

# СТАТИСТИЧЕСКИЕ МЕТОДЫ В ФИТОЦЕНОЛОГИИ НА РУБЕЖЕ ТЫСЯЧЕЛЕТИЙ (К 50-летию выхода монографии П. Грейг-Смита)

Г.С. Розенберг

Институт экологии Волжского бассейна РАН, Тольятти, Россия

## Введение

Математическая статистика как прикладная наука прошла долгий и противоречивый путь становления начиная с 1900 г., когда Карл Пирсон стал издавать журнал «*Biometrika*», посвященный статистическим методам в биологии (практически одновременно работами А.Я. Гордягина начались биометрические исследования и в России; см. Боголюбов, 2002). Шли десятилетия, появлялись новые подходы к обработке данных, новые алгоритмы и технологии. По свидетельству ряда специалистов (см., например, Орлов, 2004), в статистике, как и в большинстве других отраслей наук, происходит смена парадигм: предыдущая была сформирована к 50–60-м гг., а развитие современной парадигмы связано с созданием новых разделов и непривычных ранее идеологий (Шитиков и др., 2005) – статистики нечисловых данных, нейросетевого и эволюционного моделирования, логики нечетких (англ. *fuzzy*) множеств, статистики интервальных данных и пр.

Конец 60-х гг. прошлого века ознаменовался выходом в свет двух важных для отечественных исследователей работ по количественной геоботанике: перевода монографии «Количественная экология растений» (Грейг-Смит, 1967) и «Статистические методы в геоботанике» (Василевич, 1969); через десять лет появилась еще одна монография в том же ключе (Миркин, Розенберг, 1978) и два обзора (Василевич, 1972; Миркин, Розенберг, 1979). Пожалуй, других монографических обобщений статистических методов в геоботанике на русском языке больше не появлялось (наиболее полный на сегодняшний день обзор статистических методов исследования с примерами из гидроэкологии можно найти в монографии В.К. Шитикова с соавторами, 2005). Цель данного сообщения – сделать краткий обзор современного состояния именно статистических методов (рассмотрение методов математического моделирования в фитоценологии в широком смысле в рамках одной статьи невозможно (Розенберг, 1984), но некоторые экскурсы все-таки будут сделаны). За основу такого обзора я взял структуру монографии П. Грейг-Смита (Greig-Smith, 1957, 1964, 1983; Грейг-Смит, 1967; подразделы настоящей статьи по названию дословно совпадают с главами этой монографии), которая на про-

тяжении 50 лет (40 – для отечественных исследователей) является одной из основных монографических работ по количественной геоботанике.

Питер Грейг-Смит родился в 1922 г. в Бирмингеме (Великобритания). После окончания Кембриджского университета (на него большое влияние оказали работы А. Тэнсли: Arthur George Tansley, 1871–1955, и А. Уатта: Alexander Stuart Watt, 1892–1985) он некоторое время работал в сельскохозяйственных исследовательских учреждениях, где и познакомился с применением статистики при количественных исследованиях растительности. Более углубленное изучение количественных методов применительно к геоботаническим исследованиям проходило в Манчестерском университете, где он встретился с Э. Эшби (Ashby Eric, 1904–1992), который одним из первых обобщил имевшиеся на то время данные об использовании статистики в геоботанике (Ashby, 1935, 1936, 1948). С 1952 г. и до выхода на пенсию Грейг-Смит работает в университетском колледже Уэльса в г. Бангоре (The University of Wales College, Bangor). В 1964 г. он стал редактором «The Journal of Ecology». В 1994 г. на VI Международном экологическом конгрессе (г. Манчестер, Великобритания) он был отмечен как выдающийся эколог-статистик (вместе с ним эту награду получили David W. Goodall, Robert May, Laszlo Orloci, Evelyn Chrystalla [Chris] Pielou, Dan Simberloff, Robert Sokal и др.). Умер профессор Питер Грейг-Смит от лейкемии в феврале 2003 г. (некролог был опубликован в журнале «The Bulletin of the British Ecological Society». 2003. Vol. 34. № 1).

Наконец, следует назвать переводчика «Количественной экологии растений» кандидата географических наук В.Д. Утехина и редакторов этого издания профессоров Т.А. Работнова (1904–2000) и А.А. Уранова (1901–1974).

Завершить это краткое введение мне хотелось бы словами самого Питера Грейг-Смита из его предисловия к первому изданию «Количественной экологии растений», которые и сегодня не потеряли своей актуальности: «Переход от качественного к количественному подходу характерен для развития любой отрасли науки. Когда достигается определенная степень понимания явлений в общих чертах, внимание исследователей, естественно, сосредотачивается на более тонких деталях структуры или поведения, наблюдаемые различия которых менее значительны и могут быть оценены лишь с помощью измерения. Неудивительно, что количественный подход развился ранее в большинстве отраслей наук о неживой природе, чем в биологической науке. По-видимому, одно из наиболее значимых различий между физикой и химией, с одной стороны, и биологией, с другой – состоит в том, что в первых обычно можно временно изолировать одну переменную величину для изучения, в то время как в биологии

это редко бывает осуществимо. Таким образом, в физике и химии относительно простая программа качественных исследований позволяет легко установить общие черты явлений и подготовить путь для применения более точных количественных методов. В биологии не только редко удается изолировать для изучения отдельные переменные, но и объекты изучения сами обычно настолько сложны, что плохо поддаются измерению» (Грейг-Смит, 1967, с. 13).

## 1. Количественное описание растительности

Любое количественное исследование начинается с планирования эксперимента. Именно в этом, казалось бы, досконально исследованном аспекте математической статистики сегодня разгораются дискуссии (правда, чаще всего на страницах непрофильных журналов, где легче замаскировать, какие утверждения обоснованы на научном уровне, а какие являются субъективными мнениями отдельных исследователей и где меньше риск получить квалифицированную отповедь). Таковыми мне видятся дискуссионные статьи М.В. Козлова (2003) и М.В. Козлова и С. Хёлберта (2006), эксплуатирующие традиционные для классической статистики понятия «повторности измерений», их «независимости» и вводящие новое понятие «мнимая повторность» (Hurlbert, 1984)<sup>1</sup>.

Точный смысл понятия «независимые испытания» построен на детально проработанной в математической статистике вероятностной модели порождения данных. В этой модели выборки из наблюдаемых значений обычно рассматривают как реализацию независимых одинаково распределенных случайных величин  $X_1(\omega), X_2(\omega), \dots, X_n(\omega), \omega \in \Omega$ . При этом считается, что полученные в эксперименте конкретные значения  $x_1, x_2, \dots, x_n$  соответствуют определенному элементарному событию  $\omega = \omega_0$ . В других условиях (например, при действии некоторого внешнего фактора среды) могут быть получены иные наблюдаемые значения  $y_1, y_2, \dots, y_m$ , соответствующие другому элементарному событию  $\omega = \omega_1$ . Предполагается также, что наблюдения в одной выборке не зависят от наблюдений в другой, поэтому выборки называют независимыми.

---

<sup>1</sup> Стюарт Хёлберт (Hurlbert, 1984) сделал критический анализ 156 экспериментальных экологических статей, опубликованных в течение 1960–1980 гг. в ведущих англоязычных журналах. Оказалось, что в 27% случаев исследуемое воздействие применялось в *одной* повторности, либо использовавшиеся повторности *не были статистически независимыми*; тем не менее для доказательства эффекта экспериментального воздействия использовались статистические методы. Ошибки такого рода получили у него название «мнимых повторностей» (*pseudoreplication*).

Одна из задач статистической обработки состоит в том, чтобы по результатам эксперимента (т.е. наборам из  $m$  и  $n$  действительных чисел, соответствующим различным элементарным событиям  $\omega = \omega_0$  и  $\omega = \omega_1$ ) проверить гипотезы об однородности порождаемых данных. Или в иной формулировке: требуется проверить, есть ли статистические различия между выборками или они принадлежат к одной генеральной совокупности. Техника такого статистического анализа десятилетиями отработывалась в медико-биологических исследованиях для выявления эффекта воздействия, и в большинстве случаев корректность полученных выводов не вызвала никаких сомнений.

В чем же состоит точка зрения С. Хёлберта (Hulbert, 1984) на проблему «мнимых повторностей»? По сути, в «сухом остатке» остается лишь два совершенно бесспорных тезиса (см. Козлов, 2003):

- не всегда выводы, полученные при изучении частных выборок, корректно распространять на всю генеральную совокупность;
- оценка степени влияния фактора может оказаться ошибочной, если изучаемое воздействие должным образом не локализовано, а сопоставляемые данные взяты из недостаточно рандомизированных источников.

Но эти основные принципы статистических оценок давно и широко известны отечественным математикам и экологам. Еще в 1925 г. Р. Фишер (1958) подвел итог длительной полемики о принципиальной возможности делать заключения относительно генеральной совокупности по частным выборкам<sup>2</sup>: «... *теоретически можно говорить лишь о достоверной вероятности выводов или уровне правдоподобия проверяемых гипотез*. И, если соблюдены основные предпосылки вероятностной модели порождения данных (случайность и независимость), индуктивные выводы вполне корректно могут быть сделаны на основе статистических критериев существенности».

Далеко не нов и второй тезис – *эксперимент должен быть спланирован таким образом, чтобы включить максимально возможную вариацию всех влияющих факторов*, что чаще всего подразумевает современную стандартную методику изучения объекта исследования. Задолго до Стюарта Хёлберта у российских ученых были настольными книги В.В. Налимова и Н.А. Черновой (1965), Ю.П. Адлера (1969), В.В. Фёдорова (1971) и других, ставших классиками в области матема-

---

<sup>2</sup> «Практическое использование теоретически обоснованных методов статистических оценок настоятельно требует, чтобы исследователь при интерпретации своих результатов и при изучении результатов других экспериментаторов умел видеть различие между заключениями, соответствующими природе наблюдаемых фактов, и такими выводами, которые обусловлены просто применением неправильных методов оценки» (Фишер, 1958, с. 236).

тических методов планирования оптимальных многофакторных экспериментов. Сюда же следует отнести работы П. Грейг-Смита (1967, с. 31–57) и В.И. Василевича (1969), которые хорошо понимают важность соблюдения двух выше названных тезисов (так, Грей-Смит специально подчеркивал [с. 49]: «... нередко считают, что достаточно ходить по площадке, бросая квадрат через плечо, производя ту же операцию с закрытыми глазами или каким-либо другим способом... Если положение квадратов проанализировать на случайность, почти всегда обнаруживается, что пробы распределены по площади не случайно... Поэтому дополнительная затрата сил на применение более объективного метода рандомизации, как правило, оправдывает себя»). Проблема того, как должны между собой соотноситься конкретные задачи исследования, природа исследуемых факторов (*treatment*), объекты (*experimental unit*), участвующие в эксперименте, и регистрируемые показатели, долгие годы была в центре внимания как методической литературы, так и отраслевых практических рекомендаций, нормативов и стандартов. Поэтому убеждать читателя в необходимости **тщательно продумывать конечные цели, план и способ реализации наблюдений** было бы пустопорожней говорильней.

В естествознании принципиально невозможна постановка «идеального опыта», не несущего погрешности эксперимента. При любом самом оптимальном плане в реальных условиях трудно получить стерильные выборки случайных величин, напрямую связанные с уровнями воздействия факторов, изолированные от влияния посторонних эффектов и имеющие функцию распределения, согласующуюся с закономерностями генеральной совокупности. Как нельзя дважды войти в одну реку, так и нельзя обеспечить абсолютную идентичность двух местообитаний для сравниваемых растительных сообществ. Иными словами, можно категорически утверждать, что не в 27%, как считает С. Хёлберт, а во всех 100% случаев использовавшиеся повторности реального экологического эксперимента *не были статистически независимы*.

Однако носят ли эти отклонения от исходных предпосылок *статистически значимый* характер (*t*-статистика какого-нибудь параметра распределения случайной величины должна быть состоятельной, несмещенной, эффективной и достаточной; Г. Корн, Т. Корн, 1968, с. 536), определяющий суть получаемых выводов? Разумеется, возможность применения вероятностной модели порождения данных в каждой конкретной реальной ситуации требует обоснования. Но обычно при правильно спланированном эксперименте независимость и случайность результатов наблюдений, входящих в выборки, обеспечивается самой схемой постановки опытов и не требует постоянной проверки

математических соотношений. В сомнительных случаях можно осуществить измерения в нескольких контрольных точках и выполнить оценку статистических гипотез независимости и одинаковой распределенности на основе соответствующих критериев (см., например, Большев, Смирнов, 1983). В любом случае, если при описании статистических результатов использованы ритуальные фразы: «Наблюдаемые величины по условиям эксперимента считаем независимыми и одинаково распределенными в пределах каждой выборки», автор обладает некоторой «презумпцией невиновности», т.е. доказывать, что наблюдения в одной выборке зависят от наблюдений в другой, юридически необходимо уже «Хёлберту». Кстати, процесс доказательства «обвинительной» гипотезы о статистической взаимозависимости измеряемых показателей и «мнимости» используемой группировки данных в каждом конкретном случае не является тривиальным и немислим без кропотливого анализа всего исходного массива наблюдений. Фактически на языке математической статистики мы имеем дело с ошибками первого (гипотеза отвергается при условии, что она справедлива) и второго рода (гипотеза принимается при условии, что она ошибочна). При этом отличительной особенностью определения ошибки первого рода является тот факт, что мы сами задаем (пусть и достаточно малую) вероятность такой ошибки. А вот для вычисления ошибки второго рода (доказательство «обвинительной» гипотезы) необходимо знать распределение *t*-статистики при условии, что гипотеза несправедлива (т.е. *t* не подчиняется закону распределения Стьюдента). А найти этот закон – очень не простая задача.

