WIERTNICTWO NAFTA GAZ • TOM 25 • ZESZYT 1 • 2008

Zbigniew Fąfara*

ANALIZA ZMIENNOŚCI PRĘDKOŚCI FILTRACJI WĘGLOWODORÓW W GRUNCIE METODĄ ANOVA**

1. WSTĘP

Wykorzystując przygotowane laboratoryjne stanowisko pomiarowe do badania migracji substancji ropopochodnych w fizycznych modelach ośrodka gruntowo-wodnego [1], określono prędkość filtracji pionowej trzech wybranych produktów naftowych. Uzyskane wartości charakteryzują się dużym zróżnicowaniem [1]. Można je wyjaśnić na podstawie ogólnie znanego prawa Darcy'ego – prędkość filtracji pionowej płynu w ośrodku porowatym pod wpływem grawitacji jest wprost proporcjonalna do współczynnika przepuszczalności fazowej tego płynu i jego gęstości objętościowej oraz odwrotnie proporcjonalna do współczynnika lepkości dynamicznej. Interesującym zagadnieniem stała się jednak ocena istotności wpływu właściwości przygotowanych fizycznych modeli gruntu i właściwości użytych mieszanin węglowodorów na zmienność zebranej populacji danych. W badaniach wykorzystano jednoczynnikową analizę wariancji. Jej wnioski mogą być podstawą do przeprowadzenia w dalszej kolejności analizy korelacji i regresji.

2. PODSTAWY TEORETYCZNE JEDNOCZYNNIKOWEJ ANALIZY WARIANCJI

Analiza wariancji (ANOVA – *ANalysis Of VAriance*) jest jedną z podstawowych metod statystycznych pozwalających na sprawdzenie, czy rozważane czynniki klasyfikujące wywierają wpływ na obserwowane wyniki pomiaru zmiennej objaśnianej. W najprostszej wersji bada się wpływ tylko jednego czynnika klasyfikującego (klasyfikacja jednoczynnikowa). Rozwinięcie metody (klasyfikacja wieloczynnikowa) umożliwia równoczesne badanie wpływu wielu czynników na wartości zmiennej objaśnianej, wraz z efektem wzajemnej korelacji czynników. Analiza wariancji jest szczególnie przydatna przy próbie opisu wpływu jakościowych czynników klasyfikujących [5, 6].

* Wydział Wiertnictwa, Nafty i Gazu AGH, Kraków
** Praca zrealizowana w ramach badań własnych WWNiG AGH

Niech zmienna objaśniana X jest określona za pomocą populacji n niezależnych wyników pomiaru

$$x_i,\ \text{dla}\ i = 1, 2, ...n \tag{1}$$

Niech czynnik klasyfikujący A posiada (pozwala na wydzielenie) k poziomów

$$a_j,\ \text{dla}\ j = 1, 2, ...k,\ \text{gdzie}\ k \geq 2\ \text{i}\ k < n \tag{2}$$

Wówczas czynnik klasyfikujący A dzieli populację zmiennej objaśnianej X na k populacji odpowiadających konkretnym wartościom czynnika A o liczebnościach n_j. Obliczając średnią $\bar{x}$ dla całej populacji zmiennej X oraz średnie $\bar{x}_j$ dla populacji odpowiadających poszczególnym wartościom zmiennej klasyfikującej, można wyrazić zmienność wartości x_i w następujący sposób

$$\sum_{j=1}^{k}\sum_{i=1}^{n_j}(x_{ji}-\bar{x})^2 = \sum_{j=1}^{k}\sum_{i=1}^{n_j}(x_{ji}-\bar{x}_j)^2 + \sum_{j=1}^{k}\sum_{i=1}^{n_j}(\bar{x}_j-\bar{x})^2 \tag{3}$$

Wyrażenie po lewej stronie równania (3) reprezentuje całkowitą sumę kwadratów (SS_{total}). Pierwsze wyrażenie po prawej stronie stanowi wewnątrzgrupową sumę kwadratów i jest miarą zmienności losowej, czyli błędu przypadkowego (SS_{error}). Drugie wyrażenie prawej strony równania oznacza międzygrupową sumę kwadratów i jest miarą systematycznych różnic między grupami (SS_{effect}).

