

## WYBRANE ZAGADNIENIA WALIDACJI METOD ROZPOZNANIA PODŁOŻA GRUNTOWEGO

### Streszczenie

Niniejsza praca ma na celu oderwanie od spojrzenia na wzory korelacyjne w sposób ścisły. Wykorzystywanie takich wzorów, a także ich porównywanie oraz walidacja powinny brać pod uwagę niepewność jaką zawsze obarczone są wyniki badań. Skupiono się przede wszystkim na problematyce stosowania w metodach rozpoznania podłoża gruntowego różnych wzorów korelacyjnych dotyczących oszacowania tego samego parametru. Nie skupiono się jednak na analizie rozbieżności istniejących wzorów, a na przyczynach tych rozbieżności, ocenie istotności różnic i ich znaczeniu w kontekście walidacji metod rozpoznania podłoża gruntowego. Dokonano symulacji eksperymentów polegających na uzyskaniu wzorów korelacyjnych. Następnie omówiono ideowo sposób postępowania w ocenie dokładności uzyskanych metod. Za miarę dokładności przyjęto szerokość przedziału ufności oszacowania pojedynczego wyniku pomiaru. Przedstawiona symulacja i porównanie wyników miały na celu zobrazowanie istotności różnic w wynikach uzyskiwanych na podstawie poszczególnych korelacji.

### WSTĘP

Walidacja metod rozpoznania podłoża gruntowego polega na ustaleniu parametrów charakteryzujących sprawność działania i ograniczenia metod badawczych oraz sprawdzeniu ich przydatności do określonych celów tj. rozpoznania podłoża gruntowego dla budownictwa drogowego. Ze względu na specyficzny rodzaj danych opracowanie procedury walidacji wymaga powiązania wielu zagadnień związanych z metodyką prowadzonych badań. W przypadku niektórych metod badawczych pomierzone cechy gruntu wykorzystuje się bezpośrednio do obliczenia poszukiwanego parametru gruntu. Na przykład na podstawie laboratoryjnego badania wskaźników porowatości (maksymalnego  $e_{max}$ , minimalnego  $e_{min}$  i naturalnego  $e_n$ ) wyznacza się stopień zagęszczenia gruntu  $I_D$  [1, 2, 5]. Ten sam parametr można często wyznaczyć inną metodą, w sposób przybliżony, posługując się wzorami korelacyjnymi [3, 7, 8]. Na przykład wyżej wymieniony stopień zagęszczenia można wyznaczyć pośrednio na podstawie wyników sondowań statycznych CPT, lub dynamicznych DPSH. Generalnie walidacja będzie polegała na zastosowaniu odpowiednich narzędzi statystycznych do sprawdzenia wiarygodności metod przybliżonych (określanych jako metody walidowane) poprzez porównanie ich wyników z wynikami uzyskanymi metodami definiującymi dany parametr (określane jako metody odniesienia). Jak jednak wiadomo, istnieje więcej niż jeden wzór korelacyjny pozwalający na określenie stopnia zagęszczenia na podstawie określonego rodzaju sondowania. Nasuwa się zatem pytanie: który wzór należy zastosować? Należy podkreślić, że niniejsza praca skupia się na statystycznym ujęciu przyczyn różnic we wzorach korelacyjnych tego samego parametru oraz znaczeniu tych różnic w procedurze walidacyjnej. Przedstawiony przykład nie odwołuje się do konkretnej metody, lecz opisuje problem w możliwie ogólny i uproszczony sposób.

### 1. RÓŻNICE WE WZORACH KORELACYJNYCH

Na początku pracy przytoczono znany przykład parametru ( $I_D$ ), który można wyznaczyć różnymi metodami wskaźnikowymi, stosując różne korelacje. Jak wynika z przedstawianego w literaturze porównania [9], wyniki uzyskane różnymi metodami nie pokrywają się wzajemnie. Podobna sytuacja ma miejsce również podczas wyznaczania innych parametrów gruntu metodami wskaźnikowymi.