К сожалению, при анализе вероятностных моделей в практических условиях во все времена существовала (и будет существовать?) основная проблема прикладной статистики: приходится обрабатывать данные, опираясь на предпосылки, которые всегда более или менее нарушаются. Например, кроме требования независимости определены два еще более жестких условия применимости статистических критериев для оценки однородности выборок:

- результаты наблюдений должны иметь нормальные распределения с математическими ожиданиями  $m_1$  и  $m_2$  и дисперсиями  $\sigma_1^2$  и  $\sigma_2^2$  в первой и во второй выборках соответственно;
- дисперсии результатов наблюдений в первой и второй выборках должны совпадать.

Если хотя бы одно из этих условий не выполнено, то формально нет никаких оснований считать, что *t*-статистика имеет распределение Стьюдента, и поэтому применение традиционного метода проверки равенства средних, строго говоря, не обосновано.

Поскольку априори нет оснований предполагать нормальность распределения результатов геоботанических (экологических) наблюдений, следовательно, эту предпосылку надо проверять. Однако проверка нормальности – еще более сложная и трудоемкая статистическая процедура, чем проверка однородности. Например для того, чтобы гарантировать вывод, что функция распределения результатов наблюдений отличается от нормальной не более чем на 0,01, требуется порядка 2500 наблюдений (Орлов, 2004).

Среди сравнительно новых статистических оценок укажу еще на один показатель – *планируемое число наблюдений для получения средней с заданной точностью* (Розенберг, 1976; Усманов, 1984). Во всех пособиях по биометрии можно найти следующую формулу:

$$N_b = V(n)^2/p^2,$$

где  $N_b$  – планируемый объем выборки;  $n$  – объем первоначальной выборки;  $V(n)$  – коэффициент вариации;  $p$  – точность определения планируемой средней. Недостаток этой классической формулы (она не учитывает случайный характер определения коэффициента вариации  $V$ ) преодолевается путем расчета уточненного планируемого объема выборки (Розенберг, 1976):

$$N_R = N_b n [1 + S_t \sqrt{2/n}] / (n - 1),$$

где  $S_t$  – показатель функции Лапласа (для  $t$  – степени надежности получаемого результата, меняющегося от 0,08 до 0,95, значение  $S_t$  будет меняться от 1,28 до 1,96). Замечу, что при высоком значении  $t$  надежность «вступает в конфликт» с экономичностью исследования, так как для больших  $t$  будут получаться и большие  $N_R$ , следовательно, выбор величины  $t$  должен заметно зависеть от цели исследования.

Поэтому в практических условиях приходится выбирать между двумя возможностями. Либо использовать *t-статистику* на основе нереалистического предположения о нормальности распределения результатов наблюдений, надеясь, что отклонения от нормальности незначительно отразятся на качестве выводов, либо использовать непараметрические критерии или асимптотические методы, где существуют свои ограничения. Процедура проверки гипотез входит в «порочную спираль»...

Тем не менее исследованиями (Орлов, 2004) показано, что при больших объемах выборок требование нормальности ослабевает (нужный эффект обеспечивается центральной предельной теоремой), а при близком объеме выборок не требуется также и равенства дисперсий. Другими словами, если объемы двух выборок достаточно велики (несколько десятков) и равны, то проверка равенства математических ожиданий с помощью критерия Стьюдента дает правильные результаты независимо

от того, выполнены ли предпосылки нормальности и равенства дисперсий или нет.

На основе этого напрашивается следующий вывод. В практических условиях статистические выводы всегда характеризуются принципиальной неопределенностью, поскольку не ясно, как нарушение исходных предпосылок влияет на качество конечных результатов. Однако *эффект нарушенных предпосылок* чаще всего имеет место в случаях, когда влияние изучаемого фактора проявляется не слишком отчетливо (в пограничной полосе неопределенности), а объем обрабатываемого экспериментального материала недостаточен. Поэтому как аналитическая статистика, так и фитоценология (экология) – это не только **наука**, но и **искусство** чувствовать не мнимую, а истинную проблематику (Oksanen, 2001). При этом чтобы избежать скоропалительных выводов, надо предвзительно набрать необходимый интеллектуальный багаж, приобрести опыт размышлений и практической работы.

Позитивным моментом настоящей дискуссии следует признать тезис, последовательно подчеркиваемый М.В. Козловым и С. Хёлбертом и не попавший до сих пор акцентированно в поле зрения отечественных экологов, о необходимости *всегда четко в эксперименте и в последующей статистической обработке отражать уровни, на которых взяты анализируемые повторности*. Например, если изучается индивидуальная реакция растения на действующий фактор, то выборочные случайные величины, несомненно, составляют измерения у каждой особи (см., например, Кононов, 1981). Если анализируется влияние фактора на интегральные популяционные характеристики (например, биоразнообразие или способность к размножению), тогда, бесспорно, в качестве выборочных единиц выступают некоторые популяционные группы. Можно привести примеры, когда элементами выборок могут являться ландшафты или даже биомы.

Совершенно правы М.В. Козлов и С. Хёлберт (2006, с. 148): «Повторности не могут рассматриваться как истинные, либо мнимые сами по себе – *лишь их использование в статистическом анализе может расцениваться как правильное либо ошибочное*» (*курсив автора* – Г.Р.). И тут же они противоречат сами себе, объявляя без разбора и анализа целей опытов некорректными все эксперименты с аквариумами, теплицами, делянками, клетками с мышами и др., как основанные на «мнимых повторностях». Было бы больше пользы, если бы М.В. Козлов и С. Хёлберт (возможно, с профессиональными математиками) вместо бесплодных перепалок, проработали бы строгую методологию *практической оценки степени мнимости повторностей*. Например, А.Н. Колмогоровым (1987) было формализовано понятие случайности на основе

теории информации (грубо говоря, числовая последовательность является случайной и независимой, если ее нельзя заметно сжать без потери информации). Распространяя эти идеи на тему настоящей дискуссии, можно предположить, что *если при переходе к более агрегирующим экспериментальным единицам происходит заметная потеря информации, то наблюдаемая последовательность не является мнимой* (Шитиков и др., 2007).

## 2. Отбор образцов и сравнение

«Цели, для которых применяется количественная оценка растительности, укладываются в рамки одной из трех категорий: а) общая характеристика состава растительности на определенном участке, причем производится также сравнение с другими территориями или с тем же участком в другое время; б) изучение изменений в составе растительности в пределах одного участка; в) взаимосвязь изменений состава растительности с изменениями одного или нескольких факторов среды» (Грейг-Смит, 1967, с. 43). «Одной из задач статистики ... является сведение большого числа исходных данных к нескольким показателям с сохранением возможно большей части информации, содержащейся в первоначальном материале» (Василевич, 1969, с. 20).

Определение средних, дисперсий, ошибок и других моментов разного порядка, по-видимому, наиболее простой и часто используемый метод обработки результатов наблюдений. Однако следует еще раз подчеркнуть, что исчисление средних – не механическая счетная операция. Выбор средней должен согласовываться с предполагаемым распределением варьирующего и усредняемого признака (кроме того, следует учитывать и цель определения средней тенденции). В этом отношении показателен следующий пример (Розенберг и др., 1993, с. 24).

На скамейке сидело пять человек. Двое были бродягами, их имущество оценивалось по 25 центов, третий был рабочим с имуществом в 2000 долларов, четвертый владел 15000, а пятый был миллионером с чистым доходом в 5 млн дол. Таким образом, мода (положение максимума распределения) равна 25 центам (точно описывает 40% выборки), медиана (значение распределения, при котором описано 50% выборки, центральный член ряда) равна \$2000, а средняя арифметическая – \$1003400,1. Таким образом, если взять за основу характеристики этих странных соседей на скамейке моды, то общий актив этой группы должен быть оценен примерно в 1–1,5 доллара (что обидно для всех, кроме бродяг). Остальные оценки выглядят еще хуже.

Фактически это пример того, что имеются системы объектов, для которых *отсутствует центральная тенденция*. В таких случаях рекомендуется использовать более сложные устойчивые статистики, например, **средневзвешенную** (один из основных показателей прямого градиентного анализа; Миркин, 1971; Миркин, Розенберг, 1978) или **бивес-оценку** (Мостеллер, Тьюки, 1982; Розенберг и др., 1994).

**Среднее** является *неустойчивой статистикой* (статистика называется *устойчивой*, если при изменении малой доли данных, неважно какой и сколь сильно, существенных изменений в суммирующей статистике не происходит, Мостеллер, Тьюки, 1982). **Медиана** является примером *устойчивой статистики* (на медиану не влияют величины «больших» и «малых» значений: она терпима к нарушениям нормальности на «хвостах распределения»). Однако, кроме *устойчивости (робастности) к предпосылкам* «хорошая» статистика должна обладать еще и свойством *устойчивости (робастности) к эффективности* (высокая эффективность оценивания должна гарантироваться при широком варьировании ситуаций; табл. 1).

Таблица 1. Характеристики некоторых статистик центральной тенденции  
(по: Мостеллер, Тьюки, 1982, с. 214)

Статистика	Объем выборки	Устойчивость	Гауссова устойчивость	Робастность к предпосылкам	Робастность к эффективности
Арифметическое среднее	Малый	Нет	100%	Плохая	Плохая
	Большой	Нет	100%	Очень плохая	Очень плохая
Медиана	Малый	Да	Высокая	Высокая	Высокая
	Большой	Да	100%	Высокая	Умеренная
Бивес-оценка	Малый	Разумно	Неплохая	Высокая	Высокая
	Большой	Да	90%	90%	Высокая

Таким образом, если пренебречь совсем малыми выборками, бивес-оценка обладает всеми желаемыми свойствами и может быть рекомендована для практики. В ситуациях, когда для целей исследования достаточно умеренной эффективности, а также в случаях малых выборок лучше использовать медиану. Среднее же следует использовать очень осторожно, когда нет «выбросов» данных, «хвосты» распределения коротки и т.п. Иными словами, *сложившаяся практика сравнения с помощью среднего в фитоценологических и экологических исследованиях очень часто неадекватна реальной ситуации*.

Бивес-оценка (занимая, как правило, промежуточное место между средним и медианой) рассчитывается следующим образом (Мостеллер, Тьюки, 1982):

$$b = \frac{\sum \omega_i x_i}{\sum \omega_i},$$

где

$$\omega_i = \begin{cases} [1 - ((x_i - b)/[c\mu_s])^2]^2, & \text{если } ((x_i - b)/[c\mu_s])^2 < 1; \\ 0 & \text{в противном случае;} \end{cases}$$

$\mu_s$  – медиана абсолютных отклонений  $|x_i - b|$ ,  $c$  – константа, которая берется равной 6 или 9. Поскольку  $\mu_s$  является оценкой примерно  $2/3\sigma$  ( $\sigma$  – стандартное отклонение), то при расчете бивес-оценки  $b$  не учитываются «хвосты» нормального распределения, т.е. измерения, превышающие  $4\sigma$  (при  $c = 6$ ) или  $6\sigma$  (при  $c = 9$ ). Так как мы не можем непосредственно вычислить  $b$ , не зная вектора весовых коэффициентов  $\omega_i$ , и в то же время не можем найти веса, пока не знаем  $b$ , бивес-оценка рассчитывается по приведенным формулам с использованием итеративной процедуры.

В качестве примера приведем расчет некоторых средних показателей для распределения встречаемости одного из видов (табл. 2).

Таблица 2. Распределение встречаемости *Vaccinium myrtillus* по классам pH (Emmett, Ashby, 1934; цит. по: Грейг-Смит, 1967, с. 180)

pH	4,8	4,9	5,0	5,1	5,2	5,3	5,4	5,5	5,6	5,7	5,8	5,9	6,0
	2	2	2	5	13	17	7	44	78	16	7	9	7
	2	2	2	4	8	11	3	28	54	12	1	3	1
	100	100	100	80	61,5	65	43	64	69	75	14	33	14

Мода –	5,4
Медиана –	4,9
Средняя –	5,57
Средневзвешенная –	5,25
Бивес-оценка –	5,05

### 3. Размещение растений в пределах сообществ

«Для геоботаников и фитогеографов представляет большой интерес выяснение причин, определяющих характер размещения изучаемых ими объектов – от размещения особей на небольшой площадке до размещения типов растительности и систематических групп растений на поверхности Земли... Поскольку изучение факторов, т.е. причин, определяющих распределение растений и растительности, является основной задачей геоботаники, любая методика, которая может способствовать их выявлению, очевидно, представляет ценность. В то же время нужно подчеркнуть, что обнаружение и анализ неслучайного

распределения есть исходная точка дальнейшего исследования факторов, лежащих в его основе, а не самоцель» (Грейг-Смит, 1967, с. 90, 92). «Наблюдая растительность любого участка, мы можем легко обнаружить, что особи одного вида распределены более или менее равномерно по площади, особи другого вида образуют скопления в определенных местах, а третий вид образует более или менее плотные пятна, разделенные промежутками, где особи этого вида почти полностью отсутствуют» (Василевич, 1969, с. 39).

Исследование горизонтальной структуры фитоценозов включает собственно *анализ распределения* видов (без учета конкретного положения пробных площадок на местности) и *анализ размещения* (с учетом такого положения). В первом случае методы анализа практически остались прежними (многочисленные индексы гомогенности и гомотонности сообществ были рассмотрены ранее: Миркин и др., 1989). Что касается анализа размещения, то его методы, основанные на сравнении матриц сходства «*a la* цепи Маркова», весьма корректны (Миркин, Розенберг, 1976; Миркин, Янтурин, 1981), но, к сожалению, пока все еще не нашли широкого применения в фитоценологических исследованиях (Миркин, Розенберг, 1976, 1977; Миркин и др., 1976; Конов, Розенберг, 1978).

