

## SEKWENCJE MODALNE W ANALIZIE PROFILI LITOLOGICZNYCH – SPOJRZENIE METODYCZNE

### Modal sequences in lithological profiles analysis – methodological approach

**Marek DOKTOR & Andrzej J. KRAWCZYK**

*Akademia Górniczo-Hutnicza, Wydział Geologii, Geofizyki i Ochrony Środowiska;  
al. Mickiewicza 30, 30-059 Kraków;  
e-mail: doktor@agh.edu.pl, akraw@geolog.geol.agh.edu.pl*

**Treść:** Powszechnie stosowane w sedymentologii metody teorii łańcuchów Markowa nie pozwalają na w pełni wiarygodne wnioskowanie o cykliczności procesu sedymentacji, której przejawem może być występowanie sekwencji modalnych, powtarzających się w profilu wielokrotnie i posiadających jednoznaczną interpretację genetyczną. W pracy przedstawiono propozycję statystycznego testu permutacyjnego, przeznaczonego do weryfikacji hipotezy o losowości występujących w profilu litologicznym sekwencji warstw. Posługując się symulacyjną techniką Monte Carlo, szacuje się prawdopodobieństwo, że w losowym „profilu” złożonym z tych samych warstw co profil badany znajdzie się dana liczba poszczególnych sekwencji. Pozwala to na odróżnienie sekwencji modalnych, wskazujących na istotne cechy procesu sedymentacji, od takich, których częste występowanie w profilu spowodowane jest wyłącznie dużą liczbą warstw danej odmiany litologicznej.

**Słowa kluczowe:** cykliczność sedymentacji, sekwencje modalne, test permutacyjny

**Abstract:** The Markov chains theory is a tool commonly applied to sedimentological studies. Unfortunately, this method does not provide fully credible conclusions concerning the cyclicity of sedimentation. Such cyclicity is documented, among others, by the appearance of modal sequences, i.e., sequences many times repeated in the lithological profile and having unequivocal genetic interpretation. The paper proposes the statistical permutation test, which can verify the hypothesis of the randomness of layers succession in a given lithological profile. Applying the Monte Carlo simulation method, the probability is estimated that in a random sequence composed of the same layers as the studied profile the given number of particular sequences of layers will occur. Such attempt allows the researcher to distinguish the modal sequences, i.e., those indicating the important features of sedimentation process, from sequences whose frequent appearance results exclusively from the dominance of particular lithological varieties of rocks.

**Key words:** cyclicity of sedimentation, modal sequences, statistical permutation test

## WPROWADZENIE

Jednym z istotnych elementów analizy profili litologicznych jest poszukiwanie w nich takich sekwencji warstw (odmian litologicznych, litofacji), którym można przypisać jednoznacznie interpretację genetyczną. Są one zwykle nazywane cyklotemami lub cyklami sedymentacyjnymi (depozycyjnymi) i odzwierciedlają określone następstwo zdarzeń składających się na proces sedymentacji.

Przystępując do tego typu analizy profilu, można przyjąć jedną z dwóch strategii. Pierwsza polega na tym, że z góry zakładamy, jakie sekwencje nas interesują, i sprawdzamy, czy (i jak często) występują one w badanym profilu; procedura taka nie nastęrcza żadnych trudności, jest jednoznaczna i precyzyjnie zdefiniowana. Drugie podejście jest czysto empiryczne: nie czynimy żadnych założeń wstępnych, poprzestając na zliczeniu sekwencji rzeczywiście występujących w profilu; w tym przypadku konieczne jest oczywiście przyjęcie klucza identyfikującego wyróżniane cykle (może to być np. liczba elementów albo określony element graniczny).

Tak czy inaczej, jako wynik otrzymujemy zestawienie częstości występowania w profilu wszystkich wyróżnionych sekwencji; dalsza interpretacja koncentruje się oczywiście na tzw. sekwencjach modalnych, czyli tych, które występują najczęściej. I właśnie w tym miejscu pojawia się problem, będący przedmiotem naszego zainteresowania. Chodzi mianowicie o to, że jakieś sekwencje najczęstsze pojawią się również w profilach czysto losowych. Ważne jest zatem stwierdzenie, czy zaobserwowana częstość występowania danej sekwencji może być dziełem przypadku, czy też wskazuje na rzeczywiste jej uprzywilejowanie.

