

## КОРЕЛАЦИОНА АНАЛИЗА КАО МЕТОД ИЗДВАЈАЊА ГЕОХЕМИЈСКИХ АСОЦИЈАЦИЈА ЕЛЕМЕНАТА У РУДНОМ ПОЉУ КАЛУДРА (ЈЗ СРБИЈА)

од

Ненада Лончаревића\* и Бошка Стајевића\*

При хемијској проспекцији потенцијално рудоносних средина, дефинисање асоцијације елемената-индикатора је један од почетних задатака. Овај проблем се нарочито истиче при изучавању комплексних типова орудњења у чијој грађи учествује већи број рудно-минералних компоненти, као што је полиметалични тип орудњења.

Компаративна анализа добијених резултата, уз остала релевантна геолошко-геохемијска обележја омогућава поузданије дефинисање геохемијске асоцијације индикаторних елемената.

У раду се приказују три статистичка метода анализе корелације међу потенцијалним елементима-индикаторима при геохемијским истраживањима и даје компаративна анализа на тај начин добијених резултата. Као пример, обрађени су резултати геохемијских истраживања полиметаличног рудног поља Калудра на југоисточним обронцима Голије.

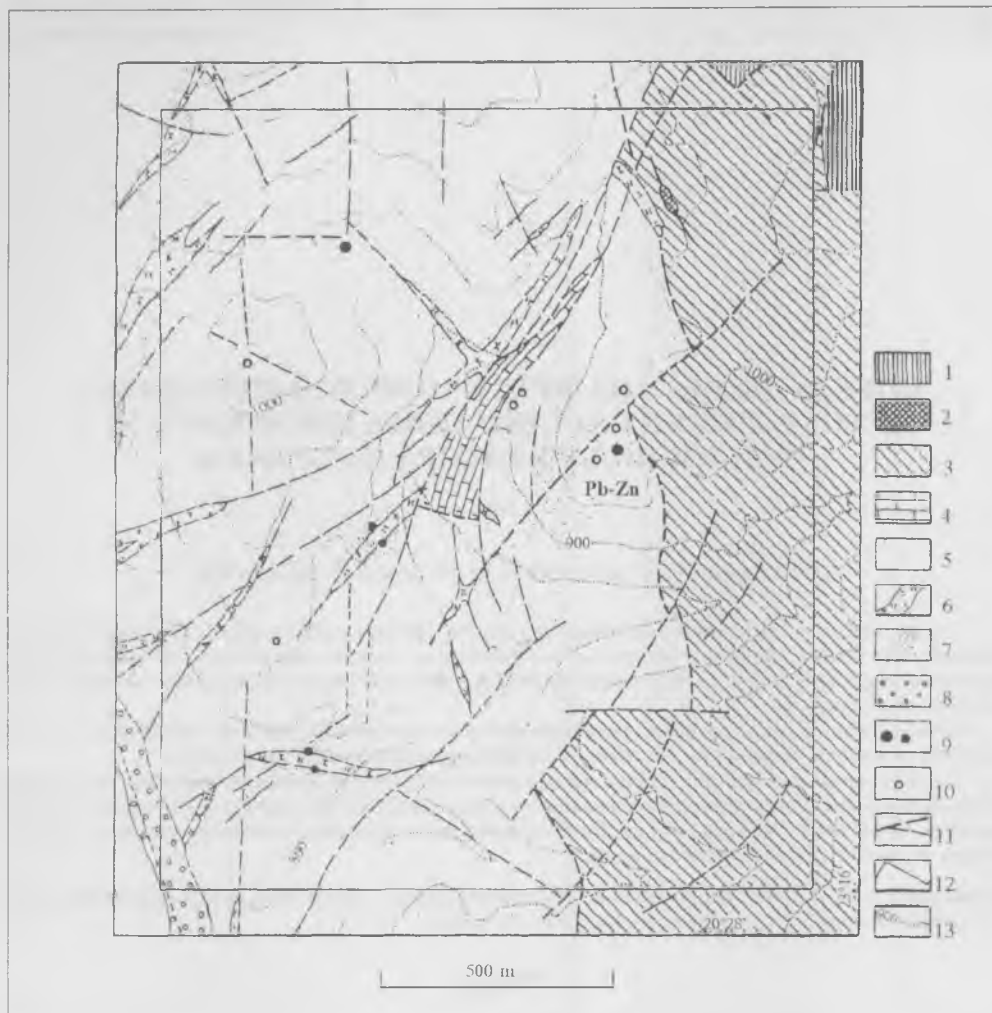
**Кључне речи:** литогеохемијска проспекција, статистичка анализа, корелација, факторска анализа, полиметалично орудњење, рудно поље Калудра, ЈЗ Србија.

