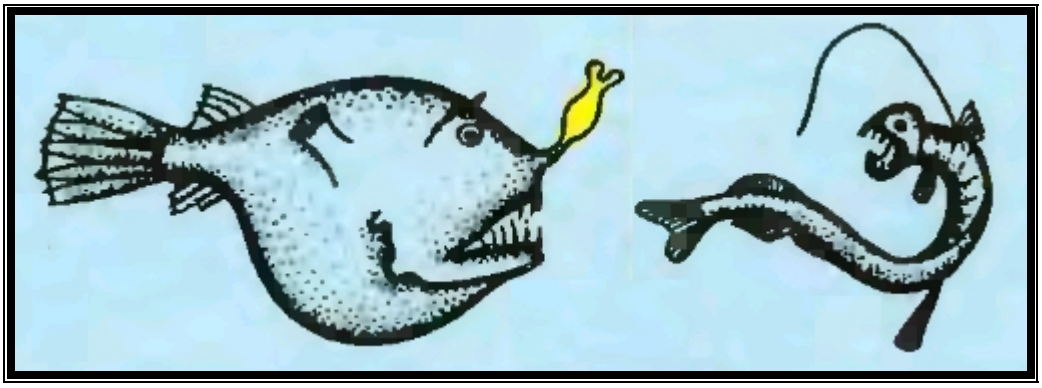
В.К. Шитиков,
Т.Д. Зинченко,
Г.С. Розенберг

МАКРОЭКОЛОГИЯ РЕЧНЫХ СООБЩЕСТВ:

концепции, методы, модели



Российская академия наук

Институт экологии Волжского бассейна

В.К. Шитиков, Т.Д. Зинченко,

Г.С. Розенберг

МАКРОЭКОЛОГИЯ РЕЧНЫХ СООБЩЕСТВ:

концепции, методы, модели

Тольятти 2011

Шитиков В.К., Зинченко Т.Д., Розенберг Г.С. Макроэкология речных сообществ: концепции, методы, модели. – Тольятти: Кассандра, 2011. – 255 с.

ISBN

В книге рассматриваются современные теоретические концепции анализа видовой структуры бентосных сообществ лотических экосистем. Подробно представлены структурные схемы статистической обработки данных гидроэкологического мониторинга с использованием различных методов макроэкологии. Особое внимание уделяется описанию типов и способов применения нуль-моделей и непараметрических методов проверки статистических гипотез на основе процедур рандомизации и ресамплинга (бутстреп, «складной нож», перестановки и т. д.). Методическая часть подробно иллюстрируется расчетами, выполненными на основе единого массива данных гидробиологической съемки донных организмов на малых реках степной и лесостепной зон Среднего Поволжья.

Описаны схемы разложения показателей видового богатства и разнообразия на альфа- и бета-составляющие, а также методы статистически корректной оценки их величин по моделям разрежения в зависимости от величины выборочного усилия. Сделан подробный анализ различных метрик вложенности (nestedness) структурных фрагментов и статистических методов оценки взаимно согласованной встречаемости таксономических групп. Обсуждаются современные подходы к моделированию ранговых распределений обилия видов. Подробно рассматриваются методы ординации и другие процедуры, позволяющие оценить изменчивость видовой структуры гидробионтов по шкале продольного градиента водотока, чтобы проверить статистическую гипотезу о закономерном континууме донных сообщества против альтернативы о стохастической «мозаике пятен». Вводится понятие «экологических последовательностей местообитаний» и обсуждаются проблемы их статистического анализа.

Книга ориентирована на широкий круг специалистов в области теоретической экологии, гидробиологии, биогеографии, рационального использования ресурсов водоемов, а также может быть использована в качестве учебного пособия по статистическим методам для студентов и аспирантов высших учебных заведений, изучающих и специализирующихся по экологии.

Табл. 36, ил. 59. Библиогр. 626 назв.

Рецензенты: д.б.н., профессор Д.Б. Гелашвили (г. Нижний Новгород),
д.б.н., профессор И.А. Евланов (г. Тольятти)

Рекомендовано к печати Ученым советом Института экологии Волжского бассейна РАН (протокол № 6 от 12 мая 2011 г.).

445003, Россия, Самарская обл., г. Тольятти, ул. Комзина, 10
Институт экологии Волжского бассейна РАН
Тел., факс: (8482) 489-504; E-mail: ievbras2005@mail.ru

© ИЭВБ РАН, 2011 г.
© В.К. Шитиков, Т.Д. Зинченко, Г.С. Розенберг, 2011 г.