Przyczyny mogą być różne. Badania prowadzące do opracowania poszczególnych wzorów kalibracyjnych były prowadzone w różnych warunkach. Różne czynniki, niebrane we wzorze pod uwagę, miały wpływ na wyniki (lokalizacja, wilgotność gruntu, technik wykonujący badanie itp.). W związku z brakiem możliwości ilościowego uwzględnienia powyższych czynników należy uznać, że wyniki badań obarczone są błędem losowym zależnym od tych czynników. Znajomość rozkładu błędów pozwoliłaby wnioskować o wiarygodności danej korelacji. Niestety, w przypadku kiedy nie jest możliwe przebadanie całej populacji, można jedynie szacować parametry rozkładu na podstawie pewnej ograniczonej próby badań, co wpływa na dalsze możliwości wnioskowania [10]. Dlatego też badając jedną populację kilka razy można uzyskać różne wzory korelacyjne. Po opracowaniu wyników badacze zwykle udostępniają je w postaci gotowych wzorów bez informacji o przedziale dokładności metody (miarą dokładności może być na przykład szerokość przedziału ufności). Nawet jeżeli informacja taka znajduje się w publikacji źródłowej to zwykle zanika w trakcie jej dalszego upowszechniania (w zaleceniach i wytycznych). W efekcie może funkcjonować kilka wzorów empirycznych opisujących to samo zjawisko, które w swojej ścisłej formie dają różne wyniki. Po uwzględnieniu rozrzutu wyników okazałoby się jednak, że w znacznym zakresie się pokrywają. Aby lepiej zilustrować to zagadnienie przeprowadzono poniższą symulację.

### 2. PRZYKŁAD OBLICZENIOWY

Praca ma na celu jedynie przedstawienie wpływu ograniczonej liczności próby na uzyskiwane wzory korelacyjne oraz na dokładność oszacowania rozbieżności między metodą odniesienia i metodą walidowaną. Aby uczynić przykład bardziej czytelnym zostaną poczynione pewne uproszczenia. Po pierwsze eksperyment zostanie przeprowadzony poprzez symulację wyników o znanym rozkładzie i z góry założonej zależności. Ponadto zostaną zastosowane proste, liniowe wzory korelacyjne. Dlatego też analiza zostanie przeprowadzona w oderwaniu od konkretnej metody, czy parametrów gruntu. Założmy, że istnieje ścisła zależność pomiędzy parametrem gruntu  $M$ , wyznaczanym w badaniu I oraz poszukiwanym parametrem gruntu  $G$ , wyznaczanym w badaniu II. Zależność dana jest wzorem (1).

$$G = 0,5 + 1,5M \quad (1)$$

Parametr G jest bezpośrednio wykorzystywany w projektowaniu, jednak badanie I jest kosztowne i czasochłonne. Badacz spodziewający się, że istnieje zależność pomiędzy parametrami M i G, może zaproponować sposób wyznaczania parametru G na podstawie wyników badania II. Wówczas badanie I będzie metodą wzorcową dla parametru G. Natomiast badanie II będzie metodą wskaźnikową, gdyż w jego wyniku otrzymuje się parametr M, a dopiero na podstawie korelacji parametr G. Korzyścią jaką można otrzymać jest zmniejszenie kosztów i czasu badań w przyszłości, jeżeli metoda wskaźnikowa okaże się tańsza i jednocześnie dostatecznie dokładna. Zaprojektowanie metody wskaźnikowej wymaga przeprowadzenia badań, w których zostaną pomierzone obydwa parametry gruntu. Ponieważ na wyniki badań wpływ mają różne czynniki ukryte, lub celowo pomijane w analizie, rezultaty obciążone są błędem  $\varepsilon_i$ . Na potrzeby symulacji badań założono, że błąd ma rozkład normalny o znanych parametrach (średni błąd wynosi 0, a odchylenie standardowe  $\sigma=0,2$ ). Zatem otrzymane wyniki będą postaci (2).

