UDC 550.42:553.98 DOI: https://doi.org/10.15407/geotm2022.163.005

METHOD OF CLUSTERIZATION OF C6 COAL SEAM ZONES OF DIFFERENT THICKNESS IN THE DNIPROVSKA MINE FIELD BY GERMANIUM CONCENTRATION

[1]Ishkov V.V., [2, 3]Kozii Ye.S.

[1]Institute of Geotechnical Mechanics named by N. Poljakov of National Academy of Sciences of Ukraine, [2]Dnipro University of Technology, [3]State Agrarian and Economic University

Abstract. In the article, based on the conducted research and analysis of the results of clustering, the most effective method of creating objective typification of zones of coal seam c_6 of the Dniprovska Mine of different thickness according to germanium concentrations was established. The typification procedure is a systematization of objects according to a priori specified features. Cluster analysis, taxonomy, pattern recognition, and factor analysis are usually used for this purpose. The authors of the article carried out clustering using various methods implemented in the most different professional statistical software platforms, performed their analysis and justified the choice of the most optimal of them. The optimal method of clustering areas of different thicknesses of the coal seam was determined. The analysis of the dendrogram of the results of clustering by the weighted centroid-median method of the sites of the c_6 seam by germanium content, unlike others, allows not only to achieve the most stable division of the entire set of sites under consideration, but also to maximize the visualization of their division by classes at different scale levels in the absence of a priori hypotheses regarding the number clusters and their forms. At the same time, the structure of clusters is clearly distinguished, regardless of the scale level of their formation, and the sequence of combining individual deposits and their groups into the resulting cluster is clearly traced. These advantages make it possible to make maximum use of already existing information for the development of natural typifications of c_6 coal seam areas by germanium content and to interpret the obtained results in geological terms. It was established that the weighted centroid median method of cluster analysis is the most optimal for the subjectivity-free researcher to divide the sections of coal seam c_6 of the Dniprovska Mine field according to germanium content into taxa. The constructed dendrograms of the clustering of deposits by germanium content can be used as a basis for the development of a natural typification of the coal seams of the Dniprovska Mine for their subsequent geological and economic assessment. This, in turn, will make it possible to make maximum use of already available information and interpret the obtained results in geological and genetic concepts, which will provide the opportunity to use it for the complex use of mineral raw materials and to solve strategic issues of sustainable development of Ukraine.

Keywords: cluster analysis, coal seam, germanium, clusters, taxa, dendrogram of clustering results, weighted centroid method.

1. Introduction

The research relevance of the content of germanium in coal seams is due to the possibility of its industrial extraction and use as a valuable accompanying component [1–3]. Currently, coal is the main estimated source of germanium in Ukraine, China, Uzbekistan, as well as in Russia, and germanium coal deposits are also being developed in England, Canada, the USA, and other countries [4–6].

The research carried out is particularly relevant to the decision of the National Security and Defense Council of Ukraine dated July 16, 2021 "About stimulation of the search, extraction and enrichment of minerals that are of strategic importance for the sustainable development and defense capability of the state" and Decree of the President of Ukraine No. 306/2021, which introduces this decision in effect. In these documents, germanium ores are included in the list of strategic importance for the sustainable development and defense capability of the state.

For an objective geological and economic assessment of the possibility of simultaneous extraction of germanium from coal, waste and products of its processing and planning of the most effective organizational and technical measures in this regard, it is first of all necessary to have information about the nature of distribution and the level of concentration of this element in the coal. In order to obtain such information, the authors performed detailed studies of the distribution of germanium over the area and in the section of the coal seam c_6 of the Dniprovska Mine field.

Analysis of previous studies. Previously, the peculiarities of the distribution of "elements–impurities" in the coal seams of some mines and geological and industrial areas of Donbas were investigated [7–29]. The methods of natural typification of coal deposits by the content of accompanying elements and oil deposits of the Dnipro-Donetsk basin by the content of metals were substantiated [30]. At the same time, the analysis of the results of the clustering of coal seam sections of different strength by germanium concentrations in coal seam c_6 of the Dniprovska Mine field has not been performed before.

The purpose of the research is to establish the most effective method of creating objective typification of sections of the coal seam c_6 of the Dniprovska Mine of different capacity according to germanium concentrations based on the analysis of the clustering results.

2. Methods

The factual basis of the work was the results of 347 germanium analyzes [31] performed after 1981 in the central certified laboratories of production geological exploration organizations of Ukraine from the material of reservoir samples obtained by production and scientific research enterprises and organizations and measurements of reservoir thickness. In a number of cases, they were supplemented with analyzes of reservoir samples taken by the furrow method from duplicate cores and mine workings [32] with the participation of the authors and employees of the geological service of the coal mining enterprise and production geological exploration organizations in the period from 1981 to 2013.

Germanium content was determined by quantitative emission spectral analysis. Seven percent of duplicate samples were sent to internal laboratory control; 10% of duplicate samples were subjected to external laboratory control. The quality of the results of the analyzes (correctness and reproducibility) was evaluated as the significance of the average systematic error, which is tested using the Student's criterion, and the significance of the average random error, which is tested using the Fisher criterion. Since the above-mentioned errors at the significance level of 0.95 are not significant, the quality of the analyzes is recognized as satisfactory.

