

Geostatystyka dla niematematyków. Przewodnik praktyczny – część I

Wojciech Naworyta¹



Geostatistics for non-mathematicians. A practical guide – Vol. I. Prz. Geol., 72: 341–349; doi: 10.7306/2024.18

Abstract. Geostatistical tools are useful and even necessary in many fields, not only in geology. The obstacle to their widespread use is the seemingly difficult mathematical foundations with which the teaching of the subject usually begins. Many years of experience and observations allow me to claim that even if geostatistical methods are used, they are often incomplete and imperfect. In a series of three articles, I would like to introduce potential non-mathematicians to the most important methods and tools from the arsenal of spatial statistics and how to use them properly. I will indicate the areas where they can be used, explain whether they can always be used, show what decisions should be made during calculations and how to interpret the obtained results. It will not be a compendium, but rather a pocket guide facilitating the reader's first contact with geostatistics. In the first article, in the introduction, I will explain where the title comes from; I will show what geostatistics can be useful for; what is interpolation, is it always possible, when does it make sense and what does its accuracy depend on? I will show how imperfect the research material available to a geologist is and why we should use sophisticated software to solve seemingly simple problems. Using a non-obvious example, I will try to explain the phenomenon of autocorrelation, which is important in geostatistics. I will also ask a few questions, the answers to which will be in the next article in this series.

Keywords: geostatistics, interpolation, autocorrelation, range of autocorrelation

Mam wielki szacunek dla matematyków. Swego czasu naczytałem się o lwowskiej szkole matematycznej i trudno mi nie żywić podziwu dla pięknych umysłów – Banacha, Steinhausa, Ulama, Auerbacha, Mazura i wielu innych. W przeszłości często zdarzało się, że abstrakcyjne rozważania matematyków znajdowały zastosowanie dopiero po wielu latach, gdy człowiek i technika dojrzały do wykorzystania efektów ich teoretycznej pracy. Dziwnym zbiegiem okoliczności krótko po studiach zostałem immatrykulowany na Wydziale Matematyki Stosowanej w Akademii Górniczej we Freibergu, gdzie w ramach Kolegium Doktoranckiego Räumliche Statistik usiłowałem się doktoryzować. Kilka razy w roku spotykaliśmy się na wyjazdowych szkołach w malowniczych Rudawach, gdzie każdy z nas referował postępy swojej pracy i przedstawiał plany na przyszłość. Bardzo męczyły mnie te spotkania. Już wtedy dość dobrze mówiłem po niemiecku, znałem angielski, ale kompletnie nie wiedziałem, o czym oni – moi koledzy matematycy – mówią. Na podstawie jednej prostej formułki potrafili gawędzić i dyskutować godzinami. Wtedy poznałem bezmiar mojej niewiedzy. Z geostatystyką zetknąłem się właśnie w Niemczech podczas studiów doktoranckich. Symulację geostatystyczną zastosowałem do modelowania parametrów górotworu w celu prognozowania deformacji powstających pod wpływem podziemnej eksploatacji złóż. Podstawy tej dziedziny wykładał mój promotor – matematyk. Po semestrze spędzonym na wykładach profesora byłem tak samo mądry jak pół roku wcześniej. Profesor próbował wytłumaczyć meandry geostatystyki, poczynając od dość skomplikowanej strony matematycznej (Menz, 2000). Nie był to błąd, było to wręcz klasyczne podejście, stosowane w wielu szkołach do tej pory. Niestety, moje wykształcenie matematyczne nie było wystarczające, abym mógł podążać za myślą wykładowcy. Z geostatystyką zaprzyjaźniłem się w końcu dzięki samodzielnej lek-

turze różnych źródeł, również tych pisanych po polsku (Journel, Huijbregts, 1978; Namysłowska-Wilczyńska, 1986; Isaaks, Srivastava, 1989; Mucha i in., 2007). Doświadczałem przy tym wielokrotnie tego, co Niemcy określają mianem *Aha!-Efekt*. Dotarło do mnie, że w gruncie rzeczy to wszystko jest dość proste, wystarczy tylko nieco inaczej o tym opowiedzieć. Po latach, kiedy sam wykładałem w Akademii Górniczo-Hutniczej przedmiot o nazwie modelowanie złóż, ugruntowałem się w przekonaniu, że aby przekazać studentom konieczną wiedzę, nie trzeba wcale zaczynać od skomplikowanych podstaw teoretycznych. Wystarczy nauczyć podstawowych narzędzi i pokazać, jak one działają, jak należy przygotować dane, jakie wnioski można wyciągnąć na podstawie wstępnych analiz i wreszcie, jak należy interpretować uzyskane wyniki. Przecież, aby poprawnie jeździć samochodem, nie trzeba zaczynać nauki od budowy silnika ani od składu chemicznego oleju. Ta wiedza nie jest niezbędna, by w bezpieczny sposób korzystać z urządzenia, jakim jest auto. Aby posługiwać się młotkiem, nie zaczynamy nauki od podstaw mineralogii, metod wytopu stali czy propedeutyki kowalstwa. Można się bez tej wiedzy obejść i z powodzeniem wbić gwoździe.

Dlaczego tak się upieram, aby pominąć tło matematyczne i idąc niejako na łatwiznę, przejść od razu do wykorzystania narzędzi i interpretacji wyników? Otóż narzędzia geostatystyczne wymyślili i nadal próbują udoskonalać matematycy, ale w praktyce i z powodzeniem stosują je geolodzy, mineralodzy, gleboznawcy, specjaliści od ochrony środowiska, nawet leśnicy i meteorolodzy. A przecież nie są oni matematykami, nie znają, bo i nie muszą znać, całego aparatu pojęciowego, który jest wymagany, aby zrozumieć język królowej nauk. Przy tym matematycy, nie będąc specjalistami w wyżej wymienionych dziedzinach, nie stosują na co dzień narzędzi geostatystycznych.

¹ Wydział Inżynierii Lądowej i Gospodarki Zasobami, Akademia Górniczo-Hutnicza im. St. Staszica, al. Adama Mickiewicza 30; 30-059 Kraków; naworyta@agh.edu.pl; ORCID ID: 0000-0003-4569-3907

Czyli produkują narzędzia, którymi właściwie się nie posługują. To poważne zastrzeżenie. Bo w konsekwencji powstają różne modyfikacje i pozorne ulepszenia już istniejących i sprawdzonych narzędzi, które wcale nie poprawiają znacząco efektów pracy, przeciwnie, głównie ją komplikują.

Na podstawie moich długetletnich obserwacji doszedłem do wniosku, że nauczanie geostatystyki według klasycznej szkoły bardzo szybko zniechęca niedoszłych użytkowników i w konsekwencji prowadzi do nikłej znajomości bardzo przydatnych w wielu dziedzinach narzędzi. W takiej opinii nie jestem odosobniony. Podobne spostrzeżenie wyraził znacznie wcześniej Tadeusz Cyrul (1992). Pomysł cyklu artykułów na ten temat rodził się w mojej głowie przez lata. Prowadząc zajęcia dydaktyczne, próbowałem dotrzeć z wiedzą do studentów i wiem, że przynajmniej w kilku przypadkach udało mi się zrealizować powyższy zamiar. Postanowiłem pokazać podstawowe narzędzia od strony praktycznej, tak jak wspomniane auto albo młotek. Chciałbym, aby po lekturze tego tekstu Czytelnik mógł powiedzieć: *przecież to bardzo proste!*

Wzorem Stephena Hawkinga postanowiłem ograniczyć liczbę cytowanych w moich artykułach wzorów. Ten genialny fizyk we wstępie do swojej bestsellerowej książki pt. *Krótką historią czasu* stwierdził, że każdy wzór zmniejszy liczbę czytelników o połowę. Oczywiście, nie obejdzie się bez przedstawienia kilku najbardziej podstawowych formuł, ale postaram się je wytłumaczyć tak, jakbym sam chciał, aby były mi one wytłumaczone.

