

Primjena faktorske analize u statističkoj obradi podataka geokemijske prospекcije na Papuku i Krndiji

Jasna DRAVEC-BRAUN

Institut za geotehniku, Hinkovićeva 7, YU-42000 Varaždin

Ključne riječi: Faktorska analiza, Geokemijska prospекcija

U radu su prikazani rezultati primjene faktorske analize u obradi podataka geokemijske prospекcije koja je provedena na Papuku i Krndiji u 1980. godini. Primjenjeno je nekoliko modela faktorske analize koje su proizvele slične rezultate. Na taj način definirana međuzavisnost ispitivanih elemenata pružila je korisne informacije kod istraživanja potencijalnih ležišta urana.

Key words: Factor Analysis, Geochemical exploration

In this article, the results of the application of several factor analysis models on the results of reconnaissance geochemical exploration on the Mt. Papuk and Mt. Krndija, are reviewed. It could be concluded, that they produced very similar results which could be useful in uranium ore exploration.

Uvod

Statističke metode obrade analitičkih rezultata u geokemiji su se uvodile postupno, paralelno s razvojem same metodologije geokemijske prospекcije, analitičkih metoda kao i usavršavanjem statističkih modela i tehnika. U tome su nesumnjivo vrlo važnu ulogu odigrali računari, posebno računari velike brzine, što je ubrzalo izračunavanje statističkih parametara i omogućilo primjenu viševarijantnih metoda, koje su, zbog obrade velikog broja podataka bez računara praktički neizvedive (faktorska analiza, diskriminacijska analiza i dr.).

U SR Hrvatskoj su krajem sedamdesetih godina započela intenzivna istraživanja nuklearnih sirovina u kojima je jedna od primjenjenih metoda bila i geokemijska prospекcija. Regionalna geokemijska i radiometrijska prospекcija na području Papuka i Krndije izvođena je od 1979. do 1981. godine. U ovom radu obrađeni su rezultati regionalne geokemijske prospекcije dobiveni u jednogodišnjem istraživanju, u 1980. godini, na području Papuka izgrađenog od naslaga »Psunjске tektonske jedinice« i naslaga »Radlovačke tektonske jedinice«. Geokemijskom prospекcijom te godine prikupljeno je 247 uzoraka dolinskog detritusa koji su analizirani na U, Cu, Zn, Pb, Ni, Co, Fe i Mn. Prikupljeno je i isto toliko uzoraka površinskih voda koji su analizirani samo na uran.

Analize uzoraka izvedene su u Institutu za geološko-rudarska istraživanja i ispitivanja nuklearnih i drugih mineralnih sirovina u Beogradu pod rukovodstvom mr. Z. Červenjaka (U u detritusu i U u vodi primjenom laserske fluorescencije), te na Rudarsko-geološko-naftnom fakultetu u Zagrebu gdje je analize Pb, Zn, Cu, Fe, Mn, Co i Ni izveo dr. L. Palinkaš primjenom atomske apsorpcione spektrometrije.

Rezultati analiza obrađeni su statistički, metodama deskriptivne statistike (Braun, 1980; Braun i dr.

1981), dok su metode inferencijalne statistike primjenjene samo dijelom (Dravec, 1984).

Kako se složenije statističke metode inferencijalne statistike u toj fazi istraživanja nisu koristile, ostala je mogućnost njihove naknadne primjene. To je ujedno bio i poticaj za izradu ovog rada.

Općenito o faktorskoj analizi

Faktorska analiza se u geologiji počela primjenjivati sredinom ovog stoljeća.

Osnovna svrha faktorske analize je na što jednostavniji način s manjim brojem faktora objasniti promatrani strukturalni model varijabli, kod čega svaki faktor predstavlja novo svojstvo promatranog skupa u kojem se signifikantnim smatraju faktorska opterećenja jednakaka ili veća od 0,5.

Varijable mogu biti sve vrste opažanja ili mjerena na najrazličitijim geološkim objektima. Faktorska analiza karakterizira čitavu geološku populaciju i prikazuje odnos između varijabli (»R – mod«) ili odnos između uzoraka (»Q – mod«).

Odnos oba moda faktorske analize mogu se prikazati slijedećom shemom:

	varijable					
	x1	x2	x3	x4	xi
u	1					
z	2					
o	3					
r	.					
c	.					
i	n					

Korelacija daje Q
mod

Korelacija daje R – mod

Varijable su prikazane u matričnom obliku. Polazna matrica iz koje se ekstrahiraju faktori je korelaciona matrica ili matrica varijanca-kovarijanca u čiju glavnu dijagonalu se uvrštavaju različite vrijed-

nosti – dijelovi varijance. U radu sa standardiziranim vrijednostima varijabli korelaciona matica se podudara s matricom varijanca-kovarijanca. Ekstrakcija faktora može se postići s nekoliko metoda kao što su centroidna metoda, metoda glavnih osi (komponentnih), Jacobijeva metoda utvrđivanja svojstvenih vrijednosti, metoda tridiagonalizacije matrice, Housholderov postupak dijagonalizacije i Francisov Q-R postupak dijagonalizacije (Fulgozi, 1979.).

Postoji također i nekoliko modela faktorske analize kao što su faktorska analiza glavnih komponenata, faktorska analiza zajedničkih faktora, Joreskogova analiza image-a, faktorska analiza najveće vjerojatnosti, Spearmanova faktorska analiza, kanonička faktorska analiza i alfa faktorska analiza (Fulgozi, 1979.).