#### 4. Сопряженность между видами

##### *Коэффициенты связи*

«Четкие связи между видами выявляются без применения специальных методов, однако часто встречаются менее тесные взаимоотношения, обнаружить которые не так легко...» (Грейг-Смит, 1967, с. 145). «Растительные сообщества относятся к числу систем, характеризующихся довольно слабой целостностью... В растительном сообществе мы, как правило, встречаемся не с функциональными зависимостями, а со стохастическими» (Василевич, 1969, с. 70).

Число индексов (коэффициентов) взаимосвязи видов весьма велико, что свидетельствует об успешном развитии в рассматриваемый период «индексологии». Правда, большинство из этих показателей так и остаются на уровне «хорошо что-нибудь с чем-нибудь сложить и на что-то поделить...».

Здесь остановимся только на одном показателе – трансформированном коэффициенте Дайса (ТКД; Миркин и др., 1972; Миркин, Наумова, 1974; Розенберг и др., 1978):

$$\text{ТКД}(A,B) = [a - \min(b,c)] / [a + \min(b,c)] ,$$

где  $a$  – число совместных встреч видов А и В;  $b$  – число встреч вида В;  $c$  – число встреч вида А. Для этого показателя были определены ошибка и  $\chi^2$ -критерий достоверности. Преимущество этого индекса (именно для этого он и создавался) – это его независимость от так называемого  $d$ -эффекта (т.е. ТДК не зависит от числа совместного отсутствия видов  $d$ )<sup>3</sup>. Причем этот эффект достигается либо, когда виды действительно «не хотят» встречаться по каким-то экологическим соображениям, либо у них существенно различается встречаемость в данном типе растительности (например, пальма и береза в Тольятти). Тем не менее в статье В.М. Ефимова (1976) был высказан ряд обвинений в адрес этого показателя, что потребовало сформулировать собственные контраргументы (Розенберг и др., 1978). Высокую эффективность и работоспособность этого коэффициента подтвердили и построенные, например, с его помощью количественные классификации луговой растительности Якутии (Кононов, Сафронеева, 1974; Кононов и др., 1976).

Хочется еще раз подчеркнуть своего рода «трезвость» при выборе того или иного показателя взаимосвязи – каждый из них *обладает своей областью применимости и не существует единого и пригодного для всех случаев индекса*. Во все более четком определении этих областей применимости и видится путь дальнейшего развития «индексологии».

### ***Коэффициенты сходства***

«Ясно, что любая попытка объективного сравнения списков видов из разных сообществ (как бы полезен ни был этот метод при сравнении флор разных регионов) основывается на таком количестве недоказанных предположений, что едва ли стоит предпринимать ее. Это тем более справедливо, что необходимо принимать во внимание неточность флористических списков как критериев при характеристике сообществ» (Грейг-Смит, 1967, с. 209). «Проблема качественной оценки или количественного определения сходства между объектами чрезвычайно важна в любой науке» (Василевич, 1969, с. 134).

Прогресс в этом разделе количественной геоботаники (Грейг-Смитом он совершенно справедливо рассматривается в разделе «Растительные сообщества. I. Описание и сравнение») за последние 30–40 лет невелик. Можно смело утверждать, что основными показателями сходства, наиболее часто используемыми в процедурах количественной классификации, остаются *коэффициенты Сьёрсенса и евклидово расстояние* (по так или

---

<sup>3</sup> Например, коэффициент линейной корреляции для четырехпольной таблицы (см., например, Грейг-Смит, 1967, с. 70) зависит от клетки  $d$ , и в случае очень большого числа совместно отсутствующих сравниваемых видов корреляция может быть близка к единице.

иначе нормированному пространству признаков-видов). Здесь для примера назову лишь практические работы по оценке сходства растительности аласов Центральной Якутии (Скрябин, Бурцева, 1976) и в подзоне северной тайги Республики Коми (Галанин, 1981, 1983).

## 5. Соотношение растительности с факторами местообитания

«Поскольку различия в растительности определяются различиями в условиях среды, следует ожидать, что они коррелируют друг с другом. Если различия в растительности или в факторах местообитания малы, выявить корреляцию с помощью качественного исследования очень трудно, и существующую связь можно установить лишь путем объективной оценки соответствующих количественных данных» (Грейг-Смит, 1967, с. 170).

Наиболее часто цитируемой и в то же время наиболее «идеологической» расплывчатой областью экологии (фитоценологии) является некоторая совокупность методов, называемая «биоиндикацией» («геоботанической индикацией»). Хотя истоки наблюдений за индикаторными свойствами биологических объектов можно найти в трудах естествоиспытателей самой глубокой древности (правда, П. Грейг-Смит не рассматривает проблемы количественной геоботанической индикации, а в монографии В.И. Василевича (1969) они уже обсуждаются), до сих пор отсутствуют стройная теория и адекватные методы биоиндикации.

Относительно благополучно дело обстоит с описательным объяснением терминов. Например, согласно определению Н.Ф. Реймерса (1990, с. 44) «Биоиндикатор: группа особей одного вида или сообщество, по наличию, состоянию и поведению которых судят об изменениях в среде, в том числе о присутствии и концентрации загрязнителей... Сообщество индикаторное – сообщество, по скорости развития, структуре и благополучию отдельных популяций микроорганизмов, грибов, растений и животных которого можно судить об общем состоянии среды, включая ее естественные и искусственные изменения». Безусловно, объективные факты свидетельствуют о существовании тесного влияния факторов среды на биотические процессы экосистемы (плотность популяций, динамику видовой структуры, поведенческие особенности). Такие факторы среды, как свет, температура, водный режим, биогенные элементы (макро- и микроэлементы) и другие, имеют функциональную важность для организмов на всех основных этапах жизненного цикла. Однако можно использовать обратную закономерность и судить, например, по видовому составу организмов о типе физической среды. Поэтому «биоиндикация – это определение биологически значимых нагрузок на основе реакций на них живых

организмов и их сообществ. В полной мере это относится ко всем видам антропогенных загрязнений» (Кривоуцкий и др., 1988, с. 11). С этих позиций основной задачей биоиндикации является разработка методов и критериев, которые могли бы адекватно отражать уровень антропогенных воздействий с учетом комплексного характера загрязнения и диагностировать ранние нарушения в наиболее чувствительных компонентах биотических сообществ.

Биоиндикация, как и мониторинг, осуществляется на различных уровнях организации биосферы: макромолекулы, клетки, органа, организма, популяции, биоценоза (Биоиндикация: теория., 1994). Очевидно, что сложность живой материи и характер ее взаимодействия с внешними факторами возрастают по мере повышения уровня организации. В этом процессе биоиндикация на низших уровнях организации должна диалектически включаться в биоиндикацию на более высоких уровнях, где она предстает в новом качестве и может служить для объяснения динамики более высокоорганизованной системы.

Рассмотрим частную задачу биоиндикации, в которой оценивается степень влияния произвольного фактора среды на некоторую совокупность измеренных показателей экосистемы. Если рассматривать ее формальную сущность, то тут фигурируют две векторные переменные: 1)  $Y$  – характеризует состояние экологического объекта (растительности или гидробиологического сообщества); 2)  $X$  – состояние среды (например, почвы или воды). Будем считать, что обе эти переменные являются *номинальными*, т.е. измерены в некоторых упорядоченных шкалах. Это позволяет оценить плотность распределения вероятности  $p_0(x)$  значений  $y_i$  для каждого интервала значений  $X$  (хотя с точки зрения классической математической статистики эта процедура относится к классу некорректных задач). Если распределение вероятности параметра экосистемы по оси значений фактора является равномерным, то можно говорить о том, что параметр не является индикатором фактора  $X$ .

Пусть задан некоторый критерий –  $c$ -мера отклонения закона распределения вероятности, восстановленного по эмпирическим данным, от равномерного распределения. Тогда при  $c$ , превышающем некоторое пороговое значение  $c_{пор}$ , можно говорить об индикаторной значимости переменной  $Y$ , а диапазоны значений  $X$ , где функция распределения принимает экстремальные значения, можно считать индицируемыми диапазонами фактора.

С точки зрения математики поставленная задача биоиндикации в реальных условиях относится к *классу плохо формализуемых задач*, поскольку характеризуется следующими особенностями:

- существенной *многомерностью* факторов среды и измеряемых параметров экосистем;
- сильной *взаимобусловленностью* всего комплекса измеренных переменных, не позволяющей выделить в чистом виде функциональную связь двух индивидуальных показателей  $F(y,x)$ ;
- *нестационарностью* большей части информации об объектах и среде;
- *трудоемкостью* проведения всего комплекса измерений в единых координатах пространства и времени, в результате чего обрабатываемые данные имеют обширные пропуски.

В связи с этим нахождение адекватной связи индикаторов и индицируемых факторов является типичной операцией с «размытыми» множествами (Заде, 1974, 1976), а следовательно, характеризуется существенной неопределенностью (стохастичностью). В то же время к настоящему моменту сложились условия, позволяющие преодолеть некоторую математическую «ущербность» биоиндикации:

- сформированы банки многолетних данных по наблюдениям за природными экосистемами;
- разработан и апробирован ряд методов и математических моделей интегральной оценки состояния сложных систем различного типа, позволяющих, по терминологии А.П. Левича и А.Т.Терехина (1997, с. 329), осуществлять «поиск детерминации и распознавание образов в многомерном пространстве экологических факторов для выделения границ между областями нормального и патологического функционирования экосистем»;
- развиваются аппаратные и программные информационные компьютерные технологии, позволяющие анализировать необходимые массивы экологических данных;
- существует огромный объем неформальных знаний высококвалифицированных специалистов, частично сконцентрированный в методических разработках.

П. Грейг-Смит (1967, с. 174), характеризуя возможность использования количественных методов для оценки связей в системе «растительность – среда», исходит из четырех ситуаций, когда растительность и среда описываются либо количественными, либо качественными данными. Основными методами анализа тогда (да и во многом сегодня) были: *сравнение средних, характер распределения параметров растений* (например, встречаемости) *вдоль градиента среды и регрессионный анализ.*

Рассмотрим в этой связи некоторые специальные методы и инструментальные средства построения автоматизированных и неавтомати-

зированных экспертных систем для решения задач геоботанической индикации.

### ***Прямой градиентный анализ как основа количественных биоиндикационных исследований***

Учитывая непреходящую важность прямого градиентного анализа, воспользуемся случаем (забегая несколько вперед; см. далее раздел 7) и напомним как «классические», так и некоторые «современные» его составляющие (Розенберг, 1998).

В начале 50-х гг. XX века американский эколог и фитоценолог Роберт Уиттекер (1980) предложил «интуитивно-статистический» метод прямого градиентного анализа применительно к растительности горных систем, сущность которого сводилась к следующему. Горная система представлялась в двухмерной схеме, где по одной оси откладывался комплексный градиент высоты над уровнем моря (гидротермический градиент), а по другой – топографический (различные экспозиции склона). В этих осях координат строились распределения отдельных популяций растений и некоторые характеристики видообразия и продуктивности, что позволило подтвердить концепцию континуума Раменского – Глисона. Следует заметить, что эти пионерные работы по прямому градиентному анализу встретили достаточно острую критику (см. обзор: Миркин, Наумова, 1998), которая в первую очередь касалась произвольности (субъективности) выбора осей ординации. Р. Уиттекер (Whittaker, 1973) убедительно показал, что, «субъективно» выбирая эти оси, исследователь опирается на значительный багаж знаний об экологии видов и осуществляет ординацию по бесспорно ведущим градиентам.

Другой недостаток метода Уиттекера (а именно его интуитивно-статистический характер) был исправлен в количественной модификации прямого градиентного анализа, разработанного уфимскими геоботаниками под руководством Б.М. Миркина (1971; Миркин, Розенберг, 1979; Миркин и др., 1989). Построение графиков изменения наблюдаемого признака (вес, встречаемость, обилие и пр.) для отдельных видов в зависимости от изменения выбранного для ординации видов фактора сопровождается расчетом некоторых статистических характеристик:

- *средневзвешенная напряженность фактора:*

$$\bar{X}_i = \sum_{j=1}^r p_{ij} X_j ;$$

- *средневзвешенная дисперсия:*

$$D_i = \sum_{j=1}^r p_{ij} (X_j - \bar{X}_i)^2,$$

где  $X_j$  – середина  $j$ -й градации исследуемого фактора;  $r$  – число этих градаций,

$$p_{ij} = \frac{m_{ij}}{n \sum (m_{ik}/n_k)}, \quad \sum n_j = N,$$

где  $n_j$  – число наблюдений в  $j$ -й градации фактора;  $m_{ij}$  – число встреч вида  $i$  в этой градации,  $N$  – общее число наблюдений:

- кроме того, рассчитывается *сила влияния исследуемого фактора на данный вид*, получаемая из однофакторного дисперсионного анализа ( $\eta^2$ ).

С помощью средневзвешенных значений фактора каждый вид «привязывается» к определенному отрезку градиента, а малое значение  $D_i$  и большое  $\eta^2$  свидетельствуют о высокой индикаторной роли вида  $i$ , причем преимущество следует отдавать оценке по значениям  $D_i$ , (Миркин, 1971) – чем меньше  $D_i$ , тем вид занимает более узкий участок по оси данного фактора и, следовательно, является хорошим его индикатором.

Думается, что будут весьма полезны некоторые сравнительно новые представления о силе влияния фактора. В конце 60-х гг. XX века прошла дискуссия по вопросу об оценке показателя силы влияния фактора  $\eta^2$  при дисперсионном анализе, в которой участвовали Н.А. Плохинский, В.Ю. Урбах и Э.Х. Гинзбург, позднее интерес к этой проблеме проявился вновь (Плохинский, 1978; Розенберг, Долотовский, 1988; Розенберг и др., 1993). Сравнению подвергались в основном три показателя – квадрат корреляционного отношения Пирсона (**P**), рекомендованный для биологических исследований Н.А. Плохинским, показатель Миллса – Лукомского (**ML**), пропагандируемый В.Ю. Урбахом, и коэффициент Снедекора (**S**), отстаиваемый Э.Х. Гинзбургом. Аргументы за и против в этой дискуссии были подробно проанализированы (Розенберг, Долотовский, 1988).

