

$\Delta V_{2-3} = \Delta V_{3-2} = 0,17$ млн·м³; в оз. Медведно — 0,01 млн·м³). Следовательно, пятиметровая толща воды в оз. Потех и шестиметровая в оз. Медведно в течение суток участвует во внутреннем вертикальном водообмене. Нижние слои озер — консервативные с минимальной долей участия в водообмене с другими слоями.

Количественная оценка целостной гидродинамической обстановки озер с учетом объема их водных масс, годовой динамики в горизонтальной и вертикальной плоскостях выполнены по показателю гидродинамического объема (V_q): $V_q = q_s \cdot q_v \cdot V_{оз}$, м³, где q_s , q_v — соответственно гидравлическая и динамическая нагрузки; $V_{оз}$ — объем озера, м³. В зависимости от причинно-следственных связей гидродинамических факторов озерных процессов соотношение гидродинамических (V_q) и естественных ($V_{оз}$) объемов изменяется: для оз. Ильменок как 1:30, для оз. Потех 1:6, для оз. Медведно 1:2. Высокие значения гидравлической и динамической нагрузок оз. Ильменок обусловили сопоставимость показателя V_q озера с гидродинамическим объемом оз. Потех, естественный объем которого в семь раз больше (см. таблицу). В группе сравниваемых озер самоочищающие возможности оз. Медведно, обусловленные гидродинамическими процессами, ограничены ($V_q = 3,1$ млн·м³).

Таким образом, проведенная оценка гидродинамических факторов озерных процессов на примере трех эвтрофных неглубоких озер указывает на неоднозначность экологической обстановки водоемов в пределах одного генетического подтипа, что вызывает необходимость учета причинно-следственных связей гидродинамических факторов, количественной оценки целостной гидродинамической обстановки малых водоемов и их частей в прогнозировании реакции водных объектов на дополнительное поступление питательных веществ с водосбора.

ЛИТЕРАТУРА

1. Vollenweider R. A.— Schweiz. Z. Hydrol., 1975, Bd., 37, 1, S. 53.
2. Власов Б. П., Романов В. П.— С сб.: История озер в СССР: Тез. докл. VI Всесоюзного совещания. Таллин, 1983, т. 1, с. 56.
3. Базыленко Г. М. и др.— Вестн. Белорусского ун-та. Сер. 2, хим., биол., геогр., 1983, № 2, с. 41.
4. Давтян Н. А.— Труды ГГИ, Л., 1974, вып. 210, с. 155.
5. Хомскис В., Жукайте Э.— В сб. статей для XXIII международного гео-
6. Тихомиров А. И. Термика крупных озер.— Л., 1982.
7. Хомскис В., Жукайте Э.— В сб. статей для XXIII международного географического конгресса. Вильнюс, 1976, с. 137.
8. Як ушко О. Ф. Озероведение. География озер Белоруссии.— Минск, 1981.

УДК 911.63

А. М. ТРОФИМОВ, В. А. РУБЦОВ

В ЕДИНОЙ СИСТЕМЕ МЕТОДОВ АВТОМАТИЧЕСКОГО РАЙОНИРОВАНИЯ И КЛАССИФИКАЦИИ

Проблема районирования, базирующаяся на принципах комплексного образования,— одна из основных проблем географии, имеющих теоретическое и прикладное значение. Функциональное (содержательное) районирование отражает существующий характер географической территориальной дифференциации и как метод исследования весьма широко применяется в предплановых исследованиях и прогнозных заключениях при комплексном районировании различных условий в определенных точках пространства. Однако сам принцип комплексности может быть выдержан в том случае, когда исследование оперирует показателями многомерного пространства объективно. При этом и возникают сложности, связанные, во-первых, с необходимостью перебора множества значений различных (комбинированных) показателей в каждой точке пространства; во-вторых, с объективной их оценкой, не зависящей от расположения объектов, их предпочтения и т. д. В противном случае необ-