W procesie modelowania zmienności interesującego nas parametru geologicznego wykonujemy tysiące obliczeń. Każdy węzeł siatki interpolacyjnej, czyli coś jakby piksel na bitmapie, powstaje w wyniku rozwiązania osobnego systemu równań, a tych punktów na jednej mapie mogą być setki tysięcy. Obliczenia oparte na równaniach macierzowych są żmudne i czasochłonne. To oczywiste, że aby je wykonać, nie będziemy prowadzić rachunków ręcznie ani stukając w klawiaturę najlepszego nawet inżynierskiego kalkulatora. Do obliczeń wykorzystuje się specjalistyczne, dedykowane do tego celu programy komputerowe. Dlatego w swoim tekście będę się odwoływał do pojęć i nazw, jakie w tych programach funkcjonują. *Software* geostatystyczny, mimo że prosty w obsłudze, wymaga jednak podstawowej wiedzy operatora. Nie są to programy intuicyjne i nie sposób nauczyć się ich obsługi przez przypadkowe klikanie myszą. Operacje są wykonywane w kilku etapach i wymagają od operatora podejmowania świadomych wyborów. Każda decyzja jest brzemienna w skutkach. Próba ominięcia pewnych etapów prowadzi na manowce. Co prawda w programach stosuje się czasem ścieżkę na skróty, dedykowaną mniej świadomym operatorom. Takie podejście jest możliwe, ale nierekomendowane. W tym przypadku zamiast podejmować świadome decyzje, możemy po załadowaniu danych użyć przycisku: *Skip to end*. Program, wykorzystując automatyczne, uproszczone funkcje, wykona za nas działania i wskaże wynik. Wynik ten będzie efektem kompromisu. Mapy wykonane w ten sposób będą wyglądały ładnie i pozornie poprawnie, jednak nie będą dokładne, mogą zawierać poważne błędy prowadzące do podjęcia nietrafionych decyzji biznesowych albo błędnych wniosków naukowych. Korzystanie z takich dróg na skróty można przyrównać do sytuacji, kiedy zasiadając za kierownicą nowoczesnego lamborghini, zaprzęgamy do auta

parę koni zamiast uruchamiać silnik. Jedziemy, to prawda, ale czy po to kupowaliśmy drogie auto?

W swojej praktyce wykonałem wiele map różnych zjawisk przestrzennych, wykorzystując różne rodzaje krigingu albo metody symulacji geostatystycznej. Za jednym kliknięciem myszy komputer rozwiązywał setki równań macierzowych w ułamkach sekund. Niemal natychmiast otrzymywałem wynik w postaci siatki interpolacyjnej lub od razu kolorowej mapy. Taka szybkość obliczeniowa nikogo już dzisiaj nie dziwi, szczególnie studentów, którzy należą do pokolenia urodzonego w XXI w., czyli w czasach, kiedy komputery zawsze były. Obsługując program, warto jednak pamiętać, że podstawy teoretyczne, które stoją za działaniem programów geostatystycznych, wymyślono w latach 50. XX w., w czasach, gdy komputery o mocy obliczeniowej dzisiejszych smartfonów zajmowały kilka pomieszczeń biurowych. To również stąd bierze się mój podziw i szacunek dla matematyków. Produkty ich umysłów wyprzedzają czas.

Specjalistyczny *software* przeznaczony do rozwiązywania problemów geostatystycznych jest oprogramowaniem niszowym. W związku z tym nie jest tłumaczony na wszystkie języki świata. Jego obsługa wymaga podstawowej znajomości języka angielskiego. Dlatego w moich artykułach obok polskich pojęć będę stosował również angielskie terminy używane w tych programach. To może być pomocne, bo znaczeń słów stosowanych w geostatystyce próżno szukać w popularnych słownikach, nawet w translatorach internetowych.

GEOSTATYSTYKA – CZYM JEST I DO CZEGO MOŻE BYĆ PRZYDATNA?

Gdy po raz pierwszy zetknąłem się ze słowem geostatystyka, byłem przekonany, że chodzi o coś, co ma związek z geomechaniką. Szybko okazało się, że byłem w błędzie. Nie ma to żadnego związku ze statyką, jest natomiast gałęzią statystyki mającą zastosowanie do rozwiązywania problemów w dziedzinie *geo*, czyli Nauk o Ziemi. Czasem mówi się też o statystyce przestrzennej, bo każdy pomierzony parametr, będący przedmiotem analiz, obok wartości ma przypisane współrzędne. To umiejscowienie liczb w przestrzeni różni geostatystykę od klasycznej statystyki, z którą ta pierwsza ma jednak wiele wspólnego. Pierwotnie narzędzia geostatystyczne stworzono do modelowania złóż kopalin, czyli rozwiązywania zagadnień górniczo-geologicznych. Dzisiaj jednak mają znacznie szersze zastosowanie, np. są wykorzystywane do rozwiązywania wielu problemów z różnych obszarów leżących poza geologią i górnictwem. Najbardziej oczywistym przykładem jest zastosowanie pierwotne, w wyniku którego geostatystyka powstała, czyli modelowanie parametrów złóż. Złoże to nagromadzenie kopaliny, które może być przedmiotem ekonomicznie uzasadnionej eksploatacji. Najczęściej złoże leży pod innymi osadami geologicznymi, a więc gołym okiem go nie widać. Nie ma też wciąż cudownych metod, za pomocą których moglibyśmy górotwór prześwietlić, tak jak ludzkie ciało promieniami Roentgena. Aby poznać złoże, trzeba wykonać roboty geologiczne, najczęściej w postaci wierceń. W wyniku przeprowadzonych badań pozyskuje się informacje z otworów, np. w formie rdzeni. Z każdego rdzenia można pozyskać informacje o parametrach strukturalnych lub jakościowych złoże: grubości pokładu, głębokości

zalegania, chemicznych właściwościach skały itd. Każdy otwór jest umiejscowiony w przestrzeni za pomocą współrzędnych. Problem polega na tym, że informacja z otworów geologicznych, jakkolwiek bogata, jest informacją dyskretną, punktową, nie obejmuje całego złoża. Odwiert to jakby punkcja ciała złożowego. Tymczasem geolog, a potem górnik, chciałby znać rozkład albo zmienność parametrów w całej przestrzeni występowania złoża. Złoże ma, przypuśćmy, powierzchnię kilku kilometrów kwadratowych, a w tabeli mamy informację z kilkudziesięciu otworów rozrzuconych w nieregularnej sieci pomiarowej. Aby móc wnioskować o badanym złożu, a potem podejmować decyzje biznesowe, należy na podstawie posiadanej informacji wykonać mapę, która zobrazuje, jak rozkładają się poszczególne parametry. Moje doświadczenia wskazują, że obserwacji z otworów zawsze jest za mało, ale to niestety wszystko, czym tu i teraz dysponujemy. I tu właśnie w sukurs przychodzi geostatystyka. Stosując wyrafinowane narzędzia z jej arsenału, najpierw dowiemy się, czy na podstawie dostępnych danych możemy w ogóle wykonać jakąkolwiek mapę zmienności parametrów złoża, czy interpolacja parametrów jest możliwa i czy ma sens. Jeśli tak, to stosując jedną z metod interpolacyjnych, np. kriging, wykonamy mapę. Jako bonus albo jako produkt uboczny interpolacji przeprowadzonej metodą krigingu otrzymamy mapę wiarygodności modelu, czyli mapę rozkładu teoretycznego błędu, z jakim został wykonany model parametru w każdym punkcie złoża. To bardzo ważny element analizy, odróżniający geostatystyczne metody interpolacyjne od innych. O interpolacji metodą krigingu napiszę więcej w trzecim artykule tego cyklu, pokazując meandry i pułapki, jakie czyhają na niecierpliwego operatora *software'u* na każdym etapie pracy. Bez tej wiedzy łatwo o pomyłkę i błędne wnioski, których konsekwencje w dalszym postępowaniu górniczym mogą być kosztowne.

