

Zwraca się również uwagę na zakłócenia wyników testu wynikające z tzw. efektu brzegowego (por. np. I. Hodder, C. Orton 1976, s. 41-43). Wzór na obliczenie średniej odległości do najbliższego sąsiedniego zdarzenia (wzór 3) oczekiwanej przy losowym rozkładzie Poissona został bowiem wyprowadzony dla przestrzeni nieograniczonej. Jeżeli na taką przestrzeń nałożymy obszar skończony, co musimy zrobić, aby określić pole powierzchni obszaru i liczebność populacji, to granica obszaru, jeśli nie jest to granica naturalna występowania danego zjawiska, prawdopodobnie przerwie połączenia pomiędzy niektórymi punktami (zdarzeniami) wewnątrz obszaru, a ich „najbliższymi sąsiadami” poza granicą. Jeśli zatem obliczamy odległości pomiędzy najbliższymi sobie zdarzeniami wewnątrz obszaru, to niektóre z tych miar będą większe niż by to miało miejsce w przypadku uwzględnienia „najbliż-

szych sąsiadów” poza obszarem. Efekt ten będzie największy, jeśli populacja zdarzeń jest mało liczna, ponieważ wówczas frakcja punktów oderwanych od „najbliższych sąsiadów” może być wysoka. Prostym wyjściem z tej sytuacji jest pominięcie w analizie wszystkich punktów leżących bliżej granicy niż swego „najbliższego sąsiada” wewnątrz obszaru badanego, taki zabieg powoduje jednak częstokroć znaczne zubożenie liczby analizowanych zdarzeń.

Wreszcie test Clarka-Evansa pozostaje całkowicie „nieczuły” wobec pewnych alternatyw rozkładu Poissona, a w szczególności, co ważne przy badaniu przestrzennych struktur osadnictwa — wobec układów liniowych, które postrzegane są jako losowe (por. dalej). Dla wykrycia takich układów zaproponowano pewne, bardziej skomplikowane statystyki (por. L. J. King 1969, s. 98; D. A. Pinder, M. E. Witherick 1975; S. Broadbent 1980).

Test ten dzieli również wiele wad z innymi metodami analizy przestrzennej układów punktowych, a w szczególności brak uwzględnienia zróżnicowania zdarzeń. Test ten oparty jest bowiem na założeniu, że każdy z punktów układu reprezentuje zdarzenie o tej samej treści społeczno-kulturowej, a więc np. osadę o tej samej wielkości. Wprowadzenie rang w celu zróżnicowania zbioru punktów jest możliwe (por. I. Hodder 1977, s. 238-239), ale wówczas pojawiają się dodatkowe trudności obliczeniowe. W miejsce prostej odległości Euklidesowej wprowadzić bowiem należy wówczas do obliczeń odległość ważoną (por. T. Marek, C. Noworol 1983, s. 22). Natomiast brak uwzględnienia rang poszczególnych zdarzeń powodować może, że złożone mechanizmy deterministyczne mogą być postrzegane jako losowe (por. wyniki analizy rozmieszczenia grodów w południowej i zachodniej Anglii — I. Hodder, C. Orton 1976, s. 44-46).

Dla przewyciężenia niektórych spośród ograniczeń testu Clarka-Evansa, a w szczególności efektu brzegowego, zaproponowano w ostatnich latach dwa sposoby korekcy wyników.

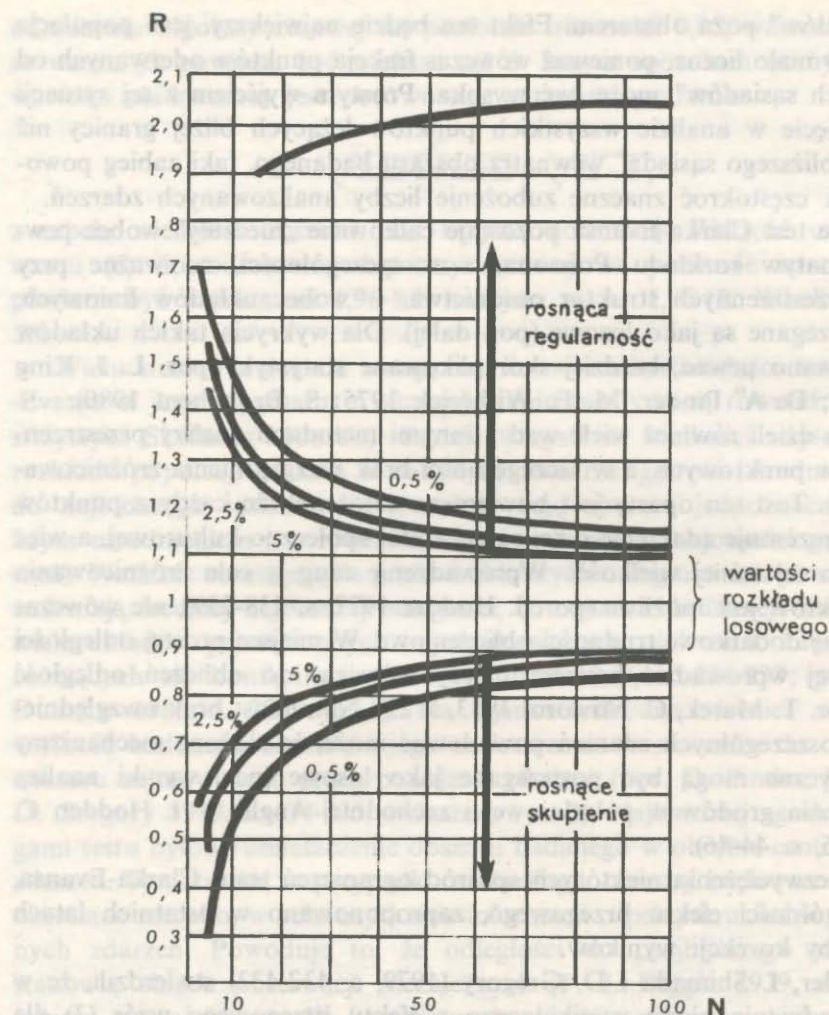
D. Pinder, I. Shimada i D. Gregory (1979, s. 432-433) stwierdzili, że w celu uwzględnienia błędu wynikającego z efektu brzegowego wzór (3) dla obliczania RE powinien mieć postać:

$$R_E = (0,497 + 0,127 \sqrt{\frac{4}{N}}) \cdot \sqrt{\frac{A}{N}}. \quad (9)$$

Statystyka R nie przyjmuje wówczas stałych wartości w przedziale $0 \dots 2$, 15, niezależnie od liczebności zbioru zdarzeń, ale tworzy pewne strefy wartości podlegające interpretacji przy danych poziomach istotności i danej liczbie analizowanych punktów (ryc. 1).

Inną modyfikację obliczeń, minimalizującą efekt brzegowy, zaproponował K. P. Donnelly (1978), który postuluje obliczanie wskaźnika sumarycznej odległości do „najbliższych sąsiadów” (TNND) według wzoru:

$$TNND = \frac{T - E(T)}{\sqrt{(var T)}}, \quad (10)$$



Ryc. 1. Diagram służący do interpretacji skorygowanej wartości statystyki Clarka-Evansa

Wg D. Pindera, I. Shimady i D. Gregory'ego 1979, s. 439 ryc. 5

Fig. 1. Graph for interpretation a value of corrected Clark and Evans statistics

After D. Pinder, I. Shimada, D. Gregory 1979, p. 439 fig. 5

gdzie:

$$E(T) = 0,5 \sqrt{N \cdot A} + \left(0,051 + \frac{0,041}{\sqrt{N}} \right) \cdot L, \quad (11)$$

$$\text{var } T = 0,07 \cdot A + \frac{0,037 \cdot L}{\sqrt{\frac{A}{N}}} \quad (12)$$

gdzie: T – suma odległości do „najbliższych sąsiadów”; L – obwód badanego obszaru.

TNND testuje się przez porównanie z rozkładem normalnym (test dwustronny), a więc dla odrzucenia hipotezy zerowej na poziomie istotności równym 0,05, wartość uzyskana ze wzoru (10) powinna być większa niż 1,96 lub mniejsza niż -1,96.

Nawet po tych korekcjach test Clarka-Evansa przeznaczony jest dla obszarów o prostych granicach, najlepiej zbliżonych do kwadratu, ponieważ wraz ze wzrostem nieregularności obszaru zwiększa się długość granic, co potęguje efekt brzegowy.

Niektóre spośród wskazanych wyżej ograniczeń metod szacujących najbliższe odległości między zdarzeniami, przewyżniają metody oparte na ocenie odległości pomiędzy losowo rozmieszczonymi w badanym obszarze punktami a najbliższymi im zdarzeniami (są to tzw. metody punkt-zdarzenie).

Spośród tych pracochłonnych metod, z których wiele wymaga testowania symulacyjnego (por. P. Holgate 1965; P. J. Diggle, J. Besag, J. T. Gleaves 1976; C. Orton 1982, s. 18-19), najprostszy jest tzw. test Pielou (E. C. Pielou 1959). Istotą testu, podobnie jak wszystkich z tej grupy, jest losowe rozmieszczenie w badanym obszarze pewnej liczby punktów losowych. Obliczane są następnie najbliższe odległości pomiędzy każdym z tych punktów, a najbliższym mu zdarzeniem. Statystyka:

$$2\pi \frac{N}{A} \cdot \Sigma d^2, \quad (13)$$

może być porównana z rozkładem chi-kwadrat (test dwustronny) przy $2F$ stopniach swobody (d – odległość między punktem losowym a najbliższym sąsiednim zdarzeniem; F – liczba punktów losowych). Wadą testu jest brak sprecyzowania hipotezy alternatywnej, jak również konieczność estymacji gęstości zdarzeń.

