

Использование кластерного анализа для классификации горных пород месторождений полезных ископаемых Азербайджана по их физико-техническим свойствам

Зарифа Джахангир гызы ЭФЕНДИЕВА*,

Азербайджанский государственный университет нефти и промышленности, Азербайджанская Республика, Баку

Актуальность исследования. Выбор оптимальной технологии и рациональной разработки месторождений полезных ископаемых зависит от достоверности знаний о физических свойствах горных пород разрабатываемых месторождений. Поэтому нам пришлось уделить большое внимание созданию классификации горных пород в месторождениях полезных ископаемых по физико-техническим свойствам.

Методология. Для решения поставленной задачи применен один из методов математической статистики – кластерный анализ. В качестве меры кластеризации принято евклидово расстояние, которое образуется при сравнении значений физических величин. Применяв кластерный анализ, нам удалось провести группировку месторождений по физико-техническим свойствам. Используя признаки сходства можно снизить объем горных работ и повысить эффективность разработки полезных ископаемых, что важно, особенно когда нужно ввести в эксплуатацию большое количество месторождений с разнообразными физическими показателями.

Выводы. Таким образом, применение методов кластерного анализа позволяет проводить группировку месторождений по физическим показателям, особенно когда изучается большое количество месторождений и показателей и нужен общий анализ их состояния. На основе группировки и степени сходства месторождений для обеспечения эксплуатации можно искать рациональные пути. По физическим параметрам сходство месторождений можно считать обоснованным. Используя признаки сходства, применяя ту же технологию для разработки подобных месторождений Азербайджана, можем снизить объем горных работ и увеличить рациональность эксплуатации месторождений полезных ископаемых.

Ключевые слова: месторождения полезных ископаемых, кластерный анализ, классификации месторождений, метод математической статистики, меры кластеризации, степени сходства месторождений.

Введение
Известно, что большинство месторождений полезных ископаемых характеризуются сложным строением и различными формами залегания. Они изучаются с помощью различных горных выработок и бурения поисковых скважин с отбором необходимых проб для комплексного исследования. Но с увеличением глубины скважин извлечение керновых проб затрудняется, а также изменяются физические свойства кернов (минеральный состав, строение, пористость, твердость и др.), и без учета этих изменений эффективное ведение горных работ невозможно.

Возможные изменения пород по минеральному составу, строению и по их физическому свойству с глубиной залегания могут быть решены на основании применения теории вероятностей и математической статистики [1].

Иначе для получения надежных результатов физических параметров горных пород требуется проведение многочисленных измерений по каждому типу породы, все это сопровождается проведением достаточно большого объема работы и требует больших материальных расходов. Поэтому в этом случае применение методов математической статистики дает надежные результаты и позволяет сэкономить материальные ресурсы [2, 3].

Методы исследования

Иногда требуется группировка разрабатываемых месторождений полезных ископаемых по физико-техническим свойствам в регионе или на рудниках. В этом случае можно использовать метод кластерного анализа в качестве эффективного математического метода.

Кластерный анализ является одним из многомерных методов для классификации горных пород, вмещающих полезные ископаемые. Такой подход представляет собой совокупность методов, позволяющих классифицировать многомерные наблюдения, каждое из которых описывается набором исходных переменных X_1, X_2, \dots, X_n и позволяет разбить исходную совокупность объектов на группы схожих, близких между собой объектов [4–7].

Основной целью в кластерном анализе является выделение сравнительно небольшого числа групп объектов, как можно более схожих между собой внутри группы и как можно более отличающихся в разных группах. Этот вид анализа широко используется в информационных системах при решении задач классификации [8, 9].

Методы кластерного анализа можно разделить на две группы: иерархические и неиерархические.

