

Jarosław Zawadzki

## Badania korelacji przestrzennych zawartości wybranych pierwiastków śladowych w glebach Warszawy i okolic

Siłę współzależności zmiennych losowych określa się najczęściej przy pomocy tradycyjnych wielkości statystycznych, takich jak współczynnik korelacji liniowej Pearsona, współczynnik korelacji  $\tau$  Kendalla, współczynnik korelacji rang Spearmana, współczynnik zbieżności Czuprowa oraz stosunki korelacyjne Pearsona [1,2]. Jakkolwiek miary te przeznaczone są do badań zmiennych losowych, które nie zależą od współrzędnych przestrzennych, często stosuje się je także w badaniach związanych ze środowiskiem naturalnym, rolnictwem, hydrologią itp., gdzie wymiar przestrzenny odgrywa istotną rolę w interpretacji otrzymanych wyników. W takiej sytuacji korzystanie z powyższych miar statystycznych może prowadzić do zbyt dużych uproszczeń lub nawet fałszywych wniosków.

Badaniami zjawisk losowych, z uwzględnieniem wymiaru przestrzennego, zajmuje się autonomiczna gałąź metod statystycznych, zwana geostatystyką, której początki wywodzą się z badań geologicznych. Pierwsze prace z geostatystyki ukazały się na przełomie lat 50. i 60. XX wieku [3–5]. Geostatystyka dostarcza wielu oryginalnych metod opisu ciągłości przestrzennej, będącej zasadniczą cechą wielu zjawisk przyrodniczych. W związku z tym obserwuje się ostatnio gwałtowny wzrost liczby prac poświęconych zastosowaniu geostatystyki do rozwiązywania zagadnień związanych ze środowiskiem naturalnym [6]. Do podstawowych zagadnień, którymi zajmuje się geostatystyka należy również interpretacja oraz przewidywanie (estymacja, symulacja) rozkładów przestrzennych badanych zjawisk [7–13].

W geostatystyce używa się tzw. zmiennych zregionalizowanych, związanych z położeniem, które mają właściwości pośrednie pomiędzy zmienną losową i deterministyczną. Szczególne uznanie metody geostatystyczne znalazły w górnictwie, gdzie problem podejmowania właściwych decyzji, np. dotyczących lokalizacji złóż, wiąże się ze znacznymi kosztami. Również w naszym kraju w ostatnim ćwierćwieczu opublikowano wiele prac, w których zastosowano te metody w takich dziedzinach, jak geologia górnicza, geofizyka, hydrogeologia i ochrona środowiska (np. [14–16]). Większość zastosowań geostatystyki ograniczyła się jednak do krigingu, tj. metody estymacji zapewniającej minimum wariancji błędów szacowanych rozkładów. Kriging stosowany jest do wykonywania dokładniejszych map niż przy użyciu alternatywnych metod, jak na przykład metoda najmniejszej krzywizny, czy metoda odwrotnych odległości.

W niniejszej pracy zaprezentowano systematyczny opis współczesnych metod wykorzystywanych do badania korelacji zmiennych zregionalizowanych. W szczególności opisano najczęściej wykorzystywane w geostatystyce miary do opisu korelacji, tj. semiwariancję wzajemną (krzyżową), zwaną również wariogramem wzajemnym (krzyżowym) oraz wzajemną (krzyżową) funkcję korelacji, zwaną również korelogramem wzajemnym (krzyżowym). Dodatkowo omówiono rolę i zastosowanie liniowego modelu koregionalizacji. Następnie wykorzystano powyższe miary i metody do systematycznego opisu korelacji przestrzennych zawartości wybranych pierwiastków śladowych (As, Zn, Cd, Co, Mn, Cu, Ni i Pb) w glebach Warszawy i okolic. W pracy porównano również wyniki tradycyjnych obliczeń współczynnika korelacji Pearsona ze współczynnikiem korelacji otrzymanym drogą modelowania wariogramów oraz wariogramów krzyżowych (wzajemnych).

### Pobór próbek i metody analityczne

Podstawą do obliczeń geostatystycznych były pomiary zawartości arsenu, cynku, kadmu, kobaltu, manganu, miedzi, niklu i ołowiu w próbkach gleb. Pomiary te zostały wykonane w ramach prac naukowych, przeprowadzonych przez Państwowy Instytut Geologiczny, mających na celu opracowanie „Atlasu geochemicznego Warszawy i okolic” [17]. Analizą geostatystyczną objęto obszar badań zbliżony kształtem do prostokąta, który został oznaczony na mapie województwa mazowieckiego (rys. 1). Rzeczywiste wymiary tego obszaru wynoszą około 41 km w kierunku wschód-zachód oraz około 37 km w kierunku północ-południe. Zbiór danych wykorzystanych w pracy zawierał 1358 pomiarów zawartości każdego z pierwiastków w glebie. Siatka punktów pomiarowych była nieregularna, jednakże odległości pomiędzy najbliższymi punktami były podobne. Średnia odległość między punktami pomiarowymi wynosiła 1,1 km.

Próbki gleb o masie około 0,5 kg pobrano ręczną sondą o średnicy około 8 cm z głębokości 0,0±0,2 m. Następnie próbki wysuszono w temperaturze pokojowej, przesiano przez sito o oczkach 1,0 mm i poprzez kwartowanie uzyskano próbkę analityczną o masie około 100 g. Pomiary zawartości pierwiastków śladowych w próbkach gleb z obszaru Warszawy wykonano metodą absorpcji atomowej na spektrometrze firmy Pye Unicam SP 9-80, natomiast w próbkach z zachodniej części badanego obszaru pomiary wykonano metodą ICP za pomocą spektrometru emisyjnego ze wzbudzeniem plazmowym firmy Philips 8060.



Rys. 1. Położenie badanego obszaru na tle województwa mazowieckiego

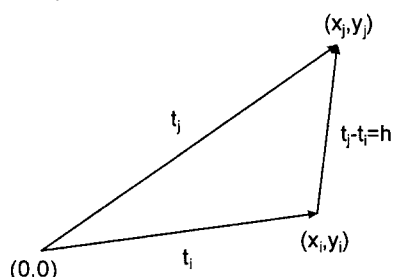
Zgodność wyników oznaczeń uzyskanych tymi dwiema metodami potwierdzono odpowiednimi badaniami porównawczymi. Szczegóły dotyczące poboru i analizy próbek zawarte są w pracy [17].

### Podstawy teoretyczne geostatystycznej analizy korelacji

Badania korelacji przestrzennej przeprowadzono wykorzystując podstawową w geostatystyce miarę ciągłości przestrzennej, tj. semiwariancję (wariogram) oraz semiwariancję krzyżową (wzajemny wariogram krzyżowy). Semiwariancja jest powiązana z innymi miarami ciągłości przestrzennej, takimi jak funkcja kowariancji lub funkcja korelacji (korelogram) [18,19], i dostarcza podobnych informacji. Jednakże ze względów historycznych najczęściej wykorzystywaną miarą jest semiwariancja, zwana zwyczajowo wariogramem.

### Wykresy rozrzutu h i wariogramy

Podstawą do wyznaczenia wariogramu (jak również innych miar ciągłości przestrzennej) są tzw. wykresy rozrzutu h, które pokazują w prostokątnym układzie współrzędnych wszystkie możliwe pary pomiarów (danych) jednakowo odległych w danym kierunku (rys. 2).



Rys. 2. Notacja wektorowa wykorzystana do opisu wykresów rozrzutu h

Odkładając na osi odciętych układu współrzędnych wartości  $v(t_i)$  analizowanej zmiennej „v” w punktach wyznaczonych przez wektory  $t_i$  (indeks „i” przebiega po wszystkich punktach zbioru danych), zaś na osi rzędnych wartości  $v(t_i+h)$  tej zmiennej w punktach wyznaczonych przez wektory  $t_i+h$ , otrzymuje się wykres rozrzutu. Wektor  $h$  nazywany jest wektorem przesunięcia (*lag vector*). Bardzo rzadko punkty pomiarowe tworzą regularną siatkę prostokątną. Zazwyczaj mamy do czynienia z siecią nieregularną, co związane jest najczęściej z możliwościami pomiarowymi w terenie. Obliczając wariogram, za jednakowo odległe uważamy te punkty, które leżą w pewnym zakresie tolerancji długości wektora  $h \pm \Delta h$  oraz w pewnym zakresie tolerancji kierunku tego wektora  $\Phi \pm \Delta \Phi$ .

