

## Оценка на кластер-класификация на геохимични данни от Боровишките вулкански скали — Източни Родопи

*Методи Караджов*

K a r a d j o v, M. 1991. Evaluation of cluster classification of geochemical data concerning the Borovitsa volcanic rocks, Eastern Rhodopes. — *Geochem., Miner. and Petrol.*, 27, 108-115. †

Ten rock samples of the Borovitsa volcanic series have been analysed by instrumental methods for 34 chemical elements. The data obtained are interpreted by cluster analysis for the purposes of geochemistry.

The study compares five different classification strategies in order to find a cluster which represents most reliably the geochemical nature of the magmatic process. The following conclusions can be drawn from the clusters obtained:

1) Data clustering based on a chemical elements by rocks matrix cannot be used for the purposes of geochemistry; 2) When that matrix is transposed, i. e. rearranged to form a rock by chemical elements matrix, its clustering using the Euclidian distance as a similarity measure is the most reliable approach to geochemical classification; 3) The use of at least two clustering strategies is obligatory; 4) Information on the rock genesis can be gained by mathematical treatment of the untransposed matrix.

*Key words:* geochemical data, cluster analysis, Borovitsa volcanic area.

*Address:* Bulgarian Academy of Sciences, Geological Institute, 1113 Sofia.

### Увод

Аналитичните данни са основата, върху която се развиват геохимичните и петрохимичните изследвания. Един опит да се извлече максимална информация от тези данни с цел изучаване на даден природен обект представлява т. нар. кластер-анализ.

Кластерирането представлява пакет от стратегии, с които се постига определена класификация и се разкриват някои връзки между получените аналитични параметри или изследваните обекти. В литературата (M a s s a r t, K a u f m a n, 1983) по-често са кластерирани различни обекти на базата на съдържанията на химичните елементи или други физико-химични параметри, но това в повечето случаи разрешава петрохимически проблеми.

Геохимичната задача е изясняване на изменението на концентрациите на химичните елементи в изучавания тук геоложки процес.

Интерпретацията на аналитичните данни за целите на геохимията чрез кластер-анализ може да се постигне по два начина. 1) Геохимичните параметри от нормалната матрица с данни се кластерират по коефициента на корелация. 2) Матрицата от данни се транспонира и за класифициране на елементите се използва друга мярка на подобие —  $\cos \theta$ , евклидово разстояние и др.

Целта е да се сравни кластерирането по двата начина на работа и да се покаже доколко формалната математическа логика съвпада с нашите познания за геохимическата същност на магмения процес. Разглежда се добре изучен район на Източнородопската вулканска област — Боровишкият.

## Петрохимични данни

Обект на изследването са представителни скали от Боровишката вулканска серия — Източни Родопи. Това е Еоцен-олигоценска шошонитова серия, изградена главно от средни и кисели скали. На базата на петрографски, минераложки, геохимични и изотопни данни М а р ч е в (1985) предполага, че скалните разновидности на серията са били образувани в резултат на фракционна кристализация, придружена от селективна горнокорова контаминация, смесване на магми и газов пренос. Фракционният механизъм е потвърден от разпределението на редкоземните елементи в скалите (D a i e v a, M a r c h e v, 1986). Изследвани са шошонити, латити, кварцлатити, нискосилициеви и високосилициеви риолити.

## Химико-аналитични данни

С инструментални методи са анализирани 34 химични елемента. Чрез рентгено-флуоресцентен анализ са определени петрогенните елементи, както и съдържанията на Y, Rb, Sr, Ba, Zr, V, Mn. Подготовката на пробите, измерването и обработката на резултатите са аналогични на описаните от Б а х т и а р о в (1985) и M i l l e r (1972). Неутронно-активационно са измерени La, Ce, Sm, Eu, Tb, Yb, Lu, Sc, Cs, Hf, Ta и Th при условия на измерването съгласно P e t e r s e n (1987). Атомноабсорбционният анализ на Li, Co, Ni, Zn, Pb и Cu следва методиката на С т о л я р о в и Ф и л а т о в а (1981), а златото е измервано след екстракционно концентриране (K o s h, K o s h-D e d i c, 1964). За контрол Li, Rb и Cs са анализирани и чрез пламъчна фотометрия, така че е постигната задоволителна достоверност на получените резултати.