$$G_i = 0,5 + 1,5M_i + \varepsilon_i \quad (2)$$

Należy przypomnieć, że badacz projektujący metodę wskaźnikową nie zna zależności opisanej równaniem 2, a jedynie będzie próbował ją ustalić na podstawie ograniczonej liczby badań. Najprostszą i jednocześnie bardzo popularną zależnością jest zależność liniowa i zwykle od niej zaczyna się poszukiwanie związku pomiędzy czynnikami. Analizą pozwalającą na określenie takiej liniowej zależności na podstawie losowych wyników badań może być regresja liniowa (normalna metoda najmniejszych kwadratów) [4]. Wówczas poszukujemy prostej postaci:

$$y = a \cdot x + b \quad (3)$$

gdzie:

x - zmienna objaśniająca,

y - przewidywana wartość zmiennej objaśnianej,

a, b - współczynniki regresji dobrane tak, aby suma kwadratów błędów (S) pomiędzy wartościami przewidywanymi i mierzonymi była najmniejsza.

$$S = \sum_{i=1}^N e_i^2 = \sum_{i=1}^N (y_i - y_i')^2 \quad (4)$$

S - suma kwadratów błędów,

y<sub>i</sub> - przewidywana przez model wartość zmiennej objaśnianej odpowiadająca i-tej wartości zmiennej,

y<sub>i</sub>' - wynik pomiaru odpowiadający i-tej wartości zmiennej,

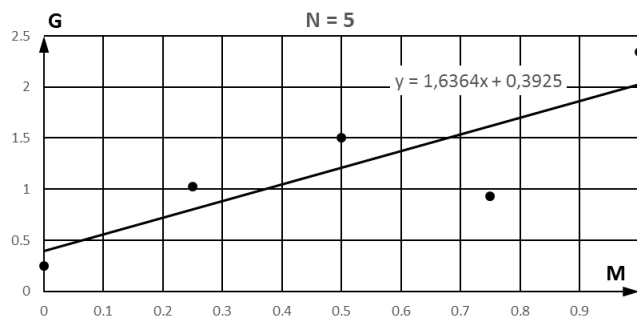
N - liczba wyników pomiarów.

Miarą jakości dopasowania otrzymanego metodą najmniejszych kwadratów wzoru do danych doświadczalnych jest współczynnik R<sup>2</sup>. Współczynnik ten przejmuje wartości z przedziału <0,1>. Gdzie wartość 1 oznacza najlepsze dopasowanie. Współczynnik ten wyznaczony jest przy pomocy wzoru:

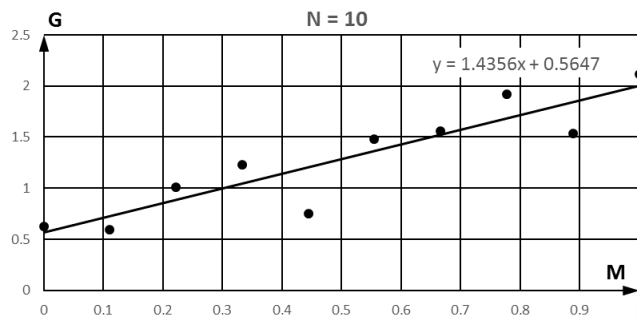
$$R^2 = \frac{\sum_{i=1}^N (y_i - \bar{y})^2}{\sum_{i=1}^N (y_i' - \bar{y})^2} \quad (5)$$

$\bar{y}$  - średnia wartość wyników dla zmiennej objaśnianej.

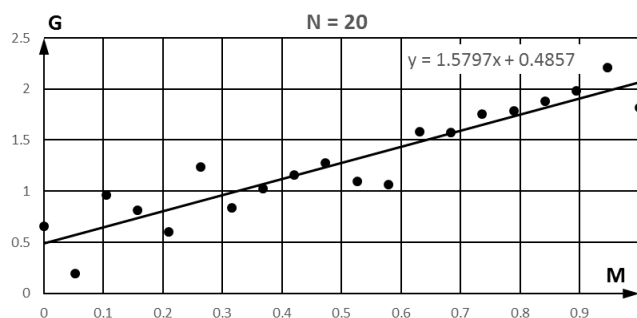
Dokonano symulacji badań mających na celu ustalenie przybliżonej zależności między parametrem M wyznaczanym metodą przybliżoną a parametrem G wyznaczanym metodą wzorcową. Symulacja polegała na dobraniu 2N wyników badań (jedno badania dla G i jedno dla M, teoretycznie w tym samym miejscu) obarczone losowym błędem zgodnie z powyższymi założeniami. Symulację przeprowadzono 4-krotnie uzyskując wykresy rozrzutów składające się z N punktów. Wyniki przedstawiono na rys. 1-4.