О Г Л А В Л Е Н И Е

ВВЕДЕНИЕ 5

Глава 1. **КОНЦЕПЦИИ МАКРОЭКОЛОГИИ И АНАЛИЗ ВИДОВОЙ СТРУКТУРЫ СООБЩЕСТВ** 9

1.1. Модели динамики популяций и «колодки» экологического мышления 9
1.2. Закономерности пространственного распределения гидробионтов в речных экосистемах (развитие теоретических концепций) 13
1.3. Донные сообщества лотических систем как объект и материал исследования 24
1.4. Методы макроэкологии и анализ видовой структуры сообществ 31
1.5. Нуль-модели в экологии 38

Глава 2. **МЕТОДЫ ОЦЕНКИ ВИДОВОГО БОГАТСТВА СООБЩЕСТВ** 44

2.1. Взаимосвязь между локальным видовым богатством и региональным пулом видов 44
2.2. Оценка видового богатства: исходные данные, задачи, подходы, решения 52
2.3. Непараметрические методы экстраполяции полного богатства видов 61
2.4. Модели разрежения и сравнение биоразнообразия сообществ с учетом выборочного усилия 67

Глава 3. **СХОДСТВО И РАЗНООБРАЗИЕ ВИДОВОЙ СТРУКТУРЫ СООБЩЕСТВ** 73

3.1. Развитие представлений о видовом разнообразии: от энтропии к фракталам 73
3.2. Оценка сходства видовой структуры сообществ 82
3.3. Иерархия уровней альфа и бета-составляющих видового разнообразия 93

Глава 4. **АНАЛИЗ СТАТИСТИЧЕСКИХ ЗАКОНОМЕРНОСТЕЙ ОРГАНИЗАЦИИ ВИДОВОЙ СТРУКТУРЫ** 100

4.1. Взаимодействия видов и их влияние на таксономическую структуру сообществ 100
4.2. Меры взаимосвязи видов с использованием структурных элементов 102
4.3. Инкорпорация видов и оценка «вложенности» (nestedness) структурных композиций 104
4.4. Статистические методы оценки структурированности матриц и согласованной встречаемости видов 108
4.5. Нуль-модели для проверки статистической значимости мер согласованности видов 112
4.6. Обсуждение результатов на примере сообществ макрозообентоса малых рек 114

Глава 5. **МОДЕЛИ РАНГОВОГО РАСПРЕДЕЛЕНИЯ ОБИЛИЯ ВИДОВ** 124

5.1. Классические модели распределения обилия и оценка роли редких видов 124
5.2. Модель Ципфа – Мандельброта и *H*-распределения видов 132

5.3. Стохастические модели распределения: идеология и идентификация параметров 141
5.4. От распределения ресурсов ниши к концепции «нейтральности» видов в сообществе С. Хаббела 146
5.5. Динамика обилия видов с позиций *Унифицированной нейтральной теории* (UNTB) 150

Глава 6. **МЕТОДЫ ОРДИНАЦИИ ВИДОВ И МЕСТООБИТАНИЙ** 154

6.1. Актуальные проблемы современной фитоценологии 154
6.2. Статистические методы оценки «верности» видов и формирования ценотических групп 162
6.3. Группировка видов и местообитаний в биогеографических описаниях 169
6.4. Многомерные методы ординации 176
6.5. Ординация сообществ с использованием нечетких множеств 190
6.6. Оценка надежности и достоверности результатов классификации 195
6.7. Анализ компонентов экосистемы с использованием теста Мантеля 202

Глава 7. **МЕТОДЫ СРАВНЕНИЯ КОМПОЗИЦИЙ ВИДОВ В ЭКОЛОГИЧЕСКИХ ПОСЛЕДОВАТЕЛЬНОСТЯХ МЕСТООБИТАНИЙ** 205

7.1. Модели распределения видов по экологическим градиентам 205
7.2. Оценка статистической значимости вариации экологических показателей по течению водотоков 212
7.3. Оценка динамики видовой структуры с использованием индексов скорости и метрик сходства 221

ЗАКЛЮЧЕНИЕ 229

СПИСОК ЛИТЕРАТУРНЫХ ИСТОЧНИКОВ 233

ВВЕДЕНИЕ

Все в мире существует только для
того, чтобы закончиться Книгой…
Стефан Малларме (1842-1898 гг.)

Мир экологии, как науки о сообществах живых организмов, полон идей и гипотез, часть которых постепенно эволюционирует до уровня «закона или теоретической конструкции» (Розенберг и др., 1999; Краснощеков, Розенберг, 2002). Иногда в истории науки несколько основных публикаций предопределяют дальнейшее направление развития целой области. Так и в экологии сообществ, после того как Дж. Даймонд (Diamond, 1975) изложил правила сосуществования видов в сообществе, основываясь на распределении популяций птиц островов архипелага Бисмарка, прозвучал резонный вопрос Э. Коннора и Д. Симберлоффа: «А как можно отличить результат конкурентного исключения видов от простого случайного совпадения?» (Connor, Simberloff, 1979).