U ovom radu su rezultati geokemijske prospexije obradeni faktorskog analizom glavnih komponenata na osnovi korelacione matrice, faktorskog analizom zajedničkih faktora i alfa faktorskog analizom za sve podatke. Posebno su također faktorskog analizom glavnih komponenata zasebno obradeni podaci koji se odnose na »Radlovačku tektonsku jedinicu« i »Psunjsku tektonsku jedinicu«. Odnosi između varijabli analizirani su pomoću »R-modra«.

Rezultati statističke obrade podataka geokemijske prospexije na području Papuka i Krndije primjenom faktorske analize

Faktorskog analizom glavnih komponenata na bazi korelacione matrice, pomoću programa FACTOR 3 (Madison Academic Computing Center), izdvojena su četiri faktora, za svojstvene vrijednosti koje su veće od 1 (Kaiserov kriterij), koji objašnjavaju 73,1% totalne varijance. Kod toga prvi faktor objašnjava 37%, a ostala tri faktora 35% totalne varijance. Shvaćajući faktore kao novo svojstvo dolinskog detritusa može se reći da se njime manifestira povezanost Fe, Ni, Co, Zn i Cu u dolinskom detritusu (prvi faktor), negativna korelacija U i Mn u dolinskom detritusu (drugi faktor), samostalna distribucija urana u dolinskom detritusu (treći faktor) te negativna korelacija između U u detritusu i Pb s jedne strane i U u vodi s druge strane.

Programom Madison Academic Computing Centra faktorskog analizom zajedničkih faktora ekstrahirano je 5 faktora. Na prvi faktor kod toga otpada 83% faktorske varijance i 32% totalne varijance, a signifikantna faktorska opterećenja nalaze se samo kod prvog faktora. Iz toga slijedi da u ovoj analizi za promatrani skup podataka korisne informacije daje samo prvi faktor dok bi se svi ostali trebali zanemariti. Prvim faktorom objašnjavaju se međusobne veze između Co, Ni, Fe, Zn i Cu u dolinskom detritusu.

Alfa faktorskog analizom također je dobivena faktorska matica koja se sastoji od četiri faktora kod čega prvi faktor doprinosi 74,3% faktorske varijance, odnosno objašnjava 33% totalne varijance, dok preostala tri faktora objašnjavaju ukupno svega 8% totalne varijance i nemaju značajnih faktorskih opterećenja niti na jednoj varijabli. U takvom slučaju se preostala tri faktora mogu zanemariti kao i u

prethodnoj analizi. Najveća faktorska opterećenja prvim faktorom i u ovoj analizi imaju Fe, Ni, Co, Zn i Cu.

Iz grafičkog prikaza konfiguracije u faktorskog prostoru koji je definiran osima prvog i drugog faktora još su jasniji odnosi elemenata i faktora (slika 1).

Pomoću programa Microgass koji je razvijen na Queens University u Kingstonu, Ontario, za potrebe Međunarodne agencije za atomsku energiju, postepenom (»stepwise«) faktorskog analizom glavnih komponenata također je dobiven sličan rezultat. Ekstrahirana su četiri faktora od kojih prvi faktor objašnjava 28,7% totalne varijance, što je nešto manje nego što je dobiveno prethodno opisanim analizama, drugi faktor 11,5%, treći 14,7% i četvrti 2,3%. Ovaj program osim samih faktora izračunava i koeficijente faktorskih bodova te faktorske bodove (»factor scores«) za sve izlučene faktore i svaki uzorak. Faktorski bodovi su statistički analizirani i podijeljeni u klase s vrijednostima x-2s do x-1,5s, x-1,5s do x-s, x-s do x, x do x+s, x+s do x+1,5s, x+1,5s do x+2s i veće od x+2s (gdje je x srednja vrijednost, a s standardna devijacija). Pripadnost uzorka pojedinoj klasi označena je na karti različitom veličinom oznaka. Prvi faktor definira asocijaciju Fe, Ni, Co, Cu, drugi ukazuje na izdvojeno ponašanje Mn kao i treći za Zn dok u četvrtom faktoru nema signifikantnih faktorskih opterećenja.

Iz pregleda rezultata izvedenih analiza evidentno je da su za sva tri, odnosno četiri, modela faktorske analize oni bili sukladni pa je zato u daljim obradama podataka geokemijske prospexije provedene na području Papuka i Krndije korištena samo faktorska analiza glavnih komponenata.

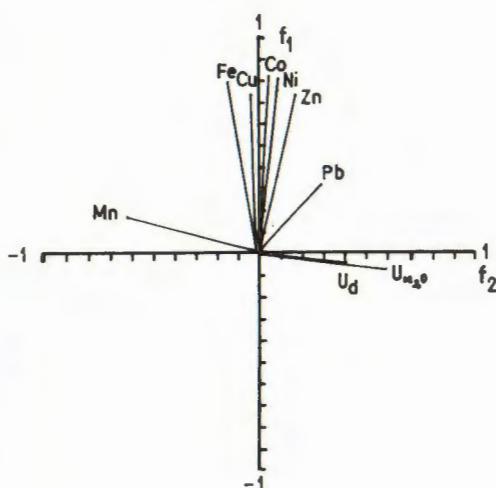
Tako su tom metodom zasebno analizirani podaci dobiveni u »Psunjskoj tektonskoj jedinici« (grupa »N«) kao i podaci dobiveni u »Radlovačkoj tektonskoj jedinici« (grupa »R«). Faktorskog analizom podataka grupe »N« ekstrahirana su četiri faktora koji ukupno objašnjavaju 60,5% totalne varijance od čega prvi faktor objašnjava 35,6%, a drugi faktor 13,6% totalne varijance. Faktorska opterećenja prvim faktorom su najveća suksesivno na slijedećim elementima: Ni, Co, Fe, Zn i Cu. Drugim faktorom se objašnjava negativna korelacija Mn s jedne strane i U u detritusu i vodi s druge strane. Kod trećeg faktora imamo potpuno suprotnu sliku odnosno pozitivnu korelaciju Mn i U u dolinskom detritusu.

