

Krzysztof Labus*

INTERPRETACJA CHEMIZMU WÓD KOPALNIANYCH METODAMI ANALIZY DANYCH ZŁOŻONYCH

1. Wstęp — podstawy metod Analizy Danych Złożonych

Metody analizy danych złożonych (ang. CDA — *Compositional Data Analysis*) tworzą zespół procedur statystycznego opracowywania danych, po raz pierwszy przedstawiony w pracy Aitchisona [1]. Pozwalają one uzyskać wgląd w naturę relacji pomiędzy zmiennymi oraz ułatwiają interpretację trendów tych relacji. Dane złożone można przedstawić jako macierz wektorów (kompozycji) typu [1]

$$x_1 + \dots + x_D = 1$$

Formuła ta oznacza w terminologii CDA kompozycję x o D — składnikach i sumie 1 (jednostkowej). W rezultacie składniki powyższego równania są od siebie zależne, gdyż ich suma jest wartością stałą.

Każdy z wektorów (kompozycji o nieujemnych elementach x_1, \dots, x_D) reprezentuje udział poszczególnych cech (zmiennych) opisujących dany obiekt (próbkę). Dane hydrogeologiczne można niejednokrotnie zdefiniować jako złożone, co w sensie CDA oznacza, iż stanowią one populację przypadków, których wartości zmiennych sumują się do 100%, np. procentowe udziały jonów głównych w próbkach wód.

Informacja zawarta w wektorach jest związana ze względną (wobec całkowitej sumy) zawartością składnika, zatem problemy kompozycyjne mogą być wyrażane w formie proporcji.

Charakterystyczne cechy danych złożonych są następujące [13]:

- zestaw danych złożonych (próbki reprezentujące populację) można przedstawić w formie macierzy;
- każdy rząd macierzy danych odnosi się do pojedynczej próbki (np. wody);

* Wydział Górnictwa i Geologii, Politechnika Śląska, Gliwice

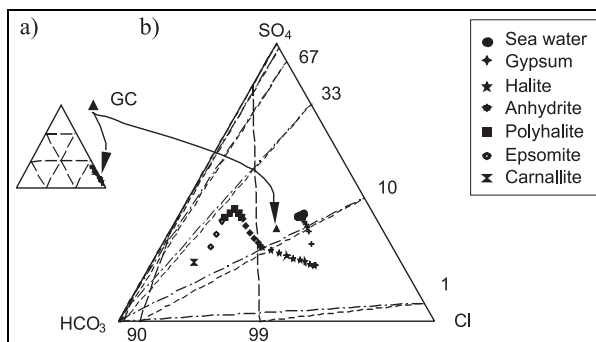
- kolumny macierzy danych przedstawiają pojedyncze składniki (np. jony);
- każdy składnik macierzy jest nieujemny;
- rzędy macierzy sumują się do 1 (np. proporcje) lub do 100 (np. udział procentowy);
- współczynnik korelacji pomiędzy składnikami zmienia się, jeśli jedną ze zmiennych usuniemy z macierzy (lub do niej dodamy) i ponownie zsumujemy rzędy do 100 (operacja subkompozycji).

Proporcje pomiędzy poszczególnymi zmiennymi są zatem „przywiązane” do sumy jednostkowej i niezależne od wielkości próbki. Oznacza to np., iż wzrost udziału danego składnika pociąga za sobą spadek udziałów pozostałych składników próbki [11]. Relacje między udziałami jonów w wodzie podziemnej nie są przecież związane z objętością badanej próbki ani z bezwzględными wartościami stężeń tych jonów. CDA pozwala zatem na interpretację nie tyle wartości poszczególnych zmiennych, ile natury proporcji pomiędzy nimi. Proporcje zmiennych tworzące pewną sumę jednostkową nie poddają się interpretacji klasycznymi metodami np. analizy korelacji. Z drugiej strony istnieją efektywne metody wnioskowania hydrogeologicznego, wykorzystującego zestawy danych złożonych, np. interpretacja procesów hydrochemicznych na podstawie proporcji pomiędzy udziałami jonów głównych (diagramy Piper). W przedstawionej pracy zawarto przykład zastosowania metod CDA w interpretacji i wizualizacji danych dotyczących chemizmu wód kopalnianych.

