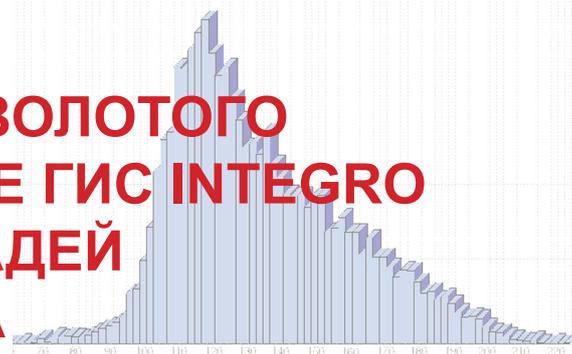


УДК: 553.411:681.3(470.6)

© Е.Д. Давыденко

*Е.Д. Давыденко*

# ПРОГНОЗИРОВАНИЕ ЗОЛОТОГО ОРУДЕНЕНИЯ В СРЕДЕ ГИС INTEGR0 НА ОДНОЙ ИЗ ПЛОЩАДЕЙ СЕВЕРНОГО КАВКАЗА



В настоящее время большая часть специалистов склоняется к мнению, что уже практически выявлена большая часть месторождений, которые выходят на земную поверхность и которые легко обнаружить стандартными методами. Поэтому основной целью прогноза и поисков становятся погребенные месторождения. Обнаружению таких объектов могут способствовать передовые космические и геоинформационные технологии, использующие математические методы обработки информации. Для прогнозирования геологических объектов по материалам дистанционного зондирования закрытых районов традиционно используется геоиндикационное дешифрирование, основанное на изучении возможных связей изображений земной поверхности с геологическими объектами непосредственно не наблюдаемыми [1]. В южном научном центре РАН на протяжении последних лет осуществляется внедрение технологии прогнозирования золоторудных месторождений по результатам математических трансформаций многозональных космических снимков. Конечным этапом этих преобразований является проведение кластер-анализа [2]. Представленная статья посвящена описанию концептуальной схемы математических процедур в среде ГИС INTEGR0 для геокластеризации признакового пространства в условиях естественных горных ландшафтов.

Основными этапами решения задачи прогнозирования с использованием кластер-анализа являются:

- формирование понятийной модели;
- подготовка массива исходных данных;
- создание числового массива признаков, описывающих (с регулярным шагом) всю исследуемую площадь;
- факторный анализ и выбор информативных факторов;
- кластер-анализ и выбор типов, соответствующих рудным образованиям.

*Понятийная модель* является основой для формирования набора прогностических признаков, полученных в результате преобразования число-

вых характеристик отражательной способности ландшафтов (зафиксированы на спектрально-анализируемом регионе). Основным положением используемой понятийной модели является рассмотрение отражательной способности (ОС) как производной от взаимодействия эндо- и экзофакторов. В состав группы экзокомпонент ОС включаем высотную зональность ландшафта, различные виды растительности (хвойные, лиственные леса, луга) в пределах одной зоны, а также рассматриваем различную степень освещенности территорий (затененные и ярко освещенные участки) при конкретных азимуте и угле наклона солнца в момент съемки. Последнее не является непосредственно ландшафтной особенностью, но влияние этого фактора особенно значимо в горных районах с интенсивной расчлененностью рельефа. Кроме того, на отражательной способности ландшафта также сказывается ориентация склонов относительно сторон света, так как южные склоны вследствие более длительного воздействия солнечных лучей получают и сохраняют больше тепла, чем северные.