Оказалось, что уже недостаточно, просмотрев экспедиционный журнал, подсчитать среднее обилие видов или разработать подходящий индекс и, ориентируясь на их результат, объявить о существовании некоторого гипотетического процесса или влияющего фактора. Возникла необходимо показать, что полученные эмпирические величины метрик статистически значимо отличаются от нуль-модельных интервалов, рассчитанных с использованием тех же мониторинговых данных, но при предполагаемом отсутствии изучаемого эффекта. После этого всесторонняя проверка гипотез о достоверности связей в ассоциациях стала одной из ключевых проблем изучения сообществ, а «научных откровений» о механизмах взаимодействия между видами стало существенно меньше.

Начался период широкой экспансии в экологию нуль-моделей, нейтральных моделей и различных приемов рандомизации данных, основанных на методах Монте-Карло. Полученные результаты часто оказывались неоднозначными, что предопределило оживленную 30-летнюю дискуссию в научном сообществе. При этом сама целесообразность использования нуль-моделей никогда не отрицалась, а осуществлялся лишь активный поиск более эффективных алгоритмов, условий и форм их применения. Поскольку отечественная литература крайне мало уделяла внимания этой проблематике, то **первой** целью настоящей монографии является систематизированный и полный обзор использования нуль-моделей и методов рандомизации данных в макроэкологических исследованиях.

В конце XX в. на стыке экологии сообществ, эволюционной экологии и биогеографии сформировалось такое научное направление, как ***макроэкология***[1], изучающая общий мультипараметрический характер процессов, лежащих в основе развития взаимосвязанных видовых ансамблей и других компонент биосферы. Она стремится построить объяснение сложности этих процессов на основе анализа статистических выборок больших множеств популяций видов, распределенных в географическом пространстве и экологическом времени. Типичными теоретическими конструкциями макроэкологии стали модели распределения обилия видов (SAD – species abundances distribution), пространственное распределение видового богатства (SAR – species area relationship), функции зависимостей для размеров тела животных (BSD – body size distribution), соотношения между региональными и локальными пулами видов и широтные градиенты распределения видового богатства, компоненты и факторы, определяющие биоразнообразие и т. д. Если рассматривать иерархию масштабных уровней организации экосистем, то область интересов макроэкологии, вопреки мнениям о её глобальном характере, простирается в пространственной шкале от

[1] Сам термин «макроэкология» был введен в 1989 г. в статье Джеймса Брауна и Брайена Морера, опубликованной в журнале "Science" (Brown, Maurer, 1989).

1000 м^2 (что соответствует протяженности участков малых рек около 200 м) до континентальных размеров.

Работы российских экологов в этом направлении практически ограничиваются учебным пособием Т.А. Акимовой с соавторами (2005). Для его авторов это – очередной повод пространно поговорить о, несомненно, важных, но общеизвестных концепциях экологии: *аксиомах Коммонера*, *принципе Ле-Шателье*, *законе константности живого вещества* и прочих проблемах преимущественно философского или социально-природоохранного характера. И совершенно игнорируется квинтэссенция макроэкологии – ***количественные методы и модели, позволяющие выделить частные или универсальные закономерности сосуществования сообществ, опираясь на большие массивы данных мониторинга.***

Значительное количество традиционно предлагаемых научных гипотез либо вообще не опирается на результаты мониторинга, либо использует их в чисто иллюстративных целях. Ученые трудолюбиво перетасовывают кучи экспериментального материала, чтобы найти «жемчужные зерна» (истины или подделки), – т. е. фрагменты данных, которые подтверждали бы их научную позицию, без сожаления отбрасывая неподходящие факты. Поэтому **вторая** цель нашей монографии – рассказать «*Как это можно делать?*», основываясь на анализе зарубежных литературных источников по макроэкологии двух последних десятилетий. Большинство описываемых задач сопровождаются примерами расчетов и иллюстрациями на основе единого тестового набора данных, содержащего результаты гидробиологической съемки сообществ макрозообентоса в малых реках на территории Средней Волги.

Фундаментальной задачей экологии и гидробиологии является выявление закономерностей распределения гидробионтов и их сообществ по продольному профилю водотоков. Общепринятые понятия пространственной организации текучих водных экосистем были сформулированы и представлены в виде концепций *речного континуума* (Vannote et al., 1980) и *динамики пятен* (Townsend, 1989). После их создания появился целый ряд работ, подтверждающих либо опровергающих описанные закономерности, а также весьма удачные попытки их синтеза (Богатов, 1995).