Для создания классификации горных пород по физико-техническим свойствам нами был выбран иерархический метод. Группировка наблюдений осуществляется на основе сходства. Чаще всего в качестве такой меры используется евклидово расстояние, сущность которого раскрыта в следующей формуле:

$$d_{xy} = \sqrt{\sum_{i=1}^m (x_i - y_i)^2},$$

где m – количество свойств; x_i, y_i – расстояние между объектами; d_{xy} – значения i -го свойства объектов x и y . В двух- или трехмерном случае это прямая, соединяющая данные точки.

Евклидово расстояние является самой популярной метрикой в кластерном анализе. Оно является геометрическим расстоянием в многомерном пространстве [10, 11].

Далее представлен кластерный анализ для месторождения мраморизованных известняков. Здесь горные породы признаны объектами исследований.

✉ efendi2005@rambler.ru

Физические свойства горных пород характеризуют следующие параметры: плотность ρ_0 , коэффициент крепости f , пористость P , предел прочности при сжатии $\sigma_{сж}$, модуль упругости E , коэффициент Пуассона ν , коэффициент теплопроводности λ , удельная теплоемкость C_m , влажность ω , коэффициент хрупкости k_x и т. д.

Результаты

На территории Республики Азербайджан находятся многочисленные месторождения мраморизованных известняков, в том числе Велидагское, Шушинское, Ханкендинское и др. [12–15]. Имеется 14 месторождений (таблица), где необходимо провести кластерный анализ. Эти месторождения характеризуются 9 признаками [16].

Классификационный способ в изучении этих месторождений дал следующий результат: в республике соединены в три группы месторождения мраморизованных известняков по их физико-техническим свойствам (рис. 1).

В первый отдельный кластер вошли 7 месторождений: Велидагское, Шушинское, Ханкендинское, Гочазское, Ходжавандское, Заринбахское, Гюлаблинское, во второй – 5 месторождений: Горовское, Шеллинское, Охчичайское, Дияллинское, Карамурадское, в третий – остальные месторождения: Агчайское, Халачское, причем первый и второй кластеры во многом сходны.

Из дендрограммы видно, что Шушинское и Ханкендинское месторождения объединились в один кластер. Проходит другая цепочка. Гочазское и Ходжавандское также объединились в один кластер, а Заринбахское, Гюлаблинское и Велидагское месторождения пытаются присоединиться к ним. Значит, эти 7 месторождений сгруппировались.

Во второй группе в первом шаге объединяются Горовское и Шеллинское месторождения, расстояние между которыми $d_{ij} = 1,4$, далее к объектам присоединяется Дияллинское месторождение. На евклидовом расстоянии $d_{ij} = 2,3$ к ним присоединяются Охчичайское и Карамурадское месторождения и составляют одну группу. Соединение первой и второй групп происходит по евклидову расстоянию $d_{ij} = 9,9$.

Можно отметить, что Агчайское и Халачское месторождения, несмотря на близость объектов, по степени сходства с другими месторождениями очень слабые. В процессе разработки месторождений полезных ископаемых эти степени разности должны учитываться.

На евклидовом расстоянии ($d_{ij} = 12,4$) происходит присоединение между этими группами и получается один кластер.

Для месторождений мраморизованных известняков рассчитываем евклидово расстояние, оценку их сходства и различия.

Существует большое разнообразие алгоритмов кластеризации [17, 18]. Воспользуемся агломеративным иерархическим алгоритмом классификации. Принято 9 основных физико-технических параметров мраморизованных известняков [19].

Наиболее распространенный способ – вычисление евклидова расстояния между двумя точками i и j на плоскости, когда известны их координаты x и y :

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}. \tag{1}$$

Примечание: чтобы узнать расстояние между двумя точками, надо взять разницу их координат по каждой оси, возвести ее в квадрат, сложить полученные значения для всех осей и извлечь квадратный корень из суммы.