### УВОД

Рудно поље Калудра представља таксономски нижу целину у оквиру рудног региона Голије и лоцирано је на њеним југоисточним обронцима. У његовој геолошкој грађи највећим делом учествују сенонски пешчарско-алевролитични, глиновито-лаиоровити седименти и јурско-кредни меланжирани кластити у које су утиснути дајковци кварцлатитско-феноандезитског састава терцијарне старости (сл. 1). Утискивање магматских стена је праћено хидротермалним изменама околних стена и полиметаличном импрегнационо-жичном минерализацијом Pb-Zn-Sb-Ag.

У склопу проспекцијских геохемијско-металометријских испитивања узете су литогеохемијске пробе секундарних ореола расејавања по правоугаоној мрежи 100×40 m при чему су оријентација и густина мреже усклађене са геолошком грађом терена и претпостављеним димензијама могућих геохемијских аномалија. Укупно је узето 648 проба, које су анализирани на 12 рудних и пратећих елемената (Pb, Zn, Ag, Sb, Ba, Sr, Cu, Mo, Cd, W, Bi). Геохемијско поље Калудре детаљније је приказао Стајевић (1996).

\* Рударско-геолошки факултет Универзитета у Београду, Бушина 7. 11 000 Београд.



Сл. 1. Карта геолошког склопа Калудре (Стајевић, 1996). Легенда: 1. метаморфити "Студеничке серије" (калкшисти, зелени шкриљци, метадијабази и филитомикашисти), 2. серпентинити Вардарске зоне, 3. кластити дијабаз-ројначке формације (глинци, лапорци и дијабази) / $K_2^3$ - $J_{2,3}$ /, 4. кречњаци усложени и банковити / $J_{2,3}$ /, 5. флишни седименти (аренити, алевролити, лапорци у пешчари) / $K_2^3$ /, 6. дајкови терцијарних вулканита (кварцлатита, латита, фенозндезита и андезита), 7. хидротермално промењене стене (пиритисане, каолинисане, силификоване, пропицитисане и др.), 8. неvezани кластити алувиона, 9. рудне појаве и минерализације Zn, Pb, Sb, 10. стари рударски радови (пинге, свртњеви, халде), 11. руптуре, 12. спољне контуре површине обухваћене литохемијском проспекцијом и 13. изохинсе површине рељефа.

Fig. 1. Geologic map of Kaludra (Stajević, 1996). Legend: 1. "Studenica Series" metamorphic rock (calc-schist, greenschist, metadiabase and phyllite micaschist); 2. Vardar zone serpentinites; 3. Diabase-chert formation (shale, marlstone, diabase) / $K_2^3$ - $J_{2,3}$ /; 4. Bedded and thick limestone / $J_{2,3}$ /; 5. Flysch deposits (arenite, siltstone, marlstone, sandstone) / $K_2^3$ /; 6. Tertiary volcanite dykes (quartz latite, latite, phenoandesite, andesite); 7. Hydrothermally altered rocks (pyritized, kaolinized, silicified, propylitized, etc.); 8. Alluvial incoherent clastics; 9. Ore occurrences and Zn, Pb, Sb mineralizations; 10. Old mine workings (caved-in shafts, dumps); 11. Fractures; 12. Outer contours of lithochemical prospect; 13. Topographic.

У раду се, применом три различита метода корелационе анализе, приказују њихове могућности при издвајању геохемијских асоцијација елемената путем квантификације њихових узајамних односа и даје упоредна анализа тако добијених резултата.