Każda z wyróżnionych sum cechuje się określoną liczbą stopni swobody df (*Degrees of Freedom*). Z definicji df jest równa liczbie wszystkich pomiarów pomniejszonej o liczbę wszystkich powiązań. W tym przypadku wartości te wynoszą odpowiednio:

$$\begin{aligned} &df_{total} = n-1, \quad df_{error} = n-k, \quad df_{effect} = k-1 \\ &df_{total} = df_{error} + df_{effect} \end{aligned} \tag{4}$$

Mając sumy (3) i stopnie swobody (4), można określić średnie sumy kwadratów (MS)

$$MS_{error} = \frac{SS_{error}}{df_{error}} = \frac{SS_{error}}{n-k}, \quad MS_{effect} = \frac{SS_{effect}}{df_{effect}} = \frac{SS_{effect}}{k-1} \tag{5}$$

Wartości średnich sum kwadratów są wykorzystywane w metodzie ANOVA do weryfikacji hipotezy zerowej, która brzmi

$$H_0: \bar{x}_1 = \bar{x}_2 = ... = \bar{x}_k \tag{6}$$

co oznacza brak istotnych różnic wartości średnich dla populacji wyników zmiennej objaśnianej, wydzielonych przez wartości zmiennej klasyfikującej.
Hipoteza alternatywna ma postać

H_1: co najmniej dwie średnie populacyjne różnią się między sobą (7)

Weryfikacja hipotezy zerowej opiera się na badaniu ilorazu średnich sum kwadratów (MS_{effect}/MS_{error}), który w przypadku jej prawdziwości powinien być bliski jedności.

W przeciwnym wypadku, gdy iloraz ten będzie się odchylał wyraźnie w górę, prawdziwa okaże się hipoteza alternatywna. Do wnioskowania statystycznego wykorzystywany jest test F w postaci

$$F_{(k-1,n-k)} = \frac{MS_{effect}}{MS_{error}} \tag{8}$$

Statystyka ta przy założeniu hipotezy zerowej ma rozkład F (Fishera–Snedecora) o stopniach swobody: $k - 1$ i $n - k$.

Możliwość prowadzenia analizy wariancji opiera się na wielu założeniach. Poniżej podano treść podstawowych założeń.

1) Analizowana zmienna objaśniana jest mierzalna.
2) Pojedyncza próba jest pobierana z populacji w sposób losowy (I zasada randomizacji).
3) Próby zostały podzielone na k porównywanych grup w sposób losowy (II zasada randomizacji).
4) Wszystkie rozważane wyniki pomiarów są niezależne.
5) Wyniki pomiarów zmiennej objaśnianej mają rozkład normalny w obrębie wszystkich k porównywanych grup.
6) Rozkłady te mają jednakową wariancję (założenie jednorodności wariancji).

Cztery pierwsze założenia to podstawowe warunki prowadzenia analizy statystycznej; z oczywistych względów są one w rozważanym przypadku spełnione. Założenie piąte nakazuje, by wyniki pomiaru zmiennej objaśnianej podlegały rozkładowi normalnemu w obrębie każdej grupy. W literaturze zaznacza się, że niewielkie odchylenie rozkładów od normalności nie ma wpływu na dostarczanie dobrych wyników [2, 4]. Dotyczy to zwłaszcza sytuacji, gdy losowane próby są równoliczne.

W statystyce opracowano wiele testów normalności rozkładu. Zazwyczaj badają one wartość współczynnika skośności. Wszystkie z nich stawiają hipotezę zerową

H_0: rozkład wartości zmiennej w populacji jest zgodny z rozkładem normalnym (9)

Następnie określana jest wartość statystyki kontrolnej testu. Jeżeli jest ona istotna, wówczas prawdziwa staje się hipoteza alternatywna

H_1: rozkład wartości zmiennej w populacji nie jest zgodny z rozkładem normalnym (10)

Najbardziej popularną metodą badania zgodności z rozkładem normalnym, uznawaną do dzisiaj za najlepszą, jest test Shapiro–Wilka [3, 4]. W razie stwierdzenia wyraźnego odchylenia rozkładu danych od normalności, objawiającej się zwykle lewostronną lub prawostronną asymetrią, można przekształcić wartości zmiennej objaśnianej.