Признавая состоятельность всех трех оценок силы влияния фактора в дисперсионном анализе, участники дискуссии стремились в большей степени доказать «нерботоспособность» противного подхода и часто «уходили» от ответов на вопрос о той или иной особенности «своего» показателя. Еще одной отличительной особенностью этой дискуссии было привлечение в качестве основного аргумента моделирования различных случайных дисперсионных комплексов на ЭВМ вместо попытки строгого аналитического анализа предлагаемых показателей.

Было доказано (Розенберг, Долотовский, 1988), что показатели Плохинского и Миллса – Лукомского полностью эквивалентны, так как оказываются связанными линейной зависимостью – их различие состоит только в границах изменений:  $P$  меняется от 0 до +1, а  $ML$  – от  $-(r-1)/(N-r)$  также до +1. Отсюда следует, что показатель  $ML$  неизбежно должен давать отрицательные значения в тех случаях, когда  $P < (r-1)/(N-r)$ . Иными словами, показатель  $ML$  полностью дублирует  $P$ ; границы изменения последнего от 0 до +1 делают его предпочтительнее в силу удобства. Преимуществом показателя  $S$  является разработанная Э.Х. Гинзбургом оптимальная схема планирования структуры дисперсионного комплекса. При малых  $n$  величина и достоверность показателей существенно зависят от структуры исследуемого материала, что обуславливает необходимость вообще осторожно использовать любые биометрические методы; для больших  $n$  все показатели практически близки.

При расчете силы влияния фактора необходимо учитывать следующий факт. В эксперименте, как справедливо подчеркивает В.И. Василевич (1969, с. 8), трудно добиться «...изменения одного фактора при постоянстве других». Таким образом, более корректным представляется использование 2-, 3-факторных и т.д. комплексов, которые построить на практике довольно сложно. Поэтому некоторые авторы при оценке  $P = \eta^2$  прибегают к ряду эвристических процедур. Один из способов состоит в сглаживании влияния неконтролируемых факторов на изучаемый фактор (Розенберг, 1984; Розенберг и др. 1993; Биоиндикация: теория..., 1994).

Пусть нам задан однофакторный равномерный дисперсионный комплекс для количественных признаков. Количество градаций одного комплекса равно  $r$ , а число наблюдений в каждой градации –  $n$ . Анализ дисперсионного комплекса начинается с рассмотрения ряда частных средних для каждой градации фактора:

$$\bar{X}_j = (\sum X_{ij}) / n,$$

где  $X_{ij}$  – значение исследуемого признака при  $i$ -м наблюдении в  $j$ -й градации ( $j = 1, r$ ).

Ряд частных средних может быть задан линией регрессии. Следует отметить, что «течение» эмпирической линии регрессии почти никогда не бывает плавным, так как большое влияние на изменение частных средних оказывают всевозможные «шумы», в роли которых выступают как случайные (неучтенные) факторы, так и тесно скоррелированные с изучаемым. Таким образом, кривые встречаемости признака в зависимости от изучаемого фактора оказываются многовершинными.

Наблюдаемая многовершинность противоречит представлениям о существовании только одного экологического оптимума распределения ви-

да по градиенту среды (см., например, Уиттекер, 1980) и резко завышает значение силы влияния фактора  $\mathbf{P}$  за счет увеличения дисперсии, объясняющей влияние фактора ( $C_x$ ). Наличие многовершинности для кривой встречаемости признака можно объяснить двумя причинами: либо большим числом градаций фактора, либо большей степенью зависимости изучаемого фактора от неучтенных показателей. Замечу, что это понимал и П. Грейг-Смит (1967, с. 182): «С этой трудностью (*распределение значений фактора среды для всей совокупности описаний значительно отличается от нормального.* – Г.Р.) пришлось бы столкнуться, если бы кривая значений фактора имела бы два отчетливо выраженных максимума, что указывало бы на возможное наличие в выборке образцов из двух совершенно разных местообитаний».

В первом случае путем укрупнения градаций фактора можно добиться одновершинной кривой частных средних (действительно, для двух градаций фактора кривая частных средних всегда будет одновершинной, а для трех – в восьми случаях из девяти). Однако слишком крупные градации фактора не позволяют вскрыть его «тонкое» влияние, тем самым огрубляя и заметно понижая значение  $\mathbf{P}$  и соответственно снижая общую эффективность дисперсионного анализа. На рис. 1 показана некоторая теоретическая модель зависимости частных средних от изучаемого фактора, разбитого на шесть градаций. Если теперь представить, что фактор разбит на две градации, то нетрудно убедиться, что для этого случая значения частных средних совпадут с общей средней и, следовательно,  $C_x = 0$ , а отсюда и  $\mathbf{P} = 0$ , что заведомо неверно.

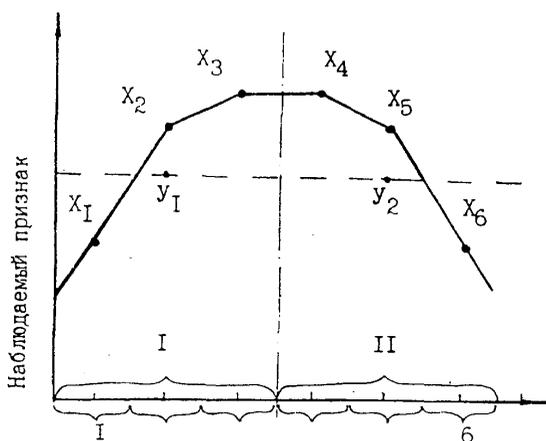


Рис. 1. Теоретическая модель зависимости частных средних от изучаемого фактора

В случае зависимости изучаемого фактора от других, применяя схему многофакторного дисперсионного анализа, можно представить общую дисперсию как сумму дисперсий, вызванных воздействием отдельных составляющих и их сочетаний. Однако если эксперимент не планировался для многофакторного дисперсионного анализа, то отсутствуют исходные данные для построения дисперсионного комплекса и, следовательно, отсутствует возможность воспользоваться результатами этого анализа.

Одним из простейших методов, позволяющих элиминировать случайные колебания эмпирической линии регрессии, вскрыть зависимость признака от изучаемого фактора, так сказать, в чистом виде и получить более плавную и одновершинную теоретическую кривую, является метод выравнивания способом скользящей средней. Будем определять средние после выравнивания по следующим формулам:

$$\overline{\overline{X}}(i) = [\overline{X}(i-1) + \overline{X}(i) + \overline{X}(i+1)]/3$$

для всех  $1 < i < r$ ; крайние значения ряда выравненных частных средних определим следующим образом:

$$\overline{\overline{X}}(1) = [2\overline{X}(1) + \overline{X}(2)]/3 \quad \text{и} \quad \overline{\overline{X}}(r) = [\overline{X}(r-1) + 2\overline{X}(r)]/3 ,$$

где  $\overline{X}(i)$  – частные средние исходного дисперсионного комплекса,  $\overline{\overline{X}}(i)$  – выравненные значения частных средних.

### *Изучаемый фактор*

Эти же данные можно представить как некоторый новый теоретический однофакторный дисперсионный комплекс, где в качестве наблюдений, соответствующих  $i$ -й градации фактора, выступают частные средние исходного комплекса (по три значения, приведенных выше, для каждой градации). Легко убедиться, что общие средние исходного и теоретического дисперсионного комплексов равны между собой. Из этого факта легко получается следующая зависимость между дисперсиями:

$$C_x = n\overline{C}_y / 3 ,$$

где  $\overline{C}_y$  – общая дисперсия теоретического комплекса. После преобразований общая дисперсия исходного комплекса  $C_y$  представляется в следующем виде (зависимость между дисперсиями для безграничных связанных рядов в общем случае получена Е.Е. Слуцким, 1927):

$$C_y = n\overline{C}_x / 3 + \overline{C}_z ,$$

где  $n\bar{C}_x/3$  и  $\bar{C}_z$  – это соответственно дисперсии, связанные с влиянием изучаемого фактора после выравнивания ряда частных средних и воздействием случайных и неучтенных факторов. Тогда сила влияния фактора после выравнивания определяется по следующей формуле:

$$P = \eta_x^2 = t C_x / C_y,$$

где  $t$  – коэффициент, прямо пропорциональный числу наблюдений в каждой градации комплекса, являющегося исходным для выравнивания ( $n$ ), и обратно пропорциональный количеству частных средних, по которым производится выравнивание (для данного частного случая  $t = n/3$ ).

Выравнивание способом скользящей средней можно повторять многократно. Последним следует считать тот шаг  $h$ , после которого кривая частных средних по градиентам фактора становится достоверно одновершинной, что проверяется сравнением средних по критерию Стьюдента. При количестве выравниваний  $h > 1$  значение  $t = 1$ , так как исходным для получения теоретического комплекса  $h$  является комплекс  $(h - 1)$ , для которого число наблюдений в каждой градации равно количеству частных средних, по которым проводится выравнивание.

Получение на практике равномерных комплексов возможно лишь в немногих случаях, когда градации фактора можно спланировать до проведения эксперимента (например, пространственные градиенты). Для количественных неравномерных комплексов (чаще всего это эоклины) значительно усложняется процесс нахождения коэффициента  $t$ , так как он является функцией от числа наблюдений. В качестве оценки величины  $t$  можно использовать среднее значение числа наблюдений в каждой градации фактора.

В качестве примера процедуры выравнивания рассмотрим распределение проективного покрытия *Puccinellia tenuiflora* и *Plantago canescens* в зависимости от засоления (Намский район Якутии; Бурцева, 1978, с. 145–149). Реальное распределение этих видов имеет двухвершинный характер с «провалом» в классах засоленности 2,5–3 и 2–2,5 соответственно. После процедуры однократного выравнивания были получены результаты, представленные в табл. 3.

Таблица 3. Сравнение силы влияния фактора засоления на распределение двух видов солончаковых лугов Средней Лены до и после выравнивания

Вид	До выравнивания		После выравнивания	
	Сила влияния ( $\eta^2$ )	Достоверность	Сила влияния ( $\eta^2$ )	Достоверность
<i>Puccinellia tenuiflora</i>	0,40	46,7	0,17	14,3
<i>Plantago canescens</i>	0,25	23,2	0,08	6,1

Таким образом, после выравнивания влияние фактора засоленности заметно снизилось (в 2–3 раза), хотя и осталось достоверным. Это делает более объективным наши представления о характере связи этих видов с данным фактором.

Изложенный выше метод позволяет, минуя схему многофакторного анализа, извлечь более объективную информацию о влиянии одного фактора на изучаемый признак. Однако следует отметить, что данный алгоритм является приближенным и может использоваться только тогда, когда по каким-либо причинам невозможно построение многофакторных комплексов. Таким образом, *при планировании того или иного эксперимента исследователь должен не только иметь рабочую гипотезу, но и четко представлять тот математический метод, с помощью которого он будет подтверждать или опровергать ее.* Если условия корректности эксперимента, напрямую или косвенно содержащиеся в любом математическом методе, соблюдены, то и использование данного метода можно считать обоснованным.

Из «свежих» работ по прямому градиентному анализу остановлюсь на монографии С.И. Мироновой (2000). Основные ее результаты количественной обработки сосредоточены в главе «Ординационный анализ техногенной растительности», которая может быть названа «Увертюрой к однофакторному прямому градиентному анализу», континуум изменений растительности развернут в хроноклины, которые проанализированы в традиционных подходах школы уфимских фитоценологов (построение графиков распределения проективного покрытия видов во времени, определение средневзвешенной и силы влияния фактора времени на распределение видов). При этом анализ проведен для отвалов Западной и Южной Якутии по пяти группам видов («сквозного» распространения, тяготеющих к ранним, средним и более поздним стадиям сукцессии, имеющим двухвершинное распределение, т.е. находящихся под существенным влиянием не только фактора времени); кроме того, различался верхний (поверхность и склоны отвалов) и нижний уровни («пазухи» и берега искусственных водоемов). Теперь становится ясным, почему данный однофакторный анализ хроноклинов можно назвать лишь «увертюрой» – если бы С.И. Миронова довела обработку до конца (а фактически материал предполагает трехфакторную схему градиентного анализа: сукцессионное время, запад–юг, верх–низ), то можно было бы с полным правом назвать эту главу «Гимном прямому градиентному анализу». Правда, в последнем разделе главы С.И. Миронова предпринимает попытку качественного сравнения хроноклинов разных вариантов и приходит к ряду интересных выводов (закономерное изменение фитосоциологического спектра по пяти классам сообществ, снижение гамма-разнообразия низких ме-

стообитаний за счет выпадения неустойчивых луговых видов, различная роль в ходе сукцессии злаковых и бобовых компонент техногенных экосистем, наконец, контрастность видового состава разных этапов сукцессии, оцененная по степени сходства).

### *Распознавание образов как метод биоиндикации*

С позиций П. Грейг-Смита (1967, с. 174; перебор использования качественных и количественных данных о растительности и среде) методы распознавания образов (см., Миркин, Розенберг, 1978; Розенберг, 1984, 1994, 1998) следует отнести к методам, которые используют качественные данные о среде (или «балльно-количественные») и качественные (или количественные) данные о растительности. Отличие от прямого градиентного анализа состоит в том, что биоиндикация методами распознавания осуществляется не по видам-индикаторам, а по растительным сообществам в целом.

