

PORÓWNANIE WYBRANYCH STATYSTYCZNYCH METOD ANALIZY DANYCH HYDROGEOCHEMICZNYCH NA PRZYKŁADZIE KENOZOICZNYCH POZIOMÓW WODONOŚNYCH W OBSZARZE ZLEWNI GWDY

COMPARISON OF SELECTED STATISTICAL METHODS OF HYDROGEOCHEMICAL DATA ANALYSIS, EXEMPLIFIED BY CENOZOIC AQUIFERS WITHIN THE GWDA RIVER BASIN

TOMASZ KOTOWSKI¹

Abstrakt. W pracy zaprezentowano porównanie wybranych metod wnioskowania statystycznego, stosowanych w badaniach hydrogeochemicznych. Metody statystyczne zostały zastosowane w celu wstępnej identyfikacji procesów formowania składu chemicznego wód poziomów mioceńskiego i głębokiego plejstoceniowego na obszarze zlewni Gwdy. W rezultacie przeprowadzonej analizy skupień otrzymano trzy grupy (skupienia) czynników kształtujących skład chemiczny wód podziemnych. W wyniku analizy składowych głównych i analizy czynnikowej wyróżniono zbiory dwóch czynników kształtujących skład chemiczny wód podziemnych. Wszystkie zastosowane metody okazały się wrażliwe na występowanie anomalii hydrogeochemicznych. Z tego względu, w przypadku analizy skupień i analizy składowych głównych, niemożliwe było dokonanie w pełni poprawnego opisu chemizmu wód podziemnych w obszarze badań. Najlepsze rezultaty osiągnięto przy zastosowaniu analizy czynnikowej. Wyniki analiz statystycznych odniesiono do rezultatów dodatkowych badań: składu izotopowego wody, stężeń siarczanów, trytu oraz badań mikrobiologicznych. Zastosowane metody statystyczne okazały się pomocne, jednakże interpretacja uzyskanych, szczególnie zróżnicowanych danych, jest utrudniona.

Słowa kluczowe: hydrogeochemia, analiza skupień, analiza składowych głównych, analiza czynnikowa.

Abstract. Comparison of selected methods of hydrogeochemical statistical analysis is presented in the paper. Statistical methods were applied to initially identify the formation processes of groundwater chemical composition within the Gwda river basin. As a result of a cluster analysis, three cluster sets of factors forming the chemical composition of the groundwater were distinguished. A set of two factors forming the chemical composition of the groundwater was distinguished as a result of a principal component analysis and a factor analysis. All the statistical methods used were found out to be susceptible upon hydrogeochemical anomalies. The observed susceptibility of the cluster analysis and the principal component analysis allowed a fully correct description of the formation of the groundwater chemical composition in the study area. The best results were achieved using the factor analysis, but this method is also susceptible upon the anomaly cases. The results of the statistical analysis were compared with the research results of additional factors such as the isotopic composition of water and concentrations of sulfates and tritium. The statistical methods appeared to be helpful. However, the interpretation of research results without the additional factors is intractable, especially in the case of hydrogeochemical anomalies.

Key words: hydrogeochemistry, cluster analysis, principal component analysis, factor analysis.

WSTĘP

Celem pracy jest porównanie często wykorzystywanych w analizie danych hydrogeochemicznych metod wnioskowania statystycznego: hierarchicznej analizy skupień, anali-

zy składowych głównych oraz analizy czynnikowej. Metody wnioskowania statystycznego wykorzystano pomocniczo w trakcie identyfikacji procesów kształtujących skład che-

¹ GSMiE PAN, ul. Wybickiego 7, 31-261 Kraków; e-mail: kotowski.tomasz@gmail.com

miczny wód podziemnych głębokich poziomów wodonośnych w obszarze zlewni Gwdy (Pomorze Zachodnie). Większość oznaczeń parametrów fizykochemicznych wód podziemnych wykonano w jednym laboratorium z wykorzysta-

niem jednolitej metodyki i w niewielkim przedziale czasowym. Interpretację wyników wnioskowania statystycznego przeprowadzono przy zastosowaniu dodatkowych metod badawczych: analizy izotopowej i badań mikrobiologicznych.

CHARAKTERYSTYKA ANALIZOWANYCH DANYCH

Chemizm wód podziemnych w obszarze zlewni Gwdy określono na podstawie wyników analiz fizykochemicznych wykonanych podczas realizacji arkusza Mapy hydrogeologicznej Polski (MhP) 1:50 000 w obszarze zlewni Gwdy: Bobolice (Kreczko, Prusak, 2004); Biały Bór, Koczała (Kreczko, 2004a, b); Barvice (Wiśniowski, 1998); Czarne, Przechlewo, Rzeczenica, Szczecinek (Prussak, 2004a–d); Czaplinek (Bielecka, 2004); Łubowo (Bielec, Operacz, 2004); Sulinowo (Rubin, Rubin, 2004); Debrzno, Okonek (Prussak, Prussak, 2004a, b); Człuchów (Balcer, Jankowski, 2000); Mirosławiec (Wojciechowska, 2004); Nadarzyce, Szwecja, Wąlcz (Marcinek, 2004a–c); Jastrowie (Krawiec, 2004); Wy-

soka, Złotów (Kachnic, Kotowski, 2004a, b); Tuczo (Malinowska-Pisz, 2004); Piła, Stara Łubianka, Śmiłowo (Wijura, 2004a–c); Krajenka (Pomianowska, 2004) oraz Szamocin (Stryczyński, 2000). Analizy chemiczne wykonano w niewielkim przedziale czasowym, większość została wykonana w latach 2003–2004 (fig. 1). Podczas weryfikacji danych odrzucono wyniki analiz chemicznych obciążone błędem bilansu jonowego przekraczającym 5%, wyniki zdublowane oraz uzyskane dla próbek pobranych w sposób niewłaściwy. W rezultacie utworzono bazę danych obejmującą 101 wyników rozszerzonych analiz chemicznych (28 wskaźników) wykonanych głównie (ponad 90%) przez Centralne Labora-