Analiza złóż jest pierwotnym przykładem wykorzystania geostatystyki. To właśnie stąd wzięła się potrzeba, dla której stworzyli tę specjalizację – D. Krige i G. Matheron – począwszy od lat 50. XX w. wymyślili narzędzia powszechnie stosowane dzisiaj w światowej geologii i górnictwie (Matheron, 1962). Metody geostatystyczne mają jednak znacznie szersze zastosowanie, m.in. do badań zanieczyszczenia gleby, np. metalami ciężkimi na podstawie próbek pobranych punktowo w terenie. Jeżeli po zakończeniu badań terenowych wydaje się, że próbek jest za mało, to oczywiście można pobrać dodatkowe próbki, czyli wykonać dodatkowe obserwacje. To znacznie łatwiejsze i tańsze niż wykonanie dodatkowych otworów geologicznych. Niemniej jednak, każde pozyskanie dodatkowych obserwacji, czyli pobranie próbek w terenie, to dodatkowa praca, która jest czasochłonna, która kosztuje, której dalszy ciąg przebiega w laboratorium z wykorzystaniem kosztownych odczynników i aparatury. Dostarczenie nowych danych zawsze jest problematyczne. O wiele łatwiej jest usiąść przed komputerem, zaimportować posiadane dane do odpowiedniego programu i w przemyślany sposób zastosować *software*, aby z dostępnych obserwacji uzyskać jak najwięcej, możliwie najdokładniejszych informacji o zanieczyszczeniu gleby na całej powierzchni analizowanego terenu. Do tego właśnie nadaje się geostatystyka.

Uogólniając – geostatystyka jest przydatna wszędzie tam, gdzie na podstawie ograniczonej liczby dyskretnych

obserwacji trzeba wyciągnąć wnioski o zjawisku przestrzennym w całym obszarze jego występowania. Stosując jej narzędzia, możemy uzyskać informacje o badanym zjawisku, które są zawarte w obserwacjach i które wynikają nie tylko z samych liczbowych miar pojedynczych obserwacji, ale również z wzajemnego ich usytuowania w przestrzeni. Biegły analityk wyposażony w najprostszy, specjalistyczny *software* wyciśnie z dostępnej informacji wszystko, co tylko jest możliwe, łącznie z informacją o błędach obserwacji terenowych albo, jeżeli już nie można inaczej, o konieczności uzupełnienia obserwacji w terenie.

Nie wszędzie stosowanie narzędzi geostatystycznych ma sens. Jeśli tak, to gdzie nie ma sensu? Na przykład do wykonywania modelu powierzchni terenu. Tu już nie posługujemy się dyskretnymi danymi, obserwacjami wykonanymi w terenie, np. za pomocą teodolitu. Kiedyś owszem, byłoby to zasadne, ale stosowanie geostatystyki i tak było ograniczone do powierzchni terenu o niewielkiej zmienności. Wąwozy, doliny, żleby, lokalne wzniesienia, większe i nieregularne deniwelacje nie są łatwe do modelowania. A co mamy w zamian? Otóż do modelowania powierzchni terenu, który można bez przeszkód obserwować, są przydatne metody fotogrametryczne. W tym przypadku mamy do czynienia ze zjawiskiem odwrotnym niż w przykładach przytoczonych wcześniej z obszaru geologii lub ochrony środowiska. W efekcie opracowania zdjęć lotniczych powstaje chmura punktów, po obróbce której można otrzymać bardzo dokładny model terenu. Chmura punktów to ogromna liczba obserwacji o dużym zagęszczeniu. W procesie opracowania danych terenowych wymagana jest ich redukcja, wycieszczenie z tzw. szumów lub obserwacji, które zniekształcają obraz. Tymczasem w przytoczonych przykładach z dziedziny geologii i ochrony środowiska obserwacji jest najczęściej za mało. Chciałoby się tę sieć obserwacji zagęścić, mieć ich więcej. Stosowanie geostatystyki ma sens tam, gdzie badane zjawisko przestrzenne nie jest widoczne inaczej jak tylko za pomocą dyskretnych obserwacji.

OBSERWACJE TERENOWE – DALECE NIEDOSKONAŁY PUNKT WYJŚCIA DO ANALIZ GEOSTATYSTYCZNYCH

Informacje o zjawiskach przestrzennych pozyskane w terenie są dalece niedoskonałe. Niemal zawsze jest ich za mało, przeważnie zawierają błędy. I co najgorsze, niemal zawsze to jedyna informacja, którą dysponujemy i na podstawie której musimy wykonywać mapy służące do wnioskowania i podejmowania wiążących decyzji biznesowych lub naukowych.

Problem niedoskonałości obserwacji terenowych spróbuję zilustrować na rzeczywistym przykładzie z obszaru geologii i górnictwa. W procesie rozpoznania i dokumentowania jednego ze złóż węgla brunatnego odwiercono łącznie aż 730 otworów. Wiercenia przeprowadzono w kilku kampaniach rozpoznawczych w ciągu kilkunastu lat. Dla geologa tyle otworów w jednym złożu to gratka – wydaje się, że złoże zostało rozpoznane niemal doskonale. Średnia statystyczna miąższość głównego pokładu złoża to ok. 10 m. Średnio każdy z 730 rdzeni wyciągniętych z górotworu w procesie rozpoznania zawierał ok. 0,13 Mg węgla brunatnego. To znaczy, że w procesie rozpoznania wydobyto na powierzchnię ok. 95 Mg kopaliny. Taka ilość węgla

mieści się w sześciu pełnych ciężarówkach. Na podstawie tej masy węgla wnioskowano o całym złożu, którego zasoby w toku prac geologicznych oszacowano na ok. 1,6 mld Mg. Gdy zestawimy łączną masę próbek wydobytych z otworów rozpoznawczych z oszacowaną masą całego złoża to okaże się, że próbka stanowi część masy złoża, która wyraża się liczbą 0,00000006. Liczba 6 znajduje się na ósmym miejscu po przecinku! A gdybyśmy wykonali dodatkowy wysiłek fizyczny i finansowy i podwoili liczbę otworów w złożu, czy to by coś zmieniło? Zamiast 730 obserwacji mielibyśmy ich 1460, a to przełożyłoby się na 190 Mg wydobytej w trakcie rozpoznania kopaliny. W zestawieniu z ogólną oszacowaną masą węgla w złożu byłby to stosunek wyrażający się liczbą 0,00000012.

Kiedyś przyrównałem dokumentowanie złóż do sondażu preferencji politycznych, np. w sprawie wyborów prezydenckich. Chciałem pokazać, w jakim zakresie niepewności obracają się geolodzy, wnioskując o właściwościach złoża na podstawie dyskretnych próbek, jakimi są informacje z otworów wiertniczych. Wydawało mi się wtedy, że porównanie do sondażu jest mocno przesadzone, bo przecież wiadomo, że sondaże są bardzo niepewne, a wyniki wyborów są często zaskakująco odmienne od tego, co było oczekiwane na podstawie publikowanych wcześniej sondaży przedwyborczych. Poza tym ankietowane osoby wcale nie muszą mówić prawdy, równie dobrze mogą zmienić swoje preferencje, mogą też celowo wprowadzać ankietera w błąd. Sprawdźmy zatem, jak mają się sondaże do prac geologicznych. Typowy sondaż preferencji politycznych w Polsce przeprowadza się telefonicznie na próbie liczącej około tysiąca respondentów. Osoby te stanowią grupę reprezentującą przekrój społeczeństwa. Ten tysiąc osób w stosunku do ok. 30 mln dorosłych uprawnionych do głosowania to odsetek wyrażający się liczbą 0,00003. Porównanie rozpoznania geologicznego do sondażu społecznego okazuje się więc rzeczywistości nietrafione. Próba sondażowa jest znacząco większa od próby geologicznej pobranej z przykładowego złoża, różni się od niej aż o trzy rzędy wielkości!