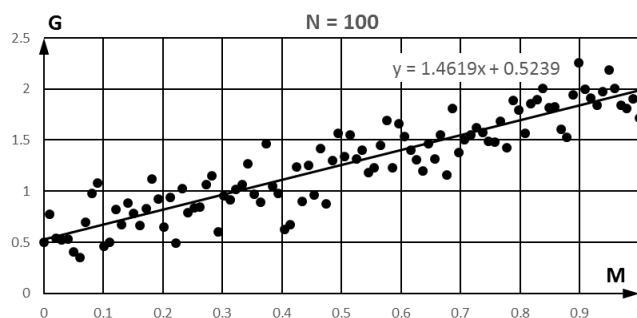
Rys. 1. Wykres przedstawiający prostą uzyskaną dla populacji równej 5 pomiarom



Rys. 2. Wykres przedstawiający prostą uzyskaną dla populacji równej 10 pomiarom



Rys. 3. Wykres przedstawiający prostą uzyskaną dla populacji równej 20 pomiarom



Rys. 4. Wykres przedstawiający prostą uzyskaną dla populacji równej 100 pomiarom

Otrzymane wyniki zestawiono również w tab. 1. Współczynnik kierunkowy prostej został najdokładniej wyznaczony dla próby składającej się ze 100 pomiarów, parametr ten został dokładnie wyznaczony dla próby składającej się z 10 pomiarów niż z pomiarów 20. Stosunkowo najmniej dokładna okazała się próba dla N = 5, zarówno pod względem otrzymanego współczynnika determinacji (R<sup>2</sup>), jak i estymatora odchylenia standardowego (S<sub>d</sub>). Współczynnik b jest najbliższy wartości właściwej w przypadku próby składającej się z 20 pomiarów. Warto zwrócić uwagę na wartość parametru R<sup>2</sup>,

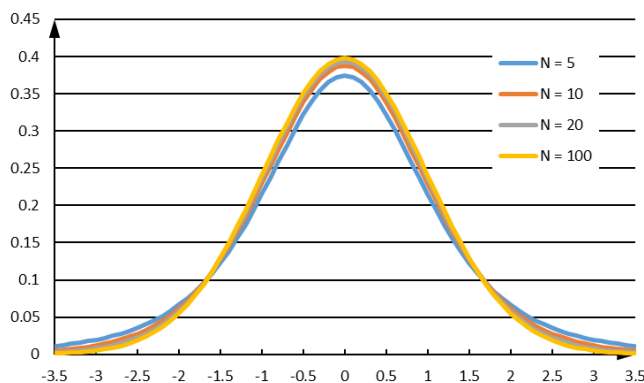
który obrazuje jak dokładnie otrzymany wzór opisuje dane doświadczalne. Ma on największą wartość dla 20 pomiarów. Zbliżone duże wartości parametru osiągamy również dla  $N = 10$  i  $N = 100$ . Jeżeli wartość parametru  $R^2$  przekracza 0,8 dopasowanie można uznać za dobre.