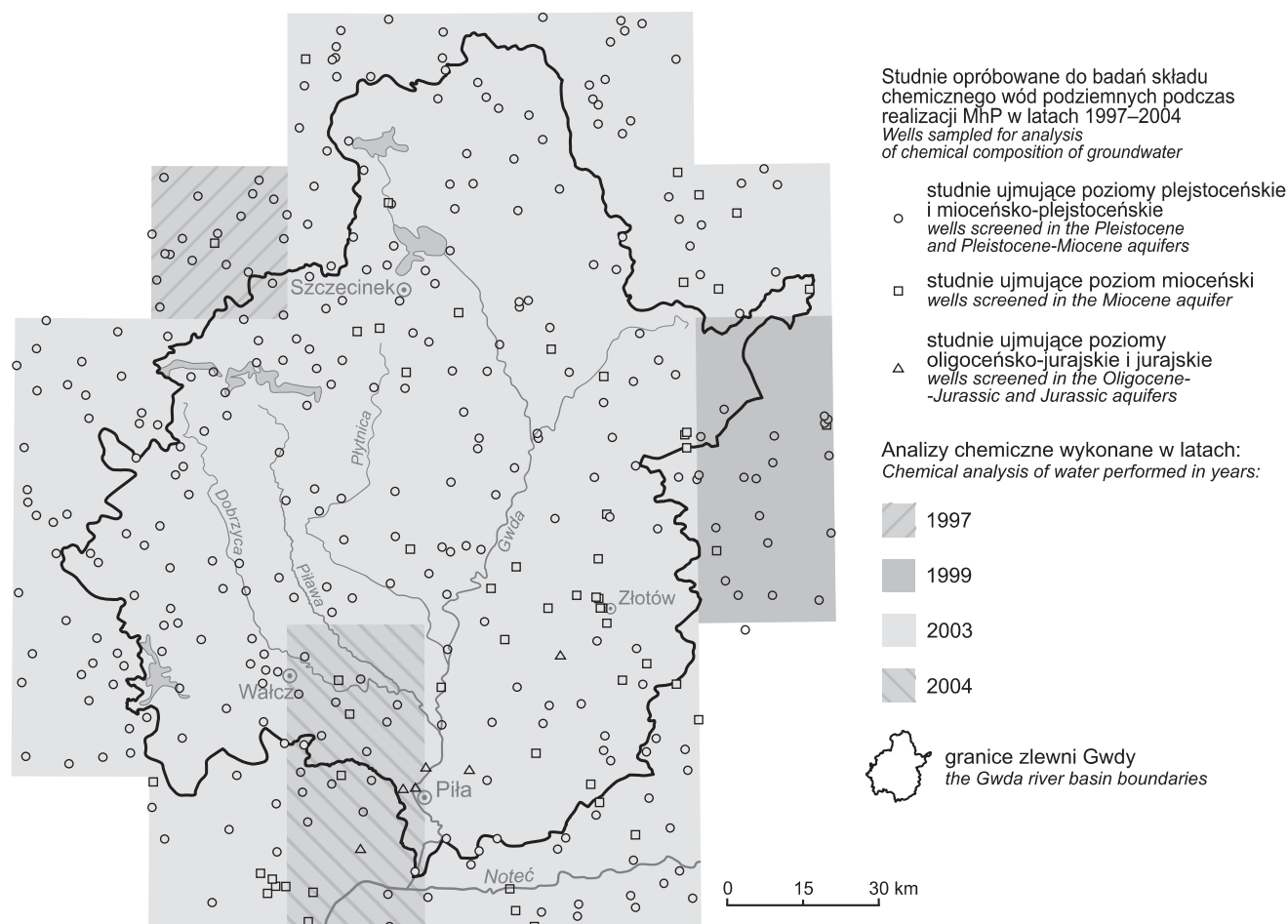


Fig. 1. Lokalizacja obszaru badań i punktów poboru próbek

Location of study area and sampling sites

torium Chemiczne Państwowego Instytutu Geologicznego. Baza danych zawiera wyniki analiz z głębokich poziomów wodonośnych, których strop występuje przeważnie na głębokości przekraczającej 100 m. Jest to poziom dolnoplejstoceni pozostający w bezpośrednim kontakcie hydrau-

licznym z poziomem mioceni oraz poziom mioceni i mioceni-plejstoceni. Szersza charakterystyka warunków hydrogeologicznych została przedstawiona we wcześniejszych pracach, dotyczących obszaru zlewni Gwdy (Kotowski, 2005; 2008; Kotowski, Kachnic, 2007).

CHEMIZM WÓD POZIOMÓW WODONOŚNYCH NEOGENU I DOLNEGO PLEJSTOCENU

Tło hydrogeochemiczne wód podziemnych poziomów neogenu i dolnego plejstocenu określono, korzystając z wykresu prawdopodobieństwa (fig. 2). Należy dodać, iż nie

wszystkie zbiory wyników, uwzględnionych na wykresie prawdopodobieństwa, charakteryzują się rozkładem normalnym lub logarytmiczno-normalnym. Dla wszystkich zbiorów

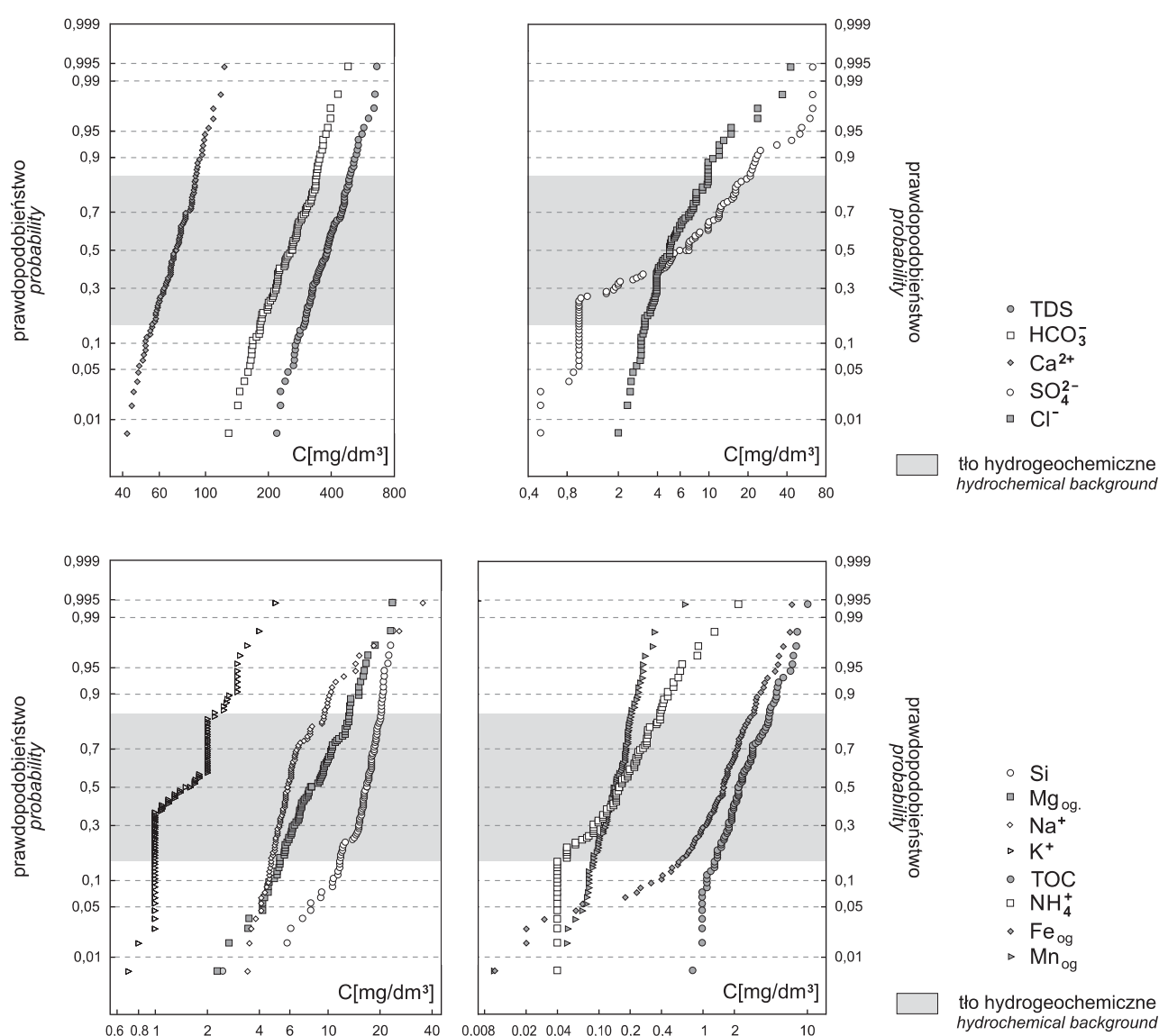


Fig. 2. Zmienność stężeń głównych jonów i zawartości węgla organicznego (TOC) na wykresie prawdopodobieństwa w wodach poziomów miocenu i dolnego plejstocenu

Variability of major ion concentrations and Total Organic Carbon (TOC) in groundwater of the Miocene and lower Pleistocene aquifers, shown on the probability curve plot

rów wyników oznaczeń wykonano testy chi-kwadrat Kołmogorowa-Smirnova z poprawką Lillieforsa (Lilliefors, 1967) oraz Shapiro-Wilka (Shapiro i in., 1968). Wyniki analiz wskazują, że jedynie zbiory wartości stężeń jonów HCO_3^- i Ca^{2+} oraz zbiór wartości mineralizacji ogólnej wód (TDS) charakteryzują się rozkładem nieodbiegającym znacząco od normalnego przy założonym poziomie istotności $\alpha = 0,05$ i obliczonym poziomie prawdopodobieństwa testowego $p = 0,05$. Niektóre zbiory wyników oznaczeń jonów charakteryzują się bardzo wyraźnym, dwu- lub wielomodalnym rozkładem częstości występowania. Spowodowane jest to m.in. dużą liczbą jednakowych wyników. Są to najczęściej wartości

granicy oznaczalności i/lub wykrywalności metody analitycznej zastosowanej do wykonania oznaczeń.

