

Wykorzystanie wielowymiarowych metod statystycznych w geologicznej interpretacji profilowań geofizyki otworowej w miocenijskich utworach z rejonu złoża gazu Dzików w SE Polsce

Monika Kaźmierczuk¹



Application of statistical methods to geological interpretation of well logs of the Miocene formation in the area of Dzików gas field in SE Poland. *Prz. Geol.*, 57: 1004–1014.

A b s t r a c t. The Miocene formation in the area of Dzików gas field is built of very thin sandy-shaly layers and thick sandstone layers so-called Dzików sandstones. Log data from wells: Dzików-12, Dzików-13 and Dzików-15 were processed with statistical methods. Principal Component Analysis was used for grouping and specifying well log data. It contributes to reducing the number of variables from a dozen of logs to four variables. Clustering, discrimination and classification are statistical tools facilitating data arrangement and preliminary data grouping according to natural petrophysical features of analysed rocks. Applied statistical methods are helpful tool assisting interpretation of log data.

Keywords: Carpathian Foredeep, gas deposits, statistical methods, Dzików sandstones

Metody statystyczne są pomocnym narzędziem w geologicznej interpretacji profilowań geofizyki otworowej. U podstaw stosowania tych metod leży zróżnicowanie ośrodka skalnego pod względem litologii i nasycenia przestrzeni porowej oraz związanych z tym zróżnicowanych wyników profilowań geofizyki otworowej zarejestrowanych w odmiennych warunkach. Proponowane metody zostały wykorzystane do porządkowania profilowań (analiza głównych składowych), grupowania danych w skupienia różniące się między sobą litologią i nasyceniem (analiza skupień) oraz do określania przynależności do poszczególnych grup wzdłuż profilu otworu wiertniczego.

Profilowania geofizyki otworowej dostarczają wielowymiarowego obrazu ośrodka skalnego. Informacja geologiczna, jaką niosą różne profilowania, może się w nich powtarzać. Dlatego, aby uporządkować tę informację i zredukować wymiar zagadnienia, stosuje się analizę głównych składowych (Moss & Seheult, 1987; Kaźmierczuk & Jarzyna, 2006; Kaźmierczuk, 2008). Zmniejszenie wymiaru polega na zastąpieniu kilkunastu profilowań geofizyki otworowej kilkoma składowymi głównymi (*Principal Components* — PC), które stanowią kombinacje liniowe wejściowych profilowań.

Celem analizy skupień jest podział zbioru danych na grupy różniące się litologią i nasyceniem. Podział danych odbywa się tak, aby obserwacje w obrębie grupy były do siebie jak najbardziej podobne i jednocześnie aby jak najbardziej różniły się od pozostałych grup pod względem wykształcenia litologicznego i właściwości zbiornikowych. Analizę skupień przeprowadzono na zredukowanym zbiorze profilowań geofizyki otworowej dla reprezentatywnego zbioru losowo wybranych głębokościowych punktów pomiarowych.

Analiza dyskryminacyjna została wykorzystana do predykcji wzdłuż całego profilu otworu wiertniczego grup o zróżnicowanej litologii i nasyceniu. Ze względu na szybkość uzyskiwania wyniku, obrazującego rozmieszczenie wzdłuż profilu otworu horyzontów o korzystnych właściwościach zbiornikowych, wielowymiarowe analizy

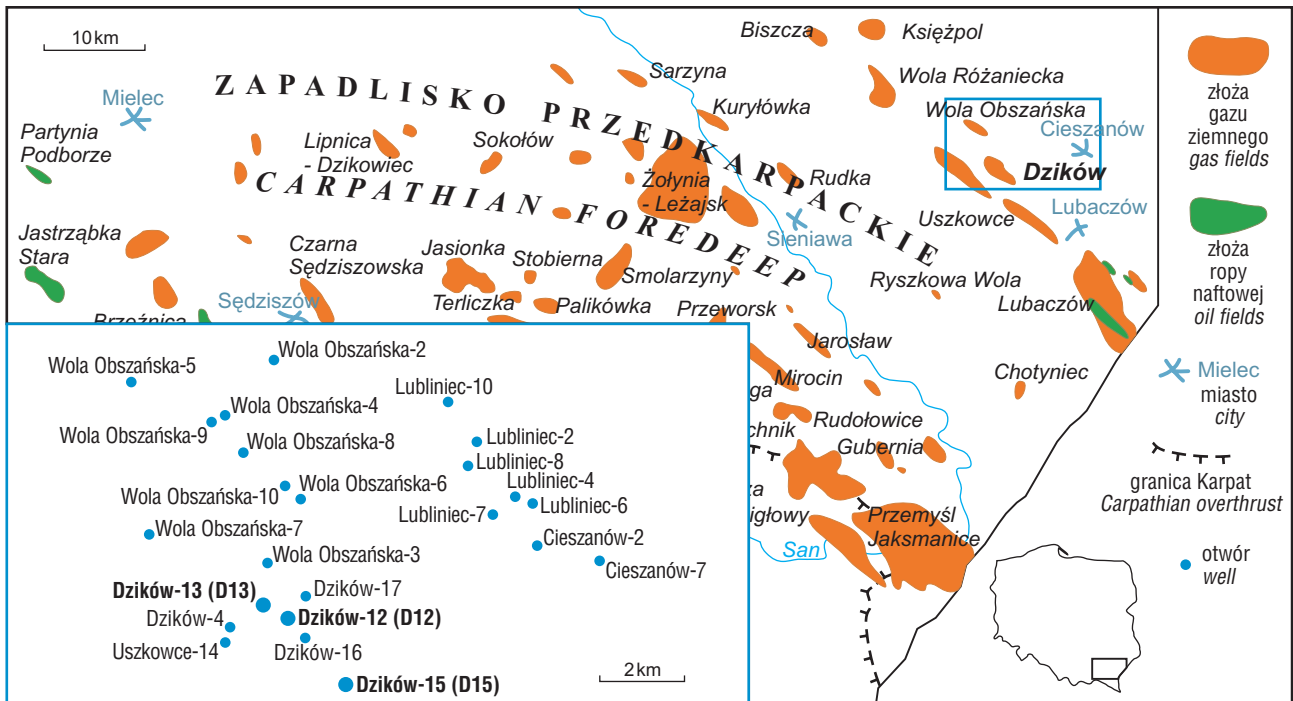
statystyczne mogą być wykorzystane do typowania poziomów do opróbowania w otworze wiertniczym.

Przedmiotem badań były osady sarmatu zapadliska przedkarpackiego, a w szczególności utwory cienko warstwowe oraz mięjsze warstwy piaskowców dzikowskich. Analizy statystyczne przeprowadzono dla profilowań geofizyki otworowej wykonanych w trzech otworach odwierconych w rejonie złoża gazu ziemnego Dzików (ryc. 1). Skałami zbiornikowymi są tutaj utwory piaskowcowe i mułowcowe, a rolę warstw uszczelniających pełnią ropy, łupki ilaste i osady ilasto-mułowcowe (Myśliwiec, 2004; Myśliwiec i in., 2004). W niektórych horyzontach akumulacje gazu są ograniczone poprzez zmiany litologiczne. Najlepszymi skałami zbiornikowymi dla gazu ziemnego w rejonie badań są tzw. piaskowce dzikowskie, wykształcone w postaci grubych pakietów piaskowcowych przeławionych mułowcami. Nasycenie gazem jest związane nie tylko z samymi piaskowcami dzikowskimi, ale także z cienkowarstwowymi utworami sarmatu. Zastosowane metody statystyczne wyraźnie pokazały strefy występowania gazu w profilu cienkowarstwowym utworów sarmatu.

Omówienie danych

Zbiór danych obejmował profilowania geofizyki otworowej wykonane w trzech otworach (Dzików-12 — D12, Dzików-13 — D13 i Dzików-15 — D15) odwierconych w złożu gazu ziemnego Dzików. Do analiz statystycznych włączono: pomiar naturalnej promieniotwórczości — GR (API), profilowanie porowatości neutronowej — NPHI (%), pomiar gęstości objętościowej — RHOB (g/cm^3), profilowanie akustyczne czasu interwałowego — DT ($\mu\text{s/m}$), profilowanie potencjałów polaryzacji naturalnej — SP (mV), elektryczne profilowania oporności pozornej — LL3 (Ωm), EL28 (Ωm), EL14 (Ωm), E64N (Ωm) i E16N (Ωm). Wykorzystano pomiary oporności pozornej wykonane różnymi typami sond o zróżnicowanej długości i radialnym zasięgu badań. Profilowania zostały wykonane aparaturą CAG-1 (*Dokumentacja otworów D12, D13, D15*). Do analiz statystycznych włączono również krzywą DCAL, która stanowi różnicę pomiędzy wynikami profilowania średnicy otworu — CALI (mm) a wartością średnicy

¹Wydział Geologii, Geofizyki i Ochrony Środowiska, Akademia Górniczo-Hutnicza, al. Mickiewicza 30, 30-059 Kraków; kazmierczuk@geol.agh.edu.pl



Ryc. 1. Rozmieszczenie złóż ropy i gazu ziemnego oraz wybranych otworów wiertniczych w zapadlisku przedkarpackim (Karnkowski, 1993, zmodyfikowana)

Fig. 1. Distribution of oil and gas fields, and wells in the Carpathian Foredeep (Karnkowski, 1993, modified)

nominalnej — BSM (mm). Wyniki analiz statystycznych profilowań porównano z wynikami kompleksowej interpretacji profilowań geofizyki otworowej, czyli ze współczynnikiem porowatości ogólnej — PHI, zawartością substancji piaszczystej — VSA i ilastej — VSH, współczynnikiem nasycenia wodą — SW oraz współczynnikiem przepuszczalności — K (mD). Ze względu na duży przedział zmienności parametru K oraz ze względu na liniową zależność między $\log K$ a innymi wielkościami występującymi w analizach posłużono się dziesiętnym logarytmem współczynnika przepuszczalności — $\log K$. W tabeli 1 pokazano, jakie profilowania oraz wyniki interpretacji były dostępne dla trzech analizowanych otworów. Na profilowaniach geofizyki otworowej zaznaczają się mięszsze pakiety piaskowców dzikowskich. Piaskowce te można łatwo wyróżnić na profilowaniach geofizycznych dzięki blokowemu kształtowi krzywych GR, SP, RHOB i NPHI (ryc. 2). Na rycinie 2 przedstawiono wyniki profilowań geofizyki otworowej w wybranym interwale głębokościowym w otworze D13, w którym pojawiają się trzy mięszsze warstwy piaskowców. W otworze D15 nie obserwuje się piaskowców dzikowskich.