Grafički prikaz faktorskih opterećenja varijabli prvim i drugim faktorom (slika 2a) sličan je istovjetnom prikazu koji se odnosi na sve podatke (slika 1a).

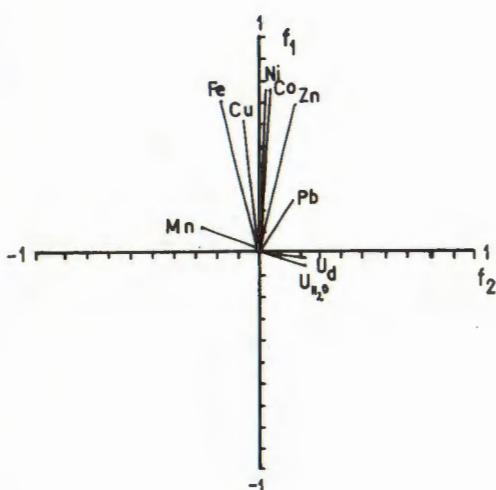
Faktorskog analizom glavnih komponenata podataka koji se odnose na »Radlovačku tektonsku jedinicu« ekstrahirana su tri faktora. Prvi faktor objašnjava 44,6% totalne varijance, drugi 12,9% i treći 11,7%. Faktorska opterećenja prvim faktorom najveća su na Pb i treći na U u vodi i U u dolinskom detritusu s pozitivnom korelacijom. Odnosi između varijabli još su jasniji na grafičkom prikazu konfiguracije u faktorskog prostoru prva dva faktora (slika 2b). Za razliku od prethodne analize, u ovoj, Mn se nalazi u asocijaciji elemenata zajedno s Fe, Co,

PSUNJSKA I RADLOVAČKA TEKTONSKA JEDINICA - Faktorska opterećenja varijabli prvim i drugim faktorom

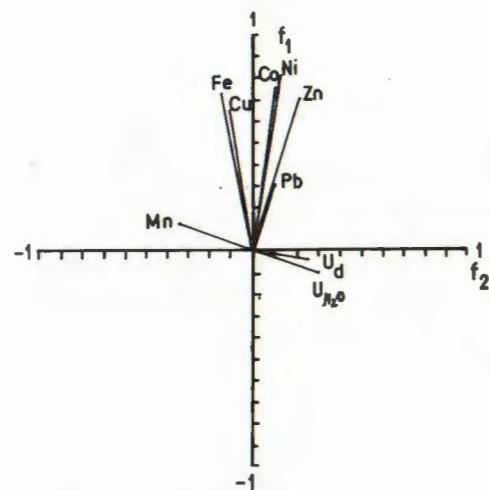
PSUNJ AND RADLOVAC TECTONIC UNITS - Plot of loadings on two first factors



a) Analiza glavnih komponenata
Principal components analysis



b) Faktorska analiza zajedničkih faktora
Principal factor analysis

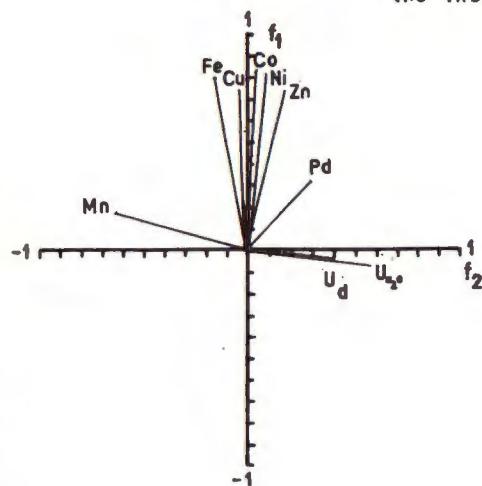


c) Alfa faktorska analiza
Alpha factor analysis

FAKTORSKA ANALIZA GLAVNIH KOMPONENTA

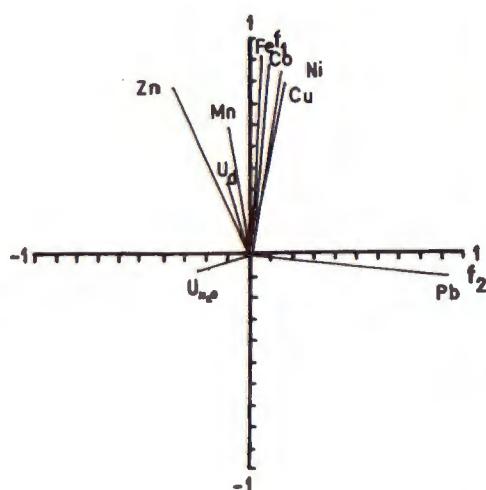
Faktorska opterećenja varijabli prvim i drugim faktorom

PRINCIPAL COMPONENTS ANALYSIS -Plot of loadings on two first factors



a) Grupa "N"

Group "N"



b) Grupa "R"

Group "R"

Slika 2
Fig. 2

Ni, Zn i Cu što bitno razlikuje »Radlovačku tektonsku jedinicu« od »Psunjiske tektonske jedinice«.

