

В.В. Ішков
Є.С. Козій
О.І. Чернобук
В.Л. Хоменко

РЕЗУЛЬТАТИ КЛАСТЕРИЗАЦІЇ ДІЛЯНОК РІЗНОЇ ПОТУЖНОСТІ ВУГІЛЬНОГО ПЛАСТА C_{10}^B ШАХТИ «ДНІПРОВСЬКА» ЗА ВМІСТОМ ГЕРМАНІЮ

Мета роботи – на основі аналізу результатів кластеризації встановити найефективніший метод створення об'єктивної типізації ділянок вугільного пласта C_{10}^B шахти «Дніпровська» різної потужності за концентраціями германію.

Методика роботи. Як відомо, процедура типізації – це систематизація об'єктів по апріорно заданим ознакам. Зазвичай для цього використовуються кластерний аналіз, таксономія, розпізнавання образів, факторний аналіз. Для досягнення поставленої в роботі мети у процесі досліджень було здійснено кластеризацію різними методами, які реалізовані у найпопулярніших професійних статистичних програмних платформах «STATISTICA» та «SPSS»; виконано їх аналіз та обґрунтовано вибір найбільш оптимального з них. У роботі використовувалися версії програм Excel 2016, STATISTICA 13.3 та IBM SPSS Statistics 22.

Основні результати досліджень полягають у виборі оптимального методу кластеризації ділянок різної потужності вугільного пласта. Аналіз дендрограми результатів кластеризації зваженим центроїдним медіанним методом ділянок пласта C_{10}^B 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 масштабового 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_{10}^B за вмістом германію та інтерпретувати отримані результати у геологічних поняттях.

Наукова новизна результатів виконаних досліджень полягає у встановленні зваженого центроїдного медіанного методу кластерного аналізу, який запроваджено у професійних статистичних програмних платформах «STATISTICA» та «SPSS» як найбільш оптимального для вільного від суб'єктивізму дослідника поділу ділянок вугільного пласта C_{10}^B шахти «Дніпровська» за вмістом германію на таксони.

Практична значимість результатів роботи полягає у тому що побудовані дендрограми кластеризації родовищ за вмістом германію можуть бути використані як основа для розробки природної типізації вугільних пластів 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: германій, кластерний аналіз, вугільний пласт, таксони, кластери, зважений центроїдний метод, дендрограма результатів кластеризації.

Вступ.

Актуальність дослідження вмісту германію у вугільних пластах обумовлена можливістю його промислового вилучення та використання в якості цінного попутного компонента [1 – 3]. Зараз вугілля є основним оціненим джерелом Ge в Україні, Китаї, Узбекистані, а також в Росії, а Ge-вугільні родовища крім того розробляються в Англії, Канаді, США та інших країнах [4 – 6].

Особливу актуальність виконаним дослідженням надає рішення Ради національної безпеки та оборони України від 16 липня 2021 року «Про стимулювання пошуку, видобутку та збагачення корисних копалин, які мають стратегічне значення для

сталого розвитку та обороноздатності держави» та Указ Президента України №306/2021, який вводить в дію це рішення. В цих документах руди Ge включені до переліку, що мають стратегічне значення для сталого розвитку та обороноздатності держави.

Для об'єктивної геолого-економічної оцінки можливості попутного вилучення германію з вугілля, відходів і продуктів його переробки та планування найбільш ефективних організаційно-технічних заходів з цього приводу, перш за все необхідно мати відомості про характер розподілу і рівень концентрації цього елемента у вугіллі. З метою одержання такої інформації авторами

були виконані детальні дослідження розподілу германія по площі і в розрізі вугільного пласта c_{10}^B поля шахти «Дніпровська».

Останні досягнення.

Раніше були досліджені особливості розподілу елементів - домішок у вугільних пластах деяких шахт та геолого-промислових районів Донбасу [7–12]. Були обґрунтовані методи природної типізації вугільних родовищ за вмістом супутніх елементів та родовищ нафти Дніпровсько-Донецької западини за вмістом металів [13]. У той же час, аналіз результатів кластеризації ділянок вугільного пласта різної потужності за концентраціями Ge у вугільному пласті c_{10}^B шахти «Дніпровська» раніше не виконувався.

Мета досліджень – на основі аналізу результатів кластеризації встановити найефективніший метод створення об'єктивної типізації ділянок вугільного пласта c_{10}^B шахти «Дніпровська» різної потужності за концентраціями Ge.

Методи дослідження.

Фактологічною основою роботи були результати 347 аналізів Ge [14], виконаних після 1981р. в центральних сертифікованих лабораторіях виробничих геологорозвідувальних організацій України з матеріалу пластових проб отриманих виробничими і науково-дослідницькими підприємствами і організаціями та вимірювань потужності пласту. У ряді випадків вони доповнювались аналізами пластових проб відібраних борозновим методом із дублікатів керна і гірничих виробках [15] за участю авторів та співробітників геологічної служби вугледобувного підприємства і виробничих геологорозвідувальних організацій в період з 1981 по 2013 рік.