Bardziej jeszcze „czułym” testem losowości rozkładu przestrzennego zdarzeń wydaje się być współczynnik skupienia Hopkinsa (B. Hopkins 1954), obliczany według wzoru:

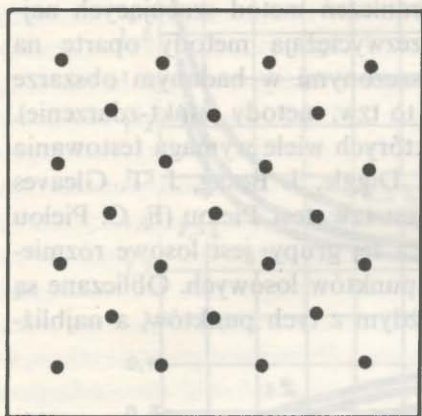
$$A = \frac{F \cdot \Sigma d^2}{N \cdot \Sigma r^2} \quad (14)$$

Może być on testowany rozkładem statystyki F przy $2F$ i $2N$ stopniach swobody. Test Hopkinsa wymaga wylosowania z populacji zdarzeń pewnej liczby elementów i obliczenia odległości od każdego z nich do „najbliższego sąsiada”. P. Holgate wskazał (1965, s. 347), że jeśli, tak jak w oryginalnym sformułowaniu Hopkinsa, $F = K$, tzn. jeśli próbka wylosowanych zdarzeń jest równoliczna z próbką punktów losowych umieszczonych w badanym obszarze, to wówczas statystyka Hopkinsa obliczana być może według wzoru:

$$A' = \frac{\Sigma d^2}{\Sigma d^2 + \Sigma r^2} \quad (15)$$

Wartość A' ma rozkład beta przy $2F$ i $2N$ stopni swobody. Oznacza to, że wartości istotnie większe niż 0,5 wskazują na skupienie rozkładu zdarzeń, natomiast wartości istotnie mniejsze niż 0,5 – na regularność (P. J. Diggle, J. Besag, J. T. Gleaves 1976, s. 661).

Metoda Hopkinsa wymaga jednak dużej ilości obliczeń i może być stosowana na szerszą skalę raczej przy użyciu elektronicznych maszyn cyfrowych.

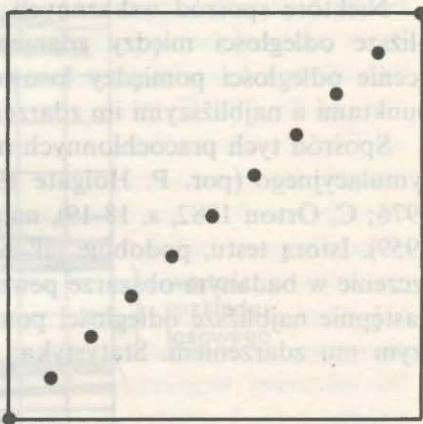


Ryc. 2. Punktowy układ regularny

Oprac. i rys. Z. Kobyliński

Fig. 2. Regular point pattern

Designed by Z. Kobyliński



Ryc. 3. Punktowy układ liniowy

Oprac. i rys. Z. Kobyliński

Fig. 3. Linear point pattern

Designed by Z. Kobyliński

Powyżej zaprezentowano sześć testów „odległościowych” pozwalających wykryć nieprzypadkowy wzór rozmieszczenia zdarzeń. Nie są to oczywiście wszystkie znane testy (por. C. Orton 1982), a jedynie te, które nie wymagają testowania symulacyjnego i są najłatwiejsze pod względem obliczeniowym. „Moc” przedstawionych testów nie jest przy tym jednakowa. Porównanie tej „mocy” przedstawić najlepiej na przykładach.

Rozpatrzmy najpierw układ o niemal doskonałej regularności rozmieszczenia zdarzeń (ryc. 2). Wyniki przeprowadzonych analiz przedstawiono w tabeli 1. Jak widać, jedynie test Pielou okazał się niezdolny do wykrycia wzoru rozmieszczenia zdarzeń.

Rozpatrzmy następnie układ liniowy (ryc. 3). W tym przypadku wartości statystyk Clarka-Evansa (również w modyfikacji Pindera i in.) oraz Skellama-Moore'a nie są w stanie wykryć nielosowości rozmieszczenia zdarzeń (tabela 2). Natomiast statystyki Pielou i Hopkinsa wskazują znaczące skupienie (!) punktów (niektóre wyniki symulacji wskazują losowość).

Weźmy następnie pod uwagę dwa rozkłady bardziej złożone: rozkład o tendencji do skupiania, należący do rodziny rozkładów „epidemii prawdzi-

Tabela 1. Wyniki analizy punktowego układu regularnego (ryc. 2)

Table 1. „Event-event” and „point-event” nearest neighbour analysis of regular point pattern (fig. 2)

Powierzchnia obszaru A	Liczba zdarzeń N	R	R skorelowane	TNND	Statystyka Skellama-Moore'a	Statystyka Pielou		Statystyka Hopkinsa	
						F = 10	F = N	F = 10	F = N
100	25	1,7678 *	1,6135 *	6,0052 *	122,7185 *	9,4845	28,2276	0,1932 *	0,2302 *
						10,1221	24,3234	0,2062 *	0,1982 *
						8,5400	35,2644	0,1740 *	0,2874 *
						10,3424	31,3243	0,2107 *	0,2553 *
						12,7486	30,0778	0,2597 *	0,2451 *

* Wartości pozwalające na odrzucenie hipotezy zerowej

* Values permitting the rejection of null hypothesis

Tabela 2. Wyniki analizy punktowego układu liniowego (ryc. 3)

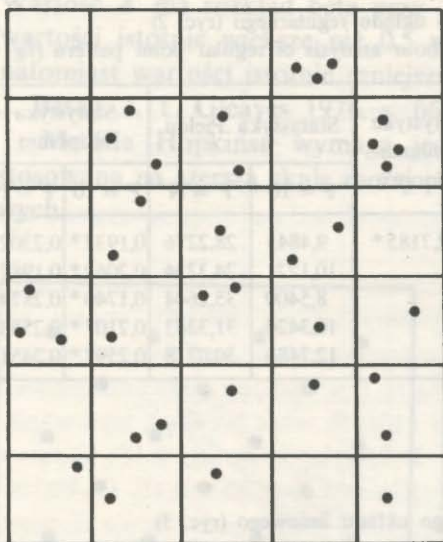
Table 2. Nearest neighbour analysis of linear point pattern (fig. 3).

Powierzchnia obszaru A	Liczba zdarzeń N	R	R skorelowane	TNND	Statystyka Skellama-Moore'a	Statystyka Pielou		Statystyka Hopkinsa	
						F = 6	F = N	F = 6	F = N
100	11	0,9381	0,8177	-0,4001 *	15,2053	33,8287 *	78,5133 *	4,0788 *	5,1635 *
						33,5796 *	94,4125 *	4,0488 *	6,2092 *
						42,4024 *	81,3223 *	5,1125 *	5,3483 *
						49,6177 *	81,6215 *	5,9825 *	5,3680 *
						16,6416	24,9680	2,0065	1,6421

wej”, a mianowicie rozkład Neymana typu A (ryc. 4) i rozkład o tendencji do regularności, a mianowicie regularny rozkład Poissona, opisany przez M. F. Dacey (ryc. 5). W pierwszym z tych przypadków nie można było wykryć nielosowego charakteru rozkładu (tabela 3). Natomiast w drugim przypadku (tabela 4) wszystkie statystyki pozwoliły wykryć regularność rozkładu, jedynie statystyka Pielou w niektórych symulacjach wskazywała na losowość.

Podsumowując wyniki powyższego prostego porównania, trzeba zwrócić uwagę na fakt, że chociaż zdaniem większości statystyków metody „punkt losowy – zdarzenie”, a w szczególności test Hopkinsa, są lepszym narzędziem analitycznym niż testy „zdarzenie – zdarzenie” (por. P. Holgate 1965, s. 347–353; P. J. Diggle, J. Besag, J. T. Gleaves 1976, s. 661; P. J. Diggle 1977, s. 390; C. Orton 1982), to jednak uzyskane przeze mnie wyniki wskazują na znaczną „moc” statystyki TNND.

Ogólnie stwierdzić trzeba, że z grupy sześciu omówionych testów najlepszymi narzędziami analizy przestrzennej są: statystyka TNND i statystyka Hopkinsa.

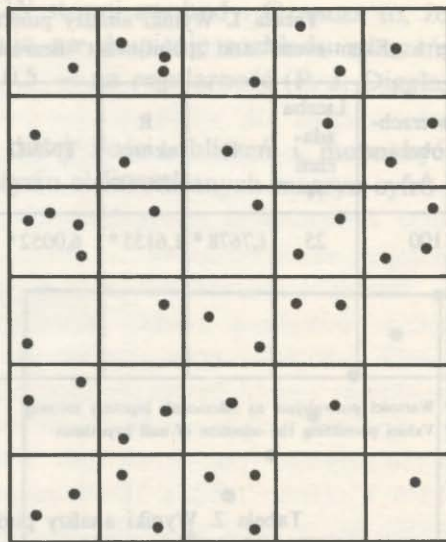


Ryc. 4. Rozmieszczenie punktów stanowiące realizację procesu Neymana typu A

Wg J. J. Wooda 1971, s. 80 ryc. 3

Fig. 4. Point distribution being a realization of Neyman type A process

After J. J. Wood 1971, p. 80 fig. 3



Ryc. 5. Rozmieszczenie punktów stanowiące realizację procesu Daceya (regularnego procesu Poissona)

Wg J. J. Wooda 1971, s. 82 ryc. 4

Fig. 5. Point distribution being a realization of Dacey's regular Poisson process

After J. J. Wood 1971, p. 82 fig. 4

Jak wskazywano wyżej, założeniem wszystkich testów „odległościowych” jest to, że analizowany obszar stanowi próbkę pobraną losowo z całej przestrzeni zajmowanej przez interesujący nas układ obiektów. Równocześnie jednak jest to założenie, które pozostaje w sprzeczności z naturalnym dążeniem każdego badacza do uzyskania informacji o całym układzie. Rozwiązaniem tego dylematu może być wykonanie analizy dla układu kompletnego, a następnie dla szeregu podobszarów, najlepiej zlokalizowanych losowo w obrębie tego układu. Procedura ta jest szczególnie pożądana w stosunku do testów zależnych od wielkości obszaru badanego. Dzięki takiej procedurze

Tabela 3. Wyniki analizy punktowego układu stanowiącego realizację procesu Neymana typu A (ryc. 4)

Table 3. Nearest neighbour analysis of Neyman type A distribution (fig. 4)