## ПРИМЕЊЕНИ МЕТОДИ

За презентацију могућности корелационе анализе у издвајању геохемијских асоцијација елемената изабрана су три метода:

1. Пирсонов коефицијент корелације ( $r$ ) из домена параметарских метода статистике;

2. Спирманов коефицијент корелације ранга ( $R$ ) из домена непараметарских метода статистике и

3. Метод анализе главних компоненти из домена факторске анализе.

**Пирсонов коефицијент корелације ( $r$ )** је у геолошкој јавности најшире познат и примењиван начин квантификације и испитивања односа између две величине. Он спада у тзв. параметарске методе статистике, што значи да се прорачунава на основу претходно израчунатих (процењених) параметара дате статистичке популације, у овом случају коваријансе два хемијска елемента и њихових стандардних девијација:

$$r = \frac{COV(x, y)}{\sigma_x \sigma_y}$$

Ова чињеница чини Пирсонов коефицијент корелације осетљивим на тип дистрибуције датих елемената, што може да доведе до нереалних резултата. Као један од потенцијалних недостатака овог метода треба навести да он представља искључиво *линеарну* корелацију између две величине. Због тога се, при проучавању корелације, у почетној фази препоручује израда "*дијаграма расипања*" или "*scatter diagrams*" како се називају у енглеској литератури не би ли се уочило евентуално постојање *нелинеарне* корелације између две променљиве.

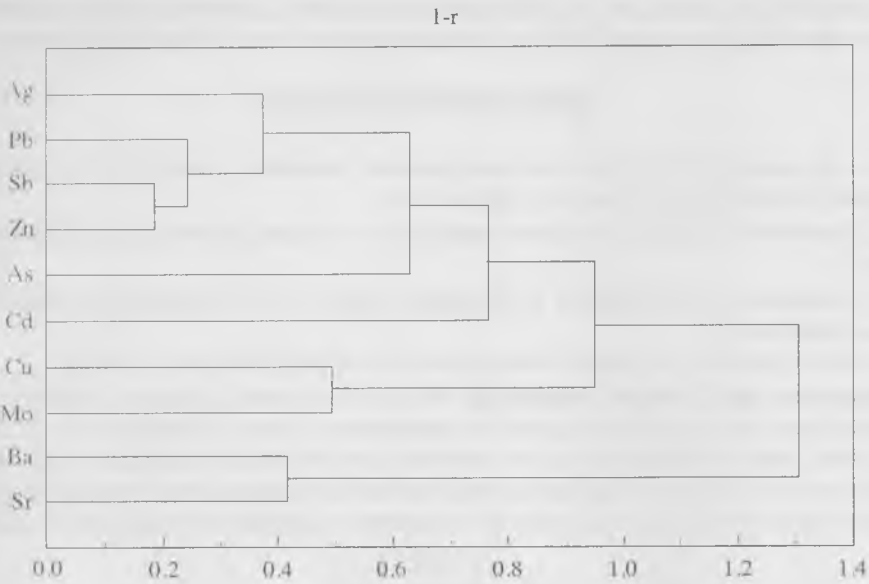
**Спирманов коефицијент корелације ранга ( $R$ )** спада у методе непараметарске статистике, што га у великој мери чини имуним на тип дистрибуције података. Оригинални подаци се трансформишу рангирањем, а затим се израчунава коефицијент корелације новодобијених података.

**Анализа главних компоненти** спада у домен факторске анализе. У математичком смислу она представља поступак налажења тзв. сопствених (карактеристичних) вредности и вектора матрице варијанси-коваријанси података, чиме се добија  $n$  међусобно ортогоналних вектора у и димензионалном простору који представљају варијабилност оригиналних података. У идеалном случају,  $n$  је знатно мање од броја оригиналних променљивих, тако да поједини вектори представљају по неколико међусобно зависних променљивих, те се преко њих може изразити корелација међу подацима.