Распознавание образов (обучение с «учителем») базируется на сочетании человека и ЭВМ. В этом случае задача распознавания решается в два этапа: первый – нахождение признаков-видов, характеризующих принадлежность описания-сообщества к классу среды, указанному «учителем» (в принципе эту задачу решает прямой градиентный анализ), второй этап – построение решающего правила для такого разбиения градиента среды на классы, чтобы качество распознавания было наилучшим (рекомендуется использовать правило близости и дихотомическое деление градиента исследуемого фактора; см. Миркин, Розенберг, 1978, с. 179–180). На этом этапе (рис. 2) экзаменационный объект (С) относится к одной из двух частей градиента (F или G); в определении принадлежности объекта к классам первого порядка участвуют I(A) и V(B) классы обучения. Координаты точки D определяются по классической теореме Пифагора:

$$AD = (AB^2 - CB^2 + AC^2) / (2 AB) .$$

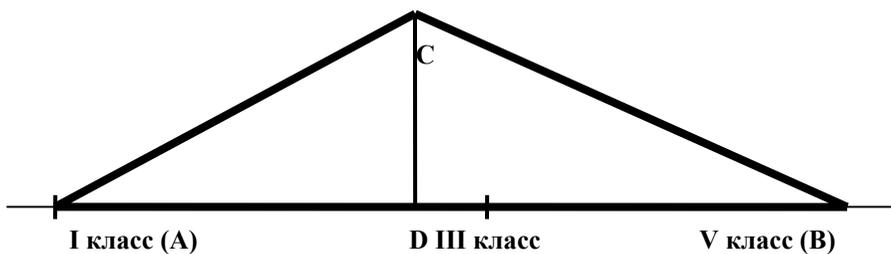


Рис. 2. Схема реализации правила близости

После того как объект отнесен к одной из частей градиента (на рис. 2 – к классу I), происходит новое разбиение на классы второго порядка (в данном случае – между классами I и III; при этом в качестве обучающих объектов участвуют классы I(A) и III(B); объект ближе к III классу). В принципе путем увеличения числа обучающих классов можно добиться сколь угодно мелкого разбиения градиента фактора; однако с биологической точки зрения вряд ли целесообразно делать более трех итераций (разбиение на  $2^3 = 8$  классов; естественно, необходимо учитывать протяженность оцениваемого градиента и амплитуды распределения видов-индикаторов). По «обучающей» последовательности, следует провести укрупнение полученных классов с тем, чтобы вероятность правильного распознавания в целом (средняя вероятность правильного распознавания) была как можно выше (например, более 0,7). В табл. 4 приведены некоторые результаты распознавания условий среды по растительности (Миркин, Розенберг, 1978; Бурцева и др., 1978).

**Таблица 4. Результаты распознавания условий среды по растительности**

Объект	Фактор	Число достоверно распознаваемых градаций	Объем каждого из пяти классов обучения (число описаний)	Средняя относительная частота правильного распознавания
Солончаковатые луга Башкирского Предуралья	Засоление почвы	3	15	0,726
Солончаковатые луга Башкирского Зауралья	Засоление почвы	4	12	0,724
Солончаковатые луга поймы р. Лены	Засоление почвы	2	15	0,705
Горные луговые степи Северо-Восточного Хангая (МНР)	Гумусированность	3	12	0,795
	Скелетность	2	12	0,793
	Кислотность (рН)	3	15	0,731

Например, для солончаковатых лугов поймы р. Лены (Бурцева и др., 1978) с вероятностью выше 0,7 по растительности было определено две градации засоления (менее 1,2% плотного остатка и более). При этом малое число распознаваемых градаций в условиях солончаковатых лугов Якутии представляется интересным экологическим феноменом (эколого-фитоценотический диапазон выборки, полученной на лугах Намского района Якутии, укорочен в силу злостности сульфатно-хлоридного засоления, адаптации растений к солончаковатости почвы (растительность лугов поймы Средней Лены представлена солестойкими экотипами даже тех видов, которые принято считать гликофитами), резкой переменностью климата Якутии, что ведет к переменности режима засоления).

## *Методы самоорганизующегося и эволюционного моделирования*

Классический регрессионный анализ (см., Шитиков и др., 2005) как эмпирико-статистический метод индикации с развитием кибернетических подходов претерпел заметные изменения. В математической кибернетике различают два вида итеративных процессов развития систем (Ивахненко и др., 1976):

- *адаптация*, при которой экстремум (цель движения системы) остается постоянной;
- *эволюция*, при которой движение сопровождается изменением положения экстремума.

*Самоорганизация* связана только с адаптационными механизмами подстройки реакций системы (например, изменением значений весовых коэффициентов). Под термином «самоорганизация» понимается «процесс самопроизвольного (спонтанного) увеличения порядка или организации в системе, состоящей из многих элементов, происходящий под действием внешней среды» (Ивахненко и др., 1976, с. 28).

Основы самоорганизации были предметом исследования многих выдающихся ученых: Дж. фон Неймана, Н. Винера, У.Р. Эшби и др.; большой вклад в развитие этого направления внесли работы украинских кибернетиков под руководством А.Г. Ивахненко (1975, 1982 и др.), разработавших целый класс адаптивных самоорганизующихся моделей (англ. *self-organization models*), который можно было бы назвать «интеллектуальным обобщением» эмпирико-статистических методов.

Можно отметить следующие принципы самоорганизации математических моделей (Ивахненко, 1975, 1982):

- принцип *неокончателных решений* (предложен Д. Габором и заключается в необходимости сохранения достаточной «свободы выбора» нескольких лучших решений на каждом шаге самоорганизации);
- принцип *внешнего дополнения* (базируется на теореме К. Гёделя и заключается в том, что только внешние критерии, основанные на новой информации, позволяют синтезировать истинную модель объекта, скрытую в зашумленных экспериментальных данных);
- принцип *массовой селекции* (предложен А.Г. Ивахненко и указывает наиболее целесообразный путь постепенного усложнения самоорганизующейся модели с тем, чтобы критерий ее качества проходил через свой минимум).

Для возникновения самоорганизации необходимо иметь исходную структуру, механизм случайных ее мутаций и критерии отбора, благодаря которому мутация оценивается с точки зрения полезности для улучшения качества системы. Таким образом, исследователь задает только *исходную*

*организацию, список переменных, критерии качества*, формализующие цель оптимизации, и *правила*, по которым модель может изменяться (самоорганизовываться или эволюционировать). Причем сама модель может принадлежать самым различным типам: линейная или нелинейная регрессия, набор логических правил или любая другая модель.

Самым традиционным алгоритмом создания самоорганизующихся моделей является *метод группового учета аргументов (МГУА)*.

Принцип массовой селекции, используемый в алгоритмах МГУА, как и многие другие идеи кибернетики, заимствует действующие природные механизмы и схематически повторяет методы селекции растений или животных. Так называемое полное описание объекта  $Y = f(x_1, x_2, x_3, \dots, x_m)$  заменяется несколькими рядами «частных» описаний:

- 1-й ряд селекции:  $y_1 = \varphi(x_1, x_2), y_2 = \varphi(x_1, x_3), \dots, y_s = \varphi(x_{m-1}, x_m)$ ;

- 2-й ряд селекции:  $z_1 = \varphi(y_1, y_2), z_2 = \varphi(y_1, y_3), \dots, z_p = \varphi(y_{s-1}, y_s)$ , где  $s = C_n^2$ ,  $p = C_s^2$  и т.д.

Входные аргументы и промежуточные переменные сопрягаются попарно, и сложность комбинаций на каждом ряду обработки информации возрастает (как при массовой селекции), пока не будет получена единственная модель оптимальной сложности. Поскольку каждое частное описание является функцией только двух аргументов, его коэффициенты легко определить по данным обучающей последовательности.

Из ряда в ряд селекции «пропускается» только некоторое количество самых регулярных переменных, степень регулярности которых оценивается по специальным критериям (см.: Розенберг и др., 1994). Ряды селекции наращиваются до тех пор, пока регулярность повышается. Как только достигнут минимум ошибки, селекцию во избежание «инцухта» следует остановить. Практически рекомендуется остановить селекцию даже несколько раньше достижения полного минимума, как только ошибка начинает падать слишком медленно, это приводит к более простым и более достоверным уравнениям.

Самоорганизующиеся модели служат в основном для прогнозирования поведения и структуры экосистем, так как по самой логике их построения участие исследователя в этом процессе сведено к минимуму. Можно привести ряд конкретных примеров использования алгоритмов МГУА (ссылки на эти работы см.: Шитиков и др., 2005): для долгосрочных прогнозов экологической системы оз. Байкал, системы «хищник–жертва», прироста деревьев, прогнозирования токсикологических показателей поллютантов, оценки динамики численности сообществ зооплankтона и др.

*Эволюционное моделирование* (Фогель и др., 1969; Букатова, 1979; Букатова и др., 1991) представляет собой существенно более универсальный способ построения прогнозов макросостояний системы в условиях, когда полностью отсутствует апостериорная информация, а априорные данные задают лишь предысторию этих состояний. Общая схема алгоритма эволюции выглядит следующим образом:

- задается исходная организация системы (в эволюционном моделировании в этом качестве может фигурировать, например, конечный детерминированный автомат Мили; см.: Букатова, 1979);
- проводят случайные «мутации», то есть изменяют случайным образом текущий конечный автомат;
- отбирают для дальнейшего развития ту организацию (тот автомат), которая является лучшей в смысле некоторого критерия, например, максимальной точности предсказания последовательности значений макросостояний экосистемы.

Критерий качества модели в этом случае мало чем отличается, например, от минимума среднеквадратической ошибки на обучающей последовательности метода наименьших квадратов (со всеми вытекающими отсюда недостатками). Однако в отличие от *адаптации* (самоорганизации) в *эволюционном моделировании* структура решающего устройства мало меняется при переходе от одной мутации к другой, то есть не происходит перераспределение вероятностей, которые бы закрепляли мутации, приведшие к успеху на предыдущем шаге. Поиск оптимальной структуры происходит в большей степени случайным и нецеленаправленным образом, что затягивает процесс поиска, но обеспечивает наилучшее приспособление к конкретным изменяющимся условиям.

Наконец отмечу дальнейшее развитие методов эволюционного моделирования, нашедшее свое выражение в разработке моделей временных рядов экологических переменных с использованием трехэтапного эволюционного предсказывающего алгоритма (Розенберг и др., 1994) и автоматов Мили для многосимвольных целочисленных временных рядов (Морозов, 2000).

Собственно в экологии растений и растительности методы самоорганизации используются сравнительно редко (апробированы в основном мной на различных объектах; см. Розенберг, 1984), чему есть вполне объективное объяснение: для эффективности этих подходов требуются сравнительно большие временные (экологические) ряды. В частности, представляет интерес сравнение результатов описания и прогнозирования урожайности аласных сенокосов Якутии по гидрометеорологическим данным (Кононов, Розенберг, 1981; Розенберг, 1984, с. 68–70 и 159–160) с помощью классической линейной регрессии и МГУА. Для построения

прогнозирующих моделей была взята средняя хозяйственная урожайность аласов ( $Y$ , в ц/га) для Чурапчинского района Якутии и 13 гидрометеорологических параметров ( $X_i$ ,  $i=1,13$ ) с 1941 по 1975 г. Вся исходная информация была разделена на две части: обучающую (до 1953 г.) и экзаменационную последовательности. В табл. 5 (см. также: Розенберг, 1984, с. 159) приведены средние относительные ошибки прогнозирования урожайности разными методами. При этом был получен весьма удовлетворительный прогноз по МГУА (ошибка – 15% при прогнозировании на 5 лет и 20% – на 10 и 20 лет) и определены наиболее важные для прогнозирования гидрометеорологические факторы – ими оказались суммы осадков за июнь, июль и весь теплый период, суммы температур от 0 до 5°C и абсолютный минимум температуры за июнь.

*Таблица 5. Средние относительные ошибки прогноза (в%) урожайности аласных сенокосов по гидрометеорологическим факторам*

Модель	Прогноз по «обучению»	«Экзамен»		
		Прогноз на 5 лет	Прогноз на 10 лет	Прогноз на 20 лет
Классический регрессионный анализ (линейная модель)	9,2	27,6	54,9	222,0
МГУА	10,0	14,6	18,1	20,0

Естественно, подтверждение работоспособности модели, построенной по МГУА, требует специальных дополнительных исследований на сроки после 1975 г. К сожалению, такой работы не проводилось, но мне удалось воспользоваться результатами близких исследований Е.И. Бурцевой (Бурцева, 2000; Бурцева, Десяткина, 2000) по геоэкологическим аспектам развития регионов Крайнего Севера. По данным на 1995 г. (прогноз по рекуррентной модели, представленной в работе Розенберга (1984, с. 160), различия прогнозируемой и реальной урожайности составляли 16,4% от последней, что следует признать вполне удовлетворительным результатом. Если дополнить ряды наблюдений (например, с 1941 по 1995 г.) и увеличить обучающую последовательность (с 13 до 20–25 лет), то, возможно, удастся улучшить и качество прогнозирующей модели.

## **6. Растительные сообщества. I. Описание и сравнение**

«Сравнение сообществ только по составу имеющихся видов без какого-либо указания на обилие – грубое и нечувствительное средство их характеристик» (Грейг-Смит, 1967, с. 203).

На саммите ООН в Жанейро (1992 г.) биоразнообразие было определено как «вариабельность живых организмов из всех источников, включающих

*inter alia* (лат. *среди прочих*) наземные, морские и прочие водные экосистемы и экологические комплексы, частью которых они являются: это включает разнообразие в пределах вида, разнообразие видов и разнообразие экосистем». Это официальное определение с точки зрения буквы закона, поскольку вошло в Конвенцию ООН по вопросам биоразнообразия (статья 2).

П. Грейг-Смит практически не обсуждает проблем количественного измерения биоразнообразия, более того, само понятие «биоразнообразия» (в контексте *показатель биоразнообразия*) используется им всего два раза: при обсуждении способа измерения отклонения от случайного распределения с применением индекса Симпсона (с. 108) и при описании соотношения между числом видов и числом особей с применением индекса Уильямса (с. 205–206, 232). Именно в изучении биологического разнообразия как теоретических проблем (прежде всего работы Р. Уиттекера), так и методических (наверное, самым интересным и «модным» является фрактальный подход к изучению видовой структуры сообществ) за последние 50 лет достигнуты наибольшие успехи, в том числе и в фитоценологии.