2. Narzędzia wizualizacji danych złożonych

Jednym z narzędzi wizualizacji używanych w obrębie CDA jest diagram trójkątny. Dla bardziej złożonych kompozycji nie istnieją konwencjonalne i równocześnie efektywne sposoby obrazowania ich zmienności. Zaletą metod CDA jest możliwość graficznej prezentacji analizowanych zbiorów danych na diagramach trójkątnych po sprowadzeniu do przestrzeni trójwymiarowej (dzięki wyodrębnieniu subkompozycji) [2, 3]. Przestrzenią danych złożonych jest simpleks S^D [1]; jeśli $D = 3$ — simpleks można graficznie przedstawić w formie diagramu trójkątnego. Dla przekształcania danych w przestrzeni simpleks przydatne są operacje perturbacji (*perturbation*) oraz transformacji potęgowej (*power transformation*) [4, 8]. Perturbacją kompozycji (wektora) $x = (x_1, x_2, x_3)$ w S^3 przez wektor $p = (p_1, p_2, p_3)$ w S^3 jest nowa kompozycja (wektor) $p \oplus x = C(p_1x_1, p_2x_2, p_3x_3)$ w S^3 , gdzie C oznacza operację tzw. „zamykania” (*closure*). W trakcie tej operacji każda składowa wektora $p \oplus x$ jest dzielona przez sumę wszystkich jego składowych. Perturbacja wektora x przez jego odwrotność $x^{-1} = (1/x_1, 1/x_2, 1/x_3)$ daje w rezultacie element neutralny $e = C(1, 1, 1) = (c/3, c/3, c/3)$, reprezentowany na diagramie trójkątnym przez punkt leżący w jego środku geometrycznym. Ma to znaczenie praktyczne. Perturbacja zestawu danych złożonych, przez odwrotność jej środka geometrycznego (g^{-1}) doprowadza do przeskalowania, prowadzącego do zgrupowania danych wokół środka geometrycznego w diagramie trójkątnym. Umożliwia to zobiektywizowanie stałej przekształcenia, która to zależy od rozproszenia danych wokół średniej geometrycznej kompozycji. Operacja ta pozwala ponadto na transformację linii prostych (np.

linii siatki w pierwotnym diagramie trójkątnym) również na linie proste w diagramie przeskalowanym. Ilustracją graficznych rezultatów perturbacji jest rysunek 1.



Rys. 1. Pierwotny (a) i poddany perturbacji (b) diagram trójkątny subkompozycji anionów dla wody morskiej na różnych etapach ewaporacji (wg [10]): GC — środek geometryczny subkompozycji (0,12; 0,01; 0,87)

Kolejną z możliwych operacji jest transformacja potęgowa. Efektem transformacji potęgowej kompozycji $x = (x_1, x_2, x_3)$ w S^3 przez pewną stałą k jest nowa kompozycja $p \otimes x = C(x_1^k, x_2^k, x_3^k)$ w S^3 , gdzie C oznacza opisaną już operację „zamykania”.

Operacje perturbacji oraz transformacji potęgowej mogą być stosowane łącznie w celu zdefiniowania linii opisujących trendy występujące w populacjach danych złożonych. Dowolny tego rodzaju trend, zwany kompozycyjnym trendem liniowym (*compositional linear trend*), zdefiniowany jest przez pewną kompozycję początkową — a , oraz jednostkową kompozycję — p , określającą kierunek trendu.

W ten sposób dowolna kompozycja y_n znajdująca się w obrębie trendu liniowego może zostać zdefiniowana w S^3 przez wartość stałej k w sposób następujący:

$$y = a \oplus (k \otimes p) = C(a_1 x_1^k, a_2 x_2^k, a_3 x_3^k).$$

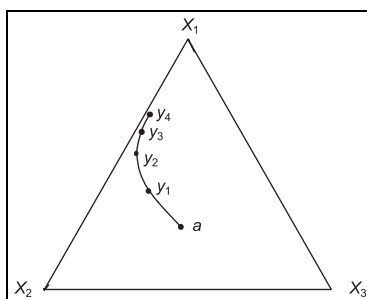
Daje to możliwość tworzenia modeli procesów zmian początkowego składu kompozycji a , prowadzących do uzyskania kompozycji y_n .

Jeżeli pewna kompozycja y_n jest efektem przekształcania początkowej kompozycji a przez kilka procesów o liniowej charakterystyce, możliwa jest interpretacja trendu obserwowanych zmian jako składającego się z kilku nakładających się na siebie trendów indywidualnych [9] (rys. 2) np.:

$$y = a \oplus (k_a \otimes p_a) \oplus (k_b \otimes p_b)$$

gdzie:

- a — kompozycja początkowa,
- k_a i k_b — stałe (krotności) przekształcania,
- p_a i p_b — wektory przekształcania.



Rys. 2. Przykład kompozycyjnego trendu liniowego; y_1 do y_4 — kompozycje powstałe wskutek perturbacji kompozycji początkowej — a (0,25; 0,40, 0,35) przez wektor p (0,50; 0,35; 0,15)

Poszukiwanie trendu liniowego dla populacji danych złożonych może opierać się na analizie składowych głównych. Na diagramach trójkątnych reprezentowane one będą przez osie, które należy wówczas interpretować jako linie regresji w modelu regresji dla danej kompozycji.