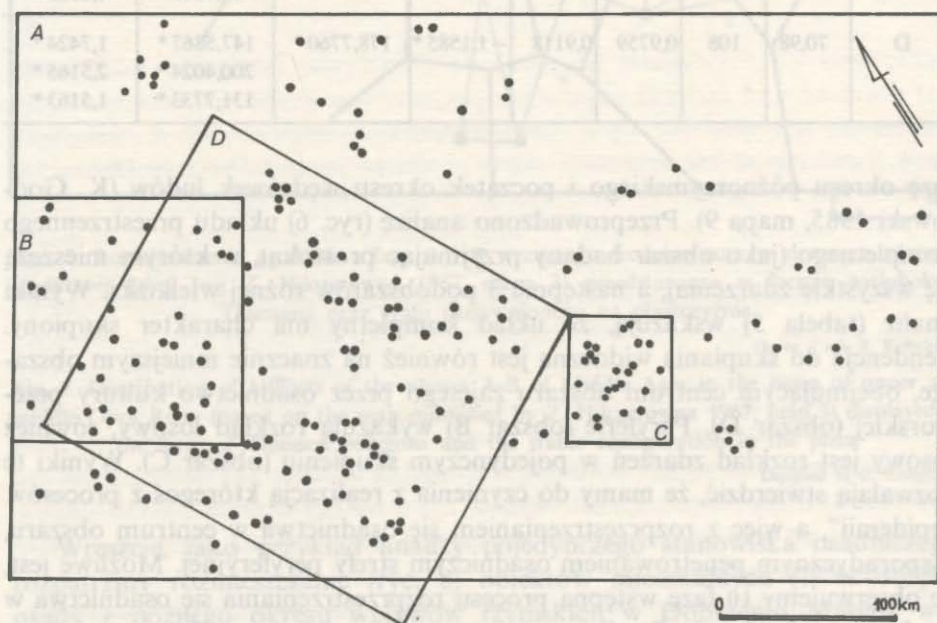
Powierzchnia obszaru A	Liczba zdarzeń N	R	R skorelowane	TNND	Statystyka Skellama-Moore'a	Statystyka Pielou		Statystyka Hopkinsa	
						F = 20	F = N	F = 20	F = N
141,5	35	1,1108	1,0286	0,2769 *	76,6457	48,3519	96,8477	1,3614	1,2636
						43,7614	56,5100	1,3144	0,7373
						37,8739	70,6733	1,0458	0,9221

Tabela 4. Wyniki analizy punktowego układu stanowiącego realizację regularnego procesu Dacey'a (ryc. 5)
Table 4. Nearest neighbour analysis of Dacey regular Poisson distribution (fig. 5)

Powierzchnia obszaru A	Liczba zdarzeń N	R	R skorelowane	TNND	Statystyka Skellama-Moore'a	Statystyka Pielou		Statystyka Hopkinsa	
						F = 20	F = N	F = 20	F = N
141,5	45	1,3833 *	1,2931 *	12,9222 *	156,5400 *	58,0886 *	91,2211	0,3711 *	0,5827 *
						81,4599 *	77,4089	0,5204 *	0,4945 *
						49,3805	70,1642	0,3154 *	0,4482 *

można uzyskać bardziej szczegółowe informacje o charakterze rozkładu przestrzennego.

Zastosowanie takiej procedury w konkretnej sytuacji archeologicznej najlepiej przedstawić na przykładzie. Rozpatrzmy najpierw rozmieszczenie stanowisk archeologicznych kultury przeworskiej datowanych na końcową



Ryc. 6. Rozmieszczenie stanowisk archeologicznych kultury przeworskiej z młodszej fazy okresu wpływów rzymskich i starszej fazy okresu wędrówek ludów (na podstawie mapy opublikowanej przez K. Godłowskiego 1985, mapa 9) z zaznaczonymi obszarami testowymi rozpatrywanymi w niniejszym artykule

Oprac. i rys. Z. Kobyliński

Fig. 6. Distribution of the Przeworsk culture archaeological sites of the late phase of Roman influences period and early phase of migration period (based on the map published by K. Godłowski 1985, map 9), with test areas marked

Designed by Z. Kobyliński

Tabela 5. Wyniki analizy układu punktów reprezentujących stanowiska kultury przeworskiej z fazy D okresu wpływów rzymskich (ryc. 6)

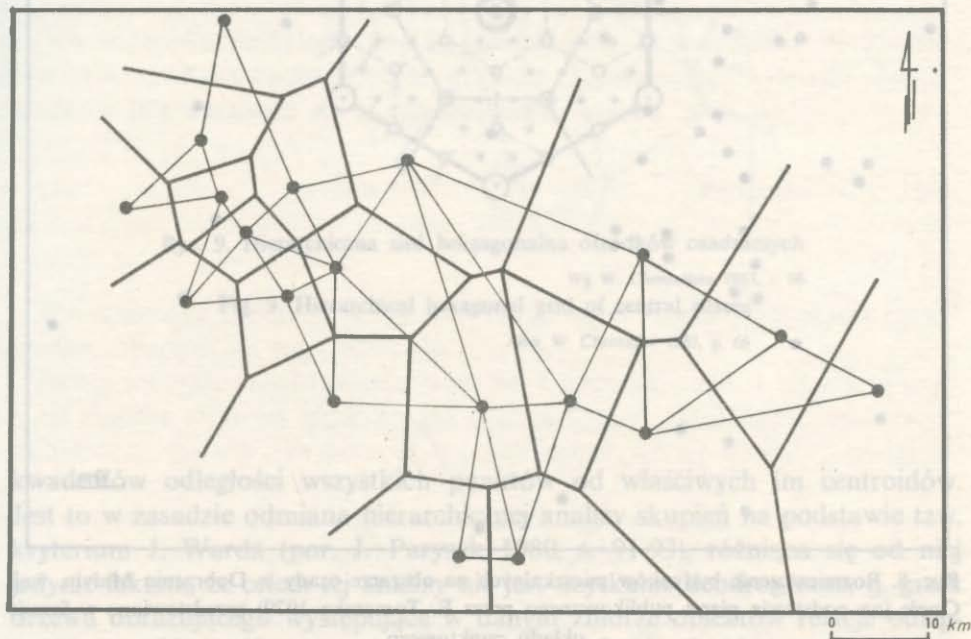
Table 5. Nearest neighbour analysis of spatial arrangement of the Przeworsk culture archaeological sites of the late phase of Roman influences period (fig. 6)

Obszar	Powierzchnia A	Liczba zdarzeń N	R	R skorygowane	TNND	Statystyka Skellama- Moore'a	Statystyka Pielou F = N	Statystyka Hopkinsa F = N
A	256,25	197	0,7391 *	0,7174 *	- 7,0556 *	218,5752 *	349,9321 * 633,7398 * 427,4053 *	5,2382 * 9,1683 * 9,1031 *
B	27,54	35	1,0330	0,9566	- 0,3462 *	72,6650 *	38,9777 35,5545 26,3121	0,6751 1,2368 0,7827
C	6,72	24	1,1233	1,0233	- 5,7653 *	58,9967 *	16,7121 17,2813 19,7193	0,5642 0,5153 0,6885
D	70,98	108	0,9759	0,9118	- 1,1585 *	178,7760 *	147,5867 * 200,4024 * 131,7733 *	1,7424 * 2,5165 * 1,5163 *

fazę okresu późnorzymskiego i początek okresu wędrówek ludów (K. Godłowski 1985, mapa 9). Przeprowadzono analizę (ryc. 6) układu przestrzennego kompletnego (jako obszar badany przyjmując prostokąt, w którym mieszczą się wszystkie zdarzenia), a następnie 3 podobszarów różnej wielkości. Wyniki analiz (tabela 5) wskazują, że układ kompletny ma charakter skupiony. Tendencja do skupiania widoczna jest również na znacznie mniejszym obszarze, obejmującym centrum obszaru zajętego przez osadnictwo kultury przeworskiej (obszar D). Peryferie (obszar B) wykazują rozkład losowy, również losowy jest rozkład zdarzeń w pojedynczym skupieniu (obszar C). Wyniki te pozwalają stwierdzić, że mamy do czynienia z realizacją któregoś z procesów „epidemii”, a więc z rozprzestrzenianiem się osadnictwa w centrum obszaru, a sporadycznym penetrowaniem osadniczym strefy peryferyjnej. Możliwe jest, że obserwujemy tu fazę wstępną procesu rozprzestrzeniania się osadnictwa w strefach peryferyjnych. Rzadko rozmieszczone punkty osadnicze w tych strefach mogłyby stać się z kolei punktami wyjściowymi dla rozprzestrzeniania się osad potomnych.

Weźmy obecnie pod uwagę układ o odmiennej charakterystyce. Rozmieszczenie grodów z faz A–B wczesnego średniowiecza w dorzeczu górnej i środkowej Obry (Z. Hilczerówna 1967, mapa 2) charakteryzuje całkowita losowość w świetle wartości testu Clarka-Evansa ($R = 1,16$; R skorygowane = 1,02), natomiast fragment tego układu, obejmujący 10 grodów w okolicach Bonikowa, zawarty w kole o promieniu 15,5 km, wykazuje wyraźną regular-

ność ($R = 1,52$; R skorygowane $= 1,25$). Na tym obszarze zatem sieć grodów jest realizacją deterministycznego procesu regularnego. Sugeruje to bądź planową akcję lokalizowania grodów, bądź, co może bardziej prawdopodobne, konkurencyjność grup ludzkich użytkujących grody (ryc. 7).