Исследование фрактальных структур – одно из молодых, но очень быстро развивающихся направлений современной математики. Фундаментальные и естественнонаучные приложения фрактальной геометрии впервые нашли отражение в трудах Бенуа Мандельброта (Mandelbrot, 1975). Широкое проникновение идей фрактальной геометрии в биологические дисциплины позволило поставить вопрос о фрактальной структуре одного из ключевых объектов исследования современной экологии – биотического сообщества (Маргалев, 1992; Пузаченко Ю., Пузаченко А., 1996; Азовский, Чертопруд, 1997; Пузаченко, 1997; Harte et al., 1999; Иудин и др., 2003; Iudin, Gelashvily, 2003; Гелашвили и др., 2004 и др.).

«Фракталы вокруг нас повсюду, и в очертаниях гор, и в извилистой линии морского берега. Некоторые из фракталов непрерывно меняются, подобно движущимся облакам или мерцающему пламени, в то время как другие, подобно деревьям или нашим сосудистым системам, сохраняют структуру, приобретенную в процессе эволюции» – из книги Х.О. Пайгена и П.Х. Рихтера (1993, с. 7). Вот еще одна цитата из работы В.Н. Беклемишева (1964, с. 37): «... живой организм (*и экосистема*. – Г.Р.) не обладает постоянством материала – форма его подобна форме пламени, образованного потоком быстро несущихся раскаленных частиц; частицы сменяются, форма остается». Иными словами, подобно мольеровскому Журдену, экологи давно говорят прозой (о фракталах), не догадываясь об этом...

Родоначальник теории фракталов Б. Мандельброт предложил такое определение фрактала: «Фракталом называется структура, состоящая из частей, которые в каком-то смысле подобны целому» (цит. по: Федер, 1991, с. 9). Однако строгого и полного определения фракталов пока не существует.

Не останавливаясь на достаточно строгом математическом описании фракталов, мультифракталов, фрактальных размерностей и пр.<sup>4</sup>, постараюсь дать «концептуальную составляющую» этих представлений.

Итак, диагностические признаки фрактального объекта:

- **самоподобие** – возможность воспроизвести объект путем увеличения какой-либо его части;
- **степенная зависимость** числа структурных элементов от масштаба, поскольку математическим выражением самоподобия являются степенные законы –  $f(x) = cx^a$  ;
- **масштабная инвариантность** (скейлинг; англ. – *scaling*) – возможность воспроизводить объект при изменении масштабов;
- **строгое отличие** фрактальной размерности ( $\alpha$ ), которая может быть как целочисленной, так и дробной, от топологической.

Степенные законы, описывающие зависимость видового богатства ( $S$ ) от «выборочного усилия», выраженного через площадь обследованной территории ( $A$ ; species-area relationship, SAR) или объем выборки ( $N$ ) были предметом многочисленных исследований (назову лишь несколько имен ботаников: Hewett Cottrell Watson, 1804–1881; Paul Jaccard, 1868–1944; Alvar Palmgren, 1880–1960; Olof Wilhelm Arrhenius, 1896–1977; Henry Allan Gleason, 1882–1975 и др.) и нашли свое логическое завершение в рамках *равновесной теории островной биогеографии* (MacArthur, Wilson, 1967).

Ю. Одум (1975, с. 196–197), обсуждая проблему определения видового богатства, указывает: «Графический анализ имеет два преимущества перед показателями: 1) сглаживается влияние различий в величине проб; 2) не делается никаких специальных предположений о характере математической зависимости. Однако, подобрав уравнения для таких кривых, можно помочь выяснить, какими математическими «законами» определяется зависимость между  $S$  и  $N$ ».

Р. Маргалеф (1992, с. 143), отстаивая использование показателей-индексов, писал: «...зависимость между  $S$  и  $N$  можно выразить следующим образом:

$S = N^0$	$S = N^k$	$S = N^1$
хемостат	обычная экосистема	музейная экспозиция

Степень  $k$  – прекрасный индекс разнообразия. Он находится в пределах между 0 и 1. Он не характеризует детали, но может выражать *фрактальную самоорганизацию внутри системы*» (выделено мной. – Г.Р.).

<sup>4</sup> Только от перечисления математических терминов (размерность Хаусдорфа-Безиковича, обобщенные размерности Реньи, преобразование Лежандра, индекс сингулярности и пр.) дух захватывает...

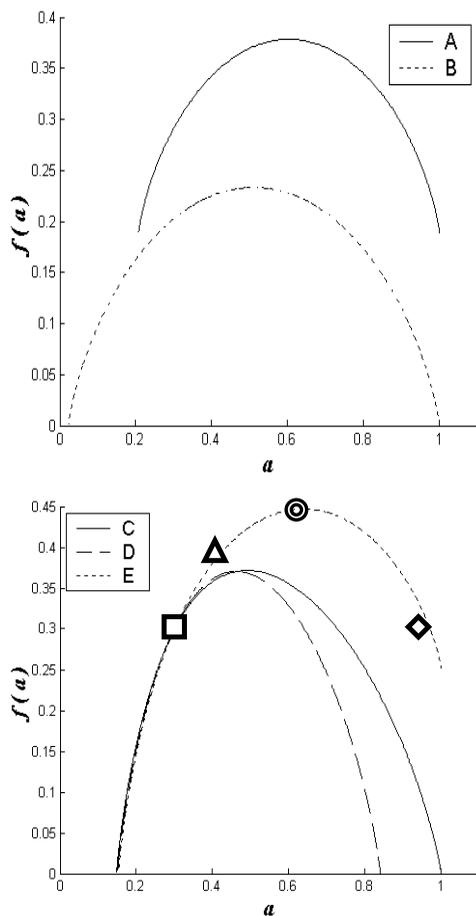
Степенная зависимость в ортогональных координатах натуральных значений  $S$  и  $N$  легко преобразуется в линейную в билогарифмических координатах  $\ln S = k \cdot \ln N$ , где  $k = \ln S / \ln N$ ,  $0 \leq k \leq 1$  имеет смысл *индекса видового разнообразия Маргалефа*. Это выражение показывает, что видовая структура сообщества (в данном случае видовое богатство) инвариантна относительно преобразования его численности. Величину  $k$  можно трактовать как соответствующую фрактальную размерность: число элементов видовой структуры (число видов сообщества) меняется по степенному закону с показателем  $k$  при увеличении размеров системы, то есть с ростом численности сообщества.

Однако, как и любая фрактальная размерность, индекс  $k = \ln S / \ln N$  не дает исчерпывающего количественного представления о видовой структуре биотического сообщества: формула Маргалефа констатирует лишь сам факт наличия вида в выборке, но не содержит, например, сведений о распределении видов по численности или о степени их доминирования. Наиболее общее описание внутреннего устройства самоподобных объектов позволяет дать *теория мультифракталов*, характеризуемых бесконечной иерархией размерностей (Федер, 1991; что-то вроде соотношения *средней и средневзвешенной*). **Именно мультифрактальный спектр дает нам возможность представить обобщенный геометрический образ видовой структуры сообщества.** Для каждого сообщества можно рассчитать некоторые параметры мультифрактальной структуры, которые называются «индекс сингулярности»  $a$  и неотрицательная функция  $f(a)$  – «спектр сингулярностей». В работе В.Н. Якимова (2007) построены мультифрактальные спектры для пяти гипотетических сообществ (А – Е; см. рис. 3).

Дадим биологическую интерпретацию различным частям мультифрактального спектра. При данном подходе сообщество рассматривается как множество, состоящее из отдельных фрактальных подмножеств, которые можно интерпретировать как **совокупности особей, относящихся к видам со сходной представленностью**. Так, например, можно выделить подмножество доминирующих видов, подмножество субдоминантов и т.д. Для таких подмножеств можно вычислить фрактальную размерность, которая и будет характеризовать **видовое разнообразие**. Именно такой смысл имеет **ордината** точек на графике мультифрактального спектра. **Абсцисса** характеризует представленность (долю) видов того или иного подмножества. Ее также можно интерпретировать как **скорость убывания доли вида** (типичного представителя данного подмножества) при росте объема выборки.

Наличие на графике спектра точек, лежащих по оси «а» близко к **нулю**, означает присутствие в сообществе **сильных доминантов** (чем меньше абсцисса точек, тем сильнее доминирование); наличие же точек, лежащих близко к **единице**, означает присутствие в сообществе **редких видов**. Таким образом, **ширину спектра** можно интерпретировать как

**показатель выравненности видов** в сообществе: **чем шире спектр, тем меньше выравненность, и наоборот**. Крайним вариантом выравненности является равнопредставленность видов; при этом весь спектр «схлопывается» в единственную точку.



**Рис. 3. Мультифрактальные спектры пяти гипотетических сообществ:**  
 A = (100 ; 100; 100; 10; 10; 10; 1; 1; 1), B = (100; 10; 1) C = (40; 15; 10; 10; 1),  
 D = (40; 15; 10; 10; 2), E = (40; 15; 10; 10; 1; 1; 1)

- ⊙ – индекс Маргалефа;    Δ – коэффициент Шеннона – Уивера;  
 □ – индекс Симпсона;    ◇ – представленность редких видов

Очень интересной особенностью мультифрактального спектра является тот факт, что практически все известные индексы разнообразия «находят» на нем место. Иными словами, когда мы определяем какой-либо из индексов, мы получаем только точку на графике, которая не позволяет дать интерпретацию структуры сообщества в целом. Прежде всего **именно в этом и есть преимущество фрактального подхода к описанию структуры биологического сообщества.**

Еще одной важной особенностью мультифрактального подхода к анализу структуры сообществ является возможность «единообразного» представления целого спектра индексов разнообразия. Введем моменты распределения особей по видам и выясним характер их асимптотического поведения при увеличении численности  $N$ :

$$M_q = \sum_{i=1}^S p_i^q \sim N^{\tau(q)},$$

где  $-\infty \leq q \leq +\infty$  – называется *порядком момента*, а показатель  $\tau(q)$  характеризует скорость изменения соответствующего момента при увеличении размера выборки. Тогда обобщенная размерность Реньи ( $D_q$ ) имеет вид:

$$D_q = \lim_{N \rightarrow \infty} \left\{ \frac{1}{1-q} \frac{\ln M_q}{\ln N} \right\} = \frac{\tau(q)}{1-q}.$$

На практике величины  $D_q$  можно оценить, используя несколько различающихся значений  $N$ , по более простой формуле:

	$D_q = \frac{1}{1-q} \frac{\ln M_q}{\ln N}$	Обобщенная размерность Реньи	Alfréd Rényi (1921–1970)
$q = 0$	$D_0 = \frac{\ln S}{\ln N} = k$	$k$ – индекс Маргалефа	Ramon Margalef (1919–2004)
$q = 0,5$	$D_{1/2} = \frac{\ln \mu}{\ln N}$ или $\mu = N^{D_{1/2}}$	$\mu$ – индекс Животовского	Лев Анатольевич Животовский (г.р. 1942)
$q = 1$	$D_1 = \frac{-\sum_{i=1}^S p_i \ln p_i}{\ln N} = \frac{H}{\ln N}$	$H$ – индекс Шеннона – Уивера	Claude Elwood Shannon (1916–2001) Warren Weaver (1894–1978)
$q = 2$	$D_2 = -\frac{\ln \sum_{i=1}^S p_i^2}{\ln N} = -\overline{\ln C}$	$C$ – индекс Симпсона – Джини	Edward Hugh Simpson (г.р. 1927) Corrado Gini (1884–1965)

Теперь приведу несколько геоботанических примеров.

Дж. Грин с соавторами (Green et al., 2003) исследовал участок растительности в калифорнийском заповеднике на серпентинной почве. Это бедные по составу почвы (в анализ было включено более 37 тыс. отдельных растений 24 видов), поэтому там практически отсутствуют корневищные травы, что позволяет четко идентифицировать отдельные особи. Самыми обильными были *Vulpia microstachys* (10792 особи), *Bromus madritensis* (6990), *Allium serra* (5989). Участок 8x8 м был разбит на 256 квадратов (16x16, по 0,25м<sup>2</sup>), в каждом из которых была определена численность каждого вида. Такая схема сбора материала максимально соответствует возможности верификации фрактальных гипотез. Анализ показал, что в исследованном сообществе в конкретном диапазоне масштабов (0,5–8 м) видовое богатство распределено фрактально (SAR имеет классический степенной вид), большая же часть видов фрактального распределения не имеет.

Б. Босют и М. Херми (Bossuyt, Hermy, 2004) представили данные об организации растительности прибрежных дюнных сообществ побережья Бельгии и северной Франции. Схема сбора материала в значительной степени аналогична схеме Дж. Грина с соавторами. Было заложено 14 участков размером 2x2 м, каждый из которых разбит на 64 квадрата (8x8 по 0,0625 м<sup>2</sup>). Общий результат также аналогичен: распределение отдельных видов отличается от фрактального, при этом сообщество в целом фрактальной структурой обладает.

Еще одно исследование (Laurie, Perrier, 2006) было направлено на изучение мультифрактальной структуры зависимости видового богатства видов семейства Proteaceae (протейные) во флористической области мыса Доброй Надежды (Южная Африка; эта область характеризуется самым высоким процентом эндемиков) от площади (SAR-анализ).

