

ZASTOSOWANIE ANALIZY STATYSTYCZNEJ W PODZIALE NA JEDNOSTKI INŻYNIERSKO-GEOLOGICZNE

UKD 624.131:311.17:519.2

Celem prognozowania statystyczno-przestrzennego jest wnioskowanie o własnościach określonej części środowiska geologicznego i jej zmienności na podstawie odpowiedniej ilości obserwacji w celu wydzielenia jednorodnych, np.: inżyniersko-geologicznych jednostek przestrzennych (2).

Historia zastosowania analizy statystycznej w naukach o Ziemi sięga końca XIX w. Badania dotyczące zastosowania tej analizy zapoczątkowali uczeni amerykańscy, francuscy, niemieccy, angielscy, skandynawscy i od lat dwudziestych bieżącego stulecia uczeni radzieccy. Wymienić tu należy przede wszystkim prace: K. Pearsona, K. Schmida, F. J. Laevinson-Lessinga, G. Simpsona, C. Wentwortha, W. Krumbeina, J. S. Kahna, R. L. Millera, I. Szarapowa i in. Początkowo dotyczyły one paleontologii paleobiometrii, petrologii, geochemii, stratygrafii, a następnie sedimentologii. Okres po drugiej wojnie światowej do chwili obecnej wykazał znaczną ekspansję analizy statystycznej we wszystkie, praktycznie biorąc, dziedziny geologii. Zarówno w świecie, jak i w ostatnim dziesięcioleciu w Polsce obserwuje się pojawienie wielu prac poświęconych całkowicie lub częściowo omówieniu możliwości przystosowania analizy statystycznej do różnych działów geologii. Podkreślić tu należy, iż prace te nie obejmowały dotychczas analizy zastosowania aparatu statystycznego w całości badań kompleksowych, a dotyczyły jedynie pewnych wybranych ocen. Celem niniejszego artykułu jest próba przedstawienia ogólnego przeglądu wybranych metod analizy statystycznej w zastosowaniu do prawidłowego wydzielenia jednorodnych pod względem inżyniersko-geologicznym jednostek.

Pierwszym problemem, jaki należy tu rozwiązać jest rozwinięcie i opracowanie modelu geologicznego, następnie inżyniersko-geologicznego, a na jego tle zdefiniowanie interesującej nas populacji oraz wybranie zmiennych, które mogą być mierzone, a także źródeł zmian, jakie powinny być wzięte pod uwagę.

Drugim problemem jest przetransponowanie modelu koncepcyjnego na model statystyczny, w którym matematyczna struktura obserwacji obejmuje explicite szereg źródeł wariancji.

Trzeci problem, to system pobierania prób: statystycznej dostosowanej do modelu statystycznego.

Czwarty i ostatni, to problem właściwych przybliżeń do populacji na podstawie próbek.

Przy omawianiu poszczególnych problemów od strony teoretycznej przedstawiono pewne sugestie stosowalności poszczególnych metod w praktycznej działalności inżyniersko-geologicznej na tle opracowanej przez zespół Katedry Geologii Inżynierskiej UW Mapy inżyniersko-geologicznej miasta Płocka.

Rozwinięcie i opracowanie koncepcyjnego modelu geologicznego poprzedzać powinny zawsze szczegółowe studia podstawowe danego interesującego obszaru, dotyczące: stratygrafii, geomorfologii, tektoniki, hydrologii i hydrogeologii oraz klimatu. Na podstawie materiałów wyjściowych i szczegółowej wizji lokalnej terenu geolog projektuje rozmieszczenie podstawowych punktów obserwacyjnych, które powinny objąć całą strefę przyszłego modelu. Wynikające z przesłanek merytorycznych określenie ilości i lokalizacji tych punktów przy zastosowaniu pew-

nych metod opartych na rachunku prawdopodobieństwa może znacznie ograniczyć nakłady ekonomiczne i czasowe na całość prac oraz wpłynąć na jakość opracowania. W. Krumbein i F. Graybill (3) podają pewne sugestie, dotyczące zarówno zagadnienia projektowania wierceń, jak i pobierania próbek. Wskazują oni na metodę randomizacji, jako na tę, dzięki której każda jednostka ma jednakową szansę znalezienia się w zasięgu obserwacji. Dla geologa inżynierskiego metoda randomizacji, czyli całkowitej losowości jest możliwa do zastosowania w tylko ograniczonym zakresie i może wkroczyć dopiero w dalszej fazie rozpoznania, kiedy to zostanie rozwiązany problem wstępnego podziału modelu na pewne obszary, wynikające już z ogólnie rozpoznanej budowy geologicznej terenu.