## РЕЗУЛТАТИ И ДИСКУСИЈА

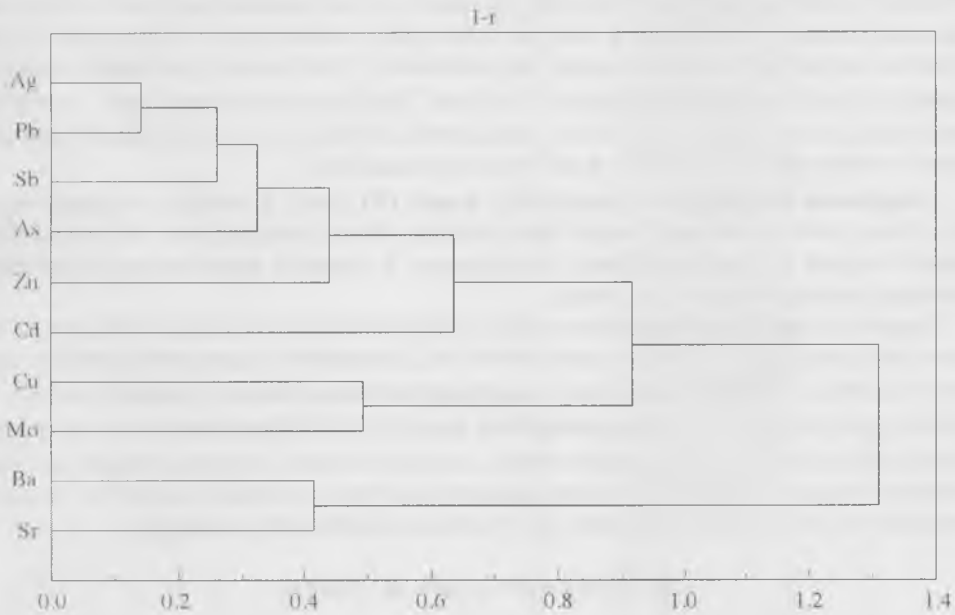
Од дваесет испитиваних хемијских елемената, за статистичку обраду је одабрано њих 10. Садржаји  $V_i$  и  $W$  су у превеликом проценту проба били испод границе осетљивости аналитичког поступка те се не могу сматрати репрезентативним.

Испитивањем момената расподеле анализираних хемијских елемената се долази до закључка да су дистрибуције  $Pb$ ,  $Zn$ ,  $Ag$ ,  $As$ ,  $Sb$  и  $Cd$  изразито лево асиметричне



Сл. 2. Дендрограм кластер анализе на основу Пирсоновог коефицијента корелације ветрансформисаних података.

Fig. 2. Tree diagram of cluster analysis based on Pearson's correlation coefficient of original data.



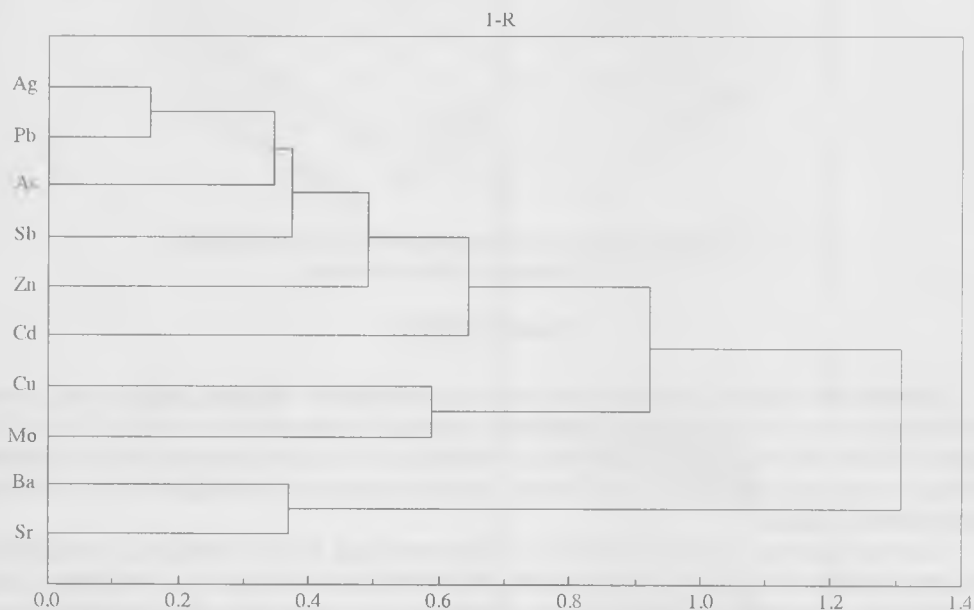
Сл. 3. Дендрограм кластер анализе на основу Пирсоновог коефицијента корелације трансформисаних података.