## Кластер-анализ

Аналитичните резултати са оформени в матрица от 34 химични елемента от 10 скални обекта. Химичните елементи от нормалната матрица са групирани чрез анализ на главните компоненти и кластер-анализ с мярка на подобие коефициента на корелация (*R*-кластер). Класификацията на химичните елементи от транспонираната матрица е провеждана с кластер-анализ с мярка на подобие  $\cos \theta$  (*T*-кластер) и кластер-анализ с мярка на подобие евклидово разстояние (*E*-кластер). Освен това аналитичните ре-

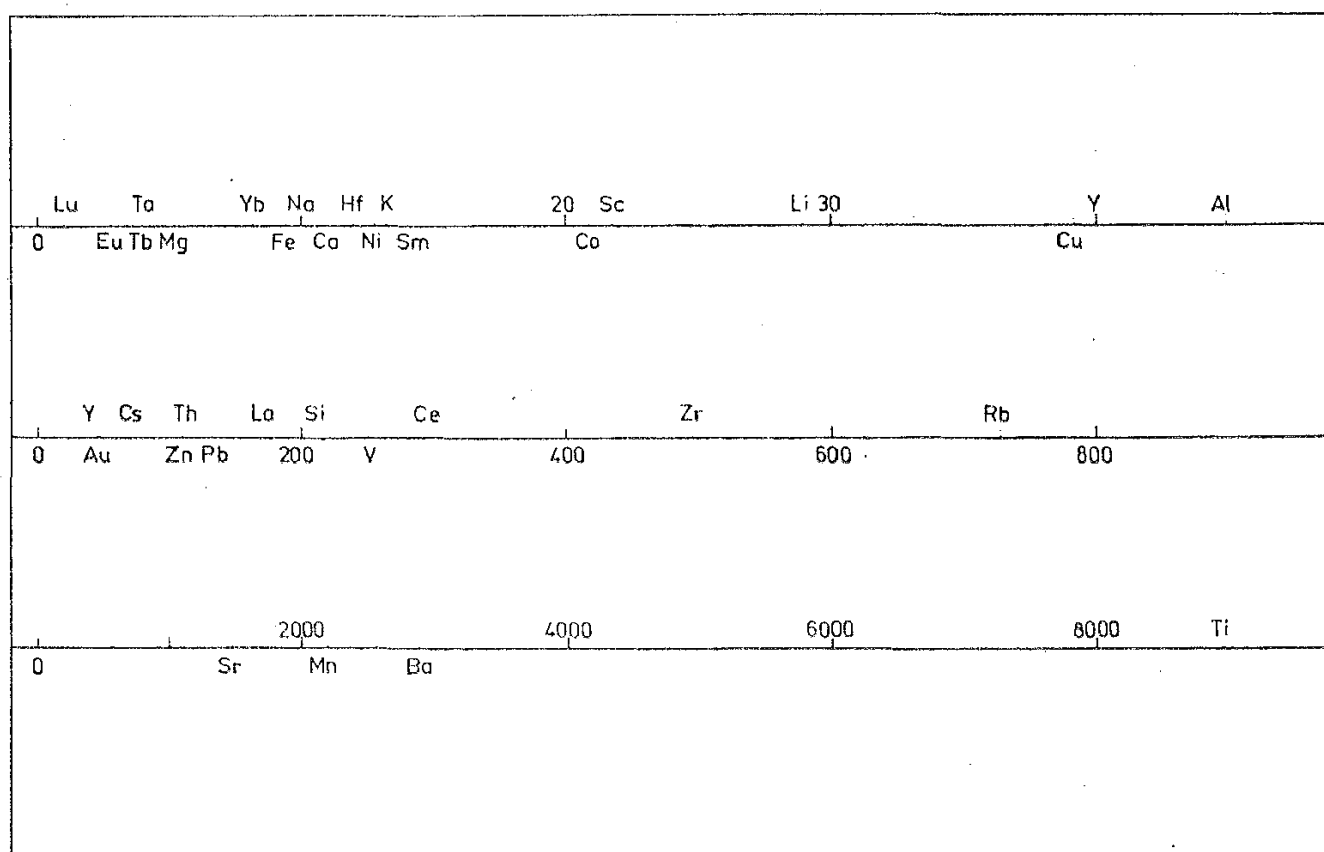
зултати са обработвани и с кореспондентен анализ. Програмното осигуряване е създадено от А. Андреев от Геологическия институт при БАН.