Группа ученых из Канады и Китая (Zhang et al., 2006) также провели исследование пространственной структуры лесных сообществ горы Донглишан (западный Тянь-Шань). Сбор материала проводился по методике, специально разработанной для последующего осуществления мультифрактального анализа: на 10 западных склонах были заложены трансекты шириной 10 м от основания к вершине, трансекты были разбиты на квадраты 10x10 м, в каждом квадрате проведен количественный учет растений древесного, кустарникового яруса, а также травостоя. В результате анализа авторы построили спектры обобщенных размерностей для трех рассмотренных ярусов. Никакой акцентированной интерпретации этих спектров авторы не дают, оставляя эту задачу как одно из направлений будущих исследований. Среди особенностей этой работы отметим также, что обобщенные размерности здесь определялись как наклон графиков

информации Реньи ( $I_q$ ) от площади. Такая модификация общей схемы анализа не имеет принципиального значения.

## 7. Растительные сообщества. II. Классификация и ординация

### *Выделение однородных групп (автоматическая классификация)*

«Классификация состоит в объединении насаждений в классы, члены каждого из которых имеют некоторое количество общих характеристик, благодаря чему они оказываются отграниченными от членов других классов. Классификация растительности – не новая область исследований, хотя до последнего времени она базировалась на данных, которые были в значительной степени субъективными» (Грейг-Смит, 1967, с. 233). «После того как тем или иным методом определено сходство каждой пары площадок, можно приступать к выполнению следующего этапа работы: выделению групп сходных описаний» (Василевич, 1969, с. 159).

Последний (из известных мне) обзор количественных методов классификации (Миркин, Розенберг, 1979) содержал чуть более 20 алгоритмов. За прошедшие 25 лет в этом направлении наметился заметный прогресс, что связано с «переходом на идеологию» обработки баз фитоценологических данных, таких как «Turbo (Veg)» (Hennekens, 1996); в 60–80-х гг. под эгидой Р. Уиттекера стало развиваться математическое обеспечение количественных ординационно-классификационных методов анализа растительности. Прежде всего это пакеты программ М. Хилла – TWINSPAN (**T**wo-way **I**ndicator **S**pecies **A**nalysis; Hill, 1979) и DECORANA (**D**etrended **C**orrespondence **A**nalysis); Hugh G. Gauch, Jr. (ORDIFLEX<sup>5</sup>, COMCLUS<sup>6</sup>; Gauch, 1977, 1982) и др.; имеются и отечественные разработки – «Фитоценолог» (Голуб и др., 1995) и др.

### *Методы анализа растительных континуумов (ординация)*

«Ординация связана с концепцией растительности как континуума... При изучении ординации насаждений каждое из них стараются расположить в некоторой системе координат (образованной одной или большим числом осей) таким образом, чтобы иметь возможность получить максимальную информацию о его составе» (Грейг-Смит, 1967, с. 233). «За последние два десятилетия (во второй половине XX века – Г.Р.) в геоботани-

---

<sup>5</sup> RDIFLEX – a flexible computer program for four ordination techniques: weighted averages, polar ordination, principal component analysis, and reciprocal averaging.

<sup>6</sup> OMPCLUS – «composite clustering».

ке получил широкое распространение взгляд на растительность как не непрерывно варьирующую систему – континуум» (Василевич, 1969, с. 188).

Здесь я не буду вдаваться в теоретические основы становления континуальной парадигмы и отошлю читателя к монографии Б.М. Миркина и Л.Г. Наумовой (1998).

Практически весь список методов ординации – прямой градиентный анализ (о котором говорилось выше), композиционная (висконсинская) ординация, факторный анализ и пр., которые достаточно широко используются в фитоценологии. Для иллюстрации приведем пример композиционной ординации видов аласов Центральной Якутии по фактору увлажнения (Гоголева и др., 1987; см. табл. 6).

Среди многомерных методов ординации стабильным успехом, начиная со ставшей классической работы Д. Гудола (Goodall, 1954), пользуются методы факторного анализа, и в первую очередь метод главных компонент; работ такого плана – огромное количество.

**Таблица 6. Обработка результатов композиционной ординации по фактору увлажнения (отобрано по два вида в каждую из экологических групп)**

Виды	Средневзвешенное (балльное) значение фактора	Сила влияния ( $\eta^2$ )	Экологические группы
<i>Psathyrostachys juncea</i>	1,29	0,92	Ксерофиты
<i>Stipa krylovii</i>	2,97	0,87	
<i>Galium verus</i>	2,85	0,42	Мезоксерофиты
<i>Festuca lenensis</i>	4,30	0,65	
<i>Koeleria cristata</i>	5,58	0,13	Ксоромезофиты
<i>Agrostis trinitii</i>	6,61	0,57	
<i>Puccinellia tenuiflora</i>	5,90	0,68	Мезофиты
<i>Saussurea amara</i>	7,46	0,34	
<i>Alopecurus arundinaceus</i>	7,58	0,29	Гигромезофиты
<i>Caltha palustris</i>	8,89	0,12	
<i>Scirpus lacustris</i>	8,95	0,12	Мезогигрофиты
<i>Glyceria triflora</i>	9,71	0,30	
<i>Stellaria palustris</i>	9,93	0,47	Гигрофиты
<i>Phragmites australis</i>	10,30	0,17	

## 8. Количественный подход в геоботанике

«На современном этапе развития геоботаники, когда все более очевидной становится ценность количественного подхода, наблюдается тенденция либо применять количественные методы с чрезмерным увлечением, либо применять их к проблемам, к которым они не применимы вообще или по крайней мере в настоящее время... Важно сохранять чувство перспективы и учитывать, приведет ли трата времени на использование

количественных методов вместо качественных к соответствующему увеличению ценности получаемых результатов» (Грейг-Смит, 1967, с. 302).

«Есть две другие опасности, с которыми сталкиваются геоботаники, использующие количественные методы: 1) настолько большое увлечение методикой, что цели исследования оттесняются на задний план, и 2) стремление втиснуть геоботанические данные в формальные математические рамки» (Грейг-Смит, 1967, с. 303).

«Количественные наблюдения следует проводить лишь при наличии определенной цели, хотя этой целью может быть только ответ на очень простой и предварительный вопрос, например: «есть ли разница в численности особей определенного вида между участками А и В? Тенденция отношения к сбору численных данных как к самоцели не заслуживает особого одобрения» (Грейг-Смит, 1967, с. 304).

«Изучение литературы показывает, что отбор образцов – наименее хорошо понимаемый аспект количественной геоботаники. Не так редко наблюдения остроумно планируются и данные тщательно анализируются, однако результаты обесцениваются неудовлетворительным способом отбора образцов... Такие ошибки, возможно, возникают в связи с тем, что при написании большинства руководств по биологической статистике главным образом имелась в виду экспериментальная работа, а при такой работе проблема отбора образцов гораздо проще; поэтому этот предмет излагается довольно поверхностно» (Грейг-Смит, 1967, с. 304–305).

Я готов подписаться под всеми этими положениями, не потерявшими своей актуальности. Для меня системный подход и развитие количественных методов в экологии и фитоценологии соединены в очень простой слоган<sup>7</sup>: «Думать надо!» И здесь, как нельзя, кстати, подходят слова В.И. Василевича (1969, с. 218), которыми он заканчивает свою монографию: «Не следует думать, что тот математический аппарат, который в настоящее время используют геоботаники, дает наилучшее решение. С одной стороны, его необходимо углублять и дополнять, совершенствовать уже существующие методы, а с другой стороны, необходимо искать принципиально новые пути решения проблем. Авторитет старых методов так же опасен, как и авторитет старых теорий».

## **9. Динамика растительности (экологическое прогнозирование)**

Этого раздела, к сожалению, ни в книге П. Грейг-Смита, ни в книге В.И. Василевича нет.

---

<sup>7</sup> Понятие «слоган» (*sluagh-ghairm*) восходит к галльскому языку и означает «боевой клич». В 1880 г. понятие «слоган» было впервые использовано в современном значении.

Среди традиционных методов анализа динамики растительности чаще всего используется все тот же прямой градиентный анализ – анализ хроноклинов. В качестве примера использования градиентного анализа хроноклинов приведем выборочные данные из работы К.Е. Кононова и А.И. Неустроевой (1976) по разногодичной изменчивости (1969–1974 гг.) максимальной высоты растений луговых сообществ поймы Средней Лены (табл. 7). «Наблюдения в течение ряда лет над высотами растений на различных типах лугов поймы р. Лены показывают, что высота луговых трав весьма сильно подвержена разногодичной изменчивости. При этом, как и следовало ожидать, большие колебания высот характерны растениям остепненного луга, в меньшей степени настоящего и еще меньшей степени – болотистого» (Кононов, Неустроева, 1976, с. 119–121).

Таблица 7. Сила влияния разногодичной изменчивости на максимальную высоту растений луговых сообществ поймы Средней Лены

Виды	Сила влияния ( $\eta^2$ ), достоверность $\Phi_{0,95} > 5,8$	Мода распределения в см (в скобках – модальный год)
<b>Остепненный луг</b>		
<i>Pulsatilla flavescens</i>	0,74	35,8 (1969)
<i>Plantago canescens</i>	0,65	30,9 (1971)
<i>Myosotis suaveolens</i>	0,84	28,0 (1971)
<i>Potentilla stipularis</i>	0,73	23,3 (1971)
<i>Oxytropis strobilacea</i>	0,74	28,3 (1971)
<i>Onobrychis sibirica</i>	0,85	80,9 (1971)
<i>Kobresia filifolia</i>	0,75	31,9 (1971)
<i>Bromus ircutensis</i>	0,69	70,0 (1974)
<i>Campanula glomerata</i>	0,86	46,1 (1971)
<i>Pedicularis venusta</i>	0,91	40,7 (1969)
<i>Poa subfastigiata</i>	0,68	57,3 (1971)
<b>Настоящий луг</b>		
<i>Sanguisorba officinalis</i>	0,74	69,0 (1974)
<i>Gentiana barbata</i>	0,68	25,8 (1970)
<i>Vicia cracca</i>	0,66	36,4 (1971)
<i>Hordeum brevisubulatum</i>	0,83	94,6 (1969)
<b>Болотистый луг</b>		
<i>Caltha palustris</i>	0,81	35,0 (1969)
<i>Parnassia cartilaginea</i>	0,75	57,8 (1969)
<i>Poa palustris</i>	0,95	47,6 (1973)

Сравнительно новый и интересный метод экологического прогнозирования, основанный на построении самоорганизующихся регрессионных моделей, обсуждался выше.

## Заключение

Приходится констатировать, что за последние 50 лет развития статистических методов в фитоценологии успехов значительно меньше, чем за весь вековой период становления этих методов. Наверное, это вполне объяснимо. Некоторые методы «устоялись» и стали обязательными (прямой градиентный анализ), другие остаются достаточно экзотическими и используются в основном авторами (распознавание образов и самоорганизация при прогнозировании). Развитие количественных методов пошло в направлении разработки математических моделей и создания разного рода баз знаний и данных. Этот процесс не простой и ожидать здесь быстрых успехов не приходится.

И еще один момент, на котором хотелось бы заострить внимание. Новые успехи фитоценологии следует ожидать **не в направлении разработки каких-то новых методов анализа растительности, а в выдвигании новых представлений о структуре и характере динамики растительных сообществ**; иными словами, должна произойти очередная смена парадигм.

## ЛИТЕРАТУРА

- Адлер Ю.П.* Введение в планирование эксперимента. М., 1969. 158 с.
- Азовский А.И., Чертопруд М.В.* Анализ пространственной организации сообществ и фрактальная структура литорального бентоса // Докл. АН. 1997. Т. 356, № 5. С. 713–715.
- Беклемишев В.Н.* Об общих принципах организации жизни // Бюл. МОИП. Отд. биол. 1964. Т. 69. Вып. 2. С. 22–38.
- Биоиндикация: теория, методы, приложения* / Под ред. Г.С. Розенберга. Тольятти, 1994. 266 с.
- Боголюбов А.Г.* К столетию начала биометрических работ в России // Изв. СамНЦ РАН. 2002. Т. 4, № 2. С. 189–196.
- Большев Л.Н., Смирнов Н.В.* Таблицы математической статистики. М., 1983. 416 с.
- Букатова И.Л.* Эволюционное моделирование и его приложения. М., 1979. 232 с.
- Букатова И.Л., Михасев Ю.И., Шаров А.М.* Экоинформатика. Теория и практика эволюционного моделирования. М., 1991. 206 с.
- Бурцева Е.И.* Опыт количественного анализа растительности солончаковых лугов в пойме Средней Лены // Экология и ценология лугов Центральной Якутии. Якутск, 1978. С. 128–164.
- Бурцева Е.И.* Методические подходы к оценке кормовых угодий Центральной Якутии. 1. Основные положения // Проблемы северного земледелия: селекция, кормопроизводство, экология. Новосибирск, 2000. С. 128–135.

*Бурцева Е.И., Десяткина Л.И.* Методические подходы к оценке кормовых угодий Центральной Якутии. 2. Оценочные шкалы // Проблемы северного земледелия: селекция, кормопроизводство, экология. Новосибирск, 2000. С. 135–141.

*Бурцева Е.И., Кононов К.Е., Розенберг Г.С.* Статистический анализ растительности солончаковых лугов поймы реки Лены. II. Геоботаническая индикация // *Иол. науки*. 1978. № 6. С. 132–139.

*Василевич В.И.* Статистические методы в геоботанике. Л., 1969. 232 с.

*Василевич В.И.* Количественные методы изучения структуры растительности // *Итоги науки и техники. Ботаника*. М., 1972. Т. 1. С. 7–83.

*Галанин А.В.* Эколого–ценотическая информативность видов и измерение флористического сходства растительных сообществ // Флористические критерии при классификации растительности. Уфа, 1981. С. 23–24.

*Галанин А.В.* Сравнение флористического состава коренной и нарушенной растительности в подзоне северной тайги Коми АССР // *Ботан. журн.* 1983. Т. 68, № 2. С. 221–227.