Wariogram (semiwariancję) definiuje się jako średnią kwadratów różnic pomiędzy odpowiednimi parami danych:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{(i,j)_{h_j=h}} (v_i - v_j)^2 \quad (1)$$

Wyrażenie  $N(h)$  oznacza liczbę par punktów wyznaczonych w opisany powyżej sposób, przy czym wartości wariogramu zależą od długości i kierunku wektora  $h$ .

Czasami oblicza się wariogramy biorąc do obliczeń pary danych położonych względem siebie w określonych odległościach  $l_{hl}$ , bez względu na kierunek wektora  $l_{hl}$ . Obliczone w ten sposób wielkości nazywamy izotropowymi, czyli niezależnymi od kierunku. Wariogramy zależne od kierunku, noszą nazwę wariogramów kierunkowych.

Często wykonuje się mapy wariogramowe, na których przedstawia się wartości funkcji wariogramu, obliczone dla wszystkich możliwych wektorów  $h$  (dla wszystkich kierunków i zwrotów). Mapa wartości funkcji wariogramu stanowi powierzchnię, która może pomóc przedstawić anizotropię.

### Krzyżowe wykresy rozrzutu h i wariogramy wzajemne

Krzyżowe wykresy rozrzutu  $h$  otrzymuje się w analogiczny sposób, jak opisane wyżej wykresy rozrzutu  $h$ . Jedyną różnicą polega na tym, że w wypadku wykresów krzyżowych, zamiast łączenia w pary wartości tej samej zmiennej w punktach odległych o wektor  $h$ , łączy się w pary wartości dwóch różnych zmiennych „u” i „v” w tych punktach (np. zawartości dwóch różnych metali). Wariogram wzajemny (krzyżowy) dany jest równaniem:

$$\gamma_{uv}(h) = \frac{1}{2N(h)} \sum_{(i,j)_{h_j=h}} (v_i - v_j)(u_i - u_j) \quad (2)$$

Oznaczenia w równaniu (2) są analogiczne do oznaczeń stosowanych do opisu wariogramu (1).

### Funkcja korelacji (korelogram) oraz wzajemna (krzyżowa) funkcja korelacji (korelogram wzajemny)

Alternatywnymi miarami ciągłości przestrzennej w stosunku do wariogramu (wariogramu krzyżowego) są funkcja kowariancji (funkcja kowariancji krzyżowej) oraz funkcja korelacji, zwana również korelogramem (funkcja korelacji krzyżowej określana również korelogramem krzyżowym). Funkcja kowariancji  $C(h)$  definiowana jest równaniem:

$$C(h) = \frac{1}{N(h)} \sum_{(i,j)_{h_j=h}} (v_i v_j - m_{-h} m_{+h}) \quad (3)$$

Wartości danych oznaczone są poprzez  $v_1 \dots v_n$ . Sumowaniu podlega tylko  $N(h)$  par danych, odseparowanych o wektor  $h$ , zaś  $m_{-h}$  jest średnią ze wszystkich wartości danych odseparowanych o wektor  $-h$ :

$$m_{-h} = \frac{1}{N(h)} \sum_{i|h_j=h} v_i \quad (4)$$

Analogicznie,  $m_{+h}$  jest średnią ze wszystkich wartości danych odseparowanych o wektor  $h$ , zdefiniowaną jako:

$$m_{+h} = \frac{1}{N(h)} \sum_{j|h_j=h} v_j \quad (5)$$

Częściej stosuje się funkcję korelacji  $\rho(h)$ , będącą standaryzowaną funkcją kowariancji:

$$\rho(h) = \frac{C(h)}{\sigma_{-h} \sigma_{+h}} \quad (6)$$

w której odchylenie standardowe  $\sigma_{-h}$ , obliczane na podstawie wszystkich wartości danych odległych o wektor  $-h$ , dane jest równaniem:

$$\sigma_{-h}^2 = \frac{1}{N(h)} \sum_{i|h_j=h} v_i^2 - m_{-h}^2 \quad (7)$$

Analogicznie, odchylenie standardowe  $\sigma_{+h}$ , obliczane na podstawie wszystkich wartości danych odseparowanych o wektor  $h$ , jest definiowane jako:

$$\sigma_{+h}^2 = \frac{1}{N(h)} \sum_{j|h_j=h} v_j^2 - m_{+h}^2 \quad (8)$$

Funkcja kowariancji wzajemnej definiowana jest następująco:

$$C_{uv}(h) = \frac{1}{N(h)} \sum_{(i,j)|h_{ij}=h} u_i v_j - m_{-u} m_{+v} \quad (9)$$

w której:

$$m_{v_{-h}} = \frac{1}{N(h)} \sum_{j|h_{ij}=h} v_j \quad (10a)$$

oraz:

$$m_{v_{+h}} = \frac{1}{N(h)} \sum_{j|h_{ij}=h} v_j \quad (10b)$$

Funkcja korelacji wzajemnej dana jest równaniem:

$$\rho_{uv}(h) = \frac{C_{uv}(h)}{\sigma_{u_{-h}} \sigma_{v_{+h}}} \quad (11)$$

w której:

$$\sigma_{u_{-h}}^2 = \frac{1}{N(h)} \sum_{i|h_j=h} u_i^2 - m_{u_{-h}}^2 \quad (12a)$$

oraz:

$$\sigma_{u_{+h}}^2 = \frac{1}{N(h)} \sum_{i|h_j=h} v_j^2 - m_{v_{+h}}^2 \quad (12b)$$

Warto zauważyć, że dla  $h=0$  wartość korelogramu krzyżowego jest równa współczynnikowi korelacji liniowej Pearsona pomiędzy zmiennymi „ $u$ ” i „ $v$ ”. Dla  $h \neq 0$  wartość korelogramu krzyżowego może być traktowana jako wartość współczynnika korelacji liniowej Pearsona, obliczanego dla zmiennych „ $u$ ” i „ $v$ ” w punktach odległych o wektor  $h$ .

## Modelowanie wariogramów

Wariogramy stanowią efektywną charakterystykę ciągłości przestrzennej, występującej w badanym zbiorze danych. Niestety, w wielu sytuacjach trudno jest wykorzystać wariogramy empiryczne. Przykładem może być estymacja metodą kriginu, należąca do podstawowych metod estymacji w geostatystyce. Metoda ta wymaga znajomości wartości funkcji wariogramu dla wielu długości i kierunków wektora  $h$ , które nie zawsze występują w wariogramie eksperymentalnym. Dane doświadczalne, na przykład ze względu na ich rozmieszczenie, nie pozwalają często na wyliczenie tych wszystkich niezbędnych wartości funkcji wariogramu, ułatwiają jednak zbudowanie modelu wariogramu i określenie jego parametrów. W oparciu o model wariogramu można wyliczyć jego wszystkie konieczne wartości.

Funkcje używane do modelowania wariogramu nie mogą być dowolne, lecz muszą być określone dodatkowo [2,20], gdyż jedynie dodatkowo określony model wariogramu zapewnia, że równania kriginu dają jedno – stabilne – rozwiązanie. Najprostszą, a jednocześnie najbardziej praktyczną, metodą otrzymywania dodatkowo określonych wariogramów jest stosowanie do ich budowy kilku funkcji, o których wiadomo, że są dodatkowo określone (tzw. modeli podstawowych). Kombinacja liniowa funkcji dodatkowo określonych jest również funkcją określoną dodatkowo (przy założeniu, że wszystkie wagi tej kombinacji  $w_i$  są nieujemne):

$$\gamma(h) = \sum_{i=1}^n w_i \gamma_i(h) \quad (13)$$

Modele wariogramu, których wykresy osiągają nasycenie, nazywa się przejściowymi, przy czym wartość tego nasycenia określa się w geostatystyce terminem próg (*sill*), zaś odległość od początku układu odniesienia, na jakiej osiągnię jest nasycenie nazywa się zasięgiem oddziaływania (*range of influence*). Wykresy niektórych modeli przejściowych osiągają próg asymptotycznie. W tym wypadku jako zasięg oddziaływania przyjmuje się odległość, na jakiej wariogram osiąga 95% wartości wariancji progowej. Najczęściej używanymi modelami podstawowymi są:

– **model samorodków:**

$$\gamma(h) = 0 \text{ dla } |h|=0 \quad (14a)$$

$$\gamma(h) = C \text{ dla } |h| \neq 0 \quad (14b)$$

w którym  $C$  jest wartością wariancji progowej (prog),

– **model sferyczny:**

$$\gamma(h) = C \left[ 1,5 \frac{|h|}{a} - 0,5 \left( \frac{|h|}{a} \right)^3 \right] \text{ dla } |h| \leq a \quad (15a)$$

$$\gamma(h) = C \text{ dla } |h| > a \quad (15b)$$

w którym stała  $a$  jest wartością zasięgu oddziaływania,

– **model wykładniczy:**

$$\gamma(h) = C \left[ 1 - \exp\left(-\frac{3|h|}{a}\right) \right] \quad (16)$$

– **model gaussowski:**

$$\gamma(h) = C \left[ 1 - \exp\left(-\frac{3|h|^2}{a^2}\right) \right] \quad (17)$$

– **model potęgowy:**

$$\gamma(h) = C|h|^a \quad (18)$$

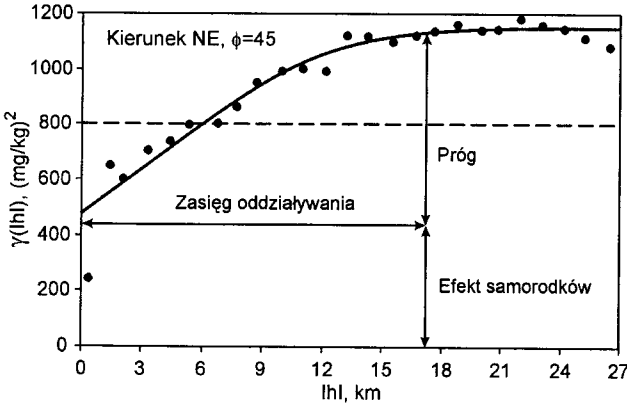
w którym  $a$  jest liczbą z przedziału od 0 do 2,

– model liniowy:

$$\gamma(h) = C h^a \quad (19)$$

który jest szczególnym przypadkiem modelu potęgowego dla  $a=1$ , wydzielony ze względu na częste zastosowania.

Przykładowy wariogram i jego model przedstawiono na rysunku 3 (jakość dopasowania określa suma kwadratów odchyleń wariogramu empirycznego od jego modelu w punktach węzłowych).



Rys. 3. Zapis modelowania wariogramu kierunkowego dla zawartości ołowiu

Modelowanie wariogramów izotropowych i wariogramów wzajemnych oraz inne obliczenia geostatystyczne przeprowadzono za pomocą oprogramowania geostatystycznego Variowin 2.2. Program ten został napisany i jest ciągle rozbudowywany w Uniwersytecie w Lozannie (Szwajcaria). Variowin 2.2 jest pakietem czterech programów (Prevar2D, Vario2D with PCF, Model oraz Grid Display), służących do wszechstronnych obliczeń geostatystycznych [20]. Pozostałe obliczenia statystyczne wykonano z wykorzystaniem programu Statistica 5.1.

### Liniowy model koregionalizacji

Dotychczas omówiono modelowanie jednej zmiennej przy pomocy wariogramu (autowariogramu – przedrostek auto- opuszcza się w sytuacjach, gdy nie prowadzi to do nieporozumienia). Bardzo często występuje jednak sytuacja, gdy bada się dwie lub więcej zmiennych. Wówczas każdą zmienną opisuje się jej własnym autowariogramem, zaś dowolną parę tych zmiennych opisuje się odpowiednim wariogramem wzajemnym. Poniżej zaprezentowano najprostszy przypadek dla dwóch zmiennych.

Metodę modelowania autowariogramów oraz wariogramów wzajemnych dla zmiennych zregionalizowanych, zapewniających dodatnią określoność dowolnej liniowej kombinacji tych zmiennych, nazywa się liniowym modelem koregionalizacji. W celu zapewnienia dodatniej określoności w każdym autowariogramie lub w wariogramie wzajemnym, takie same modele podstawowe muszą występować w identycznych pozycjach. Na przykład, jeżeli występują dwie zmienne „u” i „v”, to modele autowariogramów tych zmiennych, tj.  $\gamma_{uu}(h)$ ,  $\gamma_{vv}(h)$  oraz wariogramu wzajemnego  $\gamma_{uv}(h)$  między nimi, są następujące:

$$\begin{aligned} \gamma_{uu}(h) &= u_0 \gamma_0(h) + u_1 \gamma_1(h) + \dots + u_m \gamma_m(h) \\ \gamma_{vv}(h) &= v_0 \gamma_0(h) + v_1 \gamma_1(h) + \dots + v_m \gamma_m(h) \\ \gamma_{uv}(h) &= w_0 \gamma_0(h) + w_1 \gamma_1(h) + \dots + w_m \gamma_m(h) \end{aligned} \quad (20)$$

Konieczność korzystania z takich samych modeli podstawowych powoduje, że modelowanie autowariogramów

i wariogramów wzajemnych, występujących w modelu koregionalizacji, może być trudne. Często jeden z modelowanych autowariogramów lub wzajemnych wariogramów krzyżowych nie pasuje wystarczająco dobrze do wzajemnych wariogramów empirycznych, podczas gdy pozostałe pasują zadowalająco. Szczegóły liniowego modelu koregionalizacji podane są w pracy [7].

W notacji macierzowej układ równań autowariogramów i wariogramów wzajemnych można zapisać w postaci:

$$\Gamma(h) = B \gamma(h) \quad (21)$$

w której B jest macierzą współczynników.

Wielkością będącą miarą korelacji pomiędzy badanymi zmiennymi zregionalizowanymi jest tzw. współczynnik kodispersji (*codispersion coefficient*), zdefiniowany jako:

$$C_{uv}(h) = \frac{\gamma_{uv}(h)}{\sqrt{\gamma_{uu}(h) \gamma_{vv}(h)}} \quad (22)$$

Powyższy model można oczywiście rozszerzyć na większą liczbę zmiennych.

Warto podkreślić, że współczynnik kodispersji uwzględnia wiele różnych, jednocześnie występujących, udziałów do korelacji pomiędzy zmiennymi. Udziały te, związane z różną skalą odległości (zasięgu), reprezentowane są poprzez różne „modele gniazdowe” wariogramów. Na przykład, wykorzystanie w równaniu (22) wyłącznie składników losowych (*nugget effect*) opisywałoby korelacje w mikroskali (tzn. dla  $h=0$ ). W takim wypadku współczynnik kodispersji jest równy współczynnikowi liniowemu Pearsona. Inne udziały do korelacji związane są ze zmiennością w większych skalach odległości, które są określane przez zasięgi modeli występujących w „strukturach gniazdowych” wariogramu.

Stosunkowo częstą sytuacją jest występowanie w wariogramie dwóch struktur gniazdowych, np. składnika losowego (efektu samorodków) oraz modelu sferycznego. W wypadku występowania takich wariogramów w wyrażeniu (22) współczynnik kodispersji odzwierciedla korelacje w mikroskali ( $h=0$ ). W zakresie odległości większych niż klasa podstawowa wariogramu, a mniejszych od maksymalnego zasięgu modeli sferycznych występujących w równaniu (22), występuje zmiana tego współczynnika, związana ze zmianami korelacji w makroskali. W odległościach większych niż maksymalny zasięg modeli sferycznych, występujących w wyrażeniu (22), współczynnik kodispersji będzie miał wartość stałą, która wynika z udziałów korelacji w mikro- i makroskali.

Szczególnie prosta sytuacja występuje, jeżeli wszystkie autokorelogramy, jak również korelogramy wzajemne (krzyżowe), występujące w wyrażeniu (22) są proporcjonalne do tego samego wariogramu podstawowego, tzn.:

$$\gamma_{ij} = b_{ij} \gamma(h) \quad ij = \{u, v\} \quad (23)$$

w którym  $b_{ij}$  jest odpowiednim współczynnikiem proporcjonalności.

