

М. М. Лабушев

*Сибирский федеральный университет
пер. Вузовский, 3, Красноярск, 660025, Россия*

mlabushev@yandex.ru

ОТ ПАРАМЕТРИЧЕСКОГО КЛАССИФИЦИРОВАНИЯ МИНЕРАЛОВ И ГОРНЫХ ПОРОД К ОБЩЕМУ РЕШЕНИЮ ПРОБЛЕМЫ КЛАССИФИЦИРОВАНИЯ ВЕЩЕСТВА ПО ХИМИЧЕСКОМУ СОСТАВУ

Предложено классифицирование минералов и горных пород на основе уравнений теории информации для расчета взаимной информации двух величин. Методом Монте-Карло из большого массива атомных масс, соответствующих химическому составу минерала или горной породы, извлекается 8 атомных масс и формируется двумерный массив 3×3 , элементами которого становятся восемь атомных масс химических элементов и их сумма. Определяется информационный коэффициент пропорциональности как алгебраическая сумма трех неопределенностей – распределения атомных масс по строкам, столбцам и двумерному массиву. Средние значения и стандартные отклонения таких распределений используются как классификационные показатели. Для расчетов создана программа Agemarket с открытым исходным кодом. Дано описание этой программы с примерами расчетов, которые иллюстрируют базовые классификационные возможности и возможности определения относительного возраста горных пород на примере известняков.

Ключевые слова: классифицирование минералов и горных пород, теория информации, информационные коэффициенты пропорциональности, относительный возраст горных пород.

Памяти Зинаиды Николаевны Файнберг посвящается

Введение

Целью данного исследования является определение новых алгоритмов классифицирования горных пород. Гипотезой исследования принимаем возможность использования атомных масс химических элементов для классифицирования химических соединений, минералов и горных пород с получением однородных количественных показателей, которые можно обрабатывать методами математической статистики.

В настоящее время используются десятки приемов отображения химизма горных пород. Это методы пересчетов CIPW, П. Ниггли, А. Ритмана, система А. Н. Заварицкого для расчета индексов, пересчеты по методам Т. Барта, А. А. Предовского и А. А. Маракушева, методы Л. С. Бородина, А. Миша, В. П. Флоренского, О. М. Розена, RNA и др. Несмотря на это, не удалось существенно уменьшить неопределенность разделения разновидностей пород [1].

Для достижения поставленной цели определим требования к расчетам. Первым требованием укажем их связь с расчетом энтропии химического состава минералов и горных пород. Второе требование связано с учетом содержания всех химических элементов в составе мине-

Лабушев М. М. От параметрического классифицирования минералов и горных пород к общему решению проблемы классифицирования вещества по химическому составу // Вестн. НГУ. Серия: Информационные технологии. 2017. Т. 15, № 2. С. 47–58.

ралов и горных пород. В качестве третьего требования определим то, что классификационные показатели должны быть однородными количественными для возможности дальнейшей обработки методами математической статистики.

Эти требования в совокупности можно охарактеризовать как требования фундаментальности, полноты и однородности.

Как известно, принятые в Петрографическом Кодексе [2] иерархические классификации кристаллических горных пород учитывают разнородные, количественные и качественные, генетические, фациальные, химические, минералогические и структурно-текстурные признаки. Эти данные неоднородны и поэтому не соответствуют указанным требованиям.

В наибольшей степени этим требованиям соответствует метод РНА [3], в нем вычисляются информационная энтропия и ранговая формула содержания химических элементов. Неоднородность значений информационной энтропии при ее расчетах для разного количества химических элементов, возможно, является наиболее существенным недостатком метода. Так, например, методами математической статистики нельзя совместно обрабатывать показатели информационной энтропии для пирита и халькопирита, так как уже для содержания двух и трех химических элементов получаем неоднородные результаты.

Дополнение указанных расчетов в РНА ранговой формулой не приводит к однородным результатам. Все сказанное характеризует расчеты в РНА как информативные, но не рациональные для классификации минералов и горных пород. Необходима разработка новых вариантов расчетов.

Параметрическое классифицирование горных пород и минералов

Предлагается параметрическое классифицирование по химическому составу. При таком классифицировании проводятся расчеты информационных коэффициентов пропорциональности I_p , которые предложены как математическое обобщение понятия «коэффициент пропорциональности» [4; 5] на основе уравнений теории информации для расчета взаимной информации двух величин.

В расчетах I_p вместо вероятности совместных событий используются атомные массы химических I_p элементов, и рассчитывается известная в теории информации алгебраическая сумма трех неопределенностей, в нашем случае – распределения атомных масс в матрице 3×3 по строкам, столбцам и в целом. Важно то, что перестановка элементов матрицы во многих случаях приводит к изменению результирующего I_p , что делает расчеты более информативными.