Przykład ten miał zilustrować stopień niepewności i niewiedzy, z którym geolog musi sobie radzić w swojej pracy. W procesie dokumentowania o złożu wiemy bardzo niewiele, ale ta wiedza musi nam wystarczyć do opracowania danych, np. wykonania map, obliczania zasobów, wyceny złoża, przygotowania planów eksploatacji. Teraz Czytelniku chyba już nie masz wątpliwości, dlaczego mówiąc o obserwacjach geologicznych, twierdzę, że jest to informacja dalece niedoskonała. Muszę dodać, że złoża, które posłużyły mi do ilustracji tego problemu, należy do bardzo dobrze udokumentowanych.

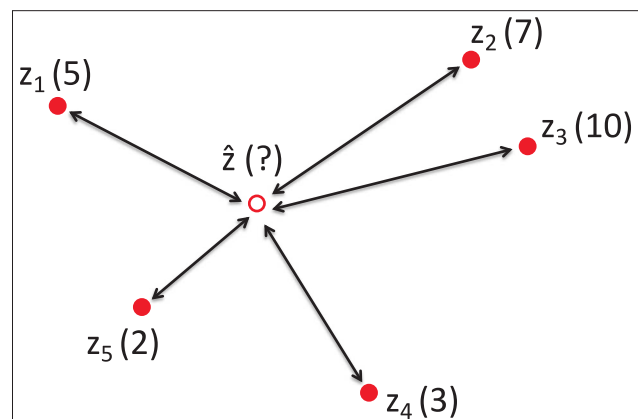
Problem niedoskonałości danych podkreśliłem z premedytacją. Musimy zrobić wszystko, aby wycisnąć z nich całą zawartą informację, a ta kryje się nie tylko w pojedynczych otworach wiertniczych, ale również w relacjach pomiędzy nimi. Dlatego właśnie poznanie i posługiwanie się wyrafinowanymi narzędziami geostatystycznymi ma głęboki, praktyczny sens.

KILKA SŁÓW O INTERPOLACJI

Interpolacja jest oszacowaniem wartości parametrów pomiędzy obserwacjami. Nie jest to proste obliczenie wartości, tylko estymacja, przewidywanie, prognoza, szacowanie. Będę tych słów używał wymiennie. Oczywiście

estymacja opiera się na obliczeniach z użyciem określonych algorytmów matematycznych, ale wartości będące wynikiem tych obliczeń to wartości prognozowane. To oczywiste, bo przecież to, co naprawdę wiemy o analizowanym zjawisku, to tylko to, co zostało przez nas zmierzone, czyli obserwacje w otworach wiertniczych, czy w miejscach poboru próbek, np. gleby. Możemy jedynie przewidywać, co występuje pomiędzy obserwacjami. Weźmy trywialny przykład. Chcemy wykonać interpolację pomiędzy dwiema obserwacjami miąższości pokładu węgla kamiennego. W jednym otworze zmierzylismy pokład o grubości 10 metrów, w drugim tylko 8. W środku odcinka łączącego obydwie otwory oczekujemy, że pokład węgla będzie miał miąższość ok. 9 metrów. Wszystko na to wskazuje i jest to wartość najbardziej prawdopodobna. Jest ona równie podobna do jednej i drugiej sąsiadującej obserwacji. Oczekujemy, że miąższość pokładu stopniowo zmniejsza się z tych 10 m w jednym otworze do 8 m w drugim – tylko, że nie mamy gwarancji, że parametr zmienia się stopniowo. Równie dobrze w środku odcinka łączącego dwa otwory grubość pokładu może wynosić 12 metrów. Możliwe? Ależ oczywiście, że możliwe, bo zjawisko, jakim jest osadzanie się materii organicznej i powstawanie z niej pokładów węgla w długotrwałych procesach geologicznych wcale nie musiało przebiegać w sposób regularny i liniowy. Dlatego interpolację zawsze traktujemy jako prognozę, która charakteryzuje się określonym prawdopodobieństwem. To bardzo ważne, bo w mojej praktyce zauważyłem tendencję do nadmiernej wiary w obliczenia komputerowe. Jeżeli w wyniku zastosowania algorytmów obliczeniowych komputer narysuje mapę, to jest ona traktowana jak prawda objawiona. Nie i jeszcze raz nie! Jedyne, co jest pewne, to są tzw. twarde dane pozyskane w toku obserwacji w terenie. Wszystko, co powstaje w wyniku obliczeń na podstawie tych obserwacji, jest tylko i wyłącznie prognozą, estymacją, przewidywaniem i tak to należy traktować.

Klasyyczny przykład szacowania wartości w dowolnym punkcie na podstawie obserwacji z najbliższego otoczenia przedstawiłem na ryc. 1. Intuicyjnie czujemy, że wpływ pojedynczej obserwacji z_i na wartość szacowaną \hat{z} zależy od odległości między nimi. Owszem, ale obserwacje z_1 i z_4 leżą niemal w takiej samej odległości od \hat{z} , podobnie z_2 i z_3 . Która zatem wartość będzie dominująca? Rozwiązanie



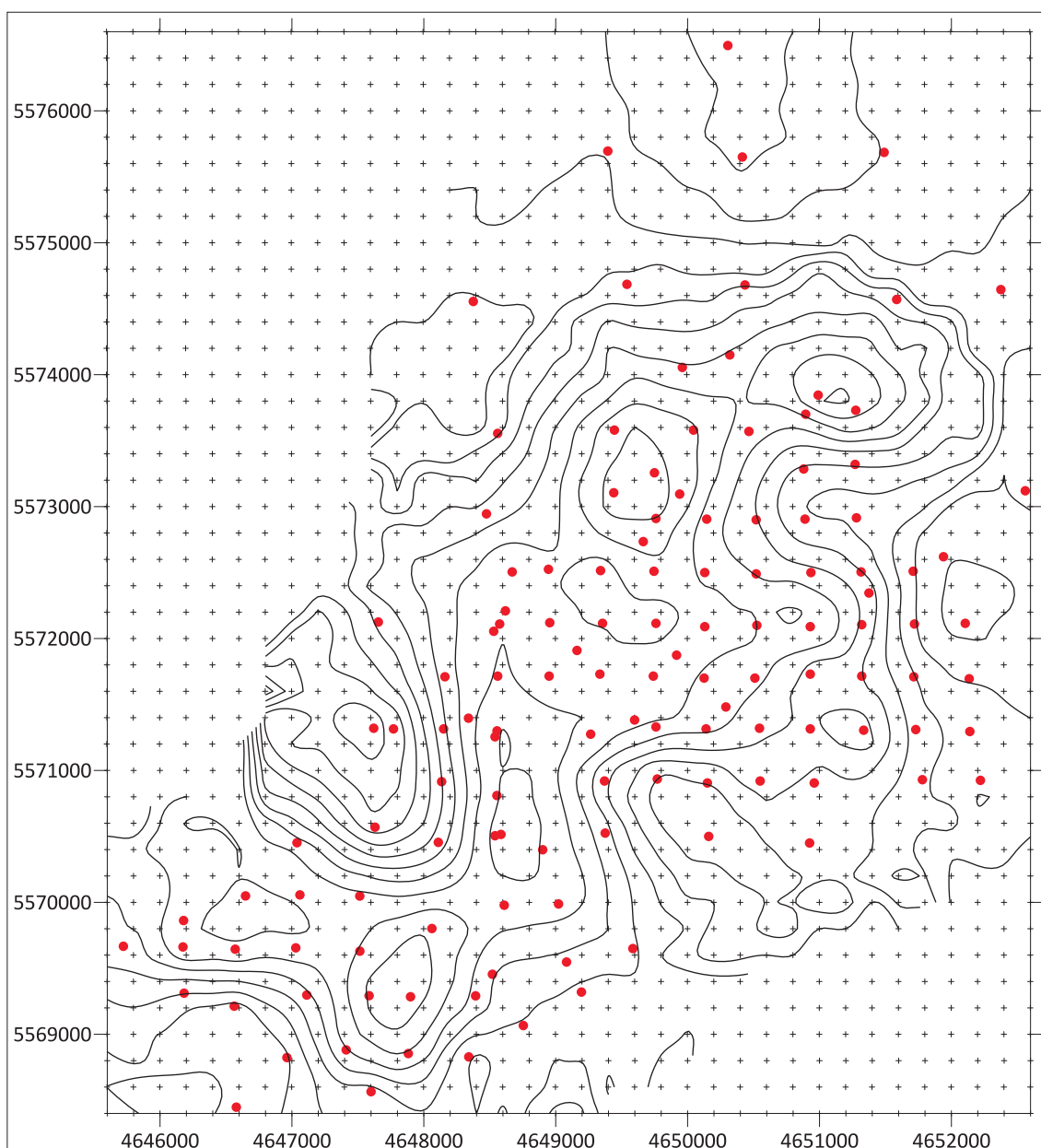
Ryc. 1. Idea interpolacji – oszacowanie wartości nieznanej na podstawie obserwacji z sąsiedztwa