Fig. 3. Tree diagram of cluster analysis based on Pearson's correlation coefficient of transformed data.

(трећи моменат расподеле већи од 6), те је одлучено да се при прорачуну Пирсоновог коефицијента корелације и при анализи главних компоненти користе подаци трансформисани логаритмовањем, док су код осталих елемената коришћени оригинални подаци. Ово је учињено зато што су код лево асиметричних дистрибуција блиских логнормалним, геометријска средина и варијанса логаритмованих вредности ваљани параметри расподеле, за разлику од аритметичке средње вредности и варијансе. Као пример разлике између коришћења сирових и логаритмованих података код оваквих дистрибуција, дат је дендрограм кластер-анализе при чему је као мера удаљености узет Пирсонов коефицијент корелације тзв. сирових података (сл. 2) и логаритмованих података за елементе са дистрибуцијом блиском логнормалној (сл. 3).

Може се запазити да коефицијент корелације сирових података није успео да "ухвати" везу између As с једне и Pb, Zn, Ag и Sb с друге стране, за разлику од коефицијента корелације рачунатог са логаритмованим подацима за елементе са лево-асиметричном дистрибуцијом.

Дендрограм на основу Спирмановог коефицијента корелације ранга (сл. 4) има доста сличности са сл. 3, што указује на оправданост логаритамске трансформације, с обзиром на то да је Спирманов коефицијент веома робустан метод "имун" на тип дистрибуције сирових података.



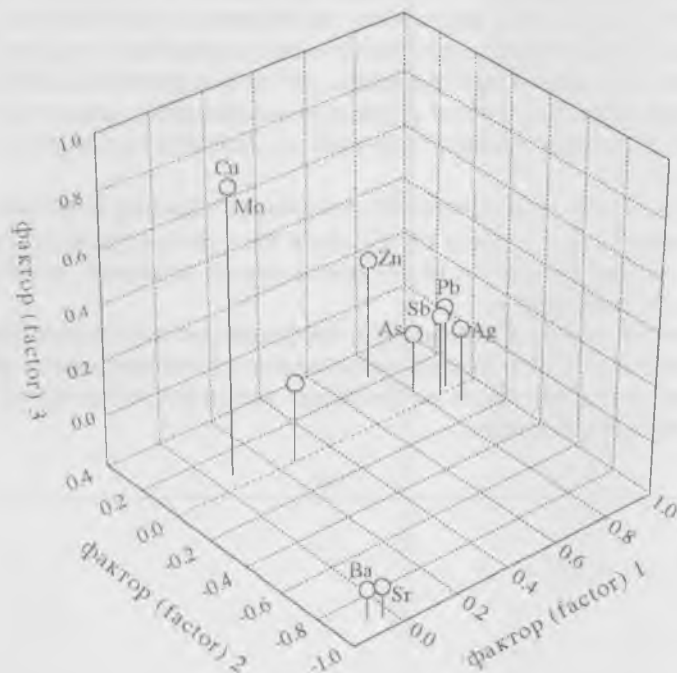
Сл. 4. Дендрограм кластер анализе на основу Спирмановог коефицијента корелације ранга.

Fig. 4. Tree diagram of cluster analysis based on Spearman's rank correlation coefficient.

Анализом главних компоненти издвојена су четири фактора: један обухвата Pb, Zn, Ag, Sb и As, други Ba и Sr, трећи Mo и Cu, и четврти само Cd (сл. 5).

Сва три метода су успешно детектовала геохемијску асоцијацију елемената-индикатора полиметаличног орудњења: Pb-Ag-Zn-Sb-As±Cd. Веза кадмијума са осталим елементима може на основу презентованих дијаграма изгледати дискутабилна, мада је чињеница да се аномалије овог елемента веома добро поклапају са

аномалијама осталих елемената, поготово  $Zn$ . Међутим, аномалије  $Cd$  имају знатно уже распрострањење док су у осталом делу терена оне у нивоу кларка те је вероватно због тога дошло до извесног "расплњавања" корелације између овог и осталих елемената па је због тога везу између њих теже детектовати.