3. Przykład zastosowania analizy danych złożonych do interpretacji chemizmu wód kopalnianych KWK „Chwałowice”

Hydrogeochemiczne środowisko kształtowania się wód kopalnianych KWK „Chwałowice” wyróżnia się specyfiką na tle pozostałych kopalń ROW. Wynika ona z oddziaływania na chemizm wód trzeciorzędowego złoża soli, którego fragment znajduje się w zasięgu obszaru górniczego tej kopalni. Wnioskowaniu poddano 322 reprezentatywne zestawy oznaczeń. Wstępna interpretacja, przeprowadzona za pomocą aglomeracji metodą Warda, pozwoliła na wyróżnienie czterech wyraźnych hydrochemicznych grup wód. Założenie o istnieniu co najmniej takiej liczby grup stanowiło podstawę dalszych obliczeń prowadzonych metodą rozmytych k -średnich [5]. Wyróżniono cztery grupy wód. Najbardziej wyraźny podział populacji próbek wód otrzymano wydzielając grupy na podstawie pH, procentowych udziałów jonów głównych oraz wskaźników: $r_{Na/Cl}$, $r_{Ca/Na}$, $r_{Mg/Na}$, r_{Ca/HCO_3} i $r_{Ca/Ca + SO_4}$ (tab. 1).

Grupa 1 charakteryzuje się najniższymi średnimi wartościami: %Mg i $r_{Ca/Ca + SO_4}$ oraz najwyższymi: pH, %K, %HCO₃, $r_{Na/Cl}$ (~1), i obejmuje występujące na poziomach 390 i 550 m wody typu Cl-Na oraz sporadycznie, na poziomie 390 m, typu Cl-SO₄-Na.

Grupa 2 — najwyższe wartości średnie: %Mg, %Ca, %SO₄, $r_{Ca/Na}$; najniższe: mineralizacja, pH, %Na, %Cl i $r_{Na/Cl}$. Zawiera ona odbierane na poziomie 390 m wody dopływające do szybów. Są to wody typu Cl-SO₄-Na-Ca oraz Cl-Na-Ca i Cl-Ca.

Grupa 3 — najwyższe wartości: mineralizacja, %Cl i $r_{Ca/Ca + SO_4}$; najniższe: %SO₄ i %HCO₃. Należą tu głównie wody typu Cl-Na z poziomów od 390 do 700 m.

Grupa 4 — najwyższa wartość średnia: %Na, najniższe: %K, %Ca, $r_{Ca/Na}$; obejmuje stwierdzenia wód typu Cl-Na, z poziomów 390 i 550 m.

TABELA 1

Wartości średnie i odchylenia standardowe parametrów dla grup wód KWK „Chwałowice”

Zmienna	Grupa 1 (n = 21)		Grupa 2 (n = 7)		Grupa 3 (n = 125)		Grupa 4 (n = 169)	
	średnia	odch. std	średnia	odch. std	średnia	odch. std	średnia	odch. std
pH	7,584	0,411	6,850	0,328	7,069	0,522	7,208	0,314
%Na	89,849	2,255	41,820	20,933	82,193	2,504	90,136	1,923
%K	0,880	0,346	0,713	0,327	0,616	0,191	0,470	0,221
%Mg	4,853	1,084	13,108	2,911	9,495	1,296	5,022	1,141
%Ca	4,418	1,708	44,358	18,933	7,695	1,866	4,373	0,744
%Cl	85,075	5,938	78,746	7,703	97,961	2,154	97,103	1,398
%SO ₄	12,352	5,226	19,051	7,816	1,718	1,897	2,353	1,197
%HCO ₃	2,574	1,958	2,202	0,588	0,321	0,297	0,544	0,284
rNa/Cl	1,058	0,112	0,549	0,291	0,842	0,044	0,933	0,028
rCa/Na	0,050	0,021	2,153	2,593	0,094	0,027	0,049	0,009
rCa/Ca+SO ₄	0,270	0,071	0,676	0,152	0,834	0,137	0,666	0,099

Chemizm wód grup 1–4 przedstawiono za pomocą zapisu Kurlowa. W nawiasach ukazano zakresy stężeń wybranych składników swoistych (mg/dm³):

— grupa 1:

$$M^{2,3-65,4} \frac{Cl^{63,8-91,3} SO_4^{6,9-28,8} HCO_3^{0,6-7,9}}{Na^{82,7-93,9} Ca^{2,2-10,9} Mg^{2,6-6,8} K^{0,5-1,4}}$$