**Tab. 1.** Otrzymane wartości parametrów dla różnej ilości pomiarów.

| Lp. | N   | a    | b    | $S_d$ | R    |
|-----|-----|------|------|-------|------|
| 1   | 5   | 1,64 | 0,39 | 0,38  | 0,69 |
| 2   | 10  | 1,44 | 0,56 | 0,21  | 0,82 |
| 3   | 20  | 1,58 | 0,49 | 0,19  | 0,86 |
| 4   | 100 | 1,46 | 0,52 | 0,19  | 0,83 |

### 3. ANALIZA WPLYWU LICZEBNOŚCI PRÓBY NA DOKŁADNOŚĆ OSZACOWANIA ROZBIEŻNOŚCI MIĘDZY METODĄ ODNIESIENIA I METODĄ WALIDOWANĄ

Należy zwrócić uwagę na to, że zgodność wyżej otrzymanych wzorów korelacyjnych z rzeczywistą zależnością jest wynikiem pewnego eksperymentu losowego. Nie zależała ona ściśle od liczby pomiarów. Może być tak, że w przypadku mniejszej liczby pomiarów otrzymamy dokładniejszy wzór. Jednak nie jesteśmy tego w stanie ocenić gdyż w rzeczywistości ścisła zależność określona równaniem 1 nigdy nie jest znana. Intuicją podpowiada, że oszacowanie na podstawie większej liczby pomiarów jest lepsze. Można to uzasadnić w ten sposób, że wraz ze zwiększeniem liczby pomiarów zwiększa się prawdopodobieństwo, że otrzymane parametry będą zbliżone do poszukiwanych. Dokładność oszacowania rozbieżności między wynikami uzyskanymi metodą odniesienia i metodą można ocenić stosując estymację przedziałową. Wówczas ilościowy wpływ liczby pomiarów na jakość wyników będzie uwzględniony przez zastosowanie odpowiedniego rozkładu. W tym celu możemy zastosować rozkład t-Studenta, który zbiega do rozkładu normalnego wraz ze wzrostem liczby stopni swobody  $n$  zależnej od liczby pomiarów. Na rys. 5 przedstawiono wykresy gęstości prawdopodobieństwa standaryzowanych rozkładów t-Studenta otrzymanych dla  $N = (5, 10, 20, 100)$  wyników pomiarów.

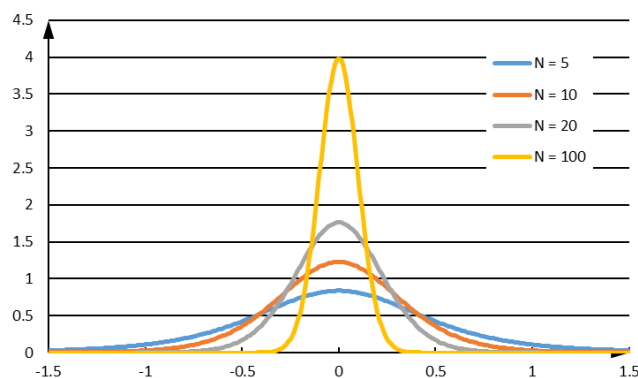


**Rys. 5.** Rozkład t-Studenta różnic poszczególnych oszacowań dla rozważanych wielkości prób

W tym przypadku zostało zastosowane wnioskowanie bezpośrednie, to znaczy, że na podstawie parametrów rozkładu wnioskujemy o przedziale, w którym znajdzie się przyszły wynik [6]. Szukamy takiego przedziału wartości zmiennej  $y$ , że dla danego  $x$  wynik pomiaru znajdzie się w tym przedziale z określonym prawdopodobieństwem (na przykład 95%).

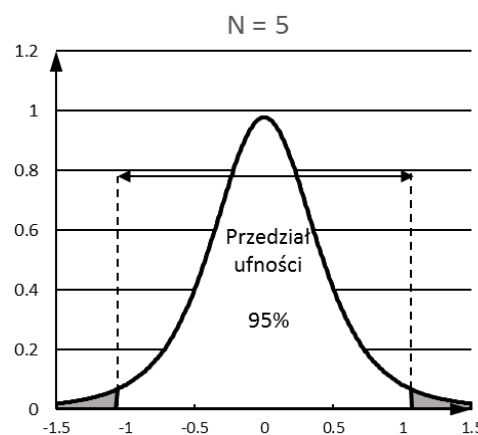
Inne podejście należy zastosować, aby ocenić czy średnia z wyników (w tym przypadku różnice pomiędzy wartością przewidywaną, a zmierzoną) znajdzie się w określonym przedziale. W takim

przypadku rozkład będzie dużo bardziej zależny od liczby pomiarów, co przedstawiono na rys. 6.