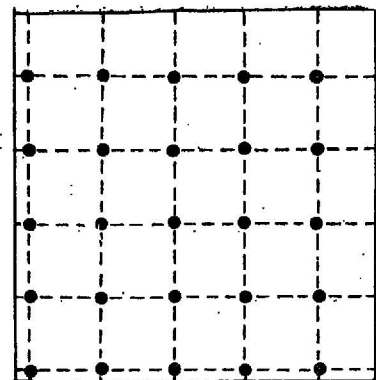
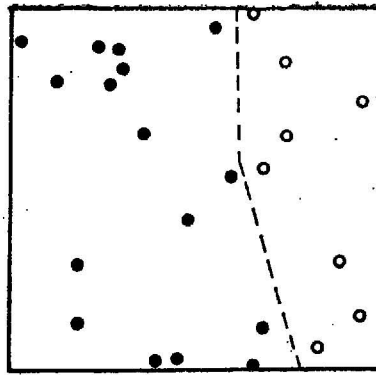
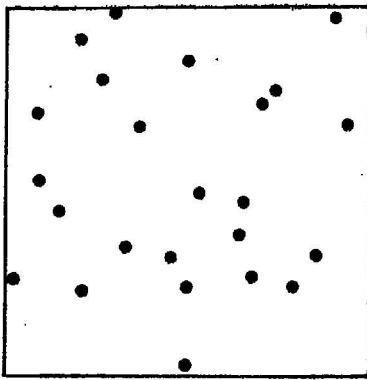
W początkowej fazie opracowywania Mapy inżyniersko-geologicznej miasta Płocka teren został podzielony na kilka części, takich, jak: obszar wysoczyzny i obszar doliny Wisły, w których obrębie wydzielono obszary mniejsze, np.: tarasy denudacyjne, równiny morenowe, równiny młodoplejstocenijskiej akumulacji wodno-łodowcowej, równiny akumulacji zastoiskowej, obszary akumulacji antropogenicznej. Projektowanie lokalizacji uzupełniających podstawowych punktów obserwacyjnych wynikało z ww. podziału. W celu potrzebnej dla danego zadania szczegółowości rozpoznania wydaje się możliwe stosowanie metody randomizacji dla oznaczenia granic litologicznych i wydzielenie jednostek litofacjalnych w przypadku, w którym przesłanki geomorfologiczne zawiodą. Korzystając z wykonanych już na każdym wydzielonym obszarze danych z podstawowych punktów obserwacyjnych geolog uzyskuje pewien pogląd na rozmiary skomplikowania układu jednostek litofacjalnych.

Zmienność tę, (np. dla układu litofacji) można wyrazić liczbowo w ogólnie stosowanego w statystyce matematycznej wzoru, określającego współczynnik zmienności jako stosunek odchylenia standardowego S do średniej y . Jeśli za zmienną y przyjmiemy oznaczenie ilościowe poszczególnych litofacji na pewną jednostkę miary, to miara ta da nam już wstępne sugestie dotyczące konieczności projektowania ilości podstawowych punktów obserwacyjnych, uzupełniających. Podobnie należy potraktować inne czynniki wpływające na stopień skomplikowania analizowanego modelu środowiska inżyniersko-geologicznego.

Dla przykładu można tu podać profil układu litofacji (tak poziomy, jak i pionowy) obszaru, wydzielonego na terenie Płocka, jako denudacyjne równiny morenowe, gdzie na podstawie wierceń wykazano następujące orientacyjne zróżnicowanie w serii gliny zwalowej szarej i brązowej:

grunty mało spoiste	51
„ średnio spoiste	247
„ spoiste ciężkie	163
„ bardzo spoiste	23
razem	484

odchylenie standardowe $S = 0,72$,
współczynnik zmienności $W = 0,31$, czyli 31%.