Подреждането на химическите елементи от петте статистически обработки е представено на фиг. 1—5. Кластер-връзките при дендрографите (фиг. 2—4) са разкъсвани на ниво 0,3 (критичната стойност на корелацията е 0,33 за 34 елемента, при ниво на значимост  $\alpha=0,05$ ) и на ниво 0,7 (критичната стойност е 0,63 за 10 проби (Д ъ о р ф е л, 1987)) за отделяне на групите елементи със значима корелация.

## Резултати и обсъждане

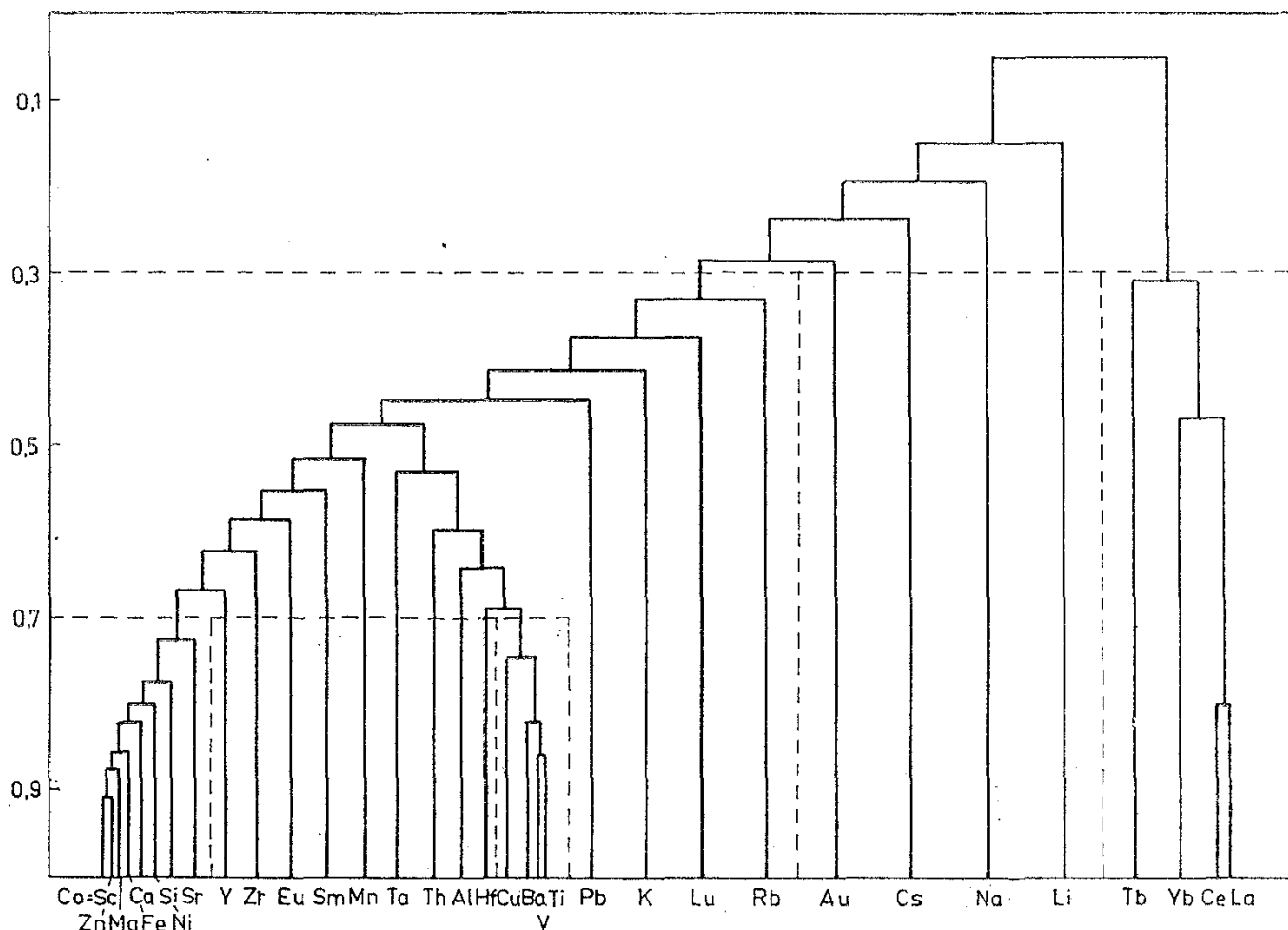
Използуваната математическа обработка трябва да отчита корелациите или ковариациите между химическите елементи и да ги подрежда в групи с подобни геохимични и кристалохимични свойства. Например би трябвало да се разграничат сидерофилните (Fe, Mg, Ca, Sc, Ni) елементи от литофилните (Si, K, Na, Th, Ta, TR), които се натрупват в по-киселите скали. Или, казано по друг начин, елементите, на които концентрацията нараства в хода на магмената диференциация, трябва да се отделят статистически от тези, които намаляват.