## METODY POSZUKIWANIA SEKWENCJI MODALNYCH

Sformułowane wyżej zagadnienie było przedmiotem zainteresowania geologów w drugiej połowie ubiegłego wieku (np. Duff & Walton 1962, Zeller 1964, Dembowski & Unrug 1970, Schwarzacher 1975, Krawczyk 1978, Radomski & Gradziński 1979, 1981, Doktor *et al.* 1997), ale później zostało praktycznie zarzucone. Zamiast tego zainteresowano się badaniem następstwa warstw w profilach metodami teorii łańcuchów Markowa, które z czasem stały się niemal standardem w badaniach sedymentologicznych i które do pewnego stopnia zastąpiły klasyczną analizę cykliczności. Rozpowszechniła się mianowicie procedura pośredniego wyszukiwania sekwencji modalnych za pośrednictwem macierzy częstości przejść, pokazującej, jak często ponad warstwą litofacji *X* pojawia się warstwa litofacji *Y*. Jest to jednak procedura zawodna, prowadząca często do fałszywych wniosków.

Aby to wyjaśnić, wyobraźmy sobie profil złożony wyłącznie z powtarzających się sekwencji  $A \rightarrow B \rightarrow D \rightarrow B \rightarrow C \rightarrow A \rightarrow B \rightarrow E \rightarrow B \rightarrow C$ . Macierz częstości przejść dla takiego profilu będzie miała postać

$$W = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0.25 & 0.25 & 0.25 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix},$$

a weryfikacja jej losowości (przy odpowiedniej długości profilu) wskaże, iż statystycznie najbardziej istotne (częstsze od oczekiwanych) są przejścia  $A \rightarrow B$ ,  $B \rightarrow C$  i  $C \rightarrow A$ . Można by więc wnioskować, że podstawową sekwencją modalną jest sekwencja  $A \rightarrow B \rightarrow C$ , która w profilu w ogóle nie występuje.

Jak się wydaje, jedną z przyczyn zaniechania bezpośredniej analizy cykliczności profili litologicznych i zastąpienia jej analizą metodami łańcuchów Markowa mógł być brak odpowiedniego aparatu statystycznego (w odróżnieniu od metod markowowskich, dla których jest on obszerny i dobrze uzasadniony). Chodzi tu zwłaszcza o obiektywną odpowiedź na pytanie, czy zaobserwowane w profilu sekwencje najczęstsze są rzeczywiście sekwencjami modalnymi, informującymi nas o specyfice procesu sedymentacji, czy też po prostu dziełem przypadku. Trzeba bowiem uświadomić sobie, w każdym profilu – również w profilu czysto losowym – częstość występowania różnych sekwencji będzie zależała od łącznej liczby warstw poszczególnych odmian litofacjalnych: sekwencje złożone z odmian liczniejszych będą z natury rzeczy częstsze od takich, które zawierają litofacje pojawiające się w profilu tylko sporadycznie. Nie wystarczy zatem znaleźć sekwencje najczęstsze: trzeba jeszcze upewnić się, że jest ich „za dużo” – więcej niż można by oczekiwać w profilu losowym.

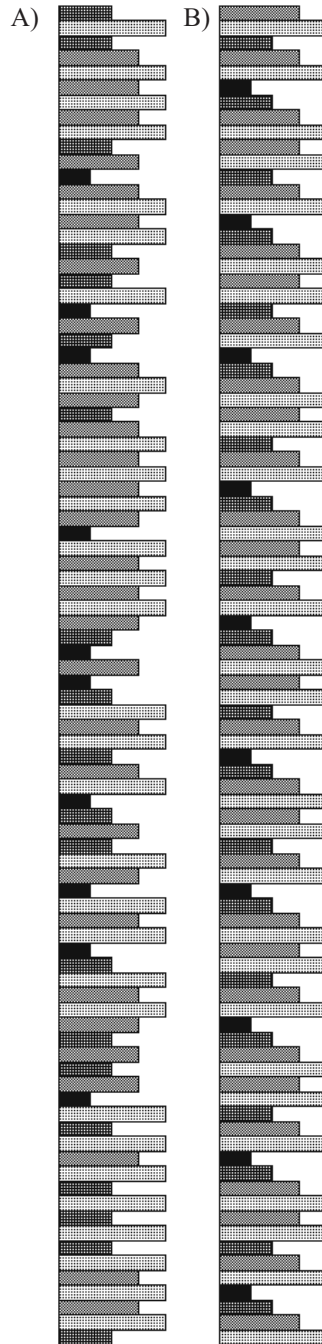
Zilustrujemy to prostym przykładem. Weźmy pod uwagę 90 warstw czterech odmian litologicznych: po 30 warstw litofacji A i B, 20 warstw litofacji C i 10 warstw litofacji D. Utwórzmy z tych warstw dwa modelowe profile – jeden czysto losowy (losujemy bezwrotnie kolejne warstwy z zachowaniem zasady, że żadne dwie kolejne warstwy nie mogą reprezentować tej samej litofacji), a drugi ściśle cykliczny (ustawiamy wszystkie warstwy w kolejności  $A \rightarrow B \rightarrow C \rightarrow D$ ); profile te pokazuje figura 1 (miąższość warstw nie ma dla naszych rozważań żadnego znaczenia i dlatego wszystkie warstwy mają taką samą grubość), a występujące w nich sekwencje dwu-, trzy- i czteroelementowe zestawione są w tabelach 1 i 2.