Tego typu korelacja nazywana jest czasami korelacją wewnętrzną (*intrinsic correlation*) i wówczas:

$$C_{ij}(h) = \frac{b_{ij}}{\sqrt{b_{ii}(h) b_{jj}(h)}} \frac{\gamma(h)}{\gamma(h)} = r_{ij} \quad (24)$$

tzn. korelacje pomiędzy zmiennymi nie zależą od odległości, zaś współczynnik kodispersji jest równy współczynnikowi korelacji.

Tabela 1. Statystyki podstawowe zawartości pierwiastków śladowych w glebach Warszawy i okolic

Pierwiastek	Średnia mg/kg	Mediana mg/kg	Minimum mg/kg	Maksimum mg/kg	Kwartył d. mg/kg	Kwartył g. mg/kg	Kwartył 95% mg/kg	Rozstęp mg/kg	Odch. stand. mg/kg	Skośność	Kurtoza
As	5,3	2,0	2,0	2830,0	2,0	4,0	8,0	2,0	68,9	40,3	1650,1
Cd	0,3	0,2	0,2	19,1	0,2	0,2	0,6	0,0	0,6	24,4	770,0
Co	1,9	1,0	1,0	10,0	1,0	2,0	4,0	1,0	1,2	2,1	6,5
Cu	10,0	6,0	1,0	560,0	3,0	11,0	30,0	8,0	21,5	14,4	301,6
Mn	185,2	148,0	1,0	2172,0	76,0	228,0	496,0	152,0	183,0	3,5	21,7
Ni	4,8	4,0	1,0	31,0	2,0	6,0	12,0	4,0	4,0	2,1	6,7
Pb	19,5	13,0	1,0	401,0	9,0	20,0	56,0	11,0	25,7	6,6	66,4
Zn	54,3	34,0	1,0	1432,0	21,0	57,0	159,0	36,0	76,4	7,2	88,3

## Dyskusja wyników

Pierwszym etapem analizy wyników pomiarów zawartości badanych pierwiastków śladowych w glebach Warszawy i okolic było wykonanie obliczeń podstawowych miar statystyki opisowej. Obliczono najczęściej stosowane miary średnie, miary rozproszenia, miary asymetrii i zawartości pierwiastków (tab. 1). Wykonano również wykresy ramkowe (*box and whiskers plots* – rys. 4) zawartości badanych pierwiastków śladowych w glebie, które pozwalają na szybką, wizualną ocenę danych.

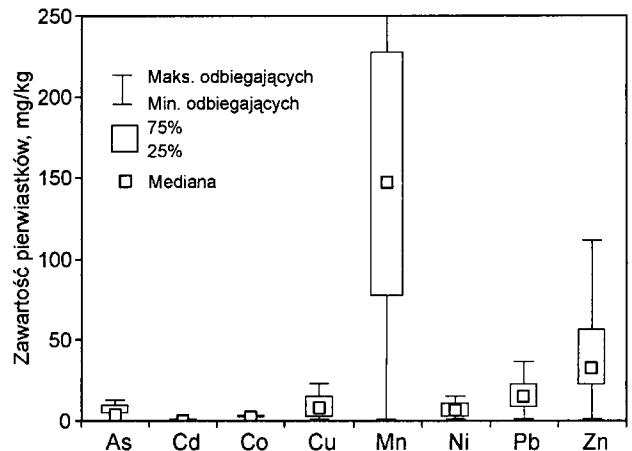
Dla porównania, w tabeli 2 podano graniczne zawartości wybranych pierwiastków śladowych w glebach. Otrzymane rezultaty wykazały, że przeciętna zawartość pierwiastków śladowych w glebach na badanym obszarze nie przekroczyła granicznej zawartości naturalnej, chociaż wartości maksymalne zawartości kadmu, miedzi i cynku mieściły się w klasie V. W wypadku cynku, średnia zawartość nieznacznie przekroczyła zawartość naturalną jedynie w wypadku gleby typu „a”. Mangan powszechnie występuje w skorupie ziemskiej przeciętnie w ilości 0,08%, natomiast naturalna zawartość arsenu w glebie wynosi 0,1+40 mgAs/kg (śr. 6 mgAs/kg) [22].

W dalszej kolejności wyliczono odpowiednie współczynniki korelacji liniowej Pearsona ( $r$ ) pomiędzy zawartościami analizowanych pierwiastków w glebie (tab. 3).

Tabela 2. Graniczne zawartości metali ciężkich w różnych grupach gleb w warstwie o grubości 0+20 cm [21]

Metal	Grupa gleb	Zawartość pierwiastka w glebie, mg/kg					
		0	I	II	III	IV	V
Ołów	a	30	70	100	500	2500	>2500
	b	50	100	250	1000	5000	>5000
	c	70	200	500	2000	7000	>7000
Miedź	a	15	30	50	80	300	>300
	b	25	50	80	100	500	>500
	c	40	70	100	150	750	>750
Nikiel	a	10	30	50	100	400	>400
	b	25	50	75	150	600	>600
	c	50	75	100	300	1000	>1000
Kadm	a	0,3	1,0	2	3	5	>5
	b	0,5	1,5	3	5	10	>10
	c	1,0	3,0	5	10	20	>20
Cynk	a	50	100	300	700	3000	>3000
	b	70	200	500	1500	5000	>5000
	c	100	300	1000	3000	8000	>8000

0 – zawartość naturalna, I – zawartość podwyższona, II – słabe zanieczyszczenie, III – średnie zanieczyszczenie, IV – silne zanieczyszczenie, V – bardzo silne zanieczyszczenie



Rys. 4. Wykresy ramkowe rozkładu zawartości pierwiastków śladowych w glebach Warszawy i okolic

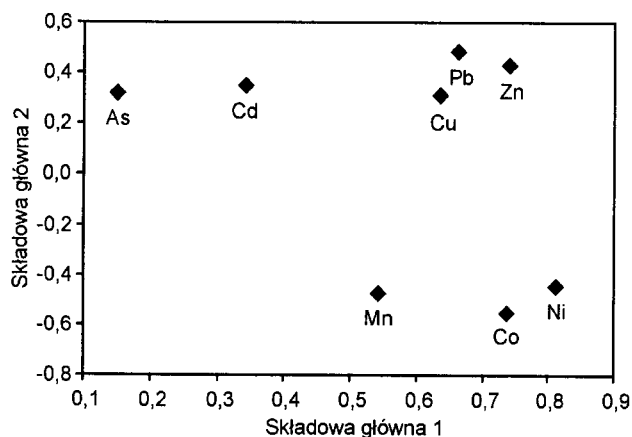
Otrzymane rezultaty wykazały silną korelację między zawartością kobaltu i niklu ( $r=0,85(0,80)$ ) oraz cynku i ołowiu ( $r=0,68(0,73)$ ) (liczby w nawiasach oznaczają wartości współczynnika korelacji, obliczone dla danych mniejszych od odpowiednich kwantyli 95%). Wyraźna korelacja wystąpiła między zawartością miedzi i cynku ( $r=0,51(0,77)$ ), manganu i kobaltu ( $r=0,48(0,53)$ ), manganu i niklu ( $r=0,45(0,55)$ ) oraz ołowiu i miedzi ( $r=0,44(0,71)$ ), natomiast mniejsza dla cynku i niklu ( $r=0,38(0,51)$ ), niklu i miedzi ( $r=0,34(0,62)$ ) oraz ołowiu i miedzi ( $r=0,30(0,35)$ ). Jak można było się spodziewać, odrzucenie pomiarów nietypowych – większych od kwantyli 95% – zwiększyło na ogół wartości współczynników korelacji. Efekt ten był szczególnie wyraźny w wypadku korelacji zawartości arsenu z pozostałymi pierwiastkami (z wyjątkiem kadmu). Wynikało to z faktu, iż w wypadku arsenu stwierdzono wiele wyraźnie odbiegających wartości, co zaobserwowano na odpowiednich wykresach rozrzutu. Istotny wzrost współczynnika korelacji, po odrzuceniu wartości odbiegających, można zaobserwować również w wypadku miedzi.