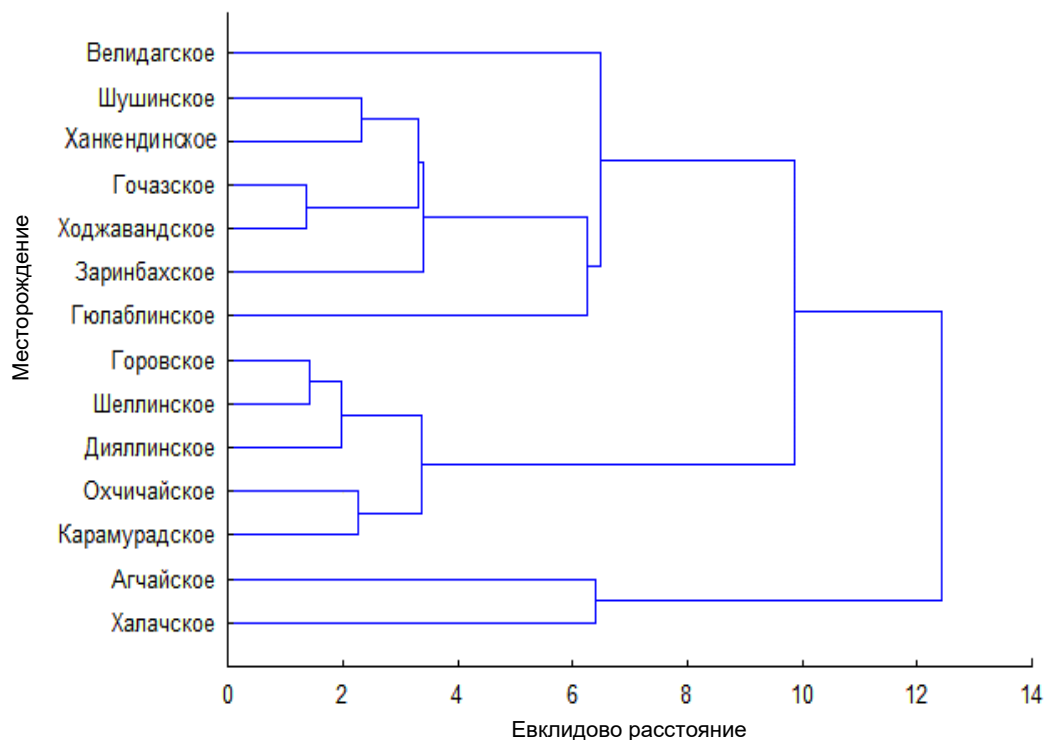


Рисунок 1. Дендрограмма 14 исследуемых месторождений мраморизованных известняков по их физико-техническим свойствам. Figure 1. Dendrogram of 14 studied deposits of marbled limestone by their physical and technical properties

**Физико-технические свойства горных пород месторождений мраморизованных известняков.
Physical and technical properties of rocks of deposits of marbled limestones.**

Месторождение	ρ , т/м ³	f	P , %	$\sigma_{сж}$, МПа	$E \cdot 10^{-5}$, МПа	λ , Вт/К	C_m , Дж/(кг × К)	ω , %	k_x , %
Велидагское	2,66	4,68	1,52	81,2	5,40	1,17	1,07	0,96	0,64
Шушинское	2,55	3,90	1,54	75,0	5,20	1,16	1,10	0,82	0,70
Ханкендинское	2,72	3,84	1,95	44,0	5,38	1,15	1,06	0,73	0,80
Гочазское	2,60	3,88	1,60	55,0	5,60	1,17	1,08	0,78	0,79
Ходжавендское	2,71	4,30	1,70	81,2	5,30	1,21	1,30	0,97	0,67
Заринбахское	2,73	4,87	1,64	89,5	5,50	1,21	1,25	0,91	0,63
Гюлаблинское	2,66	5,00	1,50	98,3	5,40	1,20	1,12	1,00	0,62
Горовское	2,70	4,90	1,65	82,0	5,30	1,17	1,06	0,91	0,69
Шеллинское	2,56	3,94	1,83	58,4	5,26	1,18	1,10	0,89	0,80
Дияллинское	2,64	3,87	1,80	60,0	5,44	1,16	0,99	0,92	0,74
Охчайское	2,68	4,13	1,82	60,0	5,24	1,12	1,06	0,78	0,79
Карамурадское	2,67	4,15	1,71	65,2	5,26	1,19	0,98	0,76	0,77
Агчайское	2,70	3,97	1,78	45,0	5,31	1,18	0,99	1,00	0,74
Халачское	2,71	4,30	1,60	81,2	5,30	1,21	1,08	0,87	0,67