Подобно приблизително групиране се постига при обработката на данните от нормалната матрица, но се наблюдава и смесване на елементи с различно поведение. От представените на фиг. 1 проекции на векторите на кон-



Фиг. 1. Проекции на 34 химични елемента върху главния принципен компонент, разделен на три части. Размери на частите: 0—50; 0—1000; 0—10 000

Fig. 1. Projections of 34 chemical elements on the main component of principle, divided in three parts. Dimensions of the parts: 0—50; 0—1000; 0—10 000



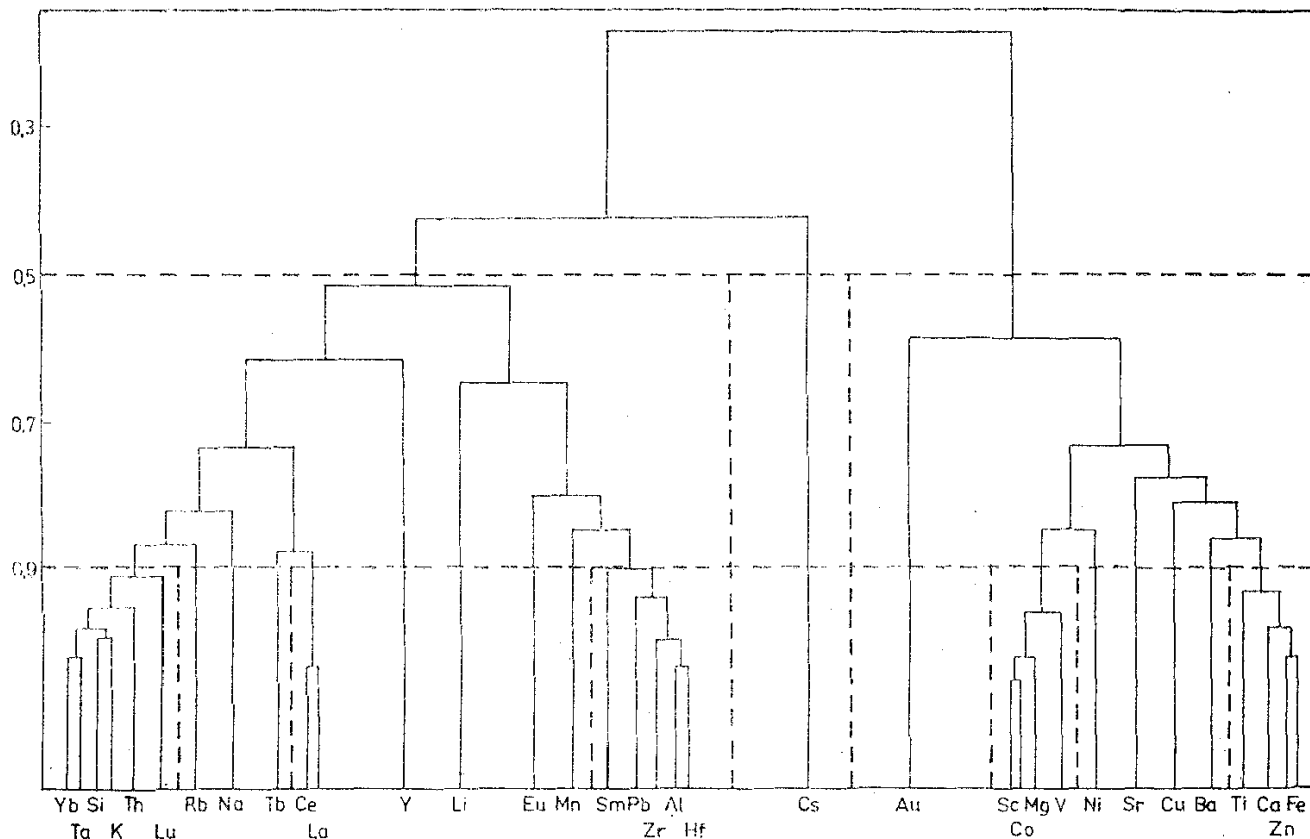
Фиг. 2. Групиране на 34 химични елемента от кластер с мярка на подобие коефициента на корелация (*R*-кластер)