Сл. 5. Тродимензионални дијаграм резултата факторске анализе.

Fig. 5. 3D diagram of factor loadings.

## ЗАКЉУЧАК

Параметри и компаративни резултати корелационе анализе дискретних аналитичких података геохемијске проспекције ореола полиметаличног рудног поља Калудра уз примену статистичких метода омогућују да се на релативно егзактан начин потврде и искомбацију одређене геохемијско-металогенетске специфичности локалних рудоносних средина.

Сва три метода се могу успешно користити као метод издвајања асоцијација елемената индикатора при геохемијским истраживањима. Важно је, међутим, напоменути да је код примене Пирсоновог коефицијента корелације (као и метода параметарске статистике уопште) и факторске анализе која се заснива на анализи матрице варијанси-коваријанси потребно обратити пажњу на тип дистрибуције оригиналних података те их подвргнути евентуалним трансформацијама како би се њихова расподела што је могуће више приближила нормалној.

Из поменутих разлога, при реализацији и интерпретацији података геохемијске проспекције, сугерише се шири примена статистичких метода.

|   |    |         |   |
|---|----|---------|---|
| Геол. ан. Балк. полуос.<br>Ann. Geol. Penins. Balk. | 62 | 455–463 | Београд, децембар 1998<br>Belgrade, Decembre 1998 |
|---|----|---------|---|

UDC 550.42:553.44/497.2(497.11–15)

Original scientific paper

## CORRELATION ANALYSIS AS A METHOD OF DETERMINING GEOCHEMICAL ASSOCIATIONS OF ELEMENTS IN THE ORE FIELD KALUDRA (SW SERBIA)

by

Nenad Lončarević\* and Boško Stajević\*

In the course of geochemical prospecting of potential ore-bearing terrains, defining indicator elements associations is one of the primary tasks. This problem is especially emphasized in the investigation of complex types of mineralizations comprising several ore-forming components, i.e. polymetallic ore occurrences.

Comparative analysis of results along with other relevant geological-geochemical features facilitates reliable determination of geochemical associations of indicator elements.

This paper presents three statistical methods of analysis of correlation between potential indicator elements during geochemical exploration and comparative analysis of results thus obtained. Results of geochemical investigations of the ore field Kaludra on the southeast flanks of Golija Mt. were taken as an example.

**Key words:** lithochemical prospecting, statistical analysis, correlation, factor analysis, polymetallic mineralization, ore field Kaludra, SW Serbia.

### INTRODUCTION

The Kaludra ore field is a taxonomically lower unit within the framework of Golija ore region and is located on its southeastern slopes. Geologically, it is built mostly of Senonian sandstone-allevrolitic, clayey-marly sediments and Jurassic-Cretaceous melanaged clastites into all of which have intruded quartzlatite-phenoandesitic dykes of Tertiary age (Fig. 1). Intrusions were followed by hydrothermal alterations of surrounding rocks and polymetallic impregnation-veinlet type Pb-Zn-Sb-Ag mineralization.

As a part of geochemical investigations of secondary dissemination haloes of this mineralization, 648 samples on a rectangular 100×40 m grid were analysed for 12 ore and accessory elements (Pb, Zn, Ag, Sb, Zn, Ba, Sr, Cu, Mo, Cd, W, Bi). The orientation and spacing of the exploratory grid were adapted to the geological build of the terrain and the expected size of geochemical anomalies. The geochemical field of Kaludra was presented by Stajević (1996).

---

\* Univeristy of Belgrade, Faculty of Mining and Geology, Djusina 7, 11000 Belgrade.

This paper presents three different methods of correlational analysis and their characteristics as means of separating geochemical associations of elements by quantification of their interrelations as well as a comparative analysis of the methods used.

### APPLIED METHODS

In order to present the ability of correlational analysis to separate geochemical associations, three methods were chosen:

1. Pearson's correlation coefficient ( $r$ ), as a representative of parametric statistical methods;
2. Spearman's rank correlation coefficient ( $R$ ), from the domain of non-parametric statistics and
3. Principal components analysis (PCA) as a part of factor analysis.