Таковыми расчетами установлено свойство атомов химических элементов находиться в составе минералов и горных пород в определенных пропорциях, которые характеризуются унимодальными распределениями I_p их атомных масс. Атомные массы химических элементов берутся в расчеты пропорционально вероятности присутствия соответствующих атомов в составе минерала или горной породы.

Методом Монте-Карло из большого массива атомных масс формируется двумерный массив 3×3 , элементами которого становятся восемь атомных масс химических элементов и их сумма. Алгебраическая сумма трех неопределенностей обеспечивает распознавание перестановок атомных масс в строках и столбцах массива.

В одном из расчетов I_p для сильвина случайно выбрана атомная масса калия 39.0983 и семь атомных масс хлора 35.45, первая атомная масса и суммарный элемент случайно расположились в двумерном массиве (матрице) как крайние диагональные элементы. Построчные суммы атомных масс 109.9983, 106.3500 и 358.1483 при таком расположении равны суммам столбцов, а сумма всех элементов массива равна 574.4966. Результаты расчета I_p :

$$I_p = -2 \cdot (109.9983/574.4966 \cdot \ln(109.9983/574.4966) + \\ + 106.35/574.4966 \cdot \ln(106.35/574.4966) + \\ + 358.1483/574.4966 \cdot \ln(358.1483/574.4966)) -$$

$$\begin{aligned}
 & -(-7 \cdot (35.45/574.4966 \cdot \ln(35.45/574.4966))) - \\
 & -39.0983/574.4966 \cdot \ln(39.0983/574.4966) - \\
 & -287.2483/574.4966 \cdot \ln(287.2483/574.4966) = 0.11409938.
 \end{aligned}$$

В предлагаемых расчетах все дроби представляют собой обычные коэффициенты пропорциональности. Если умножить все элементы расчетного массива на любую положительную константу, результат расчетов не изменится, подобно тому, как не меняется значение коэффициента пропорциональности при умножении числителя и знаменателя на константу. Таким образом, все I_p по этому свойству соответствуют коэффициентам пропорциональности (табл. 1).

Результаты расчетов полностью определяются значениями элементов массива и их суммами по строкам и столбцам. Поэтому I_p из приведенного выше примера не изменит свое значение, если атомную массу калия и атомную массу любого смежного элемента массива поменять местами. Но если поместить ее в одну строку или столбец с суммарным элементом, то значение I_p изменится.

После определения большого массива I_p получаются их унимодальные распределения, которые удовлетворительно симметризируются извлечением квадратного корня из каждого I_p с последующим определением их среднего значения I_{av} . Это позволяет использовать для обработки результатов параметрические методы математической статистики. Сама классификация характеризуется как параметрическая, так как I_{av} в первом приближении можно рассматривать как один из двух параметров нормального распределения.

Предполагается, что пропорциональность атомных масс, так же как и сами атомные массы, имеет фундаментальную природу и определяет предельное число минералов, неорганических и органических соединений как число сочетаний по две, три и четыре атомные массы из 95 атомных масс природных химических элементов [4].

Высказана гипотеза, что распределение I_p атомных масс для каждого минерала, неорганического и органического химических соединений по значениям I_{av} совпадает с одним из распределений таких коэффициентов соответственно для двух, трех или четырех атомных масс химических элементов из 95. Природа такой закономерности еще ждет своего объяснения.

Подобные соотношения распределений I_p покажем на следующем примере. Если бросать много раз куб и тетраэдр с пронумерованными по образцу игральной кости гранями, то в результате получим следующие распределения вероятностей «число (вероятность появления при бросании)»: 1(1/6), 2(1/6), 3(1/6), 4(1/6), 5(1/6), 6(1/6) и 1(1/4), 2(1/4), 3(1/4), 4(1/4). Затем рассчитаем средние значения (математические ожидания) двух полученных распределений: $1 \cdot 1/6 + 2 \cdot 1/6 + 3 \cdot 1/6 + 4 \cdot 1/6 + 5 \cdot 1/6 + 6 \cdot 1/6 = 3.5$ и $1 \cdot 1/4 + 2 \cdot 1/4 + 3 \cdot 1/4 + 4 \cdot 1/4 = 2.5$. Если числа на гранях куба заменить на 0.6, 1.2, 2, 2.4, 4.2, 4.6 то по среднему значению результатов бросаний, равному 2.5, такой куб будет соответствовать тетраэдру.