Нет сомнения, что основную роль в распределении сообществ гидробионтов играют абиотические факторы среды, среди которых ведущее положение занимает скорость течения, которая закономерно изменяется по продольному профилю водотоков и во многом определяет остальные физико-химические характеристики. Сформулировано достаточно вольное определение «*продольной сукцессии*» (Fisher et al., 1983), как не зависящей от времени закономерной последовательности смены сообществ в проточных водах от истоков до устья. Однако под влиянием определенных комбинаций локальных биотопических особенностей и факторов среды в любой точке водотока появляются «процессные зоны», в которых развитие отдельных функциональных или экологических групп гидробионтов стимулируется или угнетается. В результате этого могут возникать самые различные сочетания «сгущений жизни» и биотические лакуны.

Несмотря на солидный теоретический багаж, до сих пор не представлены четкие методические аспекты практического анализа продольного распределения гидробионтов на основе данных мониторинга. «Являются ли статистически значимыми различия видового состава в двух сравниваемых створах водотока, или они определяются случайными популяционными флуктуациями и ошибкой гидробиологической съемки?», «Как построить количественную модель русловой динамики таксономической структуры и оценить её адекватность?» – ответы на эти вопросы еще ждут своего решения. Поэтому **третья** цель нашей книги – проанализировать, в какой мере традиционные математические приемы макроэкологии могут способствовать практической реализации этой проблемы.

Методы и модели анализа видовой структуры пресноводных сообществ были рассмотрены нами на примере средней реки Сок вместе с её верхним притоком р. Байтуган.

Это – весьма типичный и в наибольшей степени исследованный равнинный водоток, достаточно протяженный (около 400 км), чтобы включить в себя необходимое разнообразие биотопических и ландшафтных элементов, и находящийся в зоне умеренной антропогенной нагрузки. В качестве модельного объекта биологического мониторинга нами рассматривались сообщества макрозообентоса, которые обладают признанными биоиндикационными свойствами и являются наиболее стабильным и разнообразным компонентом лотических экосистем. В расчетах использовался материал на основе 147 гидробиологических проб, взятых на 22 станциях за период с 1990 по 2007 г., в результате чего было обнаружено 375 видов, принадлежащих к различным таксономическим группам.

«Таксономическая структура водных сообществ не является строго зафиксированной композицией видов и, в зависимости от свойств области исследования, характеризуется различной степенью пространственной и временной изменчивости». Это утверждение выглядит обыденной банальностью, но… мало кто воспринимает его как руководство к практическому действию. Огромные пласты научных работ, от простейших графиков и таблиц до изощренных диаграмм канонического анализа соответствий, полны свидетельств типа «На станции А водотока Б был зафиксирован следующий состав гидробионтов с показателями обилия…». Здесь нет ошибки, если на станции А была взята только одна гидробиологическая проба. Однако если проб было несколько, то простой вопрос «А какой механизм использовался для усреднения проб?» имеет далеко не простой подтекст. Теоретически рекомендуемых способов обработки много – от расчета средних и медиан до экспертного оценивания, но ни один из них не является корректным в случае отчетливого «негауссового» распределения данных об обилии видов в пробах.

В этой связи **четвертой** целью монографии явился всесторонний анализ зависимости традиционных показателей видовой структуры водных сообществ (число видов, индексы разнообразия, кривые ранг–обилие и др.) от «выборочного усилия», т. е. количества обработанных гидробиологических проб. Нами были предприняты попытки разработать универсальные алгоритмы оценивания изменчивости видового состава местообитаний при неэквивалентном выборочном усилии, которые выносятся на суд читателя.

Все четыре цели мы пытались достичь, рассматривая комплекс связанных с ними проблем в определенной последовательности по следующей схеме:

- рассматривается та или иная задача макроэкологии и описываются методы её решения по литературным данным;
- обсуждаются нуль-модели, используемые для оценки значимости искомых метрик или факторов;
- анализируется устойчивость решения и выводов при изменении выборочного усилия;
- полученные результаты рассматриваются с точки зрения, насколько они могут опровергнуть нулевую гипотезу о видовой однородности сообщества макрозообентоса на всем протяжении от истоков р. Байтуган до устья р. Сок.

Настоящая монография написана в жанре «научно-популярной литературы для студентов и специалистов». В связи с этим в ней нет ни абсолютных, ни относительных истин, а лишь точки зрения разработчиков тех или иных методов. Убедительно просим читателя не отождествлять нас с авторами цитируемых работ – наши взгляды на проблему могут серьезно расходиться. В книге сделана попытка, по возможности методически полно, бережно и лаконично, передать цели, смысл и даже терминологию авторов проанализированных статей и монографий, снабдив их лишь некоторыми скупыми комментариями. Даже если изложенные мысли или полученные результаты казались нам мало симпатичными, но включали в себя некоторый элемент нетривиальности научных подходов, мы старались беспристрастно представить спектр всех точек зрения, полагая, что время покажет, а читатель – сам разберется.

В итоге материал, представленный на ваш суд, образует весьма «размытое множество» догадок, сомнений и противоречий. Однако наука – это *систематическое расшире-*

ние области человеческого незнания. Поэтому всё изложенное в книге можно и нужно критиковать, но лучше на примере своих собственных разработок.