Obliczenia współczynników korelacji liniowej Pearsona przeprowadzono w typowy sposób, tzn. do obliczeń wzięto wyniki pomiarów zawartości badanej pary pierwiastków, w kolejnych punktach pomiarowych. Warto zwrócić uwagę na fakt, że jeżeli będą występowały bardzo wysokie, lecz odseparowane od siebie o nawet niewielką odległość zawartości dwóch metali, to współczynnik korelacji liniowej nie uwzględni tych pomiarów. Z tego względu, współczynnik korelacji liniowej Pearsona nie jest dobrą miarą korelacji zmiennych regionalizowanych.

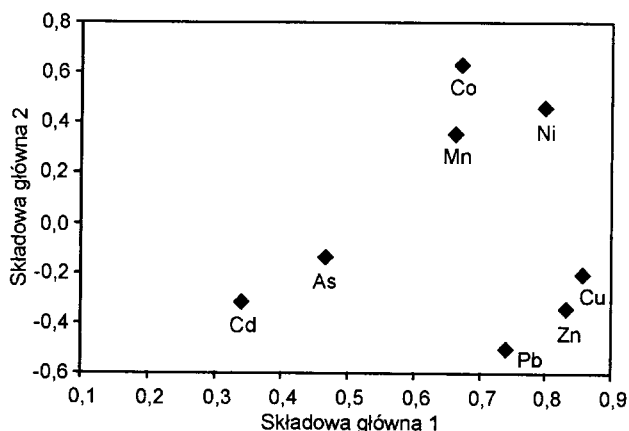
Tabela 3. Współczynniki korelacji liniowej Pearsona ( $r$ ) pomiędzy zawartościami metali ciężkich w glebie (współczynniki korelacji są istotne na poziomie istotności  $p=0,05$ )

Współczynniki korelacji dla wszystkich danych								Współczynniki korelacji dla danych mniejszych od kwantyli 95%							
Cd	0,26							Cd	0,03						
Co	0,01	0,07						Co	0,17	0,13					
Cu	0,09	0,16	0,25					Cu	0,31	0,24	0,45				
Mn	0,04	0,12	0,48	0,14				Mn	0,33	0,13	0,53	0,41			
Ni	0,03	0,16	0,85	0,34	0,45			Ni	0,22	0,17	0,80	0,62	0,55		
Pb	0,10	0,18	0,19	0,44	0,14	0,30		Pb	0,34	0,29	0,21	0,71	0,31	0,35	
Zn	0,06	0,24	0,27	0,51	0,19	0,38	0,68	Zn	0,33	0,27	0,33	0,77	0,40	0,51	0,73
	As	Cd	Co	Cu	Mn	Ni	Pb		As	Cd	Co	Cu	Mn	Ni	Pb

W celu wychwycenia dodatkowych związków pomiędzy zawartościami badanych pierwiastkami wykonano również wykres rozrzutu analizowanych danych w układzie dwóch pierwszych głównych składowych (rys. 5 i 6). Zastosowanie metody składowych głównych pozwoliło na zredukowanie liczby zmiennych wyjaśniających (przyczyn), poprzez utworzenie nowych zmiennych, będących liniowymi kombinacjami zmiennych silnie skorelowanych (w tym celu wykorzystywana jest macierz współczynników korelacji – tab. 3). Takie zmienne, zwane głównymi składowymi, mogą objaśnić zróżnicowanie pomiędzy analizowanymi pierwiastkami prawie tak dobrze, jak cały zbiór zmiennych wyjaśniających [23].



Rys. 5. Wykres rozproszenia badanych pierwiastków śladowych w układzie dwóch pierwszych składowych głównych, wykonany na podstawie wszystkich pomiarów (pierwsza składowa główna wyjaśnia 35,7% zmienności, druga składowa główna wyjaśnia 15,8% zmienności)

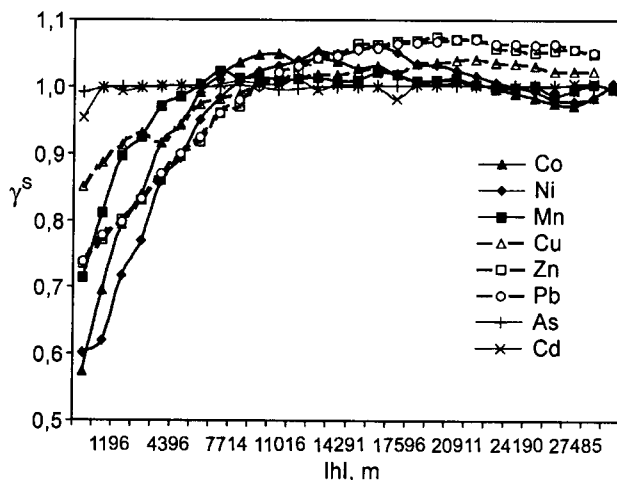


Rys. 6. Wykres rozproszenia badanych pierwiastków śladowych w układzie dwóch pierwszych składowych głównych, wykonany na podstawie pomiarów mniejszych od kwantyli 95% (pierwsza składowa główna wyjaśnia 48,1% zmienności, druga składowa główna wyjaśnia 15,8%)

Metoda składowych głównych jest często stosowana w celu graficznego przedstawienia na płaszczyźnie zależności pomiędzy wieloma zmiennymi. Można ją zastosować np. do analizy pomiarów zanikczyszczenia gleby [24]. Wykres przedstawiony na rysunku 5 jest wykonany na podstawie wszystkich pomiarów, zaś wykres przedstawiony na rysunku 6 wykonany został na podstawie pomiarów mniejszych od kwantyli 95%.

Jak widać na rysunkach 5 i 6, w zbiorze badanych pierwiastków można wyróżnić wyraźne trzy podzbiory: pierwszy zawiera arsen i kadm, drugi kobalt i nikiel, zaś trzeci miedź, ołów oraz cynk. Taki podział badanych pierwiastków jest zgodny z ich właściwościami fizyczno-chemicznymi oraz z danymi literaturowymi na temat ich występowania [21]. Nieco trudności mogłoby sprawić zakwalifikowanie manganu, który nie jest pierwiastkiem śladowym należącym do którejś z wyżej wymienionych grup. Po odrzuceniu wartości nietypowych widać, iż mangan leży na wykresie najbliżej kobaltu i niklu.

W celu ustalenia stopnia korelacji przestrzennej analizowanych pierwiastków na badanym obszarze zastosowano omówione powyżej metody geostatystyczne. Analizę geostatystyczną rozpoczęto od obliczenia wariogramów izotropowych (niezależnych od kierunku) zawartości badanych pierwiastków w glebie. Wariogramy izotropowe można w pierwszym przybliżeniu traktować jako średnią z wariogramów kierunkowych. Obliczenia wariogramów, ze względu na nieregularną sieć pomiarową, wymagały właściwego wyboru odpowiednich parametrów (i ich tolerancji) związanych z wektorem  $h$ . W wypadku analizowanych danych najbardziej regularne i stabilne wariogramy otrzymano stosując krok wektora  $h$ ,



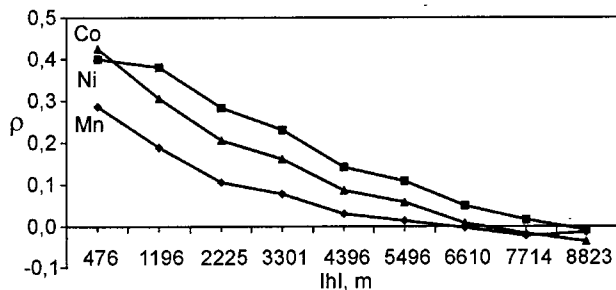
Rys. 7. Standardyzowane wariogramy izotropowe zawartości badanych pierwiastków śladowych w glebach Warszawy i okolic

równy 1100 m (średnia odległość pomiędzy próbkami) z tolerancją  $\Delta h$  tego kroku, równą 550 m (połowa średniej odległości pomiędzy punktami). W celu łatwiejszego porównania wariogramów standaryzowano je, dzieląc przez odpowiednie wariancje. Otrzymane standaryzowane wariogramy izotropowe przedstawiono na rysunku 7.