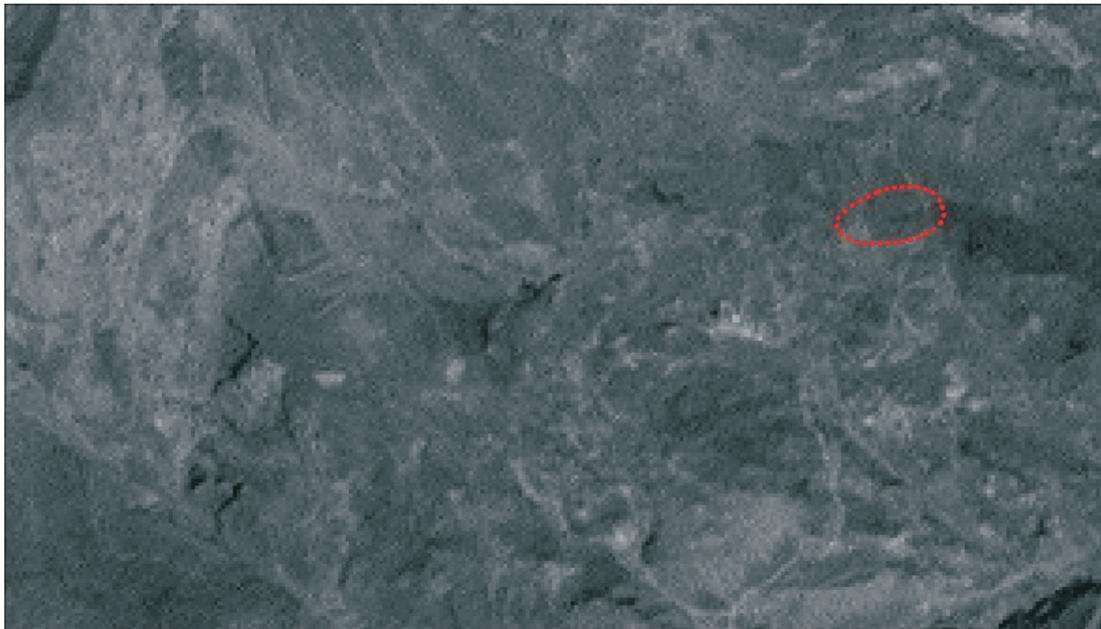
Переходя к рассмотрению вопроса об эндокомпонентах принимаемой модели, следует отметить, что особенности геологического пространства влияют на газово-геохимическую среду почвенного слоя, в связи с этим изменяются яркости окраски растительности [1]. Эти изменения обуславливают отображение эндофакторов в отражательной способности ландшафта. По результатам предшествующих геологических работ большая часть площади перекрыта отложениями терригенной толщи юрской системы. На юге широко распространены гранитоиды палеозойского возраста. Выходы трахитов и риолитов юрского вулканического комплекса наблюдаются в центральной части участка. Рудообразование связано с процессом формирования вулканического комплекса. Критерии, позволяющие априори определить возможность отображения элементов геологического строения в отражательной способности ландшафта, пока не установлены. Поэтому решение этих вопросов осуществляется эмпирически. Совокупность

эндо- и экзокомпонент рассматривается в качестве понятийной модели признакового пространства.

В качестве исходных данных использованы космоснимки (цифровые) с американского спутника Landsat-7 (рис. 1), выполненные в восьми зонах спектра по нескольким годам (август 1999, июнь 2000, сентябрь 2002, май 2003 года), и цифровая модель рельефа (рис. 2) SRTM (Shuttle Radar Topography

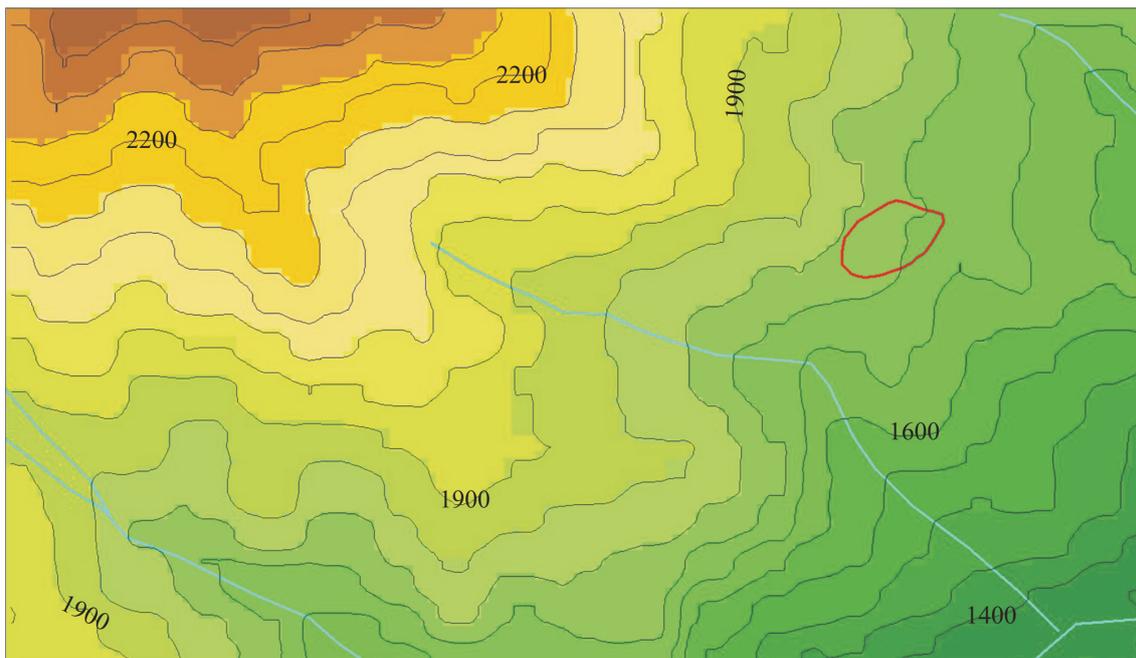
Mission). Математическая обработка осуществлялась в среде ГИС INTEGR0.