Do badania jednorodności wariancji opracowano także wiele testów sprawdzających. Wszystkie one służą do weryfikacji hipotezy zerowej o równości wariancji

$$H_0:\ \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_k^2 \tag{11}$$

wobec hipotezy alternatywnej

$$H_1:\ \sigma_i^2 \neq \sigma_j^2 \text{ dla co najmniej jednej pary } (i, j)\text{, gdzie } i \neq j \tag{12}$$

Dla przypadku jednowymiarowego (jednoczynnikowa ANOVA) zazwyczaj wykorzystuje się test Hartleya, test C Cochrana i test Bartletta. W przypadku gdy dochodzi do jawnego naruszenia zasady normalności rozkładu porównywanych populacji, dobre wyniki daje test Levene'a [2, 4].

Test Hartleya [3, 4], bardzo prosty w swej naturze, porównuje wariancję największą z najmniejszą. Wymaga on w przybliżeniu jednakowej liczby obserwacji w każdej grupie wyznaczonej wartościami zmiennej klasyfikującej. Jest bardzo wrażliwy na odchylenie rozkładów populacji od normalności.

Test C Cochrana [3, 4], równie prosty, porównuje wariancję maksymalną z ich sumą. Wymaga on również w przybliżeniu jednakowej liczby obserwacji w każdej grupie wyznaczonej wartościami zmiennej klasyfikującej. Jest także bardzo wrażliwy na odchylenie rozkładów populacji od normalności.

Test Bartletta [3, 4] o bardzo skomplikowanej postaci porównuje ważoną średnią arytmetyczną z wariancji z ważoną geometryczną średnią z wariancji. Jego zaletą jest to, że porównywalne grupy nie muszą być jednakowe pod względem liczebności, nadal jednak jest wrażliwy na odchylenie rozkładów populacji od normalności.

Test Levene'a [3, 4] bada bezwzględne odchylenia wyników pomiarów od średnich z populacji. Zastosowana metoda przeprowadzania tego testu sprawia, że jest on w znacznym stopniu odporny na odchylenia rozkładu od normalności i jednorodności wariancji. Gdy rozbieżności te są jednak bardzo duże, należy posłużyć się innymi metodami, np. testem Kruskala–Wallisa.

3. OBLICZENIA

Celem badań jest analiza wpływu różnych czynników na tempo migracji pionowej węglowodorów w fizycznym modelu gruntu. Wykorzystano dane eksperymentalne uzyskane na specjalnie przygotowanym laboratoryjnym stanowisku pomiarowym do badania migracji produktów naftowych w ośrodku gruntowo-wodnym [1]. Rolę zmiennej zależnej (objaśnianej) odgrywa zmierzona prędkość filtracji pionowej substancji ropopochodnej w gruncie.

Uwzględniono następujące czynniki klasyfikujące:

- Model gruntu – przygotowano trzy modele gruntu, wyraźnie różniące się między innymi właściwościami filtracyjnymi. Poziomy wartości czynnika to [1]:
  1) FM1 – fizyczny model gruntu piaszczystego,
  2) FM2 – fizyczny model gruntu piaszczysto-pylastego,
  3) FM3 – fizyczny model gruntu piaszczysto-pylastego.
- Substancja zanieczyszczająca – badania wykonano z użyciem trzech substancji ropopochodnych o odmiennych właściwościach, między innymi gęstości właściwej i lepkości dynamicznej. Poziomy wartości czynnika [1]:
  1) ET – etylina bezołowiowa,
  2) ON – olej napędowy zimowy,
  3) RG – lekka ropa kopalniana z kopalni Grobla k. Bochni.

– Strefa głębokości – przygotowane modele gruntu cechują się zmieniającym się z głębokością profilem zawilgocenia; zawartość wody w przestrzeni porowej zmienia się od kilku do ponad trzydziestu procent [1]. Intensywność tych zmian jest różna. Początkowo zawilgocenie wzrasta powoli, natomiast na większej głębokości wzrost jest gwałtowniejszy. Poziomy wartości czynnika:
  1) S1 – strefa modelu gruntu o zawilgoceniu poniżej 9% (powolny wzrost),
  2) S2 – strefa modelu gruntu o zawilgoceniu od 9% do 20% (pośredni wzrost),
  3) S3 – strefa modelu gruntu o zawilgoceniu powyżej 20% (intensywny wzrost).