With the help of statistical software platforms, at the initial stage of processing primary geochemical information, the values of the main descriptive statistical indicators were calculated, frequency histograms of the content were constructed and the distribution law of germanium was established. During the construction of graphs, all values of germanium concentrations and thickness of c_6 seam were normalized to

bring the sample to the same scale, regardless of the units of measurement and the scale of the samples.

Typification procedure is the systematization of objects according to a priori given features. Cluster analysis, taxonomy, pattern recognition, and factor analysis are usually used for this purpose.

It is important that, unlike other methods used in solving typification problems, cluster analysis does not require a priori assumptions about the data set, which does not impose restrictions on the presentation of the studied objects, allows analyzing natural indicators of various types of data (interval data, frequencies, binary data, etc.). The use of cluster analysis for the purpose of typification has a number of advantages, as it allows the division of a large number of studied objects and features into groups or clusters that are homogeneous in the relevant sense, as well as to reveal the internal structure at different hierarchical levels of the sample population. At the same time, like any other method, cluster analysis has certain disadvantages. In particular, the composition and number of clusters depends on the selected grouping criteria ("classification strategies"), and the application of different methods corresponding to different conceptual approaches to the selection of taxa to the same samples can lead to significantly different results. Thus, a characteristic feature of cluster analysis, unlike other methods of multivariate statistics, is the strong dependence of the obtained results on the a priori assumptions of the researcher at the substantive level. In our case, a priori assumptions include: lack of hypotheses regarding the number of clusters, their structure on forms; achieving maximum visualization of the breakdown of coal seam sections of different capacity by classes at different scale levels; establishment of a clustering method (algorithm) for the most stable division of the entire set of samples under consideration.

In the cluster analysis, it is considered that: a) the selected characteristics allow, in principle, the desired division into clusters; b) the units of measurement (their scale) are chosen correctly. Thus, the choice of scale in classification procedures plays a significant role. To bring it to the same scale, the raw data were normalized according to the algorithm given in works [33–36].

To achieve the goal set in the work, in the process of the research, clustering was carried out using various methods in different professional statistical software platforms. Their analysis was performed and the choice of the most optimal of them was substantiated.

3. Results and discussion

In order to develop a methodology for choosing the most effective method for creating objective (natural) typification of sections of the c_6 coal seam of different strength according to germanium concentrations, the general sample, taking into account the number of samples, was divided into 10 private samples, the main characteristics of which are shown in Table 1. To perform cluster analysis in the professional software platforms, a family of hierarchical agglomerative methods, two-input pooling and iterative divisive methods of averages is proposed.

Method of two-input association is relatively rarely used for simultaneous clustering of both observations and variables. In this case, it is expected that both observations and variables simultaneously contribute to the detection of clusters, which are further interpreted in geological terms. The main drawback of the method is problems with the conceptual interpretation of the results, which are a consequence of the fact that the distance between different clusters can be determined by differences in variables. The ambiguity of the interpretation of the results of the analysis does not allow for its optimal use.

Table 1 – Main characteristics of private samples of germanium concentrations

Private samples	1	2	3	4	5	6	7	8	9	10
Coal thickness interval, m	≤ 0.45	0.45 – 0.50	0.50 – 0.55	0.55 – 0.60	0.60 – 0.65	0.65 – 0.70	0.70 – 0.75	0.75 – 0.80	0.80 – 0.85	≥ 0.85
Number of samples	36	42	22	21	28	43	57	46	29	23
Normalized median values of Ge content	0.76	0.57	0.49	0.43	0.35	0.27	0.21	0.17	0.15	0.11

The use of the iterative divisive method of K-means for the optimal typification of sites by germanium content also has significant drawbacks. It is characterized by the problem of suboptimal solutions inherent in all iterative divisive methods, which consist in the initial breakdowns of sample populations. Its use implies the existence of a priori hypotheses regarding the number of clusters, and the result of clustering, presented in the form of a system of tables, does not allow visual visualization of the internal structure of the data.