Fig. 1. The idea of interpolation – estimation of the unknown value based on observations from the neighbourhood

tego zagadnienia wykracza już poza ramy kilku prostych obliczeń, które dałoby się wykonać ręcznie. Tu trzeba zaprzęgnąć jakiś algorytm. Sprawa komplikuje się wtedy, gdy obserwacji jest więcej, występują w nieregularnej sieci, czyli właśnie tak, jak w praktyce dzieje się najczęściej. Dlatego poświęcę temu nieco więcej miejsca.

W specjalistycznych programach interpolacja to przyporządkowanie w węzłach siatki interpolacyjnej prognozowanych wartości analizowanego parametru na podstawie obserwacji. Na rycinie 2. przedstawiłem lokalizację otworów rozpoznawczych (czerwone punkty) oraz węzły siatki interpolacyjnej (krzyżyki). Siatka interpolacyjna (*grid*) składa się z regularnie rozmieszczonych punktów określonych mianem węzłów (*node*). Cechy siatki, tj. zakres czy gęstość, zależą od decyzji operatora. Najczęściej nie ma powodu, aby odległości między węzłami siatki (*spacing*) w kierunku X i Y były od siebie różne, oczka siatki mają więc kształt kwadratów. W każdym węźle siatki algorytm matematyczny obliczy wartość na podstawie obserwacji i ich położenie względem każdego węzła. W wyniku tych obliczeń powstanie regularna sieć punktów z przypisanymi wartościami analizowanego parametru. Dopiero na kanwie tej wypełnionej liczbami regularnej sieci zostanie wykreślona mapa izoliniowa, czyli obraz zmienności analizowanego fenomenu (miąższości złoża, zawartości składnika chemicznego, zanieczyszczenia gleby itp.). Co ważne, w procesie kreślenia mapy izoliniowej nie będą już wykorzystywane obserwacje. Ich rola skończyła się w procesie interpolacji. Będzie to miało niekorzystne konsekwencje w tzw. efekcie wygładzania (*smoothing effect*). Wartości niektórych danych, zwłaszcza te skrajnie wysokie lub skrajnie niskie, będą odstawały od wygenerowanej powierzchni

ku X i Y były od siebie różne, oczka siatki mają więc kształt kwadratów. W każdym węźle siatki algorytm matematyczny obliczy wartość na podstawie obserwacji i ich położenie względem każdego węzła. W wyniku tych obliczeń powstanie regularna sieć punktów z przypisanymi wartościami analizowanego parametru. Dopiero na kanwie tej wypełnionej liczbami regularnej sieci zostanie wykreślona mapa izoliniowa, czyli obraz zmienności analizowanego fenomenu (miąższości złoża, zawartości składnika chemicznego, zanieczyszczenia gleby itp.). Co ważne, w procesie kreślenia mapy izoliniowej nie będą już wykorzystywane obserwacje. Ich rola skończyła się w procesie interpolacji. Będzie to miało niekorzystne konsekwencje w tzw. efekcie wygładzania (*smoothing effect*). Wartości niektórych danych, zwłaszcza te skrajnie wysokie lub skrajnie niskie, będą odstawały od wygenerowanej powierzchni



Ryc. 2. Izolinie miąższości przykładowego złoża jako efekt interpolacji wartości z obserwacji (czerwone punkty) na siatkę interpolacyjną (krzyżyki)

Fig. 2. Isopachs of an example deposit as a result of interpolation of observation values (red points) into the interpolation grid (crosses)

modelu. *Smoothing effect* sprawia, że w modelu będą miały zawsze zawężony zakres w stosunku do danych źródłowych. Najmniejsza wartość modelu będzie nieco wyższa niż najniższa wartość zaobserwowana i odwrotnie w obszarze wartości najwyższych. Zjawisko wygładzania jest charakterystyczne dla wszystkich metod interpolacji, geostatystycznych również. Należy być tego świadomym. Do interpolacji można zastosować różne algorytmy. W programach komputerowych są ich dziesiątki. Opierają się na różnych założeniach, różne są też równania matematyczne stosowane do obliczeń. Wśród nich można wymienić metody: odwrotność odległości do potęgi (*inverse distance to a power*), minimum krzywizny (*minimum curvature*), średnią kroczącą (*moving average*), regresję wielomianową (*polynomial regression*) i wiele innych. Od wyboru algorytmu interpolacyjnego zależy wynik. Mapy wykonane na podstawie tego samego zestawu obserwacji za pomocą różnych metod interpolacyjnych będą się od siebie różnić, nieraz znacznie. Te różnice ilustrują, w jakim zakresie niepewności porusza się geolog modelujący złożo. Co więcej, nikt nie jest w stanie powiedzieć z całkowitą pewnością, która z map lepiej przedstawia prawdziwy obraz złoża. Powtórzę jeszcze raz – jedyne co o złożu wiemy na pewno to obserwacje, wszystko co poza nimi to prognoza.

Ze względu na główny temat artykułu ograniczę się do opisanego algorytmów wynikających z założeń geostatystycznych. Wśród nich należy wymienić kriging i jego odmiany: zwyczajny, uniwersalny, cokriging, kriging indyktorowy i inne oraz symulacje geostatystyczne. Przewagą metod geostatystycznych nad pozostałymi jest informacja o wiarygodności albo inaczej – o teoretycznym błędzie modelu. Wykonywanie obliczeń metodą krigingu jest kilkustopniowe, wymaga wprowadzenia kilku ważnych pojęć, dlatego poświęcę temu więcej miejsca w trzecim artykule mojego cyklu.

Grid to w sensie matematycznym tabela. Każdy węzeł siatki zajmujący osobny wiersz w tabeli ma co najmniej trzy wartości. Współrzędne płaskie X i Y oraz cechę Z, czyli wartość liczbowa opisującą analizowany parametr. Dzięki współrzędnym X i Y położenie każdego węzła jest jasno określone w układzie współrzędnych.

Wielkość oczek siatki interpolacyjnej, czyli odległość sąsiadujących ze sobą węzłów siatki, zależy od decyzji operatora. Jak więc dobrać gęstość siatki interpolacyjnej, czym się kierować? Przez węzły siatki będą przechodziły izolinie (ryc. 2). Z jednej strony jeżeli węzłów będzie mało, to izolinie mogą nie być dość regularne, mogą się załamywać, wyglądać sztucznie, nienaturalnie. Izolinie będą miały regularny kształt, gdy węzłów będzie więcej, gdy siatka interpolacyjna będzie gęsta. Z drugiej strony wartość w każdym węźle wymaga osobnego rozwiązania systemu równań; im więcej węzłów, tym więcej bajtów zajmie wypełniona liczbami siatka interpolacyjna. Być może ten ostatni argument nie trafia do dzisiejszego użytkownika komputerów. Pamiętajmy jednak, że takich gridów do modelowania pojedynczego zjawiska można wykonać wiele. W przypadku symulacji geostatystycznej jednego zjawiska jednym kliknięciem myszy będziemy generować dziesiątki albo nawet setki gridów. Dysk komputera może się szybciej zapełnić niż nam się wydaje. Warto więc i to wziąć pod uwagę.