Когда осей больше двух, расстояние рассчитывается таким образом: сумма квадратов разницы координат состоит из столько слагаемых, сколько осей (измерений) присутствует в нашем пространстве. Например, если нам нужно найти расстояние между двумя точками в пространстве трех измерений (такая ситуация представлена на рис. 2), формула (1) приобретает вид:

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}.$$

Тогда в соответствии с этой формулой расстояние между физическими параметрами горных пород составляет:

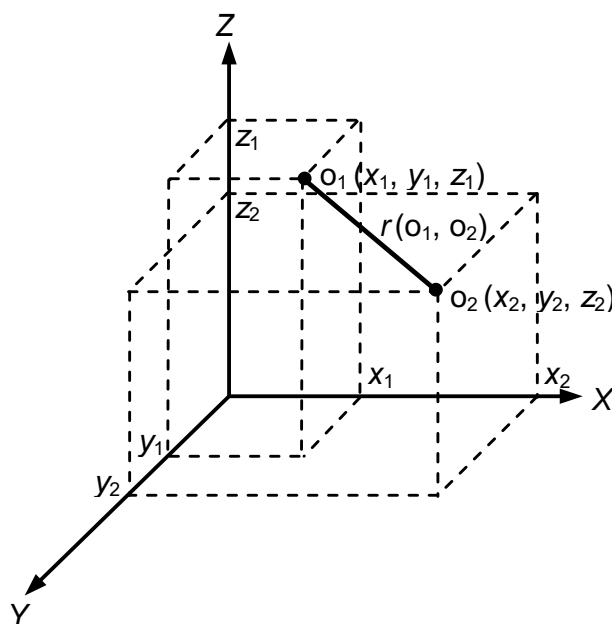
$$d_{xy} = \sqrt{(\rho_{0x} - \rho_{0y})^2 + (P_x - P_y)^2 + (\sigma_{сж,x} - \sigma_{сж,y})^2 + (\lambda_x - \lambda_y)^2 + (E_x - E_y)^2 + (C_{mx} - C_{my})^2 + (\omega_x - \omega_y)^2 + (k_{xx} - k_{xy})^2}.$$

Расстояние между объектами зависит от масштаба по осям.

Данный эффект можно объяснить тем, что уровни значений переменных физико-технических свойств горных пород очень сильно отличаются друг от друга. Согласно формуле евклидовой меры, переменная, имеющая большие значения, практически полностью доминирует над переменной с малыми значениями.

**Оригинальные значения данных.
Original data values.**

ρ_0	f	P	$\sigma_{сж}$	λ	E	C_m	ω	k_y
2,55	3,84	1,50	44,0	1,15	5,20	0,98	0,73	0,62
2,73	5,00	1,95	98,3	1,21	5,60	1,30	1,00	0,80



**Рисунок 2. Расстояние между двумя точками в пространстве трех измерений.
Figure 2. Distance between two points in the space of three dimensions.**

$$d_{\rho_0}^2 = (2,73 - 2,55)^2 = 0,0324;$$

$$d_f^2 = (5,00 - 3,84)^2 = 1,3456;$$

$$d_p^2 = (1,95 - 1,50)^2 = 0,2025;$$

$$d_{\sigma_{сж}}^2 = (98,3 - 44,0)^2 = 2948,49;$$

$$d_{\lambda}^2 = (1,21 - 1,15)^2 = 0,0036;$$

$$d_E^2 = (5,60 - 5,20)^2 = 0,16;$$

$$d_{C_m}^2 = (1,30 - 0,98)^2 = 0,1024;$$

$$d_{\omega}^2 = (1,00 - 0,73)^2 = 0,0729;$$

$$d_{k_x}^2 = (0,80 - 0,62)^2 = 0,0324.$$

$$d_{xy} = \sqrt{0,0324 + 1,3456 + 0,2025 + 2948,49 + 0,0036 + 0,16 + 0,1024 + 0,0729 + 0,0324} = 54,3.$$