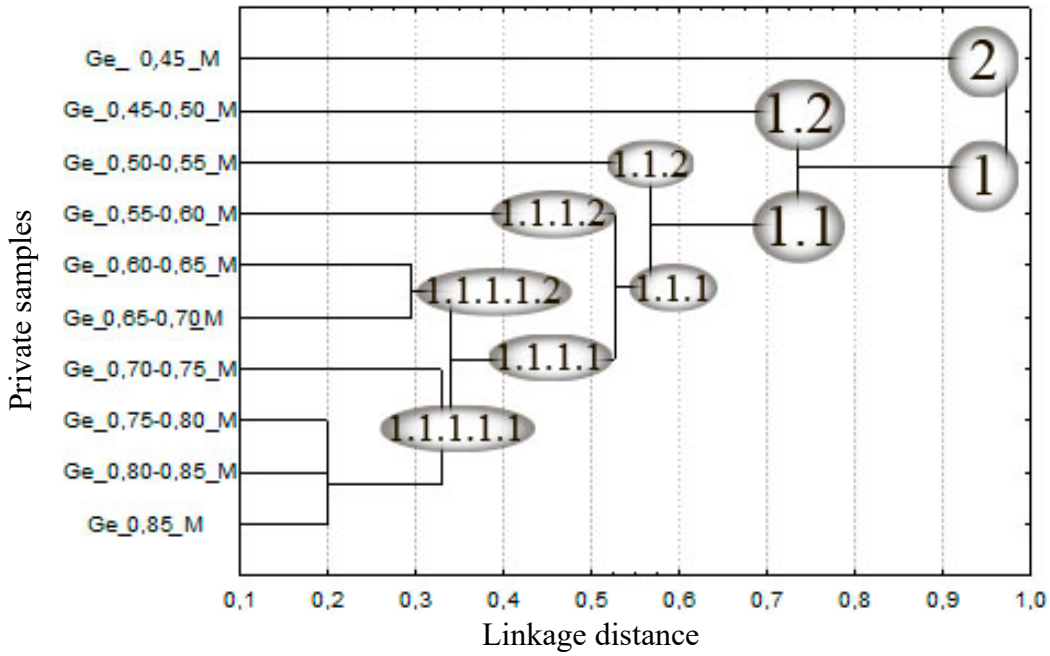
The family of hierarchical unifying methods implemented in the used programs refers to the cluster analysis procedures that are most often used. They all consist of a sequential combination of the most similar objects, which can be visualized in the form of a tree diagram - a dendrogram, which graphically displays the hierarchical structure of the object similarity matrix. This clarity of clustering results is a significant advantage of these methods. As a rule, the objects that are clustered are indicated horizontally in the dendrogram, and the value of the interclass distances at which they are combined (merging coefficient or the merging distance) is indicated vertically. At the same time, as a result of the analysis, non-overlapping groups of clusters are formed, and each cluster is an element of a wider cluster at a higher level of similarity. According to the method of grouping, all hierarchical agglomerative methods in the used programs are divided into the following: the method of single connection ("nearest neighbor"); full connection method ("farthest neighbor"); varieties of the "average link" method - the unweighted "average link" method ("unweighted pairwise average") and the weighted "average link" method; weighted centroid method and Ward's method. In addition, interclass distances can be used in all of the above methods - Euclidean distance (or its square), Manhattan distance ("distance of city blocks"), Chebyshev and Minkovsky metrics, linear correlation coefficient. With re-

gard to the features of the solved grouping problem, the most optimal is the use of the Euclidean distance as a measure of similarity.

Let us consider the possibilities of applying hierarchical agglomerative clustering methods to the construction of typification of c_6 reservoir sections by germanium content in more detail.

The single connection method forms clusters based on the principle of having at least one connection between objects. As a result of the operation of this method, clusters appear as long "chains" "linked together" only by individual elements that happened to be closer than others to each other. Despite the fact that its results are invariant to monotonic transformations of the similarity matrix and the use of the method does not limit the presence of "coincidence" in the data, its practical application for the purpose of typing causes certain difficulties.

On the example of the results of the clustering of the c6 argilla sections by germanium content (Fig. 1), it can be seen that the entire set of deposits as it approaches the end of the clustering process forms clusters 1 and 2, which in turn consist of the correspondingly nested subclusters 1.1, 1.2, 1.1.1 , 1.1.2, etc. Analysis of fig. 1 does not allow to determine the number and structure of clusters contained in the source data, and it also significantly complicates the establishment of the number and structure of clusters at other scale levels.

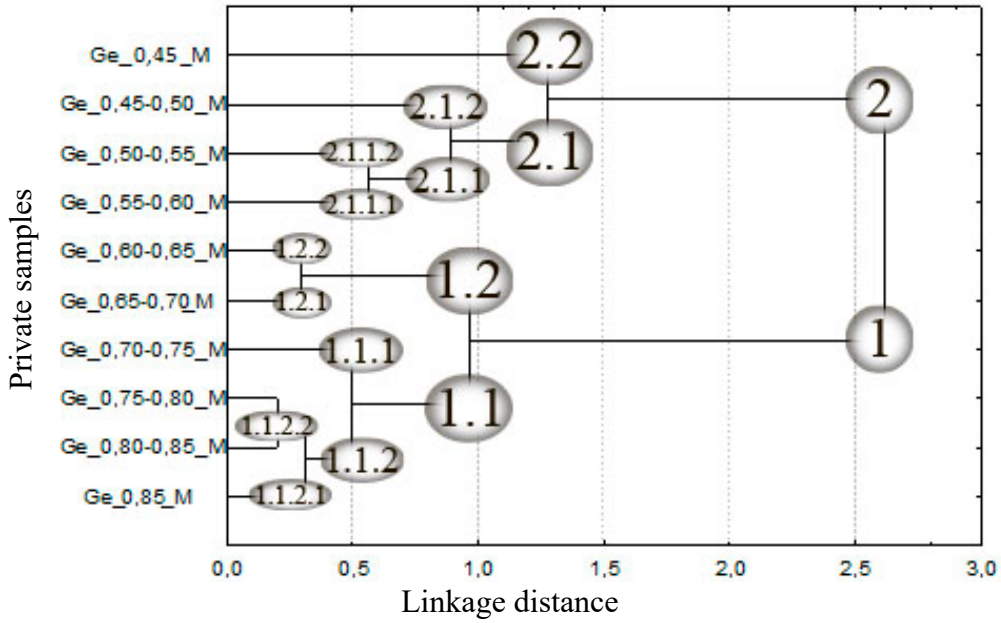


Legend: 1, 2, 1.1, 1.2, 1.1.1, 1.1.2, 1.1.1.1, 1.1.1.2, 1.1.1.1.1, 1.1.1.1.2 – clusters

Figure 1 – Dendrogram of the results of clustering by the method of single connection of zones by germanium content

The full connection method (Fig. 2), unlike the method discussed above, imposes stricter requirements for combining objects into one cluster. In this case, there is a tendency to detect relatively compact hyperspherical (in multidimensional space) clusters combining similar objects. Here, the distances between clusters are deter-

mined by the largest distance between any two objects in different clusters (i.e. the most distant neighbors).