Ryc. 1.
Ryc. 2.
Ryc. 3.
Ryc. 4.

Jak widać z powyższych danych stopień skomplikowania, rzutujący na ilość potrzebnych do właściwego rozpoznania podstawowych punktów obserwacyjnych, jest wysoki. Ustawienie zależności funkcyjnej opartej na rachunku prawdopodobieństwa, a określającej potrzebną ilość podstawowych punktów obserwacyjnych dla danej (założonej przesłankami merytorycznymi) dokładności rozpoznania jest aktualnym przedmiotem badań i w przypadku rozpoznania inżyniersko-geologicznego zagadnieniem jeszcze nie rozwiązany. W. Krumbeln i F. Graybill proponują cztery wersje metody randomizacji.

Jedną z nich, która została określona jako metoda randomizacji prostej przedstawiona jest na ryc. 1. Lokalizację podstawowych punktów obserwacyjnych opracowuje się tu przy użyciu tablicy liczb losowych. Znając rozmiar powierzchni interesującego nas wydzielonego uprzednio obszaru i liczbę wierceń przypadającą na ten obszar dzieli się go na jednostki miary powierzchniowej, oznaczwszy je kolejnymi liczbami. Z tablicy liczb losowych, znając ogólną ilość jednostek powierzchni, wybiera się, arbitralnie selekcyjując miejsca na tablicy, ciąg tylu liczb ile jest wierceń, otrzymując w ten sposób ich lokalizację. Ryc. 1 przedstawia lokalizację. Ryc. 2 przedstawia lokalizację 24 wierceń, wybranych z pewnego obszaru, przeprowadzoną metodą randomizacji prostej.

Druga wersja (ryc. 2) stanowi pewną wariację wersji pierwszej. Jeśli na wydzielonym na podstawie przesłanek geomorfologicznych obszarze istnieje pewne predyspozycje dalszego podziału, wówczas hipotetycznie zakłada się granice, a wiercenia lokalizuje na obu obszarach niezależnie drogą losowania z tablic liczb losowych.

Trzecia wersja (ryc. 3) obrazuje lokalizację wg założonej siatki kwadratów. Czwarta (ryc. 4), określona jako gniazdowa, zakłada podział wydzielonego obszaru na pewne jednostki powierzchni, na których wiercenia lokalizujemy jak gdyby gniazdowo wg pewnego schematu, jaki przykładowo przedstawiono na ryc. 4.

Należy podkreślić, iż zdaniem autorki stosowanie metody randomizacji w zastosowaniu do badań inżyniersko-geologicznych, szczególnie dla rozwiązywania np. zagadnień litofacyjnych, ogranicza się wyłącznie do obszarów o minimalnym stopniu rozpoznania lub trudnych do powierzchniowego rozwiązania, gdzie nie ma predyspozycji bardziej merytorycznie uzasadnionego, np. kierunkowego usytuowania wierceń.

Uzyskane z podstawowych punktów obserwacyjnych informacje powinny już pozwolić na określenie bardziej uściślonego modelu geologicznego i profilu dla każdego z wydzielonych obszarów, z ogólnym przeglądem układu litofacji zarówno w profilu poziomym, jak i pionowym.

Następnym zagadnieniem do rozwiązania jeszcze w pierwszym problemie tej fazy badań jest zdefiniowanie pojęcia populacji i modelu środowiska inżyniersko-geologicznego. Model ten będzie wymagał

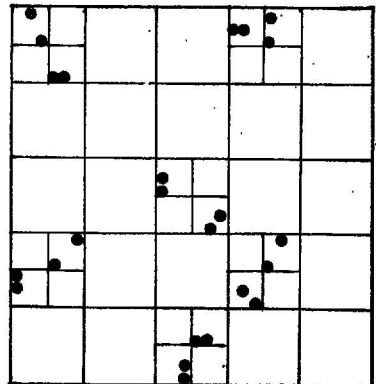


Fig. 1.
Fig. 2.
Fig. 3.
Fig. 4.

dokonania podziału terenu na jednostki przestrzenne różnego rzędu różniące się między sobą sumą warunków: geologicznych, hydrogeologicznych, geomorfologicznych, geodynamicznych, geomechanicznych, a więc inżyniersko-geologicznych i zagospodarowania.