В этом примере расстояние между объектами будет определяться в основном разницей содержаний $\sigma_{сж}$. Другие параметры практически не будут учтены при выделении кластеров, поскольку их содержания на порядок меньше, чем содержания $\sigma_{сж}$.

Чтобы сделать масштабы по осям одинаковыми, перед кластеризацией исходные величины нормируют, используя следующую формулу:

$$x_{\text{норм}} = \frac{x}{\bar{x}} \text{ (нормировка на среднее).}$$

Стандартизированные значения данных.
Standardized data values.

ρ_0	f	P	$\sigma_{сж}$	λ	E	C_m	ω	k_x
2,55	3,84	1,50	44,0	1,15	5,20	0,98	0,73	0,62
2,73	5,00	1,95	98,3	1,21	5,60	1,30	1,00	0,80
2,64	4,42	1,72	71,1	1,18	5,40	1,14	0,86	0,71

Если мы производим такое преобразование для переменных физико-технических свойств горных пород, получаем:

Нормированные значения.
Normalized value.

ρ_0	f	P	$\sigma_{сж}$	λ	E	C_m	ω	k_x
1,03	1,15	1,14	1,61	1,02	1,03	1,16	1,17	1,14
0,96	0,88	0,88	0,72	0,97	0,96	0,87	0,86	0,88

$$d_{\rho_0}^2 = (1,03 - 0,96)^2 = 0,0049;$$

$$d_f^2 = (1,15 - 0,88)^2 = 0,0729;$$

$$d_p^2 = (1,14 - 0,88)^2 = 0,0676;$$

$$d_{\sigma_{сж}}^2 = (1,61 - 0,72)^2 = 0,7921;$$

$$d_{\lambda}^2 = (1,02 - 0,97)^2 = 0,0025;$$

$$d_E^2 = (1,03 - 0,96)^2 = 0,0049;$$

$$d_{C_m}^2 = (1,16 - 0,87)^2 = 0,0841;$$

$$d_{\omega}^2 = (1,17 - 0,86)^2 = 0,0961;$$

$$d_{k_x}^2 = (1,14 - 0,88)^2 = 0,0676.$$

После нормализации значения физических свойств горных пород составляют:

– нормированные значения

$$d_{xy} = \sqrt{0,0049 + 0,0729 + 0,0676 + 0,7921 + 0,0025 + 0,0049 + 0,0841 + 0,0961 + 0,0676} = 1,3;$$

следовательно, у нас теперь есть нормализованные наборы данных в сравнимых шкалах;

– оригинальные значения

$$d_{xy} = \sqrt{0,0324 + 1,3456 + 0,2025 + 2948,49 + 0,0036 + 0,16 + 0,1024 + 0,0729 + 0,0324} = 54,3.$$

Таким образом, мы провели самый простой кластерный анализ. Мы выбрали такой вид графического представления, с помощью которого можно было бы отчетливо распознать группирование в кластеры.

Выводы

Применение методов кластерного анализа позволяет проводить группировку месторождений по физико-техническим свойствам, особенно когда изучается большое количество месторождений и показателей и нужен общий анализ их состояния. Используя признаки сходства, применяя ту же технологию для разработки подобных месторождений Азербайджана, можно снизить объем горных работ и увеличить рациональность эксплуатации месторождений полезных ископаемых.