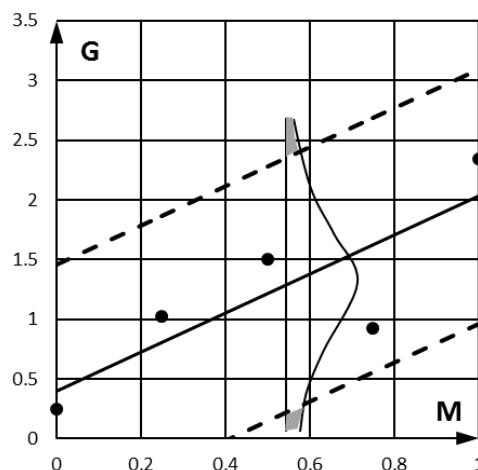


**Rys. 6** Rozkład t-Studenta średnich różnic poszczególnych oszacowań dla rozważanych wielkości prób

Na rys. 7 zilustrowano interpretację 95% przedziału ufności w przypadku wnioskowania bezpośredniego.

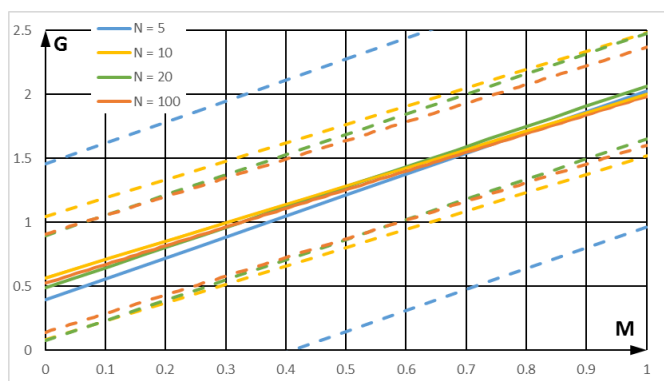


**Rys. 7** Przedział ufności na wykresie gęstości prawdopodobieństwa



**Rys. 8** Prosta regresji dla  $N = 5$

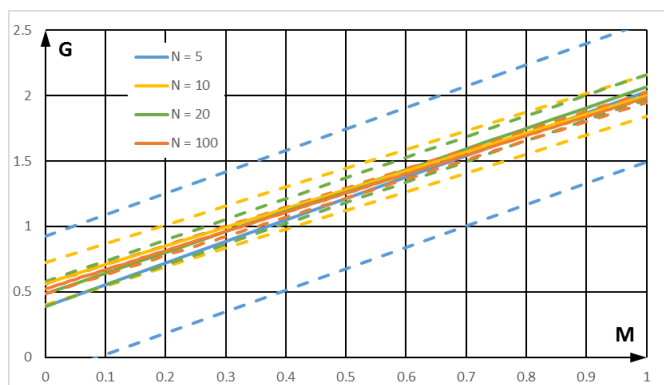
Zakładając, że błąd oszacowania nie zależy od wartości zmiennej  $x$ , a w tym konkretnym przykładzie parametru  $M$ , można przyjąć, że granice przedziału ufności będą liniami równoległymi do prostej regresji. W ten sposób można porównać ze sobą przedziały ufności uzyskane na podstawie wyników badań o różnej liczności. Porównanie przedstawiono na rys. 9.



**Rys. 9** Uzyskane proste regresji i przedziały ufności pojedynczych pomiarów dla rozważanych liczebności prób

Zestawienie przedziałów ufności pozwala na lepsze porównanie uzyskanych wzorów korelacyjnych. Z wykresów liniowych przedstawiających przewidywane wartości trudno odczytać czy wyniki różnią się istotnie. Natomiast analizując przedziały ufności widać w jakim zakresie się pokrywają. Jeżeli estymator odchylenia standardowego jest podobny, to większą pewność wyników uzyskujemy w przypadku większej liczby próbek (wykres gęstości prawdopodobieństwa rozkładu t-Studenta jest bardziej spłaszczony przy mniejszej liczbie pomiarów). Jak widać z rys. 5 przy większej liczbie wyników pomiarów wykresy gęstości prawdopodobieństwa nie różnią się znacząco i decydujący wpływ na niepewność oszacowania będzie miało odchylenie standardowe (a właściwie jego estymator).