Fig. 2. Dendrogram of 34 chemical elements based on a similarity matrix of the correlation coefficients (*R*-cluster)

центрациите върху първия главен компонент се вижда, че Ta и K са много отдалечени от групата на Si, Ce, La, Th и Y. Освен това те са разделени от принадлежащия на друга група скандий. По същия начин около Mg, Fe, Ca, Ni, Co са се наредили геохимически различните Eu, Tb, Sm.

Много по-значимо смесване се наблюдава при резултатите от *R*-кластера (фиг. 2). При добре оформената група на т. нар. лиофобни елементи — Co, Sc, Zn, Fe, Mg, Ca, Ni, е класиран и съвсем противоположният по свойства силиций. Подобно групиране е получил и Y a n e v et al. (1988) при аналогична математическа обработка. По-внимателното разглеждане на корелационната матрица и кластера показва, че корелационни коефициенти с близки положителни и отрицателни стойности се групират един до друг. Тази класификация, която има и други слабости (лиофилните елементи са много разпръснати, цирконият и хафният са значимо отделени), разкрива, че кластерирането с мярка на подобие коефициенти на корелация не е подходящо за геохимична интерпретация на аналитични данни.

Статистическата обработка на данните от нормалната матрица не дава значимите геохимически групи от елементи, но може да се използва за решаване на петрохимични проблеми. Например при компонентния анализ (фиг. 1) на първия главен компонент (Карунен-Льоеве трансформация)



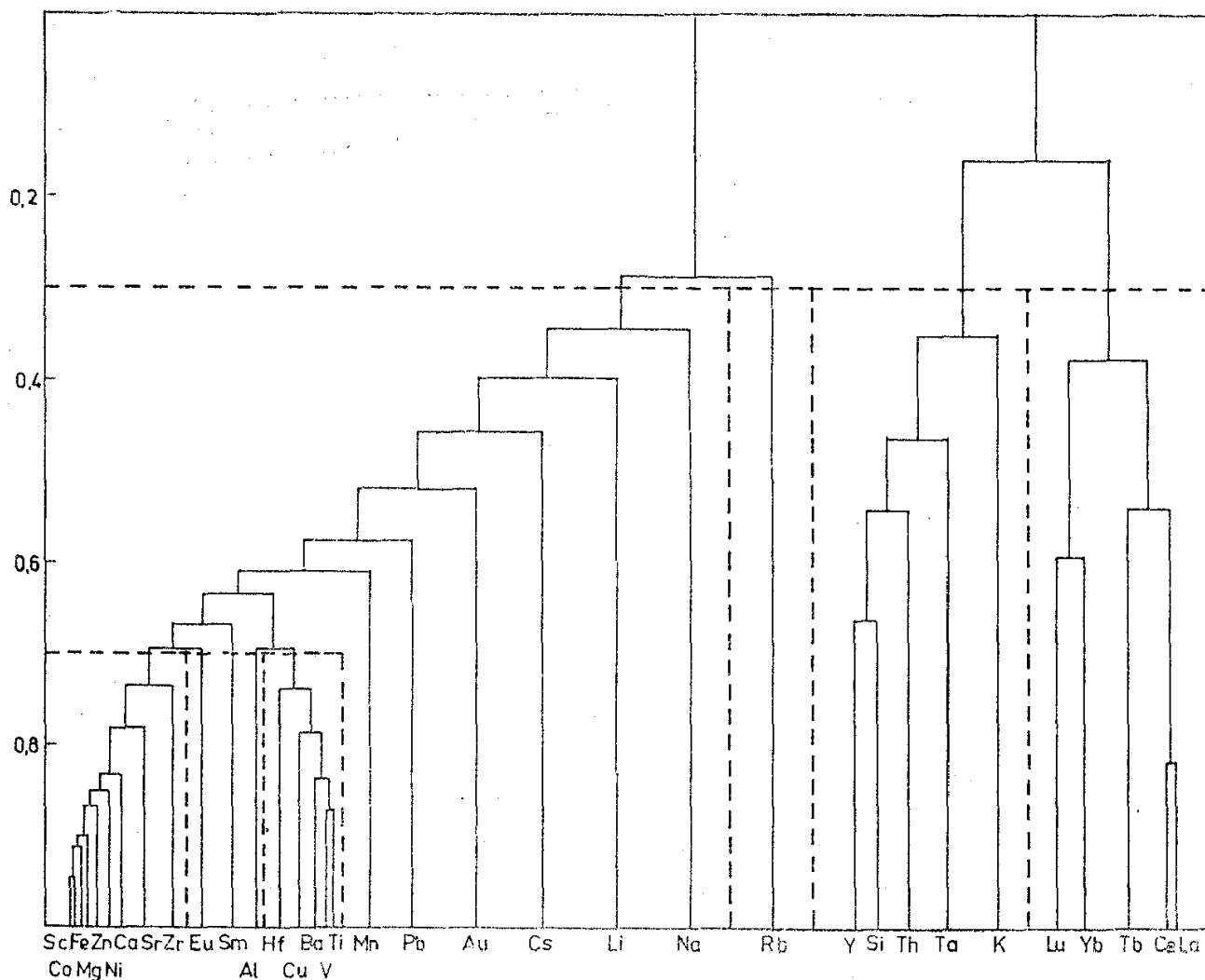
Фиг. 3. Групиране на 34 химични елемента от кластер с мярка на подобие  $\cos \theta$  ( $T$ -кластер)