Metodyka badań

Analizy statystyczne wykonano oddzielnie dla każdego otworu. W otworach D12 i D13 analizy statystyczne przeprowadzono w interwałach głębokościowych obejmujących cienko warstwowane utwory sarmatu oraz mięszsze pakiety piaskowców dzikowskich. Natomiast

Tab. 1. Profilowania geofizyki otworowej i wyniki interpretacji wykorzystane do analiz statystycznych

Table 1. Well logs and results of interpretation used in statistical analyses

Otwór Well	D12	D13	D15
Analizowany interwał głębokościowy [m] Analyzed depth interval [m]	210,00–1012,25	208,25–1045,75	161,40–1041,00
Profilowania geofizyki otworowej Well logs	GR, NPHI, RHOB, DT, SP, CALI, LL3, EL14, EL28, E16N, E64N	GR, NPHI, RHOB, DT, SP, CALI, EL14, EL28, E16N, E64N	GR, NPHI, RHOB, DT, SP, CALI, LL3, EL14, EL28, E16N, E64N
Wyniki interpretacji Porosity-lithology solution	PHI, K, SW, VSA, VSH	PHI, K, SW, VSA, VSH	PHI, K, SW, VSA, VSH

GR — profilowanie gamma, *gamma ray log*

NPHI — profilowanie porowatości neutronowej, *neutron porosity log*

RHOB — profilowanie gęstości objętościowej, *bulk density*

DT — profilowanie czasu interwałowego, *transit time interval*

SP — profilowanie potencjałów polaryzacji naturalnej, *spontaneous potential*

CALI — profilowanie średnicy otworu, *caliper log*

LL3, EL14, EL28, E16N, E64N — profilowania oporności pozornej, *resistivity logs*

PHI — współczynnik porowatości ogólnej, *porosity index*

K — współczynnik przepuszczalności, *permeability index*

SW — współczynnik nasycenia wodą, *water saturation index*

VSA — objętościowa zawartość piaskowca, *sandstone volume*

VSH — objętościowa zawartość minerałów ilastych, *shale volume*

w otworze D15 analizy statystyczne zostały wykonane dla cienko warstwowych utworów sarmatu.

Analiza głównych składowych posłużyła do redukcji liczby zmiennych. W jej wyniku otrzymano nowe zmienne, nie tracąc przy tym istotnych informacji o ośrodku skalnym. Te nowe zmienne zastąpiły wejściowe profilowania i w kolejnych etapach badań (analizie skupień i analizie dyskryminacyjnej) obliczenia były przeprowadzone na głównych składowych. W następnym kroku, w wyniku losowania głębokościowych punktów pomiarowych, wyłoniono reprezentatywny zbiór danych, na którym wykonano analizę skupień i wyróżniono grupy charakteryzujące się zróżnicowaną litologią i nasyceniem przestrzeni porowej. Grupy te w analizie dyskryminacyjnej posłużyły do predykcji litofacji i rodzaju nasycenia skał wzdłuż całego profilu otworu wiertniczego. Obliczenia wykonano przy użyciu programu *Statistica* (*StatSoft, Inc.*, 2005).

Analiza głównych składowych

Profilowania geofizyki otworowej dostarczają wielowymiarowego obrazu ośrodka skalnego wzdłuż otworu wiertniczego. Interpretacja tego obrazu jest problemem skomplikowanym i czasochłonnym. Ponieważ część profilowań geofizycznych z natury swej niesie podobne informacje o ośrodku skalnym, można zmniejszyć wymiar przestrzeni danych poprzez zastąpienie wielu wyjściowych profilowań mniejszą ilością nowych zmiennych, tzw. składowych głównych (Moss & Seheult, 1987; Kaźmierczuk & Jarzyna, 2006; Kaźmierczuk, 2008). W metodzie tej wyniki profilowań geofizyki otworowej traktuje się jako zmienne losowe. Aby wyniki różnych profilowań mogły być ze sobą porównywane, przed przystąpieniem do analiz profilowania geofizyki otworowej zostały poddane standaryzacji, tak by ich wartości średnie były równe zeru, a wariancja równa jedności. Na podstawie standaryzowanych wartości profilowań wyznacza się nowe zmienne, tzw. główne składowe (PC), które stanowią kombinacje liniowe standaryzowanych profilowań geofizyki otworowej:

$$PC_j = a_{1j} \cdot X_{1st} + a_{2j} \cdot X_{2st} + \dots + a_{mj} \cdot X_{mst}$$

gdzie:

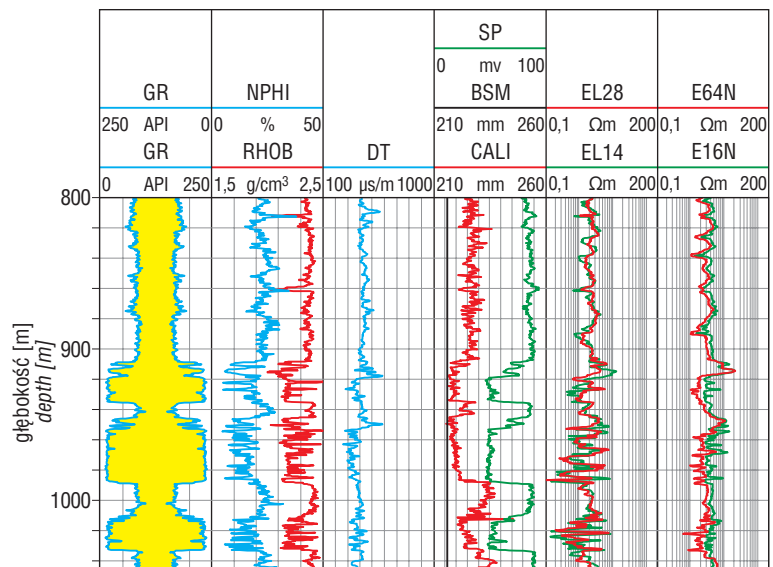
PC_j — poszczególne składowe główne, $j = 1, 2, \dots, m$,

a_{mj} — współczynniki określające kierunki osi nowego układu,

X_{mst} — standaryzowane profilowania geofizyki otworowej,

m — liczba profilowań wykorzystanych do obliczenia głównych składowych.

Miarą zmienności charakteryzującej ośrodek skalny, a wyjaśnianej przez poszczególne składowe, jest wariancja. Kierunki osi nowych zmiennych pokrywają się z kierunkami maksymalnej zmienności w zbiorze danych, a więc z kierunkami maksymalnej wariancji. Samo obliczenie głównych składowych nie powoduje zmniejszenia wymiaru zagadnienia. Można uzyskać tyle głównych składowych, ile profilowań zostało włączonych do analizy. Aby zmniejszyć liczbę zmiennych z kilkunastu wejściowych



Ryc. 2. Profilowania geofizyki otworowej w otworze D13
Fig. 2. Well logs in D13 well

profilowań do kilku składowych głównych, należy zrezygnować z głównych składowych charakteryzujących się najmniejszą wariancją. Redukcja wymiaru odbywa się w sposób kontrolowany przez zastosowanie różnych kryteriów, np. kryterium Kaisera (Dobosz, 2004), zgodnie z którym do dalszej analizy pozostawia się te składowe, których wartości własne są większe od jedności. O budowie nowych zmiennych — składowych głównych — świadczą ładunki czynnikowe, czyli współczynniki korelacji pomiędzy poszczególnymi składowymi a profilowaniami.

