

## Физико-географические исследования

УДК 551.4.013(47-924)

*А.Ф. Гилязов*

### КЛАСТЕРНЫЙ АНАЛИЗ КАК ИНСТРУМЕНТ РАЙОНИРОВАНИЯ ТЕРРИТОРИИ ПО КРУПНОСТИ РЕЧНЫХ НАНОСОВ (НА ПРИМЕРЕ БАССЕЙНА ВОЛГИ)

Произведено районирование бассейна Волги по крупности взвешенных наносов и донных отложений с использованием двух (наиболее часто применяемых в исследованиях) методов кластерного анализа – агломеративного метода иерархического подхода и метода  $k$ -средних вероятностного подхода. Несмотря на различие в алгоритмах этих методов, в обоих случаях получаются непротиворечащие друг другу картины. Метод  $k$ -средних стремится уменьшить разнородность внутри групп, из-за чего могут формироваться территориально разобщенные однородные классы; агломеративный метод в противовес ему отказывается от деталей и формирует компактные группы с более пестрым содержанием. На картах, построенных по данным этих разновидностей кластерного анализа, можно проследить влияние некоторых факторов на крупность речных наносов.

*Ключевые слова:* кластерный анализ, районирование, крупность речных наносов, взвешенные наносы, донные наносы, бассейн Волги.

Крупность речных наносов обуславливается влиянием множества факторов. Значение этих факторов неодинаково: в разных природно-антропогенных условиях одни являются главными, определяющими основные черты гранулометрического облика речных наносов, другие же создают более тонкие вариации этих основных черт [1]. Есть основание полагать, что на каждой отдельно взятой территории формируется свой уникальный комплекс природных и антропогенных факторов; а это, в свою очередь отражается на крупности переносимого и накапливаемого реками материала: даже в смежных бассейнах могут формироваться различные по гранулометрическому составу наносы. При картировании подобных явлений можно получить весьма пеструю, мозаичную структуру исследуемой территории. Подобные картины сложны для интерпретации даже опытному исследователю. Поэтому в географических исследованиях нередко возникает необходимость в обобщении полученных результатов, абстрагировании от деталей и генерализации исходного материала с одной лишь целью – получить наиболее наглядную и легко интерпретируемую схему районирования исследуемой территории по изучаемому признаку. В последнее время среди большого разнообразия методов районирования широкое применение получил кластерный анализ.

Основная цель анализа состоит в выделении в исходных многомерных данных таких однородных подмножеств, в которых объекты были похожи в известном смысле друг на друга, а объекты из разных групп – не похожи. Под «похожестью» понимается близость объектов в многомерном пространстве признаков, и тогда задача сводится к выделению в этом пространстве естественных скоплений объектов, которые и считаются однородными группами [2].

Кластерный анализ активно используется при изучении различных объектов, входящих в область интересов наук о Земле. В.А. Кузьмичёнок [3] отмечает, что использование математического аппарата кластерного анализа для выделения типов внутригодового распределения атмосферных осадков привело к удовлетворительным результатам. В.И. Пономарев с соавторами [4] заключают, что применение кластерного анализа дает возможность выделить различные совокупности (группы) близких по структуре временных рядов, характеризующих климатическую систему, определить, как изменяется содержание групп от одного исторического периода к другому, а также установить взаимосвязи между климатическими параметрами. Анализ дендрограмм, внутрирайонных коэффициентов корреляции Спирмэна и вероятностей внутри- и межкластерных связей с учетом природной зональности позволил М.Н. Бубину [5] установить существование гидрологических районов с синхронными многолетними колебаниями зимнего стока рек и выполнить районирование территории Челябинской области. Ю.О. Кочеткова [6; 7] применила кластерный анализ с целью группировки местностей по совокупности эколого-геоморфологических показателей, влияющих на размещение населения, и выделила территории, различающиеся степенью комфортности проживания в их пределах. По данным А.В. Чашкова и В.М. Киселева [8], кластеризация данных позволяет заметно повысить надежность и достоверность определения литотипов, а также пористости и проницаемости пород.

Реализация кластерного анализа в вышеперечисленных работах осуществлялась с применением различных алгоритмов. В.Б. Бериков и Г.С. Лбов [9] отмечают, что общепринятой классификации методов кластерного анализа не существует, но можно выделить несколько подходов кластеризации.

В настоящей работе предпринята попытка оценки возможности использования двух (наиболее часто применяемых в исследованиях) методов кластерного анализа – агломеративного метода иерархического подхода и метода  $k$ -средних вероятностного подхода – для районирования бассейна Волги по крупности взвешенных наносов и донных отложений.

Перед началом кластеризации агломеративным методом все объекты считаются отдельными кластерами, в ходе алгоритма они объединяются. Вначале выбирается пара ближайших кластеров, которые объединяются в один кластер. В результате количество кластеров становится равным  $N - 1$ , где  $N$  – общее число классифицируемых объектов. Процедура повторяется, пока все классы не объединятся [10].

В кластерном анализе для количественной оценки сходства вводится понятие метрики. Сходство или различие между классифицируемыми объектами устанавливается в зависимости от метрического расстояния между ними. Если каждый объект описывается  $j$  признаками, то он может быть представлен как точка в  $j$ -мерном пространстве и сходство с другими объектами будет определяться как соответствующее расстояние. Существует множество способов определения близости между объектами: линейное расстояние, евклидово расстояние, квадрат евклидова расстояния, обобщенное степенное расстояние Минковского, расстояние Чебышёва, расстояние городских кварталов (Манхэттенское расстояние) [11; 12]. Забегая вперед, отметим, что в работе дистанция между кластерами рассчитывалась с применением евклидова расстояния, которое является самой популярной метрикой в кластерном анализе. Оно попросту является геометрическим расстоянием в многомерном пространстве. Геометрически оно лучше всего объединяет объекты в шарообразных скоплениях [11; 12].