Вміст Ge визначався кількісним емісійним спектральним аналізом. На внутрішній лабораторний контроль направлено 7% дублікатів проб. Зовнішньому лабораторному контролю піддано 10% дублікатів проб. Якість результатів аналізів (правильність і відтворюваність) оцінювалася як значимість середньої систематичної похибки, яка перевіряється за допомогою критерію

Стьюдента і значимість середньої випадкової похибки, яка перевіряється за допомогою критерію Фішера. Оскільки вказані вище похибки при рівні значимості 0,95 є не значимими, якість аналізів визнано задовільною.

За допомогою програм Excel 2016 і Statistica 13.3 на початковому етапі обробки первинної геохімічної інформації розраховувалися значення основних описових статистичних показників, виконувалась побудова частотних гістограм вмісту і встановлення закону розподілу германію. В ході побудови графіків всі значення концентрацій Ge та потужності пласта c_{10}^B нормувались для приведення вибірки до одного масштабу незалежно від одиниць виміру та розмаху вибірок.

Як відомо, процедура типізації – це систематизація об'єктів по апріорно заданим ознакам. Зазвичай для цього використовуються кластерний аналіз, таксономія, розпізнавання образів, факторний аналіз.

Важливо що на відміну від інших методів, які використовують при вирішенні задач типізації, кластерний аналіз не вимагає апріорних припущень про набір даних, що не накладає обмеження на подання досліджуваних об'єктів, дозволяє аналізувати природні показники різних типів даних (інтервальних даних, частот, бінарних даних, тощо). Використання кластерного аналізу з метою типізації має ряд переваг, оскільки дозволяє виконати розбиття безлічі досліджуваних об'єктів і ознак на однорідні у відповідному розумінні групи або кластери, а також виявити внутрішню структуру на різних ієрархічних рівнях вибіркової сукупності. У той же час, як і будь-який інший метод, кластерний аналіз має певні недоліки. Зокрема, склад і кількість кластерів залежить від обраних критеріїв угруповання («стратегії класифікації»), а застосування різних методів, що відповідають різним концептуальним підходам до виділення таксонів, до одних і тих же вибірок, може призвести до суттєво відмінних результатів. Таким чином, характерною особливістю кластерного аналізу, на відміну від інших методів багатомірної статистики, є сильна залежність одержуваних результатів від апріорних установок дослідника на змістовному рівні. У нашому випадку, до

апріорних установок відносяться: відсутність гіпотез щодо числа кластерів, їх структури на форми; досягнення максимальної візуалізації розбиття ділянок вугільного пласта різної потужності по класам на різних масштабних рівнях; встановлення методу (алгоритму) кластеризації для найбільш стійкого поділу всієї сукупності проб, що розглядаються.

У кластерному аналізі вважається, що: а) вибрані характеристики допускають в принципі бажане розбиття на кластери; б) одиниці виміру (їх масштаб) обрані правильно. Таким чином, вибір масштабу у процедурах класифікації відіграє значну роль. Для приведення до одного масштабу вихідні дані були нормовані за алгоритмом, що наведено у роботах [16-18].

Для досягнення поставленої в роботі мети у процесі досліджень було здійснено кластеризацію різними методами, які реалізовані у найпопулярніших професійних статистичних програмних платформах «STATISTICA» та «SPSS»; виконано їх аналіз та обґрунтовано вибір найбільш оптимального з них.

У роботі використовувалися версії програм Excel 2016, STATISTICA 13.3 та IBM SPSS Statistics 22.

Результати дослідження та їх обговорення.

Для розробки методики вибору найефективнішого методу створення об'єктивної (природної) типізації ділянок вугільного пласта c_{10}^B різної потужності за концентраціями Ge загальна вибірка із врахуванням кількості проб була поділена на 10 приватних вибірок, основні характеристики яких наведені у таблиці 1.

Для виконання кластерного аналізу в професійних програмних платформах

«STATISTICA» та «SPSS» пропонується сімейство ієрархічних агломеративних методів, двовходового об'єднання та ітеративний дивізімний метод середніх.

Метод двовходового об'єднання відносно рідко використовується для одночасної кластеризації як спостережень так і змінних. У цьому випадку очікується, що і спостереження і змінні одночасно роблять внесок у виявлення кластерів, які далі інтерпретуються в геологічних поняттях. Головним недоліком методу є проблеми з понятійною інтерпретацією результатів, які є наслідком того, що відстань між різними кластерами може визначатися відмінностями у змінних. Неоднозначність інтерпретації результатів аналізу не дає можливості для його використання як оптимального.

Для нього характерна властива всім ітеративним дивізімним методам проблема субоптимальних рішень, які полягають у вихідних розбиттях вибіркової сукупностей. Його використання має на увазі існування апріорних гіпотез щодо числа кластерів, а результат кластеризації, представлений у вигляді системи таблиць, не дозволяє наочно візуалізувати внутрішню структуру даних.

Сімейство ієрархічних об'єднуючих методів реалізованих у використаних програмах відноситься до процедур кластерного аналізу, які найчастіше застосовуються. Вони всі полягають у послідовному об'єднанні найбільш схожих об'єктів, яке можна візуалізувати у вигляді деревоподібної діаграми – дендрограми, яка графічно відображає ієрархічну структуру матриці подібності об'єктів. Така наочність результатів кластеризації є суттєвою перевагою цих методів.