Analizę głównych składowych w otworze D12 przeprowadzono dla jedenastu profilowań geofizyki otworowej. W tabeli 2 przedstawiono ładunki czynnikowe wraz z procentem wariancji wyjaśnianej przez poszczególne składowe, a także skumulowany procent wariancji dla trzech otworów. Pierwsza składowa główna w otworze D12 wyjaśnia największą część zmienności w zbiorze danych (ok. 38%). Druga składowa opisuje ok. 24% zmienności, trzecia zaś ok. 11%. Kolejne składowe główne wyjaśniają coraz mniejszą część zmienności. Do dalszych analiz statystycznych pozostawiono cztery składowe, które w sumie wyjaśniają ok. 82% wariancji w zbiorze danych (tab. 2). Pierwsza składowa główna (PC1) w otworze D12 jest zbudowana przede wszystkim z potencjałowych profilowań oporności oraz profilowania LL3 (tab. 2). Za istotne uważa się ładunki, których moduł jest większy bądź równy 0,7. Ładunki czynnikowe dla tych profilowań w pierwszej składowej przyjmują dodatnie wartości: 0,95 dla E16N; 0,92 dla LL3 i 0,86 dla E64N. Duży wkład w PC1 ma również profilowanie NPHI, przy czym ładunek czynnikowy przyjmuje wartość ujemną (−0,59). Druga główna składowa (PC2) jest zbudowana przede wszystkim z profilowań: SP, GR, DCAL. Duży udział w tej składowej ma również profilowanie RHOB i NPHI. Ładunki czynnikowe wymienionych profilowań w drugiej głównej składowej przyjmują dodatnie wartości (tab. 2). Trzecia składowa zawiera profilowanie akustyczne DT, dodatnio skorelowane z PC3. Czwarta składowa stanowi kombinację liniową gradientowych profilowań oporności: EL14 i EL28.

Największy udział profilowań oporności w PC1 w otworze D12 wskazuje, że zmiany litologiczne oraz występowanie gazu w przestrzeni porowej prowadzą do wyraźnej zmiany

oporności skał. Wpływ nasycenia jest podkreślony przez duże ładunki czynnikowe, wyznaczone na podstawie oporności mierzonej w strefie bliskiej (EN16) i dalekiej (EN64 oraz LL3). Zróznicowanie oporności w tych strefach wskazuje na migrację gazu w przestrzeni porowej na skutek wnikania filtratu płuczki do skały gazonośnej i zmianę oporności mieszaniny filtratu i wody złożowej w skałach wodonośnych. Zailenie skał w otworze D12 jest odzwierciedlone w budowie składowej PC2 przez udział GR i SP. Obecność DCAL w tej składowej z dużym współczynnikiem korelacji wskazuje, że zwiększenie średnicy, nawet niewielkie w stosunku do średnicy nominalnej, w utworach ilastych oraz obecność osadu ilastego w skałach porowatych i przepuszczalnych uzupełnia informację o litologii i nasyceniu. Wysoki współczynnik korelacji między PC3 i DT ukazuje rolę porowatości ogólnej w charakterystyce badanych skał. W ostatniej składowej, PC4, największą rolę odgrywa oporność, zarejestrowana sondami gradientowymi.

W otworze D13 redukcję wymiaru przeprowadzono dla dziesięciu profilowań geofizyki otworowej. Wymiar zredukowano do czterech głównych składowych, które wyjaśniają w sumie ok. 82% zmienności (tab. 2). Pierwsza główna składowa wyjaśnia ok. 35% zmienności i jest zbudowana głównie z profilowań: SP (ładunek czynnikowy wynosi 0,94), GR (0,89) i NPHI (0,78). Drugą główną składową, opisującą ok. 25% wariancji, tworzą potencjałowe profilowania oporności, dla których ładunki czynnikowe przyjmują wartość 0,90 dla E16N i 0,93 dla E64N. Trzecia główna składowa jest ujemnie skorelowana z profilowaniem RHOB (ładunek czynnikowy wynosi -0,79). Duży udział w tej składowej ma również profilowanie DT. PC3 wyjaśnia ok. 13% zmienności w zbiorze danych. Do czwartej głównej składowej, wyjaśniającej ok. 8% wariancji, weszły gradientowe profilowania oporności, EL14 i EL28.

Budowa składowej PC1 w otworze D13 wskazuje na dominację ilastych składników litologicznych. Obserwuje się znaczną różnicę w procencie zmienności wyjaśnianej

przez składową PC1 w otworach D12 i D13. Jednak wyniki podstawowych statystyk dotyczących zawartości piaskowca, VSA, zailenia, VSH, porowatości ogólnej, PHI, i współczynnika nasycenia wodą, SW, w badanych otworach nie są wyraźnie zróżnicowane (tab. 3). Udział profilowań oporności EN16 i EN64 w konstrukcji składowej PC2 jest wynikiem znaczącego wpływu litologii i nasycenia na wyniki tych pomiarów w skałach piaskowcowo-ilastych. Większy, niż w otworze D12, udział profilowania RHOB może świadczyć o większym wpływie gazu na zróżnicowanie gęstości objętościowej w warstwach nasyconych wodą i gazem w otworze D13.

Analizę głównych składowych w otworze D15 przeprowadzono dla jedenastu profilowań geofizyki otworowej. Wymiar przestrzeni danych zmniejszono do czterech głównych składowych, które razem wyjaśniają ok. 83% zmienności w zbiorze danych (tab. 2). Każda z czterech składowych wyjaśnia odpowiednio ok. 32%, 25%, 16% i 10% wariancji. Pierwsza składowa jest zbudowana głównie z potencjałowych profilowań oporności (E16N i E64N) oraz z profilowania LL3. Druga składowa jest związana z profilowaniami GR, NPHI i SP oraz DCAL. Kolejną zmienność wyjaśniają profilowania DT (ładunek czynnikowy wynosi 0,83) i RHOB (-0,89). Gradientowe profilowania oporności weszły w skład czwartej głównej składowej.

Podsumowując analizę składowych głównych w trzech badanych otworach, stwierdzono, że taka sama wartość średnia zailenia w otworach D12 i D15 odzwierciedliła się podobną budową składowych PC1, PC2 i PC4. Czas interwału wykazał wysoki współczynnik korelacji ze składową PC3 w otworach D12 i D15. W otworach tych obliczono mniejszą średnią porowatość w porównaniu do najwyższej wartości w otworze D13. W otworze D12, w którym obliczono najwyższe średnie nasycenie gazem (równie 0,18) przy najwyższej wartości wyznaczonego nasycenia gazem równej 0,86; gęstość objętościowa, RHOB, wykazała małą bezwzględną wartość współczynnika korelacji, podczas gdy w otworach D13 i D15, gdzie

Tab. 2. Ładunki czynnikowe dla głównych składowych w otworach z rejonu złoża Dzików

Table 2. Factor loadings for principal components for wells in Dzikow gas field

Profilowania geofizyki otworowej <i>Well logs</i>	D12				D13				D15			
	PC1	PC2	PC3	PC4	PC1	PC2	PC3	PC4	PC1	PC2	PC3	PC4
GR	-0,13	0,93	0,10	0,08	0,89	-0,08	-0,27	0,05	-0,11	0,81	-0,44	0,13
NPHI	-0,59	0,56	0,20	0,01	0,78	-0,38	-0,07	-0,09	-0,12	0,73	0,31	0,16
RHOB	-0,05	0,58	-0,54	-0,01	0,25	-0,06	-0,79	-0,01	0,01	0,26	-0,89	0,04
SP	-0,03	0,95	0,04	0,07	0,94	0,05	-0,09	0,07	-0,12	0,86	0,03	-0,03
DT	0,10	0,11	0,88	0,10	0,51	0,19	0,68	0,08	0,11	0,24	0,83	0,18
EL14	0,46	0,01	0,09	0,80	0,02	0,47	0,16	0,76	0,35	0,07	0,07	0,87
EL28	0,14	0,02	0,06	0,95	0,03	0,12	0,06	0,94	0,14	0,06	0,05	0,95
E16N	0,95	-0,16	0,04	0,21	-0,23	0,90	0,09	0,25	0,94	-0,16	0,02	0,18
E64N	0,86	0,01	0,19	0,22	0,00	0,93	0,12	0,17	0,94	0,02	0,04	0,22
LL3	0,92	-0,15	0,04	0,18	-	-	-	-	0,95	-0,09	0,05	0,12
DCAL	-0,13	0,76	-0,07	-0,10	0,38	-0,07	-0,66	-0,18	0,05	0,73	-0,06	0,00
% ogółu wariancji % of total variance	38,44	24,35	10,71	8,61	35,46	24,80	13,07	8,59	31,97	24,95	15,82	10,06
skumulowany % cumulative %	82,11				81,92				82,80			

oznaczone ładunki czynnikowe są >0,70 lub <-0,70
loadings marked by bold are >0,70 or <-0,70

Tab. 3. Podstawowe statystyki dla wyników interpretacji w otworach D12, D13 i D15

Table 3. Basic statistics for results of lithology-porosity interpretation for D12, D13 and D15 wells

Profilowania geofizyki otworowej <i>Well logs</i>	Liczba danych <i>Number of observations</i>	Interwał głębokościowy [m] <i>Depth interval [m]</i>	Średnia <i>Mean</i>	Minimum <i>Minimum</i>	Maksimum <i>Maximum</i>	Odchylenie standardowe <i>Standard deviation</i>
DZIKÓW-12						
VSA	3210	210,00–1012,25	0,37	0,00	0,99	0,12
VSH			0,52	0,01	1,00	0,15
PHI			0,10	0,00	0,38	0,07
SW			0,82	0,14	1,00	0,19
DZIKÓW-13						
VSA	3351	208,25–1045,75	0,40	0,00	0,97	0,11
VSH			0,49	0,03	1,00	0,14
PHI			0,11	0,00	0,34	0,07
SW			0,87	0,21	1,00	0,18
DZIKÓW-15						
VSA	8796	161,40–1041,00	0,41	0,00	0,65	0,06
VSH			0,52	0,18	1,00	0,09
PHI			0,07	0,00	0,29	0,06
SW			0,97	0,52	1,00	0,08

średnie nasycenie gazem jest niższe (odpowiednio 0,13 i 0,03), gęstość objętościowa weszła z dużym współczynnikiem korelacji do składowej PC3.