На первом шаге, когда каждый объект представляет собой отдельный кластер, расстояния между этими объектами определяются выбранной мерой. Однако необходимо правило объединения или связи для двух кластеров. Выделяют следующие методы объединения кластеров: расстояние «ближайшего соседа» (одиночная связь), расстояние «дальнего соседа» (полная связь), невзвешенное попарное среднее, взвешенное попарное среднее, невзвешенный центроидный метод, взвешенный центроидный метод (медиана), метод Варда. В работе был использован метод Варда, отличающийся от всех других методов, поскольку он использует методы дисперсионного анализа для оценки расстояний между кластерами. В этом методе в качестве целевой функции применяют внутригрупповую сумму квадратов отклонений, которая есть не что иное, как сумма квадратов расстояний между каждой точкой (объектом) и средней по кластеру, содержащему этот объект. На каждом шаге объединяются такие два кластера, которые приводят к минимальному увеличению целевой функции, то есть внутригрупповой суммы квадратов. Этот метод направлен на объединение близкорасположенных кластеров. В целом метод представляется очень эффективным, однако он стремится создавать кластеры малого размера [11; 12]. Главными достоинствами настоящего метода являются ее наглядность и отсутствие необходимости знания количества выделяемых классов. Этот метод удобен для оценки структуры данных на наличие разнородных групп. Наглядность этого метода достигается построением дендрограмм – древовидной диаграммы, содержащей  $n$  уровней, каждый из которых соответствует одному из шагов процесса последовательного укрупнения кластеров [11]. Количество выделяемых групп можно определить по агломерационному графику: существенное увеличение дистанции между объединяемыми кластерами говорит о том, что сливаются непохожие друг на друга классы; это и служит критерием определения числа кластеров.

Сущность метода  $k$ -средних заключается в разбиении множества элементов векторного пространства на заранее известное число кластеров  $k$ . На каждой итерации перевычисляется центр масс для каждого кластера, полученного на предыдущем шаге, затем векторы разбиваются на кластеры вновь в соответствии с тем, какой из новых центров оказался ближе по выбранной метрике. Алгоритм завершается, когда на какой-то итерации не происходит изменения центра масс кластеров [13]. Достоинствами этого метода являются относительная простота и быстрдействие алгоритма, а также возможность обработки больших статистических совокупностей. Недостаток этого метода заключается в том, что здесь необходимо заранее задавать количество кластеров. Эту проблему можно преодолеть проведением иерархического анализа со случайно отобранной выборкой наблюдений и, таким образом, определить оптимальное количество кластеров.

## Материалы и методы исследований

Источниками информации о гранулометрическом составе речных наносов выступили «Государственный водный кадастр СССР. Основные гидрологические характеристики» и «Ресурсы поверхностных вод СССР. Основные гидрологические характеристики» [14-22]. В вышеперечисленных справочниках данные о гранулометрическом составе наносов представлены в различные фазы водного режима (межень, летне-осенние паводки, а также подъем, пик и спад половодья) тремя характерными составами – крупный, средний и мелкий, – определяемыми визуально по интегральным кривым гранулометрического состава наносов [1]. «Ресурсы поверхностных вод СССР. Основные гидрологические характеристики» выпущены и содержат информацию за два различных периода: до 1962 г. и 1963–1970 гг. «Государственный водный кадастр. Основные гидрологические характеристики» содержит информацию за 1971-1975 гг. Каждый последующий (новый) выпуск учитывает информацию о наблюдениях за предыдущие периоды.

В настоящей работе крупность наносов выражается медианным диаметром ( $M_d$ ). Определение медианного диаметра производилось по формуле квантилей Бернштейна [23]:

$$M_d = x_1 + k \left( \frac{N_p - \sum f_{i-1}}{f_i} \right), \quad (1)$$

где  $M_d$  – величина искомого показателя,  $x_1$  – нижняя граница интервала, содержащего  $x_p$ ,  $k$  – размер интервала,  $N_p$  – 50 %,  $\sum f_{i-1}$  – суммарная частота ниже интервала, содержащего  $x_p$ ,  $f_i$  – частота интервала, содержащего  $x_p$ .

В работе были применены следующие критерии отбора к исходному материалу:

- 1) площадь водосбора 2 000–50 000 км<sup>2</sup>, что соответствует понятию средней реки, режим которой формируется под влиянием зональных факторов<sup>1</sup>;
- 2) фазы водного режима – подъем половодья, спад половодья или межень;
- 3) средний характерный состав наносов;
- 4) количество произведенных измерений гранулометрического состава в каждой из фаз водного режима – не менее трех.

Из 275 постов в бассейне Волги, в которых производились наблюдения за гранулометрическим составом наносов, после выполнения вышеперечисленных условий было отобрано 112 постов.