Układ przestrzenny litofacji jest jednym z pierwszych wskaźników wpływających na pełną charakterystykę tego środowiska. W uzasadnionych przypadkach pozwala on na przyjęcie pewnej powierzchni odniesienia. Mogą ją stanowić serie gruntów spoiстых lub sypkich, bądź skał litych znacznych miąższości, występujące ciągią pokrywając na większości opracowywanego terenu. W przypadku opracowywanej mapy inżyniersko-geologicznej Płocka na powierzchnię odniesienia przyjęto — powierzchnię stropu gruntów spoiстых — lodowcowych. Mając na względzie wymodelowany układ przestrzenny litofacji w zależności od merytorycznych przesłanek badawczych, ustalamy parametry, które powinny być analizowane, aby dokonać podziału na jednostki inżyniersko-geologiczne. W celu ustalenia tych parametrów należy ocenić czynniki pierwszorzędne dla danego celu badania.

Za najważniejsze czynniki przy opracowywaniu Mapy inżyniersko-geologicznej miasta Płocka przyjęto: ukształtowanie terenu (spadki), dopuszczalne obciążenia jednostkowe gruntów na głęb. 1 m, głębokość występowania pierwszego zwierciadła wody gruntowej, występowanie procesów geodynamicznych. Wszystkie te czynniki wymagają określenia wielu wartości, z których wynikną oceny jakościowe; możemy je wyrazić w postaci liczbowej, charakteryzującej układy zależnościowe. W tym momencie pojawia się problem określenia właściwego modelu statystycznego oraz dostosowanego do tego modelu systemu pobierania próbek.

Omówienie zastosowania odpowiedniej metody badawczej aparatu statystycznego dla każdego z tych czynników można przedstawić na przykładzie analizy określania dopuszczalnych obciążeń jednostkowych, a ściślej mówiąc, analizy określania cech fi-

zyczno-mechanicznych, które na to określenie priorytatywnie wpływają, omawianie bowiem wszystkich czynników przekroczyłoby ramy tego artykułu.

Cechy fizyczno-mechaniczne stanowią sumę pewnych zmiennych, takich, jak: rodzaj gruntu, ciężar właściwy, objętościowy, stopień i wskaźnik plastyczności, stopień zagęszczenia, ściśliwość, spójność, kąt tarcia i in. Pomiedzy tymi zmiennymi obserwujemy większe lub mniejsze korelacje, ustalenie ich jest pierwszym zadaniem, które dzięki zastosowaniu metod analizy statystycznej może pozwolić na wyodrębnienie minimalnego zespołu cech, różnicujących poszczególne jednostki ze względu na ten czynnik. Wkracza tu więc problematyka korelacyjna i regresyjna, dotycząca oceny współzależności oraz problematyka prognozy, dotycząca oceny wartości jednej cechy na podstawie wartości drugiej cechy, oczywiście w obrębie jednej litofacji.

Weźmy przykładowo zależność między wilgotnością naturalną, a stopniem plastyczności piasków gliniastych, określonego wieku i genezy w serii gliny zwalowej Płocka. Na podstawie 40 przebadanych laboratoryjnie próbek ułożono tabelę korelacyjną obrazującą współzależność między tymi dwoma cechami. Tabela ta przedstawia szeregi rozdzielcze obu mierzonych cech (zmiennych) oraz ich rozkłady brzegowe jako podstawę wszystkich przeprowadzonych obliczeń.

TABELA KORELACYJNA (DANE WYJŚCIOWE)

$I_1 \backslash w$	10,5	13,5	16,5	19,5	22,5	frakcje I_1
0,95					1	1
0,85			1			1
0,75				1		1
0,65					2	2
0,55			1	2		3
0,45		3	8			11
0,35		6	4			10
0,25			2			2
0,15	1	1			1	3
0,05	4	2				6
frakcje w	5	12	16	3	4	40

I_1 i w — oznaczenia stopnia plastyczności i wilgotności naturalnej (w %) przyjęte przez Komitet Wykonawczy V Kongresu Międzynarodowego Mechaniki Gruntów i Fundamentowania w Paryżu w 1961 r.