Analiza skupień

Celem przeprowadzonej analizy skupień jest podział zbioru danych na grupy różniące się litologią i nasyceniem. Podział danych odbywa się tak, aby obserwacje w obrębie grupy były do siebie jak najbardziej podobne i jednocześnie aby jak najbardziej różniły się od pozostałych grup pod względem przyjętych kryteriów. Spośród różnych metod grupowania wybrano metodę Warda, która wykorzystuje wariancję w zbiorze danych, tzn. dąży do minimalizacji sumy kwadratów odchylenia dowolnych skupień, które mogą zostać uformowane na każdym etapie grupowania (Ward, 1963). Punktem wyjścia jest zbiór n danych (n skupień), który w pierwszym kroku grupowania łączy się w $n - 1$ grup. Takiego łączenia dokonuje się poprzez obliczenie wewnątrzgrupowej sumy kwadratów dla każdej możliwej pary danych. Para danych, dla której wewnątrzgrupowa suma kwadratów przyjmuje najmniejszą wartość, tworzy nowe skupienie. Identyczne postępowanie powtarza się w kolejnym kroku grupowania i uzyskuje $n - 2$ grup. Procedura grupowania jest prowadzona do momentu, aż dane zostaną zgrupowane w jednym skupieniu.

Kolejne kroki analizy skupień można przedstawić na wykresie drzewkowym (dendrogramie) (ryc. 3). Na jednej osi są rozmieszczone elementy poddane aglomeracji, na drugiej zaś podaje się odległość aglomeracyjną, w tym przypadku odległość euklidesową. Na rycinie 3, obok wykresu drzewkowego, przedstawiono również wykres przebiegu aglomeracji, pomocny podczas ustalania optymalnego poziomu odcięcia wykresu drzewkowego. Oś pozioma opisuje kolejne etapy grupowania, a pionowa podaje odległość wiązania. W idealnej sytuacji optymalny poziom odcięcia wyznacza się w miejscu nagłego wzrostu odległości wiązania. Zmiennymi w analizie skupień mogą być zarówno profilowania geofizyki otworowej, jak i wyróżnione na ich podstawie główne składowe (Każmierczuk

& Jarzyna, 2006; Każmierczuk, 2008). W trzech analizowanych otworach jako zmienne wejściowe do analizy skupień wybrano główne składowe.

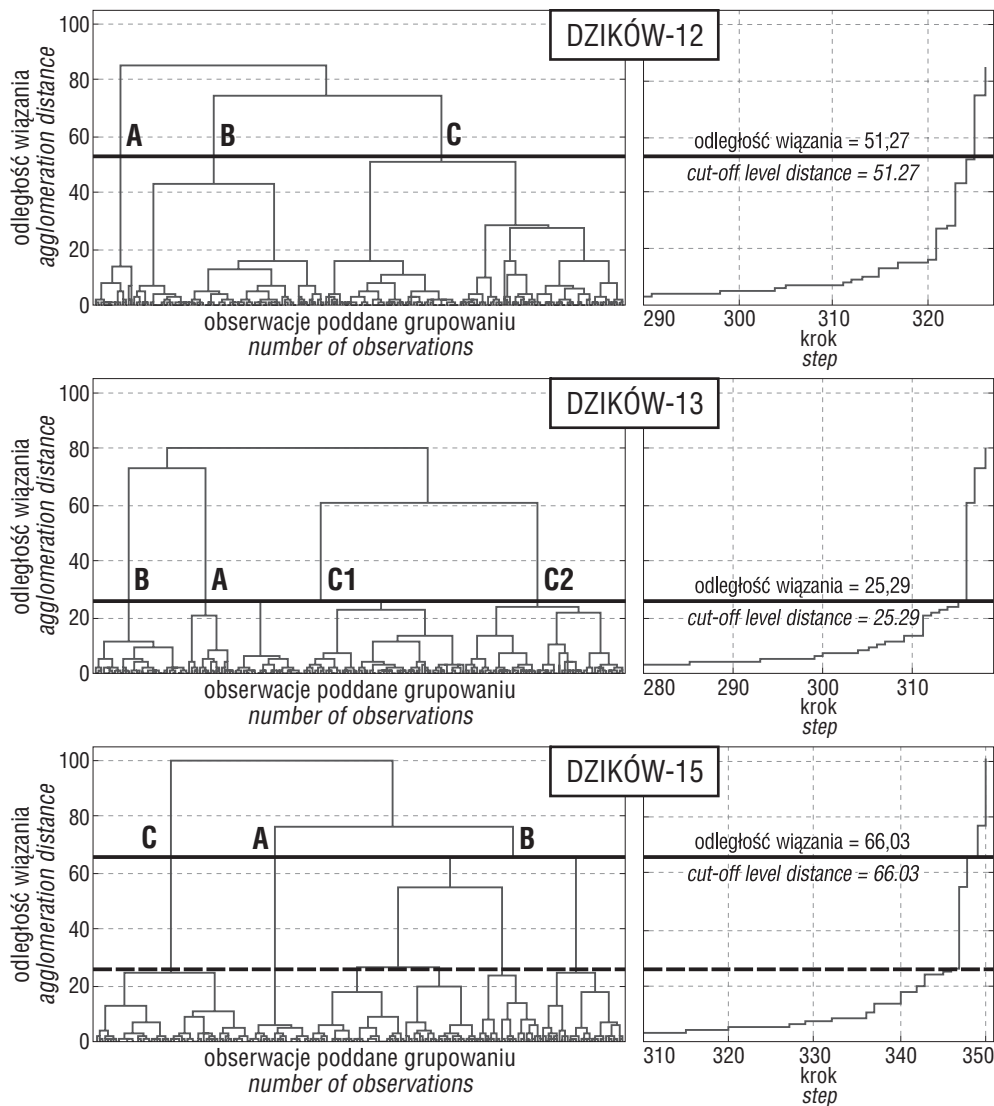
W otworze D12 analizę skupień przeprowadzono dla pełnego profilu utworów sarmatu. Ze względu na dużą liczbę (3210) głębokościowych punktów pomiarowych wylosowano 327 wyników pomiarów i dla nich przeprowadzono obliczenia. Efektem tej analizy jest podział zbioru danych na trzy grupy, które na rycinie 3 wstępnie oznaczono jako A, B i C. Te trzy grupy odzwierciedlają zróżnicowanie ośrodka skalnego pod względem litologii i nasycenia. Liczność danych w poszczególnych grupach przedstawiono w tabeli 4. Najliczniejsza okazała się grupa C, w której znalazło się 185 elementów. Następnie grupa B ze 118 wynikami i grupa A licząca 24 obserwacje. Na uwagę zasługuje fakt, że grupę A tworzą dane pochodzące z poziomu piaskowców dzikowskich. W grupach B i C procent danych pochodzących z piaskowców jest znikomy.

W celu sprawdzenia, czy wyłonione grupy różnią się pomiędzy sobą pod względem litologii i nasycenia, porównano je z wynikami kompleksowej interpretacji profilowań geofizyki otworowej. Dlatego też dla współczynnika porowatości ogólnej — PHI, logarytmu współczynnika przepuszczalności — $\log K$, współczynnika nasycenia wodą — SW oraz zawartości składnika piaszczystego — VSA przeprowadzono nieparametryczny test mediany. Test ten jest alternatywą dla jednoczynnikowej analizy wariancji w układzie międzygrupowym. Oceniana jest w nim hipoteza

Tab. 4. Liczba danych w grupach w otworze D12

Table 4. Number of data for clusters from D12 well

Grupa <i>Cluster</i>	Liczba danych <i>Number of data</i>	Liczba danych z poziomu piaskowców dzikowskich <i>Number of data from Dzików sandstones</i>
A	24	24 (100%)
B	118	3 (2,54%)
C	185	2 (1,08%)



Ryc. 3. Wykres drzewkowy wraz z wykresem przebiegu aglomeracji dla głównych składowych w trzech otworach
Fig. 3. Dendrogram and agglomeration distance plot for three wells for principal components

teza, że porównywane próby zostały pobrane z populacji o tym samym rozkładzie lub z rozkładów o tej samej medianie.