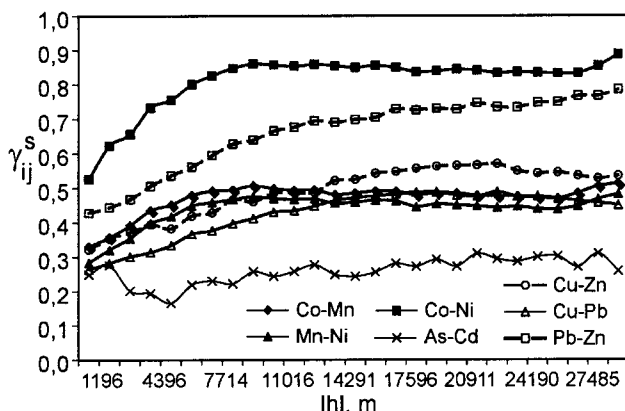
Pomimo znacznego zróżnicowania przestrzennego rozkładu zanieczyszczeń w badanym obszarze, wariogramy izotropowe zawartości większości badanych pierwiastków (z wyjątkiem arsenu i kadmu) wykazały dużą regularność, z wyraźnie zaznaczonym składnikiem losowym, zasięgiem oddziaływania i z wariancją progową. Składnik losowy związany był ze zmiennością w małej skali, tzn. występującą w odległościach mniejszych niż krok wektora  $h$ , stosowaną przy obliczaniu. Drugą przyczyną występowania składnika losowego były błędy pomiarowe. Regularny przebieg otrzymanych wariogramów izotropowych wynikał z malejących korelacji przestrzennych zanieczyszczenia tymi pierwiastkami wraz ze wzrostem odległości.

Dokładniejsza analiza wariogramów przedstawionych na rysunku 7 wykazała dużą zgodność z wykresami rozproszenia badanych pierwiastków w układzie dwóch pierwszych składowych. Wariogramy te można podzielić na wyraźne trzy grupy. Wariogramy zawartości arsenu i kadmu były typowe w sytuacji braku ciągłości przestrzennej (w rozpatrywanej skali), gdzie dominującą rolę odegrał efekt samorodków. Bardzo regularny i podobny do siebie przebieg miały wariogramy zawartości kobaltu i niklu. Nieco większa różnica wystąpiła w wypadku manganu, lecz charakter zmienności przestrzennej tego pierwiastka wykazał jego przynależność do tej samej grupy. Wariogramy zawartości ołowiu i cynku praktycznie pokrywały się. Podobny wariogram, chociaż nieco przesunięty w stosunku do wariogramów tych pierwiastków, otrzymano dla zawartości miedzi.

Dla porównania, na rysunku 8 przedstawiono autokorelogramy pierwiastków z drugiej grupy, tzn. kobaltu, niklu



Rys. 8. Autokorelogramy izotropowe zawartości manganu, niklu i kobaltu w glebach Warszawy i okolic

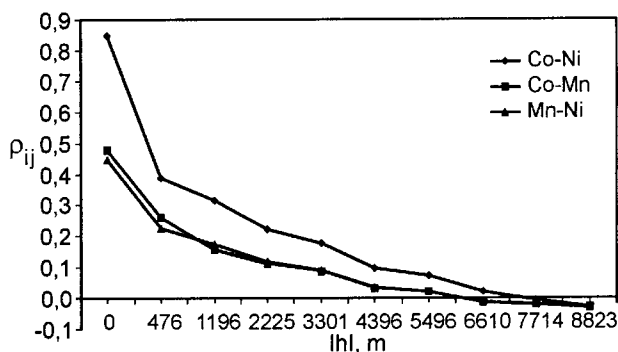


Rys. 9. Standaryzowane izotropowe wariogramy wzajemne zawartości badanych pierwiastków śladowych w glebach Warszawy i okolic

i manganu. Wykresy autokorelogramów dostarczyły podobnych informacji jak wykresy wariogramów.

Rysunek 9 przedstawia wariogramy wzajemne obliczone dla par pierwiastków, dla których współczynniki korelacji liniowej Pearsona (tab. 3) były istotnie duże ( $\rho > 0,4$  dla wszystkich danych). Dodatkowo obliczono i zamieszczono na wykresie wariogramy dla arsenu i kadmu.

Następnym etapem analizy było obliczenie standaryzowanych wariogramów wzajemnych zawartości w glebie tych par pierwiastków, dla których współczynnik korelacji (tab. 3) przekroczył wartość  $\rho = 0,4$ . Dla porównania pokazano również wykres standaryzowanego korelogramu wzajemnego arsenu i kadmu. Widać, iż największe wartości, podobnie jak współczynnik korelacji, osiągnął standaryzowany wariogram wzajemny zawartości kobaltu i niklu. Piłokształtny i w przybliżeniu stały standaryzowany wariogram wzajemny zawartości arsenu i kadmu wykazał brak zależności przestrzennych między tymi pierwiastkami. Dodatkowo, w celu porównania z wariogramami wzajemnymi, wyliczono również korelogramy wzajemne zawartości manganu, kobaltu i niklu (rys. 10).



Rys. 10. Izotropowe korelogramy wzajemne zawartości manganu, niklu i kobaltu w glebach Warszawy i okolic

Przedstawione na rysunkach 7 i 9 autowariogramy oraz wariogramy wzajemne były następnie modelowane z wykorzystaniem pakietu geostatystycznego Variowin 2.2. Wyniki modelowania przedstawiono w tabeli 4. Parametr IGF (*Indicative Goodness of Fit*), znajdujący się w ostatniej kolumnie, informuje o dokładności dopasowania przebiegu modelu

Tabela 4. Wyniki modelowania autowariogramów izotropowych oraz wzajemnych wariogramów izotropowych zawartości metali ciężkich

Pierwiastek	Efekt samorodków (mg/kg) <sup>2</sup>	Model zmienności długozasięgowej	Zasięg oddziaływania m	Wariancja progowa (mg/kg) <sup>2</sup>	Parametr dopasowania (IGF)
Mn	24661	sferyczny	4806	9746,33	$6,06 \cdot 10^{-3}$
Co	1,057	sferyczny	9770	0,58	$2,25 \cdot 10^{-3}$
Ni	8,825	sferyczny	7879	9,79	$3,91 \cdot 10^{-3}$
Mn-Ni	192,343	sferyczny	8446	173,32	$1,28 \cdot 10^{-3}$
Co-Mn	71,168	sferyczny	9960	43,74	$5,05 \cdot 10^{-3}$
Co-Ni	2,411	sferyczny	8538	2,17	$2,02 \cdot 10^{-3}$
Cu	441,54	sferyczny	8815	225,72	$4,70 \cdot 10^{-3}$
Pb	467,94	sferyczny	12560	555,21	$1,31 \cdot 10^{-3}$
Zn	3358,82	sferyczny	13535	5272,45	$2,22 \cdot 10^{-3}$
Pb-Zn	826,50	sferyczny	1343	1213,52	$2,41 \cdot 10^{-3}$
Cu-Zn	546,63	sferyczny	12523	708,22	$2,43 \cdot 10^{-3}$
Cu-Pb	161,77	sferyczny	15507	225,22	$1,43 \cdot 10^{-3}$

teoretycznego do wariogramu eksperymentalnego. Dokładną definicję IGF podano w pracy [20]. Jak wynika z tabeli 4, dobre dopasowanie eksperymentalnych autowariogramów i wariogramów wzajemnych otrzymano stosując sumę efektu samorodków i modelu sferycznego. Wartości efektu samorodków, zasięgu oddziaływania i wariancji progowej podano w tabeli 4. Następnie w oparciu o wzór (22) wyliczono zależności współczynnika kodyspersji od odległości (funkcję kodyspersji), korzystając z autowariogramów wzajemnych oraz wariogramów wzajemnych przedstawionych na rysunkach 7 i 9. Jak widać, wartości funkcji kodyspersji dla małych wartości wektora modułu wektora  $h$  były na ogół bardzo zbliżone do odpowiednich wartości współczynników korelacji liniowej przedstawionych w tabeli 3.

Należy pamiętać, że wartości autowariogramów, wariogramów wzajemnych, a zatem i funkcji kodyspersji dla małych wartości  $lhl$ , obliczono na podstawie znacznie mniejszej liczby par danych, niż dla dużych wartości  $lhl$  i dlatego mogą być obciążone dużym błędem. W celu sprawdzenia wartości funkcji kodyspersji dla  $lhl=0$  należy przeprowadzić modelowanie odpowiednich wariogramów.