Таблиця 1. Основні характеристики приватних вибірок концентрацій германію

Приватні вибірки	1	2	3	4	5	6	7	8	9	10
Інтервал потужності, м	$\leq 0,45$	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	$\geq 0,85$
Кількість проб	36	42	22	21	28	43	57	46	29	23
Нормовані медіанні значення вмісту Ge	0,76	0,57	0,49	0,43	0,35	0,27	0,21	0,17	0,15	0,11

Як правило, в дендрограмі по горизонталі вказуються об'єкти, що кластеризуються, а по вертикалі - значення міжкласових відстаней, при яких відбувається їх об'єднання (коєфіцієнт злиття або дистанція об'єднання). При цьому в результаті аналізу формуються групи кластерів, що не перекриваються, причому кожен кластер є елементом ширшого кластера на більш високому рівні подібності. За способом угруповання всі ієрархічні агломеративні методи у використаних програмах поділяються на наступні: метод одиночного зв'язку («найближчого сусіда»); метод повного зв'язку («найвіддаленішого сусіда»); різновиди методу «середнього зв'язку» – незважений метод «середнього зв'язку» («незважене попарне середнє») та зважений метод «середнього зв'язку»; зважений центроїдний метод і метод Уорда. Крім того, у всіх перерахованих методах можуть бути використані міжкласові відстані – евклідова відстань (або її квадрат), манхеттенська відстань («відстань міських кварталів»), метрики Чебишева і Мінковського, лінійний коєфіцієнт кореляції. Стосовно особливостей вирішуваної задачі угруповання найбільш оптимальним є використання як міри подібності евклідової відстані.

Розглянемо можливості застосування ієрархічних агломеративних методів кластеризації до побудови типізації ділянок пласта c_{10}^B за вмістом Ge більш докладно.

Метод одиночного зв'язку формує кластери із принципу наявності хоча б одного зв'язку між об'єктами. В результаті роботи цього методу кластери видаються довгими «ланцюжками» «зчепленими разом» лише окремими елементами, які випадково виявилися ближчими від інших один до одного. Незважаючи на те, що його результати інваріантні до монотонних перетворень матриці подібності і використання методу не обмежує присутність «збігу» в даних, практичне його застосування з метою типізації викликає певні труднощі.

На прикладі результатів кластеризації ділянок паста c_{10}^B за вмістом Ge (рис. 1) видно, що уся сукупність родовищ по мірі наближення до завершення процесу кластеризації формує кластери 1 та 2, які у свою чергу складаються з відповідно вкладених підкластерів 1.1, 1.2, 1.1.1, 1.1.2, тощо. Аналіз рис. 1 не дозволяє визначити кількість та структуру кластерів, що містяться у вихідних даних, а також суттєво ускладнюється встановлення кількості та структури кластерів на інших масштабних рівнях.

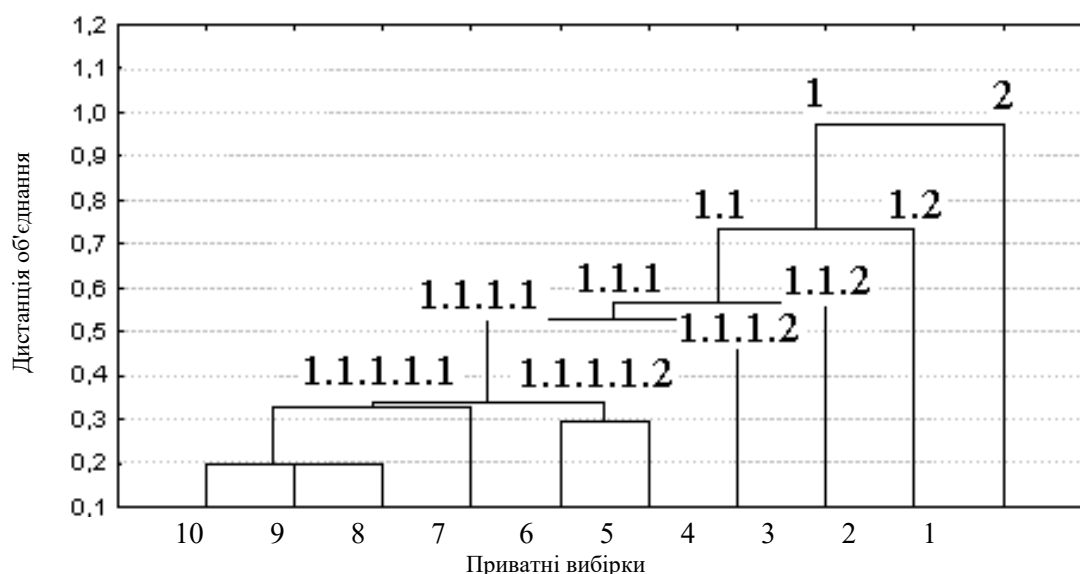


Рис. 1. Дендрограма результатів кластеризації методом одиночного зв'язку ділянок по вмісту Ge. Умовні позначення: 1, 2, 1.1, 1.2, 1.1.1, 1.1.2, 1.1.1.1, 1.1.1.2, 1.1.1.1.1, 1.1.1.1.2 – кластери

Метод повного зв'язку (рис. 2), на відміну від розглянутого вище методу накладає більш жорсткі вимоги до об'єднання об'єктів в один кластер. У цьому випадку з'являється тенденція до виявлення відносно компактних гіперсферичних (у

багатовимірному просторі) кластерів, що поєднують схожі об'єкти. Тут відстані між кластерами визначаються найбільшою відстанню між будь-якими двома об'єктами в різних кластерах (тобто найбільш віддаленими сусідами).