ходимо вводить «весовые коэффициенты», которые чаще всего несут отпечатки субъективизма исследователя. Подобные сложности обуславливают необходимость использования автоматических методов и средств районирования, как не зависящих от объекта и, следовательно, относительно объективных. Автоматизация процесса приобрела сторонников во всех областях географии: при выделении природных комплексов, экономико-географических районов, зон и ядер влияния и т. п. Однако схемы и методы большинства из них не всегда четко обоснованы, недостаточно эффективны и просты. Каждый из авторов выбирает метод, который считает наиболее приемлемым, и проводит свои исследования. Сопоставление результатов, полученных различными методами, как правило, не проводится. Между тем мы нашли, что для определенного круга географических задач существуют свои специфические методы, в меньшей степени эффективные для других групп задач. Это привело нас к мысли о создании единой оптимальной системы этапов автоматического районирования, использование которой позволит найти нужный набор этапов и методов для решения заданного класса географических задач. Автоматическое районирование (АР) становится в последнее время одним из ведущих способов получения надежных и обоснованных результатов. Им широко пользуются в различных областях географической науки. Применение традиционных методов районирования (например, наложение схем покомплексного районирования, сопряженного анализа компонентов и др.) хоть и обеспечивает достижение удовлетворительных результатов, тем не менее наводит на мысль о необходимости использования автоматических средств, особенно в тех случаях, когда исследователь имеет дело с исключительно сложными географическими объектами, состоящими из взаимодействующих частей самой различной природы (например, природных и социально-экономических). Именно это обстоятельство обусловило широкий поток работ по АР, который хотя и не привел к единству во взглядах на принципы и униформизм методов АР, однако заставил задуматься над вопросом о соотношении групп методов АР и целевым назначением и характером географических задач. Эта проблема, в свою очередь, вызвала необходимость классификации существующих методов АР, обеспечивающей возможность выделения групп методов, удовлетворительных для определенного класса задач географического районирования. Подобная классификация позволит выделить общие модули методов процесса АР и, таким образом, отработать оптимальность этапов АР.

В основу любого районирования должна быть положена целевая функция, обеспечивающая выбор и обоснование принципов районирования. Последние определяют характер и вид исходной информации. Возможность отработки исходной информации обеспечивает выбор метода. Таким образом, анализ должен включать в себя: характеристику алгоритмов АР и их классификацию, анализ методов с выделением их недостатков, схему этапов районирования.

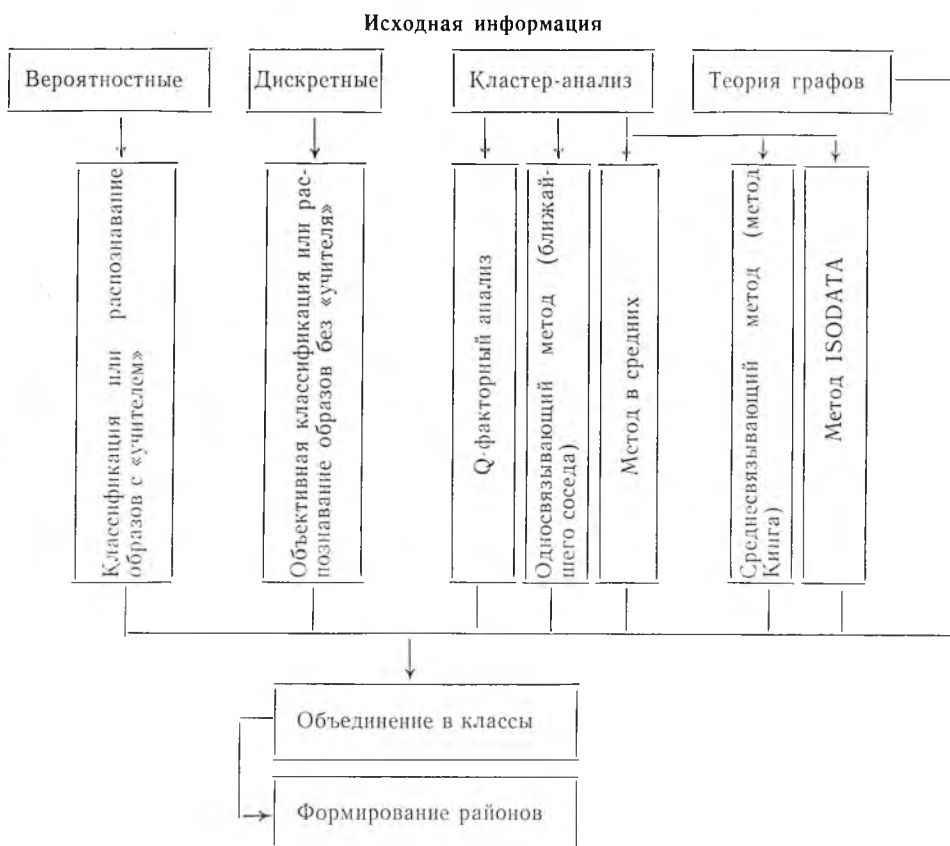
В основе существующих алгоритмов АР лежит подход к классификации как к разделению множества объектов геопространства на непересекающиеся области. Опыт показывает, что их целесообразно разделить на вероятностные и детерминированные, методы теории графов и кластер-анализа (табл. 1).