Z uwagi na znacznie mniejsze różnice właściwości filtracyjnych między modelami gruntu FM2 i FM3 w porównaniu z modelem FM1 oraz mniejsze różnice wartości współczynnika lepkości dynamicznej między ON i RG w porównaniu z ET, w pierwszym etapie badań zdecydowano się uwzględnić tylko wyniki pomiarów laboratoryjnych prędkości filtracji pionowej ET i ON w FM1 i FM2. Przy tak dużym zróżnicowaniu właściwości gruntów i substancji zanieczyszczających porównywane populacje powinny wyraźniej ulegać wpływowi czynników klasyfikujących. Z drugiej strony łatwiej będzie spełnić założenia podstawowe metody ANOVA dla bardziej jednorodnych danych. Liczebność analizowanej populacji wartości zmiennej objaśnianej w tym przypadku jest równa 28.

W dalszym etapie badań rozważania zostaną rozszerzone do całej populacji danych o liczebności 61 i każdy czynnik klasyfikujący będzie posiadał trzy poziomy grupowania.

Wszystkie obliczenia statystyczne wykonano używając pakietu STATISTICA 7.1 PL. Przyjęto standardowy poziom istotności rozważań równy 0,05.

3.1. Analiza normalności rozkładu zmiennej objaśnianej

Badanie normalności rozkładu przeprowadzono na podstawie testu Shapiro–Wilka, analizując istotność wartości statystyki W [3, 4]. Dla całej populacji rozważanych w pierwszym etapie badań wartości eksperymentalnych zmiennej objaśnianej uzyskano następujące rezultaty

$$W = 0{,}77114, \text{ co odpowiada poziomowi istotności } p = 0{,}00003 \tag{13}$$

Wynik ten ($p < 0{,}05$) oznacza, że hipotezę o normalności rozkładu prędkości filtracji należy odrzucić. Z uwagi na asymetrię rozkładu zmiennej objaśnianej zdecydowano się dokonać przekształcenia jej wartości z użyciem funkcji logarytmicznej. Operacja ta wyraźnie poprawiła zgodność rozkładu z rozkładem normalnym. Ostateczne wyniki obliczeń zamieszczono w tabeli 1.

Logarytmowanie wyników pomiaru zmiennej objaśnianej pozwoliło podnieść krytyczną wartość testu o ponad dwa rzędy wielkości. Co prawda nadal jest ona dużo niższa od założonego poziomu istotności prowadzonej analizy i należy odrzucić hipotezę o normalności rozkładu całej populacji rozważanych danych. Nie ma to jednak w tym przypadku znaczenia, ponieważ założenie podstawowe prowadzenia analizy metodą ANOVA wymaga normalności rozkładów populacji danych w grupach wydzielonych przez czynnik klasyfikujący.

Tabela 1

Wyniki badania normalności rozkładu za pomocą testu Shapiro–Wilka

Populacja	Wartość czynnika klasyfikującego	Liczność populacji	Statystyka *W* Shapiro–Wilka	Obliczony poziom istotności *p*
Całość danych	–	28	0,88154	0,00434
Model gruntu 1	Model = FM1	14	0,88634	0,07160
Model gruntu 2	Model = FM2	14	0,86723	0,03837
Etylina	Substancja = ET	14	0,86487	0,03559
Olej napędowy	Substancja = ON	14	0,88644	0,07185
Wilgotność gruntu poniżej 9%	Strefa = S1	8	0,84774	0,09036
Wilgotność gruntu od 9% do 20%	Strefa = S2	10	0,84672	0,05309
Wilgotność gruntu powyżej 20%	Strefa = S3	10	0,88487	0,14833

W tym przypadku wyróżniamy trzy rodzaje czynika klasyfikującego:

1) Czynnik klasyfikujący – fizyczny model gruntu. Wydzielone zostały dwie równoliczne populacje wartości zmiennej objaśnianej. Pierwsza z nich, dotycząca FM1, posiada nieistotną wartość statystyki W (p = 0,072) i nie ma podstaw do odrzucenia hipotezy zerowej o normalności rozkładu. Druga populacja, dotycząca FM2, cechuje się istotną wartością statystyki W na poziomie istotności 0,05, dlatego hipotezę zerową o zgodności z rozkładem normalnym należy odrzucić. Odchylenie od rozkładu normalnego jest jednak niewielkie, na co wskazuje wyliczony poziom istotności statystyki W (p = 0,038), a ponieważ rozważane populacje są równoliczne, z powodzeniem można zastosować analizę ANOVA [2, 4].
2) Czynnik klasyfikujący – substancja zanieczyszczająca. Sytuacja jest bardzo podobna do opisanej w punkcie 1). Wydzielone zostały dwie równoliczne populacje wartości zmiennej objaśnianej. Druga z nich, dotycząca oleju napędowego, posiada nieistotną wartość statystyki W (p = 0,072) i nie ma podstaw do odrzucenia hipotezy zerowej o normalności rozkładu. Pierwsza populacja dla etyliny cechuje się istotną wartością statystyki W na poziomie istotności 0,05 prowadzonej analizy, dlatego hipotezę zerową o zgodności z rozkładem normalnym należy odrzucić. Odchylenie od rozkładu normalnego jest jednak niewielkie, co pokazuje wyliczony poziom istotności statystyki W (p = 0,036), a rozważane populacje są równoliczne, więc z powodzeniem można zastosować analizę ANOVA [2, 4].
3) Czynnik klasyfikujący – strefa głębokości (zawilgocenie i jego tempo wzrostu). Każda z trzech wydzielonych populacji danych charakteryzuje się nieistotną wartością statystyki W, przy założonym poziomie istotności prowadzonej analizy, dlatego nie ma podstaw do odrzucenia hipotezy zerowej o normalności rozkładu wartości.

3.2. Analiza równości wariancji rozkładów zmiennej objaśnianej

Badanie wariancji rozkładów populacji zmiennej objaśnianej zostało przeprowadzone z wykorzystaniem testu Hartleya, testu C Cochrana i testu Bartletta. Z uwagi na nieznaczne naruszenie w dwóch przypadkach zasady zgodności z rozkładem normalnym zastosowano także test Levene'a. Wyniki obliczeń zamieszczono w tabeli 2. Dla trzech pierwszych testów (Hartleya, C Cochrana i Bartletta) w tabeli podano tylko najbardziej niekorzystną wartość poziomu istotności otrzymanych statystyk z punktu widzenia słuszności hipotezy zerowej.

Tabela 2

Wyniki badania równości wariancji rozkładów

Czynnik klasyfikujący	Parametr	Test			
		Hartleya	C Cochrana	Bartletta	Levene'a
		F max	*C*	*Chi kw.*	*F*
Model	statystyka	2,4924	0,7137	2,5243	3,4066
	poziom istotności	0,1121			0,0764
Substancja	statystyka	2,7983	0,7367	3,1771	3,7713
	poziom istotności	0,07468			0,0630
Strefa	statystyka	5,2637	0,5962	5,4062	3,2429
	poziom istotności	0,0670			0,0559

Wyniki obliczeń zamieszczone w tabeli 2 pokazują, że dla żadnego z trzech rozważanych czynników klasyfikujących nie dochodzi do naruszenia założenia o równości wariancji rozkładów zmiennej objaśnianej. W związku z tym nie ma przeszkód do przeprowadzenia jednoczynnikowej analizy wariancji w odniesieniu do uzyskanych eksperymentalnie wartości prędkości filtracji pionowej substancji ropopochodnej w gruncie.

3.3. Jednoczynnikowa analiza wariancji

Wyniki jednoczynnikowej analizy wariancji zamieszczono w tabelach 3–5. W obliczeniach wykorzystano logarytmowane wartości eksperymentalne prędkości filtracji, zgodnie z poprzednimi sugestiami.