(Br^{7,7-67,2} J^{0,6-16,4} Fe^{0,0-11,9} Mn^{0,0-1,1}),

— grupa 2:

$$M^{3,3-13,0} \frac{Cl^{69,5-88,7} SO_4^{8,6-28,7} HCO_3^{1,2-2,9}}{Na^{11,9-57,9} Ca^{30,4-73,9} Mg^{9,3-18,1} K^{0,3-1,2}}$$

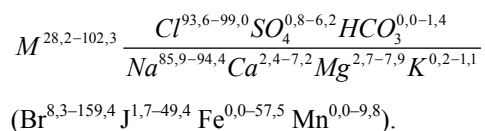
(Br^{3,0-5,6} J^{0,9-1,8} Fe^{0,5-51,5} Mn^{0,0-3,1}),

— grupa 3:

$$M^{18,6-207,2} \frac{Cl^{86,2-99,9} SO_4^{0,0-12,8} HCO_3^{0,0-1,8}}{Na^{69,9-85,9} Ca^{5,4-20,3} Mg^{6,1-13,4} K^{0,1-1,4}}$$

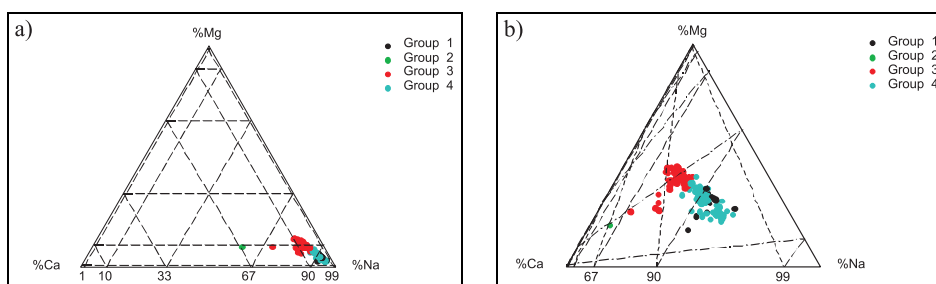
(Br^{23,2-385,0} J^{1,5-86,9} Fe^{0,0-35,3} Mn^{0,0-7,2}),

— grupa 4:

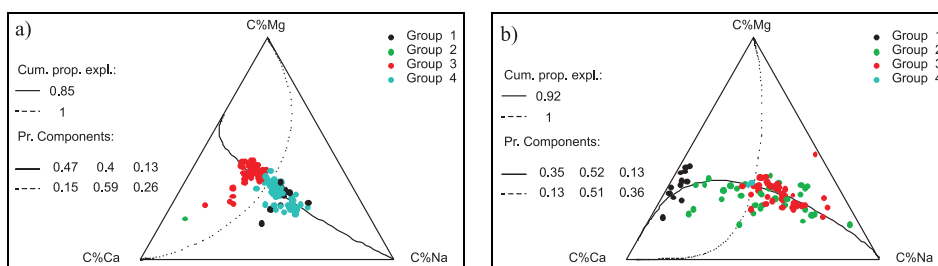


W celu przeanalizowania relacji między udziałami składników wód analizowano wybrane ich subkompozycje, sprowadzone do przestrzeni trójwymiarowej. Wizualizację na diagramach trójkątnych oraz metodą biplot uzyskano dzięki użyciu oprogramowania CoDaPack [14] autorstwa S. Thió-Henestrosa z Universitat de Girona w Hiszpanii.

Rysunek 3a odpowiada trójkątowi kationów w diagramie Pipera. Wskutek nieznacznego udziału Mg i Ca zróżnicowanie proporcji kationów w wodach kopalnianych jest tutaj słabo widoczne. Perturbacja oryginalnego diagramu przez odwrotność jej środka geometrycznego — tzw. centrowanie (rys. 3b), ułatwia wgląd w analizowane relacje, zwłaszcza że układ odniesienia również uległ przeskalowaniu. Na wszystkich diagramach zaznaczono punkty reprezentujące próbki należące do wydzielonych grup.



Rys. 3. Diagramy trójkątne przedstawiające pierwotną (a) i wycentrowaną (b) kompozycję składu wód kopalnianych KWK „Chwałowice” (subkompozycja kationów: Mg-Ca-Na). Określenia Group 1 do 4 odpowiadają wyróżnionym grupom wód



Rys. 4. Diagramy trójkątne przedstawiające składowe główne na tle punktów projekcyjnych wycentrowanej kompozycji składu wód kopalnianych KWK „Chwałowice” (a) i KWK „Rydułtowy” (b); (subkompozycja kationów: Mg-Ca-Na)