Если вместо граней куба и тетраэдра представить число атомов в двух минералах, а вместо чисел на гранях – значения I_p для соответствующих множеств атомных масс, то мы можем получить равные I_{av} для минералов с шестью и четырьмя атомами. Этот пример условен, так как число различных I_p для указанных минералов на много порядков превышает число граней как тетраэдра, так и куба.

При предполагаемых разновидностях минералов, неорганических и органических соединений, равных 4 465, 138 415 и 3 183 545 и объединенных в пакеты по 95 химических соединений [4; 5], для горных пород не может быть других сочетаний атомных масс, кроме тех же 4 465 состояний. Горные породы должны по средним значениям I_{av} соответствовать определенным минералам.

Такие мономинеральные породы, как мраморы, кварциты, альбититы, сильвиниты, можно сопоставить с кальцитом, кварцем, альбитом и сильвином. С позиций высказанной выше гипотезы такие породы проявляют тенденцию к «мономинеральности». При многообразии атомов разных химических элементов формируются полиминеральные горные породы, для сопоставления которых с определенным минералом необходимы расчеты. Например, по распределениям I_p установлено соответствие микроклина и одной из разновидностей диорита.

Таблица 1

Значения I_p и соответствующие им частоты классификационного расчета для сильвина

№	I_p	I_p Squareroot	Frequency	Approximate frequency	№	I_p	I_p Squareroot	Frequency	Approximate frequency
1	0.0958277	0.3095604	78042	1	24	0.1103618	0.3322075	312319	4
2	0.0993197	0.3151503	312118	4	25	0.1104687	0.3323683	624245	8
3	0.0995709	0.3155486	312514	4	26	0.1104884	0.3323980	624786	8
4	0.1028230	0.3206603	156643	2	27	0.1105582	0.3325030	312097	4
5	0.1028856	0.3207579	312674	4	28	0.1106181	0.3325930	12500133	16
6	0.1030199	0.3209672	623634	8	29	0.1138279	0.3373838	312307	4
7	0.1031349	0.3211462	313637	4	30	0.1139117	0.3375080	313052	4
8	0.1032229	0.3212832	624875	8	31	0.1139120	0.3375086	312604	4
9	0.1033266	0.3214446	157183	2	32	0.1140471	0.3377086	626021	8
10	0.1064641	0.3262883	312379	4	33	0.1140994	0.3377860	312514	4
11	0.1065405	0.3264054	312779	4	34	0.1141057	0.3377954	624415	8
12	0.1065408	0.3264059	312733	4	35	0.1141109	0.3378030	626100	8
13	0.1066847	0.3266262	624796	8	36	0.1142997	0.3380824	311793	4
14	0.1067140	0.3266710	312264	4	37	0.1143163	0.3381070	937036	12
15	0.1067413	0.3267128	625944	8	38	0.1175534	0.3428607	156962	2
16	0.1067459	0.3267199	625252	8	39	0.1176108	0.3429443	312342	4
17	0.1069353	0.3270097	935716	12	40	0.1177570	0.3431575	625194	8
18	0.1069513	0.3270342	312737	4	41	0.1178832	0.3433412	312583	4
19	0.1101189	0.3318417	156260	2	42	0.1179535	0.3434435	624743	8
20	0.1102188	0.3319921	624746	8	43	0.1180912	0.3436440	156370	2
21	0.1102770	0.3320798	624576	8	44	0.1214115	0.3484416	313238	4
22	0.1102804	0.3320849	625576	8	45	0.1216813	0.3488285	312016	4
23	0.1103553	0.3321977	311965	4	46	0.1252888	0.3539616	78207	1

Таким образом, использование предлагаемого классифицирования может способствовать систематизации химических соединений, минералов и горных пород. Кроме этого, классификационный показатель для горных пород при определенных условиях в пределах локальных участков, как установлено экспериментально, отражает их относительный возраст [6; 7].

Технология классификационных расчетов

В 2012 г. для предложенных расчетов была создана программа Agemarket с открытыми исходными текстами, с декабря этого же года она находится в свободном доступе на сайте www.skyproject.org. Программа позволяет дать полную количественную характеристику пропорциональности атомных масс химических элементов химических соединений, минералов и горных пород.

Для вычисленных первичных и симметризованных распределений I_p рассчитываются значения I_{av} , соответственно **Average (Ip)** и **Average (Ip Squareroot)**, дисперсия, стандартное отклонение, показатели асимметрии и эксцесса выборки I_p [6]. Предполагается, что основную информацию содержит первый показатель. Результаты расчетов записываются в формате текстового файла, который может быть использован для проверки исходных данных и для повторных расчетов.