*Гелаишвили Д.Б., Иудин Д.И., Розенберг Г.С. и др.* Степенной закон и принцип самоподобия в описании видовой структуры сообществ // *Поволжский экол. журн.* 2004. № 3. С. 227–245.

*Гоголева П.А., Кононов К.Е., Миркин Б.М., Миронова С.И.* Синтаксономия и симфосоциология растительности аласов Центральной Якутии. Иркутск, 1987. 176 с.

*Голуб В.Б., Халеев Е.А., Рухленко И.А.* Пакет программ для обработки геоботанических данных «Фитоценолог» // II совещание «Компьютерные базы данных в ботанических исследованиях»: Тез. докл. СПб, 1995. С. 13–14.

*Грейг-Смит П.* Количественная экология растений. М., 1967. 359 с.

*Ефимов В.М.* О некоторых коэффициентах межвидовой сопряженности // *Изв. СО АН СССР, сер. биол.* 1976. № 10. Вып. 2. С. 131–136.

*Заде Л.* Основы нового подхода к анализу сложных систем и процессов принятия решений // *Математика сегодня*. М., 1974. С. 5–49.

*Заде Л.А.* Понятие лингвистической переменной и его применение к принятию приближенных решений. М., 1976. 165 с.

*Ивахненко А.Г.* Долгосрочное прогнозирование и управление сложными системами. Киев, 1975. 311 с.

*Ивахненко А.Г.* Индуктивный метод самоорганизации моделей сложных систем. Киев, 1982. 296 с.

*Ивахненко А.Г., Зайченко Ю.П., Димитров В.Д.* Принятие решений на основе самоорганизации. М., 1976. 280 с.

*Иудин Д.И., Гелаишвили Д.Б., Розенберг Г.С.* Мультифрактальный анализ структуры биотических сообществ // *Докл. АН*. 2003. Т. 389, № 2. С. 279–282.

*Козлов М.В.* Мнимые повторности (pseudoreplication) в экологических исследованиях: проблема, не замеченная российскими учеными // *Журн. общ. биол.* 2003. Т. 64, № 4. С. 292–307.

*Козлов М.В., Хёлберт С.Х.* Мнимые повторности, бесплодные дискуссии и интернациональная сущность науки: ответ Д.В. Татарникову // *Журн. общ. биол.* 2006. Т. 67, № 2. С. 145–152.

*Колмогоров А.Н.* Теория информации и теория алгоритмов. М., 1987. 304 с.

Кононов К.Е. Опыт использования композиционной ординации для изучения отношения видов лугов поймы Средней Лены к увлажнению // Экология. 1981. № 1. С. 83–86.

Кононов К.Е., Софронеева Т.П. Опыт построения классификации пойменных лугов и степей реки Лены с использованием «блок-метода» // Количественные методы анализа растительности. Уфа, 1974. С. 139–141.

Кононов К.Е., Розенберг Г.С. Количественный анализ мозаичности пойменных лугов // Экология и ценология лугов Центральной Якутии. Якутск, 1978. С. 70–174.

Кононов К.Е., Розенберг Г.С. Прогнозирование урожайности аласных сенокосов по гидрометеорологическим факторам методами самоорганизации // Биол. науки. 1981. № 3. С. 99–104.

Кононов К.Е., Наумова Л.Г., Розенберг Г.С., Софронеева Т.П. Опыт использования кластерного анализа для выявления экологических групп видов травянистой растительности поймы р. Лены на территории Олекминского района // Теоретические и методические вопросы изучения лугов Центральной Якутии. Якутск, 1976. С. 72–98.

Корн Г., Корн Т. Справочник по математике для научных работников и инженеров. Определения, теоремы, формулы. М., 1968. 720 с.

Криволицкий Д.А., Степанов А.М., Тихомиров Ф.А., Федоров Е.А. Экологическое нормирование на примере радиоактивного и химического загрязнения экосистем // Методы биоиндикации окружающей среды в районах АЭС. М., 1988. С. 4–16.

Левич А.П., Терехин А.Т. Метод расчета экологически допустимых уровней воздействия на пресноводные экосистемы (метод ЭДУ) // Водн. ресурсы. 1997. Т. 4, № 3. С. 328–335.

Маргалеф Р. Облик биосферы. М., 1992. 254 с.

Миркин Б.М. Статистический анализ экоклиндов // Ботан. журн. 1971. Т. 56, № 12. С. 1772–1788.

Миркин Б.М., Наумова Л.Г. Методические вопросы применения корреляционного метода при геоботанических исследованиях // Ботан. журн. 1974. Т. 59, № 8. С. 1142–1156.

Миркин Б.М., Наумова Л.Г. Наука о растительности (история и современное состояние основных концепций). Уфа, 1998. 413 с.

Миркин Б.М., Розенберг Г.С. Анализ мозаичности травянистых растительных сообществ 1. Популяционный уровень // Биол. науки. 1976. № 12. С. 127–134.

Миркин Б.М., Розенберг Г.С. Анализ мозаичности травянистых растительных сообществ. 2. Ценотический уровень // Биол. науки. 1977. № 2. С. 121–126.

Миркин Б.М., Розенберг Г.С. Фитоценология. Принципы и методы. М., 1978. 212 с.

Миркин Б.М., Розенберг Г.С. Количественные методы классификации, ординации и геоботанической индикации // Итоги науки и техники. Ботаника. М., 1979. Т. 3. С. 71–137.

Миркин Б.М., Розенберг Г.С., Наумова Л.Г. Новый коэффициент межвидовой сопряженности, удобный для классификации пойменных лугов (трансформиро-

ванный коэффициент Дайса, ТКД) // Растительность речных пойм, методы ее изучения и вопросы рационального использования. Уфа, 1972. С. 85–86.

*Миркин Б.М., Розенберг Г.С., Наумова Л.Г.* Словарь понятий и терминов современной фитоценологии. М., 1989. 223 с.

*Миркин Б.М., Розенберг Г.С., Обыденнова Е.А., Янтурин С.И.* Опыт анализа размещения видов в луговых ценозах // Структура и динамика растительного покрова. М., 1976. С. 89–90.

*Миркин Б.М., Янтурин С.И.* Статистический анализ влияния удобрений и погодных условий на горизонтальную структуру луга. II. Популяционный уровень // иол. науки. 1981. № 6. С. 56–68.

*Миронова С.И.* Техногенные сукцессионные системы растительности Якутии (на примере Западной и Южной Якутии). Новосибирск, 2000. 151 с.

*Морозов В.Г.* Эволюционное моделирование рядов произвольной вариабильности: необходимость и методика прогнозирования // Изв. СамНЦ РАН. 2000. Т. , № 2. С. 206–215.

*Мостеллер Ф., Тьюки Дж.* Анализ данных и регрессия. М., 1982. 320 с.

*Мэгарран Э.* Экологическое разнообразие и его измерение. М., 1992. 181 с.

*Налимов В.В., Чернова Н.А.* Статистические методы планирования экстремальных экспериментов. М., 1965. 340 с.

*Одум Ю.* Основы экологии. М., 1975. 740 с.

*Орлов А.И.* Прикладная статистика. М., 2004. 671 с.

*Пайген Х.О., Рихтер П.Х.* Красота фракталов. М., 1993. 248 с.

*Плохинский Н.А.* Математическое оснащение биологов // Методы современной биометрии. М., 1978. С. 27–35.

*Пузаченко Ю.Г.* Приложение теории фракталов к изучению структуры ландшафта // Изв. РАН. Сер. геогр. 1997. № 2. С. 24–40.

*Пузаченко Ю.Г., Пузаченко А.Ю.* Семантические аспекты биоразнообразия // Журн. общ. биол. 1996. Т. 57, № 1. С. 5–43.

*Реймерс Н.Ф.* Природопользование. Словарь–справочник. М., 1990. 637 с.

*Розенберг Г.С.* Об оценке точности планируемого числа наблюдений // Биол. науки. 1976. № 3. С. 125–129.

*Розенберг Г.С.* Модели в фитоценологии. М., 1984. 256 с.

*Розенберг Г.С.* Геоботаническая индикация с использованием распознавания образов и регрессионных моделей // Биоиндикация: теории, методы, приложения. Тольятти, 1994. 130–137 с.

*Розенберг Г.С.* Количественные методы фитоиндикации // Экологический мониторинг. Методы биологического и физико–химического мониторинга. Часть III. Н. Новгород, 1998. С. 5–27.

*Розенберг Г.С., Долотовский И.М.* Еще раз о показателях силы влияния // Биол. науки. 1988. № 9. С. 105–110.

*Розенберг Г.С., Наумова Л.Г., Бурцева Е.И.* О математической корректности индекса ТКД // Экология и ценология лугов Центральной Якутии. Якутск, 1978. С. 165–169.

*Розенберг Г.С., Шитиков В.К., Мозговой Д.П.* Экологическая информатика: Учебн. пособие. Самара, 1993. 151 с.

*Розенберг Г.С., Шитиков В.К., Брусиловский П.М.* Экологическое прогнозирование (Функциональные предикторы временных рядов). Тольятти, 1994. 182 с.

*Скрябин К.И., Бурцева Е.И.* Опыт классификации растительности аласов с использованием подходов Браун–Бланке и оценки естественности единиц коэффициентом сходства // Теоретические и методические вопросы изучения лугов Центральной Якутии. Якутск, 1976. С. 59–71.

*Слуцкий Е.Е.* О сложении случайных величин как причине циклических процессов // Вопросы конъюнктуры. 1927. Т. 3, № 1. (см.: Слуцкий Е.Е. Избр. труды. Теория вероятности и математическая статистика. М.; Л., 1960. С. 99–132).

*Уиттекер Р.* Сообщества и экосистемы. М., 1980. 328 с.

*Усманов И.Ю.* К методике планирования числа наблюдений при определении морфологических структур растений // Сельхоз. биол. 1984. № 1. С. 113–116.

*Федер Е.* Фракталы. М., 1991. 254 с.

*Фёдоров В.В.* Теория оптимального эксперимента (планирование регрессионных экспериментов). М., 1971. 312 с.

*Фогель Л., Оуэнс А., Уоли М.* Искусственный интеллект и эволюционное моделирование. М., 1969. 230 с.

*Шитиков В.К., Розенберг Г.С., Зинченко Т.Д.* Количественная гидроэкология: методы, критерии, решения: в 2-х кн. М., 2005. Кн. 1. 281 с.; кн. 2. 337 с.

*Якимов В.Н.* Фрактальность видовой и пространственной структуры биологических сообществ: разработка концепции и верификация: Автореф. дис. ... канд. биол. наук. Н. Новгород, 2007. 18 с.

*Ashby E.* Quantitative analysis of vegetation // Ann. Bot. 1935. Vol. 49. P. 779–802.

*Ashby E.* Statistical ecology // Bot. Rev. 1936. Vol. 2. P. 221–235.

*Ashby E.* Statistical ecology. II. A reassessment // Bot. Rev. 1948. Vol. 14. P. 222–234.

*Bossuyt B., Hermy M.* Species turnover at small scales in dune slack plant communities // Basic and Applied Ecology. 2004. Vol. 5. P. 321–329.

*Gauch H.G.* ORDIFLEX a flexible computer program for four ordination techniques: weighted averages, polar ordination, principal component analysis and reciprocal averaging: Release B. Ithaca, 1977. 195 p.

*Gauch H.G.* Multivariate Analysis in Community Ecology. Cambridge, 1982. 314 p.

*Goodall D.W.* Objective methods for the classification of vegetation. III. An essay in the use of factor analysis // Austral. J. of Bot. 1954. Vol. 2. P. 304–324.

*Green J.L., Harte J., Ostling A.* Species richness, endemism and abundance patterns: tests of two fractal models in a serpentine grassland // Ecology Letters. 2003. Vol. 6. P. 919–928.

*Greig-Smith P.* Quantitative Plant Ecology. New York; London, 1957. 198 p.

*Greig-Smith P.* Quantitative Plant Ecology / 2nd ed. London, 1964. 256 p.

*Greig-Smith P.* Quantitative Plant Ecology / 3rd ed. Berkeley, 1983. 359 p.

*Harte J., Kinzig A.P., Green J.* Self-similarity in the distribution and abundance of species // Science. 1999. Vol. 284. P. 334–336.

*Hennekens S.M.* TURBO (VEG). Software package for input, processing, and presentation of phytosociological data. Version 1.0. IBNDLO. Lancaster, 1996. 52 p.

*Hill M.O.* TWINSPAN a FORTRAN program for arranging multivariate data in an ordered two way table by classification of the individuals and the attributes // Ecology and Systematic. Ithaca: Cornell Univ. Press, 1979. 48 p.

*Hurlbert S.H.* Pseudoreplication and the design of ecological field experiments // Ecol. Monographs. 1984. Vol. 54. P. 187–211.

*Iudin D.I., Gelashvily D.B.* Multifractality in ecological monitoring // Nucl. Instr. Meth. Phys. Res. 2003. Vol. 502. P. 799–801.

*Laurie H., Perrier E.* A multi-fractal model for the species–area relationship. 2006. [www.mth.uct.ac.za/~henri/multfrac6par.pdf](http://www.mth.uct.ac.za/~henri/multfrac6par.pdf)

*MacArthur R.H., Wilson E.O.* The Theory of Island Biogeography. Princeton (N.J.), 1967. 203 p.

*Mandelbrot B.* Les objets fractals: forme, hasard et dimension. Paris, 1975.

*Oksanen L.* Logic of experiments in ecology: is pseudoreplication a pseudoissue? Oikos. 2001. Vol. 94. P. 27–38.

*Whittaker R.* Direct gradient analysis: techniques // Handbook of vegetation science. V. 5. Ordination and classification of vegetation. Hague, 1973. P. 7–30.

*Zhang Y., Ma K., Anand M., Fu B.* Do generalized scaling laws exist for species abundance distribution in mountains? // Oikos. 2006. Vol. 115. P. 81–88.