Z rysunku 10 widać, że największe wartości funkcja kodyspersji przybrała dla kobaltu i niklu. Dla  $lhl=0$  jej wartość wynosiła  $C_{NiCo}(0)=0,8$ . Następnie, aż do około 1100 m, nastąpił lekki wzrost funkcji kodyspersji, w wyniku czego  $C_{NiCo}(1100\text{ m})=0,83$ . Dla  $lhl\approx 22000$  m funkcja kodyspersji stopniowo nasyciła się. Tego typu zachowanie można wytłumaczyć występowaniem korelacji przestrzennych w dużej skali. Powyżej 22000 m ponownie nastąpił gwałtowniejszy wzrost funkcji kodyspersji, aż do wartości  $C_{NiCo}(25000\text{ m})=0,85$ , co można prawdopodobnie wytłumaczyć ujawnieniem się trendu.

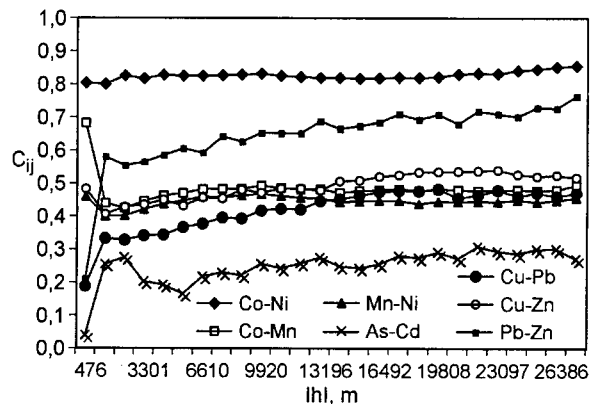
Wykresy funkcji kodyspersji dla ołowiu i cynku oraz ołowiu i miedzi różniły się nieco od wykresu funkcji kodyspersji dla kobaltu i niklu. Różnica polegała głównie na nieco szybszym i bardziej systematycznym wzroście tych funkcji. W celu lepszego zrozumienia charakteru funkcji kodyspersji przedstawionych na rysunku 11, w oparciu o wzór (22) wykonano obliczenia funkcji kodyspersji. W obliczeniach, zgodnie z wynikami przedstawionymi w tabeli 4, przyjęto, że autowariogramy i wariogramy wzajemne występujące we wzorze (22) są sumą efektu modelu samorodków oraz modelu sferycznego. Obliczone w ten sposób funkcje kodyspersji (zwane dalej teoretycznymi) porównano na rysunkach 12 i 13 z odpowiednimi eksperymentalnymi funkcjami kodyspersji. Jak można zauważyć, w większości wypadków uzyskano rozsądną zgodność, co świadczyło zarówno o właściwym dobraniu modelu jak i parametrów wariogramów występujących we wzorze (22).

Wartości teoretycznych funkcji kodyspersji, ekstrapolowane do  $lhl=0$ , przedstawiono w lewej części tabeli 5, zaś w prawej części tej tabeli przedstawiono wartości teoretycznych funkcji kodyspersji dla  $h$  równego zakresowi odpowiednich wariogramów sferycznych (a).

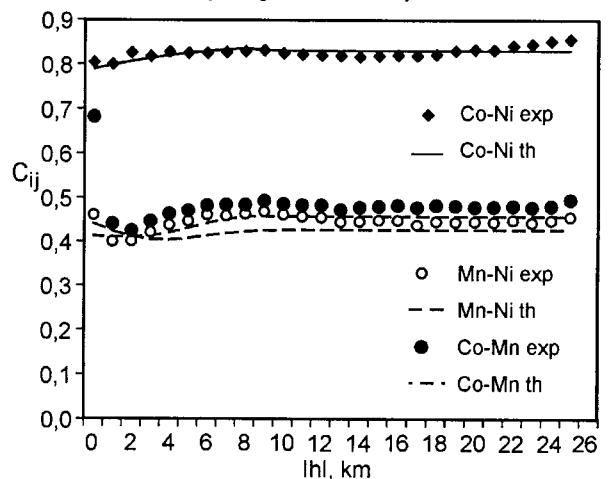
Tabela 5. Wartości funkcji kodyspersji dla wybranych par metali ciężkich w glebie, otrzymane na podstawie modelowania wariogramów (9)

$C_{ij}(lhl=0)$				$C_{ij}(lhl=a)$				
Mn	0,44			Mn	0,48			
Ni	0,80		0,41	Ni	0,83		0,46	
Pb		0,36		Pb		0,46		
Zn		0,49	0,65	Zn		0,52	0,69	
	Co	Cu	Mn	Pb	Co	Cu	Mn	Pb

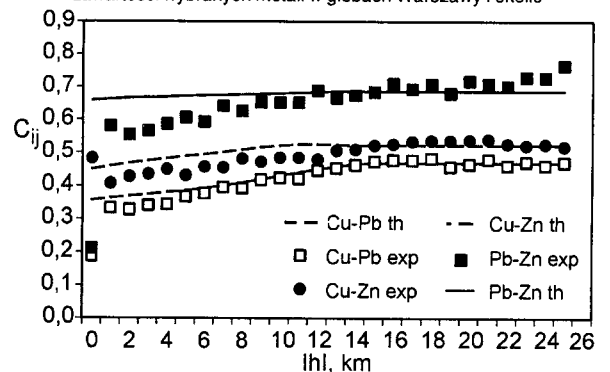
Wartości w lewej części tabelki otrzymano uwzględniając jedynie efekt samorodka dla  $lhl=0$ , natomiast wartości w prawej części tabelki otrzymano uwzględniając efekt samorodka oraz model sferyczny dla  $lhl$  równego zasięgowi oddziaływania modelu sferycznego



Rys. 11. Wykresy funkcji kodyspersji zawartości pierwiastków śladowych w glebach Warszawy i okolic



Rys. 12. Wykresy teoretycznych i eksperymentalnych funkcji kodyspersji zawartości wybranych metali w glebach Warszawy i okolic



Rys. 13. Wykresy teoretycznych i eksperymentalnych funkcji kodyspersji zawartości wybranych metali w glebach Warszawy i okolic

Porównując wartości teoretycznych funkcji kodyspersji z odpowiednimi wartościami współczynników korelacji liniowej (tab. 3), można stwierdzić na ogół dużą zgodność otrzymanych wyników. Widoczny wzrost wartości funkcji kodyspersji po prawej stronie tabeli można wytłumaczyć udziałem wynikającym z istnienia korelacji o długim zasięgu.

## Podsumowanie

Na podstawie wykresów rozrzutu (w układzie składowych głównych) rozkładów zawartości badanych pierwiastków śladowych w glebie Warszawy i okolic wykazano, że można je zgrupować w trzy podzbiory: pierwszy zawiera arsen i kadm, drugi kobalt, nikiel i mangan, natomiast trzeci miedź, ołów i cynk. Zawartość badanych pierwiastków w glebie zależała



również od położenia, a zatem należy traktować ją jako zmienną zregionalizowaną, a nie zmienną losową. Ponieważ klasyczne metody statystyczne nie wystarczają do zbadania ciągłości przestrzennej zanieczyszczenia gleb Warszawy i okolic, dlatego do zbadania tej ciągłości zastosowano metody geostatystyczne. Charakterystyka ciągłości przestrzennej zawartości cynku, ołowiu i żelaza w glebie została przedstawiona w postaci standaryzowanych wariogramów izotropowych oraz autokorelogramów. Obliczenia i wyniki modelowania wariogramów wykazały, że rodzaj ciągłości przestrzennej zawartości większości pierwiastków (z wyjątkiem arsenu i kadmu) był bardzo podobny i może być dobrze opisany sumą efektu modelu samorodków i modelu sferycznego. W wypadku arsenu i kadmu nie zaobserwowano istotnej ciągłości przestrzennej.

Badając korelację pomiędzy zawartością analizowanych pierwiastków stwierdzono, że najczęściej stosowany współczynnik korelacji liniowej Pearsona nie uwzględnia korelacji przestrzennych, a zatem nie jest dobrą miarą do obliczania korelacji związanych z rozkładami przestrzennymi w badaniach środowiska. Otrzymane rezultaty wykazały silną korelację pomiędzy zawartością kobaltu i niklu oraz cynku i ołowiu. Wyraźna korelacja istniała pomiędzy zawartością miedzi i cynku, manganu i kobaltu, manganu i niklu oraz ołowiu i miedzi, słabsza zaś dla cynku i niklu, niklu i miedzi oraz ołowiu i niklu.