Строка меню в интерфейсе Agemarket включает **File**, **Calculations** и **Help**. Меню **File – Setting** позволяет определить число потоков при вычислениях. Вначале при помощи **Calculations** необходимо сделать выбор между вводом данных с клавиатуры **New** или загрузкой их из текстового файла, созданного в Agemarket ранее **Load from results**.

Исходные данные вводятся в массовых процентах содержаний оксидов или химических элементов, также возможны другие единицы измерения, например г/т. Для одного расчета допустимо использовать данные только в одних единицах измерения. Возможен прямой ввод числа атомов (атомных масс) с их последующим автоматическим пропорциональным увеличением для расчетов.

В меню **New** необходимо при помощи переключателей выбрать **Mass %** или **Number of Atoms**. При выборе **Mass %** доступен ввод данных в таблицы оксидов и химических элементов. Если химический состав представлен как в виде оксидов, так и в виде содержаний отдельных химических элементов, данные вводятся в обе таблицы, в остальных случаях данные вводятся в одну из них.

Например, если в образце кристаллического CO_2 содержится еще и свободный углерод, общий состав должен быть отражен в двух таблицах, для расчетов Agemarket эти содержания суммирует. Подобно этому суммируются все содержания кислорода из таблицы оксидов. При выборе **Number of Atoms** как исходные данные для расчетов указываются числа атомов химических элементов. Для чистого Al_2O_3 , например, можно указать две атомные массы алюминия и три атомные массы кислорода. Расчеты с учетом статистической погрешности дадут тот же результат, что и при вводе соответствующих данных в таблицу содержаний оксидов или химических элементов.

Содержание каждого оксида пересчитывается на содержание химического элемента и кислорода. Полученное содержание суммируется с содержанием этого же элемента, введенного в таблицу содержания химических элементов. Показатели содержания кислорода в оксидах всех изучаемых химических элементов суммируются, и также к полученному значению добавляется содержание кислорода из таблицы содержания химических элементов.

Суммарное содержание каждого химического элемента делится на соответствующую ему атомную массу. В зависимости от требуемого общего числа расчетов I_p каждое частное умножается на целое число **Multiplier**. Числа атомных масс каждого химического элемента, пропорциональные содержанию этих элементов в исходной пробе, получаем после округления каждого такого произведения до целого числа. Эти числа затем умножаются на 8, так как сумма атомов всех химических элементов в расчетах должна быть кратна 8.

Изменяя значение **Multiplier**, мы изменяем общее число расчетов I_p , а если ввести нужное число таких расчетов в поле **Calculations**, автоматически изменится величина **Multiplier**. В этом диалоговом окне для построения частотного ряда распределения I_p можно указать

число интервалов группирования I_p (по умолчанию 7) и выбрать основание логарифмов для расчетов, по умолчанию используются натуральные логарифмы.

Например, если ввести только содержание углерода 12.011 %, оно будет нормировано на 12.011, и при требовании 100 расчетов I_p значение **Multiplier** будет показано как 100, так как $100 \cdot 8 \cdot 12.011 / 12.011 = 800$. Для приведенного примера число расчетов будет равно 100, хотя сами расчеты для этого примера будут проведены только с атомными массами углерода. Будет получено 100 I_p со значением 0.11012 (информационный коэффициент моноэлементной пропорциональности), как и в случае расчетов для любого чистого химического элемента.

Из примера видно, что сумма значений содержания химических элементов может отличаться от 100 %, это дает возможность делать расчеты для неполных систем атомных масс, например для макро- и микроэлементов минералов и горных пород. Фактически за 100 % в расчетах может приниматься любая часть обрабатываемого химического состава. Такие расчеты не противоречат рекомендациям Подкомиссии по систематике изверженных пород Международного союза геологических наук (МСГН), при классифицировании магматических горных пород их химические анализы пересчитываются на 100 % [2].

Для минералов и горных пород рекомендуется выполнять не менее 50 млн расчетов I_p , так как установлено, что, если в составе минерала или горной породы содержится до 7–8 химических элементов в высоких концентрациях, это обеспечивает точность результата до четырех значащих цифр, а для двух химических элементов достаточно выполнять 5 млн расчетов.

Для примера определим показатель I_{av} для чистого кварца. Вводим в **Oxides table** 100 % для SiO_2 , вместо этого можно ввести значения содержания кремния и кислорода в **Elements table**. Устанавливаем **Multiplier** 3000000 и по умолчанию принимаем **Logarithmic base: e**.