W dolnych warstwach wodonośnych plejstocenu stwierdzono występowanie podwyższonych w stosunku do tła hydrogeochemicznego stężeń jonów SO_4^{2-} i Cl^- . Najwyższe stężenia jonów SO_4^{2-} i Cl^- stwierdzono w próbkach pochodzących z miejscowości Bądecz i Wysoka, gdzie od chwili rozpoczęcia eksploatacji studni obserwuje się stały wzrost stężeń tych jonów (Kotowski, Kachnic, 2007; Kotowski, 2008). Stężenia jonów Cl^- i SO_4^{2-} mieszczą się w przedziale wartości określonych dla wód przeznaczonych do picia (Rozporządzenie..., 2007).

ZASTOSOWANE METODY STATYSTYCZNE WRAZ Z INTERPRETACJĄ OTRZYMANYCH WYNIKÓW

Jako jedną z metod pomocniczych w analizie procesów kształtujących skład chemiczny wód głębokich poziomów wodonośnych wykorzystano często stosowaną w hydrogeochemii hierarchiczną analizę skupień (Güler i in., 2002). Metoda ta umożliwia przede wszystkim rozdzielenie zbioru danych na grupy zawierające podobne zmienne. Do analizy zbioru 17 zmiennych wykorzystano metodę Warda. Polega ona na oszacowaniu odległości pomiędzy skupieniami poprzez analizę wariancji. Metoda ta jest uznawana za bardzo efektywną (Stanisz, 2007). Graficzną prezentację uzyskanych wyników przedstawiono na poniższym rysunku (fig. 3). Zgrupowanie w skupieniu G1 jonów SO_4^{2-} , Cl^- , NO_3^- oraz TOC i pH można byłoby zinterpretować jako związane z dopływem wód zanieczyszczonych antropogenicznie. W tym przypadku wpływ migracji zanieczyszczeń antropogenicznych na chemizm wód głębokich poziomów wodonośnych może zaznaczać się lokalnie, ale na pewno nie jest czynnikiem dominującym. W wodach podziemnych poziomów głą-

bokich, poza nielicznymi przypadkami, nie stwierdzono lub stwierdzono bardzo niskie stężenia innych parametrów zanieczyszczeń antropogenicznych (fosforanów, azotanów lub produktów ich redukcji). W rejonach, gdzie stwierdzono najwyższe stężenia jonów Cl^- i SO_4^{2-} , dopływ wód zanieczyszczonych antropogenicznie ma znikome znaczenie z uwagi na jednocześnie obserwowany w badanych próbkach brak podwyższonych zawartości trytu (maksymalna wynosi 1,2 TU) oraz innych wskaźników antropopresji (Kotowski, 2008). Włączenie do skupienia G1 jonów NO_3^- oraz TOC i pH związane jest z wrażliwością analizy skupień na występowanie wartości wyraźnie różniących się od pozostałych.

Skupienie G2 można interpretować jako związane głównie z procesami rozpuszczania minerałów węglanowych i glinokrzemianów. Z uwagi na charakter tej analizy, szersza interpretacja czynnika związanego z rozpuszczaniem minerałów zostanie przedstawiona w dalszej części pracy.

Zgrupowanie w skupieniu G3 zmiennych: Fe_{og} , Mn_{og} i NH_4^+ oraz głębokości stropu warstwy wodonośnej, mogą wskazywać, iż czynnikiem odpowiedzialnym za korelację tych elementów jest zmiana warunków oksydacyjno-redukcyjnych wraz z głębokością. Jednak analiza zmiennych zestawionych w skupieniu G3 wskazuje, że taka interpretacja jest nieuzasadniona (fig. 4). Brak istotnej zależności pomiędzy zmiennymi Fe_{og} , Mn_{og} i NH_4^+ a głębokością warstwy wodonośnej jest wyraźny. Można jedynie wyróżnić bardzo słabą korelację (współczynnik korelacji wynosi 0,35) dla cząstek $\text{SiO}_2 \cdot \text{H}_2\text{O}$, która jednak nie ma związku ze zmianą warunków oksydacyjno-redukcyjnych wraz z głębokością.

Jak wynika z powyższych spostrzeżeń, hierarchiczna analiza skupień umożliwiła wyodrębnienie trzech grup zmiennych: G1, G2 i G3. Jednakże interpretacja zgrupowanych zmiennych jako rezultat oddziaływania procesów kształtujących chemizm wód jest znacznie utrudniona i niejednoznaczna. Dlatego w celu dalszej analizy chemizmu wód zastosowano metodę składowych głównych. Liczebność zbioru, wynosząca 101 wyników, jest w pełni wystarczająca do przeprowadzenia analizy składowych głównych (Tabachnick, Fidell, 1996). Zastosowany test Barletta (1954)

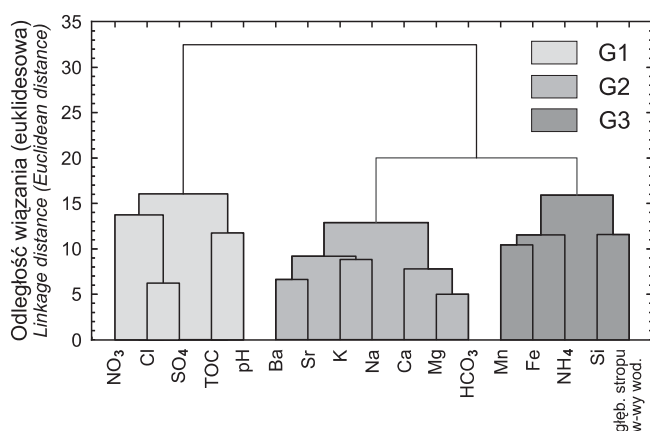


Fig. 3. Diagram wyników hierarchicznej analizy skupień dla wód z głębokich poziomów wodonośnych

Diagram showing the results of hierarchical cluster analysis of deep aquifers

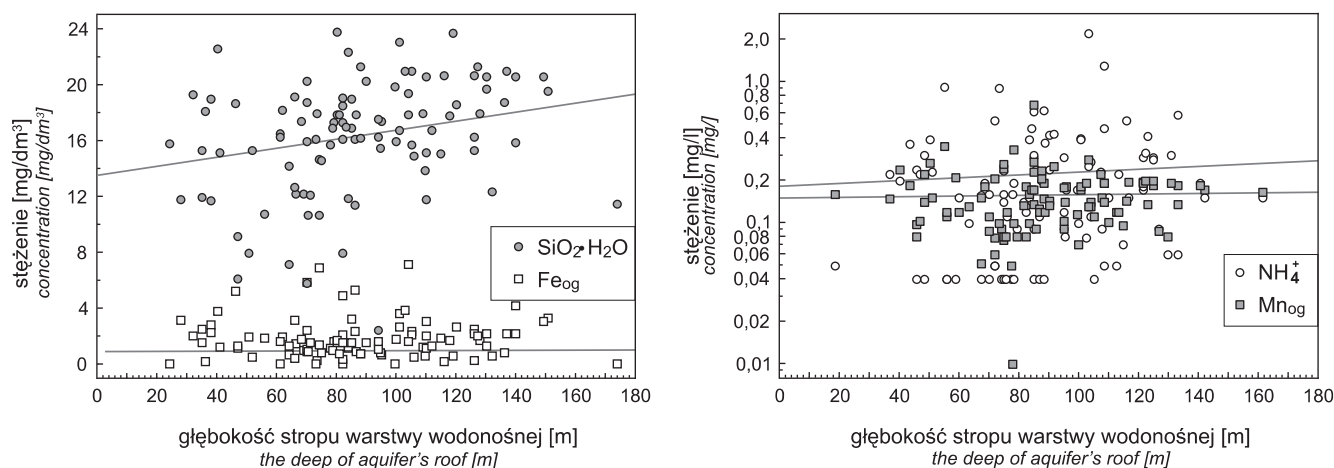


Fig. 4. Wykres stężeń jonów: Fe_{og} , Mn_{og} , NH_4^+ i drobin $\text{SiO}_2 \cdot \text{H}_2\text{O}$ w funkcji głębokości stropu warstwy wodonośnej