Aby ocenić korelacje przestrzenne pomiędzy badanymi zanieczyszczeniami obliczono izotropowe standaryzowane wariogramy wzajemne pomiędzy stężeniami tych pierwiastków, dla których znaleziono wyraźne korelacje. Dla porównania wykonano również obliczenia wybranych korelogramów wzajemnych. Zaprezentowano także wyniki modelowania otrzymanych wariogramów oraz wzajemnych wariogramów eksperymentalnych. Przebiegi prawie wszystkich autowariogramów i wariogramów wzajemnych można było aproksymować modelem złożonym z dwóch składowych, tj. modelu samorodka oraz modelu sferycznego. Wyjątkiem były autowariogramy i wariogramy wzajemne zawartości arsenu lub kadmu. W tym wypadku nie zaobserwowano zmienności o długim zasięgu.

W oparciu o liniowy model koregionalizacji wyliczono zależności współczynnika kodyspersji od odległości (funkcji kodyspersji). Współczynnik kodyspersji, będąc w geostatystyce analogiczną miarą korelacji do współczynnika korelacji liniowej Pearsona, w statystyce tradycyjnej uwzględnia indywidualne udziały odpowiednich autowariogramów oraz wariogramów wzajemnych. Przedyskutowano otrzymane wyniki obliczeń funkcji kodyspersji oraz wykonano modelowanie tych funkcji, otrzymując dobre dopasowanie, co świadczyło o właściwym dobraniu modelu oraz parametrów autowariogramów i wariogramów wzajemnych.

*Autor wyraża serdeczne podziękowania Panu dr. Józefowi Lisowi – kierownikowi Pracowni Chemii Środowiska w Państwowym Instytucie Geologicznym – za udostępnienie danych dotyczących zanieczyszczenia gleb Warszawy i okolic.*

## LITERATURA

1. M. SOBCZYK: Statystyka. PWN, Warszawa 1996.
2. M. FISZ: Rachunek prawdopodobieństwa i statystyka matematyczna. PWN, Warszawa 1967.
3. D. G. KRIGE: A statistical approach to some mine valuations and allied problems at the Witwatersrand. Masters thesis, University of Witwatersrand, South Africa, 1951.
4. G. MATHERON: *Traité de Géostatistique Appliquée*. Technip, Paris 1962–1963.
5. D. G. KRIGE: Two-dimensional weighted moving average trend surfaces for ore valuation. *Proceedings of the Symposium on Mathematics, Statistics and Computer Applications in Ore Valuation*, Johannesburg. South African Institute of Mining and Metallurgy, Johannesburg 1966.
6. S. ROUHANI; R. SRIVASTAVA, A. DESBARATS, M. CROMER, A. JOHNSON: *Geostatistics for Environmental and Geotechnical Applications*. ASTM STP 1283, 1996.
7. E. H. ISAAKS, R. M. SRIVASTAVA: *An Introduction to Applied Geostatistics*. Oxford University Press, New York 1989.
8. M. DAVID: *Handbook of Applied Advanced Geostatistical Ore Reserve Estimation*. Elsevier, Amsterdam 1988.
9. A. G. JOURNEL, C. J. HUIBREGTS: *Mining Geostatistics*. Academic Press, London 1978.
10. A. G. ROYLE: *A Practical Introduction to Geostatistics*. Mining Sciences Department, University of Leeds, Leeds 1975.
11. B. D. RIPLEY: *Spatial Statistics*. Wiley, New York 1981.
12. A. G. JOURNEL: *Fundamentals of Geostatistics in Five Lessons*. Washington DC 1989.
13. Gamma Design Software, *GS+ Geostatistics for the Environmental Software*, 1998.
14. J. A. CZUBEK, J. ŁASKIEWICZ, J. GYURSAK, A. LENDA, K. UMIASTOWSKI, T. ZORSKI: Charakterystyki geostatystyczne niektórych serii litostratygraficznych monokliny przedsudeckiej. *Mat. VI symp. „Zastosowanie metod matematycznych i informatycznych w geologii”*, AGH, Kraków 1977.
15. B. NAMYSIOWSKA-WILCZYŃSKA, A. WILCZYŃSKI: Badania geostatystyczne zanieczyszczenia gleb metalami ciężkimi w wybranych rejonach Górnego Śląska. *Ochrona Środowiska*, 1997, nr 2(65), ss. 9–18.
16. B. NAMYSIOWSKA-WILCZYŃSKA, A. WILCZYŃSKI: Multivariate estimation and simulation for environmental data modelling: processing of heavy metals concentration data in soil. *Data Science Journal*, 2002, 1(1), pp. 27–44.
17. J. LIS: *Atlas geochemiczny Warszawy i okolic*. Państwowy Instytut Geologiczny, Warszawa 1992.
18. E. H. ISAAKS, R. M. SRIVASTAVA: Spatial continuity measures for probabilistic and deterministic geostatistics. *Mathematical Geology*, 1988, Vol. 20, No. 4.
19. C. V. DEUTSCH, A. G. JOURNEL: *GSLIB Geostatistical Software Library and Users Guide*. Oxford University Press, New York 1992.
20. Y. PANNATIER: *Variowin, Software for Spatial Data Analysis in 2D*. Springer-Verlag, 1996.
21. A. KABATA-PENDIAS, T. MOTOWICKA-TERELAK, M. PIOTROWSKA, H. TERELAK, T. WITEK: Ocena stopnia zanieczyszczenia gleb metalami ciężkimi i siarką. *Ramowe wytyczne dla rolnictwa*. IUNG, Seria P(53), Puławy 1993.
22. J. R. DOJLIDO: *Chemia wód powierzchniowych*. Wydawnictwo Ekonomia i Środowisko, Białystok 1995.
23. T. KASPRZAK, M. LASEK, M. PĘCZKOWSKI: *Statystyczna analiza wielowymiarowa za pomocą komputera*. Wydawnictwa Uniwersytetu Warszawskiego, Warszawa 1997.
24. H. WACKERNAGEL: *Principal Component Analysis for Autocorrelated Data. A geostatistical perspective*. Technical Report N-22/98/G, Centre de Géostatique – Ecole des Mines de Paris, Paris 1998.

**Zawadzki, J. Spatial Correlation Studies of Trace Element Concentrations in the Soils of Warsaw and Immediate Vicinity. *Ochrona Środowiska* 2002, Vol. 24, No. 4, pp. 17–26.**

**Abstract:** Empirical measures of spatial correlation and variability were calculated and discussed using examples based on a data set which included soil concentrations of As, Cd, Co, Mn, Ni, Pb and Zn measured in Warsaw and in the neighbouring area. Chemical analyses were carried out with soil samples collected at the depth of 0 to 0.20 m. Next the descriptive statistical parameters were calculated and thereafter Principal Component Analysis was carried out. The results showed that the trace elements under study can be grouped into the following three classes: As nad Cd; Co, Mn and Ni; Pb and Zn. Conventional geostatistical measures of spatial continuity – auto-variograms and cross-variograms – were calculated and discussed in detail. For comparison, calculations were also carried out for auto-correlograms and cross-correlograms. The computations were followed by accurate modelling of the experimental auto-variograms and cross-variograms. It was found that for all the trace elements under study (except As and Cd) the best "nested

structure" was a combination of the nugget effect and spherical model. For As and Cd, nugget-like variograms were observed. The parameters of the variograms and cross-variograms describing spatial continuity and spatial correlation were also determined. A significant classical correlation described by the Pearson coefficient and a significantly high spatial correlation described by the cross-variograms were obtained for the concentrations of the following pairs of elements: Co–Ni, Pb–Zn, Cu–Zn, Mn–Ni, Cu–Pb, Ni–Zn, Ni–Cu and Ni–Pb. To gain a better understanding of the correlations observed, use was made of the linear model of coregionalization. For this purpose, the spatial dependence of the coregionalization coefficient, which is a geostatistical equivalent of the Pearson coefficient, provides information on both classical and spatial correlations. This is why it can take values greater than 1 and reach a maximum for sampling distances greater than zero ("deferred correlation"). Such behavioral pattern was detected by modelling via "nested structures", which are combinations of the nugget effect and spherical model for the probable sets of component variogram parameters.