Fig. 3. Dendrograph of 34 chemical elements based on a similarity matrix of  $\cos \theta$  ( $T$ -cluster)

съответствува 96,1% от общата дисперсия на данните. Стойностите на векторите по втория главен компонент, съдържащ 3,2% от дисперсията, не са нанасяни, а само е маркирано положението на елементите спрямо първия главен компонент. Тази стойност на дисперсията по първия главен компонент индикира (K u s m u s, P i e t s c h, 1986) общи характеристики на скалите, т. е. определя като доминиращ само един скалообразуващ процес.

На фиг. 3, 4 са представени  $T$ - и  $E$ -кластери, за които е прилагана математическа обработка на транспонираната матрица от данни. И двата кластера очертават добре при нива на корелация 0,3 и 0,7 три групи елементи: групата на лиофилните елементи, които са характерни елементи за киселия магматизъм — Si, K, Ta, Th и TR (M а р а к у ш е в, 1979; M е й с о н, 1971); групата на лиофобните елементи — Mg, Ca, Fe, Zn, Ni, Co и Sc, които влизат преимуществено в състава на по-раннообразуваните минерали (M а r c h e v e t a l., 1987); групата на Al, Pb, Au, Mn, Li и Cs, които са с по-ниски коефициенти на ковариация в изследваните магмени скали. Класифицирането на повечето от редките и разсеяните елементи може да намери геохимично обяснение. Например Eu и Sm, които като MREE имат по-особено поведение (M и н е е в, 1972), се отделят от другите редки земи. Li и Cs се обособяват, защото не образуват собствени минерали, и се концентрират предимно в биотита.

Като цяло  $E$ -кластерът очертава по-добре специфичните групи елементи. Групата на редкоземните елементи е компактна в близост до лиофилните елементи. Леофобните елементи образуват цяла неразделена от



Фиг. 4. Групиране на 34 химични елемента от кластер с мярка на подобие евклидово разстояние (*E*-кластер)