Utvrdena povezanost elemenata u dolinskom detritusu u geokemijskom smislu znači da se elementi svakog pojedinog faktora u procesu koncentracije ili dekoncentracije pokoravaju zakonu koji je specifičan za svaki faktor.

Primjenom faktorske analize po separiranim grupama (kod čega je kriterij za separaciju grupa bio isključivo geološki) pokazuju se razlike u svojstvima dolinskog detritusa vodotoka koji dreniraju »Radlovačku tektonsku jedinicu« i vodotoka koji dreniraju »Psunjisku tektonsku jedinicu«. Podudarnost rezultata analiza za sve podatke i rezultata analiza podataka iz »Psunjiske tektonske jedinice« može biti posljedica većeg broja uzoraka koji su uzeti iz te serije (193 uzorka iz vodotoka koji dreniraju »Psunjisku tektonsku jedinicu« spram 54 uzorka iz vodotoka koji dreniraju »Radlovačku tektonsku jedinicu«), ali može biti i posljedica bitno većeg intenziteta geokemijskih procesa koji su doveli do koncentracije elemenata u dolinskom detritusu »Psunjiske tektonske jedinice«. Na posljednje ukazuje raspored klasa faktorskih bodova za prvi faktor koji definira asocijaciju Fe, Ni, Co i Cu (slika 3), prema kojem se jasno može razdvojiti područje izgrađeno od naslaga »Radlovačke tektonske jedinice« od naslaga »Psunjiske tektonske jedinice«. Promatrajući asocijaciju elemenata čija je povezanost objašnjena prvim faktorom u provedenim analizama, može se pretpostaviti da navedeno objašnjenje u stvari predstavlja proces koncentracije elemenata adsorpcijom na Fe – hidroksidima koji je prema tome karakterističniji za »Psunjisku tektonsku jedinicu«. Na ovu razliku utiče i uočeno različito ponašanje distribucije Mn u svakoj od istraživanih tektonskih jedinica, koje ukazuje i na takve uvjete sredine formiranja ove asocijacije elemenata u kojima dolazi do razdvajanja distribucije Mn i Fe, a odnose se prvenstveno na pH<8 na području »Psunjiske tektonske jedinice«, odnosno na uvjete u kojima dolazi do njihovog zajedničkog obaranja – pH>8 na području »Radlovačke tektonske jedinice« (slika 4).

Negativna korelacija Mn i U u detritusu potkrepljuje ovu pretpostavku, ali isto tako objašnjava da se radi o oksidativnoj sredini koja omogućava koncentraciju Mn i istovremeno sprečava koncentraciju U u detritusu. Obaranje U u detritusu najvjerojatnije, budući da ima posve samostalnu distribuciju, može se vezati za prisutstvo organskih kompleksa (humati, fulvati) (slika 5). Određena korelativnost Mn i U u detritusu koja se javlja u trećem faktoru kod analize uzorka iz »Psunjiske tektonske jedinice« (grupa »N«), odražava litološke karakteristike područja, odnosno, pokazuje vezu urana u detritusu s uranonskim »grafitičnim« naslagama u kojima je kemijskim analizama stijena utvrđen povišen sadržaj mangana koji se isto tako odražava i na njegovom sadržaju u dolinskom detritusu. Klasirani faktorski bodovi izračunati postupnom faktorskrom analizom za drugi faktor koji reprezentira distribuciju Mn, također indiciraju rasprostranjenost »grafitičnih« naslaga (slika 6).

Slabo izražena zavisnost U u vodi i U u detritusu koja je objašnjena trećim faktorom kod analize podataka iz »Radlovačke tektonske jedinice« vjerojatno ima lokalni značaj, vezano za pH i Eh sredine obaranja u detritusu. Analizom njihovih pojedinačnih distribucija međuzavisnost nije utvrđena (Braun 1980). Ovdje također treba uzeti u obzir i nizak procenat totalne varijance koji taj faktor objašnjava.

Samostalnu distribuciju dobro koreliranu s drugim faktorom u analizi podataka iz »Radlovačke tektonske jedinice« ima i Pb, koje se vjerojatno izdvaja u formi anglezita.

Prema rezultatima provedenih analiza distribucija Zn imala bi dvojni karakter, dio koji je neovisan o ostalim elementima i dio koji se koprecipitira ili adsorbira Fe-hidroksidima (slika 7). Međutim, ili su procesi kojima se u području gdje su analizama pojedinačnih distribucija utvrđene anomalne koncentracije Zn (Braun 1980) u odnosu na procese obaranja Fe-hidroksidima slabijeg intenziteta ili se radi o matematički ipak nesignifikantnom faktoru koji definira distribuciju cinka kao samostalnu.