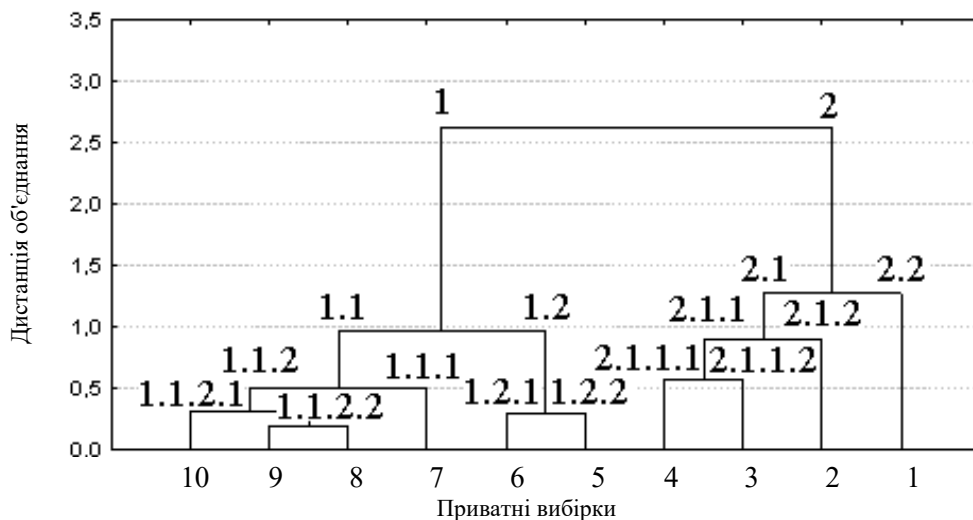


Рис. 2. Дендрограма результатів кластеризації методом повного зв'язку ділянок по вмісту Ge. Умовні позначення: 1, 2, 1.1, 1.2, 1.1.1, 1.1.2, 1.1.2.1, 1.1.2.2, 1.2.1, 1.2.2, 2.1, 2.2, 2.1.1, 2.1.2, 2.1.1.1, 2.1.1.2 – кластери

Зіставлення рис. 1 з рис. 2 дозволяє виявити ряд переваг кластеризації з використанням методу повних зв'язків. У той же час, якщо при кластеризації методом одиночного зв'язку отримана дендрограма досить переконливо вказує на кількість та структуру кластерів (рис. 1), то при кластеризації методом повних зв'язків спостерігається деяка зміна ділянок між кластерами і загалом, при порівнянні результатів обох методів виявлення остаточної структури та кількості родовищ у кластерах (рис. 1, рис. 2) не так очевидно. Крім того, в обох випадках використання тільки дендрограми без залучення первинних даних ускладнює віднесення окремих об'єктів до тих чи інших кластерів, а сама структура кластерів виявлена досить слабо.

Метод «середнього зв'язку» розроблений Сокелом та Мінченером у 1958 р. як компроміс між методами одиночного та повного зв'язку. У версії програм STATISTICA 13.3 та IBM SPSS Statistics 22 реалізовано два різновиди

методу: незважений метод «середнього зв'язку» («незважене попарне середнє») та зважений метод «середнього зв'язку». У першому різновиді методу відстань між двома кластерами обчислюється як середня відстань між усіма парами об'єктів у них, а у другій - крім того, розмір кластерів (тобто кількість об'єктів, що містяться в них) використовується як ваговий коефіцієнт. Використання кількості об'єктів, що містяться в кластері, як ваговий коефіцієнт передбачає «хорошу якість» аналізу за наявності у вибірці кластерів нерівного розміру. Результати кластеризації ділянок пласта за вмістом Ge обома методами у вигляді дендрограм наведені відповідно на рис. 3 та рис. 4. Аналіз наведених дендрограм дозволяють прийти до висновку, що у випадку з поділом ділянок пласта за вмістом Ge результати кластерного аналізу відповідно методами незваженого середнього зв'язку і зваженого середнього зв'язку у нашому випадку повністю співпадають.