Na podstawie powyższych danych obliczony współczynnik korelacji wynosi:

$$r_{wI_1} = \frac{S_{wI_1}}{S_w S_{I_1}} = 0,79$$

- r_{wI_1} — współczynnik korelacji między zmiennymi — wilgotnością naturalną a stopniem plastyczności,
 S_{wI_1} — kowariancja, jako średnia iloczynów odchyleń zmiennych: wilgotności naturalnej i stopnia plastyczności od ich średnich,
 S_w, S_{I_1} — odchylenia standardowe dla odpowiednich zmiennych.

Znak dodatni współczynnika korelacji wskazuje na zależność wprost proporcjonalną, a jego wartość reprezentuje duży stopień współzależności upoważniającej (z wystarczającą w tym przypadku dokładnością) do wyznaczenia równania prostej regresji. Równanie to uzyskano metodą najmniejszych kwadratów, stosowaną w statystyce matematycznej. Ma

ono w przypadku badanych zmiennych następującą postać:

$$I_1 = \bar{I}_1 + b_{I_1} w \quad (w - \bar{w}) = 0,0474 w - 0,395$$

- \bar{I}_1 — obliczony stopień plastyczności,
 \bar{I}_1 — średni stopień plastyczności w badanej próbie statystycznej,
 $b_{I_1 w}$ — współczynnik regresji, stanowiący miarę wzrostu I_1 , gdy wartość w wzrasta o jednostkę,
 w — wilgotność naturalna (dana badaniami laboratoryjnymi),
 \bar{w} — średnia wilgotność naturalna w badanej próbie statystycznej:

Ustalając w dalszej fazie kolejne możliwe korelacje można określić zależności funkcyjne mierzonych cech, składające się na określone wartości dopuszczalnego obciążenia na danej głębokości. Powstanie więc ogólny statystyczny model liniowy, stanowiący funkcję wielu zmiennych. Do modelu tego można dojść drogą analizy czynnikowej.

Otrzymany model matematyczny, stanowiący w tym przypadku funkcję ciągłą z określonymi granicami, ujmie w formie zależnościowej zmienne główne i wtórnie z nich wynikające. Pozwoli to w konsekwencji (przez analizę zmiennych głównych w poszczególnych podstawowych punktach obserwacyjnych na danej głębokości) ustalić izarytmy dopuszczalnych obciążeń jednostkowych. W ten sposób dochodzi się do mapy gruntów budowlanych na danej głębokości. Wydzielenie na całym opracowywanym terenie obszarów o jednakowych dopuszczalnych naprężeniach jednostkowych przy poprawnym przeprowadzaniu wszystkich omawianych dotychczas badań nie powinno nastęrczać większych trudności.

Całość obliczeń, które mają doprowadzić do wydzielenia jednorodnych (pod względem cech fizyczno-mechanicznych) jednostek wymaga również badań wstępnych związanych z określeniem rozkładu poszczególnych cech oraz określeniem wielkości próbki statystycznej, która może dostatecznie scharakteryzować populację ze względu na daną cechę.

Badacze amerykańscy, zakładając rozkład normalny jako najczęściej w przyrodzie spotykany, proponują zastosowanie następującego wzoru na określenie liczbowe wielkości próbki statystycznej:

$$n = \left[\frac{St_{\alpha/2}^{(m)}}{d} \right]^2$$

gdzie:

- S — odchylenie standardowe dla danej mierzonej cechy w próbie,
 $tu_{\alpha/2}$ — półprzedział ufności, którego procentową wartość odczytujemy z tabel prawdopodobieństwa wyprowadzonych dla rozkładu t -Studenta,
 m — stopnie swobody, jako parametr tego rozkładu uzależniony od ilości próbek, którą dysponowaliśmy przy obliczaniu wskaźników statystycznych,
 d — żądana dokładność, czyli liczbowe określenie możliwego błędu.

Jeśli $n - (m + 1)$ okaże się mniejsze lub równe zero, wówczas nie ma potrzeby powiększania próbki. Jeśli wartość ta okaże się większa od zera, w celu właściwego wnioskowania statystycznego należy powiększyć tak próbkę, by równanie to spełniło określone warunki.