Zaključak

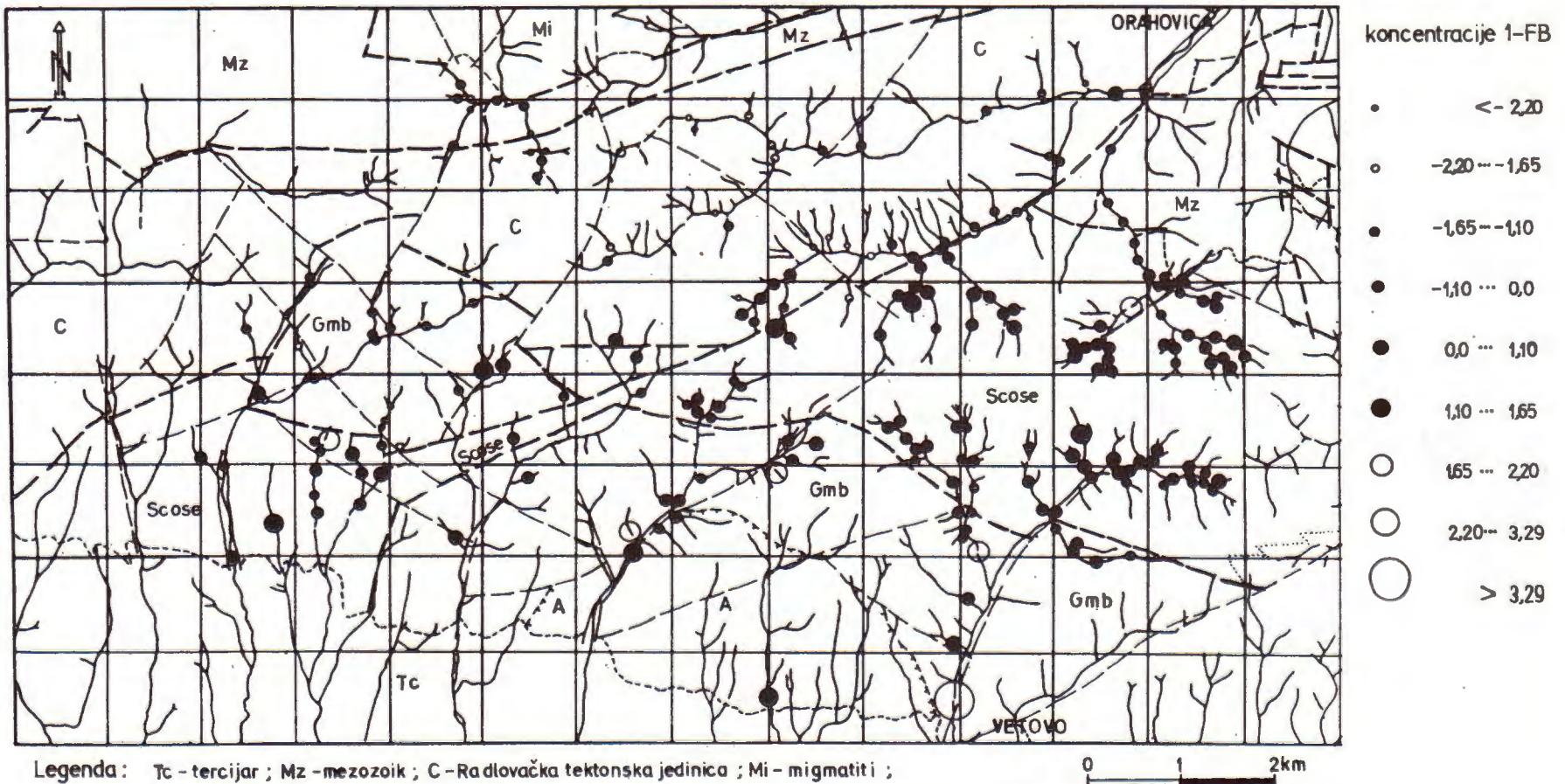
Uspoređivanjem rezultata faktorske analize programom FACTOR 3 Madison Computer Centra može se reći da su dobiveni rezultati sa sva tri modela bili sukladni. Treba međutim napomenuti da bi se možda neke veze između elemenata koje u primjenjenim modelima i tehnikama faktorske analize nisu objašnjene, daljom rotacijom faktorskih osi mogle pojaviti.

Faktorska analiza ukazala je na genezu pojedinih asocijacija elemenata, jedna vezana za koprecipitaciju s Fe-hidroksidima, jedna koja je najvjerojatnije vezana za organsku supstancu i humate te na elemente sa samostalnim distribucijama za čije položenje u detritusu postoje specifični razlozi ili uvjeti. Faktorska analiza daje i podatke na osnovu kojih se mogu razdvojiti područja različitih geoloških karakteristika koje uvjetuju i različito ponašanje pojedinih elemenata. To se na istraživanom području odnosi prvenstveno na proces obaranja elemenata Fe-hidroksidima koji su mnogo intenzivniji i karakterističniji za područje izgrađeno od naslaga »Psunjiske tektonske jedinice«, te specifičnost distribucije Mn.

Provedene različite faktorske analize dale su identične rezultate u odnosu na prvi faktor koji objašnjava 30–35% totalne varijance.