Agemarker вычисляет содержание кислорода и кремния, соответственно 53.25633 и 46.74367 %. После деления этих величин на атомные массы соответствующих элементов получаем 3.3287287 и 1.6643643. Умножением на 3000000 и округлением результата до целых чисел получаем числа атомных масс 9986186 и 4993093. Умножив их на 8, получаем 79889488 и 39944744 атомных масс кислорода и кремния соответственно и **Atomic weights (total sum)**: 119834232, что обеспечит 14979279 расчетов I_p .

Из массива атомных масс случайным образом отбираются по 8 атомных масс и располагаются в матрице 3×3 для определения I_p и корня квадратного из него **Ip Squareroot**. В ходе расчетов определяются и частоты каждого I_p . Извлечение квадратного корня из каждого I_p симметризует полученное распределение. Классификационным показателем является среднее значение I_{av} симметризованных значений I_p **Average (Ip Squareroot)**. В нашем примере может быть получено значение 0.34307.

Файл расчетов в формате txt содержит все исходные данные и результаты расчетов в виде частотного ряда распределения с указанием значения каждого I_p , квадратного корня из него и соответствующей эмпирической частоты. Далее в блоке выборочных статистик указываются средние значения, дисперсии, показатели асимметрии и эксцесса и абсолютная среднеквадратичная случайная погрешность среднего значения. Эти статистики приводятся для симметризованных данных, а затем, отдельным блоком, – для исходных. Ниже показывается частотный ряд распределения для указанного числа интервалов группирования I_p .

Файл результатов может быть загружен для нового расчета с другими установками. По умолчанию предусмотрена постановка расчетов в очередь и защита от непредвиденного прерывания расчетов. При отключении питания, например, очередь расчетов восстанавливается при следующем запуске программы.

Для оценки соответствия расчетов случайной модели выбора атомных масс можно использовать эмпирические и теоретические частоты значений I_p для восьми атомных масс одного химического элемента. Такой I_p представляет собой константу, при расчетах в натуральных логарифмах равную 0.11011891.

Приведем пример расчетов теоретических частот для минерала, состав которого определяется равным числом атомов двух химических элементов. Для расчета I_p вероятность выбора атомной массы одного из двух химических элементов при единичном выборе равна 0.5, вероятность того, что для расчета I_p будут использованы 8 атомных масс одного и того же

химического элемента, определяем по теореме умножения вероятностей для независимых событий $0.5^8 = 0.00390625$.

Эта вероятность характеризует и выбор восьми атомных масс второго химического элемента, расчеты с которыми также дают значение $I_p = 0.11011891$. Поэтому необходимо удвоить полученную вероятность. Таким образом, искомая суммарная теоретическая вероятность равна 0.0078125. Следовательно, для каждых 128 расчетов I_p для рассматриваемых двухатомных минералов, теоретическая частота встречаемости указанного коэффициента будет равна 1, как произведение вероятности на число расчетов по формуле расчета математического ожидания биномиального распределения.

В общем случае для определения теоретических частот сочетаний атомных масс при расчетах I_p используем биномиальный закон распределения. Событием считаем выбор для расчета информационного коэффициента пропорциональности атомной массы одного из двух химических элементов. Одним из параметров биномиального распределения является вероятность этого события в единичном испытании, для нашего случая она равна 0.5, а второй параметр равен 8. Вероятность выбора заданных атомных масс при этих параметрах равна 0.00390625, как и в случае применения теоремы умножения, далее удваиваем эту вероятность.

Определим теоретическую вероятность появления при расчетах значения I_p для бинарных минералов с равным количеством атомов химических элементов при случайном выборе, когда в расчетной матрице их соотношение равно 1 к 7 при расположении, аналогичном приведенному с расчетами выше. Указанным параметрам биномиального распределения и семи случаям появления события соответствует вероятность 0.03125. Такой же вероятности соответствует появление в матрице расчетов только одного атома второго химического элемента.

Вероятность распределения атомных масс двух химических элементов в расчетном массиве (в приведенном выше примере для калия и хлора), с условием неизменности всех указанных в примере сумм атомных масс по строкам и столбцам, равна 0.5, так как имеем 4 варианта расположения из 8 возможных. Вероятность совместного наступления двух охарактеризованных независимых событий составляет 0.015625 как результат умножения двух вероятностей. При расчете 256 значений I_p теоретическая частота для этого расположения атомных масс равна 4.

В расчетах суммарный элемент может быть любым элементом в любой строке и столбце матрицы. Вероятности появления при расчетах определенных значений I_p , рассчитанные для любого фиксированного положения суммарного элемента, не изменяются при любых перестановках суммарного элемента, так как число вариантов будет увеличиваться в 2, 3, ..., 8 раз относительно первоначального размещения суммарного элемента. При расчетах вероятностей матричных распределений эти множители можно сократить.