Analizując położenie średniej różnicy pomiędzy pomiarem a oszacowaniem otrzymamy przedziały dużo bardziej zależne od liczby pomiarów. zilustrowano to na rys. 10, w sposób analogiczny do przedstawionego na rys. 8, stosując jednak rozkłady otrzymane dla średnich a nie wyników pojedynczych pomiarów. Przedstawione podejście jest dużym uproszczeniem, gdyż szerokość przedziału ufności nie jest w tym przypadku stała (węższy w środku, szerszy na końcach).



**Rys. 10** Uzyskane proste regresji i przedziały ufności pojedynczych pomiarów dla rozważanych liczebności prób

Podjęcie się walidacji metod badawczych, czyli oceny ich przydatności do poszczególnych celów, należy ocenić ich dokładność. A będąc bardziej precyzyjnym należałoby raczej mówić o niepewności pomiaru [6]. Biorąc pod uwagę wyżej omówione narzędzia statystyczne oraz dostępne w literaturze informacje o metodach rozpoznania podłoża gruntowego, można wyciągać wnioski, że podstawę metodyki powinno stanowić podejście pierwsze, tzn. ocena przedziału ufności pojedynczego pomiaru. Zastosowanie podejścia drugiego jest ograniczone, gdyż do oceny przedziału ufności położenia średniej przewidywanej wartości potrzebne są informacje jak dany wzór korelacyjny był konstruowany (liczba

wyników pomiarów, ich rozrzut itd.). Zadanie to służy ocenie skonstruowanej relacji i należy do badacza, który ją opracował. Natomiast mając jedynie ścisły wzór, możemy próbować odtworzyć informację o jego dokładności wykonując dodatkowe pomiary. Pomiary te będą polegały na zmierzeniu wartości danego parametru gruntowego w tym samym miejscu metodą odniesienia i metodą walidowaną. Uzyskane różnice zostaną poddane analizie statystycznej wg podejścia pierwszego. Pozwoli to w sposób przybliżony odpowiedzieć na pytanie: „O ile wynik uzyskany metodą walidowaną (przybliżoną) może różnić się od wartości otrzymanej metodą odniesienia (referencyjną)?”. Wykonanie tej procedury dla różnych metod i wzorów pozwoli na ich uszeregowanie pod względem przydatności.

## PODSUMOWANIE

W niniejszej pracy dokonano symulacji eksperymentów mających na celu uzyskanie wzorów korelacyjnych pozwalających na oszacowanie wyników pomiarów parametru G, na podstawie badania parametru M. Opracowano cztery wzory korelacyjne na podstawie różnej liczby wyników badań ( $N=(5, 10, 20, 100)$ ). Zgodność wzoru z „prawdziwą” zależnością opisującą daną relację jest wynikiem eksperymentu losowego, więc nie zależy wprost od liczby wyników pomiarów. Nie znając właściwej zależności można jedynie wnioskować o niepewności jaką są one obciążone. Wówczas liczba pomiarów ma większe znaczenie. Na podstawie wyników symulacji sporządzono i porównano przedziały ufności poszczególnych wzorów. Takie porównanie ilustruje dużo lepiej jak istotnie różnią się wyniki uzyskiwane na podstawie poszczególnych korelacji. Dlatego też na uzyskany wzór korelacyjny nie należy patrzeć jak na ścisłe rozwiązanie, a jak na jedno z wielu możliwych mieszczących się w danym przedziale z określonym prawdopodobieństwem. Ponadto nasuwa się kolejny wniosek, że nowe wzory korelacyjne, w miarę możliwości powinny być upowszechniane wraz z informacją o rozrzucie/przedziale ufności. Wówczas, ocena niepewności związanej z wynikami badań uzyskanymi określoną metodą byłaby znacznie ułatwiona.