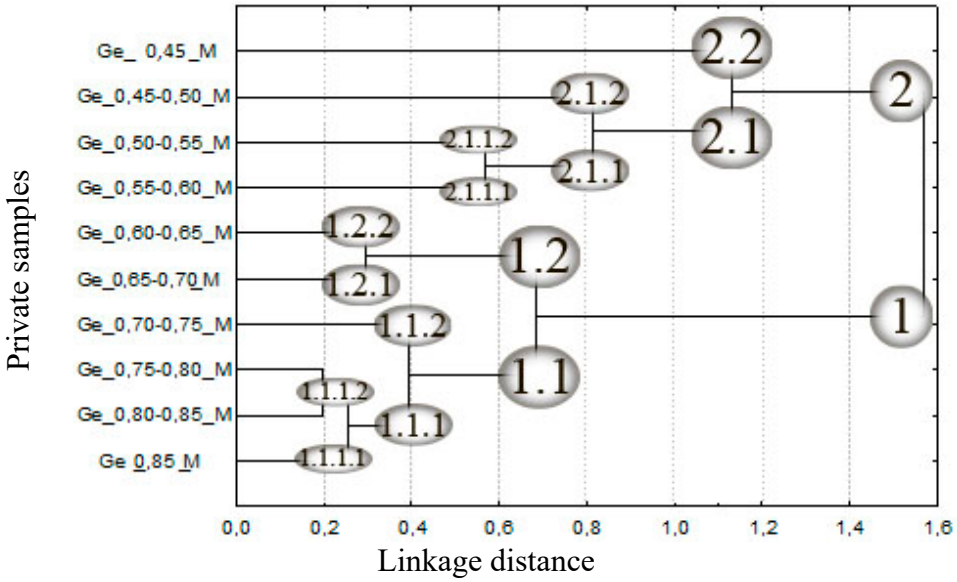


Legend: 1, 2, 1.1, 1.2, 1.1.1, 1.1.2, 1.1.2.1, 1.1.2.2, 1.2.1., 1.2.2, 2.1, 2.2, 2.1.1, 2.1.2, 2.1.1.2, 2.1.1.1 – clusters

Figure 2 – Dendrogram of the results of clustering by the method of complete connection of zones by germanium content

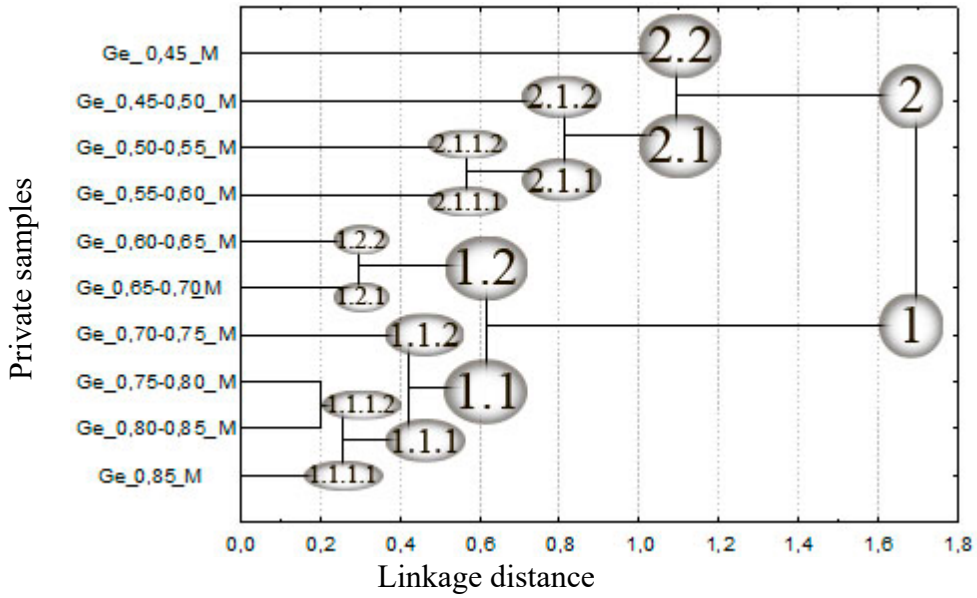
Comparison of fig. 1 and fig. 2 allows us to reveal a number of advantages of clustering using the method of complete connections. At the same time, if the dendrogram obtained during clustering by the method of single linkage quite convincingly indicates the number and structure of clusters (Fig. 1), then when clustering by the method of complete linkages, there is some change in the areas between clusters and, in general, when the compared results of both methods of identifying the final structure and number of deposits in clusters (Fig. 1, Fig. 2) are not so obvious. In addition, in both cases, the use of only the dendrogram without the involvement of primary data makes it difficult to assign individual objects to one or another cluster, and the structure of the clusters itself is rather poorly revealed.