Ryc. 7. Rozmieszczenie grodzisk z fazy A i B wczesnego średniowiecza w dorzeczu górnej i środkowej Obry (wg Z. Hilczerówny 1967, mapa 3) przedstawione w formie wieloboków Thiessena oraz grafu usztywnionego na płaszczyźnie

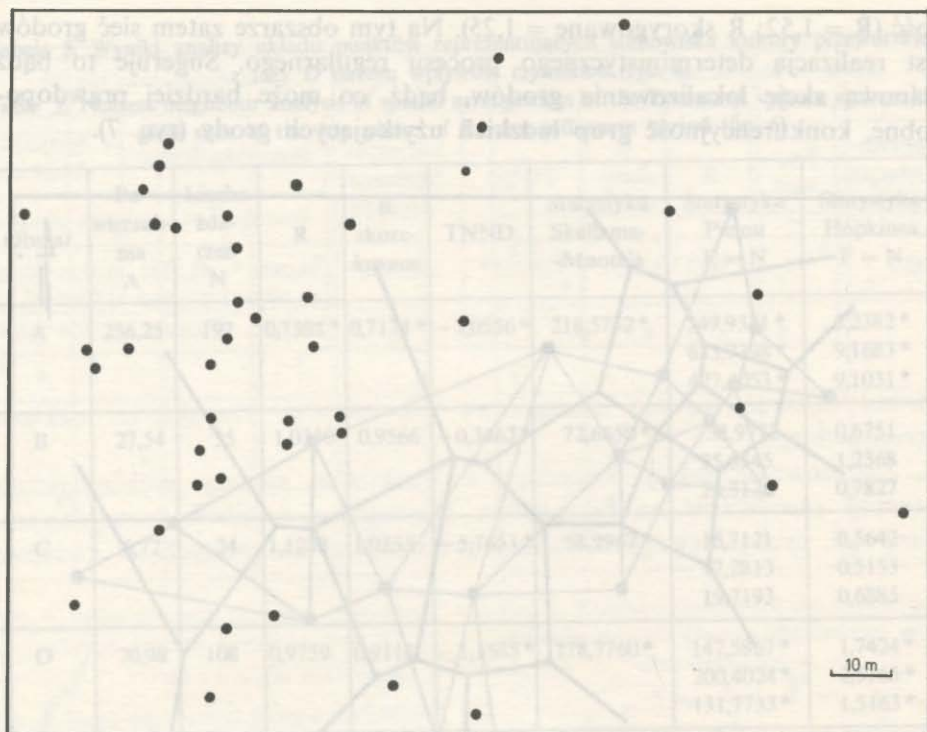
Oprac. i rys. Z. Kobyliński

Fig. 7. Distribution of hillforts of the phases A-B of Middle Ages in the basin of upper and middle Obra River (based on the map published by Z. Hilczerówna 1967, map 3) displayed in the form of Thiessen polygons and the stable configuration in the plane

Designed by Z. Kobyliński

Wreszcie, jako przykład analizy pojedynczego stanowiska osadniczego, rozpatrzmy rozmieszczenie (ryc. 8) obiektów mieszkalnych (?) w obrębie osady z późnego okresu wpływów rzymskich w Dobrzeniu Małym, woj. Opole (E. Tomczak 1979). Tutaj testy z grupy "zdarzenie – zdarzenie" okazują się być bezsilne (R skorygowane $= 1,0991$; $TNND = 1,1828$), natomiast wartości statystyki Pielou (194,3535) i Hopkinsa (3,2249) pozwalają stwierdzić skupienie układu (w badanym obszarze rozmieszczono 30 punktów losowych).

Z powyższych kilku przykładów sytuacji archeologicznych, których niepogłębioną analizę można było w tym miejscu przedstawić, wynika jasno, że dla uzyskania rzetelnej informacji o charakterze rozkładu przestrzennego



Ryc. 8. Rozmieszczenie budynków mieszkalnych na obszarze osady w Dobrzeń Małym, woj. Opole (na podstawie planu publikowanego przez E. Tomczaka 1979) przedstawione w formie układu punktowego

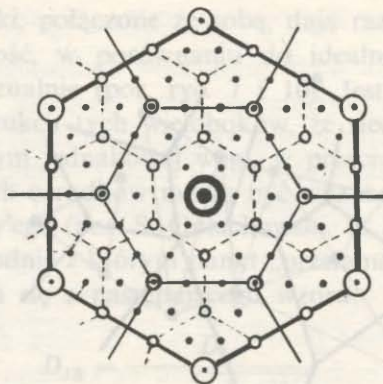
Oprac. i rys. Z. Kobyliński

Fig. 8. Distribution of dwellings in the settlement at Dobrzeń Mały, Opole voiv. (based on the plan published by E. Tomczak 1979).

Designed by Z. Kobyliński

zdarzeń konieczne jest wykonanie szeregu różnych testów, zmieniając przy tym takie parametry, jak wielkość obszaru badanego, czy liczbę punktów losowych umieszczonych w tym obszarze. Jest także oczywiste, że identyfikacja wzoru nie jest celem samym w sobie i nie powinna kończyć analizy rozkładu przestrzennego.

Metody „najbliższego sąsiada”, podobnie jak metody „kwadratowe”, mogą bowiem stać się podstawą bardziej szczegółowej analizy. Jeśli np. omówione wyżej proste testy, za pomocą których porównuje się rozkład empiryczny z rozkładem Poissona, wskazują na znaczącą statystycznie tendencję do skupiania się punktów, to wówczas można (i należy) przeprowadzić analizę skupień, np. metodą zaproponowaną przez K. Kintigha i A. Ammermana (1982) oraz J. F. Simka i R. R. Laricka (1983). Metoda ta polega na iteracyjnym wyznaczaniu ilości i współrzędnych centroidów (centralnych punktów skupień), aż do osiągnięcia minimalnej wartości sumy



Ryc. 9. Hierarchiczna sieć heksagonalna ośrodków osadniczych

Wg W. Christallera 1963, s. 66

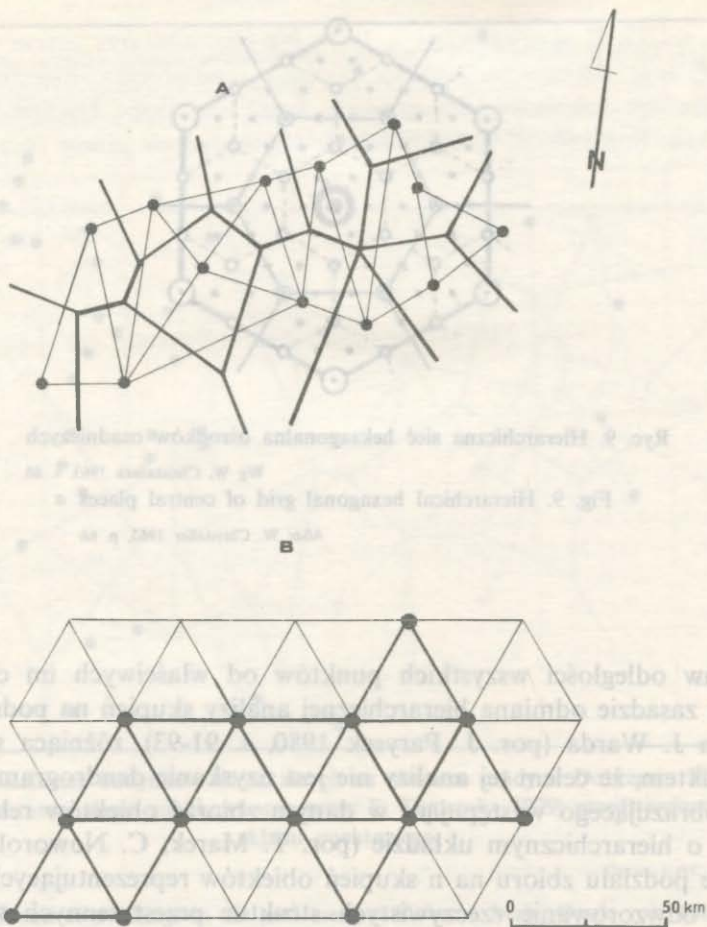
Fig. 9. Hierarchical hexagonal grid of central places

After W. Christaller 1963, p. 66

kwadratów odległości wszystkich punktów od właściwych im centroidów. Jest to w zasadzie odmiana hierarchicznej analizy skupień na podstawie tzw. kryterium J. Warda (por. J. Parysek 1980, s. 91-93), różniącą się od niej jedynie faktem, że celem tej analizy nie jest uzyskanie dendrogramu, tj. grafu drzewa obrazującego występujące w danym zbiorze obiektów relacje odległościowe o hierarchicznym układzie (por. T. Marek, C. Noworol 1983), ale uzyskanie podziału zbioru na n skupień obiektów reprezentujących, zdaniem badacza, odwzorowanie rzeczywistych struktur przestrzennych w systemie społeczno-kulturowym (np. skupienia osadnicze w przypadku analizy rozmieszczenia stanowisk archeologicznych w regionie lub obszary aktywności jednej rodziny w przypadku analizy rozmieszczenia artefaktów w obrębie osady).

Na zakończenie tego zwięzłego przeglądu podstawowych metod analizy przestrzennej układów punktowych wspomnieć trzeba o dwóch jeszcze prostych metodach, nie tyle matematycznych czy statystycznych, co graficznych, pozwalających na ocenę zbieżności rozkładu empirycznego z teoretycznym rozkładem regularnym. Metody te wykazują przy tym pewien związek z omawianymi ostatnio metodami „najbliższego sąsiada”, również bowiem opierają się na pomiarze najbliższych odległości między zdarzeniami, mogą być zatem stosowane komplementarnie wobec omówionych testów (por. ryc. 7). Metody te różnią się przy tym od omówionych wyżej tym, że jako model teoretyczny przyjmują nie rozkład Poissona, a idealnie regularną sieć christallerowską trójkątów równobocznych.

W. Christaller (1963) analizował rozmieszczenie osiedli ludzkich w południowo-zachodnich Niemczech w końcu XIX w. i stwierdził, że rozmieszczenie



Ryc. 10. Układ grodów z VII-VIII w. na Równinie Pyrzycko-Stargardzkiej (na podstawie katalogu i mapy, publikowanych przez W. Łosińskiego 1982) przedstawiony w formie: A. wieloboków Thiessena i grafu usztywnionego; B. teoretycznego grafu regularnego

Oprac. i rys. Z. Kobylński

Fig. 10. Distribution of hillforts of the 7th-8th centuries A. D. in the Pyrzyce-Stargard Plain (based on catalogue and map published by W. Łosiński 1982), displayed in the form of: A. Thiessen polygons and stable configuration; B. ideal regular pattern

Designed by Z. Kobylński

to może być przedstawione w formie deterministycznego modelu „centralnego miejsca”, w którym ośrodki wyższego rzędu otoczone są heksagonalnymi układami ośrodków niższego rzędu (ryc. 9). Jeśli, jak to ma miejsce najczęściej w badaniach archeologicznych, niemożliwe jest oszacowanie rangi punktu osadniczego, to wówczas hierarchiczna sieć heksagonalna postrzegana jest jako jednorodna sieć trójkątów równobocznych.