Zmierzona wartość zmiennej objaśnianej odpowiadająca *i-tej* obserwacji dla *j-tego* poziomu czynnika klasyfikującego może być przedstawiona za pomocą równania

$$x_{ji} = \mu + \alpha_j + \varepsilon_{ji} \tag{14}$$

gdzie:

μ – pewna nieznana wartość stała wspólna dla wszystkich populacji (w tab. 3–5 wyraz wolny);

α_j – składnik charakterystyczny dla j-tej populacji, wyrażający wpływ j-tego poziomu czynnika klasyfikującego;

ε_{ji} – błąd losowy związany z j-tym poziomem czynnika klasyfikującego i i-tą obserwacją.

Tabela 3

Jednoczynnikowa ANOVA prędkości filtracji w zależności od modelu gruntu

Efekt – model gruntu	Całkowita suma kwadratów	Liczba stopni swobody	Średnia suma kwadratów	Wartość statystyki	Poziom istotności statystyki
	SS	*df*	*MS*	*F*	*p*
Wyraz wolny	5,9123	1	5,9123	36,7301	0,000002
Effect	1,0930	1	1,0930	6,7903	0,0150
Error	4,1851	26	0,1610	–	–

Tabela 4

Jednoczynnikowa ANOVA prędkości filtracji w zależności od substancji ropopochodnej

Efekt – substancja ropopochodna	Całkowita suma kwadratów	Liczba stopni swobody	Średnia suma kwadratów	Wartość statystyki	Poziom istotności statystyki
	SS	*df*	*MS*	*F*	*p*
Wyraz wolny	5,9123	1	5,9123	42,8303	0,000001
Effect	1,6891	1	1,6891	12,2361	0,0017
Error	3,5891	26	0,1380	–	–

Tabela 5

Jednoczynnikowa ANOVA prędkości filtracji w zależności od poziomu zawilgocenia gruntu

Efekt – poziom zawilgocenia gruntu	Całkowita suma kwadratów	Liczba stopni swobody	Średnia suma kwadratów	Wartość statystyki	Poziom istotności statystyki
	SS	*df*	*MS*	*F*	*p*
Wyraz wolny	6,5001	1	6,5001	54,8078	0,000000
Effect	2,3132	2	1,1566	9,7523	0,00074
Error	2,9649	25	0,1186	–	–

W każdej rozpatrywanej sytuacji wyraz wolny cechuje się istotną wartością statystyki – poziom istotności znajduje się znacznie poniżej założonego poziomu istotności prowadzonej analizy.

Równie istotny okazał się wpływ każdego z rozważanych czynników klasyfikacyjnych rozpatrywanych z osobna. Najbardziej eksperymentalne wartości prędkości filtracji pionowej substancji ropopochodnej w gruncie zależą od zawilgocenia przestrzeni porowej, ponieważ w tym przypadku uzyskano najniższą wartość poziomu istotności obliczonej statystyki ($p = 0{,}00074$). Nieco słabiej na zmienną objaśnianą wpływa rodzaj użytej substancji ropopochodnej – różnice w gęstości właściwej i współczynniku lepkości dynamicznej ($p = 0{,}0017$). W najmniejszym stopniu na zmienną zależną wpływa fizyczny model gruntu – różne właściwości filtracyjne ($p = 0{,}0150$).

W drugim etapie badań powtórzono obliczenia dla całej zebranej populacji danych eksperymentalnych prędkości filtracji pionowej substancji ropopochodnych w fizycznych modelach gruntu [1]. Włączono dane dotyczące FM3 i RG. Z uwagi na ograniczone ramy niniejszej pracy zrezygnowano z zamieszczania wyników badania założeń podstawowych ANOVA. Rezultaty obliczeń przedstawiono w tabelach 6–8.

Tabela 6

Jednoczynnikowa ANOVA prędkości filtracji w zależności od modelu gruntu

Efekt – model gruntu	Całkowita suma kwadratów	Liczba stopni swobody	Średnia suma kwadratów	Wartość statystyki	Poziom istotności statystyki
	SS	*df*	*MS*	*F*	*p*
Wyraz wolny	5,3573	1	5,3573	46,4471	0,000000
Effect	1,4628	2	0,7314	6,3411	0,0032
Error	6,6899	58	0,1153	–	–

Tabela 7

Jednoczynnikowa ANOVA prędkości filtracji w zależności od substancji ropopochodnej