ЛИТЕРАТУРА

1. Вентцель Е. С., Овчаров Л. А. Теория вероятностей и ее инженерные приложения. М.: Наука, 1988. 480 с.
2. Рыжов П. А. Математическая статистика в горном деле. М.: Высш. школа, 1973. 287 с.
3. Шарапов И. П. Применение математической статистики в геологии. М.: Недра, 1971. 246 с.
4. Дюран Б., Оделл П. Кластерный анализ. М.: Книга по требованию, 2012. 128 с.
5. Handbook of Cluster Analysis / C. Hennig, M. Meila, F. Murtagh, R. Rocci (eds). Boca Raton: CRC Press, 2016. 754 p. <http://dx.doi.org/10.1111/rssa.12219>
6. Biernacki Ch. Introduction to cluster analysis and classification: Performing clustering. Summer School on Clustering, Data Analysis and Visualization of Complex Data, 2018. 67 p.
7. Classification and cluster / Ed. J. Wang Rayzina. N. Y.: Wiley, 1980. 390 p.
8. Kaufman L., Rousseau P. J. Finding Groups in Data: An Introduction to Cluster Analysis. Hoboken, New Jersey: Wiley & Sons, 2005. 342 p.
9. King R. S. Cluster Analysis and Data Mining: An Introduction. Herndon: Mercury Learning and Information, 2015. 300 p.
10. Brian S., Everitt, Landau S., Morven L., Stahl J. D. Cluster analysis. Hoboken, New Jersey: Wiley & Sons, 2011. 346 p.
11. Феклистова И. С. Использование кластерного анализа при оценке эффективности стратегического управления предприятиями региона // Траектория науки. 2016. № 2 (7). С. 2.1–2.15.
12. Минерально-сырьевые ресурсы Азербайджана. Баку: Озан, 2005. 807 с.
13. Geology and Mineral Resources of Azerbaijan. N. Y.: United Nations, 2000. 216 с.
14. Нагиев В. Н., Мамедов И. А. Полезные ископаемые Нахичеванской автономной республики. Баку, 2009. 276 с.
15. Геология Азербайджан. Т. VI. Полезные ископаемые. Баку: Нафта Пресс, 2003. 576 с.
16. Эфендиева З. Дж. Информационный справочник о месторождениях полезных ископаемых Азербайджана. Баку: АГУНП, 2016. 57 с.
17. Дэвис Дж. С. Статистический анализ данных в геологии. В 2 кн. / пер. с англ. В. А. Голубовой; под. ред. Д. А. Радионова. М.: Недра, 1990. 427 с.
18. Stahl D., Sallis H. Model-Based Cluster Analysis // Wiley Interdisciplinary Reviews: Computational Statistics. 2012. Vol. 4. № 4. P. 341–358.
19. Эфендиева З. Дж. Характеристика залежей и физические свойства облицовочных камней месторождений Азербайджана // Горный журнал. 2005. № 8. С. 46–47.

Статья поступила в редакцию 02 октября 2019 года

The use of cluster analysis for the classification of rocks of Azerbaijan mineral deposits by their physical and technical properties

Zarifa Jakhangir AFANDIYEVA*,

Azerbaijan State Oil and Industry University, Azerbaijan Republic, Baku

Relevance of the research. The choice of optimal technology and rational development of mineral deposits depends on the reliability of knowledge about the physical properties of rocks of the developed deposits. Therefore, we had to pay great attention to creating a classification of rocks in mineral deposits according to their physical and technical properties.

Methodology. To solve this problem, one of the methods of mathematical statistics was applied - cluster analysis. As a measure of clustering, the Euclidean distance is taken, which is formed when comparing the values of physical quantities. Using cluster analysis, we were able to group deposits according to their physical and technical properties. Using similarities, it is possible to reduce the amount of mining operations and improve the efficiency of mining, which is important, especially when you need to put into operation a large number of deposits with various physical parameters.