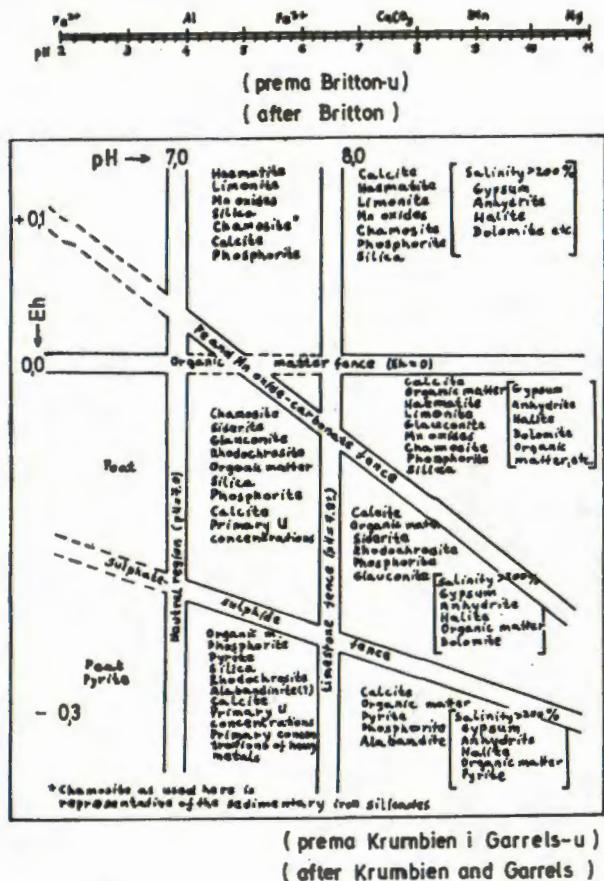
Iako je suština faktorske analize da što veći broj veza varijabli objasni jednim faktorom, pa se važećom smrta ona analiza u kojoj se prvim faktorom objašnjava 60–65% totalne varijance, na ovom primjeru je vidljivo da i analize u kojima je taj procenat za prvi faktor manji, također daju upotrebljive rezultate. U ovom slučaju se može postaviti pitanje da li metoda koja čisto matematički gledano nije dala validan rezultat može biti korištena u interpretaciji geokemijskih podataka, odnosno obratno od toga, što činiti s analizom koja je matematički dala validan rezultat, koji međutim nije primjenljiv u geokemijskoj interpretaciji, jer ako od 10 elemenata koje analiziramo, ne očekujemo da se svi zajedno grupi-

KARTA FAKTORSKIH BODOVA ZA 1. FAKTOR
FACTOR SCORES MAP FOR 1. FACTOR



Legende: Tc - Tertiary ; Mz - Mezozoic ; C - "Radlovac" - Complex ; Mi - Mygmatites ;
Scose - Phyllites ; Gmb - Gneiss ; A - Amphibolites of the "Psunj" tectonic unit

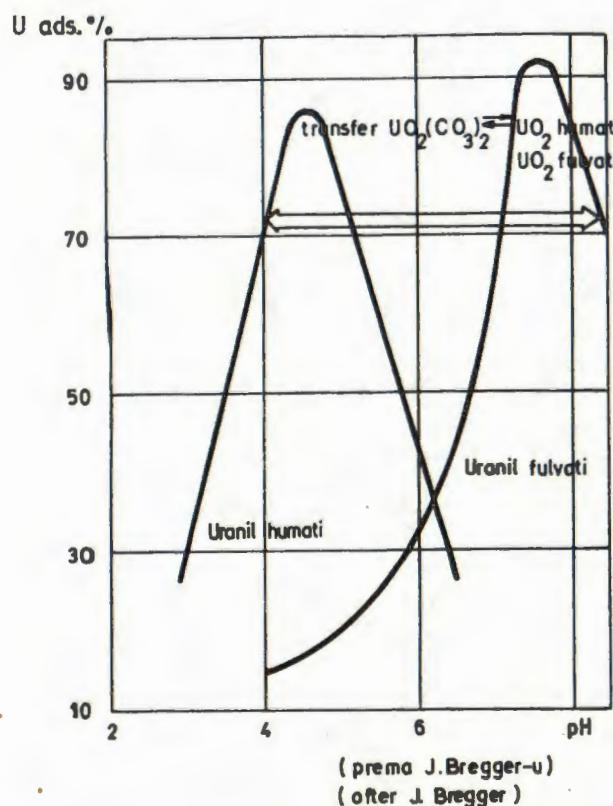
Sl. - Fig. 3.



Slika 4 - Precipitacij o hidrata Fe_2Mn , Mg, Al i CaCO_3 u ovisnosti od promjene pH otopine te konačne asocijacije kemijskih sedimenata i njihovi karakteristični uvjeti sredine formiranja obzirom na pH i Eh