Преобразование исходных данных является многоэтапным процессом и осуществляется с целью приведения их к стандартному виду и повышению информативности. Так стандартизация (нормирование и центрирование) используется для уничтожения различий в средних и дисперсии.



1

Рис. 1. Панхроматическое изображение исследуемой площади; 1 – границы эталонной области (выявленный ранее рудный участок)



а б 1 2 3

Рис. 2. Цифровая модели рельефа SRTM; 1 – перепад высот (а – максимум, б – минимум); 2 – реки; 3 – границы эталонной области

Гистограммы спектральных снимков данного района представляют собой несимметричное распределение (рис. 3), со значительной вытянутостью одной (обычно правой) ветви. В различных руководствах [2, 7, 11] по использованию математических методов рекомендуется нормализация признаков, то есть приведение функции распределения значений признака к нормальному виду. Указывается, что результаты классификаций для признакового пространства с нормализованным видом функций распределения более оптимальны, чем в случае иного вида распределения. Для нормализации исходного распределения нужно добиваться его симметрии, применяя различные функциональные преобразования. В случае левосторонней асимметрии рекомендуется применять такие функции как натуральный логарифм или квадратный корень, а правосторонней – экспоненту или возведение в квадрат. Выполненные преобразования согласно этим указаниям не дали заметных результатов, и поэтому были предприняты попытки преобразования только несимметричной части. Но в данном случае, для снимков конкретной территории, нормализация только ухудшает сигнал. Указанное связано с тем, что площади аномальных значений отражательной способности, соответствующие рудным телам, имеют небольшое распространение и на гистограмме располагаются в так называемом «хвосте». То есть для подобных случаев процедура нормализации оказалась неприемлемой.

На следующем этапе были использованы преобразования данных, задачей которых являлось повышение информативности характеристик исходных данных. Так нормирование признаков освещенностью уменьшает дисперсию (табл. 1) и корреляцию признаков с освещенностью (табл. 2). Таким образом, достигается определенное гашение помехи свет-тень в изображении ландшафта.

В связи с тем, что различные типы растительности претерпевают своеобразные изменения в течение вегетационного периода, для набора информативных признаков использованы снимки в различные сезоны. Это позволяет выделить информацию, связанную с различными типами растительности.

Также для характеристик числовой модели рельефа с указанной целью проведена частотная фильтрация исходных значений (сглаживание с различными параметрами). Это позволило получить тренд рельефа (описывает вертикальный размах, и, тем самым, высотную зональность ландшафтов), «площадной рельеф» (крупные формы), «локальный рельеф» (мелкие особенности) и градиент рельефа (резкое изменение отметок рельефа).

*Массив признаков* для описания особенностей отражательной способности ландшафтов, созданный с помощью указанных преобразований, включает 29 характеристик:

- Признаки, полученные из снимков 1999 года, нормированных стандартом и освещенностью (8 признаков).