The "average connection" method was developed by Sokel and Minchener in 1958 as a compromise between the single and full connection methods. In version STATISTICA 13.3 and IBM SPSS Statistics 22, two varieties of the method are implemented: the unweighted "linkage average" method ("unweighted pairwise average") and the weighted "linkage average" method. In the first variant of the method, the distance between two clusters is calculated as the average distance between all pairs of objects in them, and in the second, in addition, the size of the clusters (i.e. the number of objects contained in them) is used as a weighting factor. Using the number of objects contained in a cluster as a weighting factor assumes "good quality" of the analysis in the presence of clusters of unequal size in the sample. The results of the clustering of reservoir sections by germanium content by both methods in the form of dendrograms are shown in fig. 3 and fig. 4.



Legend: 1, 2, 1.1, 1.2, 1.1.1, 1.1.2, 1.1.1.1, 1.1.1.2, 1.2.1., 1.2.2, 2.1, 2.2, 2.1.1, 2.1.2, 2.1.1.2, 2.1.1.1 – clusters

Figure 3 – Dendrogram of the results of clustering by the method of unweighted average connection of zones by germanium content

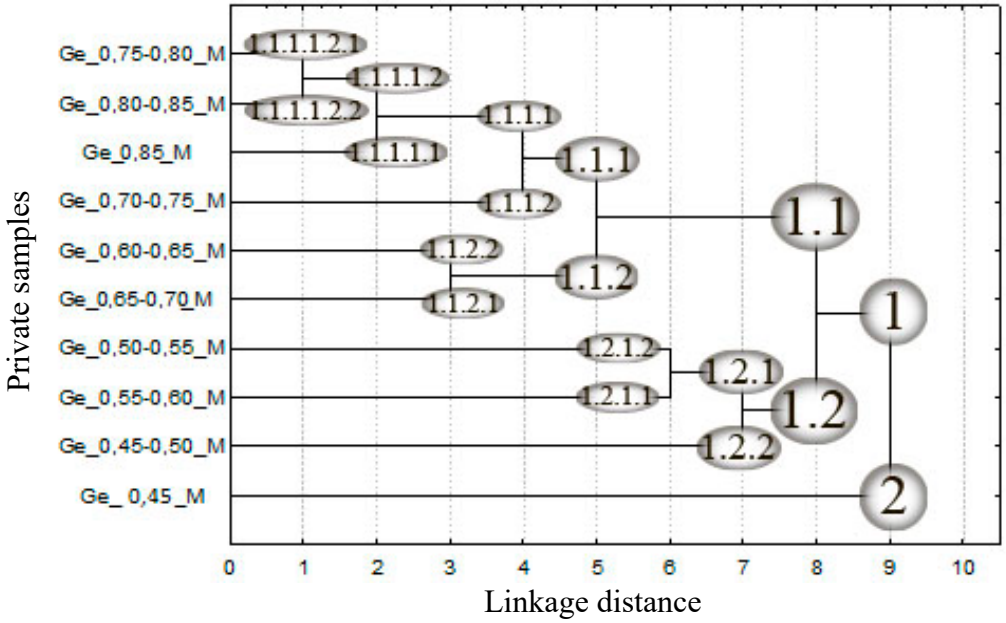


Legend: 1, 2, 1.1, 1.2, 1.1.1, 1.1.2, 1.1.1.1, 1.1.1.2, 1.2.1, 1.2.2, 2.1, 2.2, 2.1.1, 2.1.2, 2.1.1.2, 2.1.1.1 – clusters

Figure 4 – Dendrogram of clustering results using the method of weighted average connection of zones by germanium content

The analysis of the given dendrograms allows us to come to the conclusion that in the case of the division of reservoir sections by germanium content, the results of the cluster analysis, respectively, by the methods of unweighted average connection and weighted average connection, in our case, completely coincide.

The weighted centroid method uses the distance between the clusters centers of gravity as the distance between them (objects). In fig. 5, a dendrogram of the results of clustering by the weighted centroid median method of sections of the c_6 seam by germanium content is shown. The analysis of this dendrogram, in contrast to the ones presented earlier (Fig. 1, Fig. 2, Fig. 3 and Fig. 4), allows not only to achieve the most stable division of the entire set of areas under consideration, but also to maximally visualize their division by classes on different scales levels in the absence of a priori hypotheses regarding the number of clusters and their shape. At the same time, the structure of clusters is clearly distinguished, regardless of the scale level of their formation, and the sequence of unification of individual deposits and their groups into the resulting cluster is clearly traced. These advantages make it possible to make maximum use of already existing information for the development of natural typifications of areas of the c_6 coal seam by germanium content and to interpret the obtained results in geological terms.