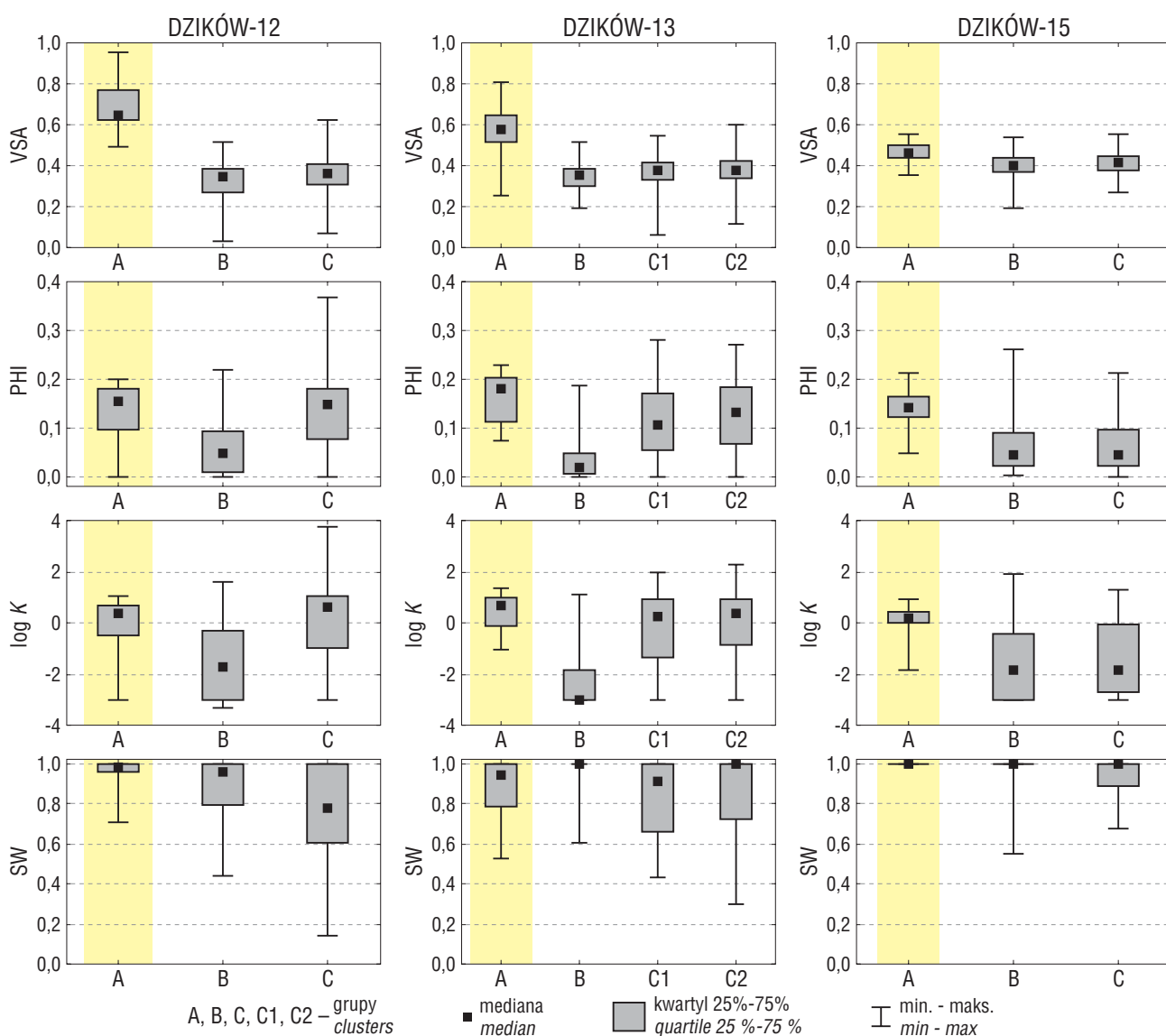
Na rycinie 4 przedstawiono wykresy ramka-wąsy dla VSA, PHI, SW i log K w poszczególnych grupach. Grupę A spośród pozostałych grup wyróżnia podwyższona zawartość składnika piaszczystego i największe nasycenie wodą. Porowatość i przepuszczalność są podobne jak w grupie C. Grupa B, pod względem litologicznym podobna do danych z grupy C (obniżona zawartość składnika piaszczystego), charakteryzuje się najmniejszą porowatością i przepuszczalnością. Najmniejsze nasycenie wodą występuje w grupie C, jednak przedział zmienności nasy-

cenia wodą dla danych w tej grupie jest większy niż w grupach A i B. Wyróżnione grupy można więc nazwać: A — piaskowce dzikowskie, B — utwory iłowcowo-piaskowcowe o obniżonej porowatości i przepuszczalności, nasycone wodą i C — utwory iłowcowo-piaskowcowe o podwyższonej porowatości i przepuszczalności, nasycone wodą i gazem.

W otworze D13 analizę skupień przeprowadzono dla składowych głównych wyróżnionych w pełnym profilu utworów sarmatu. Ze względu na dużą liczbę (3351) głębokościowych punktów pomiarowych, podobnie jak w otworze D12, wylosowano 319 elementów i dla nich przeprowadzono analizę. Efektem tej analizy jest wykres drzewkowy przedstawiony na rycinie 3. Wyboru miejsca przecięcia dendrogramu dokonano na podstawie analizy wykresu przebiegu aglomeracji (ryc. 3). Wybór odległości wiązania, równej 25,29, jest w tej sytuacji najlepszy. Począwszy od tej odległości następuje największy przyrost wartości odległości pomiędzy tworzonymi grupami. W rezultacie otrzymano podział zbioru danych na cztery grupy, które wstępnie oznaczono jako A, B, C1 i C2. Postawiono pytanie: czy te cztery grupy odzwierciedlają zróżnicowanie ośrodka skalnego pod względem litologii i nasycenia? Liczba obserwacji zakwalifikowanych do poszczególnych

Tab. 5. Liczba danych w grupach w otworze D13
 Table 5. Number of data for clusters from D13 well

Grupa Cluster	Liczba danych Number of data	Liczba danych z poziomu piaskowców dzikowskich Number of data from Dzików sandstones
A	27	25 (92,59%)
B	57	0 (0%)
C1	139	0 (0%)
C2	96	5 (5,21%)



Ryc. 4. Wykresy ramka-wąsy dla VSA, PHI, SW i log K w profilu sarmatu w otworach D12, D13 i D15
 Fig. 4. Box and whisker plots for VSA, PHI, SW, log K for the Sarmatian formation from D12, D13 and D15

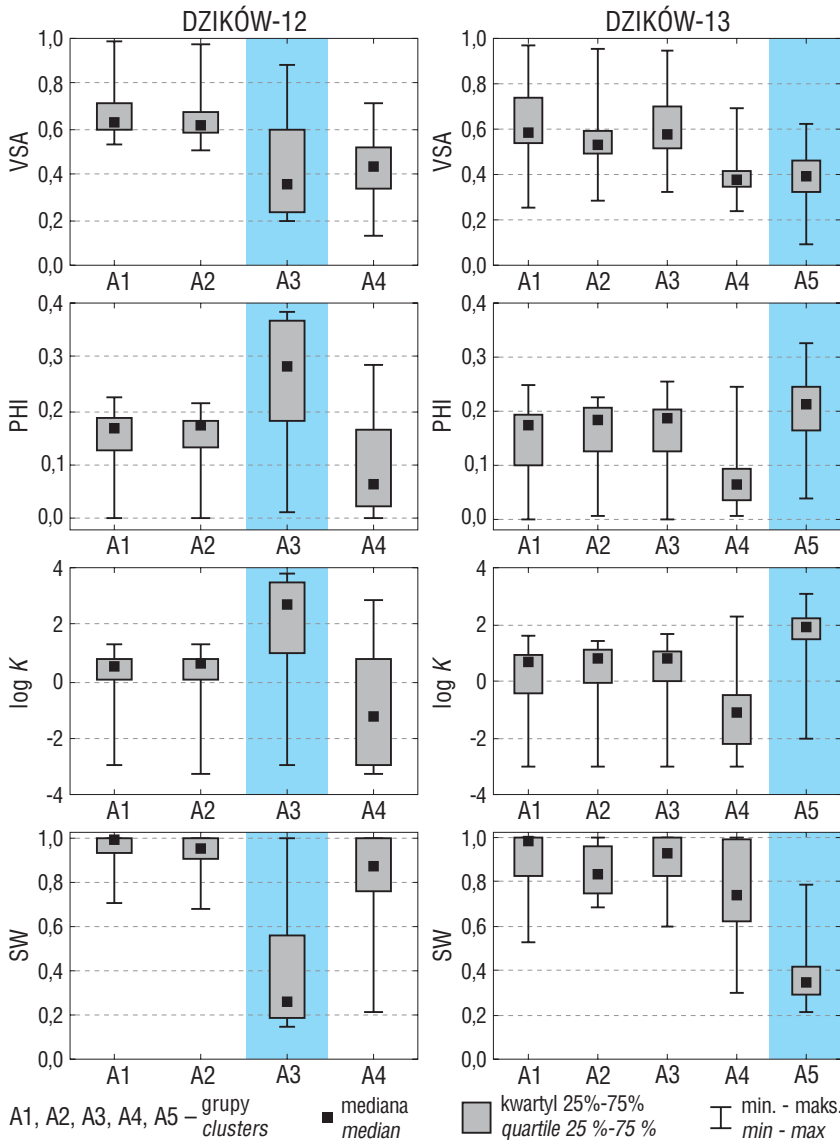
grup przedstawiono w tabeli 5. Na uwagę zasługuje fakt, że większość danych (ponad 92%) pochodząca z poziomu piaskowców dzikowskich utworzyła odrębną grupę, która na rycinach 3 i 4 została oznaczona jako grupa A. Tylko kilka obserwacji należących do piaskowców dzikowskich weszło do innej grupy (C2) (tab. 5).