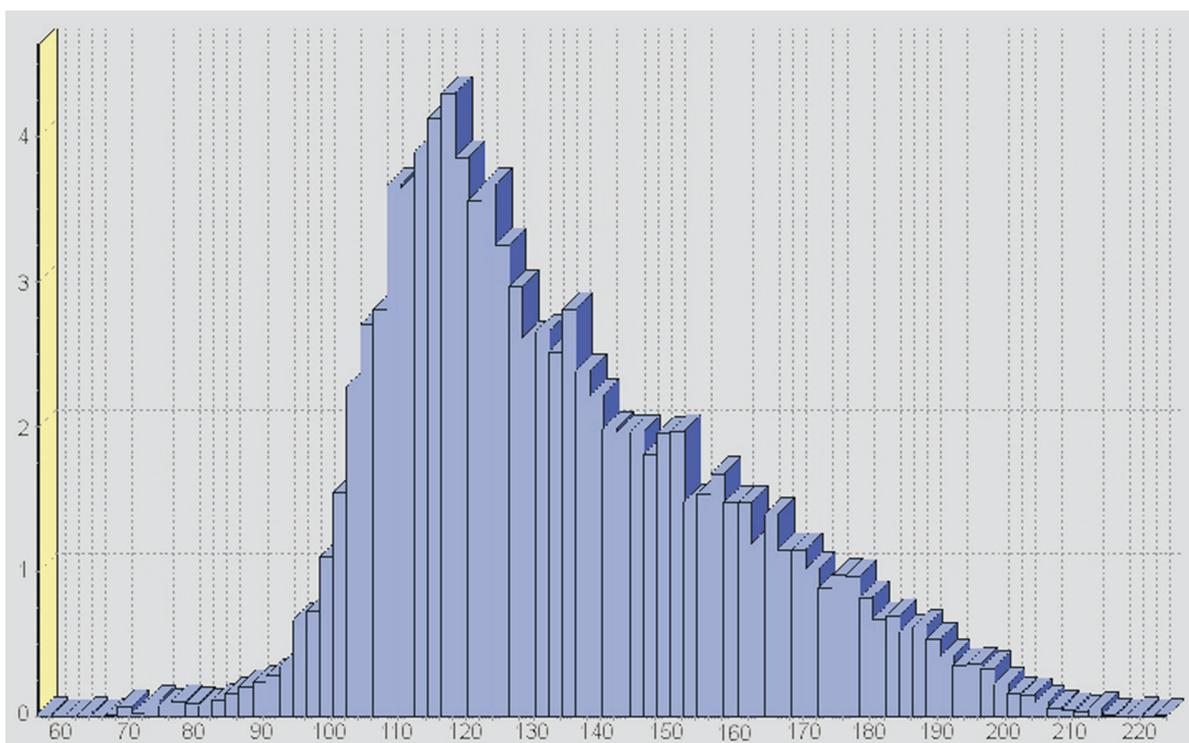


Рис. 3. Несимметричное распределение значений на примере исходного космоснимка в инфракрасной зоне спектра (гистограмма из ГИС INTEGR0)

Таблица 1

**Значения дисперсии признаков до нормирования освещенностью и после**

Съемка в спектральном диапазоне:	Дисперсия признака до нормирования освещенностью	После нормирования освещенностью
0,45-0,52 (синий-голубой)	1,00051	0,05682
0,52-0,6 (зеленый-желтый)	0,99957	0,0555
0,63-0,69 (красный)	1,00001	0,0571
0,76-0,9 (ближний ИК)	0,99983	0,05997
1,55-1,75 (коротковолновый средний ИК)	1,00018	0,05643
10,42-12,5 (дальний ИК)	1,00066	0,07066

**Примечание:** рассматриваются признаки и освещенность уже нормированные стандартом.

Таблица 2

**Коэффициенты корреляции признаков и освещенности до гашения помехи свет-тень и после гашения (таблица корреляции, полученная в программном комплексе ГИС INTEGRO)**

	10н	20н	30н	освб	10н_ос	20н_ос	30н_ос
10н	0.3000372	0.3639223	0.3074457	0.0192288	0.0884293	0.1510265	0.0884293
20н	0.3639223	0.3000372	0.3074457	0.0192288	0.0884293	0.1510265	0.0884293
30н	0.3074457	0.3639223	0.3000372	0.0192288	0.0884293	0.1510265	0.0884293
освб	0.3000372	0.3639223	0.3074457	0.0192288	0.0884293	0.1510265	0.0884293
10н_ос	0.0192288	0.0884293	0.1510265	0.0192288	0.0884293	0.1510265	0.0884293
20н_ос	0.0884293	0.0192288	0.1510265	0.0884293	0.0192288	0.0884293	0.1510265
30н_ос	0.1510265	0.0884293	0.0192288	0.1510265	0.0884293	0.0192288	0.1510265

**Примечание:** к обозначению признака (10, 20, 30) добавлена буква «н» – признак нормирован средне-квадратическим отклонением; добавлено «н\_ос» – признак дополнительно нормирован освещенностью.

- Признаки, описывающие растительный ландшафт – разности 2003 и 1999, 2000 и 2002 (всего 16 признаков).
- Четыре признака описывающие рельеф (фондовый, площадной, локальный рельеф; градиент).
- Ориентация склонов относительно сторон света.

Значительная часть из перечисленных признаков одновременно содержит информацию о большом количестве компонент (и экзо-, и эндокомпонент), влияющих на яркости ландшафта. Поэтому после создания признакового пространства принято решение обратиться к процедуре факторного анализа, чтобы подразделить используемые признаки на эндо- и экзофакторы.