Conclusions. Thus, the application of cluster analysis methods makes it possible to group deposits according to physical indicators, especially when a large number of fields and indicators are studied and a general analysis of their condition is needed. On the basis of the grouping and the degree of similarity of the deposits, lean approaches can be sought to ensure exploitation. In terms of physical parameters, the similarity of deposits can be considered reasonable. Using similarities and applying the same technology to develop similar deposits in Azerbaijan, we can reduce the amount of mining and increase the rationality of the exploitation of mineral deposits.

Keywords: mineral deposits, cluster analysis, classification of deposits, method of mathematical statistics, clustering measures, degree of similarity of deposits.

REFERENCES

1. Ventzel E. S., Ovcharov L. A. 1988, *Teoriya veroyatnostey i yeye inzhenernyye prilozheniya* [Probability theory and its engineering applications]. Moscow, 480 p.
2. Ryzhov P. A. 1973, *Matematicheskaya statistika v gornom dele* [Mathematical statistics in mining]. Moscow, 287 p.
3. Sharapov I. P. 1971, *Primeneniye matematicheskoy statistiki v geologii* [Application of mathematical statistics in geology]. Moscow, 246 p.
4. Duran B., Odell P. 2012, Cluster analysis. Moscow, Book on demand, 128 p.
5. C. Hennig, M. Meila, F. Murtagh, R. Rocci (eds). 2016, *Handbook of Cluster Analysis*. Boca Raton: CRC Press, 754 p.
<http://dx.doi.org/10.1111/rssa.12219>
6. Biernacki Ch. 2018, Introduction to cluster analysis and classification: Performing clustering. Summer School on Clustering, Data Analysis and Visualization of Complex Data. 67 p.
7. 1980, Classification and cluster. Ed. J. Wang Rayzina. N. Y.: Wiley, 390 p.
8. Kaufman L., Rousseau P. J. 2005, *Finding Groups in Data: An Introduction to Cluster Analysis*. Hoboken, New Jersey: Wiley & Sons, 342 p.
9. King R. S. 2015, *Cluster Analysis and Data Mining: An Introduction*. Herndon: Mercury Learning and Information, 300 p.
10. Brian S., Everitt, Landau S., Morven L., Stahl J. D. 2011, *Cluster analysis*. Hoboken, New Jersey: Wiley & Sons, 346 p.
11. Feklistova I. S. 2016, Application of cluster analysis in assessing the effectiveness of strategic management of enterprises in the region. *Traektoriya Nauki* [Path of Science], No 2 (7), pp. 2.1–2.15. (In Russ.)
12. 2005, *Mineral'no-syr'yevyye resursy Azerbaydzhana* [Mineral resources of Azerbaijan]. Baku, 807 p.
13. 2000, *Geology and Mineral Resources of Azerbaijan*. N. Y.: United Nations, 216 p.
14. Nagiev V. N., Mamedov I. A. 2009, *Poleznyye iskopayemye Nakhichevanskoj avtonomnoy respubliki* [Minerals of the Nakhichevan Autonomous Republic]. Baku, 276 p.
15. 2003, *Geologiya Azerbaydzhana* [Geology of Azerbaijan]. Vol. VI. Mineral resources. Baku, 576 p.
16. Efendieva Z. J. 2016, *Informatsionnyy spravochnik o mestorozhdeniyakh poleznykh iskopayemykh Azerbaydzhana* [Information guide on mineral deposits of Azerbaijan]. Baku, 57 p.
17. Davis J. C. 1990, *Statistics and data analysis in geology*. In 2 books. Transl. from English by V. A. Golubova; ed. by D. A. Radionov. Moscow, 427 p.
18. Stahl D., Sallis H. 2012, *Model-Based Cluster Analysis*. *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 4, no. 4, pp. 341–358.
19. Efendieva Z. J. 2005, Description of deposits and physical properties of facing stones of Azerbaijani deposits. *Gornyy zhurnal* [Mining Journal], no. 8, pp. 46–47. (In Russ.)

The article was received on October 02, 2019