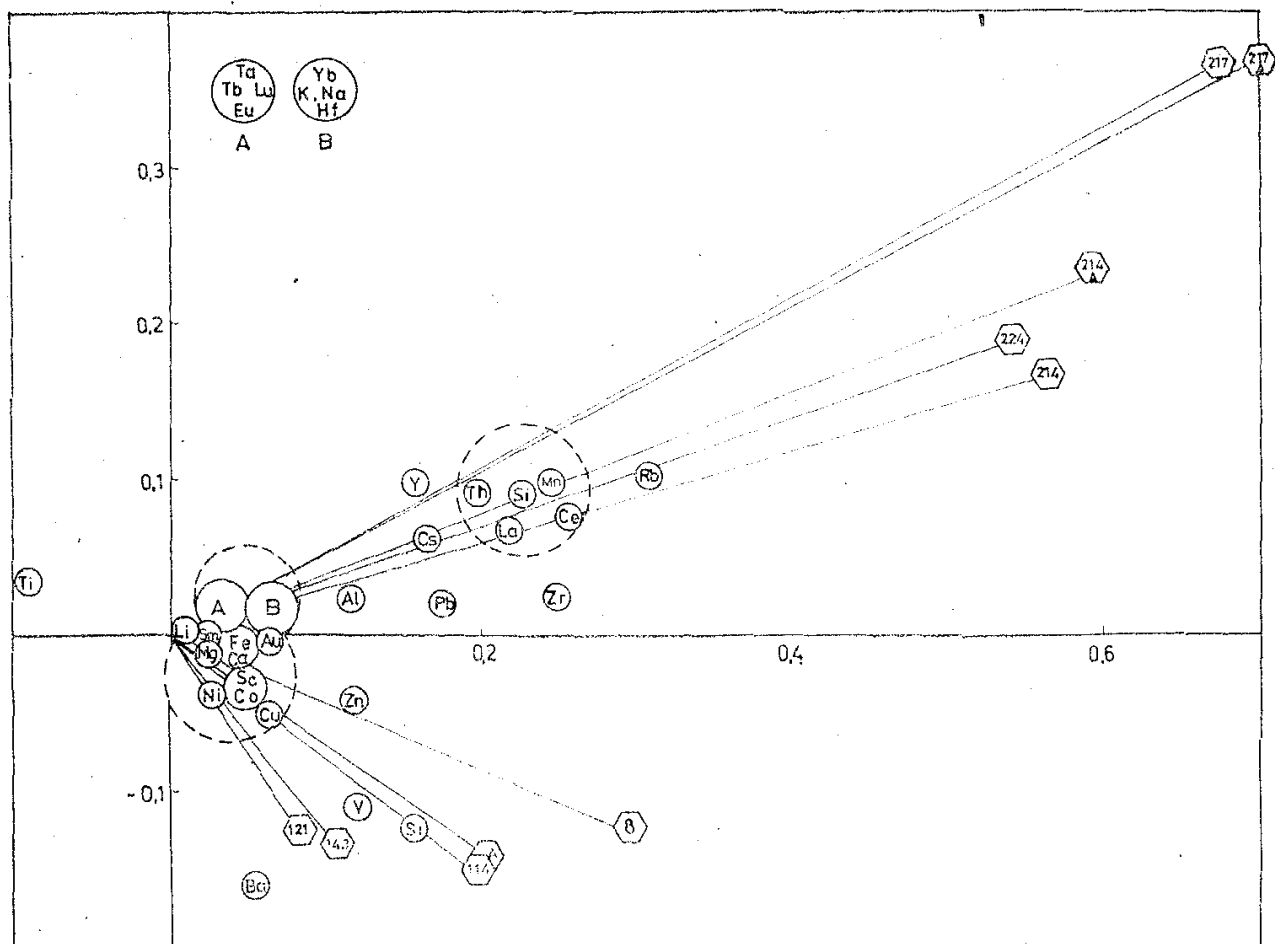
Fig. 4. Dendrogram of 34 chemical elements based on a similarity matrix of the Euclidian distances (*E*-cluster)

други елементи група. Изглежда, че кластерирането с мярка на подобие евклидово разстояние е най-подходящо за отразяването на геохимичната същност на вулканския процес.

Разпределението от кореспондентния анализ е представено на фиг. 5. На фигурата се оформят три групи елементи, разделени по местоположение и от координатните оси: I — Si, Mn, Th, La, Ce; II — Eu, Sm, Yb, Tb, Lu, Ta, K, Na, Hf; III — Mg, Fe, Ca, Co, Sc, Ni, Cu.

Когато се сравни местоположението на I и II група, се вижда, че те лежат върху векторите, описващи киселите скали. По същото направление лежат и Li, Rb и Cs, очертани подобно и на *E*-кластера. Група III заедно със скандия и стронция корелира логично със средни скали и съвпада с лиофобната група от *T*- и *E*-кластери.

Кореспондентният анализ подобно на *E*-кластера разделя Zr от Hf. Това разделяне не е много голямо (корелационното ниво е 0,6; съотношението Zr/Hf варира от 35 до 48), но се наблюдава при анализа на голям обем геохимични данни.