В общем виде факторную модель можно записать следующим образом:

$$x = f^* a + e,$$

где  $x$  – признаки;  $f$  – факторы;  $a$  – матрица факторных нагрузок (весов);  $e$  – остатки (не коррелируют с факторами).

Согласно факторной модели, признак приближается фактором, но не точно – присутствует остаток. Обычно считается, что остаток состоит из двух некоррелированных частей, одна из которых является погрешностью измерений, а вторая – это специфическая часть исходной информации. Специфическая часть содержит информацию о таких особенностях признака, как рельеф, освещенность, растительность, степень накопления тепла склонами, а также об эндогенных компонентах. Если использовать исходные данные для факторного анализа, то наиболее четко в факторах будут проявляться экзогенные составляющие ландшафта. Описанное выше расширение признакового пространства выполнено для того, чтобы снизить специфичность «помех» и осуществить разделение экзо- и эндокомпонент (добавление в набор для факторного анализа переменной, тесно связанной с какой-либо «особенностью», приводит к уменьшению специфичной части, обусловленной этой переменной).

Факторный анализ предполагает два этапа. Метод главных компонент (компоненты могут рассматриваться как факторы) является первым этапом факторного анализа. Изначально признаковое пространство под влиянием экзо- и эндопроявлений характеризуется анизотропией в различных направлениях многомерного признакового пространства. Метод главных компонент позволяет спроектировать это пространство на новые оси (главные компоненты), которые оптимально адаптированы к анизотропии этого пространства. Если рассматривать на примере двумерного пространства признаков, то главные компоненты описывают две оси эллипса, первая компонента соответствует наибольшей оси, а вторая ортогональна ей. Все исходные признаки проецируются на эти ортогональные оси. Новые признаки (компоненты, факторы) являются линейными комбинациями признаков с коэффициентами, которые можно рассматривать как веса, с которыми признаки вошли в фактор. Таким образом, чем больше значения весов, тем большее влияние оказывают соответствующие признаки на фактор [7, 11]. Как указывалось выше, общее число используемых признаков составило 29 (т.е. имеем дело с 29-мерным пространством).

Следующим этапом является вращение факторных осей (фактически в ГИС INTEGRO первый этап автоматически переходит во второй). Такое преобразование рекомендуется в случаях, когда интерпретация факторов вызывает затруднения (довольно распространенная ситуация). Процедура вращения факторов ориентирована на упрощение структуры матрицы факторных нагрузок. При этом осуществляется поворот системы координат таким образом, чтобы проекции малоинформативных факторов на новые оси стремились к нулю, а информативных – приблизились к максимальным значениям. Вращение может быть ортогональным и косоугольным.