Расчеты вероятностей других сочетаний атомных масс для рассматриваемых бинарных минералов можно провести по указанному алгоритму или получить результат прямого расчета в Agemarket (для случая равенства числа атомов двух химических элементов). Дополнительно к стандартным расчетам все частоты нормируются на значение минимальной частоты из расчетного распределения с округлением частных от деления до целых чисел.

По 80 млн атомных масс калия и хлора использованы для расчета 20 млн I_p . Полученные частоты нормируются на минимальную частоту 78042 и округляются до целых чисел, сумма которых теоретически должна быть равна 256. Аппроксимирующие значения частот в данном примере соответствуют теоретическим для 256 расчетов I_p . Сравнение частот и их аппроксимирующих значений примера позволяют увидеть статистический характер расчетов, при этом, например, теоретической частоте 78125 соответствуют эмпирические частоты 78042 и 78207.

Так, теоретическая вероятность появления при расчетах информационного коэффициента моноэлементной пропорциональности 0.1101189, как было показано выше для бинарных двухатомных химических соединений, равна 0.0078125. При 256 расчетах I_p для рассматриваемого бинарного химического соединения KCl теоретическая частота моноэлементного коэффициента будет равна 2, что совпадает с аппроксимирующим значением расчетов.

Рассмотрим результаты расчетов для стандартных образцов пород разного состава (табл. 2).

Таблица 2

Стандартные образцы из реестра ФГУП «ВИМС»

Порода	Содержание SiO ₂ , масс. %	Массив I _p	I _{av}	Стандартное отклонение I _p	Стандартная ошибка I _{av}	Номер
Гранит	72.94	5·10 ⁸	0.3452	0.0496	7.014·10 ⁻⁶	ГСО 7224-96
Аляскитовый гранит	72.10	5·10 ⁸	0.3467	0.0529	7.475·10 ⁻⁶	ОСО 353-07
Гранодиорит	62.36	5·10 ⁸	0.3480	0.0558	7.894·10 ⁻⁶	ОСО 31-84
Диорит кварцевый	60.45	5·10 ⁸	0.3489	0.0578	8.177·10 ⁻⁶	ГСО 6103-91
Габбро	46.63	5·10 ⁸	0.3513	0.0620	8.768·10 ⁻⁶	ГСО 8670-2005
Дунит	39.58	5·10 ⁸	0.3456	0.0521	7.374·10 ⁻⁶	ГСО 4233-88

Значения классификационного показателя I_{av} и стандартного отклонения I_p возрастают с уменьшением содержания SiO₂ примерно до 47 %, но для дунита фиксируется уменьшение классификационного показателя. Этот «парадоксальный» результат легко объясним. Если классификация основана на содержании только одного компонента, то она неизбежно теряет информативность при его низком содержании в изучаемой породе.

В следующем примере показаны возможности определения относительного возраста пород на примере известняков и доломитов штата Вирджиния (табл. 3). Для каждого анализа было рассчитано по 50 млн I_p .

Относительный возраст пород принимаем по закону Стено. Вверх по первым трем разрезам наблюдается возрастание значений I_{av} известняков, исключая окварцованные разности с содержанием SiO₂ более 1,5 %. Важным исключением является первый слой разреза 34, который неотличим по этому показателю от вышележащих известняков, вероятно, вследствие малой мощности слоя. Можно предположить, что для определения относительного возраста известняков их опробованные мощности должны быть не менее 10–15 м.

Если неизменные известняки различаются между собой по величине I_{av} не более чем на 0.0003, для окварцованных разностей отличие от неизменных наблюдается уже на величину от 0.0005 до 0.0015, то доломиты отличаются от измененных известняков на величину от 0.0063 до 0.007, а от неизменных известняков – на 0.0075 и более. Таким образом, окварцованные известняки по значению I_{av} гораздо ближе к неизменным известнякам, чем к доломитам.

Этот пример демонстрирует важное свойство классификационной статистики I_{av} . Небольшое изменение химического состава приводит к небольшому изменению I_{av} , а существенному изменению состава соответствует существенное изменение I_{av} . Для графического отображения результатов можно построить точечную диаграмму, на которой отображаются пары значений I_{av} – стандартное отклонение I_p , соответственно по оси абсцисс и ординат.

Обсуждение параметрического классифицирования

Выше охарактеризовано использование для классифицирования только одного параметра нормального распределения, так как во всех расчетах средние квадратические отклонения массивов I_p пропорциональны значениям I_{av} и поэтому несут в себе мало дополнительной информации.