Podobnie jak w przypadku danych w otworze D12 wydzielone grupy porównano z wynikami kompleksowej interpretacji profilowań geofizyki otworowej. Przeprowadzono nieparametryczny test mediany dla współczynnika porowatości ogólnej, logarytmu współczynnika przepuszczalności, współczynnika nasycenia wodą oraz zawartości składnika piaszczystego.

Na rycinie 4 przedstawiono wykresy ramka-wąsy dla VSA, PHI, SW i log K w poszczególnych grupach wyłonionych w analizie skupień dla utworów sarmatu w otworze D13. Grupa A (piaskowce dzikowskie) wyróżnia się podwyższoną zawartością składnika piaszczystego oraz współczynnikiem porowatości ogólnej. Grupa B charakteryzuje się najmniejszą porowatością i przepuszczalnością oraz najwyższym nasyceniem wodą. Pod względem litologicznym podobna jest do grup C1 i C2. Taki wynik

skłania, aby wyróżnione grupy nazwać: A — piaskowce dzikowskie, B — utwory iłowcowo-piaskowcowe o słabych właściwościach zbiornikowych, C1 — utwory iłowcowo-piaskowcowe o podwyższonej porowatości i przepuszczalności, nasycone wodą i gazem i C2 — utwory iłowcowo-piaskowcowe o podwyższonej porowatości i przepuszczalności, nasycone wodą. Grupy C1 i C2 są podobne do grupy C w otworze D12.

W otworze D15 metodę Warda zastosowano do składowych głównych wyróżnionych w pełnym profilu utworów sarmatu. Ze względu na dużą liczbę (8796) głębokościowych punktów pomiarowych wylosowano 351 obserwacji i dla nich przeprowadzono analizę. Efektem tej analizy jest wykres drzewkowy przedstawiony na rycinie 3. Wykres przebiegu aglomeracji sugeruje przecięcie dendrogramu w odległości równej 26,50 (wskazanej przez linię przerywaną) i podział zbioru danych na pięć grup. Jednak zdecydowano się przeciąć dendrogram w odległości równej 66,03, a tym samym podzielić zbiór na trzy grupy: A, B i C. Zbyt duża liczba grup, a w szczególności podział grupy B na trzy podgrupy, nie ujawnia żadnych



Ryc. 5. Wykresy ramka-wąsy dla VSA, PHI, SW, log K w piaskowcach dzikowskich w otworach D12 i D13

Fig. 5. Box and whisker plots for VSA, PHI, SW, log K for the Dzików sandstones from D12 and D13 wells

istotnych różnic pomiędzy tymi podgrupami w wynikach kompleksowej interpretacji geofizyki otworowej.

Spośród wyróżnionych w otworze D15 grup skupienie A charakteryzuje się podwyższoną zawartością składnika piaszczystego oraz większą porowatością i przepuszczalnością, w porównaniu z pozostałymi grupami (ryc. 3). Wszystkie obserwacje, które weszły do tej grupy, rejestrują nasycenie wodą równe 100%. Grupy C i B nie różnią się istotnie pod względem litologicznym, porowatości i przepuszczalności. Różnica pomiędzy nimi tkwi w nasyceniu wodą. Rozstęp kwartyli w grupie C dla nasycenia wodą jest większy niż w grupie B. Zatem wyróżnione grupy nazwano: A — piaskowcowo-iłowcowa o podwyższonej porowatości i przepuszczalności, nasycona wodą, B — iłowcowo-piaskowcowa o obniżonej porowatości i przepuszczalności, nasycona wodą i C — iłowcowo-piaskowcowa o obniżonej porowatości i przepuszczalności, nasycona wodą i gazem.

W analizie skupień przeprowadzonej dla danych z dwóch otworów (D12 i D13) wyłoniła się grupa piaskowców dzikowskich. Postawiono pytanie: czy piaskowce dzikow-

skie są jednorodne pod względem wykształcenia litologicznego i nasycenia? Żeby sprawdzić tę tezę, wykonano analizę skupień dla danych pochodzących z poziomów piaskowców dzikowskich w otworach D12 i D13. W rezultacie otrzymano podział piaskowców na cztery grupy w otworze D12 i pięć grup w otworze D13 (ryc. 5). Na szczególną uwagę w otworze D12 zasługuje grupa, która na rycinie 5 została oznaczona jako A3. Grupa ta obejmuje najmniejszą liczbę obserwacji (18 elementów). Pod względem litologicznym jest to grupa o obniżonej zawartości składnika piaszczystego, najwyższej porowatości i przepuszczalności oraz najniższym nasyceniu wodą (ryc. 5). Oznacza to, że na 268 wyników uzyskanych z piaskowców dzikowskich tylko 18 pochodzi z poziomów nasyconych gazem.

Sprawdzając „jednorodność” piaskowców dzikowskich pod względem wykształcenia litologicznego i nasycenia w otworze D13, przeprowadzono również analizę skupień tylko w obrębie piaskowców dzikowskich (370 obserwacji). Zbiór danych podzielono na pięć grup. Tak duża liczba wydzielonych skupień była konieczna do zbadania wewnętrznej struktury piaskowców dzikowskich. Dla wyróżnionych grup wykonano nieparametryczny test mediany, aby zbadać istotność zróżnicowania mediany w grupach. Spośród wyłonionych grup na uwagę zasługuje grupa A5 (ryc. 5), która liczy 44 obserwacje. Charakteryzuje się (podobnie jak grupa A4) obniżoną zawartością składnika piaszczystego. Oprócz tego dane w tej grupie rejestrują większe wartości przepuszczalności i najmniejsze wartości nasycenia wodą. Współczynnik porowatości jest również

podwyższony w stosunku do pozostałych grup. Piaskowce dzikowskie nie są więc utworami jednorodnymi pod względem litologii i nasycenia. Tylko 44 pomiary spośród 370 pochodzą z poziomów nasyconych gazem w piaskowcach dzikowskich w otworze D13. Zatem, biorąc pod uwagę wyniki analizy skupień w całych zbiorach danych w otworach D12, D13 i D15 oraz analizy tylko piaskowców dzikowskich, stwierdzono, że nasycenie gazem jest związane ze skałami piaskowcowo-mułowcowo-iłowcowymi. Część lamin piaskowcowych w obrębie bloków piaskowców dzikowskich jest nasycona gazem, ale istotne jest także wyłonienie warstw gazonośnych w cienkowarstwowych utworach powyżej bloków piaskowców dzikowskich.

Analiza dyskryminacyjna

Analiza dyskryminacyjna została wykorzystana do wyznaczenia funkcji dyskryminacyjnych. Ich konstrukcja polegała na znalezieniu takiej liniowej kombinacji głównych składowych, aby funkcje dyskryminacyjne najlepiej rozróżniały poszczególne typy litologiczne lub

Tab. 6. Macierz klasyfikacji dla danych testowych. Otwór D12

Table 6. Classification of testing data for D12 well

Grupa Cluster	Poprawne (%) Correct (%)	A	B	C
A	85,71	6	1	0
B	100,00	0	27	0
C	94,83	2	1	55
Razem <i>All data</i>	95,65	8	29	55

Tab. 7. Macierz klasyfikacji dla danych testowych. Otwór D13

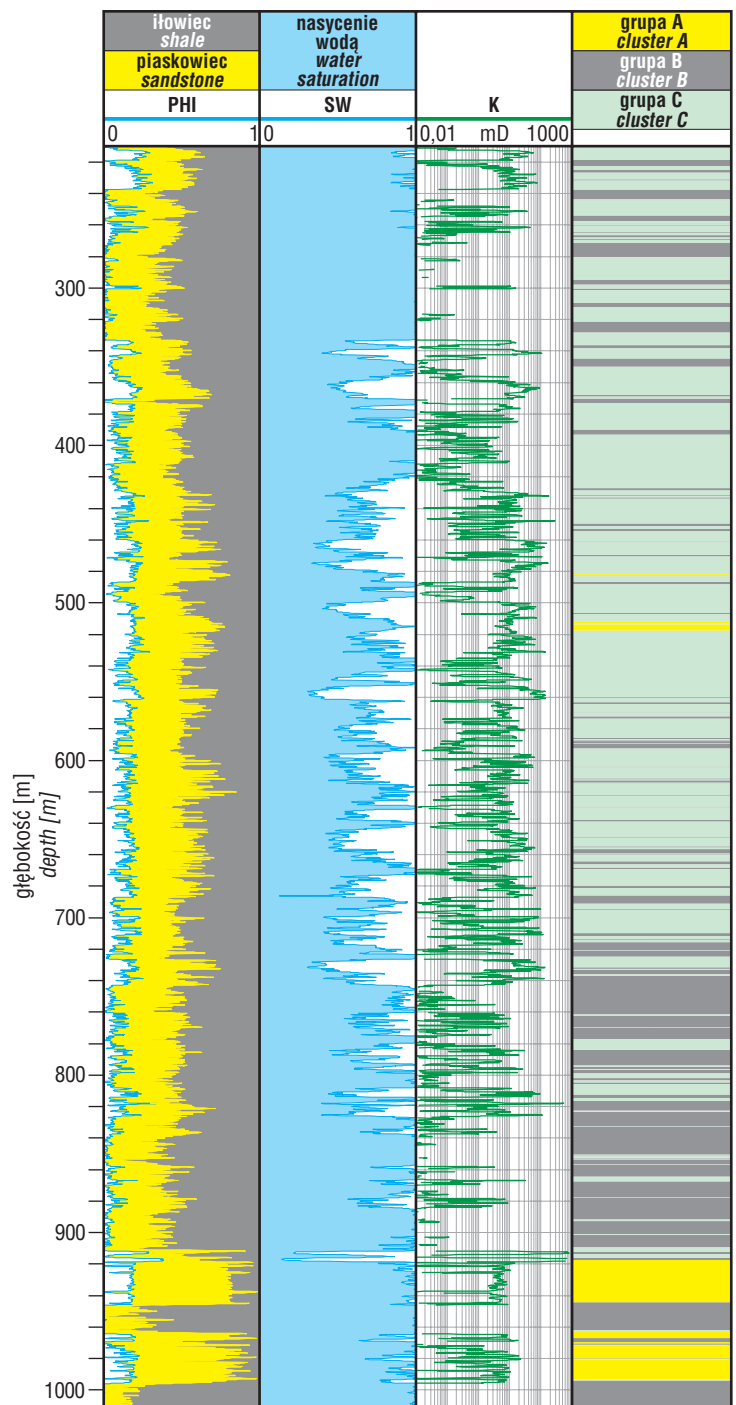
Table 7. Classification of testing data for D13 well