Efekt – substancja ropopochodna	Całkowita suma kwadratów	Liczba stopni swobody	Średnia suma kwadratów	Wartość statystyki	Poziom istotności statystyki
	SS	*df*	*MS*	*F*	*p*
Wyraz wolny	5,2946	1	5,2946	50,8139	0,000000
Effect	2,1093	2	1,0546	10,1217	0,00017
Error	6,0434	58	0,1042	–	–

Tabela 8

Jednoczynnikowa ANOVA prędkości filtracji w zależności od poziomu zawilgocenia gruntu

Efekt – poziom zawilgocenia gruntu	Całkowita suma kwadratów	Liczba stopni swobody	Średnia suma kwadratów	Wartość statystyki	Poziom istotności statystyki
	SS	*df*	*MS*	*F*	*p*
Wyraz wolny	6,4443	1	6,4443	75,1618	0,000000
Effect	3,1798	2	1,5899	18,5434	0,000001
Error	4,9729	58	0,0857	–	–

Analiza wyników obliczeń zamieszczonych w tabelach 6–8 wskazuje na wyraźne pogłębienie się wpływu czynników klasyfikujących na wartości zmiennej objaśnianej poziomy istotności wyliczonych statystyk są znacznie niższe od tych zamieszczonych w tabelach 3–5. Nadal największy wpływ na zróżnicowanie wartości prędkości filtracji pionowej substancji ropopochodnych ma zawilgocenie przestrzeni porowej gruntu ($p = 0,000001$). Nieco mniejszy udział ma tutaj rodzaj użytej substancji zanieczyszczającej ($p = 0,00017$), natomiast najsłabiej zaznacza się zależność od modelu gruntu ($p = 0,0032$).

4. PODSUMOWANIE

Zrealizowane metodą jednoczynnikową ANOVA badania wpływu wybranych czynników na eksperymentalne wartości prędkości filtracji pionowej substancji ropopochodnych w fizycznych modelach gruntu pokazały, że:

1) Wartości zmiennej objaśnianej istotnie zależą od wszystkich badanych czynników klasyfikujących: modelu gruntu, rodzaju substancji ropopochodnej i zawilgocenia przestrzeni porowej.
2) Najsilniejszy wpływ wywiera poziom zawilgocenia przestrzeni porowej i tempo jego wzrostu, nieco słabszy – rodzaj mieszaniny węglowodorów, a najsłabszy – rodzaj modelu gruntu.
3) Włączenie w drugim etapie do rozważań wszystkich zebranych danych (dotyczących modeli gruntu i substancji ropopochodnych różniących się w mniejszym stopniu niż te uwzględnione w pierwszym etapie) dało jeszcze istotniejsze wartości statystyk testu. Powodem jest ponaddwukrotny wzrost liczebności badanej populacji.

LITERATURA

[1] Fąfara Z.: *Badanie procesu migracji substancji ropopochodnych w ośrodku gruntowo-wodnym*. Uczelniane Wydawnictwa Naukowo-Dydaktyczne AGH, Rozprawy Monografie, nr 168, Kraków 2007

[2] Luszniewicz A., Słaby T.: *Statystyka z pakietem komputerowym Statistica PL. Teoria i zastosowania*. Wydawnictwo C.H. Beck, Gdynia 2003

[3] Stanisz A.: *Przystępny kurs statystyki z zastosowaniem Statistica PL na przykładach z medycyny. Tom 1. Statystyki podstawowe*. Wydawnictwo StatSoft Sp. z o.o., Kraków 2006

[4] Stanisz A.: *Przystępny kurs statystyki z zastosowaniem STATISTICA PL na przykładach z medycyny. Tom 2. Modele liniowe i nieliniowe*. Wydawnictwo StatSoft Sp. z o.o., Kraków 2007

[5] Twardowski K. et al.: *Badanie wpływu miejsca pobrania i uziarnienia prób węgla na oceny jego metanonośności*. Technicke Univerzity Ostrava, Monografie 15, LI, 2005

[6] Twardowski K. et al.: *Wykorzystanie dwuczynnikowej analizy wariancji do badania wpływu miejsca pobrania i sortymentu prób węgla na oceny jego metanonośności*. Zeszyty Naukowe Politechniki Śląskiej, seria Górnictwo, zeszyt 267, Gliwice 2005