Районирование территории производилось по двум признакам – медианному диаметру наносов и географической близости водосборов друг от друга. Как уже было сказано ранее, крупность наносов – величина дискретная и даже в соседних водосборах может существенно отличаться, тогда как медианные диаметры наносов в отдаленных друг от друга водосборах могут иметь существенное (но случайное) сходство. По этой причине вводится признак территориальной близости водосборов: с одной стороны, он способствует объединению в один кластер (т. е. район в терминологии географического районирования) смежных водосборов (пусть даже с некоторым различием гранулометрического состава наносов), а с другой – сходные по крупности наносов, но отдаленные друг от друга водосборы распределяет по разным кластерам (районам). Географическая близость выражалась через широту и долготу геометрических центров водосборов. Поскольку эти данные имеют различный масштаб, возникает необходимость в их стандартизации (нормировании), что в кластерном анализе имеет большое значение. Стандартизация приводит значения всех преобразованных переменных к единому диапазону значений путем выражения через отношение этих значений к некоей величине, отражающей определенные свойства конкретного признака [11]. В предлагаемой работе нормирование осуществлялось по формуле

$$z = \frac{a_i - a_{\min}}{a_{\max} - a_{\min}}, \quad (2)$$

где  $z$  – стандартизованная величина,  $a_i$  – исходное значение признака,  $a_{\min}$  – минимальное значение признака в выборке,  $a_{\max}$  – максимальное значение признака в выборке. Иными словами, минимальное значение признака приравнивалось к нулю, а максимальное – к единице.

<sup>1</sup> ГОСТ 19179-73. Гидрология суши. Термины и определения.

Но поскольку удаленность водосборов друг от друга как признак характеризуется двумя параметрами – широтой и долготой, а крупность наносов определяется одной величиной – медианным диаметром, то возникает необходимость уравнивать значимость (вклад) обоих признаков. С учетом выбора способа определения близости между объектами (евклидово расстояние) был определен масштаб нормирования признаков, положенных в основу кластерного анализа: медианный диаметр изменяется от 0 до  $\sqrt{2}$ , а параметры признака удаленности – широта и долгота – от 0 до 1.

### Результаты и их обсуждение

На основе проведенного кластерного анализа для среднего характерного состава наносов каждой фазы водного режима были построены карты, отражающие области формирования различных по крупности речных наносов. Кластеризация агломеративным методом выявила от 7 до 8 разнородных групп (классов). Количество задаваемых классов для метода  $k$ -средних соответствовало количеству групп, выделенных в результате проведения иерархического алгоритма (т. е. от 7 до 8); иными словами, агломеративный метод играл рекогносцировочную роль для метода  $k$ -средних.

Следует отметить, что внутри выделенных групп разброс крупности наносов может быть весьма ощутимым; однако при этом всегда можно выделить группы, размеры частиц в которых не сильно отличаются между собой (табл. 1–3). Все это объясняется влиянием признака пространственной близости, заложенного в кластерный анализ на равных условиях с медианным диаметром. Именно этот признак затушевывает различие в крупности наносов, но в то же время способствует группировке по территориальному признаку.

Поскольку основное количество взвешенных наносов проходит в половодный период, то и районирование следует проводить по данным за этот период (подъем и спад половодья). На карте районирования бассейна Волги по крупности взвешенных наносов в подъем половодья по данным агломеративного метода (рис. 1, а) выделяются 7 разнородных областей. Следует отметить, что здесь и на всех последующих картах области (районы) выделялись по преобладанию бассейнов, принадлежащих одному кластеру. По этой причине в район могли попасть единичные бассейны других кластеров. Границы между районами в большинстве случаев проводились по линиям водоразделов или долинам крупных рек, лишь в редких случаях – вкрест речным долинам. Далее в тексте нумерация выделенных районов увеличивается по мере укрупнения наносов внутри групп (медианный признак).

Таблица 1

### Основные внутригрупповые статистические характеристики крупности взвешенных наносов в подъем половодья

№	Агломеративный метод					Метод $k$ -средних							
	Кол-во постов	Avg	Md	Max	Min	Max–Min	№	Кол-во постов	Avg	Md	Max	Min	Max–Min
1	7	0,013	0,009	0,022	0,006	0,016	1	6	0,011	0,008	0,019	0,006	0,013
2	14	0,009	0,009	0,011	0,008	0,003	2	12	0,010	0,009	0,019	0,008	0,012
3	16	0,013	0,011	0,021	0,009	0,013	3	17	0,011	0,010	0,018	0,008	0,010
4	9	0,014	0,013	0,024	0,009	0,016	4	8	0,014	0,013	0,024	0,009	0,016
5	11	0,014	0,015	0,019	0,009	0,011	5	13	0,015	0,016	0,021	0,009	0,012
6	5	0,037	0,032	0,049	0,024	0,025	6	6	0,036	0,032	0,049	0,024	0,025
7	8	0,034	0,033	0,047	0,031	0,016	7	8	0,033	0,033	0,047	0,022	0,025

*Примечание.* Avg – среднее по группе, Md – медиана по группе, Max – максимальное значение по группе, Min – минимальное значение по группе, Max–Min – разброс в группе.

Медианы размерностей частиц в первых двух областях (рис. 1, а) слабо различаются (табл. 1), из чего следует, что эти группы разделены лишь по территориальному признаку. Однако соседство этих областей может служить основанием для их объединения. Крупность наносов 3 и 4 областей также не столь различима – и эти группы можно интегрировать по их территориальной близости друг к другу. В восточной части бассейна Волги выделяются две области – на юге (5 группа) и севере (6 группа): они отличаются крупными взвешенными наносами. Эти области приурочены к горному рельефу (Урал), где благодаря значительным уклонам речных русел способен транспортироваться достаточно крупный материал. В 7 группу вошли водосборы равнинной части бассейна Волги, круп-

ность наносов в которых значительно превышает таковую на остальной территории. Эта группа не имеет сколько-нибудь точных очертаний границ: водосборы разбросаны преимущественно в западной половине бассейна; по этой причине выделить самостоятельный район, соответствующий 7 кластеру, не представляется возможным. Аналогичная карта, но построенная по данным метода  $k$ -средних, отражена на рис. 1, б.