Иногда по тексту книги мы не могли удержаться от вольных рассуждений на внешне отвлеченные, но, как нам казалось, контекстуально важные темы. Мы не вняли советам выкинуть фрагменты, кажущиеся лишними (за что просим прощения у взыскательного читателя), однако снабдили их пометками «*Отступление от темы*…». И, наконец, мы пытались следовать замечанию академика П.Л. Капицы: «*Наука должна быть весёлой, увлекательной и простой. Таковыми же должны быть и учёные*».

При подготовке рукописи и проведении расчетов мы неизменно получали благожелательную помощь ученых-экологов всего мира. В связи с этим, выражаем искреннюю признательность профессору Роберту Колуэллу (Robert K. Colwell, Коннектикутский университет [США]), профессору Вернеру Ульриху (Werner Ulrich, Университет Николая Коперника, Торунь [Польша]), профессору Томасу Кристу (Thomas O. Crist, Университет в Майами [США]), доктору Петеру Солямосу (Peter Solymos, Университет провинции Альберта, Эдмонтон [Канада]), профессору Валерио Пилару (Valerio D. Pilar, Университет штата Рио Гранди ду Сул, Порту-Алегри [Бразилия]), профессору Дейву Робертсу (David W. Roberts, Университет штата Монтана [США]), доктору Луи Джосту (Lou Jost, Тунгвардиа [Эквадор]), профессору В.Н. Максимову (Московский университет им. М.В. Ломоносова), профессору В.В. Богатову (Биолого-почвенный институт ДВО РАН, Владивосток), профессору В.В. Акатову (Майкопский государственный технологический университет), д.т.н. В.В. Фуфаеву (Центр системных исследований, Москва), доктору Натану Цейтлину (Natan Zeitlin, Геттинген [Германия]), к.б.н. С.С. Крамаренко (Николаевский государственный аграрный университет [Украина]), к.б.н. В.Н. Якимову (Нижегородский университет им. Н.И. Лобачевского), к.б.н. А.Е. Бобыреву (Институт проблем экологии и эволюции им. А.Н. Северцова РАН**)**, к.б.н. А.Б. Новаковскому (Институт биологии Коми НЦ УрО РАН) и многим другим коллегам, с которыми мы обсуждали отдельные методы и разделы этой книги.

Авторы также искренне благодарны сотрудникам Института к.б.н. Л.В. Головатюк за сбор и подготовку гидробиологического материала, Э.В. Абросимовой за помощь в проведении компьютерных расчетов, к.б.н. О.Л. Носковой за техническую и редакторскую поддержку при компьютерной верстке книги.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (грант № 07-04-96610) и в рамках программ Президиума РАН «Биологическое разнообразие» (раздел «Динамика биоразнообразия и механизмы обеспечения устойчивости биосистем») и Отделения биологических наук РАН «Фундаментальные основы управления биологическими ресурсами».

ГЛАВА 1. КОНЦЕПЦИИ МАКРОЭКОЛОГИИ И АНАЛИЗ ВИДОВОЙ СТРУКТУРЫ СООБЩЕСТВ

Гипотез я не придумываю…
Исаак Ньютон (1642-1727 гг.)

1.1. Модели динамики популяций и «колодки» экологического мышления

Структурные особенности и параметры биологического разнообразия любых экологических систем (водных, растительных и др.) определяются двумя основными составляющими – числом видов и характером распределения их популяционной плотности. Под «характером распределения» может пониматься как статическое соотношение показателей обилия отдельных видов (например, количества особей, приходящихся на единицу площади или объема среды обитания), так и закономерности изменения этих численностей на определенной территории (акватории) в биологическом времени. Как правило, причинами, объясняющими динамические колебания популяционной плотности, считаются те или иные естественные или антропогенные факторы окружающей среды. Первая группа факторов объясняет изменчивость обилия видов, в частности, климатическими условиями или конкуренцией за пищевые ресурсы, а вторая – «загрязнениями природной среды, от которых и происходит всяческое зло» (Тутубалин и др., 1999, раздел 4.1).

Первые модели динамики популяций – это *ряд Фибоначчи* (1202 г.), *модель экспоненциального роста Мальтуса* (1798 г.), *модель ограниченного роста Ферхюльста* (1838 г.). В начале XX в. появились первые модели взаимодействия видов. Классической книгой современной математической экологии является труд В. Вольтерра "Математическая теория борьбы за существование" (Volterra, 1931; Вольтерра, 1976). *Модели двух популяций Лотки и Вольтерра* объясняют хронологические колебания обилия видов, в основном, межвидовыми взаимодействиями. Эти модели рассматривают несколько биологических видов и набор имеющихся ресурсов (например, запас пищи), причем о компонентах системы формулируются предположения, такие как: «хищные виды поедают жертвы, причем в единицу времени количество съеденных жертв всегда пропорционально вероятности встречи особей этих двух видов, т. е. произведению количества хищников на количество жертв». Иными словами, динамика популяций представлена здесь неограниченно продолжающимися во времени периодическими колебаниями численностей двух взаимодействующих видов – хищника и его жертвы.