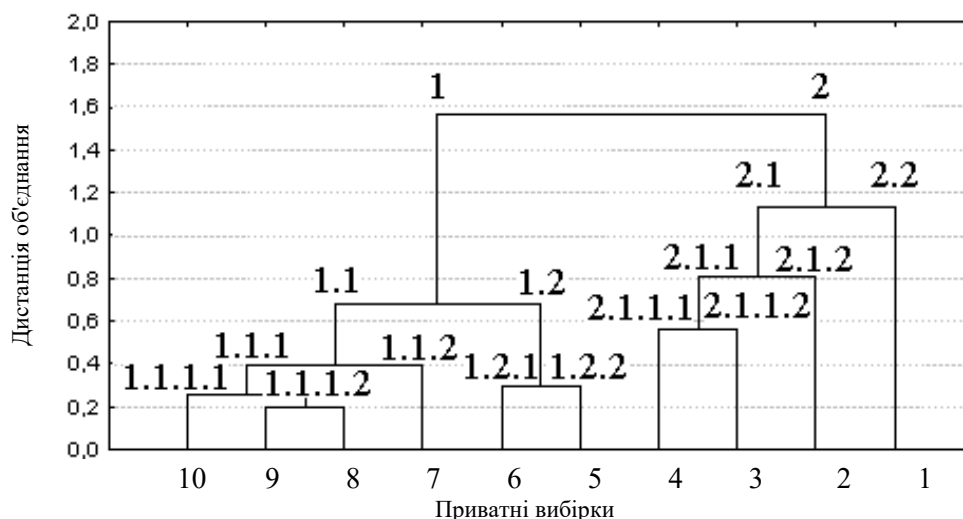


Рис. 3. Дендрограма результатів кластеризації методом незваженого середнього зв'язку ділянок по вмісту Ge. Умовні позначення: 1, 2, 1.1, 1.2, 1.1.1, 1.1.2, 1.1.1.1, 1.1.1.2, 1.2.1, 1.2.2, 2.1, 2.2, 2.1.1, 2.1.2, 2.1.1.1, 2.1.1.2 – кластери

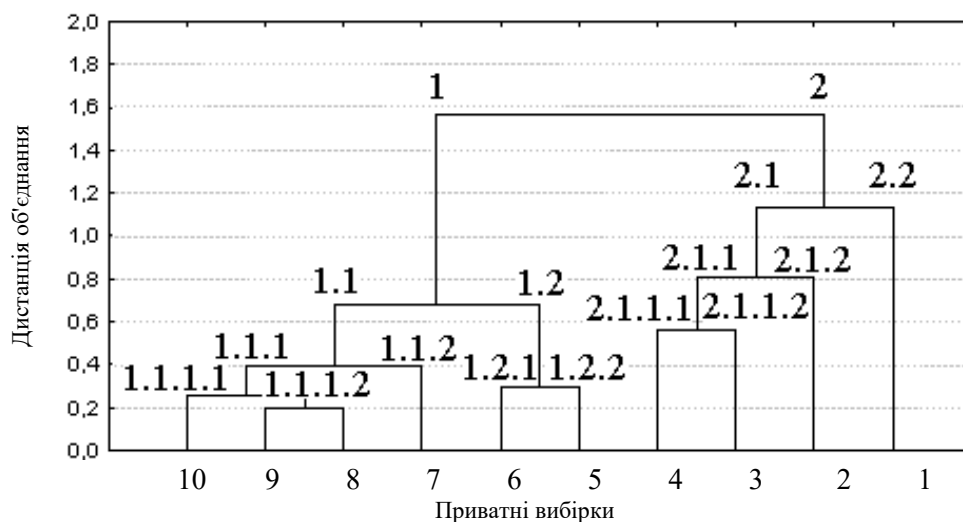


Рис. 4. Дендрограма результатів кластеризації методом зваженого середнього зв'язку ділянок по вмісту Ge. Умовні позначення: 1, 2, 1.1, 1.2, 1.1.1, 1.1.2, 1.1.1.1, 1.1.1.2, 1.2.1, 1.2.2, 2.1, 2.2, 2.1.1, 2.1.2, 2.1.1.1, 2.1.1.2 – кластери

Зважений центроїдний метод використовує як відстань між кластерами (об'єктами) відстань між їх центрами тяжіння. На рис. 5 наведено дендрограму результатів кластеризації зваженим центроїдним медіанним методом ділянок пласта c_{10}^B за вмістом Ge. Аналіз цієї дендрограми на відміну від наведених раніше (рис. 1, рис. 2, рис. 3 та рис. 4) 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_{10}^B за вмістом Ge та інтерпретувати отримані результати у геологічних поняттях.