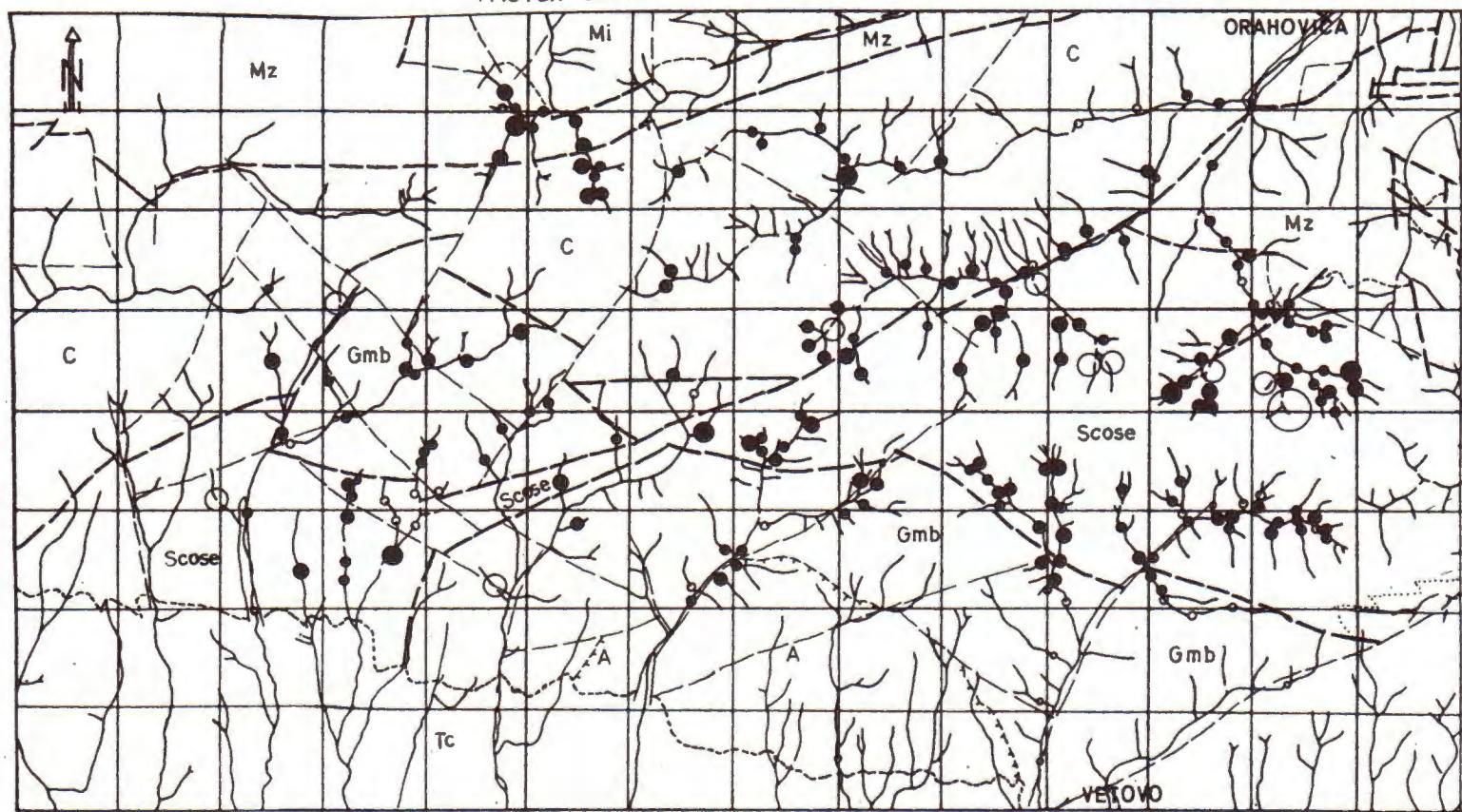
Fig. 4 - Sequence of the precipitation of the hydrates of Fe, Mn , Mg, Al and CaCO_3 with change in the pH of the solution and final associations in chemical sediments and ambient conditions characteristic of them in terms of Eh and pH

Slika 5 -
Fig. 5 -



KARTA FAKTORSKIH BODOVA ZA 2. FAKTOR

FACTOR SCORES MAP FOR 2. FACTOR



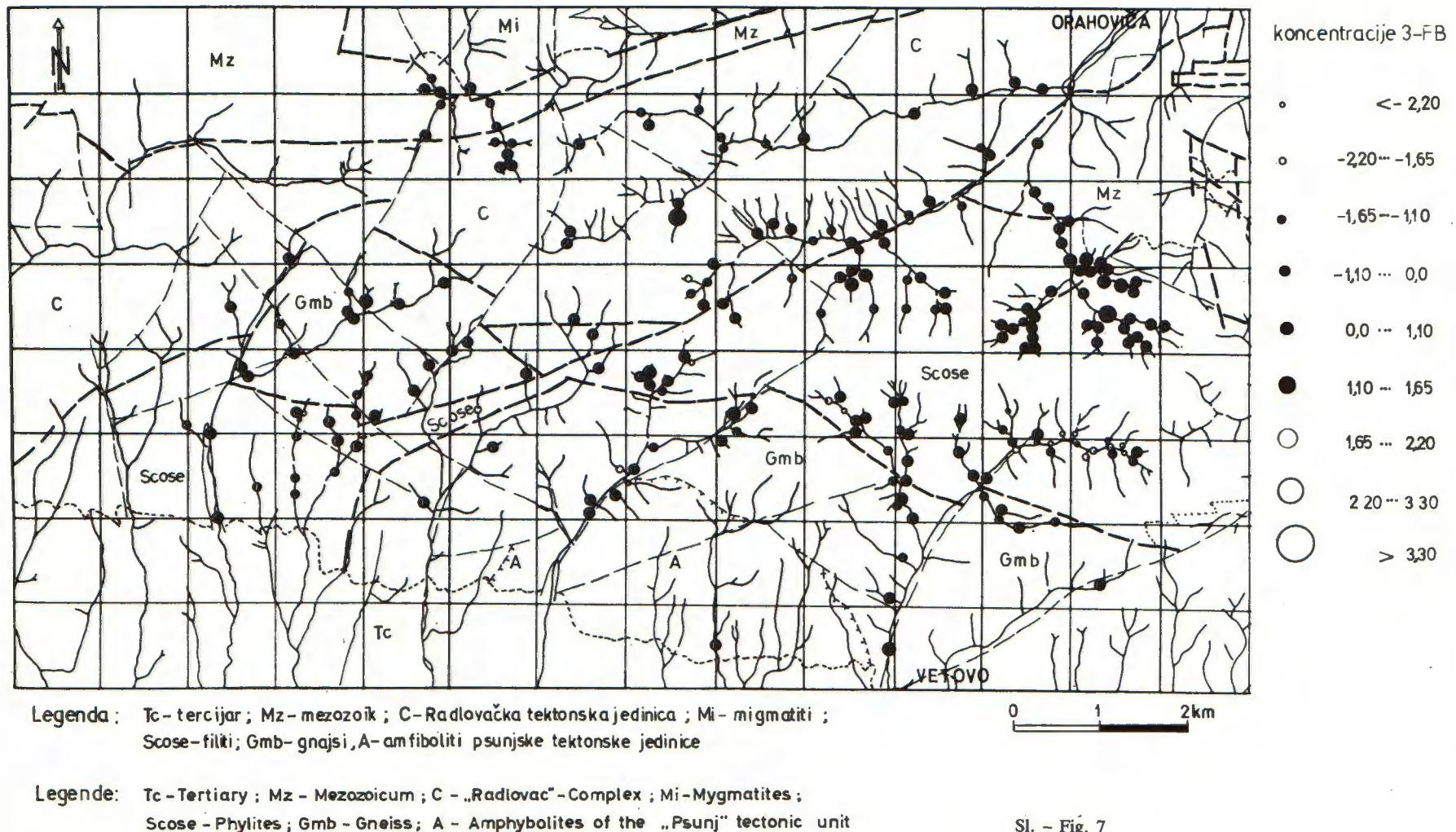
Legenda: Tc - tercijar ; Mz - mezozoik ; C - Radlovačka tektonska jedinica ; Mi - migmatiti ;
Scose-filiti, Gmb-gnajsi, A-amfiboliti psunjiske tektonske jedinice