Уравнения системы «хищник – жертва» послужили отправной точкой для создания большинства динамических моделей в экологии вплоть до сегодняшнего дня и оказали существенное влияние на все физико-математическое мышление. «Работы Вольтерра, Лотки, Гаузе и других по теории популяций принадлежат к классическим трудам общей теории систем. В них впервые была продемонстрирована возможность развития концептуальных моделей для таких явлений, как борьба за существование, которые могут быть подвергнуты эмпирической проверке» (Bertalanffy, 1962, p. 13).

В этом направлении экологии к настоящему времени разработано огромное множество самых разнообразных дискретных и непрерывных детерминистических и стохастических математических моделей (одно время в издательстве "Springer" выходила целая серия научных работ по математической биологии). При этом современные математические модели в экологии можно разбить на три класса. *Первый* – описательные модели: регрессионные и другие эмпирически установленные количественные зависимости. Они применяются обычно для описания отдельных процессов и зависимостей или включаются как фрагменты в большие имитационные модели. *Второй* – качественные динамические модели, которые воспроизводят механизмы и эффекты изучаемого процесса, такие, например,

как колебательный характер изменения биомассы или образование неоднородной в пространстве видовой структуры. *Третий* класс – имитационные модели конкретных экологических и эколого-экономических систем, учитывающие всю имеющуюся информацию об объекте. Цель построения таких моделей – детальное прогнозирование поведения сложных систем или решение оптимизационной задачи их эксплуатации. Убеждение в том, что подобные модели описывают какие-то существенные черты реальных экосистем, долгое время было широко распространено.

Обычно при математическом моделировании задача состоит в том, чтобы отследить динамику компонентов экологической системы. При этом преследуются различные цели эксплуатации моделей и делаются соответствующие исходные предположения, которые один из пионеров математической биологии А.А. Ляпунов (1972) сформировал следующим образом:

Уровень	Исходные предположения	Основные задачи	Компоненты моделей
А	Биологические характеристики компонентов неизменны, так же как и взаимоотношения между ними. Система считается однородной в пространстве.	Изучаются изменения во времени численности (биомассы) компонентов системы.	Системы обыкновенных дифференциальных уравнений с постоянными коэффициентами.
Б	Вводится предположение о закономерном изменении соотношений между компонентами.	Изучается кинетика численности компонентов в условиях многолетнего или сезонного изменения внешних условий, либо при заданном характере эволюции экоформ.	Дифференциальные и дифференциально-разностные уравнения с переменными коэффициентами.
В	Объекты считаются разнородными по своим свойствам и подверженными действию отбора. Эволюция экоформ определяется условиями существования системы.	Изучается, с одной стороны, кинетика численности компонентов, с другой – дрейф характеристик популяций. К ним относятся многие задачи популяционной генетики.	Аппарат теории вероятностей.
Г	Отказ от территориальной однородности и учет зависимости усредненных концентраций от координат.	Вопросы, связанные с пространственным перераспределением живых и косных компонентов системы (учет вертикальной стратификации или тренда по руслу водотока).	Аппарат дифференциальных уравнений в частных производных. Разбиение имитационных моделей на пространственные блоки.

Задачи пространственной организации экологических систем представляют особый интерес. Распространенное мнение исследователей, дискутируемое в литературе, состоит в том, что пространственная неоднородность плотности популяций видов связана в основном с ландшафтно-климатическими факторами. Вместе с тем, все более глубоко осознается тот факт, что пространственная структурированность экологических систем может быть обусловлена не только исходно существующей ландшафтной неоднородностью, но и спецификой локальных стохастических взаимодействий популяций между собой и с косными компонентами среды.

В последнее десятилетие в специальной литературе высказывается достаточно скептическое мнение о том, что в своем большинстве многие разработанные до недавнего времени математические модели либо просто не адекватны, либо практически бесполезны для развития теоретической биологии. Полученные с их помощью биологические результаты в достаточной мере тривиальны и могли бы быть получены без применения моделей (возможно, в несколько иной, теоретически менее безупречной форме). Такова, в частно-

сти, позиция авторов историко-методологического анализа (Тутубалин и др., 1999), которые видят исключение лишь для некоторых интересных моделей, используемых в генетике.