Legend: 1, 2, 1.1, 1.2, 1.1.1, 1.1.2, 1.1.1.1, 1.1.1.2, 1.2.1, 1.2.2, 1.2.1.1, 1.1.2.1, 1.1.2.2, 1.1.1.1.1, 1.1.1.1.2, 1.1.1.1.2.1, 1.1.1.1.2.2 – clusters

Figure 5 – Dendrogram of the results of clustering by weighted centroid median method of zones by germanium content

4. Conclusions

As a result of the conducted research, it was established that the use of the weighted centroid median method of cluster analysis, which is implemented in the professional statistical software platforms, is the most optimal for a researcher with no subjectivity to divide the sections of the coal seam c_6 of the "Dniprovska" Mine field by germanium content on taxa. The constructed dendrograms of clustering of c_6 coal seam sections by germanium content can be used as a basis for the development of natural typification of coal seams of the Dniprovska Mine for their subsequent

geological and economic assessment. This, in turn, will make it possible to make maximum use of already available information and to interpret the obtained results in geological and genetic concepts, which will provide the opportunity to use it for the comprehensive use of mineral raw materials and to solve strategic issues of sustainable development of Ukraine.

REFERENCES

1. Naumov, A.V. (2007), "World market of germanium and its prospects", *Russian Journal of Non-Ferrous Metals*, vol. 48, no. 4, pp. 265–272. https://doi.org/10.3103/S1067821207040049

2. Pavlov, A.V. (1966), "Material composition of coal ash in some areas of Western Spitsbergen", *Uch. zap. NIIGA. Regionaln. geol*., no. 8, pp. 128–136.

3. Shpirt, M.Ya., Nukenov, D.N., Punanova, S.A., Visaliev, M.Ya. (2013), "Principles for obtaining compounds of valuable metals from fossil fuels", *Himiya tverdogo topliva*, no. 2, pp. 3–8. https://doi.org/10.3103/S0361521913020110

4. Polyakov, N.P. (1973), "Patterns of the distribution of trace elements in the coals of the Aldan-Chulmansky district", *Zap. Zabaykale. Filial Geograficheskogo obschestva SSSR,* Vyip. 81, pp. 55–66.

5. Imai, N., Ando, A., Takeda, E. (1984), Minor elements in Japanese coal. *Bull. Surv. Jpan*, vol. 35, no. 7, P. 287–314.

6. Zodrow, E.L. (1986), Coal-stratigraphic geochemistry: trends in coal samples from Sydney Coalfield, Upper Carboniferous, Nova Scotia, *Can. Inst. Min. Met. Bull*., vol. 79, no. 893, pp. 83–85.

7. Koziy, E.S. (2018), "Arsenic, beryllium, fluorine and mercury in the coal of the layer $c_8^в$ of the "Dniprovska" mine of Pavlogradsko-Petropavlovskiy geological and industrial district", *Visnyk Dnipropetrovskoho universytetu. Heolohiia-Heohrafiia*, no. 26(1), pp. 113–120. https://doi.org/10.15421/111812

8. Koziy E.S. (2017), "Peculiarities of distribution of toxic and potentially toxic elements in the coal of the layer $c_{10}^в$ in the Stashkov mine of Pavlograd-Petropavlovsk geological and industrial district", *Geotekhnichna Mekhanika [Geo-Technical Mechanics]*, no. 132, pp. 157–172.

9. Ishkov V.V., Koziy E.S. (2017), "About peculiarities of distribution of toxic and potentially toxic elements in the coal of the layer $c_{10}^в$ of the Dneprovskaya mine of Pavlogradsko-Petropavlovskiy geological and industrial district of Donbass", *Geotekhnichna Mekhanika [Geo-Technical Mechanics]*, no. 133, pp. 213–227.

10. Koziy, E.S., Ishkov, V.V. (2017), "Coal classification of main working seams of Pavlohrad-Petropavlivka geological and industrial district on content of toxic and potentially toxic elements", *Geotekhnichna Mekhanika [Geo-Technical Mechanics]*, no. 136, pp. 74–86.

11. Ishkov, V.V., Kozii, Ye.S. (2020), "Peculiarities of lead distribution in coal seams of Donetsk-Makiivka geological and industrial area of Donbas", *Tectonics and Stratigraphy*, no. 47, pp. 77–90. https://doi.org/10.30836/igs.0375-7773.2020.216155

12. Kozii, Ye.S. (2021), "Toxic elements in the c_1 coal seam of the Blahodatna mine of Pavlohrad-Petropavlivka geological and industrial area of Donbas", *Geo-Technical Mechanics*, no. 158, pp. 103–116. https://doi.org/10.15407/geotm2021.158.103

13. Kozii, Ye.S. (2021), "Arsenic, mercury, fluorine and beryllium in the c_1 coal seam of the Blahodatna mine of Pavlohrad-Petropavlivka geological and industrial area of western Donbas", *Geotekhnichna Mekhanika* [Geo-Technical Mechanics], no 159, pp. 58–68. https://doi.org/10.15407/geotm2021.159.058

14. Ishkov, V.V., Kozar, M.A., Kozii, Ye.S., Bartashevskiy, S.Ye. (2022). Nickel in oil deposits of the Dnipro-Donetsk depression (Ukraine). Problems of science and practice, tasks and ways to solve them. *Proceedings of the XXVI International Scientific and Practical Conference. Helsinki, Finland*, pp. 25–26. https://doi.org/10.46299/ISG.2022.1.26

15. Kozar, M.A., Ishkov, V.V., Kozii, E.S., Strielnyk, Yu.V. (2021), "Toxic elements of mineral and organic composition of lower carbon coal Western Donbas", *Abstracts of Scientific Conference "Geological science in independent Ukraine"*, NAS of Ukraine, M.P. Semenenko Institute of Geochemistry, Mineralogy and Ore Formation, Kyiv, pp. 55–58.