Konstrukcja tzw. wieloboków Thiessena polega na wystawianiu prostopadłych do linii łączącej dwa sąsiadujące zdarzenia w połowie długości tej

linii. Prostopadłe odcinki, połączone ze sobą, dają razem pakietaż wieloboków, którego regularność, w porównaniu do idealnej sieci sześciokątów, może być oceniona wizualnie (por. ryc. 7 i 10). Jest rzeczą oczywistą, że zakłada się przy konstrukcji tych wieloboków, że sieć reprezentuje punkty jednoczasowe, a przy tym jednakowej wagi, w przeciwnym wypadku strefy wpływów poszczególnych ośrodków muszą być wyznaczone przy zastosowaniu tzw. stosunku Reilly'ego (por. S. Golachowski, B. Kostrubiec, A. Zagożdżon 1974, s. 65-66), zgodnie z którym punkt „przełamania” wpływów dwóch ośrodków (K) wyznacza się z następującego wzoru:

$$D_{JK} = \frac{D_{IJ}}{1 + \sqrt{\frac{\text{ludność I}}{\text{ludność J}}}} \quad (16)$$

W archeologii nieznaną liczbę ludności zastąpić trzeba szacunkiem rangi ośrodka, opartym na jego wielkości.

Druga metoda, przedstawiona przez M. Chilczuka (1975, s. 58-61), opiera się na analizie sztywnej struktury na płaszczyźnie, w odróżnieniu od metod „najbliższego sąsiada”, za pomocą których rozpatruje się strukturę podatną na transformacje. Twierdzenia ze statyki głoszą, że kratownica jest sztywna (statystycznie wyznaczona), jeśli ilość ramion wychodzących z każdego wierzchołka sieci (gdzie liczba wierzchołków przekracza 3) jest co najmniej o 1 większa od liczby wymiarów przestrzeni, w której kratownica jest zanurzona (S. Golachowski, B. Kostrubiec, A. Zagożdżon 1974, s. 105). Z każdego punktu analizowanej struktury wyprowadzić musimy co najmniej 3 krawędzie, a łącznie graf usztywniony obejmować powinien co najmniej $2N - 3$ odległości między punktami. Z tych odległości wyznaczana jest odległość średnia w układzie, a następnie konstruowana jest sieć regularna trójkątów równobocznych o długości boku równej tej właśnie odległości średniej. Następnie wizualnie lub matematycznie ocenić można zgodność sieci empirycznej z teoretyczną (por. ryc. 7 i 10). Ocena matematyczna opiera się na wyznaczeniu sumarycznej długości przesunięć punktów rzeczywistych w stosunku do teoretycznych.

Analogiczną metodę analizy struktur przestrzennych zaproponowali ostatnio w odniesieniu do układów osadniczych S. Evans i P. Gould (1982). Od cytowanego podejścia M. Chilczuka propozycję tych autorów różni jedynie uwzględnienie rangi poszczególnych punktów osadniczych. Wadą metody M. Chilczuka jest nieporównywalność wyników uzyskanych przy analizie dwóch różnych układów, bowiem wartość wskaźnika nieregularności uzależniona jest od wielkości układu. W celu porównania dwóch układów należy znormalizować te wielkości tak, aby średnia odległość w układzie wynosiła 1.

Zbliżoną metodę analizy grafu rozpiętego na punktach układu, stosowaną w celu wyróżnienia koncentracji punktów, przedstawił A. Różycki (1980).

METODY OCENY ZGODNOŚCI ROZKŁADU DWÓCH ZMIENNYCH

Przy rozpatrywaniu wielu problemów archeologicznych nie interesuje nas charakter rozkładu przestrzennego jednej zmiennej, a jedynie zgodność tego rozkładu, bądź jego odmienność od rozkładu innej zmiennej na tym samym obszarze. Podobnie jak w przypadku analizy rozmieszczenia zdarzeń jednego typu, dysponujemy tu metodami „kwadratowymi” i metodami „odległościowymi”.

METODY „KWADRATOWE”

Istnieje szereg metod porównujących ze sobą dwa rozkłady punktowe, opartych na podziale badanego obszaru na kwadraty, a następnie obliczeniu ilości kwadratów, w których występują punkty reprezentujące zdarzenia obu typów (X i Y), ilości kwadratów, w których wystąpiły jedynie zdarzenia X, ilości kwadratów, w których wystąpiły jedynie zdarzenia Y, i wreszcie – ilości kwadratów pustych. W ten sposób uzyskujemy tabelkę czteropolową obecności/nieobecności zdarzeń (I. Hodder, C. Orton 1976, s. 201-204). Dane z tej tabelki są podstawą dla szeregu wskaźników zgodności rozkładów. Jeżeli wymienione wyżej liczebności kwadratów oznaczmy odpowiednio literami a, b, c, d, to wówczas wskaźniki:

$$\frac{(a \cdot d - b \cdot c)^2 \cdot (a + b + c + d)}{(a + b) \cdot (c + d) \cdot (a + c) \cdot (b + d)}, \quad (17)$$

oraz:

$$\frac{\left[(a \cdot d - b \cdot c) - \frac{(a + b + c + d)}{2} \right]^2 \cdot (a + b + c + d)}{(a + b) \cdot (c + d) \cdot (a + c) \cdot (b + d)}, \quad (18)$$

testowane być mogą rozkładem chi-kwadrat, przy hipotezie zerowej głoszącej brak związku między rozkładami.

Dwa inne wskaźniki: Q i V przyjmują wartości zawierające się między -1 (całkowita segregacja), a +1 (maksymalny związek):

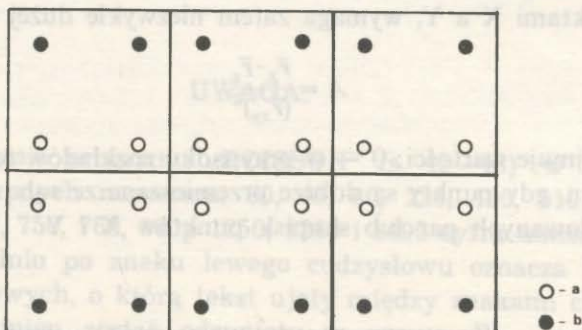
$$Q = \frac{ad - bc}{ad + bc} \quad (19)$$

$$V = \frac{ad - bc}{[(a + b) \cdot (a + c)(c + d)(b + d)]^{\frac{1}{2}}}. \quad (20)$$

Słabą stroną tych wskaźników jest silny wpływ na wynik ze strony wartości d, tzn. liczby kwadratów pustych. W przypadku rozkładu skupionego wartość d jest zazwyczaj duża. Stąd zaproponowano wskaźniki ignorujące tę wartość, np. wskaźnik G:

$$G = \frac{a}{\sqrt{(a+b) \cdot (a+c)}} \quad (21)$$

Ogólnie jednak stwierdzić trzeba, że powyższe metody oceny związku pomiędzy rozkładami punktowymi mają wiele poważnych wad. Największą z nich jest uzależnienie wyniku od wielkości kwadratów, ze względu na to, że rozkład punktów w obrębie kwadratu jest ignorowany, stąd przy kwadratach dużych rozkłady całkowicie rozdzielne mogą być uznane za całkowicie zgodne ze sobą (por. ryc. 11). Stąd, należałoby rozpatrywać kwadraty minimalnej wielkości, co z kolei stwarza trudności obliczeniowe.



Ryc. 11. Schemat ilustrujący przypadek niemożliwości wykrycia separacji rozkładów dwóch zmiennych w przypadku zastosowania metody „kwadratowej”; a – punkty rozkładu zmiennej X ; b – punkty rozkładu zmiennej Y

Oprac. i rys. Z. Kobyliński

Fig. 11. Illustration of the effect of quadrat size on the possibility to demonstrate the separation of two distributions; a – events of the class X ; b – events of the class Y

Designed by Z. Kobyliński

METODY „ODLEGŁOŚCIOWE”

Zdaniem I. Hoddera i C. Ortona (1976, s. 204-207) bardziej odpowiednie dla porównania związku pomiędzy dwoma rozkładami punktowymi są metody, którymi określa się dla każdego z punktów obu rozkładów, jaki punkt, „własny” czy „obcy”, jest jego „najbliższym sąsiadem”. Następnie, również w postaci tabelki czteropolowej, zapisać można liczebności poszczególnych przypadków: a – liczbę przypadków, w których najbliższym sąsiadem zdarzenia X jest inne zdarzenie X ; b – liczbę przypadków, w których najbliższym sąsiadem Y jest X ; c – liczbę przypadków, w których najbliższym sąsiadem X jest Y ; i wreszcie d – liczbę przypadków, w których najbliższym sąsiadem Y jest Y .

Na podstawie tak przygotowanych danych obliczyć można tzw. wskaźnik

rozdziału S , którego wartość zawiera się między $+1$ (całkowity rozdział), a -1 (całkowita zgodność). Wartość 0 oznacza przypadkowe przemieszanie punktów obu rozkładów:

$$S = 1 - \frac{(b+c)(a+b+c+d)}{(a+b)(b+d) + (c+d)(a+c)} \quad (22)$$

Wreszcie w ostatnich latach zaproponowano najdoskonalszy chyba, choć równocześnie najbardziej pracochłonny w obliczeniach, wskaźnik związku między dwoma rozkładami przestrzennymi (I. Hodder, E. Okell 1978). Wskaźnik ten opiera się na obliczeniu średniej odległości między punktami rozkładu X , średniej odległości między punktami rozkładu Y oraz średniej odległości między punktami X a Y , wymaga zatem niezwykle dużej ilości obliczeń:

$$A = \frac{\bar{r}_x \cdot \bar{r}_y}{(\bar{r}_{xy})^2} \quad (23)$$

Wskaźnik przyjmuje wartości: 0 – w przypadku rozkładów rozdzielonych; 1 – w przypadku, gdy punkty są dobrze przemieszane ze sobą; ponad 1 – w przypadku izolowanych par lub skupisk punktów X i Y .

PROGRAM MIKROKOMPUTEROWY POISSON

W celu ułatwienia zastosowania podstawowych testów statystycznych analizy rozkładów punktowych jednej zmiennej, należących do grupy testów „najbliższego sąsiada”, opracowano program w języku BASIC V2, obliczający statystyki Clarka-Evansa, korekcję Pindera i in., TNND, statystykę Skellama-Moore'a, statystykę Pielou i Hopkinsa. Publikowana wersja opracowana została dla mikrokomputera Commodore, jednakże po niewielkich zmianach dostosowujących symbole do konkretnych implementacji BASICu, nadaje się do zastosowania w dowolnym mikrokomputerze. Jest to zmieniona i rozszerzona wersja programu TEST, napisanego w języku FORTRAN IV (Z. Kobylński, J. Zagrodzki, W. Brzeziński 1984).