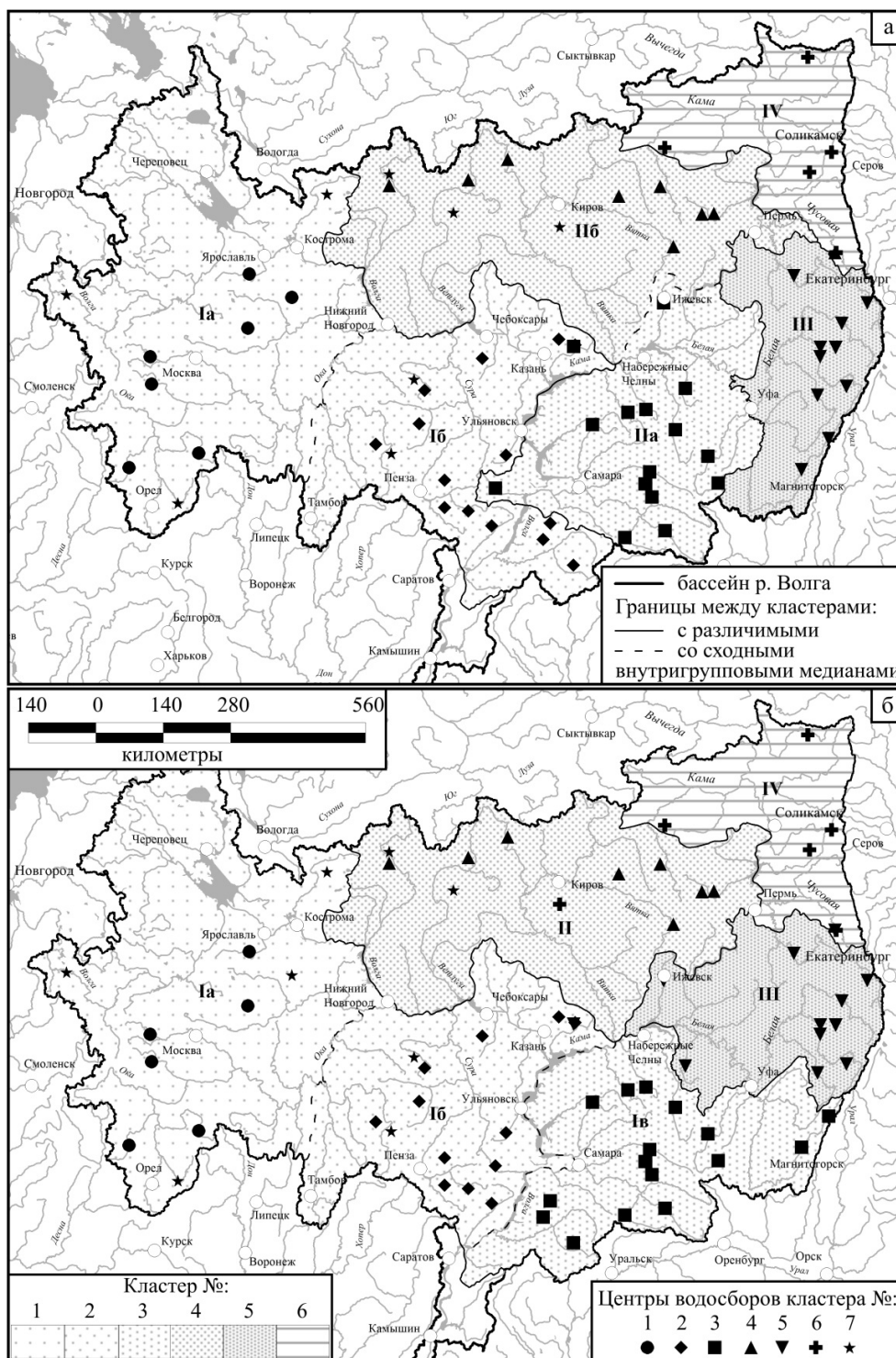


Рис. 1. Карта районирования бассейна Волги по крупности взвешенных наносов в подъем половодья, построенная с применением: а) агломеративного метода, б) метода  $k$ -средних. Номера кластеров на карте соответствуют номерам в табл. 1. Римскими цифрами с буквенными индексами показана возможная группировка кластеров в более крупные территориальные единицы.