16. Kozii, Ye. (2020), Coal height of coal seam k_5 of "Kapitalna" mine, *Proceedings of the International Forum for Students and Young Researchers "Widening Our Horizons",* pp. 399–401.

17. Ishkov, V., Kozii, Ye. (2020), "Distribution of mercury in coal seam $c_7^н$ of Pavlohradska mine field", *Scientific Papers of DONNTU Series: "The Mining and Geology"*, no. 1(23)–2(24), pp. 26–33. https://doi.org/10.31474/2073-9575-2020-3(23)-4(24)-26-33.

18. Kozii, Ye. (2021), Chromium in the coal seams of the Chervonoarmiiskyi geological and industrial area of Donbas, *Proceedings of the International Forum for Students and Young Researchers "Widening Our Horizons",* pp. 453–455.

19. Ishkov, V.V., Kozii, Ye.S. (2021), "Distribution of arsene and mercury in the coal seam k_5 of the Kapitalna mine, Donbas", *Mineralogical Journal,* no. 43(4), pp. 73–86. https://doi.org/10.15407/mineraljournal.43.04.073

20. Ishkov, V.V., Kozii, Ye.S. (2020), "Some features of beryllium distribution in the k_5 coal seam of the "Kapitalna" mine of the Krasnoarmiiskyi geological and industrial district of Donbas", *Odesa National University Herald. Geography and Geology*, vol. 25, no. 1(36), pp. 214–227. https://doi.org/10.18524/2303-9914.2020.1(36).205180

21. Mametova, L.F., Mirek, A., Kozii, Ye.S. (2020), "Pyritization of the Middle Carboniferous Sandstones of the Donbas". *Mineral. Journ. (Ukraine)*, vol. 42, no. 2, pp. 14–19. https://doi.org/10.15407/mineraljournal.42.02.014

22. Ishkov, V.V., Koziy, E.S. (2021), "Accumulation of Co and Mn on the example of the c_5 layer of the Western Donbass as a result of their migration from the weathering crusts of the Ukrainian crystalline shield", *Materialyi Mezhdunarodnogo soveschaniya po*

geologii rossyipey i mestorozhdeniy kor vyivetrivaniya «Rossyipi i mestorozhdeniya kor vyivetrivaniya XXI veka: zadachi, problemyi, resheniya». pp. 160–162.

23. Nesterovskyi, V.A, Ishkov, V.V, Kozii Ye.S. (2020), "Toxic and potentially toxic elements in the coal of the seam $c_{8н}$ of the "Blagodatna" mine of Pavlohrad-Petropavlivka geological and industrial area", *Visnyk Of Taras Shevchenko National University Of Kyiv: Geology*, no. 88(1), pp. 17–24. http://doi.org/10.17721/1728-2713.88.03

24. Koziy, E.S., Ishkov, V.V. (2017), "Coal classification of main working seams of Pavlohrad-Petropavlivka geological and industrial district on content of toxic and potentially toxic elements", *Geotekhnichna Mekhanika [Geo-Technical Mechanics]*, no. 136, pp. 74–86.

25. Ishkov, V.V., Koziy, E.S. (2013), "New data about distribution of toxic and potentially toxic elements in the coal seam c6n of the Ternovskaya mine of Pavlograd-Petropavlovsky geological and industrial area", *Zbirnyk naukovykh prats NHU*, no. 41, pp. 201–208.

26. Ishkov, V.V., Koziy, E.S. (2014), "About distribution of ash, sulfur, manganese in the coal seam c4 of the Samarskaya mine of Pavlograd - Petropavlovsk geological and industrial area", *Zbirnyk naukovykh prats NHU*. no. 44, pp. 178–186.

27. Kozar, M.A., Ishkov, V.V., Kozii, Ye.S., Pashchenko P.S. (2020), "New data about the distribution of nickel, lead and chromium in the coal seams of the Donetsk-Makiivka geological and industrial district of the Donbas", *Journ. Geol. Geograph. Geoecology*, no. 29(4), pp. 722–730. https://doi.org/10.15421/112065

28. Ishkov, V.V., Nagornyi, V.N. (2005), "About regularities of mercury accumulation in coal seams of the Krasnoarmeysky geological and industrial region", *Naukovyi visnyk Natsionalnoi hirnychoi akademii Ukrainy*, no. 2, pp. 84–88.

29. Ishkov V.V., Kozii Ye.S. (2019), "Analysis of the distribution of chrome and mercury in the main coals of the Krasnoarmiiskyi geological and industrial area", *Tectonics and Stratigraphy*, no. 46, pp. 96–104. https://doi.org/10.30836/igs.0375-7773.2019.208881

30. Yerofieiev, A.M., Ishkov, V.V., Kozii, Ye.S., Bartashevskiy, S.Ye. (2021), "Research of clusterization methods of oil deposits in the Dnipro-Donetsk depression with the purpose of creating their classification by metal content (on the vanadium example)", *Scientific Papers of Donntu Series: "The Mining and Geology"*, no. 1-25-2-26, pp. 83–93. https://doi.org/10.31474/2073-9575-2021-1(25)-2(26)-83-93

31. All-Union State Standard (1975), 10175-75: Ugli buryie, kamennyie, antratsityi, argillityi i alevrolityi. Metodyi opredeleniya soderzhaniya germaniya. M: Izd-vo standartov. 14 p.