Wymagania programu są następujące: zakłada się, że obszar badany ma postać prostokąta o bokach równoległych do osi układu współrzędnych; znane są współrzędne punktów narożnych: $X_1, Y_1; X_2, Y_1; X_1, Y_2; X_2, Y_2$; znana jest lokalizacja wszystkich punktów obszaru, podana w formie wartości dwóch współrzędnych; znana jest liczba punktów – N , pole powierzchni obszaru badanego – A i długość obwodu tego obszaru – L . Wczytywanie wszystkich danych odbywa się w trybie konwersacyjnym z klawiatury. Operator decyduje również o ilości symulacji rozmieszczania punktów losowych (zmienna SYM) i liczbie punktów umieszczanych w każdej symulacji (zmienna F). Wydruk obejmuje również podstawowe informacje ułatwiające interpretację wyników.

UWAGA:

W tabulogramie programu „POISSON” (s. 45—48) w wierszach oznaczonych numerami 30, 40, 80, 230, 400, 510, 590, 600, 640, 660, 750, 760, 800p 1250, 1260 i 1490 cyfra umieszczona bezpośrednio po znaku lewego cudzysłowu oznacza liczbę miejsc znakowych, o którą tekst ujęty między znakami cudzysłowów powinien zostać odsunięty w prawo dla uzyskania przejrzystej graficznie postaci wydruku, np.

jest: 30 PRINT „13 PROGRAM POISSON”

powinno być: 30 PRINT”

PROGRAM POISSON”

Tabulogram programu „POISSON”

```

05 REM PROGRAM POISSON
10 REM ZBIGNIEW KOBYLIŃSKI 1986
20 PRINT „SHIFT/CLS HOME”: REM CZYSZCZENIE EKRANU
30 PRINT „13 PROGRAM POISSON”
40 PRINT „5 OCENIA ZGODNOŚĆ ROZKŁADU PUNKTÓW”
50 PRINT „NA PŁASZCZYŹNIE Z ROZKŁADEM POISSONA”
60 PRINT „METODĄ NAJBLIŻSZEGO SĄSIADA TYPU”
70 PRINT „ZDARZENIE-ZDARZENIE I PUNKT LOSOWY-ZDARZENIE”
80 PRINT: PRINT: PRINT „17 ***”: PRINT
90 PRINT „WCZYTYWANIE WSPÓŁRZĘDNYCH PUNKTÓW”
100 PRINT: PRINT „PODAJ LICZBĘ PUNKTÓW (N)”: INPUT N
110 DIM P (N, 3)
120 PRINT: PRINT „PODAWAJ KOLEJNO DANE:”
130 PRINT „NUMER PUNKTU, WSPÓLRZĘDNA X, WSPÓLRZĘDNA Y”
140 PRINT „PO KAŻDEJ WARTOŚCI NACIŚNIJ RETURN”
150 REM ZAPEŁNIANIE DANYMI PÓŁ MACIERZY ZDARZEŃ P
160 FOR W1 = 1 TO N
170 FOR W2 = 1 TO 3
180 INPUT P
190 LET P (W1, W2) = P
200 NEXT W2
210 NEXT W1
220 REM WYDRUK MACIERZY ZDARZEŃ P
230 PRINT: PRINT „13 MACIERZ PUNKTÓW”
240 PRINT: PRINT „NUMER”, „X”, „Y”: PRINT
250 FOR U = 1 TO N
260 PRINT P (U, 1), P (U, 2), P (U, 3)
270 NEXT U
280 PRINT: PRINT „PODAJ POLE POWIERZCHNI (A)”: INPUT A
290 REM OBLICZANIE ŚREDNIEJ ODLEGŁOŚCI DO NAJBLIŻSZEGO SĄSIADA
300 FOR I = 1 TO N
310 ND = 999999: REM ND — ODLEGŁOŚĆ ZDARZENIA I DO NAJBLIŻSZEGO
    SĄSIADA
320 M1 = P(I, 1): X1 = P(I, 2): Y1 = P(I, 3)
330 FOR J = 1 TO N
340 M2 = P(J, 1): X2 = P(J, 2): Y2 = P(J, 3)
350 IF M2 = M1 THEN 380
360 R = SQR(((X2 - X1)↑2) + ((Y2 - Y1)↑2))
370 IF R < ND THEN ND = R
380 NEXT J
390 KR = ND↑2: REM KR — KWADRAT NAJBLIŻSZEJ ODLEGŁOŚCI
400 T = T + ND: REM T — SUMA NAJBLIŻSZYCH ODLEGŁOŚCI
410 S = S + KR: REM S — SUMA KWADRATÓW NAJBLIŻSZYCH ODLEGŁOŚCI
420 NEXT I
430 REM OBLICZANIE STATYSTYK CLARKA-EVANSA
440 RO = T/N
450 RE = 1/(2 * SQR(N/A))
460 E = 0.26136/SQR(N*(N/A))
470 REM WYDRUK WYNIKÓW

```



```

480 PRINT „SHIFT/CLS HOME”: REM CZYSZCZENIE EKRANU
490 PRINT „13 TEST CLARKA-EVANSA”
500 PRINT „RO = ”; RO, „RE = ”; RE
510 PRINT „13 R = ”; RO/RE
520 PRINT „DLA ROZKŁADU POISSONA R = 1, DLA ROZKŁADU”
530 PRINT „SKUPIONEGO DAŻY DO Ø, DLA REGULARNEGO – DO 2.15”
540 PRINT: PRINT „C = ”; C
550 IF N < 100 THEN 580
560 PRINT „WARTOŚĆ C PORÓWNAJ Z ROZKŁADEM NORMALNYM”
570 GOTO 590
580 PRINT „WARTOŚĆ C PORÓWNAJ Z ROZKŁADEM III PEARSONA”
590 PRINT: PRINT „16 ***”
600 PRINT „6 MODYFIKACJA PINDERA ET AL.”
610 DRAN = (0.497 + 0.127 * SQR(4/N)) * SQR(A/N)
620 RN = RO/DRAN: REM RN – WARTOŚĆ R W MODYFIKACJI PINDERA ET AL.
630 PRINT: PRINT „DRAN = ”; DRAN, „DOBS = ”; RO
640 PRINT „13 RN = ”; RN
650 PRINT „WARTOŚĆ RN PORÓWNAJ Z DIAGRAMEM PINDERA ET AL.”
660 PRINT „16 ***”
670 PRINT „6 TEST TNND DONNELLYEGO”
680 PRINT: PRINT „PODAJ DŁUGOŚĆ OBWODU OBSZARU (L)”; INPUT L
690 REM OBLICZANIE STATYSTYKI DONNELLYEGO
700 TM = SQR(0.07 * A + ((0.037 * L)/SQR(A/N)))
710 TL = 0.5 * SQR(N/A) + (0.051 + 0.041/SQR(N)) * L
720 TNND = (T - TL)/TM
730 PRINT: PRINT „TNND = ”; TNND
740 PRINT „DLA ODRZUCENIA HØ, TNND POWINNO BYĆ > 1.96 LUB < -1.96”
750 PRINT: PRINT „15 ***”
760 PRINT „6 TEST SKELLAMA-MOOREA”
770 PRINT: PRINT „WYNIK TESTU = ”; 2 *  $\pi$  * (N/A) * S
780 PRINT „WYNIK PORÓWNAJ Z ROZKŁADEM CHI-KWADRAT”
790 PRINT „PRZY 2N STOPNIACH SWOBODY”
800 PRINT: PRINT „16 *****”
810 PRINT „TESTY PUNKT LOSOWY-ZDARZENIE”
820 PRINT: PRINT „PODAJ WSPÓŁRZĘDNE NAROŻNYCH PUNKTÓW”
830 PRINT „OBSZARU BADANEGO”
840 PRINT „X1 = ”;
850 INPUT X1
860 PRINT „X2 = ”;
870 INPUT X2
880 PRINT „Y1 = ”;
890 INPUT Y1
900 PRINT „Y2 = ”;
910 INPUT Y2
920 REM GENEROWANIE PUNKTÓW LOSOWYCH W OBSZARZE
930 PRINT „ILE SYMULACJI ROZMIESZCZENIA PUNKTÓW”
940 PRINT „LOSOWYCH CHCESZ WYKONAĆ (SYM)?”;
950 INPUT SYM
960 PRINT „ILE PUNKTÓW CHCESZ UMIEŚCIĆ W OBSZARZE (F)?”;
970 INPUT F
980 DIM M(SYM, F, 3): REM DEKLARACJA MACIERZY PUNKTÓW LOSOWYCH
990 REM SYMULACJA ROZMIESZCZENIA PUNKTÓW
1000 FOR B = 1 TO SYM

```

```

1010 REM ZAPEŁNIANIE DANYMI PÓŁ MACIERZY PUNKTÓW LOSOWYCH
1020 FOR G = 1 TO F
1030 N1 = RND(Ø) * X2
1040 IF N1 < X1 GOTO 1030
1050 LET M(B, G, 2) = N1
1060 NEXT G
1070 FOR H = 1 TO F
1080 N2 = RND(Ø) * Y2
1090 IF N2 < Y1 GOTO 1080
1100 LET M(B, H, 3) = N2
1110 NEXT H
1120 REM OBLICZANIE ODLEGŁOŚCI OD PUNKTÓW LOSOWYCH DO NAJBLIŻ-
    SZYCH IM ZDARZEŃ
1130 KW = Ø: SK = Ø
1140 FOR K1 = 1 TO F
1150 KD = 999999: REM KD - NAJBLIŻSZA ODLEGŁOŚĆ OD PUNKTU LOSOWEGO
    DO ZDARZENIA
1160 FOR K2 = 1 TO N
1170 D = SQR(((M(B, K1, 2) - P(K2, 2)) ↑ 2) + ((M(B, K1, 3) - P(K2, 3)) ↑ 2))
1180 IF D < KD THEN KD = D
1190 NEXT K2
1200 KW = KD ↑ 2: REM KW - KWADRAT NAJBLIŻSZEJ ODLEGŁOŚCI
1210 SK = SK + KW
1220 NEXT K1
1230 REM WYDRUK WYNIKÓW
1240 PRINT „SHIFT/CLS HOME”: REM CZYSZCZENIE EKRANU
1250 PRINT „6 WYNIKI SYMULACJI”: B: PRINT
1260 PRINT „13 TEST PIELOU”
1270 PRINT: PRINT „WYNIK TESTU =”: 2 * π * (N/A) * SK
1280 PRINT: PRINT „WYNIK PORÓWNAJ Z ROZKŁADEM CHI-KWADRAT”
1290 PRINT „PRZY 2F STOPNIACH SWOBODY”
1300 IF F = N GOTO 1470
1310 REM LOSOWANIE PRÓBY O LICZEBNOŚCI F Z POPULACJI ZDARZEŃ
1320 DS = Ø
1330 FOR F1 = 1 TO F
1340 NF = INT(RND(Ø) * N) + 1
1350 XF = P(NF, 2)
1360 YF = P(NF, 3)
1370 DD = 999999: REM DD - ODLEGŁOŚĆ OD LOSOWO WYBRANEGO ZDARZE-
    NIA DO NAJBLIŻSZEGO SĄSIEDNIEGO ZDARZENIA
1380 FOR F2 = 1 TO N
1390 IF F2 = NF THEN 1420
1400 DF = SQR(((XF - P(F2, 2)) ↑ 2) + ((YF - P(F2, 3)) ↑ 2))
1410 IF DF < DD THEN DD = DF
1420 NEXT F2
1430 D2 = DD ↑ 2: REM D2 - KWADRAT ODLEGŁOŚCI DO NAJBLIŻSZEGO SĄSIA-
    DA
1440 DS = DS + D2: REM DS - SUMA KWADRATÓW NAJBLIŻSZYCH ODLEGŁOŚCI
1450 NEXT F1
1460 AH = SK/DS: HOP = SK/(SK + DS): GOTO 1490
1470 AH = SK/DS
1480 HOP = SK/(SK + S)
1490 PRINT: PRINT „13 TEST HOPKINSA”