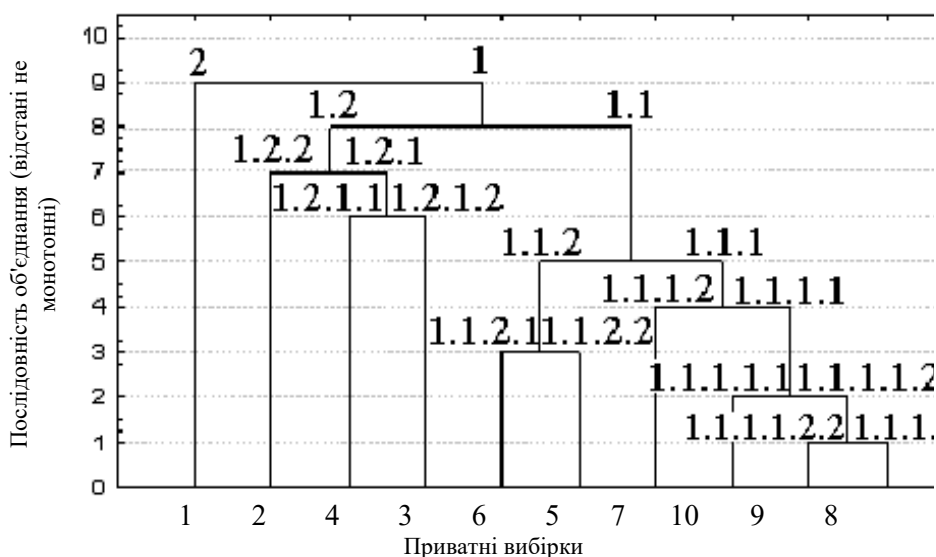


Рис. 5. Дендрограма результатів кластеризації зваженим центроїдним медіанним методом ділянок по вмісту Ge. Умовні позначення: 1, 2, 1.1, 1.2, 1.1.1, 1.1.2, 1.1.1.1, 1.1.1.2, 1.2.1, 1.2.2, 1.2.1.1, 1.1.2.1, 1.1.2.2, 1.1.1.1.1, 1.1.1.1.2, 1.1.1.1.2.1, 1.1.1.1.2.2 – кластери

Висновок. В результаті виконаних досліджень встановлено що використання зваженого центроїдного медіанного методу кластерного аналізу, який запроваджено у професійних статистичних програмних платформах «STATISTICA» та «SPSS» є найбільш оптимальним для вільного від суб'єктивізму дослідника поділу ділянок вугільного пласта c_{10}^B шахти «Дніпровська» за вмістом Ge на таксони.

Побудовані дендрограми кластеризації ділянок вугільного пласта c_{10}^B за вмістом Ge можуть бути використані як основа для розробки природної типізації вугільних пластів 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 надає можливість її для комплексного використання мінеральної сировини та вирішення стратегічних питань сталого розвитку України.

Список літератури

1. Naumov, A.V. World market of germanium and its prospects. *Russian Journal of Non-Ferrous Metals*. 2007. V. 48. № 4. P. 265-272. <https://doi.org/10.3103/S1067821207040049>
2. Павлов, А.В. Вещественный состав золы углей некоторых районов Западного Шпицбергена. *Уч. зап. НИИГА. Региональн. геол.* 1966. Вып. 8. С. 128-136.
3. Шпирт, М.Я., Нуkenov, Д.Н., Пунанова, С.А., Висалиев, М.Я. Принципы получения соединений

ценных металлов из горючих ископаемых. *Химия твердого топлива*. 2013. №2. С. 3-8.

4. Поляков, Н.П. Закономерности распределения микроэлементов в углях Алдан-Чульманского района. *Зап. Забайкалье. Филиал Географического общества СССР*. 1973. Вып. 81. С. 55-66.

5. Imai, N., Ando, A., Takeda, E. Minor elements in Japanese coal. *Bull. Surv. Jpn.* 1984.V. 35. № 7. P. 287-314.

6. Zodrow, E.L. Coal-stratigraphic geochemistry: trends in coal samples from Sydney Coalfield, Upper Carboniferous, Nova Scotia. *Can. Inst. Min. Met. Bull.* 1986. V. 79. № 893. P. 83-85.

7. Козий, Є.С., Ішков, В.В. Класифікація вугілля основних робочих пластів Павлоград-Петропавлівського геолого-промислового району за вмістом токсичних та потенційно токсичних елементів. *Збірник наукових праць «Геотехнічна механіка»*. 2017. № 136, С. 74-86.

8. Ішков, В.В., Козий, Є.С. Новые данные о распределении токсичных и потенциально токсичных элементов в угле пласта c_6^H шахты «Терновская» Павлоград-Петропавловского геолого-промышленного района. *Збірник наукових праць НГУ*. 2013. № 41. С. 201-208.

9. Ішков, В.В., Козий, Є.С. О распределении золы, серы, марганца в угле пласта c_4 шахты «Самарская» Павлоград - Петропавловского геолого-промышленного района. *Збірник наукових праць НГУ*. 2014. № 44. С. 178-186.

10. Kozar, M.A., Ishkov, V.V., Kozii, Ye.S., Pashchenko, P.S. 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. Geocology*. 2020. № 29(4). P. 722-730. <http://doi: 10.15421/112065>

11. Ішков В.В., Нагорный В.Н. О закономерностях накопления ртути в угольных пластах Красноармейского геолого-промышленного

района. *Науковий вісник Національної гірничої академії України*. 2005. № 2. С. 84-88.

12. Ішков, В.В., Козій, Є.С. Аналіз поширення хрому і ртуті в основних вугільних пластах Красноармійського геолого-промислового району. *Вид-во ІГН НАН України. Серія тектоніка і стратиграфія*. 2019. № 46. С. 96-104.

13. Єрофєєв, А.М., Ішков, В.В., Козій, Є.С., Барташевський, С.С. Дослідження методів кластеризації родовищ нафти Дніпровсько-Донецької западини з метою створення їх класифікації за вмістом металів (на прикладі V). *Наукові праці ДонНТУ. Серія Гірничо-геологічна*. 2021. №1(25)-2(26). С. 83-93. [https://doi.org/10.31474/2073-9575-2021-1\(25\)-2\(26\)-83-93](https://doi.org/10.31474/2073-9575-2021-1(25)-2(26)-83-93)

14. ГОСТ 10175-75. (1975). Угли бурые, каменные, антрациты, аргиллиты и алевролиты. Методы определения содержания германия. М: *Изд-во стандартов*. 14 с.