Особенности алгоритмов каждого типа связаны с характером обрабатываемой информации, чем и определяется применение того или иного подхода. При использовании **вероятностных алгоритмов** числовое значение признака рассматривается как величина случайная, которая при изменении условий измерений может варьироваться в некотором интервале; выводы, полученные при анализе ограниченной группы объектов, с известной степенью вероятности распространяются на большее число объектов. При использовании **детерминированных** методов каждое из анализируемых значений признака принимается за единственно возможное, окончательное [1]. Каждая из этих групп, в свою очередь, подразде-

ляется на две подгруппы: распознавание образов с «учителем» (или классификация) и распознавание образов без «учителя» (или «объективная классификация»). К первой относятся алгоритмы, в которых предполагается известным существование одного или нескольких фиксированных классов (эталонов или обучающая последовательность), состоящих из некоторого числа объектов. Сходство и различие определяются путем соотношения их характеристик с эталонными [2, 3]. Вторая подгруппа объединяет алгоритмы, в которых обучающая последовательность отсутствует. Системе предъявляется некоторое множество объектов, которые требуется разделить на неизвестное число неизвестных классов. Деление должно быть осуществлено таким образом, чтобы объекты, попавшие в один класс, были максимально сходны, а объекты разных классов — максимально различны [1, 3].

Т а б л и ц а 1

МЕТОДЫ АВТОМАТИЧЕСКОГО РАЙОНИРОВАНИЯ



В настоящее время разработано и используется в географических работах большое число **методов кластер-анализа**. Приведем наиболее употребимые.

1. Q-факторный анализ. Число кластеров определяется весами, а вхождение в кластер — наибольшей факторной нагрузкой.

2. Односвязывающий метод (метод ближайшего соседа). Группы, состоящие вначале из одного объекта, объединяются исходя из принципа минимальности расстояния до ближайшего соседа. Расстояние между группами определяется как расстояние между ближайшими членами группы.

3. Среднесвязывающий метод. О расстоянии между группами судят по расстоянию между «центрами масс» (центроидами). Объединяются клас-кластеры с ближайшими центроидами [4].

4. Метод K средних. Начинается с K любых точек. Критерием является минимальное расстояние внутри кластера относительно среднего. При включении элемента в кластер среднее пересчитывается. Процедура продолжается до полного размещения всех точек (объектов) [5].

5. Метод ISODATA. Начинают с k кластеров и соотносят все элементы в соответствии с внутрикластерным критерием минимальности. После того, как все элементы разбиты на группы, средние пересчитываются, и процедура повторяется вновь до тех пор, пока не перестанут получать улучшение внутрикластерного критерия минимальности.

Отметим, что методы 2 и 3 представляют собой иерархические процедуры кластеризации, так как они начинаются с кластеров, содержащих по одному элементу.

Методы АР, основанные на теории графов, предполагают представление исходных данных в виде графа близости [6]. Многие задачи АР решаются либо как первичное образование районов, либо как преобразование уже существующей сетки путем укрупнения ранга района [1]. Для этих целей существует множество алгоритмов АР; различие заключается в выборе мер сходства. Наиболее употребимыми являются:

1. Обобщенное расстояние или евклидова метрика [1, 7] $d_{ij} = \sum_{j=1}^m K_j (X_{ji} - X_{jl})^2$, где K — весовой коэффициент, X_{ji} и X_{jl} — объекты i и l соответственно.

2. Нормированная метрика Минковского, задаваемая выражением [8]: $(\sum |U_{jk} - C_{ik}|^{1/r})^r / (\sum U_{jk}^{1/r})^r$, где $0 \leq r \leq 1$; j, k — пары ОТЕ.