Фиг. 5. Разпределение на 34 химични елемента от 10 скални проби чрез кореспондентен анализ. Проби: № 121 и 143 — шохонити, № 8 — латит, № 114, 114 А — кварцлатити, № 214, 214 А и 224 — нискосилициевни риолити, № 217 и 217 А — високосилициевни риолити. Пробите с „А“ след номера са от контактните зони на съответните скали

Fig. 5. Distribution of 34 chemical elements of 10 rock samples by a correspondent analysis. Samples: No 121 and 143 — shoshonites, No 8 — latite, No 114, 114 A — quartz-latites, No 214, 214 A and 224 — low -Si rhyolites, No 217 and 217 A — high-Si rhyolites. Samples having “A” after the No indicate contact zones of the corresponding rocks

Кореспондентният анализ, *E*-кластерът и анализът на главните компоненти определят на титана по-особено положение. То не може да се обясни с наличната информация, с която разполагаме, но може да индикира необичайни магмени процеси.

Така представените резултати могат да се обобщят в следните изводи:

1. Кластерирането на аналитичните данни от нормалната матрица не може да се използва за нуждите на геохимията.

2. При работа с транспонираната матрица от данни кластерът с мярка на подобие евклидово разстояние дава в този конкретен случай по-достоверно групиране на химичните елементи.

3. Кореспондентният анализ в комбинация с кластер-класификация може да бъде добър критерий за верността на геохимичната информация.

4. Част от математическата обработка дава информация за генезиса на скалите, което придава практическа стойност на получените резултати.

## Заклучение

Статистическият анализ на резултатите е по-достоверен, когато се използват поне две „кластериращи“ стратегии. Той дава изходна и насочваща информация за тълкуването на геохимията и петрохимията на изследваните магмени комплекси и серии, която подпомага изучаването на генезиса, изходните магми и механизма на кристализационната диференциация.

## Литература

- Бахтиаров, А. В. 1985. Рентгено-спектрален флуоресцентен анализ в геологията и геохимията. Л., 143 с.
- Дьорфел, К. 1987. Статистика в аналитичната химия. С., Наука и изкуство. 197 с.
- Маракушев, А. А. 1979. Петрогенезис и рудообразование. М., Наука. 246 с.
- Марчев, П. Г. 1985. Кандидатска дисертация. С. 327 с.
- Мейсон, Б. 1981. Основы геохимии. М., Недра. 311 с.
- Минеев, Д. А. 1972. Составы лантанидов сосуществовающих минералов как индикаторы физических и химических условий минералообразования — В: Редкометаллические месторождения, их генезис и методы исследования. М., Недра.
- Столяров, И. А., М. П. Филатова. 1981. Атомно-абсорбционная спектрометрия при анализе минерального сырья. Л., Недра. 152 с.
- Daieva, L. I., P. G. Marchev. 1986. REE in Borovitsa Eo-Oligocene shoshonitic series. — C. R. Bulg. Acad. Sci., **39**, 79-81.
- Koch, O. G., G. A. Koch-Dedici. 1964. Handbuch der Spurenanalyse. Springer-Verlag, S. 1232.
- Kusmus, W., R. Pietsch. 1986. Cluster analysis of the elemental composition of five Austrian peats. — Anal. Chim. Acta, **191**, 487-493.
- Marchev, P., A. Andreev, P. Lillov, M. Popov, L. Daieva. 1987. Petrogenesis of oligocene silicic lavas from Borovitsa volcanic area, SE Bulgaria. — In: International symposium on petrogenesis and mineralization of granitoids. December 7-10. Guangzhou, China, 215-216.
- Massart, D. L., L. Kaufman. 1983. The interpretation of analytical chemical data by the use of cluster analysis. N. Y., J. Wiley, p. 237.
- Müller, R. O. 1972. Spectrochemical analysis by X-ray fluorescence. Plenum Press, p. 326.
- Petersen, J. S. 1987. Multielement analysis of purified mineral separates. — J. Radioanal. and Nucl. Chem., **112**, No 1, 175-192.
- Yanev, Y., Y. Yordanov, R. Boyadjieva, A. Andreev. 1988. Geochemistry of the collision related acid volcanism in the Eastern Rhodopes, Bulgaria. — Abstracts of the EAG Inter. Congress of Geochemistry and Cosmochemistry. — Chem. Geol., **71**, No 4, p. 370.

Одобрена на 29. IV. 1989 г.

Accepted April 29, 1989