*Интерпретация факторов* должна быть выполнена с целью выбора эндофакторов. Подразделение полученных 15 факторов на эндо- и экзо- составляющие в первую очередь осуществляется по нагрузкам исходных признаков на факторы. Кроме того, целесообразно получить коэффициенты корреляции факторов и признаков, которые могут оказаться такими же полезными, как и веса (нагрузки) признаков – высокие значения коэффициентов укажут на признаки, повлиявшие на фактор в наибольшей степени (рис. 4).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
10_но_нц	0.075	0.050	0.087	0.108	0.067	-0.210	-0.039	-0.075	-0.031	0.213	-0.211	-0.036	-0.447	0.093	0.526	-0.125	1.461	-0.122
20_но_нц	0.068	0.068	0.096	0.152	0.028	-0.207	-0.079	-0.101	-0.027	0.009	-0.423	-0.041	-0.421	0.215	0.544	-0.146	-0.691	-0.201
30_но_нц	0.079	0.054	0.101	0.095	-0.004	-0.145	-0.036	-0.030	-0.011	0.109	-0.195	-0.108	-0.175	0.111	0.324	-0.072	-0.037	0.927
50_но_нц	0.067	0.014	0.109	0.052	-0.125	-0.200	-0.160	-0.162	-0.215	-0.475	0.050	0.079	0.768	0.006	-0.528	0.174	-0.203	-0.333
70_но_нц	0.080	0.032	0.114	0.043	-0.067	-0.126	-0.076	-0.052	-0.079	-0.054	0.135	-0.045	0.402	0.032	-0.145	0.044	-0.027	0.406
40_нц	-0.071	0.016	-0.034	0.108	-0.068	0.011	-0.133	-0.210	0.064	-0.784	-0.734	0.135	-0.039	0.180	-0.617	0.022	0.551	0.082
61_нц	0.073	0.003	0.048	-0.105	-0.163	0.124	0.223	0.282	0.236	-0.437	0.107	0.031	-0.290	0.237	0.201	-0.501	0.088	0.048
62_нц	0.073	0.003	0.048	-0.105	-0.163	0.124	0.222	0.280	0.235	-0.437	0.110	0.029	-0.297	0.228	0.200	-0.508	0.075	0.086
10разн_ис	0.055	0.127	-0.139	-0.026	-0.031	0.027	-0.127	-0.072	0.081	-0.102	0.164	0.026	-0.201	0.082	0.198	0.994	-0.092	-0.078
20разн_ис	0.058	0.111	-0.151	-0.059	0.034	0.027	-0.108	-0.005	0.032	-0.135	0.056	-0.074	-0.095	0.034	0.241	0.950	-0.143	-0.023
30разн_ис	0.058	0.100	-0.154	-0.095	0.033	0.080	-0.098	-0.143	0.075	-0.132	0.161	0.022	-0.117	-0.112	0.080	0.567	0.129	-0.192
50разн_ис	0.049	0.113	-0.044	-0.167	0.205	0.122	0.098	-0.129	-0.129	0.167	-0.461	-0.112	0.422	0.130	-0.064	-1.024	-0.145	-0.072
70разн_ис	0.049	0.116	-0.107	-0.168	0.117	0.128	0.004	-0.295	-0.035	-0.008	-0.088	0.080	0.252	0.001	-0.191	-0.929	0.190	0.431
40разн_ис	0.015	-0.080	-0.113	-0.037	0.268	-0.155	0.095	0.656	-0.276	-0.258	-0.453	-0.430	0.418	0.276	0.440	0.459	0.051	-0.096
61разн_ис	-0.019	-0.010	0.228	-0.190	0.157	0.158	0.096	-0.184	0.100	-0.074	-0.174	0.039	0.023	0.081	0.238	0.825	-0.071	-0.044
62разн_ис	-0.020	-0.011	0.228	-0.189	0.157	0.155	0.098	-0.190	0.095	-0.069	-0.179	0.040	0.016	0.074	0.235	0.767	-0.044	-0.032
10разн_ам	-0.064	0.128	0.010	0.003	-0.049	-0.186	0.165	0.016	-0.009	0.010	0.154	0.226	0.194	0.752	0.299	-0.008	-2.379	0.732
20разн_ам	-0.064	0.115	0.000	-0.045	-0.034	-0.213	0.242	0.047	0.062	0.103	0.269	0.188	0.319	0.767	-0.106	0.231	1.961	1.045
30разн_ам	-0.071	0.121	0.006	-0.022	-0.006	-0.153	0.122	-0.011	0.008	-0.020	0.070	0.139	0.027	0.378	0.185	-0.111	0.569	-2.551
50разн_ам	-0.064	0.128	0.001	-0.026	0.017	-0.146	0.174	0.125	0.079	-0.068	-0.244	-0.133	-0.287	-0.889	-0.436	0.193	0.072	2.049
70разн_ам	-0.062	0.138	0.011	-0.004	-0.010	-0.136	0.141	0.076	0.090	-0.094	-0.214	-0.090	-0.313	-0.787	0.034	-0.237	-0.512	-1.389
40разн_ам	0.064	0.087	0.118	-0.075	-0.053	-0.110	0.130	0.218	0.004	0.098	0.254	-0.301	0.447	-0.851	-0.560	0.236	0.091	-0.684
61разн_ам	-0.036	0.113	0.096	0.144	0.194	0.188	-0.310	0.253	0.135	-0.126	0.309	-0.144	0.035	0.108	-0.052	-0.322	0.086	0.169
62разн_ам	-0.036	0.114	0.096	0.144	0.196	0.187	-0.306	0.254	0.134	-0.136	0.307	-0.140	0.034	0.121	-0.052	-0.283	0.079	0.081
Фон_нц	-0.075	0.018	0.003	-0.060	-0.188	0.040	-0.126	-0.177	-0.254	-0.271	0.151	-0.332	0.403	-0.560	1.785	-0.337	0.434	0.488
Лок_нц	-0.009	-0.004	0.030	-0.300	-0.085	-0.174	-0.504	0.382	0.140	0.187	-0.281	0.905	0.054	-0.196	0.170	-0.156	0.054	0.092
Площ_нц	-0.040	-0.007	0.030	-0.286	-0.162	-0.150	-0.310	-0.055	-0.002	0.101	-0.001	-1.110	-0.380	0.594	-0.545	-0.147	-0.113	-0.044
Осв_ю_нц	0.000	0.076	-0.006	0.146	-0.320	0.310	0.003	0.163	0.404	0.469	-0.646	-0.248	0.577	0.179	0.138	0.406	0.067	-0.162
Град_нц	-0.007	0.103	0.069	-0.011	-0.170	0.298	0.021	0.213	-0.981	0.099	-0.098	0.222	-0.403	0.147	-0.320	0.237	0.135	-0.002