Grupa Cluster	Poprawne (%) Correct (%)	A	B	C1	C2
A	100,00	7	0	0	0
B	100,00	0	14	0	0
C1	96,67	0	0	29	1
C2	86,96	0	0	3	20
Razem <i>All data</i>	94,59	7	14	32	21

odmienne nasycenie w skałach zbiornikowych. Drugim zadaniem analizy dyskryminacyjnej była klasyfikacja obserwacji, czyli wyników pomiarów profilowań geofizyki otworowej, do grup określających poszczególne typy litologiczne i zróżnicowane nasycenie. Podstawę klasyfikacji stanowi zróżnicowanie w wartościach średnich pomierzonych parametrów petrofizycznych, a co z tym jest związane, również w wartościach średnich głównych składowych, wyznaczonych na podstawie profilowań geofizyki otworowej.

W otworze D12 wzorzec stanowiły grupy wyróżnione w analizie skupień. Dane te losowo podzielono na dwa zbiory: zbiór treningowy oraz zbiór testowy. Liczność zbiorów danych treningowych i testowych ustalono w stosunku 3 : 1. Na podstawie treningowego zbioru danych wyznaczono składową główną najlepiej różnicującą ośrodek skalny. Największy wkład do ogólnej dyskryminacji w otworze D12 wniosła druga główna składowa, związana z profilowaniami GR, SP i DCAL. Poprawność klasyfikacji sprawdzono na testowym zbiorze danych (tab. 6). Taki sam sposób postępowania zastosowano do otworów D13 i D15.

W testowej grupie A w otworze D12 znalazło się 7 obserwacji, z czego poprawnie do tej grupy zakwalifikowano 6 wyników (tab. 6), co stanowi 85,71%. Poprawność klasyfikacji w grupie B wyniosła 100%, a w grupie C — 94,83%. Ogólnie w otworze D12 poprawnie zakwalifikowano 95,65% danych testowych. W otworze D13 największy wkład w dyskryminację miała pierwsza główna składowa, związana z profilowaniami GR, NPHI i SP. Poprawność klasyfikacji w testowym zbiorze danych była również wysoka i wyniosła 94,59% (tab. 7). W otworze D15



grupa A – piaskowce dzikowskie

grupa B – utwory ilowcowo-piaskowcowe o obniżonej porowatości i przepuszczalności, nasycone wodą

grupa C – utwory ilowcowo-piaskowcowe o podwyższonej porowatości i przepuszczalności, nasycone wodą i gazem

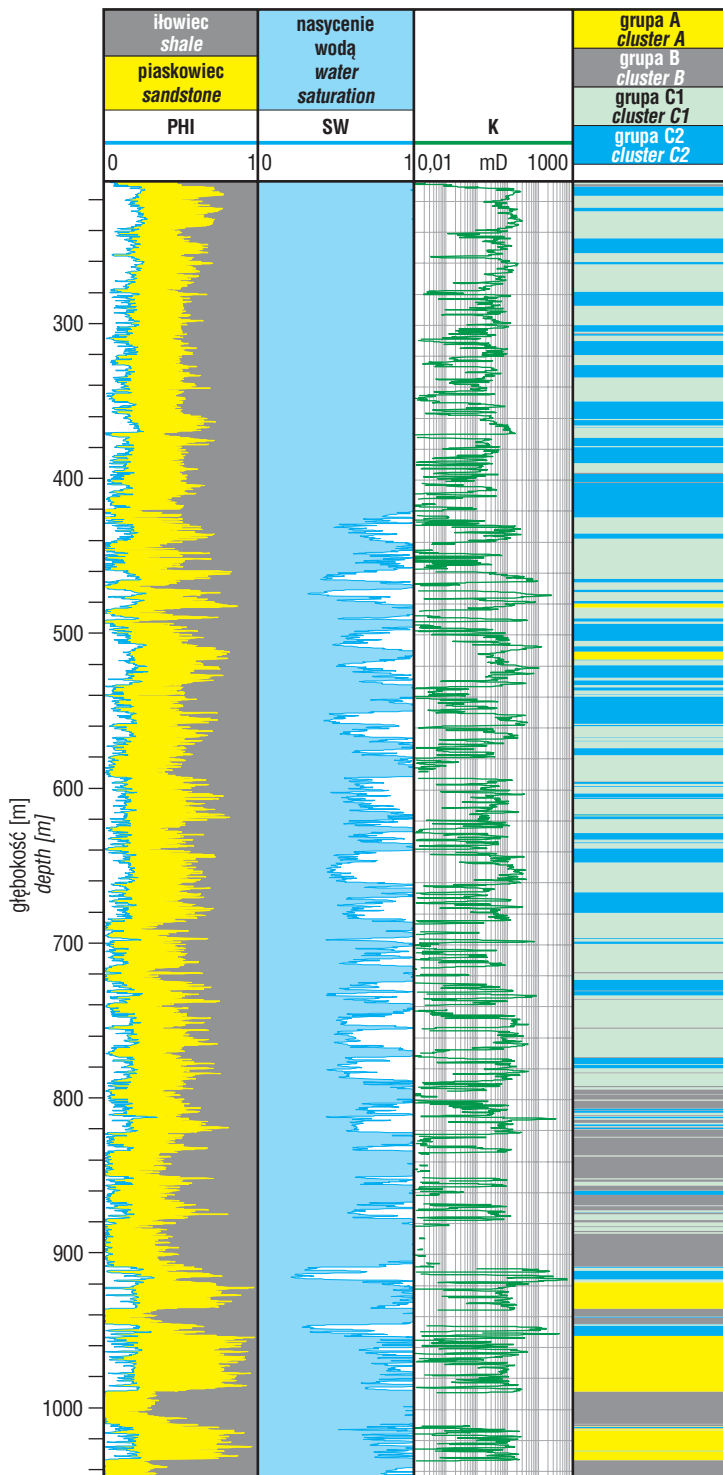
cluster A – Dzików sandstones

cluster B – sandy shale with lower porosity and permeability, water saturated

cluster C – sandy shale with higher porosity and permeability, water and gas saturated

Ryc. 6. Rozwiązanie porowatościowo-litologiczne wraz z wynikami analizy dyskryminacyjnej dla profilu sarmatu w otworze D12

Fig. 6. Porosity-lithology solution and results of discriminant analysis for Sarmatian formation for D12 well



- grupa A** – piaskowce dzikowskie
grupa B – utwory iłowcowo-piaskowcowe o słabych właściwościach zbiornikowych
grupa C1 – utwory iłowcowo-piaskowcowe o podwyższonej porowatości i przepuszczalności, nasycone wodą i gazem
grupa C2 – utwory iłowcowo-piaskowcowe o podwyższonej porowatości i przepuszczalności, nasycone wodą
cluster A – Dzików sandstones
cluster B – sandy shale with poor reservoir properties
cluster C1 – sandy shale with higher porosity and permeability, water and gas saturated
cluster C2 – sandy shale with higher porosity and permeability, water saturated

Tab. 8. Macierz klasyfikacji dla danych testowych. Otwór D15

Table 8. Classification of testing data for D15 well

Grupa Cluster	Poprawne (%) Correct (%)	A	B	C
A	100,00	9	0	0
B	97,87	1	46	0
C	81,48	0	5	22
Razem All data	92,77	10	51	22

największy wkład w budowę funkcji dyskryminacyjnych miały trzecia i druga główna składowa. Składowa PC3 była związana z profilowaniami RHOB i DT, zaś składowa PC2 — z GR, NPHI i SP. Procent poprawnie zakwalifikowanych danych w całym testowym zbiorze wyniósł 92,77% (tab. 8).