3. Функция Уорда [8] $\sum_K \sum_{ij} \sum_n (X_{ij} - \bar{X}_{ij})^2$, где K — число кластеров; t_j — число ОТЕ в кластере J ; n — число признаков; j — номер ОТЕ в кластере; X_{ij} — значение признака; $\bar{X}_{ij}$ — среднее значение признака.

4. Потенциальные функции [3] $K(X, Y) = [1 + \alpha R^\varepsilon(X, Y)]^1$ или $K(X, Y) = \beta \exp[-\alpha R^\varepsilon(X, Y)]$, где α, β — константы; ε — любое рациональное число; $R(X, Y)$ — расстояние между объектами X и Y .

5. Критерий [1] $Im = I_1(m) - I_2(m)$, где $I_1(m) = \frac{1}{m} K(D, D)$; $I_2(m) = \frac{2}{m(m-1)} \sum_{i=1}^{m-1} \sum_{d>c} K(C, D)$; $K(D, D)$ — среднее сходство между собой всех векторов, попавших в одну группу D ;

$$K(D, D) = \frac{2}{N_D(N-1)} \sum_{X=1}^{N_D-1} \sum_{Y>X} K(X, Y),$$

$$K(C, D) = \frac{1}{N_C N_D} \sum_{X \in C} \sum_{Y \in D} K(X, Y),$$

где N_C, N_D — число элементов в группах C и D . В качестве конечного результата выбирается минимальный вариант разбиения, в котором $Im = \max$.

6. Критерий Фрея [8] $k = (d_{v_{j+1}} - d_{v_j}) / (d_{s_{j+1}} - d_{s_j}) \geq 1$, где d — среднее обобщенное расстояние; v — индекс межгрупповых расстояний; s — индекс внутригрупповых расстояний; j — номер шага. После вычисления обобщенного расстояния в одну группу объединяются два вектора, расстояния между которыми минимальны, затем к ним прибавляют еще один вектор, допускающий минимальный сдвиг k , и т. д. Объединение заканчивается при достижении $k < 1$.

Перечисленные меры сходства используются в основном при получении однородных районов, которые выделяются на основании общности свойств. И уже здесь заложены принципы выбора подходов к методам

АР. Если, например, природное районирование осуществляется на основании общности свойств, то для него в полной мере могут быть использованы те или иные из названных мер. При социально-экономическом или комплексном районировании решающее значение может приобрести общность связей, либо какие-нибудь иные содержательные признаки. В таком случае объединение осуществляется либо на основе общности связей с появлением узлового (коннекционного) района, либо отрабатываются свои специфические меры сходства, позволяющие выделять ориентированный на цель район. Однако это уже область специальных исследований. Приведенные нами меры сходства, таким образом, не исчерпывают все их многообразие, в том числе и возможное, а лишь являются иллюстрацией предполагаемого вооружения географа для АР.

Следует подчеркнуть, что приведенный аппарат для определения мер сходства является математическим и требует весьма серьезного внимания со стороны географов для его доработки. Использование такой меры, как введение весовых коэффициентов, не снимает проблемы. Введение их чаще всего условное, малообоснованное и в результате привносит значительный элемент субъективизма, в целом, казалось бы, весьма строгие методы АР; то же можно сказать и о пороговых значениях.

Используя изложенные в литературе данные по анализу процесса АР, а также результаты разработок группы сотрудников Казанского университета, можно выделить основную группу недостатков, свойственных всем методам и алгоритмам АР.

Они сводятся к следующему.

Большая часть методов, относящихся к группе распознавания образов, может иметь лишь ограниченное применение в области районирования. Причина — обязательное наличие достаточно представительной обучающей последовательности объектов — эталонов, принадлежащей к известным классам. Данные методы применимы лишь при решении задач по выявлению районов-аналогов некоторого фиксированного типа [1].

Применение статистических методов возможно, если пространственная переменная в пределах одного района стационарна, поэтому статистическую однородность всегда можно определить лишь приближенно, с заданной достоверностью [1].

В большинстве методов отсутствуют критерий оптимальности для классификации и единая мера оценки граничности; кроме того, необходимо вводить пороговые значения.

Таким образом, возникает проблема оценивания и последующей классификации самих методов АР, что может привести к появлению общей схемы решения задач АР (с выделением этапов и выводом на определенные операции, которым соответствуют те или иные методы или группы методов).