The plot of Fe_T , Mn_T , NH_4^+ ions concentrations and $\text{SiO}_2 \cdot \text{H}_2\text{O}$ corpuscles vs. the depth to the aquifer

i analiza współczynnika Kaisera (1958) potwierdziły zasadność przeprowadzenia analizy składowych głównych. Badanie dotyczyło 17 zmiennych. Analiza większej ich liczby była niecelowa z uwagi na skrajnie odbiegający od normalnego rozkład ich wartości lub bardzo małe zróżnicowanie wartości dla poszczególnych zmiennych. Uwzględniając występowanie zmiennych o znacznie różniących się wartościach, analizę składowych głównych prowadzono z wykorzystaniem macierzy korelacji (Stanisz, 2007). Analizę ob-

serwowanych zależności w formowaniu się składu chemicznego wód podziemnych rozpoczęto od określenia liczby czynników wpływających na chemizm wód warstw wodonośnych miocenu i dolnego plejstocenu. W tym celu zastosowano test osypiska (Cattell, 1966) oraz, pomocniczo, kryterium Kaisera. Wykres wartości własnych macierzy korelacji uzyskany w teście osypiska Cattella (fig. 5A) wskazuje na występowanie dwóch istotnych czynników, które charakteryzują w 49,84% zmienność wariancji analizowanego

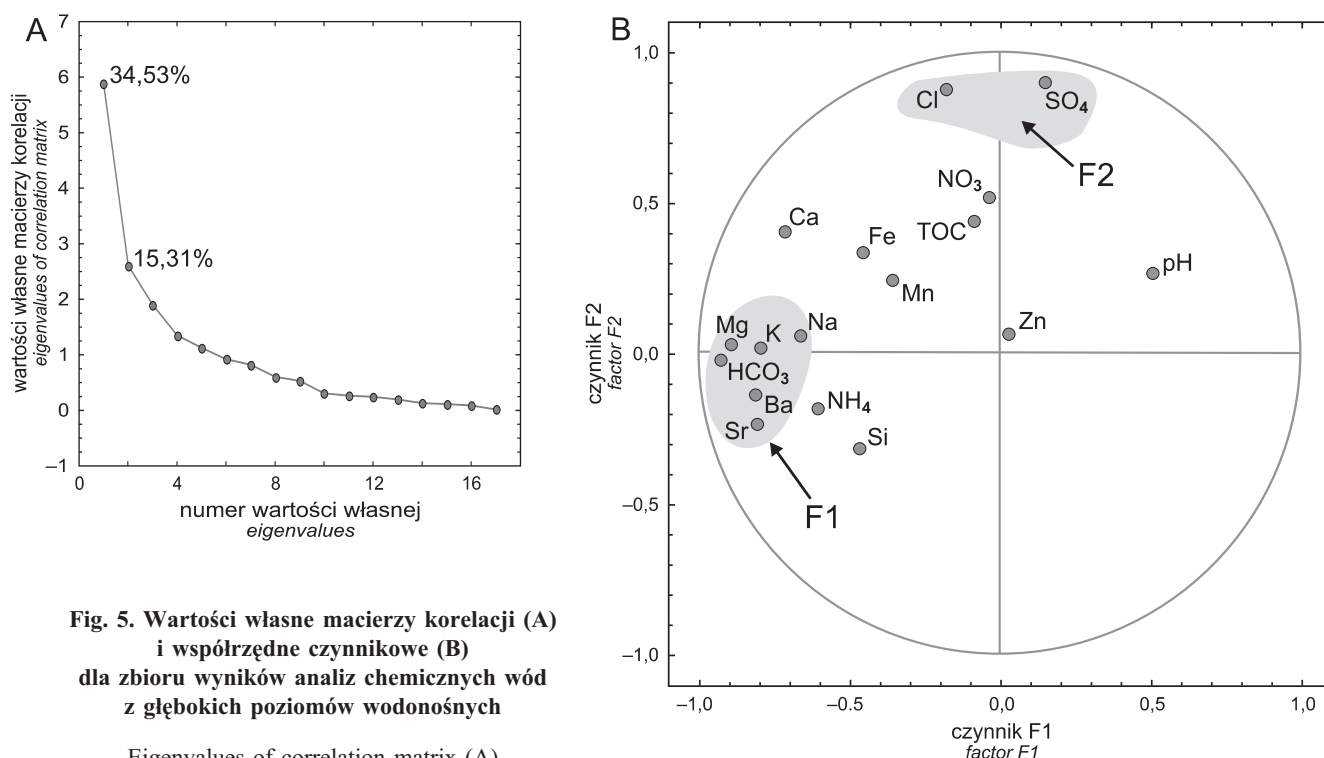


Fig. 5. Wartości własne macierzy korelacji (A) i współrzędne czynnikowe (B) dla zbioru wyników analiz chemicznych wód z głębokich poziomów wodonośnych

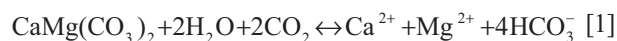
Eigenvalues of correlation matrix (A) and factor loadings (B) for a set of results of chemical analysis from the deep aquifers

zbioru danych. Ta stosunkowo niska wartość wynika z występowania w zbiorze danych „przypadków odstających”.

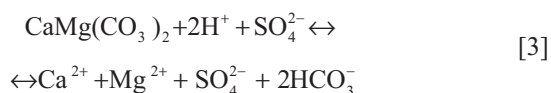
Uzyskany wykres współrzędnych czynnikowych (fig. 5B) wskazuje na występowanie dwóch czynników, które w zasadniczy sposób wpływają na korelację pomiędzy zmiennymi, tj. wartościami parametrów fizykochemicznych i stężeń poszczególnych jonów. Są to czynniki F1 i F2 o najwyższych wartościach własnych, wynoszących odpowiednio 5,87 i 2,60. Można je interpretować jako główne procesy kształtujące skład chemiczny wód warstw miocenu i dolnego plejstoenu. Czynniki F1 odzwierciedla geogeniczne procesy formowania składu chemicznego wód, związane przede wszystkim z rozpuszczaniem minerałów (głównie węglanów) występujących w matrycy skalnej ośrodka wodonośnego. Należy oczywiście pamiętać, że procesy nasycania się wód podziemnych względem węglanów z największą intensywnością zachodzą w płytszych poziomach wodonośnych. Obecność w zbiorze F1 jonów Na^+ czy K^+ wskazuje na możliwość rozpuszczania glinokrzemianów, występujących licznie w matrycy skalnej ośrodka wodonośnego lub nadkładzie izolującym poziomy wodonośny (Masłowska, 1999). Niski wskaźnik nasycenia wód względem glinokrzemianów – albitu (mediana -1,8), określony dla wód głębokich warstw wodonośnych, potwierdza taką interpretację (Kotowski, 2008).