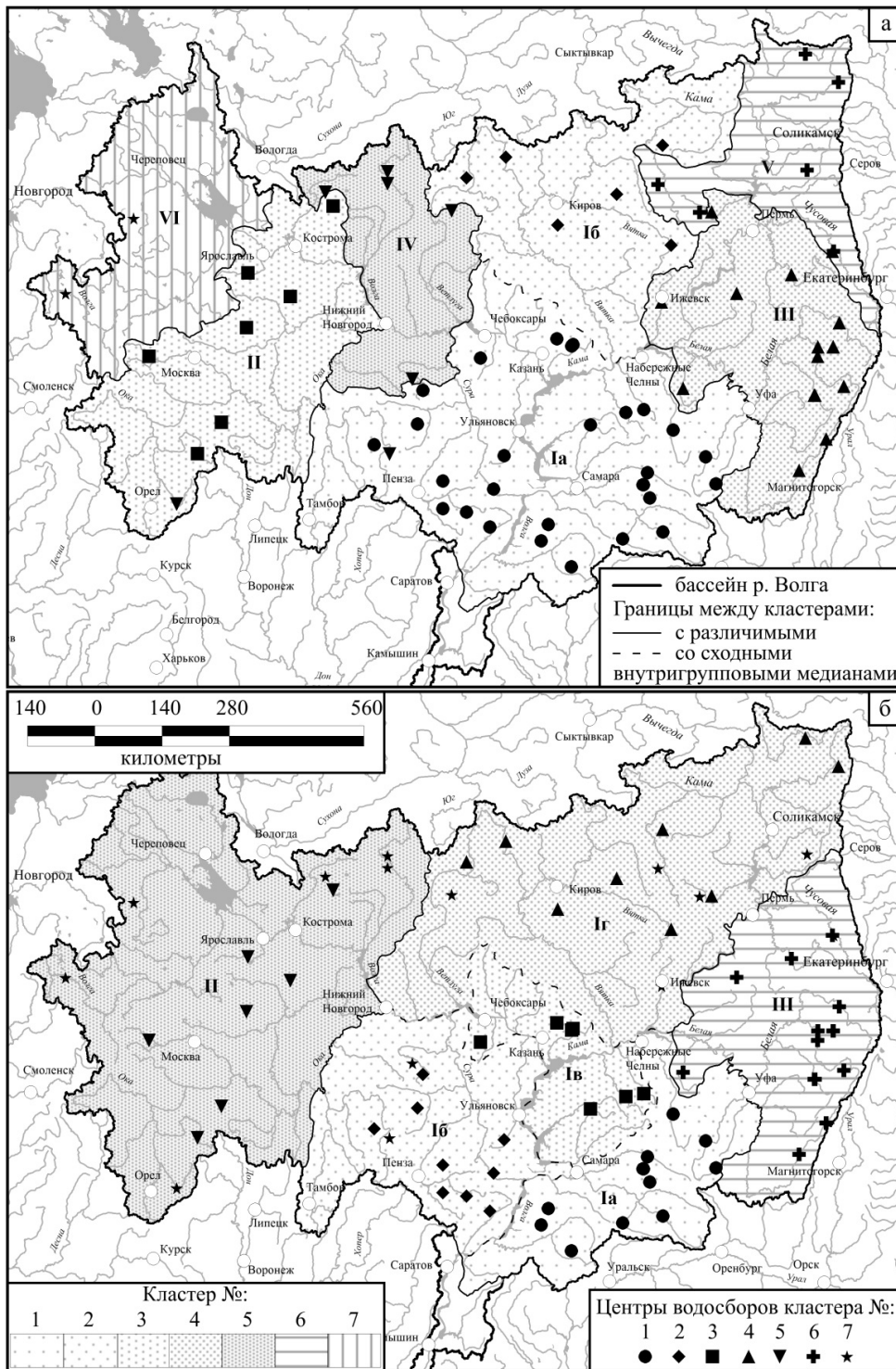


Рис. 2. Карта районирования бассейна Волги по крупности взвешенных наносов в спад половодья, построенная с применением: а) агломеративного метода; б) метода *k*-средних. Номера кластеров на карте соответствуют номерам в табл. 2. Римскими цифрами с буквенными индексами показана возможная группировка кластеров в более крупные территориальные единицы.

Медианы первых трех групп (табл. 1) различаются несущественно, а их территориальная близость является основанием слияния этих областей. Эти группы растянуты с северо-запада на юго-восток, охватывая всю южную часть бассейна. Севернее этих областей выделяется 4 группа, крупность наносов в которой несколько выше. В восточной части бассейна Волги выделяется 5 группа, а на северо-востоке – 6 группа. Материал в них крупнее, чем в вышеназванных областях. Эти группы

водосборов приурочены преимущественно к горной части исследуемой территории (Урал), но с выходом и на равнинную часть бассейна. В 7 группу также вошли водосборы равнинной части бассейна с наиболее крупным материалом; водосборы территориально разбросаны в западной части бассейна. Оба метода позволяют строить хорошо объясняемые карты районирования: по ним неплохо читается влияние орографического и зонального факторов.

Карта районирования крупности взвешенных наносов в спад половодья не отражает влияния вышеназванных факторов на размещение тех или иных областей. По данным агломеративного метода (рис. 2, а) можно объединить 1 и 2 группы, расположенные в центральной части бассейна, в силу территориальной близости и сходства величины их медиан (табл. 2). На востоке с ними соседствуют 4 и 6 группы: первая из них охватывает центральную и южную часть восточной окраины бассейна Волги, а вторая – ее северную часть; обе группы охватывают как горную, так и равнинную территории. Размерность частиц в 6 области почти в два раза больше, чем в 4 группе. Районы 3, 5 и 7 выделяются в западной части бассейна Волги: в 7 области формируются наиболее крупные наносы; эта область приурочена к северо-западной части бассейна.

Метод *k*-средних (рис. 2, б) позволяет объединить первые четыре группы, поскольку они имеют несущественные различия в медианах и территориально близки друг к другу. Эти группы практически полностью охватывают восточную половину бассейна, лишь в восточной и юго-восточной ее частях выделяется 6 группа с крупным материалом; этот район приурочен к горной области. В западной половине бассейна размещаются водосборы 5 группы с относительно крупными наносами. Наиболее крупные наносы формируются в водосборах, отнесенных к 7 группе, не имеющей территориального единства. По картам крупности взвешенных наносов в спад половодья можно сказать, что агломеративный метод выделил группы территориально сплоченные, в то время как метод *k*-средних вычленил из выборки водосборы с крупными наносами и объединил их в отдельную группу, тем самым добившись меньшей дисперсии внутри остальных групп.