A co z dokładnością mapy wyrysowanej na podstawie siatki interpolacyjnej? Która mapa będzie bardziej

dokładna – ta wykreślona na rzadkiej czy na gęstej siatce? Otóż dokładność mapy niemal wcale nie zależy od liczby oczek w gridzie. Dokładność mapy zależy przede wszystkim od dokładności materiału źródłowego, czyli od obserwacji – od liczby punktów obserwacyjnych i ich rozmieszczenia w przestrzeni oraz od dokładności oznaczenia analizowanego parametru. Wracając jednak do siatki interpolacyjnej i jej gęstości. Przygotowując mapy parametrów dużych złóż najczęściej stosuję siatkę o wymiarach kwadratowych oczek 50×50 m. Biorąc pod uwagę, że otwory geologiczne są odległe od siebie średnio o 500 m, to między takimi otworami mieści się ok. 10 węzłów siatki. Gdyby tych węzłów było tylko 5, czyli odległość między węzłami (*spacing*) wyniosłaby 100 m, to izolinie wykreślone na ich podstawie byłyby nieco mało regularne, mogłyby wyglądać sztucznie. Jednak zwiększanie gęstości, czyli zmniejszanie odległości między węzłami do np. 10 m, prowadziłoby do drastycznego zwiększenia wielkości mapy (w sensie numerycznym) bez wpływu na poprawę efektu wizualnego. Moje preferencje nie są żadną regułą. W praktyce każdy użytkownik w zależności od przedmiotu obliczeń – wielkości złoża, gęstości obserwacji – powinien sam wypracować właściwości siatki interpolacyjnej. Niepisaną regułą jest 10-krotnie mniejsza odległość między węzłami w stosunku do średniej odległości między obserwacjami. Aby siatka interpolacyjna miała sens, gęstość jej węzłów powinna być co najmniej kilka razy większa niż gęstość obserwacji, które mamy zamiar interpolować. Pamiętajmy jednak, że to nie liczba węzłów w siatce decyduje o dokładności obliczeń interpolacyjnych.

AUTOKORELACJA – CZYLI O TYM, CZY ZAWSZE MOŻNA PRZEPROWADZIĆ INTERPOLACJĘ

Nie zawsze można przeprowadzić interpolację zjawiska, które w przestrzeni zostało rozpoznane w sposób punktowy. To bardzo ważne. Wiedza ta nie jest niestety powszechna. Wręcz przeciwnie, panuje przekonanie, że jeżeli dane pomiarowe zaimportujemy do komputera, wykonamy obliczenia i otrzymamy wynik, to znaczy, że wszystko jest w porządku. Nic bardziej mylnego. Komputer jest tylko tak mądry jak jego operator, czyli jak Ty drogi Czytelniku. To Ty musisz wiedzieć, czy interpolacja jest możliwa, czy takie obliczenia mają w ogóle sens. Komputer Ci tego nie powie, ale komputer może Ci pomóc znaleźć odpowiedź na pytanie, które Ty musisz zadać. Abyś je zadał, musisz najpierw wiedzieć, że trzeba o to zapytać. Ta wiedza to już bardzo dużo.

Przypadki, kiedy interpolacja wydaje się możliwa, są dość naturalne i intuicyjne, np. rozpoznanie podstawowych parametrów złóż kopalin pokładowych, czyli dajmy na to miąższości pokładu węgla brunatnego czy kamiennego, rozkładu zanieczyszczenia gleby w pobliżu emitora, przypuśćmy obok huty żelaza. W wymienionych przypadkach zjawiska zachodzą regularnie i zależą od określonych czynników zewnętrznych. Możemy oczekiwać, że do pewnej odległości od komina zanieczyszczenie gleby będzie większe i stopniowo będzie się zmniejszać aż do odległości, gdzie ta wartość zanieczyszczeń będzie na poziomie trudnym do wykrycia. Oczekujemy również, że większe stężenie zanieczyszczeń będzie po wschodniej stronie komina, ponieważ wiatry wieją głównie z zachodu.

To dla odmiany – kiedy interpolacja nie ma sensu? Zaczniemy od przykładu, który łatwo zobrazować. Wyobraźmy sobie, że wykonaliśmy pomiary rzędnych terenu w górach o budowie alpejskiej, np. w Tatrach. Pomiary wykonaliśmy w odległości ok. 500 m. Koło punktu o rzędnej 2200 m n.p.m. w pobliżu Koziego Wierchu znajdzie się punkt o wysokości 1650 m n.p.m. w okolicach Czarnego Stawu Gąsienicowego, a zaraz za nim znowu 2100 m n.p.m., bo będzie to już w szczytowej partii Kościelca. Gdybyśmy na bazie tych pomiarów narysowali przekrój to miałby on wygląd bardzo nieregularnych zębów. Interpolacja, czyli przewidywanie wartości rzeczywistych a nie pomierzonych pomiędzy sąsiadującymi obserwacjami, jest możliwa, owszem, ale czy wynik będzie choć trochę podobny do rzeczywistości? Raczej nie. Wynik będzie powierzchnią, która zostanie dopasowana do punktów obserwacji, ale w żaden sposób nie będzie odzwierciedlała rzeczywistego zjawiska, jakim jest układ dolin i szczytów tatrzańskich. Taka mapa będzie błędna, projektowanie wycieczek na jej podstawie prędzej czy później doprowadzi do tragedii. To przykład dobrze ilustrujący problem, ale niezbyt udany, bo jak już wspomniałem, geostatystyka nie służy do tworzenia map ukształtowania powierzchni terenu. Do tego służą przecież zdjęcia lotnicze. Wróćmy więc do geologii. Wyobraźmy sobie złożę węgla kamiennego, które w wyniku ruchów tektonicznych zostało bardzo mocno poprzeracinane uskokami o zróżnicowanej wielkości zrzutu. Takie złożę, gdyby je wyjąć z otaczającego górotworu, składałoby się z kawałków płyt zawieszonych w przestrzeni na różnych wysokościach, ustawionych pod różnymi kątami w stosunku do siebie. Takie zjawisko w całości również nie nadaje się do interpolacji. Interpolacja nie radzi sobie z uskokami. Byłaby możliwa, a nawet pożądana, wyłącznie w obrębie jednego fragmentu złoża, pojedynczej płyty nie pociętej uskokami. W geologii występuje wiele zjawisk, co do których już wiemy, że występują losowo. Jednym z nich jest zawartość siarki całkowitej w złożach węgla. To samo dotyczy zawartości metali ciężkich w węglu. Próżno doszukiwać się jakiejś reguły, która rządzi zmiennością przestrzennego występowania tych pierwiastków. O ile potrafimy sobie wyobrazić zjawisko sedymentacji materii organicznej i tworzenie się pokładów węgla, to już trudniej wyjaśnić, dlaczego w pewnym obszarze osadów węglowych jest więcej siarki, a gdzie indziej nie ma jej niemal wcale.