Analiza składowych głównych nie pozwoliła w tym przypadku na wyodrębnienie wszystkich głównych jonów biorących udział w procesach rozpuszczania minerałów. Jest to widoczne w przypadku jonu Ca^{2+} . Brak wyraźnego związku jonu Ca^{2+} z czynnikami F1 i F2 (fig. 5B) jest w głównej mierze spowodowany wrażliwością analizy składowych głównych na wartości anomalne. Podwyższone

względem tła hydrogeochemicznego stężenia jonów Ca^{2+} (ponad 100 mg/dm^3) są w tym przypadku rezultatem występowania procesów intensyfikujących rozpuszczanie minerałów węglanowych. Zachodzą one równolegle z powszechnie występującymi procesami rozpuszczania minerałów węglanowych:



Procesy intensyfikacji rozpuszczania minerałów węglanowych związane są w obszarze badań głównie z utlenianiem pirytu (dostarczającym jonów SO_4^{2-}) lub, sporadycznie, dopływem wód zanieczyszczonych siarczanami pochodzenia antropogenicznego. Zachodzą one zgodnie z poniższymi reakcjami:



Współwystępowanie tych procesów ilustrują wykresy teoretycznej rozpuszczalności węglanów (fig. 6), sporządzone na podstawie zależności stechiometrycznych w równaniach reakcji [1]–[4].

Większość punktów oznaczających relacje pomiędzy jonami Ca^{2+} , Mg^{2+} i HCO_3^- w badanych próbkach wody znajduje się w pobliżu teoretycznej linii rozpuszczania węglanów

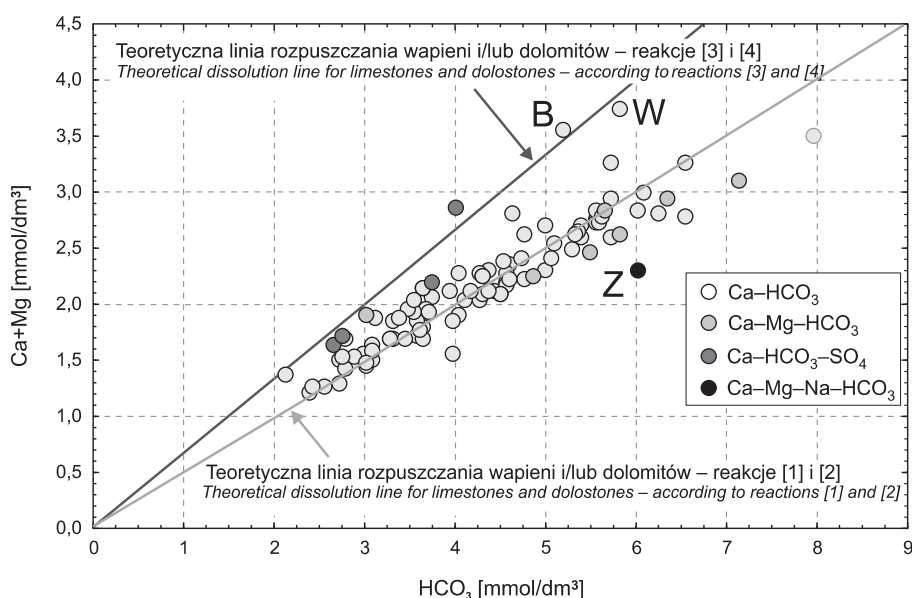


Fig. 6. Stężenia jonów: Ca^{2+} , Mg^{2+} i HCO_3^- w wodach z głębokich poziomów wodonośnych w odniesieniu do teoretycznych linii rozpuszczania węglanów

Concentrations of Ca^{2+} , Mg^{2+} and HCO_3^- ions in the deep aquifers in relation to a theoretical curve of carbonate solubility

nów. Wskazuje to, iż proces rozpuszczania opisany reakcjami [1] i [2] ma dominujące znaczenie w kształtowaniu chemizmu wód głębokich warstw wodonośnych.

Na wykresie widać także wysokie stężenia jonów Ca^{2+} i Mg^{2+} , które nie mogły być spowodowane wyłącznie procesami rozpuszczania węglanów zgodnie z reakcjami [1] i [2]. Najwyższe stężenia tych jonów obserwujemy w punktach oznaczonych jako B i W. Są to wyniki oznaczeń pochodzące z miejscowości Bądecz i Wysoka. Stwierdzono tam występowanie wód wzbogaconych w jony SO_4^{2-} w wyniku geogenicznych procesów utleniania pirytu (Kotowski, Burkowska, 2009). Punkty te znajdują się prawie na krzywej teoretycznej, wyznaczonej na podstawie procesów opisanych reakcjami [1]–[4] przy założeniu, że zachodzą równolegle i pozostają w równowadze między sobą. Współwystępowanie tych procesów powoduje także lokalny wzrost mineralizacji.

Na wykresie (fig. 6) widoczne są także punkty wskazujące na występowanie niższych stężeń jonów Ca^{2+} , niż wynikałoby to z rozpuszczania węglanów. Można to interpretować jako bardzo prawdopodobne występowanie procesów wymiany jonowej między Na^+ i Ca^{2+} , związanej z występowaniem w matrycy skalnej ośrodka wodonośnego minerałów ilastych. Taki właśnie przypadek obserwowany jest m.in. w próbce oznaczonej symbolem Z, pochodzącej z miejscowości Zamarte. Stwierdzono tam wyraźnie podwyższoną zawartość jonów Na^+ i typ hydrogeochemiczny wody Ca-Mg-Na-HCO_3 . W wyniku analizy składowych głównych występowanie procesów wymiany jonowej nie zostało wyraźnie wyodrębnione.

Czynnik F2, pomimo niewielkiej liczby zgrupowanych zmiennych (SO_4^{2-} i Cl^-), odzwierciedla wpływ na chemizm wód głębokich warstw wodonośnych procesów o odmiennej genezie, np. ascencji i/lub utleniania pirytu. Z uwagi na wystąpienie analogicznego zgrupowania jonów SO_4^{2-} i Cl^- w rezultacie zastosowania analizy czynnikowej, wspomniana korelacja zostanie szerzej omówiona w dalszej części pracy.

Zastosowano również analizę czynnikową, w celu interpretacji procesów kształtujących skład chemiczny wód głębokich warstw wodonośnych. Wykorzystano ją, uwzględniając powyższe spostrzeżenia dotyczące rezultatów zastosowania analizy składowych głównych oraz obserwowanych procesów intensyfikujących rozpuszczanie węglanów. Metoda ta polega m.in. na dekompozycji badanych zmiennych w nowy zbiór zmiennych nieskorelowanych i następnie wyodrębnienie czynników wspólnych i swoistych dla takiego zbioru. W odróżnieniu od metody składowych głównych, analiza czynnikowa zakłada istotną redukcję zmiennych (Comrey, 1973). Wykorzystanie w hydrogeologii analizy czynnikowej oraz jej wady i zalety były przedmiotem licznych publikacji (Macioszczyk, 1975; Usunoff, Guzman-Guzman, 1989; Ruiz i in., 1990; Jayakumar, Siraz, 1997). Dla zbioru danych o liczebności $n = 101$ zastosowano standaryzację, a następnie transformację metodą SMC (Squared Multiply Correlation), która pozwala na uwzględnienie wpływu czynnika swoistego zbioru (Stanisz, 2007). Przeanalizowano 17 zmiennych. Liczbę bada-

nych zmiennych i czynników określono na podstawie rozważań i zastrzeżeń analogicznych, jak w przypadku analizy składowych głównych. W celu uzyskania czytelnego obrazu układu wartości czynnikowych, uzyskane wyniki poddano rotacji znormalizowanej varimax. W wyniku obliczeń otrzymano dwa czynniki, które w celu rozróżnienia określono jako X1 i X2, ze względu na zastosowaną metodę analizy statystycznej. Czynniki te charakteryzują w 47,2% zmienność wariancji analizowanego zbioru danych. Tutaj, podobnie jak w przypadku metody składowych głównych, uzyskana niska wartość wynika głównie z wrażliwości zastosowanej metody na obecność „przypadków odstających”. Wartości własne czynników wynoszą: $X1 = 5,67$ i $X2 = 2,35$. Wyniki analizy czynnikowej w formie wykresu ładunków czynnikowych przedstawiono na figurze 7. Rezultat jest bardzo zbliżony do wyniku uzyskanego przy użyciu metody składowych głównych. Jednak w tym wypadku jon Ca^{2+} został wyraźnie, przy stosunkowo wysokiej wartości ładunku czynnikowego (ponad 0,66), zgrupowany w zbiorze X1.