Wykonano również klasyfikację danych do poszczególnych grup litologicznych wzdłuż całego analizowanego profilu w otworach D12, D13 oraz D15. Wyniki tej analizy wraz z rozwiązaniem porowatościowo-litologiczno-nasyceniowym przedstawiono na rycinach 6–8. W otworach D12 i 13 w analizach statystycznych wyłoniła się grupa piaskowców dzikowskich. Można również zauważyć dwie cienkie warstwy piaskowcowe, na głębokości ok. 500 m, o rozważanych cechach statystycznych zbliżonych do piaskowców dzikowskich (ryc. 6 i 7). Obecność poziomów nasyconych wodą i gazem (otwory D12 i D13 — grupa C) jest związana z cienkowanymi utworami sarmatu zalegającymi powyżej piaskowców dzikowskich, co potwierdziła analiza dyskryminacyjna.

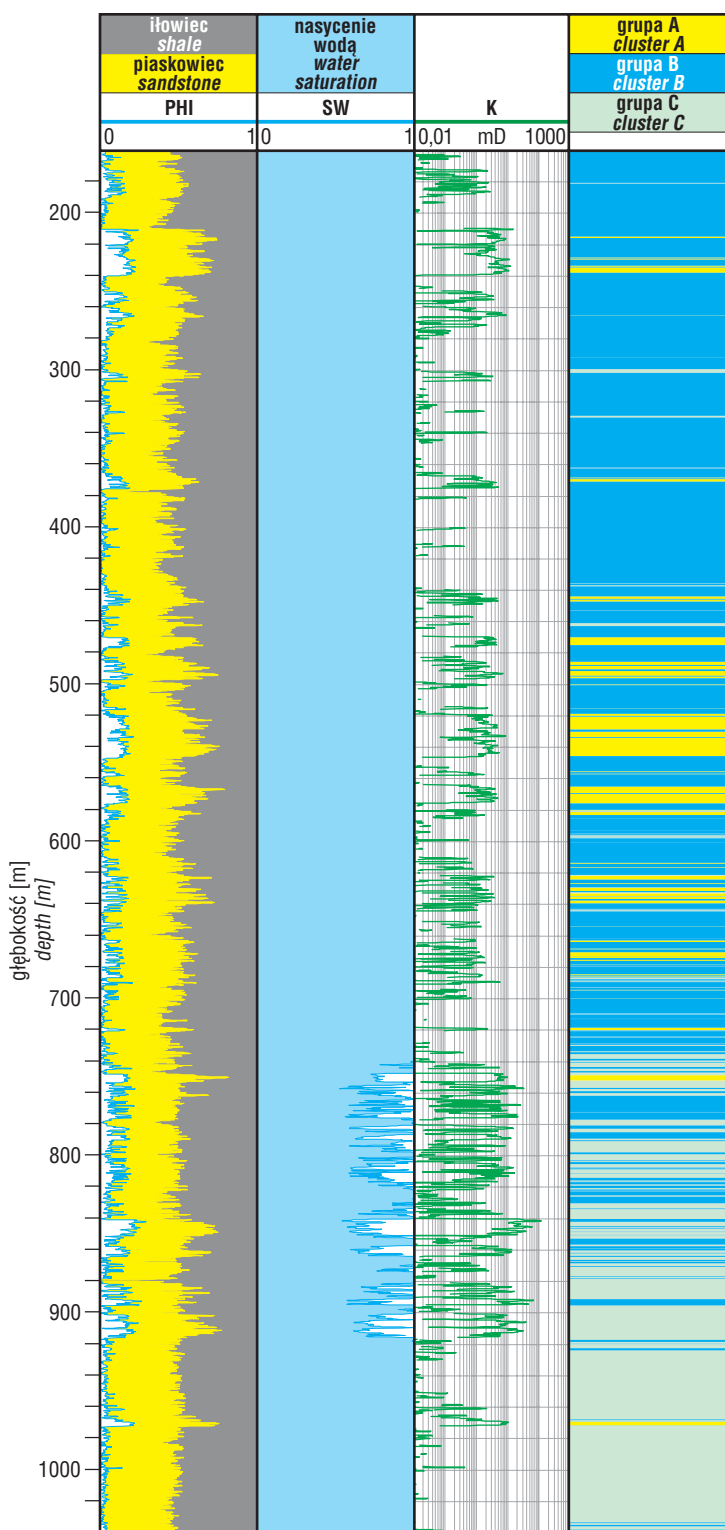
Wnioski

Proponowane metody statystyczne są pomocne podczas wstępnej geologicznej interpretacji profilowań geofizyki otworowej. Testowano je już wcześniej w cienkowanymi utworami mioceńskich złóż Ryszkowa Wola i Jasionka w zapadlisku przedkarpackim, a niniejsza praca stanowi rozszerzenie i uzupełnienie prowadzonych analiz statystycznych. Analiza głównych składowych jest skuteczną metodą porządkowania profilowań geofizyki otworowej pod względem istotności wnoszonej przez nie informacji i redukcji wymiaru zagadnienia. Niejednorodność piaskowców dzikowskich została potwierdzona dzięki analizie sku-

←

Ryc. 7. Rozwiązanie porowatościowo-litologiczne wraz z wynikami analizy dyskryminacyjnej dla profilu sarmatu w otworze D13

Fig. 7. Porosity-lithology solution and results of discriminant analysis for Sarmatian formation for D13 well



- grupa A** – utwory piaskowcowo-iłowcowe o podwyższonej porowatości i przepuszczalności, nasycone wodą
- grupa B** – utwory iłowcowo-piaskowcowe o obniżonej porowatości i przepuszczalności, nasycone wodą,
- grupa C** – utwory iłowcowo-piaskowcowe o obniżonej porowatości i przepuszczalności, nasycone wodą i gazem
- cluster A** – shaly sandstones with higher porosity and permeability, water saturated
- cluster B** – sandy shale with lower porosity and permeability, water saturated
- cluster C** – sandy shale with lower porosity and permeability, water and gas saturated



Ryc. 8. Rozwiązanie porowatościowo-litologiczne wraz z wynikami analizy dyskryminacyjnej dla profilu sarmatu w otworze D15

Fig. 8. Porosity-lithology solution and results of discriminant analysis for Sarmatian formation for D15 well

pień, którą wraz z analizą dyskryminacyjną warto wykorzystać do wstępnych badań litologii i nasycenia w profilu geologicznym otworu. Przeprowadzone testy potwierdziły, że nasycenia gazem na badanym obszarze nie należy wiązać jedynie z miąższymi pakietami piaskowcowymi (piaskowcami dzikowskimi), ale także z cienkowarstwowymi utworami sarmatu. Wyniki analiz statystycznych dobrze skorelowały się z wynikami kompleksowej interpretacji profilowań geofizyki otworowej. Metody statystyczne, ze względu na szybkość uzyskiwania poprawnego wyniku, mogą być wykorzystane do typowania poziomów nasyconych gazem do opróbowania.

Praca została wykonana w ramach projektu badawczego własnego KBN: nr N525 023 32/2308 pt. *Geologiczne rozpoznanie ośrodka skalnego na podstawie profilowań geofizyki otworowej przy wykorzystaniu metod statystycznych*, realizowanego w latach 2007–2009. Praca była prezentowana podczas I Polskiego Kongresu Geologicznego w Krakowie, 26–28.06.2008 r.

Literatura

- DOBOSZ M. 2004 — Wspomagana komputerowo statystyczna analiza wyników badań. Akademicka Oficyna Wydawnicza EXIT, Warszawa.
- Dokumentacja** otworów Dzików 12, Dzików 13, Dzików 15. Arch. Geofizyki Kraków.
- KARNKOWSKI P. 1993 — Złóża gazu ziemnego i ropy naftowej w Polsce, t.2, Karpaty i zapadlisko przedkarpackie. Towarzystwo Geosynoptyków „Geos” AGH, Kraków.
- KAŹMIERCZUK M. 2008 — Wykorzystanie metod statystycznych do identyfikacji piaskowców w miocenijskich utworach z rejonu złoża Dzików. [W:] Międzynarodowa Konf. Nauk.-Tech. GEOPETROL 2008 nt. „Nauka, technika i technologia w rozwoju poszukiwań i wydobywania węglowodorów w warunkach lądowych i morskich”. Pr. Inst. Nafty i Gazu, 150: 177–182.
- KAŹMIERCZUK M. & JARZYNA J. 2006 — Improvement of lithology and saturation determined from well logging using statistical methods. *Acta Geoph.*, 54, 4: 378–398.
- MOSS B. & SEHEULT A. 1987 — Does principal component analysis have a role in the interpretation of petrophysical data? [In:] 28th Annual Logging Symposium Transactions, Paper TT. Society Professional Well Log Analysts.
- MYŚLIWIEC M. 2004 — Miocenijskie skały zbiornikowe zapadliska przedkarpackiego. *Prz. Geol.*, 52: 581–592.
- MYŚLIWIEC M., PLEZIA B. & ŚWIĘTNICKA G. 2004 — Nowe odkrycia złóż gazu ziemnego w osadach miocenu północno-wschodniej części zapadliska przedkarpackiego na podstawie interpretacji bezpośredniego wpływu nasycenia węglowodorami na zapis sejsmiczny. *Prz. Geol.*, 52: 395–402.
- StatSoft**, Inc. 2005 — STATISTICA (data analysis software system), version 7.1. www.statsoft.com.
- WARD J.H. 1963 — Hierarchical grouping to optimize an objective function. *J. Amer. Stat. Assoc.*, 58: 236–244.

Praca wpłynęła do redakcji 9.06.2009 r.
Po recenzji akceptowano do druku 7.08.2009 r.