**Tabela (Table) 1**

Liczba wystąpień różnych sekwencji dwu-, trzy- i czteroelementowych w „profilu” losowym (Fig. 1A)

*Number of various two-, three- and four-element sequences in a random model (Fig. 1A)*

Sekwencja <i>Sequence</i>	Liczba <i>Number</i>	Sekwencja <i>Sequence</i>	Liczba <i>Number</i>	Sekwencja <i>Sequence</i>	Liczba <i>Number</i>
$A \rightarrow B$	18	$A \rightarrow B \rightarrow A$	13	$B \rightarrow A \rightarrow B \rightarrow A$	7
$B \rightarrow A$	18	$B \rightarrow A \rightarrow B$	11	$A \rightarrow B \rightarrow A \rightarrow B$	7
$A \rightarrow C$	9	$B \rightarrow A \rightarrow C$	5	$C \rightarrow A \rightarrow B \rightarrow A$	5
$C \rightarrow A$	9	$C \rightarrow A \rightarrow B$	5	$A \rightarrow B \rightarrow A \rightarrow C$	4
$B \rightarrow C$	8	$B \rightarrow C \rightarrow A$	4	$B \rightarrow C \rightarrow A \rightarrow B$	3
$C \rightarrow B$	7	$A \rightarrow C \rightarrow A$	4	$C \rightarrow B \rightarrow A \rightarrow B$	3
+ 6 innych + 6 others	20	+ 24 inne + 24 others	46	+ 44 inne + 44 others	58



**Fig. 1.** Przykładowe „profile”: losowy (A) i cykliczny (B)

**Fig. 1.** The tested model sequences: random (A) and cyclic (B)

**Tabela (Table) 2**

Liczba wystąpień różnych sekwencji dwu-, trzy- i czteroelementowych w „profilu” cyklicznym (Fig. 1B)

*Number of various two-, three- and four-element sequences in a cyclic model (Fig. 1B)*

Sekwencja <i>Sequence</i>	Liczba <i>Number</i>	Sekwencja <i>Sequence</i>	Liczba <i>Number</i>	Sekwencja <i>Sequence</i>	Liczba <i>Number</i>
A → B	30	A → B → C	20	A → B → C → A	10
B → C	20	B → C → A	10	A → B → C → D	10
C → A	10	B → C → D	10	B → C → A → B	10
C → D	10	C → A → B	10	B → C → D → A	10
D → A	10	C → D → A	10	C → D → A → B	10
B → A	9	D → A → B	10	D → A → B → C	10
		A → B → A	9	A → B → A → B	9
		B → A → B	9	B → A → B → C	9
				C → A → B → A	9

Rozkład częstości sekwencji występujących w modelu cyklicznym jest oczywiście bezpośrednią konsekwencją sposobu, w jaki zostały ustawione warstwy; gdyby był to profil rzeczywisty, rozkład ten odzwierciedlałby mechanizm rządzący procesem sedymentacji. W „profilu” losowym mechanizmu takiego nie ma, a więc widoczne w tabeli 2 zróżnicowanie częstości występowania poszczególnych sekwencji musi być dziełem przypadku. Oznacza to, że wyraźnie dominujące sekwencje A → B, B → A, A → B → A i B → A → B nie są rzeczywistymi sekwencjami modalnymi: ich przewaga jest spowodowana wyłącznie dużą liczbą warstw litofacji A i B. A przecież, gdybyśmy zaobserwowali taki rozkład częstości sekwencji w jakimś prawdziwym profilu, byłibyśmy skłonni uznać to za istotną przesłankę interpretacyjną. Powstaje zatem pytanie o obiektywną metodę odróżniania sekwencji modalnych od tych, które – choć są w profilu najczęstsze – to jednak pojawiają się w nim losowo. Poniżej proponujemy statystyczny test, pozwalający na to pytanie odpowiedzieć.