Podobne narzędzia można wykorzystać w procedurze walidacji metod pomiarowych. Nie znając dokładności poszczególnych metod można określić różnice uzyskiwane z korelacji (metoda walidowana) i metody odniesienia. Następnie, na podstawie analizy rozkładu tych różnic oszacować niepewność wyników pomiarów uzyskiwanych daną metodą. Tego typu informacje są dobrą podstawą do porównania poszczególnych metod oraz oceny ich przydatności.

## BIBLIOGRAFIA

1. Atkinson J., The Mechanics of Soils and Foundations. 2nd Ed. Routledge, London, 2007.
2. Head K.H., Manual of Soil laboratory testing, Volume 1, Soil classification and compaction test. London, 1992.
3. Kraszewski C., Szpikowski M., Badanie i ustalenie zależności korelacyjnych dla oceny stanu zagęszczenia i nośności gruntów niespoistych płytą dynamiczną, Instrukcja stosowania płyty dynamicznej do oceny stanu gruntów niespoistych w budowlanych warstwowo, Wyd. IBDiM, Warszawa, 2005.
4. Mazerski J., Statystyczna analiza wyników doświadczalnych, Wydawnictwo Malamut, Warszawa, 2009.
5. Myślińska E., Laboratoryjne badania gruntów i gleb, Wyd. UW. Warszawa, 2010.
6. Pajzderski P., Iglantowicz T., Dokładność metody pomiarowej a niepewność pomiaru-próba analizy, Pomiary, Automatyka, Kontrola, 46: 14-16, 2000



7. Patakiewicz M., Zastosowanie wskaźników krzywej zagęszczalności jako charakterystyk zagęszczalności gruntów niespoistych. Inżynieria Morska i Geotechnika (4), s. 366-370, 2012.
8. Pisarczyk S., Zależność między wskaźnikiem zagęszczenia i stopniem zagęszczenia gruntów niespoistych, Inżynieria i Budownictwo nr (5), 1975.
9. Ura M., Tarnawski M. Interpretowanie stopnia zagęszczenia gruntów niespoistych na podstawie wyników sondowań statycznych i dynamicznych, Przegląd Geologiczny, 62(10/2): 715-720, 2014.
10. Wątroba J., Praktyczne aspekty szacowania liczebności próby w badaniach empirycznych, Czytelnia StatSoft Polska, 2013.

### CERTAIN ISSUES OF THE SUBSOIL RECOGNITION METHODS VALIDATION

#### *Abstract*

*This papers goal is to show that correlation formulas should not be treated as a strict solution. The use of such formulas as well as their comparison and validation should take into account the uncertainty that is always associated with the experimental results. The*

*focus was mainly on the problem of the use of correlation formulas in the subsoil identification methods. Especially different correlation formulas concerning estimation of the same parameter. However, the analysis of differences of the existing formulas is not the main subject, but the reasons of those differences, the significance of differences and their importance in the context of the validation of subsoil recognition methods. A simulation of experiments involving obtaining the correlation formulas was conducted. Next, the concept of methods accuracy evaluation procedure was discussed. The width of confidence interval of an individual test result estimation was used to accuracy evaluation. Presented simulation and comparison of the results aimed to illustrate the significance of differences in the results obtained on the basis of individual correlations.*

dr inż. **Karol Brzeziński** – Wydział Inżynierii Lądowej; al. Armii Ludowej 16, 00-637 Warszawa, [k.brzezinski@il.pw.edu.pl](mailto:k.brzezinski@il.pw.edu.pl)

dr **Maciej Maślakowski** – Wydział Inżynierii Lądowej; al. Armii Ludowej 16, 00-637 Warszawa, [m.maslakowski@il.pw.edu.pl](mailto:m.maslakowski@il.pw.edu.pl)

mgr inż. **Paweł Tutka** – Wydział Inżynierii Lądowej; al. Armii Ludowej 16, 00-637 Warszawa, [p.tutka@il.pw.edu.pl](mailto:p.tutka@il.pw.edu.pl)