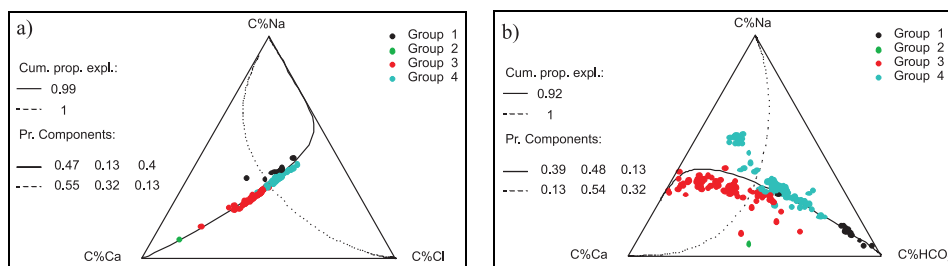
Kolejne diagramy (rys. 4a i 4b) oprócz projekcji punktów reprezentujących „wycen-trowane” subkompozycje, przedstawiają osie wyodrębnionych składowych głównych. Osie te można utożsamiać ze swojego rodzaju liniami regresji w modelu regresji dla danej kom-pozycji [6]. Parametrami każdego z wektorów regresji odpowiadają trzy sumujące się do jed-ności liczby. Oznaczenie „*Cum. prop. expl*” odnosi się do skumulowanej wariancji wyjaś-nionej przez $(n - 1)$ składowych głównych (n — liczba analizowanych zmiennych), przy założeniu, że składowe te wyjaśniają 100% wariancji.

W analizowanym przykładzie pierwsza ze składowych wyjaśnia 85% wariancji. Para-metry wektora regresji wynoszą dla zmiennych %Mg oraz %Ca odpowiednio 0,47 i 0,40. Oznacza to, iż składowa ta w podobnym stopniu odpowiada za wariancję proporcji logaryt-micznych pomiędzy Na (najniższy parametr wektora — 0,13), a każdą z wymienionych zmiennych — $\ln(\%Na/\%Ca)$ i $\ln(\%Na/\%Mg)$. Druga składowa, wyraźnie skorelowana (0,59) ze zmiennością Ca, tłumaczy 15%, czyli pozostałą część wariancji.

Składowa główna wyjaśniająca przeważającą część wariancji (85%) wyznacza w sensie statystycznym trend dla próbek należących do danej subkompozycji. Oś tej składowej mogła-by być interpretowana wprost jako linia trendu hydrogeochemicznego jedynie w przypadku, gdyby kolejne próbki stanowiły etapy przemian, prowadzących do modyfikacji pewnego składu początkowego wód. W rzeczywistości tak wyznaczony trend statystyczny opisuje zarówno próbki spełniające powyższe kryterium, jak i nienależące do trendów hydroche-micznych, a nawet i te, których subpopulacje związane są z trendami przeciwnymi.

Przydatność interpretowania głównych składowych jako modeli trendu polega zatem na możliwości wyznaczenia za ich pomocą prawidłowości odnoszących się do zbiorów pró-bek, pod warunkiem weryfikacji ich spójności hydrochemicznej. Zgodnie z tym podejściem np. wody grupy 3 i 4 (zawierające próbki o silnej mineralizacji) prezentują trend zmian udział-ów Na na rzecz Mg oraz (w drugiej kolejności) Ca. Zależność tę należy wyjaśnić procesa-mi odwrotnej wymiany jonowej; ukierunkowanie osi trendu w kierunku wierzchołka Mg jest zgodne z tendencją do spadku wskaźnika rCa/Mg wraz ze wzrostem mineralizacji wód [12]. Wyznaczenie składowych głównych daje ponadto możliwość jakościowego porównywania statystycznie zdefiniowanych, w przestrzeni kompozycyjnej, prawidłowości położenia pró-bek, reprezentujących różne środowiska hydrochemiczne. Graficznego przykładu dostarcza porównanie opisywanego diagramu kompozycji Mg-Ca-Na obrazującego wody kopalniane KWK „Chwałowice” (rys. 4a) z odpowiednim diagramem dla wód KWK „Rydułtowy” (rys. 4b).

Rysunek 5a przedstawia subkompozycję, której zmienność może być niemal w całości (99%) wyjaśniona przez jedną tylko składową, w podobnym stopniu skorelowaną z udziałami Na oraz Cl. Niska wartość parametru regresji dla Ca oznacza, iż udział tego kationu w anali-zowanej subkompozycji zmienia się w sposób znaczący, przy prawie niezmiennych propor-cjach pozostałych składników (słabo zróżnicowana wartość wskaźnika rNa/Cl). Interpreta-cja przebiegu linii regresji dla pierwszej ze składowych uzmysławia, iż zmianom proporcji udziałów Na i Ca nie towarzyszą zasadnicze wahania udziału Cl, co odzwierciedla domina-cję procesów wymiany jonowej w kształtowaniu zmienności stosunku molowego rNa/Cl .