Рис. 4. Фрагмент таблицы факторных нагрузок. Пример интерпретации факторов по значениям весов в таблице (12-й фактор интерпретирован как особенности рельефа, а 18-й – особенности растительности)

К сожалению, часто получаем факторы, принадлежность которых к какой-то группе трудно определить по коэффициентам и поэтому приходится проводить визуальный анализ, сопоставляя их с изображениями ландшафтов и геологической картой. В качестве примера на рис. 5 представлен фактор, который в значительной мере коррелируется с площадью распространения гранитов.

В сложных случаях следует сравнивать результаты кластеризации признаков пространств как с включением спорного фактора, так и без него. В результате выполненного анализа в качестве значимых эндогенных факторов принято 11 факторов. После выбора эндогенных факторов появляется возможность приступить к прогнозированию рудных тел, используя кластер-анализ.

*Кластер-анализ* космических снимков для решения задач геологии можно рассматривать как переход от геоиндикационного дешифрирования к кластер-дешифрированию. Осуществить районирование площади (классификацию элементарных площадок) по 11 признакам без применения математических методов невозможно. Построение кластеров (или иначе, групп, классов, районов) является эффективным способом представления сложных, многомерных соотношений между объектами.

Для использования кластерного анализа необходимо пространство информативных признаков, которые обязательно несут полезную информацию, хотя могут содержать и шумы. Изучаемая площадь должна быть подразделена на множество элементарных площадок, каждая из которых описана набором признаков. Согласно этому требованию было использовано исходное разбиение изображений соответственно разрешению снимков. Основными элементами кластерного анализа являются понятие меры удаленности-близости объектов и алгоритм группирования объектов на основе этого понятия. Для пояснения используемого алгоритма обычно предлагается пример, когда множество объектов по двум признакам подразделяется на два кластера объектов (два отдельных облака точек), в каждом из которых расстояние между объектами-точками внутри кластера меньше, чем расстояние между объектами-точками разных кластеров (расстояние часто вычисляется как пифагорово). Алгоритм, основанный на этом принципе, используется и для многомерных признаков пространств. Конечным результатом анализа является определение принадлежности элементарных площадок к определенным типам из общего набора выделенных здесь типов (кластеров) [2, 7, 11].



Рис. 5. Пример фактора, выбранного в качестве эндогенного (описывает площади развития гранитов).  
1 – площади гранитов в соответствии с геологической картой; 2 – аномальные значения фактора

В результате кластеризации отобранных одиннадцати факторов средствами программного комплекса ГИС INTEGRО было получено 83 типа. Площадь эталонного объекта (рудного тела, выделенного по результатам разведки) подразделяется по этим данным на 2 типа (рис. 6). Эти же типы, полученные в результатах расчетов (кластеризации), за пределами рудного тела рассматриваем в качестве потенциальных золоторудных объектов на прилегающих площадях. В результате получены сведения о возможном распространении прогнозируемых рудных объектов, которые в основном не противоречат информации о повышенных содержаниях золота в пробуренных ранее скважинах.

**Ключевые слова:** дистанционное зондирование, ГИС INTEGRО, прогнозирование золотого оруденения.

**ЛИТЕРАТУРА**

1. Аэрокосмические методы геологических исследований / под ред. А.В. Перцова. – СПб. : Изд-во СПб картфабрики ВСЕГЕИ. – 2000. – 316 с.
2. Матишов Г.Г., Парада С.Г., Давыденко Д.Б. Выбор рудоперспективных площадей по результатам дистанционной флюидоиндексации как фактор рационального недропользования в условиях горных территорий // Устойчивое развитие горных территорий. – 2010. – №3 (5). – С. 76-83.
3. Бугаец А.Н., Дуденко Л.Н. Математические методы при прогнозировании месторождений полезных ископаемых. – Л. : Недра, 1976. – 276 с.
4. Дэвис Дж.С. Статистический анализ данных в геологии : в 2 кн. – М. : Недра, 1990. – Кн. 1 – 319 с. Кн. 2 – 427 с.
5. Миллер Р., Кан Дж. Статистический анализ в геологических науках. – М. : Мир, 1965. – 481 с.