Inny przykład, tym razem polegający na znajomości genezy zjawiska. Załóżmy, że badamy stare składowisko odpadów komunalnych. Składowisko od dawna jest nieczynne, chcielibyśmy jednak wiedzieć, co tam wewnątrz zalega. Czy występują tam jakieś niebezpieczne substancje, które należałoby zneutralizować albo czy są składowane użyteczne składniki, które warto byłoby wydobyć i wykorzystać. Odwiercamy otwory, wykonujemy analizy zawartości i na tej podstawie chcielibyśmy wykonać mapę rozmieszczenia, np. żelaza. Zjawisko, które badamy, już z powodu swojej genezy nie pozwala na interpolację. Nie ma żadnych racjonalnych przesłanek do tego, aby twierdzić, że w środku odcinka między otworami, w których w jednym stężenie żelaza wynosi 10 g/m^3 , a w sąsiadującym aż 1000 g/m^3 , zawartość żelaza, jak oszacowana na podstawie interpolacji, wyniesie 505 g/m^3 . Zjawisko, które badamy, z natury rzeczy nie cechuje się regularnością. Składowisko to obiekt, w którym w dłuższym okresie uło-

kowano różne odpady pochodzące z różnych miejsc. Proces składowania odpadów cechuje się dużą przypadkowością albo losowością. W jednym miejscu będą leżały stare, stalowe taczki, których obecność wpłynie na dużą zawartość żelaza, a w drugim będą wyłącznie materiały ze szwalni, gdzie zawieruszyła się żelazna szpilka i zawartość tego pierwiastka owszem będzie, ale śladowa. Szacowanie zmienności przestrzennej zjawiska, które z natury rzeczy cechuje się dużą losowością, jest z góry skazane na porażkę. W tym przypadku muszą nam wystarczyć punktowe informacje z otworów, które można opisać syntetycznie za pomocą klasycznych narzędzi statystycznych, tj. średnia, odchylenie standardowe, zakres wartości, histogram. Więcej informacji z tych danych nie da się wycisnąć, bo interpolacja nie ma w tym przypadku najmniejszego sensu.

Co świadczy o tym, że można wykonać interpolację? Przecież przyglądamy się zjawisku wyłącznie za pomocą obserwacji punktowych. Otóż zjawisko, które odpowiada za możliwość interpolacji, określa się mianem autokorelacji. To istnienie pewnej zależności w przestrzennej zmienności parametru. Ale skąd mamy wiedzieć albo przypuszczać o jej istnieniu? Jedną z odpowiedzi już znamy. Jeżeli geneza interesującego fenomenu, np. przywołana już budowa pokładu węgla albo zawartość żelaza w odpadach komunalnych sugeruje, że zmienność przestrzenna zjawiska jest albo mocno zaburzona (np. uskokami), albo losowa na skutek samej genezy (np. taka powstała w wyniku składowania), to mamy prawo przypuszczać, że autokorelacja nie występuje, czyli nie da się wykonać wiarygodnej mapy izolinowej na drodze interpolacji. To są przypuszczenia, ale możemy je sprawdzić. W arsenale geostatystyki znajduje się narzędzie nadające się do tego celu. Jest nim wariogram. To pojęcie absolutnie kluczowe i bez niego nie byłoby żadnej geostatystyki. Sam opis wariogramu i jego zastosowanie będzie przedmiotem drugiego artykułu mojego trzyczęściowego cyklu.

Wróćmy jeszcze do autokorelacji. Jej obecność w zmienności badanego zjawiska jest warunkiem *sine qua non* dla interpolacji i w konsekwencji dla możliwości wykonania wiarygodnej mapy izolinowej. Powtórzę to, co już wyjawiałem. Komputer zawsze wyprodukuje mapę. Będzie ładnie wyglądać i możemy jej nadać barwną skalę, jednak jeżeli w zmienności ilustrowanego nią zjawiska nie występuje autokorelacja, to taka mapa będzie błędna, niewiarygodna i bezużyteczna.

AUTOKORELACJA – ILUSTRACJA ZJAWISKA NA PRZYKŁADZIE Z OBSZARU ANTROPOLOGII

Do wytłumaczenia, na czym polega zjawisko autokorelacji, wymyśliłem przykład z obszaru antropologii, na pozór całkiem odległego od nauk o Ziemi. Wyobraźmy sobie, że mamy bardzo ograniczoną wiedzę o świecie. Nie ma Internetu, nie ma nawet telewizji. Mimo to jednak wiemy, jak wyglądają mieszkańcy niektórych krajów. Wiadomo, że ludzie, nawet w obrębie tej samej nacji, bardzo różnią się między sobą, ale mają pewne cechy, które z dużą pewnością pozwalają ich przypisać do określonych grup etnicznych. Chodzi mi oczywiście o typowe cechy zewnętrzne. Trudno je nazwać, ale jest coś takiego w wyglądzie każdej nacji, że jak znajdujemy się na dużym europejskim lotnisku to od razu poznajemy, że tu na samolot czekają Skandynawowie, tam Rosjanie, tu jesteśmy

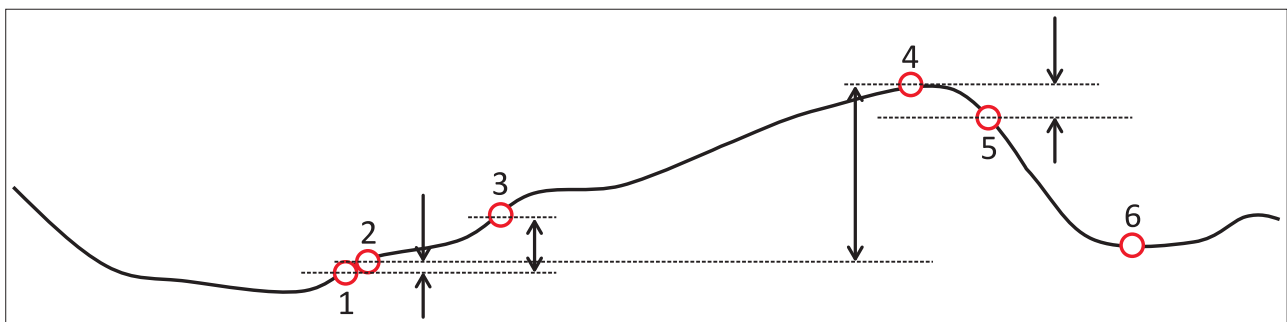
pewni, że samolot leci do Niemiec, a jak wreszcie dojdziemy do bramy, skąd odlatuje samolot do Warszawy, to polegając wyłącznie na wrażeniach wzrokowych, od razu coś nam mówi, że to nasi rodacy. Wróćmy jednak do mojego przykładu. Jako Polacy wiemy, jak wyglądają ludzie mieszkający w Polsce. Załóżmy, że znany nam jest również typowy wygląd mieszkańców Rosji w jej europejskiej części, np. w okolicach Moskwy. Nigdy nie byliśmy na Białorusi. Przypuszczamy jednak, że skoro Białoruś leży pomiędzy Rosją i Polską, to mieszkańcy tego kraju są trochę podobni zarówno do Polaków, jak i do Rosjan. I tu niewiele się mylimy, co więcej, można to sobie racjonalnie wytłumaczyć – podobną szerokością geograficzną, migracjami, a przede wszystkim relatywnie niewielką odległością w skali globu pomiędzy Warszawą, Mińskiem i Moskwą. Podobnie moglibyśmy szacować wygląd Słowaków na podstawie znajomości wyglądu Polaków i Węgrów. Wykorzystując obserwacje poczynione w przestrzeni możemy dokonać interpolacji, ponieważ zjawisko, którym jest wygląd zewnętrzny, wykazuje istnienie autokorelacji. Na podstawie tego konkretnego przypadku możemy stwierdzić, że zasięg autokorelacji tego zjawiska jest większy niż odległość między Warszawą i Moskwą. Powtórzmy to samo doświadczenie, ale przyjmując nieco większe odległości. Nasze obserwacje to znany nam wygląd Polaków oraz mieszkańców Japonii. Na tej podstawie przypuszczamy, że na wygląd mieszkańców środkowej Syberii składają się po części cechy jednych, jak i drugich. I tu okazuje się, że cechy wyglądu zewnętrznego mieszkańców Syberii, Kazachów czy Buriatów, nie są mieszkanką cech Polaków i Japończyków. Dlaczego? Bo mimo, że istnieje autokorelacja w wyglądzie ludzi zamieszkujących sąsiadujące tereny, to jednak zasięg tej autokorelacji jest znacznie mniejszy niż odległość między Warszawą i Tokio. Na podstawie tak odległych obserwacji nie można wnioskować o tym, co jest pomiędzy. *Ergo* nie możemy wykonać interpolacji tak, jak to zrobiliśmy w przypadku obserwacji bliższych, w przykładzie opisującym Polaków, Białorusinów i Rosjan albo Polaków, Słowaków i Węgrów. I znowu można to racjonalnie wytłumaczyć. Migracje na tak duże odległości, jeżeli w ogóle nastąpiły w historii, to związana z nimi wymiana genów musiała być bardzo ograniczona. Ponadto pomiędzy Polską i Japonią poza samą odległością występuje wiele trudnych do przebycia geograficznych barier, które skutecznie powstrzymywały migracje, wymianę genów i tym samym unifikację wyglądu.