Czynnik X1 odzwierciedla, analogicznie do omówionego wcześniej czynnika F1, procesy związane głównie z rozpuszczaniem minerałów lub z wymianą jonową. Niskie stężenia jonów SO_4^{2-} , niskie wartości SI określone dla gipsu (mediana –2,7) i anhydrytu (mediana –3,0) oraz analiza wskaźników hydrogeochemicznych $[\text{Ca}/(\text{Ca} + \text{SO}_4)]$ wskazują, iż głównym źródłem jonów Ca^{2+} dla wód z głębokich poziomów wodonośnych są węglany i krzemiany. Wartości wyżej wymienionych wskaźników mogą również wynikać z faktu, że w nadkładzie izolującym te poziomy nie występują ily pstre, często zawierające konkracje gipsowe i anhydrytowe.

Podobnie jak w przypadku metody składowych głównych, dla jonów Sr^{2+} zgrupowanych w czynniku X1 przeprowadzona analiza wskaźnika $\text{Ca}^{2+}/\text{Sr}^{2+}$ wskazuje na rozpuszczanie wapieni jako główne źródło pochodzenia jonów Sr^{2+} .

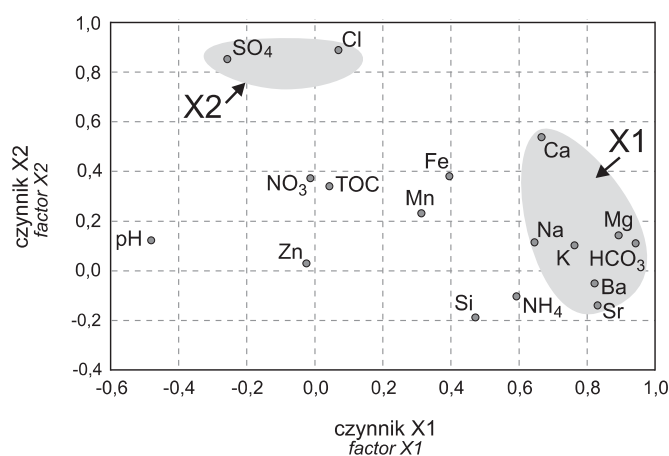


Fig. 7. Wykres ładunków czynnikowych dla wód głębokich poziomów wodonośnych

The plot of factor loadings for deep aquifers

(Kotowski, 2005). Zgrupowanie w jednym czynniku (X1) kilku procesów geogenicznych jest także związane z wrażliwością analizy czynnikowej na obecność w zbiorze danych wyraźnie różniących się od pozostałych. Podobna interpretacja procesów kształtujących chemizm wód poziomu mioceńskiego została przeprowadzona dla mniejszego obszaru, znajdującego się w granicach zlewni Gwdy (Kotowski, 2005). Zastosowanie analizy czynnikowej umożliwiło rozdzielanie procesów geogenicznych na: rozpuszczanie minerałów, wymianę jonową oraz adsorpcję i desorpcję. Jednak w analizowanym wówczas zbiorze wyniki analiz chemicznych nie były znacznie zróżnicowane.

W zbiorze X2 najwyższe ładunki czynnikowe mają jony SO_4^{2-} i Cl^- . Trudno jednoznacznie wskazać proces lub grupę procesów, które mogą decydować o zawartości tych jonów w wodach podziemnych. W świetle wyników badań izotopowych i mikrobiologicznych (Kotowski, 2007; Kotowski, Burkowska, 2009) można jednak stwierdzić, że za wzrost stężenia

jonów Cl^- w głębokich warstwach wodonośnych odpowiadają głównie procesy związane z ascencją wód z poziomów mezozoicznych lub sporadycznie — dopływ wód zanieczyszczonych antropogenicznie. Głównym zaś procesem odpowiedzialnym za wzrost stężenia jonów SO_4^{2-} w wodach głębokich warstw wodonośnych jest utlenianie piryty lub, rzadziej, dopływ wód zanieczyszczonych antropogenicznie.

Analiza wartości ładunków czynnikowych poszczególnych jonów określonych w zbiorze X2 wskazuje na występowanie stosunkowo wysokiej wartości dla jonów Ca^{2+} . Świadczy to o istotnej zależności między jonami SO_4^{2-} , Cl^- i Ca^{2+} . Z uwagi na brak w środowisku wód podziemnych interakcji pomiędzy jonami Cl^- i Ca^{2+} i/lub minerałami, które zawierają w swym składzie wspomniane jony, obserwowana zależność dotyczy w głównej mierze jonów SO_4^{2-} i Ca^{2+} . Jest to wprost proporcjonalna korelacja między nimi, związana przede wszystkim z procesami intensyfikującymi rozpuszczanie minerałów węglanowych (fig. 5).

PODSUMOWANIE

Zgodnie z przyjętymi założeniami, podczas identyfikacji procesów kształtujących skład chemiczny wód podziemnych na obszarze badań, metody statystyczne wykorzystano jako pomocnicze. Wstępne zastosowanie hierarchicznej analizy skupień miało na celu wyodrębnienie jednorodnych grup (subpopulacji), które poddaje się dalszej, osobnej analizie statystycznej. Jednakże, często wykorzystywana w badaniach przyrodniczych, hierarchiczna analiza skupień w tym przypadku nie pozwoliła na poprawne wyodrębnienie procesów kształtujących chemizm wód podziemnych. Metoda ta, jak zaobserwowano, jest zbyt wrażliwa na obecność w analizowanym zbiorze wyników analiz chemicznych, które wyraźnie różnią się od pozostałych. Dlatego dalszą analizę tą metodą zarzucono i w celu właściwej klasyfikacji obiektów nie podjęto się już zastosowania innych, niehierarchicznych odmian analizy skupień (np. często wykorzystywanej metody k-średnich).

Zastosowanie analizy składowych głównych pozwoliło na poprawne wyodrębnienie czynników, które można uznać za procesy kształtujące skład chemiczny, jednakże nie zostały prawidłowo wyodrębnione wszystkie jony biorące udział w tych procesach. Podobnie jak w przypadku analizy skupień, chociaż w nieco mniejszym stopniu, wystąpił problem „przypadków odstających”. Widoczne występowanie procesów wymiany jonowej także nie zostało wyraźnie wyodrębnione w wyniku analizy składowych głównych.

Najlepsze rezultaty wnioskowania statystycznego uzyskano w wyniku zastosowania analizy czynnikowej. Metoda ta umożliwiła poprawne wyodrębnienie czynników, które można interpretować jako grupy procesów kształtujących skład chemiczny wód podziemnych na badanym obszarze. Analiza czynnikowa okazała się najmniej wrażliwa na występowanie wartości anomalnych w analizowanych zbiorach. Podobnie jak w przypadku analizy składowych głównych,

niska wartość wyodrębnionych czynników, charakteryzujących zmienność wariancji zbioru danych, wynika głównie z wrażliwości tej metody na obecność „przypadków odstających”. Tu zapewne wspomniana wrażliwość uniemożliwiła wyodrębnienie procesów wymiany jonowej.