0 1 2km

Legende: Tc - Tertiary ; Mz - Mezozoic ; C - "Radlovac" - Complex ; Mi - Mygmatites ;
Scose - Phylites ; Gmb - Gneiss ; A - Amphibolites of the "Psunj" tectonic unit

Sl. - Fig. 6

KARTA FAKTORSKIH BODOVA ZA 3. FAKTOR
FACTOR SCORES MAP FOR 3. FACTOR



raju, tada se faktorska analiza treba koristiti na fleksibilniji način. Konačno, faktorska analiza je samo jedna od pomoćnih metoda u interpretaciji rezultata geokemijske prospekcije i geokemijskih procesa.

Može se reći da faktorska analiza u interpretaciji rezultata istraživanja ležišta mineralnih sirovina jednostavno i brzo daje mnogo korisnih informacija geologu-geokemičaru posebno ukoliko se koriste programi prilagođeni za izlaz u vidu iscrtavanja na geološkim kartama.

Zahvala

Zahvalnost dugujem I. Opiću, D. Frleti i M. Pečniku koji su mi pomogli kod kompjutorske obrade baze podataka bez koje se ovaj rad ne bi mogao ostvariti.

Primljeno: 8. I. 1990.

Prihvaćeno: 7. V. 1990.

LITERATURA

- Braun, K. (1980): Istraživanje nuklearnih sirovina u SRH, II. faza. Radiometrijska i geokemijska istraživanja na području Papuka. Fond dokumenata Instituta za geološka istraživanja, Zagreb.
- Braun, K., Dravec, J., Valković, V. (1981): Istraživanje nuklearnih sirovina u SR Hrvatskoj, II. faza. Poludetaljna i detaljna istraživanja – Papuk 1981. Fond dokumenata Instituta za geološka istraživanja, Zagreb.
- Dravec, J. (1984): Usporedba rezultata regionalne radiometrijske i geokemijske prospekcije urana na području Papuka i Krndije. Referati sa Savjetovanja: Rezultati dosadašnjih i pravci daljih istraživanja nuklearnih sirovina, 348–358, Beograd.
- Fulgozi, A. (1979): Faktorska analiza. Školska knjiga, Zagreb.

Use of factor analysis in data processing of geochemical exploration in the Mt. Papuk and Mt. Krndija region

J. Dravec-Braun

In Croatia, during the late seventies started an intensive uranium investigation, with reconnaissance geochemical survey as one of the applied methods.

This piece of work is concerned with the annual results of reconnaissance geochemical exploration (1980) in the region of Mt. Papuk and Mt. Krndija composed of the »Psunj tectonic unit« and »Radlovac tectonic unit« deposits.

In that year 247 samples of stream sediment were collected, and analysed on Cu, Pb, Zn, Ni, Co, Mn, Fe and U. The same number of surface water were collected and analysed on uranium only.

Using the laser fluorescence method uranium is analysed in stream sediments and water, while the other elements by use of atomic absorption spectrometry.

The results of the analyses were worked out statistically, using the methods of descriptive statistics, and simpler methods of inferential statistics. There is still a possibility of application of more sophisticated methods, which was the stimulus for writing this piece of paper.

Results of the reconnaissance geochemical exploration are processed by principal component analysis on the basis of the correlation matrix, principal factor analysis and alpha factor analysis according to the Madison Computer Center's program FACTOR 3, and by stepwise factor analysis (component analysis)

according to the program MICROGASS. The data were analysed by techniques of R-mod factor analysis.

Comparing the results of three models of factor analysis of Madison Computer center, it could be concluded that they are correlative, though some relations between the elements which in the applied models of factor analysis are not explained, in the further rotation of factor axes could appear.

Factor analysis pointed out the genesis of the certain association of elements, one linked to coprecipitation with Fe-hydroxide, the other is most probably related to the organic matter and humates, as well as elements with independent distribution whose precipitation is due to specific reasons and conditions. Factor analysis gives us data according to which various geological characteristics that cause different behaviours of various elements, could be distinguished. In the area of investigation, this are particularly the processes of coprecipitation and adsorption by Fe-hydroxide which are more intensive and characteristic for the area of »Psunj tectonic unit« deposit.

According to the first factor which explains 30–35% of total variance, different models of factor analysis showed the identical results. Even though, list of authors suggest minimum 60–65% explanation of the total variance by first factor, as the necessary level of significance for the whole factor analysis, it is obvious from this article, that even less percentage of total variance, explained by first factor, could be useful in interpretation of geochemical exploration data.