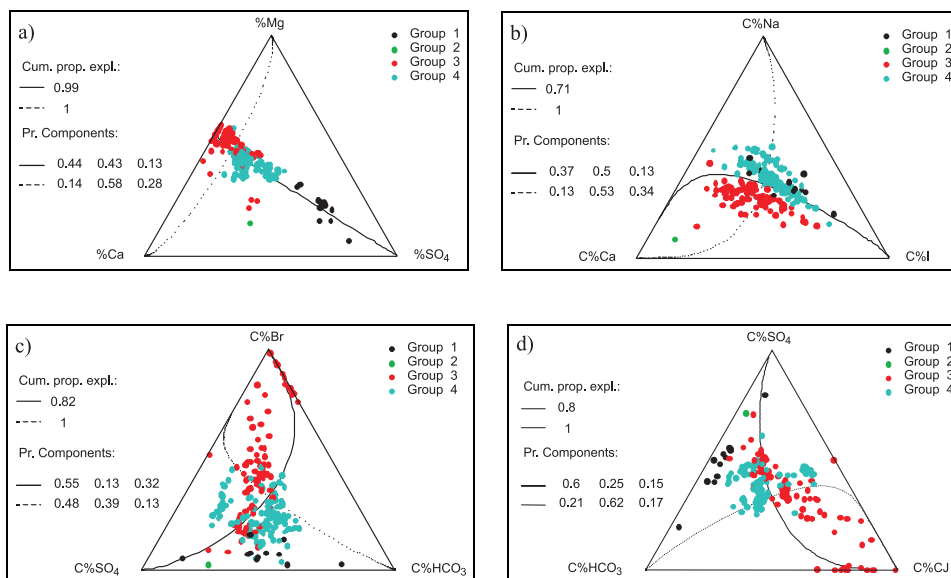


Rys. 5. Diagramy trójkątne przedstawiające składowe główne na tle punktów projekcyjnych wycentrowanych subkompozycji składu wód kopalnianych KWK „Chwałowice”:
a) Na-Ca-Cl; b) Na-Ca-HCO₃

Mimo wysokiej wartości wariancji reprezentowanej przez pierwszą ze składowych — 92%, „wycentrowały” diagram subkompozycji Na-Ca-HCO₃ (rys. 5b) ujawnia zróżnicowanie trendów zmienności składu wód. Odmiennymi trendami charakteryzują się subpopulacje zaliczane do grup 3 i 4. Próbkę zaliczane do grupy 3 cechuje wzrost udziałów Ca, podczas gdy dla grupy 4 zauważalny jest wzrost udziałów Na wraz ze spadkiem udziałów HCO₃. Subkompozycja ta pozwala na wyraźną separację wydzielonych wcześniej grup wód. Punkty projekcyjne próbek układają się w formie dwu ramion. Ramię bliższe wierzchołkowi Na reprezentuje niemal wyłącznie wody grupy 4 poziomu 390 KWK „Chwałowice”, o wskaźniku $r_{Na/Cl}$ zbliżonym do jedności i mineralizacji nie przekraczającej 100g/dm³. Są to najprawdopodobniej wody formowane pod wpływem rozpuszczania halitu. Drugie z ramion tworzą punkty grupy 4, której chemizm formowany lub modyfikowany jest przez odwrotną wymianę jonową — stąd wzrost udziału Ca wraz ze wzrostem mineralizacji i spadkiem udziału wodorowęglanów. Uzyskana linia trendu dla pierwszej ze składowych głównych jest wypadkową pomiędzy liniami regresji opisującymi powyższe procesy. Jej ukierunkowanie w stronę wierzchołka Ca wynika z liczebnej przewagi próbek należących do grupy 4.

Rysunek 6a przedstawia subkompozycję, której zmienność może być w 99% wyjaśniona przez składową jednakowo skorelowaną z pozostającymi w niezmiennych proporcjach udziałami Mg oraz Ca. Niska wartość parametru regresji dla SO₄ dowodzi natomiast, iż udział tego kationu w analizowanej subkompozycji jest silnie zróżnicowany. W subkompozycjach tworzonych przez udziały SO₄ oraz pary kationów zauważalne są stałe relacje między tymi ostatnimi. Potwierdzają to podobne wartości współczynników pierwszej ze składowych głównych. Obserwacja ta oznacza, iż zmienność udziałów jonów SO₄ jest niezależna od zmian udziałów analizowanych kationów (Na, Ca, Mg i K). Jej związek z rozpuszczaniem minerałów siarczanowych jest zatem mało prawdopodobny. Mogłaby być przypisana innym procesom np. redukcji siarczanów w środowisku wód reliktowych lub wytrącaniu BaSO₄ na skutek mieszania z wodami bogatymi w Ba.