Таблица 3

Химический состав ордовикских известняков и доломитов штата Вирджиния [8]
и показатели пропорциональности атомных масс пород

Геологический разрез (слой)	Мощность, м	Порода *	Содержание, масс. %								Сумма	I_{av}	Стандартное отклонение I_p
			CaO	MgO	CO ₂	SiO ₂	Al ₂ O ₃	Fe ₂ O ₃	P ₂ O ₅				
2 (7)	44	1	54.77	0.48	43.50	0.33	0.36	0.44	–	99.88	0.3586	0.0741	
2 (3 и 4)	28	1	54.19	0.64	43.22	0.74	0.31	0.50	–	99.60	0.3585	0.0739	
2 (1)	94	2	29.97	17.16	42.26	7.28	2.52	0.80	–	99.99	0.3508	0.0625	
8 (3)	10	1	54.62	0.41	43.31	1.28	0.20	0.16	0.063	100.04	0.3584	0.0738	
8 (2)	20	1	54.51	0.51	43.33	1.46	0.06	0.08	0.098	100.05	0.3583	0.0737	
9 (10)	21	1	54.95	0.38	43.54	0.84	0.24	0.16	0.065	100.17	0.3585	0.0739	
9 (9)	2	1	51.95	0.61	41.43	5.04	0.60	0.40	–	100.03	0.3578	0.0729	
9 (8)	26	1	54.51	0.36	43.18	1.32	0.44	0.16	0.066	100.04	0.3584	0.0737	
9 (5–7)	15	1	54.06	0.51	42.98	1.52	0.40	0.12	–	99.59	0.3583	0.0736	
34 (4 и 5)	34	1	54.71	0.45	43.43	0.43	0.69	0.26	–	99.98	0.3585	0.0739	
34 (2 и 3)	15	1	48.28	0.79	38.75	9.77	1.16	0.62	–	99.37	0.3570	0.0715	
34 (1)	8	1	54.09	0.77	43.29	1.03	0.01	0.72	–	99.91	0.3585	0.0740	

* 1 – известняки, 2 – доломиты.

Предложенное классифицирование применимо не только к химическим соединениям, минералам и горным породам, но и к сплавам, а также к химическим соединениям с примесями. Оно не является иерархическим в силу своей параметрической природы.

Параметрическое классифицирование призвано существенно дополнить такие целочисленные соотношения, как заряд ядра атома, валентность, степень окисления и др. Так как классифицирование осуществляется только по значениям I_{av} , для установления классификационных соотношений возможно использование критериев Стьюдента и Фишера (ANOVA). Так может быть решена, например, проблема числа минералов в изоморфных рядах: альбита и анортита, диопсида и геденбергита и др.

Означает ли использование параметрического классифицирования отказ от традиционных классификаций, пересчетов, отображения результатов в виде различных диаграмм? Рациональным представляется их комплексирование с обязательным использованием параметрической классификации. Перспективно совместное использование предлагаемых расчетов и метода РНА. Вместе с тем предлагаемые расчеты могут показать степень информативности классификаций на основе обработки неполной информации, например, с использованием TAS-диаграмм сумма щелочей – кремнезем [9; 10].

Заключение

Расчеты в Agemarket соответствуют самым жестким требованиям и пригодны для обработки методами математической статистики. Классификационные расчеты естественным образом могут быть дополнены систематизацией химических соединений, минералов и горных пород.

С помощью Agemarket уже составлены первые пакеты минералов [11] предполагаемой периодической системы. Все разработки в этом направлении предлагается квалифицировать в качестве общественного достояния.

Список литературы

1. Ефремова С. В., Стафеев К. Г. Петрохимические методы исследования горных пород: Справочное пособие. М.: Недра, 1985. 511 с.
2. Петрографический кодекс России. Магматические, метаморфические, метасоматические, импактные образования. 3-е изд. СПб.: Изд-во ВСЕГЕИ, 2009. 200 с.
3. Петров Т. Г. Метод РНА как решение проблемы систематизации аналитических данных о вещественном составе геологических объектов // Отечественная геология. 2008. № 4. С. 98–105.
4. Лабушев М. М. О предельно возможном числе минералов, неорганических и органических химических соединений // Journal of Siberian Federal University. Engineering & Technologies. 2008. № 3. С. 221–233.
5. Labushev M. M. The Periodic Table as a Part of the Periodic Table of Chemical Compounds. 2011. 18 p. URL: <http://arxiv.org/abs/1103.3972>
6. Лабушев М. М., Хохлов А. Н. Новый подход к определению относительного возраста магматических горных пород // Современные технологии освоения минеральных ресурсов: Сб. науч. тр. Красноярск, 2011. С. 64–68.
7. Labushev M. M., Khokhlov A. N. Relative Dating and Classification of Minerals and Rocks Based on Statistical Calculations Related to Their Potential Energy Index. 2012. 19 p. URL: <http://arxiv.org/abs/1212.2628>
8. Edmundson R. S. Industrial Limestones and Dolomites in Virginia: James River District West of the Blue Ridge. Charlottesville, Virginia, 1958.
9. Le Maitre R. W. A Proposal by the IUGS Subcommittee on the Systematics of Igneous Rocks for a Chemical Classification of Volcanic Rocks Based on the Total Alkali – Silica (TAS) Diagram // Australian J. Earth Sci. 1984. Vol. 31. P. 243–255.