Опыт исследования показывает, что целесообразность и эффективность использования тех или иных методов АР обусловлены конкретизацией исходной математической модели, постановкой задачи. В основе АР, как показано, должна лежать целевая функция, которая предопределяет выбор (обоснование) принципов районирования, обоснованных с содержательных позиций. Отбор принципов и целевая функция определяют вид и характер исходной информации, а последняя (в соответствии с тем и другим) — выбор метода или группы методов. Поэтому все методы АР мы считаем целесообразным разделить вначале на три больших группы (табл. 2):

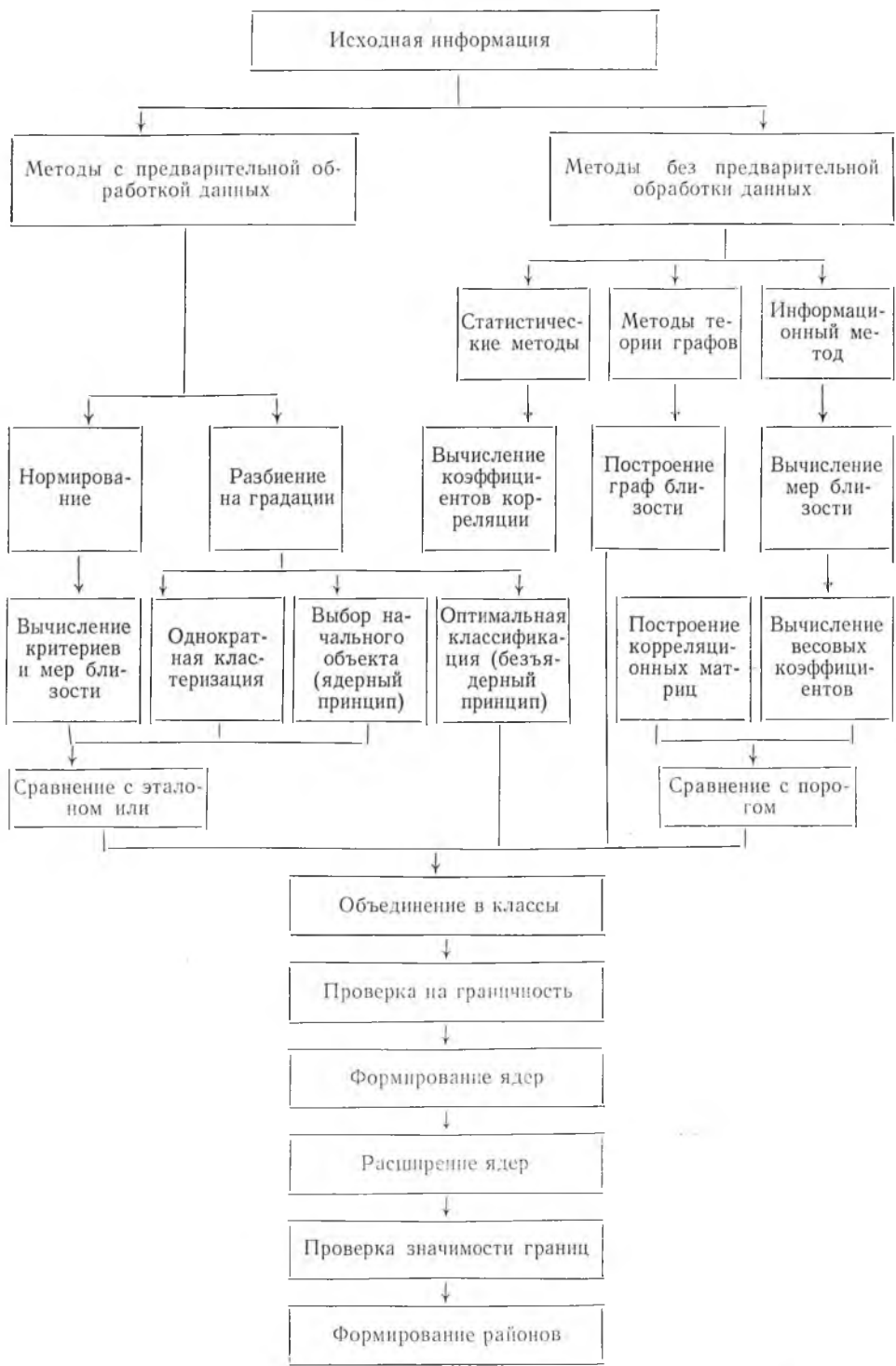
- 1) для которых не требуется первоначальная обработка исходных данных;
- 2) предполагающие первоначальную обработку исходных данных;
- 3) основанные на теории графов.