```



```

1500 PRINT: PRINT „AH =”; AH
1510 PRINT „WYNIK PORÓWNAJ Z ROZKŁADEM F PRZY 2F I 2F STOPNIACH
      SWOBODY”
1520 PRINT: PRINT „HOP =”; HOP
1530 PRINT „WARTOŚĆ PORÓWNAJ Z ROZKŁADEM BETA”
1540 PRINT „WARTOŚĆ  $\geq 0.5$  WSKAZUJE SKUPIENIE”
1550 PRINT „WARTOŚĆ  $\leq 0.5$  WSKAZUJE REGULARNOŚĆ”
1560 NEXT B: REM PRZEJŚCIE DO NASTĘPNEJ SYMULACJI ROZMIESZCZENIA
      PUNKTÓW LOSOWYCH
1570 END

```

WYKAZ CYTOWANEJ LITERATURY

Wykaz skrótów

„Am. Ant.” – „American Antiquity”, Washington.

Literatura

Broadbent N.

1979 *Coastal resources and settlement stability. A critical study of a Mesolithic site complex in Northern Sweden*, [w:] *Archaeological Studies*, A. 3, Uppsala.

Broadbent S.

1980 *Simulating the ley hunter*, „Journal of the Royal Statistical Society”, seria A, t. 143, s. 108-140.

Brzeziński W., Dulinicz M., Kobylński Z., Lichy B., Moszczyński W. A.

1985 *Rozpoznawanie stanowisk osadniczych metodą reprezentacyjną: badania w Wyszogrodzie, woj. plockie, stan. 2a*, „Sprawozdania Archeologiczne”, t. 37, s. 251-270.

Chilczuk M.

1975 *Osadnictwo wiejskie. Metody badań koncentracji zabudowy i kształtów wsi*, Warszawa.

Christaller W.

1963 *Ośrodki centralne w Południowych Niemczech*, „Przegląd Zagranicznej Literatury Geograficznej”, z. 1, s. 1-72.

Clark P. J., Evans F. C.

1954 *Distance to nearest neighbour as a measure of spatial relationships in populations*, „Ecology”, t. 35, s. 445-453.

Dacey M. F.

1964 *Modified Poisson probability law for point pattern more regular than random*, „Annals of the Association of American Geographers”, t. 54, s. 559-565.

1973 *Statistical tests of spatial association in the locations of tool types*, „Am. Ant.”, t. 38, s. 320-328.

Dacey M. F., Tung T.

1962 *The identification of randomness in point patterns*, „Journal of Regional Science”, t. 4, s. 83-96.

Diggle P. J.

1976 *Note on the Clark and Evans test of spatial randomness*, [w:] I. Hodder, C. Orton 1976, s. 246-248.

1977 *The detection of random heterogeneity in plant populations*, „Biometrics”, t. 33, s. 390-393.

1979 *On parameter estimation and goodness-of-fit testing for spatial point patterns*, „Biometrics”, t. 35, s. 87-101.

- Diggle P. J., Besag J., Gleaves J. T.
1976 *Statistical analysis of spatial point patterns by means of distance methods*, „Biometrics”, t. 32, s. 659-667.
- Donnelly K. P.
1978 *Simulations to determine the variance and edge effect of total nearest-neighbour distance*, [w:] *Simulation studies in archaeology*, I. Hodder red., Cambridge, s. 91-95.
- Dulinicz M.
1983 *Niektóre aspekty zastosowania w archeologii geograficznych metod analizy przestrzennej osadnictwa*, „Kwartalnik Historii Kultury Materialnej”, R. 31, nr 3, s. 299-315.
- Evans S., Gould P.
1982 *Settlement models in archaeology*, „Journal of Anthropological Archaeology”, t. 1, s. 275-304.
- Gifford D. P.
1980 *Ethnoarchaeological contributions to the taphonomy of human sites*, [w:] *Fossils in the making*, A. K. Behrensmeier, A. P. Hill red., Chicago, s. 93-106.
- Godłowski K.
1985 *Przemiany kulturowe i osadnicze w południowej i środkowej Polsce w młodszym okresie przedrzymskim i w okresie rzymskim*, Wrocław – Warszawa – Kraków – Gdańsk – Łódź.
- Golachowski S., Kostrubiec B., Zagożdżon A.
1974 *Metody badań geograficzno-osadniczych*, Warszawa.
- Góralski A.
1974 *Metody opisu i wnioskowania statystycznego w psychologii*, Warszawa.
- Greig-Smith P.
1964 *Quantitative plant ecology*, London.
- Hamond F. W.
1980 *The interpretation of archaeological distribution maps: biases inherent in archaeological fieldwork*, „Archaeo-Physika”, t. 7, s. 193-216.
- Hietala H. J., Stevens D. E.
1977 *Spatial analysis: multiple procedures in pattern recognition studies*, „Am. Ant.”, t. 42, nr 4, s. 539-559.
- Hilczarówna Z.
1967 *Dorzecze górnej i środkowej Obry od VI do początków XI wieku*, Wrocław – Warszawa – Kraków.
- Hines W. G. S., Hines R. J. O.
1979 *The Eberhardt statistic and the detection of non randomness of spatial point distributions*, „Biometrika”, t. 66, s. 73-79.
- Hodder I.
1977 *Some new direction in the spatial analysis of archaeological data at the regional scale (macro)*, [w:] *Spatial archaeology*, D. L. Clarke red., London, s. 223-351.
- Hodder I., Hassall M.
1971 *The non-random spacing of Romano-British walled towns*, „Man”, t. 6, s. 391-407.
- Hodder I., Okell E.
1978 *An index for assessing the association between distribution of points in archaeology*, [w:] *Simulation studies in archaeology*, I. Hodder red., Cambridge, s. 97-107.
- Hodder I., Orton C.
1976 *Spatial analysis in archaeology*, Cambridge.
- Holgate P.
1965 *Tests of randomness based on distance methods*, „Biometrika”, t. 52, s. 345-353.
- Hopkins B.
1954 *A new method of determining the type of distribution of plant individuals*, „Annals of Botany”, t. 18, s. 213-226.

Johnson G. A.

- 1972 *A test of the utility of Central Place Theory in archaeology*, [w:] *Man, settlement and urbanism*, P. J. Ucko, R. Tringham, G. Dimmbleby red., London, s. 769-785.

King L. J.

- 1969 *Statistical analysis in geography*, Englewood Cliffs.

Kintigh K. W., Ammerman A. J.

- 1982 *Heuristic approaches to spatial analysis in archaeology*, „Am. Ant.”, t. 47, nr 1, s. 31-63.

Kobyliński Z., Zagrodzki J., Brzeziński W.

- 1984 *APIS: komputerowy program próbkowania losowego i statystycznej analizy przestrzennej stanowiska archeologicznego*, „Archeologia Polski”, t. 29, z. 1, s. 41-55.

Kostrubiec B.

- 1969 *Miary koncentracji w badaniach geograficznych*, „Przegląd Geograficzny”, t. 41, z. 2, s. 229-251.

Kruk J.

- 1979 *Gospodarka w Polsce południowo-wschodniej w V-III tysiącleciu p.n.e.*, Warszawa.

Krzysztofiak M., Luszniwicz A.

- 1979 *Statystyka*, Warszawa.

Kucharczyk S., Szczepanowski H.

- 1957 *Optymalne poletko poszukiwań archeologicznych*, „Zastosowania Matematyki”, t. 3, nr 2, s. 154-161.

Łosiński W.

- 1982 *Osadnictwo plemienne Pomorza (VI-X w.)*, Wrocław – Warszawa – Kraków – Gdańsk – Łódź.

Marek T., Noworol C.

- 1983 *Wprowadzenie do analizy skupień*, Kraków.

Mead R.

- 1974 *A test for spatial pattern at several scales using data from a grid of contiguous quadrats*, „Biometrics”, t. 30, s. 295-307.

Moore P. G.

- 1954 *Spacing in plant populations*, „Ecology”, t. 35, s. 222-227.

Neyman J.

- 1939 *On a new class of „contagious” distributions, applicable in entomology and bacteriology*, „The Annals of Mathematical Statistics”, t. 10, s. 35-57.

Orton C.

- 1982 *Stochastic process and archaeological mechanism in spatial analysis*, „Journal of Archaeological Science”, t. 9, s. 1-23.

Parysek J. J.

- 1980 *Analiza skupień jako metoda klasyfikacji w geografii*, [w:] *Metody taksonomiczne w geografii*, Z. Chojnicki red., Warszawa – Poznań, s. 87-99.

Pielou E. C.

- 1959 *The use of point to plant distances in the study of pattern of plant population*, „Journal of Ecology”, t. 47, s. 607-613.

Pinder D., Shimada I., Gregory D.

- 1979 *The nearest neighbour statistic: archaeological application and new developments*, „Am. Ant.”, t. 44, nr 3, s. 430-445.