Rysunek 6b ilustruje subkompozycję, dla której wyznaczono składowe główne o zróżnicowanych parametrach wektora regresji. Punkty projekcyjne układają się w formie 3 wyraźnych zespołów, równoległych do osi diagramu, przechodzącej przez wierzchołek C%J. Sugeruje to występowanie odpowiedniej liczby grup wód o różniących się proporcjami udziałach Na i Ca przy zbliżonej zmienności udziałów trzeciego ze składników subkompozycji.

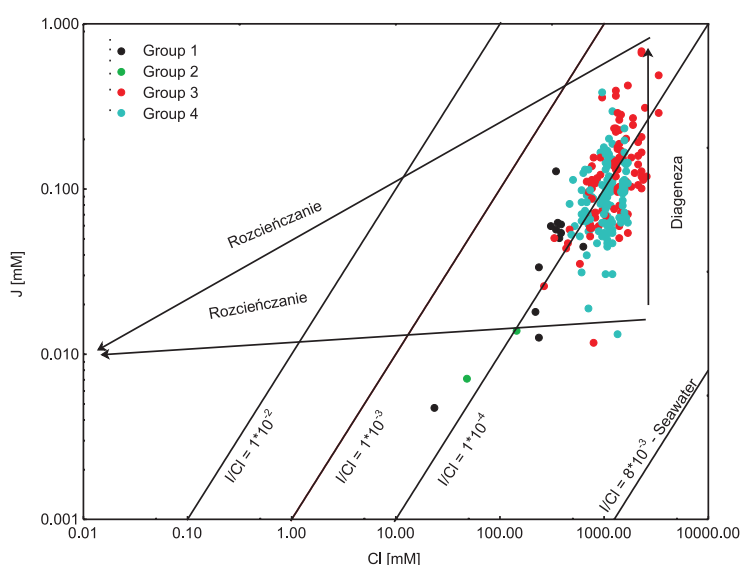


Rys. 6. Diagramy trójkątne przedstawiające składowe główne na tle punktów projekcyjnych subkompozycji składu wód kopalnianych KWK „Chwałowice”: (a) Mg-Ca-SO₄ oraz wycentrowanych subkompozycji: (b) — Na-Ca-J oraz (c) — Br-SO₄-HCO₃ i (d) — SO₄-HCO₃-Cl

Wewnętrzne zróżnicowanie subkompozycji rzutuje na relatywnie niską wartość wariancji wyjaśnionej przez każdą ze składowych głównych — odpowiednio 0,71 i 0,29. Zespół znajdujący się najbliżej wierzchołka Na odwzorowuje wody o najwyższych wartościach wskaźnika $r_{Na/Cl}$, podczas gdy środkowy odpowiada wodom o wartości $r_{Na/Cl}$ zbliżonej do 1. Te ostatnie mogą być określone jako mieszanina wód reliktowych z wodami pochodzącymi z rozpuszczania halitu, obecnego np. w profilu szybu VII KWK „Chwałowice”. Diagram dowodzi, iż zakres zróżnicowania udziałów J może być podobny w wodach o różnej charakterystyce hydrochemicznej. Oznaczałoby to również możliwość różnego udziału wód o jednakowym zaawansowaniu desorpcji jodu z warstwy wodonośnej w mieszaninach, jakimi są wody kopalniane. Procesy diagenetyczne serii węglonośnej mogą powodować wzrost koncentracji jodu bez wpływu na stężenia chlorków [7]. Diagram zależności Cl vs. J dla wód KWK „Chwałowice” zaprezentowano na rysunku 7. Wtórne zmiany stężeń obydwu składników są związane z późniejszym rozcieńczeniem i procesami mieszania wód.

Subkompozycje Br-SO₄-HCO₃ i SO₄-HCO₃-Cl (rys. 6c i 6d) prezentują odmienne relacje pomiędzy proporcjami SO₄-HCO₃ a trzecim elementem subkompozycji. W przypadku kompozycji Br-SO₄-HCO₃ trzeci z elementów — Br jest słabo zmienny w stosunku do pozostałych. Świadczą o tym wysokie wartości parametru regresji wyznaczonego dlań w obrębie obydwu składowych głównych — odpowiednio 0,55 i 0,48. Przeciwna jest prawidłowość dla kompozycji SO₄-HCO₃-Cl, gdzie trzeci z elementów — Cl — wykazuje największą zmienność (parametry 0,15 0,17). Tak wyraźne różnice pomiędzy opisanymi kom-