32. All-Union State Standard (1975), 9815-75: Ugli buryie, kamennyie, antratsit i goryuchie slantsyi. Metod otbora plastovyih prob. M: Izd-vo standartov. 6 p.

33. Borovikov, V.P. (2001), *STATISTICA: iskusstvo analiza dannyih na kompyutere* [STATISTICA: the art of data analysis on a computer. For professionals]. Sankt-Peterburg, Russia.

34. Ishkov V.V., Kozii Ye.S., Kozar M.A. (2022), "Features of vanadium geochemistry in oils from the oil and gas fields of Eastern region of Ukraine", *Collection of scientific works "Geotechnical Mechanics",* no.162, pp 85–96. https://doi.org/10.15407/geotm2022.162.085

35. Ishkov V.V., Kozii Ye.S., Chernobuk O.I., Pashchenko P.S. (2022), "The relationship of germanium concentrations and the thickness of the $c_{8н}$ coal seam of the Dniprovska coal mine", *Collection of scientific works "Geotechnical Mechanics"*, no.162, pp. 164–176. https://doi.org/10.15407/geotm2022.162.164

36. Ishkov, V.V., Kozii, Ye.S. (2014), "About classification of coal seams by the content of toxic elements using cluster analysis", *Zbirnyk naukovykh prats Natsionalnoho hirnychoho universytetu*, no. 45, pp. 209 –221.

About authors

Ishkov Valerii Valeriiovych, Candidate of Geological and Mineralogical Sciences (Ph.D.), Senior Research in Laboratory of Studies of Structural Changes Rocks, Institute of Geotechnical Mechanics named by N. Poljakov of the National Academy of Sciences of Ukraine (IGTM NAS of Ukraine), Dnipro, Ukraine, ishwishw37@gmail.com

Kozii Yevhen Serhiiovych, Candidate of Geological Sciences (Ph.D.), Director of Educational and Scientific Center for Training of Foreign Citizens, Dnipro University of Technology, Dnipro, Ukraine, Associate Professor of the Department of Civil Engineering, Construction Technologies and Environmental Protection, Dnipro State Agrarian and Economic University, Dnipro, Ukraine, koziy.es@gmail.com

МЕТОДИКА КЛАСТЕРИЗАЦІЇ ЗОН ВУГІЛЬНОГО ПЛАСТА РІЗНОЇ ПОТУЖНОСТІ C_6 ПОЛЯ ШАХТИ «ДНІПРОВСЬКА» ЗА КОНЦЕНТРАЦІЄЮ ГЕРМАНІЮ

Ішков В.В., Козій Є.С.

Анотація. У статті, на основі проведених досліджень і аналізу результатів кластеризації було встановлено найефективніший метод створення об'єктивної типізації зон вугільного пласта c_6 шахти «Дніпровська» різної потужності за концентраціями германію. Процедура типізації представляє собою систематизацію об'єктів по апріорно заданим ознакам. Зазвичай для цього використовуються кластерний аналіз, таксономія, розпізнавання образів, факторний аналіз. Автори статті здійснили кластеризацію різними методами, які реалізовані у найпопулярніших професійних статистичних програмних платфор-

мах «STATISTICA» та «SPSS», виконали їх аналіз та обґрунтували вибір найбільш оптимального з них. У процесі досліджень використовувалися версії програм Excel 2016, STATISTICA 13.3 та IBM SPSS Statistics 22. Був визначений оптимальний метод кластеризації ділянок різної потужності вугільного пласта. Аналіз дендрограми результатів кластеризації зваженим центроїдним медіанним методом ділянок пласта с6 за вмістом германію на відміну від інших дозволяє не тільки досягти найбільш стійкого поділу всієї сукупності ділянок, що розглядаються, але й максимальної візуалізації їх розбиття по класам на різних масштабних рівнях при апріорній відсутності гіпотез щодо числа кластерів та їх форми. При цьому чітко виділяється структура кластерів незалежно від масштабного рівня їх формування, чітко простежується послідовність об'єднання окремих родовищ та їх груп у результуючий кластер. Ці переваги дозволяють максимально використовувати вже існуючу інформацію для розробки природних типізацій ділянок вугільного пласта c_6 за вмістом германію та інтерпретувати отримані результати у геологічних поняттях. Було встановлено, що зважений центроїдний медіанний метод кластерного аналізу є найбільш оптимальним для вільного від суб'єктивізму дослідника поділу ділянок вугільного пласта c_6 шахти «Дніпровська» за вмістом германію на таксони. Побудовані дендрограми кластеризації родовищ за вмістом германію можуть бути використані як основа для розробки природної типізації вугільних пластів поля шахти Дніпровська для їх наступної геолого-економічної оцінки. Це в свою чергу дозволить максимально використовувати вже наявну інформацію та інтерпретувати отримані результати у геолого-генетичних поняттях, що надасть можливість її використання для комплексного використання мінеральної сировини та вирішення стратегічних питань сталого розвитку України.

Ключові слова: кластерний аналіз, вугільний пласт, германій, кластери, таксони, дендрограма результатів кластеризації, зважений центроїдний метод.