В каждую из двух первых групп входят методы, независимо от их принадлежности к статистическим, детерминированным или кластер-анализу. В свою очередь, первая группа подразделяется на две подгруппы:

а) содержащую методы, для которых исходную информацию необходимо привести к единой мере измерения; б) методы, требующие разбиения исходных данных на градации (см. табл. 2).

Таблица 2

ЭТАПЫ АВТОМАТИЧЕСКОГО РАЙОНИРОВАНИЯ



Данная классификация позволяет выделить общие модули, принадлежащие всем алгоритмам. Последнее обеспечивает поиск основных тре-

бований или условий, которые необходимо ликвидировать или, по крайней мере, уменьшить.

Схема предполагает выделение условия проверки на граничность, как отдельного особого модуля, свойственного всем методам. Проверка на граничность включает в себя ряд вычислительных алгоритмов и оценочных методов [7].

Работа по данной схеме АР показала, что после проверки на граничность окончательно формируются ядра районов, и уже эти ядра можно рассматривать как эталоны. Таким образом, вскрывается некоторая цикличность процесса АР. Ядра районов, следовательно, могут не задаваться, а автоматически возникать в процессе АР и, таким образом, избегая вводимого ранее порога, процесс АР можно вести на основе полученного первичного ядра. Последующий процесс расширения ядер, соответствующий ранее существующему циклу районирования, приводит к окончательному формированию района. Итак, с целью исключения вводимых субъективных порогов и эталонов, процесс АР должен включать в себя два этапа: связанный с поиском и выделением ядер районов, связанный с собственно процессом районирования на основе выделенных ядер.

Предложенная схема АР позволяет увязывать звенья: цель → исходная информация → методы → конечный результат и наоборот: цель → конечный результат → методы → информация. Следовательно, исходя из имеющейся информации и целей можно предположить ту или иную группу методов или метод, состоящий из нескольких модулей различных алгоритмов. Зная цель и метод, исследователь сможет заранее готовить информацию в той или иной форме.

На кафедре экономической географии Казанского университета в связи с созданием единой оптимальной схемы этапов АР разработан пакет прикладных программ, включающий как сервисные (вспомогательные, например, для подготовки данных), так и программы АР и АК на языке Фортран. Таким образом, исследователь может иметь в своем распоряжении удобное средство для реализации на ЭВМ предварительных данных для получения окончательного, оптимального, варианта разделения территории на районы по заданной цели.

ЛИТЕРАТУРА

1. Куприянова Т. П. Принципы и методы физико-географического районирования с применением ЭВМ.— М., 1977.
2. Ватанабе С., Ломберг П., Куликовский К. и др.— В кн.: Автоматический анализ сложных изображений. М., 1969.
3. Дорофеев А. А., Лумельский В. Я.— В кн.: Алгоритмы обучения распознавания образов. М., 1966.
4. King B. F.— In: J. of Amer. Statistical Assoc, 1967, N 62.
5. Matula D. W.— In: Theory and applications of graphs in American's Bicentennial Gear., 1977.
6. Матюла Д. В.— В кн.: Классификация и кластер. М., 1980.
7. MacQueen J. B.— In: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, 1967, N 1.
8. Сокал Р. Р.— В кн.: Классификация и кластер. М., 1980.

УДК 550.46; 551.48

О. Ф. ЯКУШКО, А. Л. ЖУХОВИЦКАЯ, А. Н. РАЧЕВСКИЙ

КАРБОНАТНЫЕ ОТЛОЖЕНИЯ СОВРЕМЕННЫХ ОЗЕР БЕЛОРУССИИ

В условиях Белоруссии, несмотря на запасы коренных меловых пород, существует значительный дефицит в известковых материалах [1], что представляет серьезную проблему, учитывая необходимость известкования подзолистых почв (более 6 млн. га). Исследование карбонатных озерных отложений должно способствовать разработке прогноза и использованию разведанных запасов местных карбонатных месторождений.