**Pearson's correlation coefficient ( $r$ )** is, to geologists, the most widely known and applied way of quantification and examination of relations between statistical variables. It is one of the methods of parametric statistics, meaning that it is calculated on the basis of previously estimated parameters of the given statistical population, in this case the covariance of two chemical elements and their standard deviations:

$$r = \frac{COV(x, y)}{\sigma_x \sigma_y}$$

This fact makes Pearson's  $r$  coefficient sensitive to the type of distribution of the original data which may lead to unrealistic results. One of the potential drawbacks of this method is also the fact that it expresses only *linear* correlation between two variables. That is why it is usually recommended to initially draw and examine scatter diagrams of variables in order to spot the possible existence of nonlinear correlation among variables.

**Spearman's rank correlation coefficient ( $R$ )** is one of the methods of non-parametric statistics which makes it immune to the type of distribution of data. The original data are transformed by ranking after which the correlation coefficient of transformed data is computed.

**Principal components analysis** belongs to the domain of Factor Analysis. Mathematically speaking, it is calculation of eigenvalues and eigenvectors of a variance-covariance matrix of the original data, which gives a set of  $n$  mutually orthogonal vectors that represent the variability among variables. Ideally,  $n$  is much smaller than the number of original variables, meaning that certain vectors represent several mutually dependent variables, thus expressing their correlation.

### RESULTS AND DISCUSSION

Of the 12 investigated chemical elements, 10 were chosen for the statistical analysis. Too many of the Bi and W contents were below the detection threshold of the applied analytical method for the data to be considered representative.

The skewness of distributions of Pb, Zn, Ag, Sb, As and Cd is greater than 6, which indicates that they are highly skewed to the left, approaching lognormal distri-



but. That is why it was decided that these data should be transformed i.e. replaced by their logarithms. This way the arithmetic mean and variance of these variables become true representatives of the parameters of their respective distributions. As an example of difference between using the raw as opposed to transformed data in such cases, we present a tree diagram where a measure of linkage distance is based on Pearson's correlation coefficient of the original (Fig. 2) and transformed data of the elements whose distribution approaches lognormal (Fig. 3).

It is notable that the  $r$  coefficient of raw data didn't manage to grasp the connection between As on one side and Pb, Zn, Ag and Sb on the other, as opposed to the same coefficient calculated on the basis of transformed data for the variables with skewed distributions.

The tree diagram based on Spearman's rank correlation coefficient (Fig. 4) shows much resemblance with Fig. 3, justifying the logarithmic transformation of the skewed data, since the  $R$  coefficient is a robust method immune to the type of distribution of the original data.

The PCA separated four factors: the first encompasses Pb, Zn, Ag, Sb and As, the second Ba and Sr, the third Mo and Cu and the fourth Cd (Fig. 5).

All three methods were successful at determining the geochemical association of indicator elements of the polymetallic mineralisation: Pb–Ag–Zn–Sb–As±Cd. The link of Cd to other elements can seem dubious on the basis of presented diagrams, although the fact that the anomalies of this element overlap the anomalies of others, especially zinc, remains. However, Cd anomalies are far less extensive while the other assays have Cd content of the order of Clarke which probably caused the "dilution" of the correlation of this element with the others, making it hard to detect.

## CONCLUSIONS

Correlational analysis of geochemical data of the ore field Kaludra facilitates determination and confirmation of certain geochemical–metallogenic characteristics of local ore-bearing environments.

All three methods can be used as means of determining the indicator elements associations in the course of geochemical investigations. However, we have to stress that, when using the Pearson  $r$  coefficient (or any other method of parametric statistics for that matter) as well as factor analysis which is based on the analysis of the variance–covariance matrix, it is necessary to pay attention to the type of distribution of the original data and transform them if needed in order to obtain a distribution as close to normal as possible.

That is why we suggest a wider use of statistical methods in interpretation of data obtained by geochemical prospecting.

*Translated by N. Lončarević*

## ЛИТЕРАТУРА – REFERENCES

- Стајевић Б. (=Stajevid), 1996: Геохемијско поље Калудре (ЈЗ Србија).– Геол. ан. Балк. полуос., 60/1. 305–318, Београд.