10. Шарпенко Л. Н., Костин А. Е., Кухаренко Е. А. Детализация диаграммы сумма щелочей – кремнезем (TAS) для химической классификации вулканических пород // Регион. геология и металлогения. 2008. № 35. С. 48–55.

11. Labushev M. M. Three Packets of Minerals of the Periodic Table of Chemical Elements and Chemical Compounds. 2013. 15 p. URL: <http://arxiv.org/abs/1304.1280>

Материал поступил в редколлегию 20.04.2017

M. M. Labushev

*Siberian Federal University
3 Vuzovskiy Lane, Krasnoyarsk, 660025, Russian Federation*

mlabushev@yandex.ru

FROM PARAMETRIC CLASSIFYING OF MINERALS AND ROCKS TO GENERAL SOLUTION OF THE PROBLEM OF CHEMICAL CLASSIFYING OF MATTERS

The article proposes classifying minerals and rocks based on the equations of information theory for the calculation of mutual information of two random variables. Eight atomic masses, which are extracted (using the Monte-Carlo method) from a large array of masses corresponding to the chemical composition of a mineral, form a 3-by-3 matrix, with the ninth element being their sum.

Information coefficient of proportionality is then defined as the algebraic sum of three uncertainties – the distribution of atomic masses by rows, columns and the entire matrix. Mean values and standard deviations of such distributions are used as the classification indicators. Calculations can be performed using an open source program, Agemark. Its description includes examples that illustrate basic classification possibilities, as well as the possibility of determining the relative age of rocks (through the example of limestones).

Keywords: classifying minerals and rocks, information theory, information coefficient of proportionality, relative age of rocks.

References

1. Yefremova S. V., Stafeev K. G. Petrochemical methods of rock study: Guidebook. Moscow, Nedra, 1985, 511 p. (in Russ.)
2. Petrographic Code of Russia: Magmatic, Metamorphic, Metasomatic, Impact Formations. 3rd ed. St. Petersburg, VSEGEI Publ., 2009, 200 p. (in Russ.)
3. Petrov T. G. RHA-method as a problem solution of systemizing analytical data on material composition of geological targets. *Otechestvennaya Geologiya*, 2008, № 4, p. 98–105. (in Russ.)
4. Labushev M. M. About Limit Numbers of Minerals, Inorganic and Organic Chemical Compounds. *Journal of Siberian Federal University. Engineering & Technologies*, 2008, № 3, p. 221–233. (in Russ.)
5. Labushev M. M. The Periodic Table as a Part of the Periodic Table of Chemical Compounds, 2011, 18 p. URL: <http://arxiv.org/abs/1103.3972>
6. Labushev M. M., Khokhlov A. N. A new approach to the relative age of igneous rocks determination. *Modern technologies of development of mineral resources. Collection of scientific works*. Krasnoyarsk, 2011, p. 64–68. (in Russ.)

7. Labushev M. M., Khokhlov A. N. Relative Dating and Classification of Minerals and Rocks Based on Statistical Calculations Related to Their Potential Energy Index, 2012, 19 p. URL: <http://arxiv.org/abs/1212.2628>.
8. Edmundson R. S. Industrial Limestones and Dolomites in Virginia: James River District West of the Blue Ridge. Charlottesville, Virginia, 1958, 133 p.
9. Le Maitre R. W. A Proposal by the IUGS Subcommittee on the Systematics of Igneous Rocks for a Chemical Classification of Volcanic Rocks Based on the Total Alkali – Silica (TAS) Diagram. *Australian J. Earth Sci*, 1984, vol. 31, p. 243–255.
10. Sharpenok L. N., Kostin A. E., Kukharenko E. A. Detailed elaboration of the Region. *Geology and Metallogeny*, 2008, no. 35, p. 48–55. (in Russ.)
11. Labushev M. M. Three Packets of Minerals of the Periodic Table of Chemical Elements and Chemical Compounds, 2013, 15 p. URL: <http://arxiv.org/abs/1304.1280>