Таблица 2

### Основные внутригрупповые статистические характеристики крупности взвешенных наносов в спад половодья

Агломеративный метод							Метод <i>k</i> -средних						
№	Кол-во постов	Avg	Md	Max	Min	Max–Min	№	Кол-во постов	Avg	Md	Max	Min	Max–Min
1	28	0,011	0,010	0,019	0,008	0,011	1	12	0,011	0,009	0,018	0,009	0,009
2	6	0,010	0,010	0,014	0,009	0,005	2	9	0,009	0,009	0,010	0,008	0,002
3	7	0,014	0,015	0,022	0,008	0,014	3	7	0,012	0,010	0,019	0,009	0,010
4	16	0,019	0,020	0,031	0,010	0,021	4	9	0,015	0,011	0,028	0,009	0,020
5	7	0,033	0,032	0,043	0,027	0,016	5	7	0,014	0,015	0,022	0,008	0,014
6	6	0,039	0,039	0,053	0,026	0,026	6	14	0,018	0,018	0,030	0,010	0,021
7	2	0,060	0,060	0,061	0,058	0,002	7	14	0,040	0,036	0,061	0,027	0,034

*Примечание.* Avg – среднее по группе, Md – медиана по группе, Max – максимальное значение по группе, Min – минимальное значение по группе, Max–Min – разброс в группе.

Изучение состава донных наносов, выполняемое обычно в межень, обеспечивает получение такой картины распределения ареалов донных отложений и наносов разной крупности, которая интегрирует условия прохождения руслоформирующих расходов в предшествующий многогодный период гидрологического года [24]. По этой причине карта донных отложений строилась только по данным за межженный период. Особенности руслообразующих наносов являются региональными, то есть отличаются в регионах с различными геолого-геоморфологическими и физико-географическими условиями и гидрологическим режимом рек [25]. Агломеративным методом было выделено 8 районов (рис. 3а). Наименьшей крупностью наносов характеризуется область 1, располагаемая в южной части бассейна. Группы 2 и 3, имеющие незначительные расхождения в медианах (табл. 3), занимают центральную часть бассейна. В северо-западной части бассейна выделяется 4 кластер, в котором формируются чуть более крупные наносы. На востоке 2 и 3 группы граничат с 5 и 6 районами, в которых размерность частиц несколько выше. В юго-западной части бассейна Волги располагается 7 группа с еще более крупным материалом, а в группу 8 объединены водосборы с наибольшей крупностью донных отложений – этот район приурочен к горной части территории бассейна.

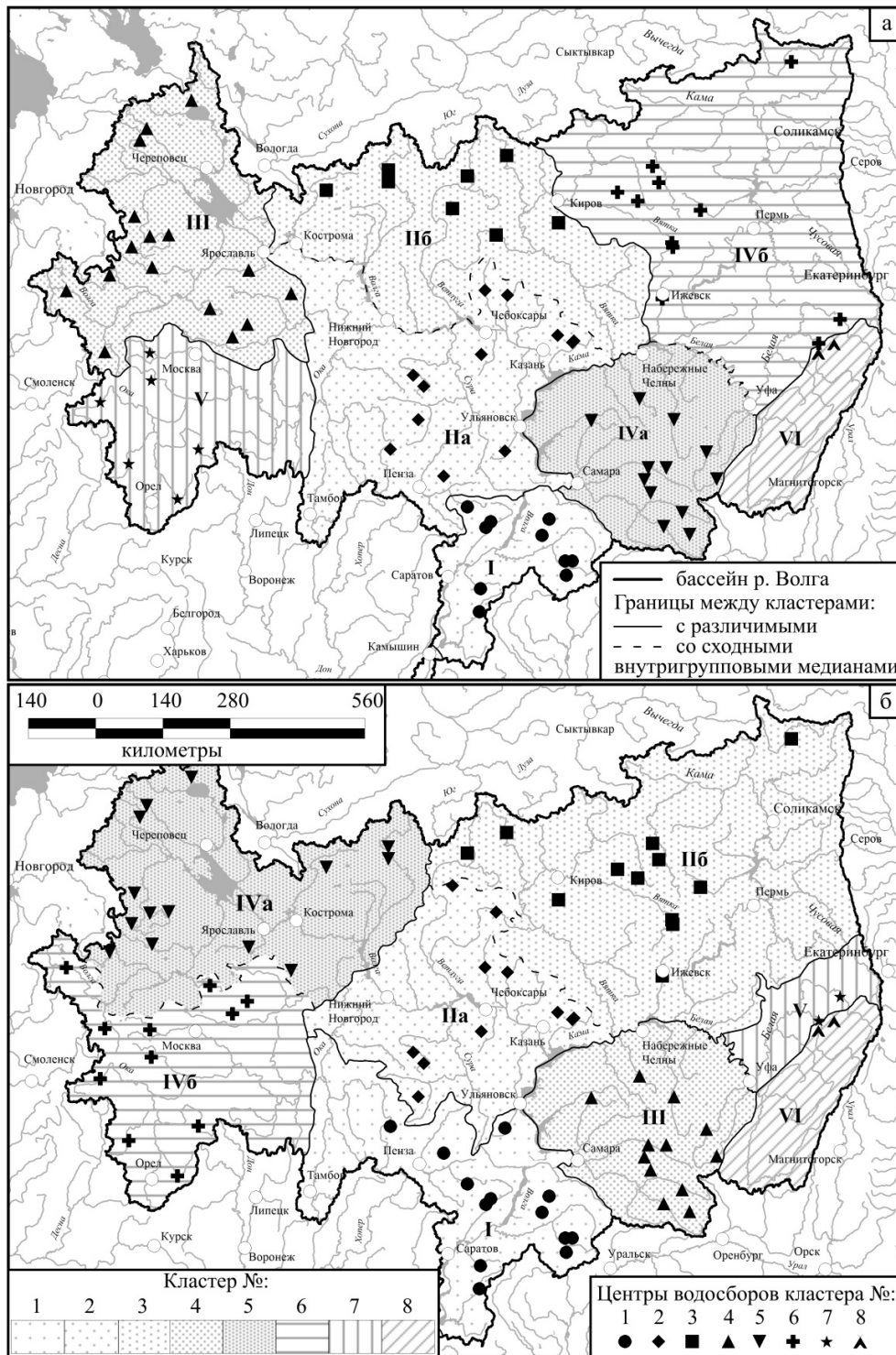


Рис. 3. Карта районирования бассейна Волги по крупности донных отложений, построенная с применением: а) агломеративного метода; б) – метода *k*-средних

Номера кластеров на карте соответствуют номерам в табл. 3. Римскими цифрами с буквенными индексами показана возможная группировка кластеров в более крупные территориальные единицы.