W przykładzie pierwszym, tj. Warszawa–Mińsk–Moskwa, na podstawie cech wyglądu Polaków i Rosjan wykonałem prognozę wyglądu Białorusinów. Dokonałem oszacowania na podstawie bliskich obserwacji. W analizowanym zjawisku

występuje autokorelacja, a jej zasięg jest większy niż odległość między Warszawą i Moskwą. Gdybym jednak na podstawie cech wyglądu Polaków i Japończyków chciał wnioskować o wyglądzie Kazachów, to moje oszacowanie cechowałoby się dużym błędem. To znaczy, że odległość między Warszawą i Tokio wykracza poza zasięg autokorelacji badanego zjawiska.

Odnieśmy to z powrotem do geologii, gdzie geostatystyka ma szersze zastosowanie. Złoża kopalin albo skały powstają w określonych warunkach geologicznych. Złoże węgla kamiennego powstaje w wyniku najpierw wzrostu roślin, gromadzenia biomasy i potem osadzania się materii organicznej. Ta materia ulega podtapianiu w okresach transgresji morskich, po czym jest przysypywana osadami ilastymi, piaszczystymi itd., aż po wielu milionach lat powstaje węgiel o określonych właściwościach. Możemy zaufać intuicji i stwierdzić, że na określonym obszarze warunki wegetacji roślin i sedimentacji biomasy były podobne. Wraz ze zwiększającą się odległością te warunki mogły się zmieniać, tak jak różne są i dzisiaj obserwowane warunki klimatyczne w zależności od miejsca na mapie. Stąd bliskie obserwacje są do siebie bardziej podobne, a dalsze mniej. Wraz ze zwiększającą się odległością podobieństwo przestaje być zależne od odległości, staje się losowe. Gdy dokonamy pomiarów miąższości złoża węgla poza zasięgiem autokorelacji, to ta miąższość może być zupełnie inna, ale równie dobrze może być całkiem podobna. Zanika obserwowana na małych dystansach zależność między odległością i podobieństwem. Odległość, dla której cechy zjawiska (tu złoża) zatracają współzależność i stają się wzajemnie niezależne albo wzajemnie wobec siebie losowe, nazywa się zasięgiem autokorelacji. Zasięg autokorelacji da się zmierzyć za pomocą wariogramu, narzędzia, o którym już wspominałem.

Jeszcze o samej nazwie. Każdy absolwent szkoły średniej wie, co znaczy korelacja. To statystyczna zależność dwóch różnych parametrów tych samych obiektów badań. Niech takim obiektem będzie grupa ludzi. Można zauważyć, że wraz ze zwiększającym się wzrostem człowieka najczęściej, czyli statystycznie, rośnie również i jego masa. To znaczy, że wzrost człowieka i jego masa są ze sobą statystycznie skorelowane. W przypadku autokorelacji chodzi o istniejącą zależność w obrębie jednego parametru, ale skorelowanie dotyczy podobieństwa obserwacji w przestrzeni. Wraz z rosnącą odległością między obserwacjami tego samego parametru maleje ich wzajemne podobieństwo albo im bliższe obserwacje tym bardziej są do siebie podobne, im dalsze obserwacje, tym bardziej się od siebie różnią. Statystycznie oczywiście. Interesuje nas jeden



Ryc. 3. Ilustracja zjawiska autokorelacji na przykładowym profilu geologicznym
Fig. 3. Illustration of the autocorrelation phenomenon on an example geological profile

parametr i jego skorelowanie w przestrzeni, dlatego nazywamy to autokorelacją.

Zjawisko autokorelacji zilustrowałem przykładowym profilem (ryc. 3). Obserwacje 1 i 2 zostały wykonane bardzo blisko i dlatego pod względem wartości niewiele się różnią. Wraz ze wzrostem odległości między parami obserwacji rośnie również różnica wartości, np. pary 1 i 3 oraz 4 i 5. Pary obserwacji bardziej odległe, np. 2 i 4 albo 4 i 6 różnią się już znacznie. Widoczna jest zależność między wzrostem odległości pomiędzy parami obserwacji a wzrostem różnicy wartości. Dla kontrastu obserwacje 2 i 6, mimo że bardzo odległe, pod względem wartości są niemal tożsame. W zależności od badanego zjawiska w pewnej odległości zależność zanika i wartości są w stosunku do siebie losowe – albo bardzo różne, albo całkiem podobne, ale są od siebie niezależne. Interpolacja między obserwacjami 1 i 4 albo 4 i 6 dałaby wynik bliski rzeczywistości, ale między 1 i 6 już nie.

PODSUMOWANIE

Staralem się przedstawić kilka najbardziej podstawowych pojęć i obszarów zastosowania geostatystyki, a także zmotywować do jej stosowania. Chciałem odpowiedzieć na pytanie, dlaczego w ogóle warto się zagłębić w tej dziedzinie. Wiele miejsca poświęciłem interpolacji i autokorelacji. Staralem się wykazać, że nie zawsze interpolacja i rysowanie mapy ma sens. W kolejnym artykule wiele miejsca poświęcę najważniejszemu pojęciu z dziedziny geostatystyki. Będę opisywał wariogram. To narzędzie geostatystyczne nie do przecenienia, o wielorakim zastoso-

waniu. Coś jak kombinerki w warsztacie. Z kształtu wariogramu można wysnuć bardzo ważne wnioski na temat analizowanego zjawiska przestrzennego. Interpolacja i rysowanie map muszą być poprzedzone analizą wariogramu, bo właśnie z jego kształtu można odczytać, czy badane zjawisko przestrzenne wykazuje autokorelację czy nie, a jeśli tak, to jaki jest jej zasięg. Wariogram będzie też niezbędny do dalszego kroku, jakim jest interpolacja metodą kriginu. O tym jednak w ostatnim artykule trzyczęściowego cyklu.

LITERATURA

- CYRUL T. 1992 – Elementy prognozowania wyrzutów skalno-gazowych i gazonośności w ujęciu geostatystycznym. Zesz. Nauk. AGH, Górnictwo, 160.
- ISAAKS E., SRIVASTAVA R.M. 1989 – An Introduction to Applied Geostatistics. Oxford University Press, New York.
- JOURNEL A.C., HUIJBREGTS CH.J. 1978 – Mining Geostatistics. Academic Press, London.
- MATHERON G. 1962 – Traité de géostatistique appliquée. Memoires du BRGM, Paris.
- MENZ J. (red.) 2000 – Angewandte Geostatistik in Bergbau, Geologie, Geophysik, Geodäsie und Umweltschutz. CPress-Verlag, Mathematische Geologie, 5.
- MUCHA J., NIEĆ M., WASILEWSKA M., SOBCZYK E., SAŁUGA P., 2007 – Dokładność szacowania zasobów węgla kamiennego jako element oceny ryzyka inwestycyjnego. Monografie Akademii Górniczo-Hutniczej im. St. Staszica w Krakowie.
- NAMYSŁOWSKA-WILCZYŃSKA B. 1986 – Zastosowanie metody wariogramów do oceny zmienności złoża rud miedzi. Pr. Nauk. Inst. Geotechniki Politechniki Wrocławskiej nr 48, Konferencje nr 20, Wrocław.

Praca wpłynęła do redakcji 8.07.2024 r.
Akceptowano do druku 10.07.2024 r.