15. ГОСТ 9815-75. (1975). Угли бурые, каменные, антрацит и горючие сланцы. Метод отбора пластовых проб. М: *Изд-во стандартов*. 6 с.

16. Боровиков, В.П. STATISTICA: искусство анализа данных на компьютере. Для профессионалов. Санкт-Петербург. 2001. 656 с.

17. Ішков В.В., Козій Є.С. Аналіз поширення хрому і ртуті в основних вугільних пластах Красноармійського геолого-промислового району / Вид-во ІГН НАН України. Серія тектоніка і стратиграфія. 2019. № 46. С. 96 - 104.

18. Ішков В.В., Козій Є.С. О классификации угольных пластов по содержанию токсичных элементов с помощью кластерного анализа. *Збірник наукових праць НГУ*. 2014. № 45. С. 209 – 221.

References

1. Naumov, A.V. (2007). World market of germanium and its prospects. *Russian Journal of Non-Ferrous Metals*. 2007. V. 48. № 4. P. 265-272.

2. Pavlov, A.V. (1966). Veschestvennyy sostav zolyi ugley nekotorykh rayonov Zapadnogo Shpitsbergena. *Uch. zap. NIIGA. Regionaln. geol.* 1966. Vyip. 8. S. 128-136.

3. Shpirt, M.Ya., Nukenov, D.N., Punanova, S.A., Visaliev, M.Ya. (2013). Printsipy polucheniya soedineniy tsennykh metallov iz goryuchih iskopaemykh. *Himiya tverdogo topliva*. 2013. № 2. S. 3-8.

4. Polyakov, N.P. (1973). Zakonomnosti raspredeleniya mikroelementov v uglyah Aldan-Chulmanskogo rayona. *Zap. Zabaykale. Filial Geograficheskogo obschestva SSSR*, 1973. Vyip. 81. S. 55 – 66.

5. Imai, N., Ando, A., Takeda, E. Minor elements in Japanese coal. *Bull. Surv. Jpan.* 1984.V. 35. № 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.* 1986. V. 79. № 893. P. 83-85.

7. Kozii, Ye.S., Ishkov, V.V. (2017). Klasyfikatsiia vuhillia osnovnykh robochykh plastiv

Pavlohrad-Petropavlivskoho heoloho-promyslovoho raionu za vmistom toksychnykh ta potentsiino toksychnykh elementiv. *Zbirnyk naukovykh prats «Heotekhnichna mekhanika»*. 2017. № 136. S. 74-86.

8. Ishkov, V.V., Koziy, E.S. (2013). Novyye dannyye o raspredelenii toksichnykh i potentsialno toksichnykh elementov v ugle plasta s6n shahtyi «Ternovskaya» Pavlograd-Petropavlovskogo geologo-promyshlennogo rayona. *Zbirnyk naukovykh prats NHU*. 2013. № 41. S. 201-208.

9. Ishkov, V.V., Kozii, E.S. (2014). O raspredelenii zolyi, seryi, margantsa v ugle plasta s4 shahtyi «Samarskaya» Pavlograd - Petropavlovskogo geologo-promyshlennogo rayona. *Zbirnyk naukovykh prats NHU*. 2014. № 44. S. 178-186.

10. 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. Geocology*. 2020. №. 29(4). pp. 722-730. <http://doi: 10.15421/112065>

11. Ishkov, V.V., Nagornii, V.N. (2005). O zakonomernostyakh nakopleniya rtuti v ugolnikh plastakh Krasnoarmeiskogo geologo-promyshlennogo raionu. *Naukovyi visnyk Natsionalnoi hirnychoi akademii Ukrainy*. 2005. № 2. S. 84-88.

12. Ishkov, V.V., Kozii, Ye.S. (2019). Analiz poshyrennia khromu i rtuti v osnovnykh vuhilnykh plastakh Krasnoarmiiskoho heoloho-promyslovoho raionu. *Vyd-vo IHN NAN Ukrainy. Seriya tektonika i stratyhrafia*. 2019. № 46. S. 96-104.

13. Yeroficiev, A.M., Ishkov, V.V., Kozii, Ye.S., Bartashevskiy, S.Ie. (2021). Doslidzhennia metodiv klasteryzatsii rodovyshch nafty Dniprovsko-Donetskoi zapadyny z metoiu stvorennia yikh klasyfikatsii za vmistom metaliv (na prykladi V). *Naukovi pratsi DonNTU. Seriya Hirnycho-heolohichna*. 2021. №1(25)-2(26). S. 83-93. [https://doi.org/10.31474/2073-9575-2021-1\(25\)-2\(26\)-83-93](https://doi.org/10.31474/2073-9575-2021-1(25)-2(26)-83-93)

14. GOST 10175-75. Ugli buryie, kamennyie, antratsityi, argillityi i alevrolityi. Metodyi opredeleniya soderzhaniya germaniya. M: *Izd-vo standartov*. 1975. 14 s.