Действительно, математические модели типа «хищник – жертва» слишком просты, чтобы адекватно описывать реально существующие природные сообщества. При более детальном изучении конкретной экосистемы каждый трофический уровень разбивается на более мелкие, но если основной единицей в экологии считать биологический вид, то общее число таких единиц в любой реальной экосистеме становится колоссальным, а трофические (и иные) взаимосвязи между ними будут невообразимо сложны. Нельзя думать, чтобы даже парные взаимосвязи компонентов экосистемы можно было описать моделью в виде дифференциальных уравнений, которые поддавалась бы решению в классическом смысле (т. е. интегрированию в формульном виде) или хотя бы качественному исследованию. Попытки как-то «расчистить» пространство переменных, устранив ненужные частности, часто приводили к ряду печальных опытов, которые убедили человечество в том, что *экосистемы нужно изучать в целом*, не ограничиваясь лишь теми видами, которые представляют, например, экономический интерес.

Прогресс вряд ли может оказать принципиальное влияние на разрешение проблемы «проклятия видовой размерности». Количественный расчет динамики природных экологических систем, видимо, выходит за пределы возможностей современной науки, даже если используется мощный вычислительный потенциал электронной техники или математические возможности новейшей теории динамических систем.

Отклонение от темы (общенаучные размышления). Как считает В.Н. Тутубалин с соавторами (1999, раздел 5.5), «наука как трансперсональное существо – эгрегор, потерпев неудачу в каком-то своем замысле, поступает, как и всякое другое существо. От того, что невозможно, она отказывается и ставит перед собой иные, более скромные задачи... Не заниматься обработкой результатов измерений совсем (кроме элементарного представления результатов в каком-то удобочитаемом виде) на том основании, что все равно не получится ничего путного – скверный выход». Поэтому приходится разрабатывать новые, более сложные модели, основанные на иных концепциях, надеясь на их адекватность. Либо решать известными средствами частные задачи изучения каких-то небольших конкретных сообществ, смягчив, однако, ту философию, которая определяет, каких успехов мы ожидаем от результатов математической обработки. «В любом случае дело кончается тем неоспоримым аргументом, что другой, более совершенной науки у нас нет» (Заключение).

В последние десятилетия в научной методологии идет активный поиск терминов, обозначающих такое понятие как, общий *шаблон* или *схему* реализации научного исследования. К таким словоформам можно отнести и латинскую *концепцию*[1] (см., например, учебное пособие "Концепции современного естествознания" [Аруцев и др., 2006]) и греческую *парадигму* (советуем ознакомиться с дискуссией о смене парадигм в экологии [Розенберг, Смелянский, 1997]). В.Н. Тутубалин с соавторами (1999) продолжает эту традицию, развернув «учение о *колодках* мышления». Выражаясь кратко, «колодка», с одной стороны, – это нечто направляющее и организующее, а с другой стороны, – ограничитель излишней свободы[2]. Авторы «учения» достаточно

[1] «Концептуальный аспект теоретического знания выражает, прежде всего, парадигмальное "сечение" последнего, задает его топику и риторику, то есть определяет релевантные области применения и способы выражения конституируемых на основе развертывания "порождающей" идеи систем понятий – базовых концептов» (http://ru.wikipedia.org).

[2] Спектр значений русского слова «колодка» включает в себя, как обувные колодки (т. е. некоторые парадигмальные схемы порождения моделей), так и тормозные колодки или колодки арестантов (с вполне понятными аналогиями).

подробно исследовали развитие колодок мышления при моделировании динамики природных экологических систем от «пророков» (А. Лотка и В. Вольтерра) и «апостолов» (Г.Ф. Гаузе) до «приходских священников».

К новым «колодкам» в математическом моделировании можно отнести следующие представления и подходы, подробное описание которых можно найти, например, в главе 6 "Поиски порядка в хаосе" (Тутубалин и др., 1999) или в "Реестре динамических моделей в экологии", подготовленном кафедрой биофизики МГУ под руководством проф. Г.Ю. Ризниченко (http://www.dmb.biophys.msu.ru/models):