## PERMUTACYJNY TEST LOSOWOŚCI SEKWENCJI MODALNYCH

Idea testów permutacyjnych zrodziła się w latach 30. ubiegłego wieku i pochodzi od Fishera (1935) i Pitmana (1937). Jej formalny, matematyczny opis zainteresowany (i odpowiednio przygotowany) Czytelnik może znaleźć w cytowanych pracach źródłowych oraz w niektórych podręcznikach statystyki; my postaramy się jedynie wyjaśnić cały tok rozumowania.

Przedstawmy badany profil litologiczny jako ciąg  $\{W_k\}$ ,  $k = 1, 2, \dots, N$ , obserwowanych w nim odmian litologicznych (litofacji)  $L_1, L_2, \dots, L_M$ . Załóżmy, że interesuje nas konkretna sekwencja  $S$ , występująca w profilu  $R_S$  razy. Aby móc uznać ją za sekwencję modalną, musimy wykazać, że w przypadku losowego następstwa warstw naszego profilu

sekwencja  $S$  pojawiałyby się rzadziej. Mówiąc bardziej precyzyjnie – musimy zbadać, jakie jest prawdopodobieństwo zdarzenia polegającego na tym, że w losowym ciągu  $\{V_k\}$ , składającym się z tych samych elementów (warstw) co ciąg  $\{W_k\}$ , znajdziemy  $R_S$  lub więcej sekwencji  $S$ . Jeżeli zdarzenie takie będzie mało prawdopodobne, to wolno nam będzie przypuszczać, że częste występowanie w badanym profilu sekwencji  $S$  nie jest przypadkowe, czyli że jest to rzeczywiście sekwencja modalna.

Poszukiwane prawdopodobieństwo oszacujemy, posługując się symulacyjną metodą Monte Carlo. Wygenerujemy mianowicie  $Q$  losowych „profilu”  $\{V_k\}$  i w każdym z nich sprawdzimy liczbę wystąpień sekwencji  $S$ . Liczba ta (oznaczymy ją przez  $T_S$ ) będzie oczywiście zmienną losową, przyjmującą różne wartości dla różnych „profilu”  $\{V_k\}$ . Mając do dyspozycji  $Q$  takich wartości, będziemy mogli oszacować jej rozkład

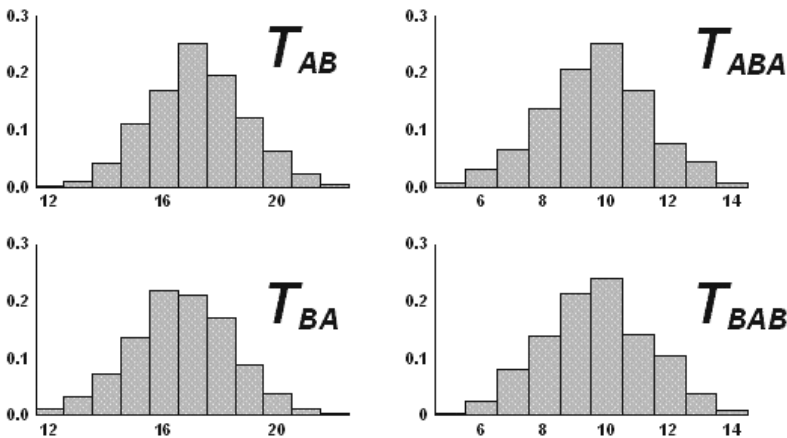
$$P(T_S = t) \approx Q_t / Q \tag{1}$$

gdzie  $Q_t$  jest liczbą „profilu”  $\{V_k\}$ , w których sekwencja  $S$  pojawiła się  $t$ -krotnie.

Zatem prawdopodobieństwo, że w „profilach” losowych wystąpi  $R_S$  lub więcej sekwencji  $S$ , będzie ostatecznie równe

$$P(T_S \geq R_S) = \sum_{t \geq R_S} P(T_S = t) \approx \frac{1}{Q} \sum_{t \geq R_S} Q_t \tag{2}$$

Aby przekonać się, jak w praktyce działa opisana procedura, zastosujemy ją teraz do naszych przykładowych profilu-modeli. W modelu losowym (Fig. 1A) zbadamy cztery najczęstsze sekwencje, tzn.  $A \rightarrow B$ ,  $B \rightarrow A$ ,  $A \rightarrow B \rightarrow A$  oraz  $B \rightarrow A \rightarrow B$ . Stosując wprowadzone wyżej oznaczenia, mamy więc  $R_{AB} = 18$ ,  $R_{BA} = 18$ ,  $R_{ABA} = 13$  i  $R_{BAB} = 11$ . Oszacowane na podstawie  $Q = 1000$  „profilu” losowych rozkłady częstości występowania poszczególnych sekwencji w pokazane są na figurze 2.