15. GOST 9815-75. Ugli buryie, kamennyie, antratsit i goryuchie slantsyi. Metod otbora plastovyih prob. M: *Izd-vo standartov*. 1975. 6 s.

16. Bоровиков, В.П. STATISTICA: искусство анализа данных на компьютере. *Dlya professionalov*. Sankt-Peterburg. 2001. 656 с.

17. Ishkov, V.V., Kozii, Ye.S. (2019). Analiz poshyrennia khromu i rtuti v osnovnykh vuhilnykh plastakh Krasnoarmiiskoho heoloho-promyslovoho raionu. *Vyd-vo IHN NAN Ukrainy. Seriya tektonika i stratyhrafia*. 2019. № 46. S. 96 - 104.

18. Ishkov V.V., Koziy E.S. (2014). O klassifikatsii ugolnykh plastov po soderzhaniyu toksichnykh elementov s pomoschy klasterного analiza. *Zbirnyk naukovykh prats NGU*. 2014. № 45. S. 209-221.

Надійшла до редакції 18.10.2022

Ішков Валерій Валерійович – кандидат геолого-мінералогічних наук, доцент кафедри геології та розвідки родовищ корисних копалин Національного технічного університету «Дніпровська політехніка», (м. Дніпро, Україна). Старший науковий співробітник лабораторії досліджень структурних змін гірських порід Інституту геотехнічної механіки ім. М.С. Полякова НАН України, (м. Дніпро, Україна).
E-mail: ishwishw37@gmail.com

Козій Євген Сергійович – кандидат геологічних наук, директор навчально-наукового центру підготовки іноземних громадян Національного технічного університету «Дніпровська політехніка», (м. Дніпро, Україна). Доцент кафедри цивільної інженерії, технологій будівництва та захисту довкілля Дніпровського державного аграрно-економічного університету, (м. Дніпро, Україна).

E-mail: kozyi.es@gmail.com

Чернобук Олександр Іванович – заступник директора департаменту стратегічного планування виробництва «Грузинський марганець», (м. Тбілісі, Грузія).

E-mail: o.chernobuk@gm.ge

Хоменко Володимир Львович – кандидат технічних наук, доцент кафедри нафтогазової інженерії та буріння Національного технічного університету «Дніпровська політехніка», (м. Дніпро, Україна).

E-mail: homenko.v.l@nmu.one

CLUSTERIZATION RESULTS OF DIFFERENT THICKNESS SECTIONS OF COAL SEAM C_{10}^B OF THE «DNIPROVSKA» MINE BY THE CONTENT OF GERMANIUM

Purpose. The purpose of the work is to establish the most effective method of creating an objective typification of sections of coal seam c_{10B} of the "Dniprovskia" mine of different thickness according to germanium concentrations, based on the analysis of the clustering results.

Methodology. 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. To achieve the goal set in the work, in the process of research, clustering was carried out using various methods, which are implemented in the most popular professional statistical software platforms "STATISTICA" and "SPSS"; their analysis was performed and the choice of the most optimal of them was substantiated. Excel 2016, STATISTICA 13.3 and IBM SPSS Statistics 22 versions were used in the work.

Results. The main results of the research consist in the selection of the optimal method of clustering of areas of different thickness of the coal seam. The analysis of the dendrogram of the results of clustering by the weighted centroid median method of the c_{10B} seam sections by germanium content, unlike others, allows not only to achieve the most stable division of the entire set of sections under consideration, but also to maximize the visualization of their breakdown 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 areas of the c_{10B} coal seam by germanium content and to interpret the obtained results in geological terms.

Scientific novelty. The scientific novelty of the results of the conducted research consists in the establishment of the weighted centroid median method of cluster analysis, which is implemented in the professional statistical software platforms "STATISTICA" and "SPSS" as the most optimal for the subjectivity-free researcher of the division of sections of the coal seam c_{10B} of the "Dniprovskia" mine by content germanium into taxa.

Practical significance. The practical significance of the work results is that 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 Dniprovskia 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 comprehensive use of mineral raw materials and to solve strategic issues of sustainable development of Ukraine.

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

Ishkov Valerii – Candidate of Geological and Mineralogical Sciences, Associate Professor of the Department of Geology and Mineral Prospecting, Dnipro University of Technology, (Dnipro, Ukraine). Senior Research Fellow of Laboratory of Studies of Structural Changes Rocks, Institute of Geotechnical Mechanics named by M.S. Poliakov of National Academy of Sciences of Ukraine, (Dnipro, Ukraine).

E-mail: ishwishw37@gmail.com

Kozii Yevhen – Candidate of Geological Sciences, 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, Ukraine).

E-mail: kozyi.es@gmail.com

Chernobuk Oleksandr – Deputy Director of Department of Strategic Production Planning, Georgian Manganese, (Tbilisi, Georgia).

E-mail: o.chernobuk@gm.ge

Khomenko Volodymyr – Candidate of Technical Sciences, Associate Professor of the Department of Oil and Gas Engineering and Drilling, Dnipro University of Technology, (Dnipro, Ukraine).

E-mail: homenko.v.l@nmu.one