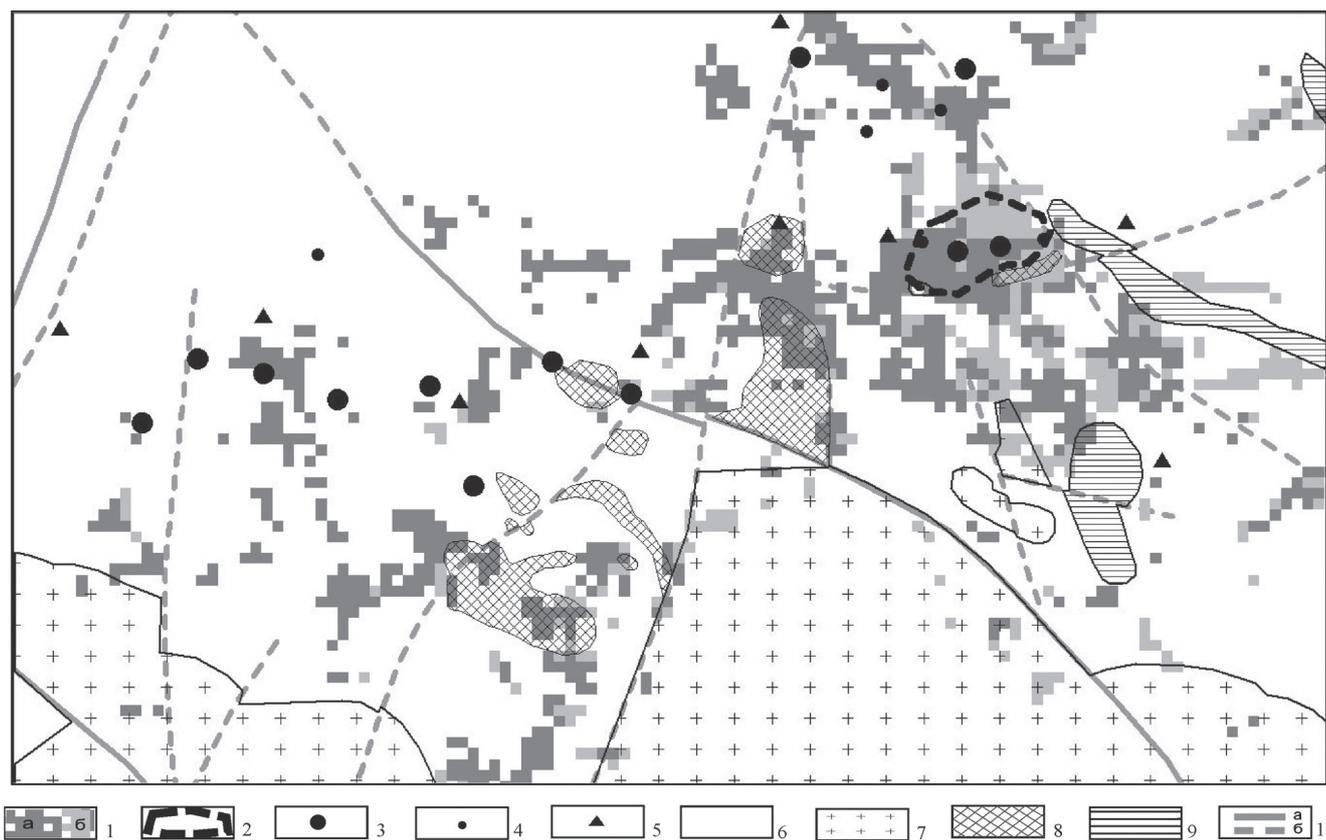


Рис. 6. Выделение площадей, перспективных на обнаружение золото-полиметаллического оруденения, по результатам кластер-анализа:

- 1 – площадь (кластеры), перспективные по результатам кластер-анализа на обнаружение золото-полиметаллического оруденения (а – богатые, б – бедные);
- 2 – проекция рудного тела (эталон);
- 3-5 – скважины с различным содержанием золота (3 – 1-10 г/т, 4 – от 0,3 до 1 г/т, 5 – менее 0,3 г/т);
- 6 – юрский вулканический комплекс; 7 – граниты; 8 – риолиты; 9 – трахиты;
- 10 – тектонические нарушения (а – установленные, б – предполагаемые)