pozycjami wskazują, iż udziały jonów chlorkowych i bromkowych w badanej populacji charakteryzuje brak ścisłej korelacji. Nie oznacza to jednak braku więzi korelacyjnych wewnątrz podzbiorów, jakimi są wydzielone wcześniej grupy. Udziały zmiennych wewnątrz obydwu analizowanych subkompozycji są wobec siebie niezależne, czego dowodzi zilustrowana diagramami struktura subkompozycji oraz wartości pierwszej składowej głównej. Wobec tego procesy sterujące udziałami jonów Br i Cl w analizowanych wodach należy także uznać za w znacznym stopniu niezależne albo też dopatrywać się istnienia zjawisk powodujących odmienne zachowanie obydwu jonów — np. rozpuszczania halitu (ubogiego w Br) lub ewaporacji wody morskiej.



Rys. 7. Diagram zależności Cl vs. J w wodach KWK „Chwałowice”

4. Podsumowanie

Zaprezentowane metody danych złożonych poszerzają możliwości interpretacji danych chemizmu wód kopalnianych, w przypadku gdy wartości zmiennych (np. udziały jonów) sumują się do 100% lub do jedności. Zespół metod CDA:

- 1) pozwala zobrazować wzajemne relacje między zmiennymi, a także pozwala na interpretację trendów tych relacji;
- 2) daje dodatkowe możliwości interpretacji wynikające z zastosowania przekształceń (przeskalowania diagramów) w operacji perturbacji lub centrowania;
- 3) umożliwia wizualizację wyników w połączeniu z interpretacją zmienności zbioru przy zastosowaniu metody składowych głównych na diagramach trójkątnych;

- 4) pozwala na analizowanie związków pomiędzy zmiennymi, których liczba jest większą od trzech (jak np. w tradycyjnych diagramach Piper).

W przedstawionym przykładzie metody CDA okazały się skuteczne we wspomaganiu interpretacji hydrochemicznych danych kopalnianych i identyfikowaniu następujących procesów, np. wymiany jonowej, rozpuszczania halitu, redukcji siarczanów, wytracania barytu, desorpcji jodu z warstwy wodonośnej.

LITERATURA

- [1] *Aitchison J.*: The Statistical Analysis of Compositional Data. Monographs on Statistics and Applied Probability, London, Chapman and Hall 1986
- [2] *Aitchison J.*: A Concise Guide to Compositional Data Analysis. Girona, CDA Workshop 2003
- [3] *Aitchison J.*: The Statistical Analysis of Compositional Data. New Jersey, The Blackburn Press 2003
- [4] *Barceló-Vidal C., Martín-Fernández J.A., Pawłowsky-Glahn V.*: Mathematical foundations of compositional data analysis. W: Ross, G. (red.) Proceedings of IAMG'01 — The sixth annual conference of the International Association for Mathematical Geology, 20, 2001
- [5] *Bezdek, J.C.*: Pattern Recognition with Fuzzy Objective Function Algorithms. New York, Plenum Press, 1981
- [6] *Billheimer D., Guttorp P., Fagan W.F.*: Statistical Analysis and interpretation of discrete compositional data. Seattle, NRCSE Technical Reports Series. No. 011, 1998
- [7] *Chagué-Goff C., Fyfe W.S.*: Geochemical and petrographical characteristics of a domed bog, Nova Scotia; a modern analogue for temperate coal deposits. Organic Geochemistry, 24, 1996, 141–158
- [8] *von Eynatten, H., Pawłowsky-Glahn V., Egozcue J.J.*: Understanding perturbation on the simplex: a simple method to better visualise and interpret compositional data in ternary diagrams. Mathematical Geology, 34(3), 2002, 246–258
- [9] *von Eynatten H., Barceló-Vidal C., Pawłowsky-Glahn V.*: Modelling Compositional Change: The Example of Chemical Weathering of Granitoid Rocks, Mathematical Geology, 35 (3), 2003, 231–251
- [10] *Labus K.*: Possibilities of Compositional Data Analysis application in hydrogeology. Proc. XII Narodni Hydrogeologický Kongres, České Budejovice 19–22 IX 2005
- [11] *Labus M.*: Compositional Data Analysis as a tool for interpretation of rock porosity parameters, Geol.Quart., 49(3), 2005, 347–354
- [12] *Płochniński Z., Ważny H.*: Wody magnezowe Polski na tle geochemii magnezu. Kwartalnik Geologiczny. 15, 1, 1971
- [13] *Reyment R.A., Savazzi E.*: Aspects of Multivariate Statistical Analysis in Geology, Elsevier 1999
- [14] *Thió-Henestrosa, S., Martín-Fernández, J.A.*: Dealing with compositional data. Mathematical Geology, 37(7), 2005, 777–797