Przystąpienie do analizy czynnikowej, jak również określenie wielkości próbki wymaga ustalenia rozkładu. Badania wielu autorów, oparte na wielkich próbkach jak i badania prowadzone w Katedrze Geologii Inżynierskiej UW wykazały, iż większość

parametrów, określających jednostki inżyniersko-geologiczne ma rozkład normalny lub log-normalny. Ustalenie rozkładu wymaga oczywiście dużej próbniki statystycznej, którą nie zawsze dysponujemy. Statystyka matematyczna daje do dyspozycji w takich przypadkach metody, za pomocą których można sprawdzić zakładany rodzaj rozkładu. Jeżeli kształt krzywej rozkładu zbliżony jest mniej lub bardziej do krzywej rozkładu normalnego, to odchylenia mogą wynikać np. z małej liczebności próbnki. Zastosowanie testu χ^2 na zgodność przy nie wyszczególnionych parametrach rozkładu może nam wyjaśnić tę kwestię.

Do oceny parametrów rozkładu przy tym teście możemy dojść drogą oceny punktowej parametrów na podstawie próbnki. Test χ^2 ma zastosowanie tylko wówczas, gdy ilość obserwacji w każdej klasie szeregu rozdzielczego jest większa od zakładanej w tym teście ilości. Różni autorzy różnie określają tę ilość. Jedni uważają, że powinna ona być większa od 5, inni, że od 10. Właściwie wydaje się przyjęcie którejś z tych wartości zależnie od tego jak dalece krzywa Gaussa odbiega od krzywej rozkładu w próbie.

Aby dojść do testowania należy w szeregu rozdzielczym danej cechy, którą określa się jako zmienną losową x lub y , przekształcić granice klas na zmienne losowe standaryzowane i obliczyć teoretyczne ilości obserwacji w danej klasie, gdyby rozkład miał charakter rozkładu normalnego. Do obliczenia teoretycznej ilości obserwacji w danej klasie dochodzi się przez obliczenie z tabel rozkładu normalnego prawdopodobieństw, czyli pola pod krzywą, jaką zajmowałaby zmienna o określonych granicach. Teoretyczne prawdopodobieństwa mnożymy przez ogólną liczbę obserwacji. Po otrzymaniu tych wartości oblicza się wartość funkcji testowej wg wzoru:

$$\chi^2 = \sum_{i=1}^c \frac{(f_i - F_i^*)^2}{F_i^*}$$

gdzie:

c — ilość klas szeregu rozdzielczego,
 f_i — liczebność poszczególnych klas w próbie,
 F_i^* — liczebności teoretyczne.

Rozkład tej funkcji zależy od ilości stopni swobody, która jest jedynym parametrem tego rozkładu. Ilość stopni swobody zależy od ilości oszacowanych parametrów z próbnki i potrzebnych do wyznaczenia teoretycznych liczebności. W przypadku testowania na zgodność z rozkładem normalnym ilość stopni swobody: $v = c - 1 - d$, gdzie: c — ilość klas, d — ilość oszacowanych parametrów rozkładu, która w tym przypadku = 2 (σ , μ). Wartość otrzymaną z obliczeń funkcji testowej porównujemy z przyjętą odpowiednio wartością procentową funkcji odczytaną z tabel rozkładu χ^2 , przy obliczonych stopniach swobody. Jeśli okaże się, że wartość ta jest większa od wartości testowej z odpowiednim ryzykiem błędu uznajemy rozkład za normalny.

Po obliczeniu i przedstawieniu graficznym wszystkich czynników wpływających na wydzielenie jednostek inżyniersko-geologicznych w profilu pionowym oraz poziomym powstaje ostateczny problem skonstruowania modelu inżyniersko-geologicznego, tj. podziału na różniące się między sobą jednostki inżyniersko-geologiczne i ich oceny jakościowej.

Zarówno podział, jak i ocena uzależnione są od celu całego opracowania. W przypadku opracowywanej mapy inżyniersko-geologicznej Płocka celem była ocena tego terenu dla bezpośredniego posadowienia obiektów. Wynikłe z tej oceny jednostki inżyniersko-geologiczne były sumą założonych czynników określonych wymiarów i korelacji.