**Fig. 2.** Rozkłady częstości sekwencji  $A \rightarrow B$ ,  $B \rightarrow A$ ,  $A \rightarrow B \rightarrow A$  i  $B \rightarrow A \rightarrow B$  (model losowy)

**Fig. 2.** Frequency distributions of successions:  $A \rightarrow B$ ,  $B \rightarrow A$ ,  $A \rightarrow B \rightarrow A$  and  $B \rightarrow A \rightarrow B$  (random model)

Jak widać, wszystkie rozkłady cechują się sporą zmiennością: sekwencje  $A \rightarrow B$  i  $B \rightarrow A$  pojawiają się w „profilach” losowych od 12 do 22 razy, zaś sekwencje  $A \rightarrow B \rightarrow A$  i  $B \rightarrow A \rightarrow B$  – od 5 do 15 razy. Są to jednak rozkłady jednomodalne, o charakterze rozkładów dwumianowych, zatem wartości skrajne są stosunkowo rzadkie. Podstawiając stosowne dane do wzoru (2), otrzymamy:  $P(T_{AB} \geq 18) = 0.410$ ,  $P(T_{BA} \geq 18) = 0.314$ ,  $P(T_{ABA} \geq 13) = 0.052$  i  $P(T_{BAB} \geq 11) = 0.295$ . Tak więc zaobserwowane w naszym profilu-modelu losowym liczby wystąpień badanych sekwencji nie są mało prawdopodobne; można by się zastanawiać jedynie nad sekwencją  $A \rightarrow B \rightarrow A$ , dla której prawdopodobieństwo 13 lub więcej wystąpień wynosi 0.052.

Dla porównania wykonamy podobne obliczenia dotyczące trzech najczęstszych sekwencji ( $A \rightarrow B$ ,  $B \rightarrow C$  i  $A \rightarrow B \rightarrow C$ ) z modelu cyklicznego (Fig. 1B). Otrzymamy  $R_{AB} = 30$ ,  $R_{BC} = 20$  i  $R_{ABC} = 20$ . Rozkłady odpowiednich zmiennych losowych pokazane są na figurze 3, a prawdopodobieństwa  $P(T_{AB} \geq 30) = P(T_{BC} \geq 20) = P(T_{ABC} \geq 20) = 0$ . Różnica w stosunku do modelu losowego jest zatem oczywista.

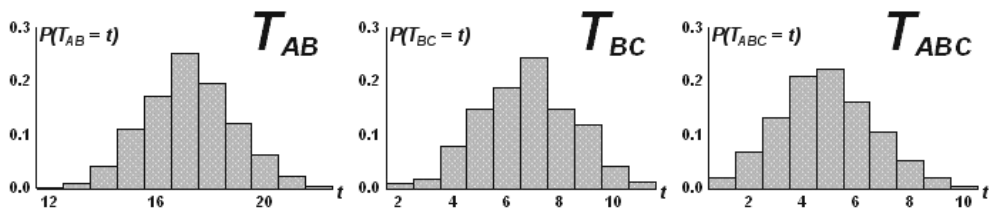


Fig. 3. Rozkłady częstości sekwencji  $A \rightarrow B$ ,  $B \rightarrow C$  i  $A \rightarrow B \rightarrow C$  (model cykliczny)

Fig. 3. Frequency distributions of successions:  $A \rightarrow B$ ,  $B \rightarrow C$  and  $A \rightarrow B \rightarrow C$  (cyclic model)

Na zakończenie pokażemy jeszcze przykład zastosowania testu permutacyjnego do poszukiwania sekwencji modalnych w rzeczywistym profilu litologicznym. Wykorzystamy w tym celu profil HEBCH1, reprezentujący utwory najwyższej części (tzw. warstw porębskich) węglonośnej sukcesji Górnośląskiego Zagłębia Węglowego w Polsce. Obejmuje on 126 wydzieleni ośmiu odmian litologicznych: SH – iłwców i mułwców ilastych (47 warstw), CS – łupków węglowych i iłwców węglistych (3 warstwy), COAL – węgla (33 warstwy), SF – piaskowców drobnoziarnistych (20 warstw), MU – mułwców (17 warstw), SM – piaskowców średnioziarnistych (2 warstwy), SC – piaskowców gruboziarnistych (1 warstwa) i HE – warstw heterolitycznych, czyli wiązek cienkich lamin mułowcowo-ilastych przekładających się z cienkimi laminami piaszczystymi (3 warstwy). Profilowanie zostało wykonane tak, że żadne dwie kolejne warstwy nie reprezentują tej samej odmiany litologicznej. Przystępując do analizy sekwencji (cykli) modalnych musimy najpierw zdefiniować sposób ich wyróżniania w badanym profilu (por. Schwarzscher 1975). Często przyjmuje się jedną z litofacji za element graniczny (początkowy lub końcowy) cyklu, ale najłatwiej jest zacząć od sprawdzenia wszystkich występujących w profilu sekwencji o danej długości. Tak też postąpimy, przy czym ponieważ naszym celem nie jest analiza merytoryczna, a tylko zilustrowanie metody, ograniczymy się do sekwencji dwu-, trzy- i czteroelementowych.