We wszystkich zastosowanych metodach wnioskowania statystycznego wystąpił problem „przypadków odstających”. Pominiecie takich wyników zniekształciłoby jednak naturalny i charakterystyczny obraz zmienności składu chemicznego wód poziomów miocenu i dolnego plejstocenu. Wyniki analiz chemicznych odbiegające od pozostałych odzwierciedlają bowiem m.in. naturalne procesy hydrogeochemiczne występujące lokalnie, ale zauważalnie wpływające na skład chemiczny wód analizowanych poziomów wodonośnych na obszarze zlewni Gwdy. Warto nadmienić, iż stężenia większości jonów (poza jonami żelaza i manganu) we wszystkich próbkach określonych jako „przypadki odstające” nie przekraczają norm dotyczących jakości wody przeznaczonej do spożycia przez ludzi.

Podczas interpretacji rezultatów wszystkich zastosowanych metod pojawił się także problem zgrupowania w jednym czynniku (skupieniu) zmiennych, których korelacje związane są z różnymi procesami zachodzącymi w ośrodku wodonośnym. Identyfikacja i wyodrębnienie tych procesów przy wykorzystaniu wyłącznie metod statystycznych są praktycznie niemożliwe.

Zastosowanie metod wnioskowania statystycznego w celu interpretacji chemizmu wód okazało się pomocne, jednakże, jak pokazano, nawet w przypadku analizy starannie wyselekcjonowanego i dużego zbioru wyników analiz fizykochemicznych wód podziemnych, ich interpretacja, szczególnie dla danych zróżnicowanych, jest utrudniona. Próba interpretacji uzyskanych wyników bez dodatkowych badań (np. izotopowych) może prowadzić do błędnych wniosków.

LITERATURA

- BALCER M., JANKOWSKI M., 2000 — Mapa hydrogeologiczna Polski w skali 1:50 000, arkusz Człuchów wraz z objaśnieniami. Centr. Arch. Geol. Państw. Inst. Geol., Warszawa [dokument elektroniczny].
- BARTLETT M.S., 1954 — A note on the multiplying factors for various chi-squared approximations. *J. Royal Statistical Soc.*, Ser. B, **16**: 296–298.
- BIELEC B., OPERACZ T., 2004 — Mapa hydrogeologiczna Polski w skali 1:50 000, arkusz Łubowo wraz z objaśnieniami. Centr. Arch. Geol. Państw. Inst. Geol., Warszawa [dokument elektroniczny].
- BIELECKA H., 2004 — Mapa hydrogeologiczna Polski w skali 1:50 000, arkusz Czaplinek wraz z objaśnieniami. Centr. Arch. Geol. Państw. Inst. Geol., Warszawa [dokument elektroniczny].
- CATTELL R.B., 1966 — The scree test for the number of factors. *Multivariate Behavior. Res.*, **1**: 245–276.
- COMREY A.L., 1973 — A first course in factor analysis. Academic Press, New York.
- GÜLER C., THYNE G.D., MCCRAY J.E., TUNER A.K., 2002 — Evaluation of graphical and multivariate statistical methods for classification of water chemistry data. *Hydrogeol. J.*, **10**: 455–474.
- JAYAKUMAR R., SIRAZ L., 1997 — Factor analysis in hydrogeochemistry of coastal aquifers — a preliminary study. *Environ. Geol.*, **31**: 174–177.
- KACHNIC J., KOTOWSKI T., 2004a — Mapa hydrogeologiczna Polski w skali 1:50 000, arkusz Wysoka wraz z objaśnieniami. Centr. Arch. Geol. Państw. Inst. Geol., Warszawa [dokument elektroniczny].
- KACHNIC M., KOTOWSKI T., 2004b — Mapa hydrogeologiczna Polski w skali 1:50 000, arkusz Złotów wraz z objaśnieniami. Centr. Arch. Geol. Państw. Inst. Geol., Warszawa [dokument elektroniczny].
- KAISER H.F., 1958 — The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, **23**: 187–200.
- KOTOWSKI T., 2005 — Chemizm wód podziemnych w północnej części GZWP 127. W: Współczesne problemy hydrogeologii (red. A. Sadurski, A. Krawiec), 12: 353–361. Wyd. UMK, Toruń.
- KOTOWSKI T., 2007 — Skład izotopowy wody oraz siarczanów w głębokim plejstocénskim poziomie wodonośnym w obrębie struktury kopalnej w pobliżu Wysokiej (Pojezierze Krajeńskie). W: Współczesne problemy hydrogeologii (red. A. Szczepański, E. Kmiecik, A. Żurek), 13, 2: 95–103. Wyd. AGH, Kraków.
- KOTOWSKI T., 2008 — Chemizm wód podziemnych w regionalnym systemie krążenia wód podziemnych w zlewni Gwdy (praca doktorska). UMK, Toruń.
- KOTOWSKI T., BURKOWSKA A., 2009 — Geneza siarczanów w głębokim plejstocénskim poziomie wodonośnym w rejonie Wysokiej na podstawie badań izotopowych i mikrobiologicznych. *Biul. Państw. Inst. Geol.*, **436**: 273–280.
- KOTOWSKI T., KACHNIC M., 2007 — Formowanie składu chemicznego wód podziemnych w warstwach miocenu i plejstocenu w rejonie występowania głębokiej doliny kopalnej w pobliżu Wysokiej (Pojezierze Krajeńskie). *Biul. Państw. Inst. Geol.*, **427**: 47–60.
- KRAWIEC A., 2004 — Mapa hydrogeologiczna Polski w skali 1:50 000, arkusz Jastrowie wraz z objaśnieniami. Centr. Arch. Geol. Państw. Inst. Geol., Warszawa [dokument elektroniczny].
- KRECZKO M., 2004a — Mapa hydrogeologiczna Polski w skali 1:50 000, arkusz Biały Bór wraz z objaśnieniami. Centr. Arch. Geol. Państw. Inst. Geol., Warszawa [dokument elektroniczny].
- KRECZKO M., 2004b — Mapa hydrogeologiczna Polski w skali 1:50 000, arkusz Koczała wraz z objaśnieniami. Centr. Arch. Geol. Państw. Inst. Geol., Warszawa [dokument elektroniczny].
- KRECZKO M., PRUSSAK W., 2004 — Mapa hydrogeologiczna Polski w skali 1:50 000, arkusz Bobolice wraz z objaśnieniami. Centr. Arch. Geol. Państw. Inst. Geol., Warszawa [dokument elektroniczny].
- LILLIEFORS H.W., 1967 — On the Kolmogorov–Smirnov test for normality with mean and variance unknown. *J. American Statist. Ass.*, **62**: 399–402.
- MACIOSZCZYK A., 1975 — Zastosowanie analizy czynnikowej, sposób R do interpretacji danych hydrogeochemicznych. *Biul. Geol. Wyd. Geol. UW*, **20**: 47–68.
- MALINOWSKA-PISZ A., 2004 — Mapa hydrogeologiczna Polski w skali 1:50 000, arkusz Tuczo wraz z objaśnieniami. Centr. Arch. Geol. Państw. Inst. Geol., Warszawa [dokument elektroniczny].
- MARCINEK U., 2004a — Mapa hydrogeologiczna Polski w skali 1:50 000, arkusz Nadarzyce wraz z objaśnieniami. Centr. Arch. Geol. Państw. Inst. Geol., Warszawa [dokument elektroniczny].
- MARCINEK U., 2004b — Mapa hydrogeologiczna Polski w skali 1:50 000, arkusz Szwecja wraz z objaśnieniami. Centr. Arch. Geol. Państw. Inst. Geol., Warszawa [dokument elektroniczny].
- MARCINEK U., 2004c — Mapa hydrogeologiczna Polski w skali 1:50 000, arkusz Wałcz wraz z objaśnieniami. Centr. Arch. Geol. Państw. Inst. Geol., Warszawa [dokument elektroniczny].
- MASŁOWSKA M., 1999 — Charakterystyka litologiczna glin zwałowych północno-zachodniej Polski. *Prz. Geol.*, **47**, 10: 920–926.
- POMIANOWSKA H., 2004 — Mapa hydrogeologiczna Polski w skali 1:50 000, arkusz Krajenka wraz z objaśnieniami. Centr. Arch. Geol. Państw. Inst. Geol., Warszawa [dokument elektroniczny].
- PRUSSAK E., 2004a — Mapa hydrogeologiczna Polski w skali 1:50 000, arkusz Czarne wraz z objaśnieniami. Centr. Arch. Geol. Państw. Inst. Geol., Warszawa [dokument elektroniczny].
- PRUSSAK E., 2004B — Mapa hydrogeologiczna Polski w skali 1:50 000, arkusz Przechlewo wraz z objaśnieniami. Centr. Arch. Geol. Państw. Inst. Geol., Warszawa [dokument elektroniczny].
- PRUSSAK E., 2004c — Mapa hydrogeologiczna Polski w skali 1:50 000, arkusz Rzeczenica wraz z objaśnieniami. Centr. Arch. Geol. Państw. Inst. Geol., Warszawa [dokument elektroniczny].
- PRUSSAK E., 2004d — Mapa hydrogeologiczna Polski w skali 1:50 000, arkusz Szczecinek wraz z objaśnieniami. Centr. Arch. Geol. Państw. Inst. Geol., Warszawa [dokument elektroniczny].
- PRUSSAK E., PRUSSAK W., 2004a — Mapa hydrogeologiczna Polski w skali 1:50 000, arkusz Debrzno wraz z objaśnieniami. Centr. Arch. Geol. Państw. Inst. Geol., Warszawa [dokument elektroniczny].
- PRUSSAK E., PRUSSAK W., 2004b — Mapa hydrogeologiczna Polski w skali 1:50 000, arkusz Okonek wraz z objaśnieniami. Centr. Arch. Geol. Państw. Inst. Geol., Warszawa [dokument elektroniczny].
- ROZPORZĄDZENIE Ministra Zdrowia z dnia 29 marca 2007 r. w sprawie wymagań dotyczących jakości wody przeznaczonej do spożycia przez ludzi (DzU nr 61, poz. 417, 2002).
- RUBIN K., RUBIN H., 2004 — Mapa hydrogeologiczna Polski w skali 1:50 000, arkusz Sulinowo wraz z objaśnieniami. Centr. Arch. Geol. Państw. Inst. Geol., Warszawa [dokument elektroniczny].