Przyuszczalnie w przyszłości do wydzielenia tych jednostek będzie można dojść drogą założenia doświadczalnych czynnikowych i analizy czynnikowej,

omówienie których kończy tę wysoce jeszcze niedoskonałą próbę analizy i możliwości zastosowania metod statystyki matematycznej w geologii inżynierskiej.

Jedną z najmniej znanych metod w geologii jest analiza czynnikowa. Operuje ona, podobnie jak analiza korelacji, liniowym modelem matematycznym. Model ten jednak przez różne przekształcenia matematyczne jest w stanie ująć wszystkie czynniki i składniki czynników występujące w badaniu i wyodrębnić zmienne główne w związku przyczynowym. Odbywa się to za pośrednictwem wielu rotacji, co jest jednak możliwe jedynie przy zastosowaniu maszyn liczących. Pierwsza rotacja (tzw. wyjściowa) doprowadza do macierzy czynników, stanowiących rzuty, czyli ładunki wszystkich pierwotnych m zmiennych na n czynników. Z wielu metod analizy czynnikowej trzy zasługują na uwagę, nazwane odpowiednio R , Q i M .

Metoda R pozwala sprawdzać hipotezy co do zmiennych wywołujących dane zjawisko, np. powierzchniowe ruchy masowe. Można założyć, że w macierzy czynników pewne zmienne mają istotne znaczenie, a uzyskane przez rotację czynniki można wykorzystać dla potwierdzenia lub odrzucenia tego rodzaju założeń.

Metoda Q przez ciąg obliczeń może doprowadzić nie tylko do znajomości czynników i względnego znaczenia zmiennych zmieniających te czynniki, ale również i do ich układu przestrzennego.

Sprawdzanie hipotez metodą M , czyli jednoczesnego programowania macierzy dotyczy zmian w czasie. Daje ona ocenę w postaci czynników określających zmiany względnego znaczenia zmiennych, składających się na te czynniki, a więc zmian oceny poszczególnych jednostek.

Analiza czynnikowa jest więc szczególnym układem metod statystycznych opartych na analizie matematycznej. W jej procedurze wyróżnia się trzy etapy: wyodrębnienie zmiennych ze zbioru korelacji, rotacji oraz układu odniesienia oraz interpretację czynników. Etap pierwszy rozpoczyna się zebraniem pomiarów wszystkich zmiennych, charakteryzujących badane środowisko inżyniersko-geologiczne. Między wynikami oblicza się wszystkie możliwe współzależności, a otrzymane współczynniki układu się w macierz korelacji. Poprzez operacje na tej macierzy dokonuje się połączenia zmiennych wyjściowych w grupy wg największego podobieństwa.

Klasyfikacje i typologie wielowymiarowe przy użyciu opisanej ogólnie metody Q , w której między przedmiotami występują wiązki korelacji, każda zaś wiązka reprezentuje zbiór przedmiotów podobnych do siebie pod względem pewnych cech i różnych od przedmiotów innego zbioru, to nic innego jak podział obszaru na części wg podobieństwa i różnicy, przeto analiza czynnikowa, określając postępowanie związane z klasyfikacją daje tym samym podstawę do regionalizacji.

Oczywiście cała bardzo tu pobieżnie potraktowana analiza czynnikowa wymaga jeszcze wielu studiów, dotyczących zarówno jej strony merytorycznej, jak również sprawdzenia możliwości jej dalszego stosowania w praktyce inżyniersko-geologicznej.

Sumując przegląd poruszonych w artykule metod statystyki matematycznej należy wyraźnie podkreślić iż stosowanie każdej z nich powinna poprzedzać zawsze analiza merytorycznej ich przydatności w każdym konkretnym przypadku. I tak: stosowanie metody randomizacji przy lokalizacji podstawowych punktów obserwacyjnych wydaje się uzasadnione w bardzo ograniczonym zakresie, głównie dla opracowań obejmujących duże obszary, w których obrębie wydzielono już na podstawie przesłanek merytorycznych pewne jednostki, a dla których nie istnieją żadne predyspozycje dotyczące lokalizacji. Stosowanie analizy korelacyjnej i regresyjnej wraz z całym statystycznym aparatem pomiarowym wydaje się słuszne i uzasadnione we wszystkich niemal

większych opracowaniach inżyniersko-geologicznych z czym oczywiście wiąże się celowość próbek statystycznej pozwalających na właściwe wnioski.