На карте крупности донных отложений в межень, построенной по данным метода *k*-средних (рис. 3, б), границы районов имеют несколько другие очертания. Наименьшие по размеру отложения формируются в центре южной части бассейна (группа 1). Севернее с переходом на северо-восток выделены водосборы с более крупным материалом (группы 2 и 3), а к востоку от группы 1 и южнее групп 2 и 3 формируются еще более крупные отложения (группа 4). В западной половине бассейна в



группах 5 и 6 образуются относительно крупные донные отложения; незначительные расхождения медиан этих групп и географическая близость служат основанием для их объединения. На юго-востоке бассейна, в горной его части, образуются самые крупные наносы (группы 7 и 8).

Таблица 3

**Основные внутригрупповые статистические характеристики крупности донных отложений в межень**

Агломеративный метод							Метод <i>k</i> -средних						
№	Кол-во постов	Avg	Md	Max	Min	Max–Min	№	Кол-во постов	Avg	Md	Max	Min	Max–Min
1	10	0,543	0,159	3,155	0,033	3,122	1	13	0,504	0,282	3,155	0,033	3,122
2	12	0,752	0,403	5,278	0,158	5,120	2	11	0,783	0,380	5,278	0,158	5,120
3	8	0,479	0,437	0,827	0,177	0,650	3	13	0,881	0,419	6,131	0,177	5,954
4	16	1,52	0,477	10,42	0,087	10,34	4	12	0,739	0,506	2,759	0,152	2,608
5	12	0,739	0,506	2,759	0,152	2,608	5	14	0,565	0,562	1,260	0,087	1,174
6	12	1,80	0,513	8,25	0,188	8,06	6	11	1,97	0,570	10,42	0,064	10,36
7	6	0,575	0,671	0,817	0,064	0,753	7	2	5,702	5,702	8,251	3,154	5,097
8	2	42,03	42,03	52,92	31,13	21,79	8	2	42,03	42,03	52,92	31,13	21,79

*Примечание.* Avg – среднее по группе, Md – медиана по группе, Max – максимальное значение по группе, Min – минимальное значение по группе, Max–Min – разброс в группе.

### Заключение

Суммируя изложенное, можно сделать некоторые выводы. В целом кластерный анализ является весьма эффективным инструментом районирования. Несмотря на различие в алгоритмах агломеративного метода и метода *k*-средних, в обоих случаях получаются непротиворечащие друг другу картины. Метод *k*-средних стремится уменьшить разнородность внутри групп, из-за чего могут формироваться территориально разобщенные однородные классы; агломеративный метод в противовес ему отказывается от деталей и формирует компактные группы с более пестрым содержанием. На картах, построенных по данным этих разновидностей кластерного анализа, можно проследить влияние некоторых факторов на крупность речных наносов. Оротографический фактор лучше отражается на карте, построенной по данным агломеративного метода, тогда как зональный фактор, проявляющийся в постепенном уменьшении крупности наносов с переходом от лесных к безлесным зонам [26], – по данным метода *k*-средних. Следует признать, что совместное применение этих методов в исследованиях способствует более детальному анализу изучаемого предмета.

### СПИСОК ЛИТЕРАТУРЫ

1. Гилязов А.Ф. Влияние природно-антропогенных факторов на крупность речных наносов рек Волжско-Камского бассейна // Материалы X семинара молодых ученых ВУЗов, объединяемых советом по проблеме эрозионных, русловых и устьевых процессов. 2014. С. 30-37.
2. Мандель И.Д. Кластерный анализ. М.: Финансы и статистика, 1988. 176 с.
3. Кузьмичёнок В.А. Кластерный анализ внутригодового распределения атмосферных осадков // Лед и снег. 2010. № 1. С. 73-80.
4. Пономарев В.И., Дмитриева Е.В., Савельева Н.И. Классификация временных рядов гидрометеорологических характеристик в Азиатско-Тихоокеанском регионе методами кластерного анализа // Вестн. ДВО РАН. 2010. № 1. С. 38-45.
5. Бубин М.Н. Районирование территории по синхронности многолетних колебаний зимнего стока рек (на примере Челябинской области) // Изв. Томск. политехн. ун-та. 2010. Т. 316, № 1: Науки о Земле. С. 137-142.
6. Кочеткова Ю.О. Методика эколого-геоморфологической оценки рельефа на региональном уровне (на примере Рязанской области) // Вестн. Брянск. гос. ун-та. 2011. № 4. Точные и естественные науки. С. 168-172.
7. Кочеткова Ю.О. Эколого-геоморфологическая оценка морфометрических особенностей рельефа Рязанской области // Вестн. Волжского ун-та им. В.Н. Татищева. Сер. Экология. 2011. 12. С. 156-163.
8. Чашков А.В., Киселев В.М. Использование кластерного анализа и аппарата искусственных нейронных сетей при интерпретации данных геофизических исследований скважин // Журнал Сибирского федерального ун-та. Сер. Техника и технологии. Т. 4, № 4. 2011. С. 453-462.