- RUIZ F., GOMIS V., BLASCO P., 1990 — Application of factor analysis to the hydrogeochemical study of a coastal aquifer. *J. Hydrol.*, **119**: 169–177.
- SHAPIRO S.S., WILK M.B., CHEN H.J., 1968 — A comparative study of various tests for normality. *J. American Statist. Ass.*, **63**: 1343–1372.
- STANISZ A., 2007 — Przystępny kurs statystyki z zastosowaniem STATISTICA PL na przykładach z medycyny, 3 — Analizy wielowariantowe. StatSoft Polska, Kraków.
- STRYCYŃSKI A., 2000 — Mapa hydrogeologiczna Polski w skali 1:50 000, arkusz Szamocin wraz z objaśnieniami. Centr. Arch. Geol. Państw. Inst. Geol., Warszawa [dokument elektroniczny].
- TABACHNICK B.G., FIDELL L.S., 1996 — Using multivariate statistics (3rd edition). Happer & Row, New York.
- USUNOFF E. J., GUZMAN-GUZMAN A., 1989 — Multivariate analysis in hydrochemistry: an example of the use of factor and correspondence analyses. *Ground Water*, **27**: 27–34.
- WIJURA A., 2004a — Mapa hydrogeologiczna Polski w skali 1:50 000, arkusz Piła wraz z objaśnieniami. Centr. Arch. Geol. Państw. Inst. Geol., Warszawa [dokument elektroniczny].
- WIJURA A., 2004b — Mapa hydrogeologiczna Polski w skali 1:50 000, arkusz Stara Łubianka wraz z objaśnieniami. Centr. Arch. Geol. Państw. Inst. Geol., Warszawa [dokument elektroniczny].
- WIJURA A., 2004c — Mapa hydrogeologiczna Polski w skali 1:50 000, arkusz Śmiłowo wraz z objaśnieniami. Centr. Arch. Geol. Państw. Inst. Geol., Warszawa [dokument elektroniczny].
- WIŚNIEWSKI Z., 1998 — Mapa hydrogeologiczna Polski w skali 1:50 000, arkusz Barwice wraz z objaśnieniami. Centr. Arch. Geol. Państw. Inst. Geol., Warszawa [dokument elektroniczny].

SUMMARY

The paper presents a comparison of exemplary methods of statistical analysis of hydrogeochemical data and their application in the case of the aquifers within the Gwda river basin. In the vicinity of Wysoka and Badecz, there is a Pleistocene buried valley intersecting the Miocene aquifer. Despite distinct lithology and origin of the aquifer deposits, the deep Pleistocene aquifer is considered together with the Miocene aquifer as a common element of the groundwater flow system. Groundwater of the deep Pleistocene aquifer contains elevated concentration of chloride and sulfate ions. Analysis of the isotopic composition of sulfates and tritium concentrations as well as hydrogeological conditions and their relation to the concentration of Cl^- and SO_4^{2-} ions observed in regions of intensive human impact indicates that anthropogenic pollutants are not a major source of additional loads of chloride and sulphate ions. Statistical methods were applied in order to identify initial processes of the formation of water chemical composition. As a results of cluster analysis, a set of three clusters forming the chemical composition of groundwater was distinguished (Fig. 3). A set of two factors forming the chemical composition of groundwater was identified as a result of factor analysis and principal component analysis (Figs 5, 7). The major factors F1 and X1 reflect the natural formation processes of the chemical composition of groundwater. These processes are associated mainly with both the solubility of minerals within the shallow Pleistocene aquifers and the percolation of groundwater from the shallow Pleistocene aquifers. The factors F2 and X2 indicate

the correlation between the concentrations of chloride and sulfate ions, and they characterize a different geochemical process. All the statistical methods used were found out to be susceptible upon a hydrogeochemical anomaly. However, refuse the cases of hydrogeochemical anomaly is incorrect because it will make distortion the natural view of chemical composition of waters. The observed of susceptible upon anomaly case in fact is mainly due to natural hydrogeochemical processes and to a lesser degree to anthropogenic pollutants. The results of the investigations and analyses made it possible to estimate the application of exemplary statistical methods to hydrogeochemical researches in the study area. Taking into account the observed susceptibility of the cluster analysis and principal component analysis, their application makes it impossible to describe relatively correctly the formation of the chemical composition of groundwater. By using the factor analysis we can achieve the best results, but this method is also susceptible upon anomaly cases. The results of statistical analysis were compared with the research results of additional factors such as the isotopic composition of water and sulfates and concentration of tritium. Based on the obtained results, the complex evaluation of the formation processes of chemical composition of the Miocene and deep Pleistocene aquifers was made. The differentiated formation processes of chemical composition of the Miocene and deep Pleistocene aquifers are of natural geological character and are related to the current groundwater flow system.