Wyniki przedstawione są w tabeli 3. Naturalnymi kandydatami na sekwencje modalne będą oczywiście w pierwszej kolejności te, które w profilu pojawiają się najczęściej, a więc dwuelementowe cykle COAL  $\rightarrow$  SH i SH  $\rightarrow$  COAL oraz trójelementowy cykl SH  $\rightarrow$  COAL  $\rightarrow$  SH. Opisany wyżej test permutacyjny pozwala na sprawdzenie, czy rzeczywiście częstość ich występowania usprawiedliwia takie przypuszczenie.

**Tabela (Table) 3**

Występujące w profilu HEBCH1 sekwencje dwu-, trzy- i czteroelementowe  
*Two-, three- and four-element sequences encountered in the HEBCH1 column*

Sekwencje 2-elementowe <i>2-element sequences</i>		Sekwencje 3-elementowe <i>3-element sequences</i>		Sekwencje 4-elementowe <i>4-element sequences</i>	
Sekwencja <i>Sequence</i>	Liczba <i>Number</i>	Sekwencja <i>Sequence</i>	Liczba <i>Number</i>	Sekwencja <i>Sequence</i>	Liczba <i>Number</i>
COAL-SH	26	SH-COAL-SH	20	COAL-SH-COAL-SH	9
SH-COAL	23	COAL-SH-COAL	11	SH-COAL-SH-COAL	8
SH-SF	11	COAL-SH-SF	8	SH-COAL-SH-SF	7
SH-MU	10	MU-SH-COAL	6	SF-SH-COAL-SH	6
SF-SH	9	SF-SH-COAL	6	MU-SH-COAL-SH	5
MU-SH	9	COAL-SH-MU	5	COAL-SH-SF-SH	4
SF-MU	6	SH-SF-SH	5	SH-SF-SH-COAL	4
+ 16 innych <i>+16 others</i>	31	+ 37 innych <i>+ 37 others</i>	63	+ 63 inne <i>+ 63 others</i>	80

Po wykonaniu testu okazuje się, że:

- prawdopodobieństwo, że sekwencja COAL  $\rightarrow$  SH wystąpi w profilu losowym 26 lub więcej razy, wynosi 0.0126;
- prawdopodobieństwo, że sekwencja SH  $\rightarrow$  COAL wystąpi w profilu losowym 23 lub więcej razy, wynosi 0.3180;
- prawdopodobieństwo, że sekwencja SH  $\rightarrow$  COAL  $\rightarrow$  SH wystąpi w profilu losowym 20 lub więcej razy, wynosi 0.0134.

Jak należy interpretować te wyniki? Przede wszystkim widać, że nie ma niczego nadzwyczajnego w częstym pojawianiu się sekwencji SH  $\rightarrow$  COAL: równie często (lub częściej) znajdziemy ją w niemal 1/3 losowych przestawień warstw naszego profilu. Inaczej przedstawia się sprawa z sekwencjami COAL  $\rightarrow$  SH i SH  $\rightarrow$  COAL  $\rightarrow$  SH, z którymi związane są znacznie mniejsze prawdopodobieństwa. Używając języka statystyki matematycznej, można powiedzieć, że test permutacyjny służy do weryfikacji hipotezy głoszącej, iż dana sekwencja litofacji pojawia się w profilu losowo, zaś obliczane prawdopodobieństwa są  $p$ -wartościami, czyli najmniejszymi poziomami istotności, przy których wynik testu pozwala na odrzucenie tej hipotezy (patrz np. Koronacki & Mielniczuk 2001). Tak więc hipotezę o losowym charakterze wystąpień sekwencji COAL  $\rightarrow$  SH i SH  $\rightarrow$  COAL  $\rightarrow$  SH możemy odrzucić na poziomie istotności 0.05, ale na poziomie 0.01 nie mamy już podstaw do jej odrzucenia (i wtedy musimy uznać, że żadna z rozpatrywanych sekwencji nie jest



sekwencją modalną, a duża częstość ich występowania wynika po prostu z dużej liczby warstw litofacji SH i COAL).