9. Бериков В.С., Лбов Г.С. Современные тенденции в кластерном анализе / Всероссийский конкурсный отбор обзорно-аналитических статей по приоритетному направлению «Информационно-телекоммуникационные системы». 2008. 26 с.
10. Методы кластерного анализа [Электронный ресурс] // Лаборатория «Информ» [Интернет-портал]. URL: <http://bug.kpi.ua/stud/work/RGR/DATAMINING/clusteranalysismethods.html> (дата обращения: 25.02.2015).
11. Буреева Н.Н. Многомерный статистический анализ с использованием ППП «STATISTICA». Нижний Новгород, 2007. 112 с.
12. Кластерный анализ // StatSoft Russia [Интернет-портал]. URL: <http://www.statsoft.ru/home/textbook/modules/stcluan.html> (дата обращения: 25.02.2015).
13. Метод *k*-средних. URL: <https://ru.wikipedia.org/wiki/K-means> (дата обращения: 25.02.2015).
14. Основные гидрологические характеристики (за весь период наблюдений по 1962 г.) // Ресурсы поверхностных вод. Т. 10: Верхне-Волжский район. Л.: Гидрометеоздат, 1967. 573 с.
15. Основные гидрологические характеристики (за 1963–1970 гг. и весь период наблюдений) // Ресурсы поверхностных вод. Т. 10: Верхне-Волжский район. Л.: Гидрометеоздат, 1976. 536 с.
16. Основные гидрологические характеристики (за 1971–1975 гг. и весь период наблюдений) // Государственный водный кадастр. Т. 10: Верхне-Волжский район. Л.: Гидрометеоздат, 1979. 483 с.
17. Основные гидрологические характеристики (за весь период наблюдений по 1962 г.) // Ресурсы поверхностных вод. Т. 11: Средний Урал и Приуралье. Вып. 1: Кама. Л.: Гидрометеоздат, 1967. 573 с.
18. Основные гидрологические характеристики (за 1963–1970 гг. и весь период наблюдений) // Ресурсы поверхностных вод. Т. 11: Средний Урал и Приуралье. Вып. 1: Кама. Л.: Гидрометеоздат, 1975. 474 с.
19. Основные гидрологические характеристики (за 1971–1975 гг. и весь период наблюдений) // Государственный водный кадастр. Т. 11: Средний Урал и Приуралье. Вып. 1: Кама. Л.: Гидрометеоздат, 1979. 300 с.
20. Основные гидрологические характеристики (за весь период наблюдений по 1962 г.) // Ресурсы поверхностных вод. Т. 12: Нижнее Поволжье и западный Казахстан. Вып. 1: Нижнее Поволжье. Л.: Гидрометеоздат, 1966. 364 с.
21. Основные гидрологические характеристики (за 1963–1970 гг. и весь период наблюдений) // Ресурсы поверхностных вод. Т. 12: Нижнее Поволжье и западный Казахстан. Вып. 1: Нижнее Поволжье. Л.: Гидрометеоздат, 1976. 332 с.
22. Основные гидрологические характеристики (за 1971–1975 гг. и весь период наблюдений) // Государственный водный кадастр. Т. 12: Нижнее Поволжье и западный Казахстан. Вып. 1: Нижнее Поволжье. Л.: Гидрометеоздат, 1980. 327 с.
23. Бутаков Г.П., Дедков А.П. Аналитическое изучение крупнообломочного материала. Казань: Изд-во Казан. ун-та, 1971. 81 с.
24. Чалов Р.С. Русловедение: теория, география, практика. Том 1. Русловые процессы: факторы, механизмы, формы проявления и условия формирования речных русел. М.: Изд-во ЛКИ. 2008. 608 с.
25. Чернов А.В. Влияние геолого-геоморфологических условий на формирование и распределение руслообразующих наносов на реках Восточной Европы // Геоморфология. 2010. № 2. С. 115-120.
26. Дедков А.П., Гусаров А.В., Мозжерин В.В. Две системы эрозии в речных бассейнах равнин Земли и их взаимная трансформация (часть II) // Геоморфология. 2008. № 4. С. 17-28.

Поступила в редакцию 30.03.15

*A.F. Gilyazov*

**CLUSTER ANALYSIS AS A TOOL OF ZONING ACCORDING TO THE SIZE OF RIVER SEDIMENTS  
(BY THE EXAMPLE OF VOLGA BASIN)**

Volga basin was zoned by the size of suspended sediments and bed materials through the use of two (most commonly used in researches) methods of cluster analysis – the agglomerative method of hierarchical approach and the method of *k*-means of probabilistic approach. Despite the difference in the algorithms of these methods in both cases the results of zoning are similar. Method of *k*-means tends to reduce the heterogeneity within groups, whereby spatially separated homogeneous classes can be formed; the agglomerative method, on the contrary, drops details and forms compact groups with a more mottled content. On the maps drawn according to these types of cluster analysis one can trace the influence of some factors on the size of river sediments.

*Keywords:* cluster analysis, zoning, size of river sediments, suspended sediments, bed material, Volga basin.

Гилязов Альберт Фаргатович, аспирант  
ФГАОУ ВПО «Казанский (Приволжский)  
федеральный университет»  
420008, Россия, г. Казань, ул. Кремлевская, 18  
E-mail: [albert\\_liberty@mail.ru](mailto:albert_liberty@mail.ru)

Gilyazov A.F., postgraduate student  
Kazan Federal University  
Kremlyovskaya st., 18, Kazan, Russia, 420008  
E-mail: [albert\\_liberty@mail.ru](mailto:albert_liberty@mail.ru)