Pinder D. A., Whiterick M. E.

- 1975 *A modification of nearest-neighbour analysis for use in linear situations*, „Geography”, t. 60, s. 16-23.

Platt C.

- 1977 *Problemy rachunku prawdopodobieństwa i statystyki matematycznej*, Warszawa.

Prochorow J. W., Rozanow J. A.

- 1972 *Rachunek prawdopodobieństwa*, Warszawa.

Redman C. L., Watson P. J.

1970 *Systematic, intensive surface collection*, „Am. Ant.”, t. 35, s. 279-291.

Ripley B. D.

1977 *Modelling spatial patterns*, „Journal of the Royal Statistical Society”, seria B, tom 39, s. 172-212.

Roper D.

1977 *Lateral displacement of artifacts due to plowing*, „Am. Ant.”, t. 41, nr 3, s. 372-375.

Różycki A.

1980 *Pojęcie zjawiska koncentracji w polu rozrzutu zabytków ruchomych na stanowiskach archeologicznych*, „Archeologia Polski”, t. 24, z. 1, s. 21-52.

Simek J. F., Larick R. R.

1983 *The recognition of multiple spatial patterns: a case study from the French Upper Paleolithic*, „Journal of Archaeological Science”, t. 10, nr 2, s. 165-180.

Skellam J. G.

1952 *Studies in statistical ecology. I. Spatial pattern*, „Biometrika”, t. 39, s. 346-362.

Tomczak E.

1979 *Wyniki ratowniczych badań wykopaliskowych w Dobrzeńcu Małym, woj. Opole (stanowisko B)*, „Sprawozdania Archeologiczne”, t. 31, s. 167-191.

Whallon R.

1973 *Spatial analysis of occupation floors. I. Application of dimensional analysis of variance*, „Am. Ant.”, t. 38, nr 3, s. 266-278.

1974 *Spatial analysis of occupation floors. II. The application of nearest neighbor analysis*, „Am. Ant.”, t. 39, nr 1, s. 16-34.

Wood J. J.

1971 *Fitting discrete probability distributions to prehistoric settlement patterns*, [w:] *The distribution of prehistoric population aggregates*, G. J. Gumerman red., Prescott, s. 63-82.

ZBIGNIEW KOBYLIŃSKI

BASIC METHODS FOR ANALYSIS OF SPATIAL POINT PATTERNS

Summary

This article is a review of basic statistical and graphic methods for analysis of spatial point patterns, both univariate and bivariate, followed by discussion of possibilities of their application in studying prehistoric settlement structures. The difficulties arising of specific nature of archaeological evidence, which usually is not a statistical representation of homogeneous spatial population, are pointed out. Point pattern studied by an archaeologist can be both reduced in relation to an original one (due to the action of depositional, postdepositional and excavational processes, cf. F. W. Hamond 1980, p. 195-198 fig. 2-8), and more numerous than the original, owing to the mechanical mixing of several samples drawn unconsciously from the chronologically and/or culturally different populations. These deformations of the original pattern can make it impossible to detect any spatial patterning. Thus, the category of archaeological evidence most useful for spatial analyses seems to be the features stably fixed to the ground, such as hillforts for example, because these are relatively „resistant” against spatial transformations.

The statistical methods themselves are also not the perfect instruments, and their informative power is recently widely discussed by mathematicians (cf. C. Orton 1982).

Basic types of discrete stochastic spatial processes generating point distributions are reviewed in the present paper, namely: the Poisson process, the cluster processes (resulting in e.g. negative binomial or Neyman type A distributions), the mixed Poisson processes (apparent contagion processes), and the regular processes. Four classes of methods for studying point processes are defined, after B. D. Ripley (1977, p. 172): „quadrat” methods, „distance” methods, „second-order” methods, and „test-set” methods. The first two classes are described in a more detailed manner.

From among the wide range of quadrat methods the Medvedkov formula for calculating entropy (M. Chilczuk 1975, p. 33-34), gravitation model (S. Golachowski, B. Kostrubiec, A. Zagożdżon 1974, p. 72-78), index of concentration (B. Kostrubiec 1969, p. 229-251), and Lexis and Steinhaus' index of habilitation (S. Golachowski, B. Kostrubiec, A. Zagożdżon 1974, p. 197-212) are mentioned. Two statistical techniques: calculating an index of dispersion (formula 2), and dimensional analysis of variance (ANOVA), are described more widely.

From among the distance methods, the technique for calculating the distance from an event to its nearest neighbour event, is described (Clark and Evans' statistic – formulas 3-5; Skellam-Moore's statistic – formula 6), and its drawbacks (arising, among, others, from the so-called edge effect; cf. C. Orton 1982; I. Hodder, C. Orton 1976, p. 41-43; B. Kostrubiec 1969, p. 236-237) are discussed. Two recently proposed corrections to Clark and Evans' statistics, namely D. Pinder, I. Shimada and D. Gregory's correction (1979) – formula 9 (Fig. 1), and TNND statistic (K. P. Donnelly 1978) – formulas 10-12, are also presented. Two more advanced methods using also the distances from randomly placed points to their nearest neighbour events (Pielou's test – formula 13, and Hopkins' test – formulas 14-15) are also described.

Informative power of these basic methods was tested against some artificially constructed examples of regular pattern (Fig. 2, Table 1), linear pattern (Fig. 3, Table 2), Neyman's distribution (Fig. 4, Table 3), and Dacey's regular Poisson distribution (Fig. 5, Table 4). The most sensitive analytical instruments appeared the TNND statistic and Hopkins' coefficient.

The same set of methods was then used to study empirical distribution of the Przeworsk culture archaeological sites of the late period of Roman influences (Fig. 6; cf. K. Godłowski 1985, map 9). At first complete pattern, and then three spatial samples, were taken into account. The results of analyses, though not quite univocal (cf. Table 5), allow to state that this distribution belongs in the class of true contagion distributions.

Difference in results obtained by means of various statistical methods is exemplified with an analysis of dwelling distribution in the settlement at Dobrzeń Mały, Opole voiv. (Fig. 8), dated at the late period of Roman influences. In that case only Pielou's and Hopkins' statistics indicated some clustering in the pattern.

On the other hand, the effect of test area size on the value of Clark and Evans' statistic is showed in the case of distribution of hillforts of the phases A–B of early Middle Ages in the basin of Obra River (Fig. 7). Here, regular patterning can be demonstrated only when a fragment of whole configuration is taken into account.

Some recent developments in spatial cluster analysis (K. Kintigh, A. Ammerman 1982; J. F. Simek, R. R. Larick 1983) are reviewed.

Two simple graphic techniques for demonstration regularity of point pattern are described: drawing the so-called Thiessen polygons (Figs. 7 and 10), and drawing the stable configuration of points in the plane (as proposed by M. Chilczuk 1975, p. 58-61; Figs. 7 and 10). Both methods assume Christaller's regular pattern (Fig. 9) as a test distribution.

Shortly described are also, after I. Hodder and C. Orton (1976, p. 201-204; also I. Hodder, E. Okell 1978) basic quadrat and distance methods for analysis the association between distributions of two different event classes in the plane (formulas 17-23). The effect of quadrat size in this case is discussed (Fig. 11).

Finally, a microcomputer program written in BASIC V2, designed for Commodore 64, is presented. The POISSON program makes easy to calculate values of six basic nearest neighbour statistics (Clark-Evans' statistics, Pinder et al. correction, TNND, Skellam-Moore's statistics, Pielou's test, and Hopkins' coefficient) for any point distribution.

Program requirements are simple: it is assumed that an area is rectangular with sides parallel to the coordinate axes; coordinates of corner points ($X_1, Y_1; X_2, Y_1; X_2, Y_2; X_1, Y_2$) are known as well as the coordinates of every point in the distribution. Input of all data is possible in conversational mode. Variables: N — number of events; A — area size; L — length of area sides; SYM — number of simulations of random points placing in the studied area; F — number of random points placed each time in the area.

Translated by Zbigniew Kobyliński

Adres Autora:

Dr Zbigniew Kobyliński

Zakład Epoki Metali

Instytutu Historii Kultury Materialnej PAN

Al. Świerczewskiego 105, 00-140 Warszawa

L. UWAGI WSTĘPNE

Krzemieniarstwo epizydurowego przykarpackiego kręgu kulturowego (zwącego dalej e.p.k.k.) jest podstawą studiów nad wytwórczością i użytkowaniem narzędzi kamiennych we wczesnej epoce brązu na ziemiach Polski i na znaczonych obszarach krajów sąsiednich. Wynika to zarówno z bogactwa bazy źródłowej, jak i rozległości terytorialnej tego kręgu. Temat ten jest obecnie przedmiotem szerszych studiów realizowanych przez autorów.

Mamy nadzieję, że zaprezentowany służy ułatwi właściwe odebranie przygotowywanego opracowania „cultura krzemieniarstwa e.p.k.k.” w którym sygnalizowane tutaj problemy zostaną rozwinięte. Przypuszczamy, że niejedną uwagę naświetlił pewne kwestie związane z wewnętrznymi relacjami kulturowymi wielkiej i trudnej do zdefiniowania jednostki, jaką jest wspomniany krąg.

U podstaw naszego stosunku do prezentowanego zagadnienia leży przekonanie o specyficznych cechach bazy źródłowej. Po pierwsze, materiał kamienny stanowi pozostałość najbardziej elementarnych czynności podlegających się nim ludzi. Jedynym rodzajem więzi pomiędzy narzędziem a jego użytkownikiem było zapotrzebowanie na dany wytwór. Naszym zdaniem narzędzia kamienne w minimalnym tylko stopniu podlegały wpływom mody oraz tradycji rozumianej jako emocjonalna więź z przeszłością. Nie mogą zmienić tego poglądu wyjątki, np. niektóre rodzaje broni, czy okazy o szczególnej wartości. Po drugie, zestaw narzędzi kamiennych wykorzystywanych przez daną społeczność był wypadkową jej potrzeb oraz możliwości i zmieniał się w zależności od tych czynników.

Podobne rozumienie krzemieniarstwa odczytujemy w niejednej pracy dotyczącej okresów starszych, a jego źródła doszukać się można w fascynującym ujęciu musterien (P. Bordes, M. Bourgon 1931; P. Bordes 1933 a; tamże 1953 b). Konsekwencją powyższego jest rozpatrywanie krzemieniarstwa w