Analogiczne obliczenia wykonane dla pozostałych sekwencji z tabeli 3 pokazują, że kandydatami na cykle modalne mogłyby być jeszcze: dwuelementowa sekwencja SF → MU i czteroelementowa sekwencja SH → COAL → SH → SF, których częstości występowania w badanym profilu są istotne na poziomie 0.05 (ale nie są już istotne na poziomie 0.01). Jest to oczywiście kwestia do analizy merytorycznej: test niczego nie przesądza, a tylko podpowiada możliwości.

Na koniec warto jeszcze poświęcić chwilę uwagi litofacjom mało licznym, pojawiającym się w profilu w kilku zaledwie warstwach. Pominęliśmy je w dotychczasowych rozważaniach i można bronić poglądu, że na tym należy poprzestać, ponieważ ich sporadyczne wystąpienia nie stanowią próbki statystycznie reprezentatywnej. Z drugiej jednak strony sprawdzenie rozmieszczenia tych nielicznych warstw w profilu może się okazać interesujące. Na przykład w profilu HEBCH1 występują trzy warstwy CS, wszystkie w sekwencjach SH → COAL → CS. Czy jest to tylko przypadek?

## PODSUMOWANIE

Przedstawione rozważania miały na celu zarysowanie problemu wiarygodności szybkiego i bezkrytycznego uznawania za istotne sekwencje modalne sekwencji najczęściej spotykanych w badanym profilu lub – co jeszcze bardziej ryzykowne – sekwencji wydedukowanych na podstawie markowowskich macierzy częstości przejść.

W przykładowym profilu HEBCH1 pojawiająca się aż 23 razy sekwencja SH → COAL wydaje się o wiele lepszym kandydatem na sekwencję modalną aniżeli występująca jedynie sześć razy sekwencja SF → MU. Tymczasem test pokazuje, że prawdopodobieństwo pojawienia się w profilu losowym sześciu par SF → MU jest znacząco mniejsze od prawdopodobieństwa pojawienia się 23 par SH → COAL; wynika to z rozkładu liczebności warstw poszczególnych odmian litologicznych. Tak więc przy sedymentologicznej interpretacji profilu zdecydowanie ważniejsze jest wytłumaczenie pojawiania się sekwencji SF → MU, które – statystycznie rzecz biorąc – nie powinny występować tak licznie.

Zwrócenie bacznej uwagi na przedstawiony problem pozwala uniknąć zbyt dużych uproszczeń w poszukiwaniu i interpretowaniu sekwencji modalnych. Mechaniczne opieranie się tylko na najczęściej występujących sekwencjach może prowadzić do błędnych wniosków, a charakterystyczne sekwencje, wymagające właściwej interpretacji, są pomijane. W szczególności może to mieć istotne znaczenie w wykrywaniu i interpretowaniu zjawisk nietypowych.

## LITERATURA

Dembowski Z. & Unrug R., 1970. Analiza statystyczna sedymentacji cyklicznej w warstwach łaziskich (Górnośląskie Zagłębie Węglowe). *Rocznik Polskiego Towarzystwa Geologicznego*, 40, 63–110.

- Doktor M., Gradziński R. & Słomka T., 1997. Cyclicity in Upper Carboniferous coal-bearing fluvial sediments: example from the Upper Silesia, Poland. *Prace Państwowego Instytutu Geologicznego*, 157, 2, 53–61.
- Duff P.McL.D. & Walton E.K., 1962. Statistical basis for cyclothems: a quantitative study of the sedimentary succession in the East Pennine coalfield. *Sedimentology*, 1, 4, 235–255.
- Koronacki J. & Mielniczuk J., 2001. *Statystyka*. WNT, Warszawa.
- Krawczyk A.J., 1978. Cykle modalne w losowych sekwencjach litostratygraficznych. *Zeszyty Naukowe AGH, Geologia*, 4, 1, 35–38.
- Radomski A. & Gradziński R., 1979. Sekwencje warstwowań w krakowskiej serii piaskowcowej Górnośląskiego Zagłębia Węglowego. *Zeszyty Naukowe AGH, Geologia*, 5, 1, 25–30.
- Radomski A. & Gradziński R., 1981. Facies sequences in the Upper Carboniferous alluvial coal-bearing deposits, Upper Silesia, Poland. *Studia Geologica Polonica*, 68, 29–41.
- Schwarzacher W., 1975. *Sedimentation models and quantitative stratigraphy*. Elsevier, Amsterdam – New York.
- Zeller E.J., 1964. Cycles and psychology. *Kansas Geological Survey Bulletin*, 169, 631–636.