Kwestia ostatecznego wydzielenia jednostek inżyniersko-geologicznych w układzie przestrzennym na podstawie statystyki w obecnym rozeznaniu jego stosowalności w tym zakresie nie jest jeszcze rozwiązana; chociaż wstępne zapoznanie się z analizą czynnikową wskazuje, iż problem ten ma szansę rozwiązania przy zastosowaniu tej właśnie metody.

LITERATURA

1. Chayes F. — On ratio correlation in Petrography. The Journal of Geology, vol. 57, No. 3. Chicago, 1949.
2. Kowalski W. C. — Jedność i różnorodność nauk geologicznych w świetle teorii prognoz. Prz. geol. 1967, nr 3.

SUMMARY

The paper deals with four statistical problems that appear during the subdivision into engineering-geological units. The problems are as follows:

1 — elaboration of geological model, definition of population examined and, on this basis, selection of variables,

2 — transposition of conception model on statistical model in which the mathematical structure of observation distinctly comprises a series of variance sources,

3 — statistical sampling according to statistical model,

4 — determination of approximations to population on the basis of a sample.

The problems mentioned above are discussed by the present author on the basis of empirical data obtained during the elaboration of the Engineering-Geological Map of Płock Area, plotted by the geologists of the Engineering Geology Department, Warsaw University. Discussing various aspects of the process of subdivision into engineering-geological units, the author takes into account some selected methods of statistical analysis, particularly the Krumbein and Greybill's method of randomisation, the correlation analysis applied in determining both population and model of engineering-geological environment, the determination of statistical samples, and the application of factor analysis in its three variants (R, Q, M). A necessity is especially emphasized by the author to base all the statistical methods on the following factors: thorough examination of non-statistical aspects of the geological problem studied, and concrete decision as concerns the usefulness of a given statistical method with respect to each definite problem of the engineering-geological researches.

3. Krumbein W. C., Graybill F. A. — An introduction to statistical models in geology. New York, 1965.
4. Linnik I. W. — Metoda najmniejszych kwadratów i teoria opracowywania obserwacji, Warszawa, 1962.
5. Miller R. L., Olsen E. C. — The statistical stability of quantitative properties as fundamental criterion for the study of environments. The Journal of Geology, vol. 63, No. 4, Chicago, 1955.
6. Parks J. M. — Cluster analysis applied to multivariate geologic problems. Ibidem, vol. 74, No. 5. Chicago, 1966.
7. Szarapow I. P. — Применение математической статистики в геологии. Москва, 1965.
8. Vistelius A. B. — Sedimentation time trend functions and their application for correlation of sedimentary deposits. The Journal of Geology, vol. 69, No. 6. Chicago, 1961.

РЕЗЮМЕ

В статье рассматриваются четыре статистические проблемы, возникающие при инженерно-геологической классификации:

1) составление геологической модели, определение исследуемой популяции и выбор переменных;

2) преобразование концепционной модели в статистическую модель, в которой математическая структура данных охватывает ряд источников дисперсии;

3) отбор статистической пробы, использованной в статистической модели;

4) определение приближений для популяции на основании пробы.

Автор рассматривает эти проблемы на основании эмпирических данных Инженерно-геологической карты г. Пlock, составленной коллективом кафедры инженерной геологии Варшавского университета. Рассматривая разные варианты инженерно-геологической классификации, автор учитывает некоторые методы статистического анализа, в частности метод рандомизации Крамбеина и Грейбилла, применение корреляционного анализа в определении популяций и модели инженерно-геологической среды, определения статистических проб и применения метода множителей в трех вариантах (R, Q, M). Автор особенно подчеркивает необходимость основания всевозможных статистических методов на тщательном изучении переменных свойств исследуемого геологического объекта и конкретном выяснении пригодности данного статистического метода для исследования определенной проблемы инженерно-геологических работ.