## Summary

Modal sequences play a crucial role as important features characterizing the lithological profiles. Modal sequences are many times repeating sequences of layers (lithological varieties, lithofacies, etc.), which can be unequivocally genetically interpreted. In last years the Markov chains theory has been commonly applied as an indirect identification method of such repeating sequences. Unfortunately, this method appears to be rather unreliable and its application frequently leads to erroneous conclusions. Therefore, the most common sequences should be directly sought in a given lithological profile and such sequences should be regarded as natural candidates to be the modal sequences.

However, objections are raised because some common successions appear also in purely random sequences. For example, figure 1 presents two hypothetical sequences (models of lithological profiles) composed of the same 90 layers of four lithologically different rocks: A, B, C and D. The first sequence (Fig. 1A) represents random succession of layers (random model) whereas in the second one (Fig. 1B) the succession is strictly cyclic ( $A \rightarrow B \rightarrow C \rightarrow D$ ) (cyclic model).

The two-, three- and four-element sequences occurring in both the models are listed in tables 1 and 2. In the cyclic model, distributions of the frequencies of sequences are obviously the result of the applied scheme of layers arrangement. In the random model such arrangement scheme does not exist, by definition. Hence, the diversity of frequencies of sequences seen in table 2 must be a random feature. Consequently, the clear dominance of sequences  $A \rightarrow B$ ,  $B \rightarrow A$ ,  $A \rightarrow B \rightarrow A$  and  $B \rightarrow A \rightarrow B$  is not the representation of true modal sequences but rather results from the dominance of A and B lithofacies.

Therefore, the question arises what is the objective method of discrimination of modal sequences from those which, although most frequent in a given lithological profile, are, in fact, random. The authors propose a statistical permutation test which provides the solution to this problem.

Let us present the studied lithological profile as a sequence  $\{W_k\}$ ,  $k = 1, 2, \dots, N$ , of observed lithological varieties (lithofacies)  $L_1, L_2, \dots, L_M$ . Furthermore, let us focus attention on a particular sequence  $S$ , which occurs  $R_S$  times in the studied profile. Our aim is to estimate the probability that in a random succession  $\{V_k\}$  composed of the same elements (layers) as the sequence  $\{W_k\}$  we will find the  $R_S$  number (or more) of sequences  $S$ . If probability of such an event is low we are allowed to conclude that the common appearance of  $S$  in the studied profile is not a random event, hence, such sequence is the modal sequence.

The probability is estimated with the Monte Carlo simulation method in which a number  $Q$  of random successions  $\{V_k\}$  is generated and in each succession a number of particular sequences  $S$  is checked. The obtained number  $T_S$  is a random variable. Having  $Q$  values of that variable we can estimate its distribution [Formula (1)] and probability that in random successions we will observe the  $R_S$  or more sequences  $S$  [Formula (2)].

If such attempt is applied to our model profiles (see Figs 2, 3) it appears that for random model the probabilities are:  $P(T_{AB} \geq 18) = 0.410$ ,  $P(T_{BA} \geq 18) = 0.314$ ,  $P(T_{ABA} \geq 13) = 0.052$  and  $P(T_{BAB} \geq 11) = 0.295$ . Hence, the numbers of particular sequences observed in this model do not show low probabilities and, consequently, these cannot be regarded as modal sequences. Similar probability calculations for three most frequent successions ( $A \rightarrow B$ ,  $B \rightarrow C$  and  $A \rightarrow B \rightarrow C$ ) encountered in cyclic model give the result  $P(T_{AB} \geq 30) = P(T_{BC} \geq 20) = P(T_{ABC} \geq 20) = 0$ . Thus, the difference in comparison with the random model is obvious.

The proposed statistical permutation test was applied to the search for modal sequences in the HEBCH-1 sequence which represents the uppermost part of the Upper Silesian coal-bearing formation (so-called Poruba Beds) (Tab. 3). The results demonstrate that in this particular lithological profile only the sequences COAL  $\rightarrow$  SH (claystone shales), SF (fine-grained sandstones)  $\rightarrow$  MU (mudstones), SH  $\rightarrow$  COAL  $\rightarrow$  SH and SH  $\rightarrow$  COAL  $\rightarrow$  SH  $\rightarrow$  SF are the modal sequences (at significance level 0.05).

The paper presented above is a methodological attempt. The authors aim to find a convenient but objective, statistical tool, which would enable the sedimentologists to select the important but rare features of studied stratigraphic sequences. The method will be applied to practical problems in the forthcoming papers.