1. Многие природные системы можно рассматривать как самоорганизованные критические системы (Jorgensen et al., 1998; Milne, 1998). Они объединяют в себе концепцию критичности с концепцией самоорганизации и основаны на свойствах неравновесных систем и механизмах обратной связи. Предполагается, что экосистемы функционируют в состоянии устойчивого динамического равновесия на границе хаоса. Хотя скорости процессов в таких системах сбалансированы таким образом, чтобы оптимально использовать доступные ресурсы, даже незначительный дисбаланс в параметрах и скоростях может приводить на отдельных промежутках времени к хаотическому поведению и коренной перестройке отдельных компонентов или всей системы в целом. Развитие этой концепции нашло свое отражение в моделях Э. Реншоу (Renshaw, 1991), представлении об экосистеме, как о «странном аттракторе» (Розенберг, 1985; Schaffer et al., 1986), моделях нелинейной демографической динамики (Dennis et al., 1995) и др.
2. Экология имеет дело с принципиально нередуцируемыми иерархическими системами (Jorgensen et al., 1998; Алимов, 2001), где каждый элемент (функциональная единица) также представляет собой систему, которая, в свою очередь, является элементом системы более высокого уровня организации. Концепция иерархической организации биологических систем (Allen, Starr, 1982) развивается в неразрывной связи с проблемой характерных масштабов, в пределах которых формируется структура определенного уровня. В работе А.И. Азовского (2001) проанализирована связь между характерными масштабами пространства [**L**] и времени [**T**] для естественных иерархических систем различной природы и показано, что имеющиеся данные хорошо аппроксимируются степенным соотношением вида $[\mathbf{T}] = a[\mathbf{L}]^b$, где b – параметр масштаба (экспонента Хёрста), a – переходный коэффициент. Показано, что для планктонных экосистем характерны значения b, варьирующие от 1 (процессы прямого переноса) до 2 (случайные диффузионные процессы типа броуновского движения).
3. Согласно *теории прерывистого равновесия Элдриджа–Гулда*, эволюционный процесс происходит скачками, перемежающимися с длительными периодами, в которых не происходит существенных изменений. Экосистемы и сообщества, являющиеся, по сути, динамическими системами, проходят через следующие четыре характерные фазы адаптивного цикла (Carpenter et al., 2001): быстрого роста и активного использования ресурсов среды (r-фаза), консервации (стагнации) (K-фаза), коллапса (разрушения, деструкции) (Ω-фаза) и обновления (реконструкции) (α-фаза). Наиболее важной особенностью этого цикла является существование относительно коротких периодов, в течение которых в системе происходят наиболее существенные изменения – фазы Ω и α. При этом исчезают и появляются новые компоненты (например, виды), изменяются система связей и интегративные свойства. Более длительным периодам постепенного эволюционного развития системы соответствуют фазы r и K. Это позволяет провести, хотя бы в принципе, разделение переменных на быстрые и медленные и упростить тем самым модельное изучение систем, выделив характерные масштабы времен.

Наш краткий экскурс в область математического моделирования экологических систем преследует цель показать, что, несмотря на длинный и ветвистый путь, пройденный этой областью науки, вопрос о практических методах изучения множественных стати-

стических связей между динамикой численности различных биологических видов и комплексом биотических и абиотических факторов до настоящего времени не был поставлен во главу угла, хотя его актуальность была вполне очевидна уже в 60-е годы XX столетия. Потребовались десятилетия, чтобы модели со случайными вмешательствами и диссипативные структуры стали систематически изучаться в экологии. Однако, остановившись перед «*проклятием таксономической размерности*», «модельеры» увлеклись более приятным делом – любоваться, например, математическими красотами детерминированного хаоса (руководствуясь, видимо, известной максимой «если в модели меньше трех переменных, то она не заслуживает внимания, а если в модели больше семи переменных, то она все равно неразрешима»).

Процессы, которые лежат в основе динамики биологического разнообразия, являются комплексными и разнонаправленными. Эти процессы простираются от генетических вариаций в пределах отдельных геномов до ландшафтно-географических градиентов видов в масштабе всей Земли. Обширное количество деталей, которые были бы необходимы, чтобы описать всю эту сложность, превышают способность любого человеческого мозга, даже оснащенного современной компьютерной техникой. Как следствие, возникла необходимость найти какие-либо упрощенные способы осмыслить эту сложность, чтобы облегчить наше понимание структуры и механизмов развития экологических сообществ. Это побудило экологов к активной деятельности, связанной с изобретением новых «колодок» разной степени эфемерности…

1.2. Закономерности пространственного распределения гидробионтов в речных экосистемах (развитие теоретических концепций)

Основная задача экологии сообществ состоит в том, чтобы выделить закономерности, которые определяют композиционный состав, разнообразие, трофическую организацию и стабильность ассоциаций видов в природных объектах, и понять процессы, которые определяют эти закономерности. В центре обсуждаемых проблем стоит дилемма: являются ли сообщества простыми сочетаниями видов, случайно извлеченными из регионального пула, или они являются ассоциациями взаимодействующих между собой таксономических групп, состав которых сильно детерминирован факторами окружающей среды.

Таким образом, в этой книге (как и в любой другой литературе по экологии) речь пойдет о *сообществах*, *местообитаниях* и методах оценки их связи между собой. Предварительно отметим, что многочисленные литературные источники различной направленности не дают строгого и однозначного определения этих терминов. М. Бигон с соавторами (1989) определяет *сообщество* как «совокупность популяций видов, встречающихся *вместе в пространстве и времени*». Т. Саутвуд (Southwood, 1986) делает важное уточнение, что эта совокупность является «*организованной*» (т. е. обладающей вполне детерминированными внутренними связями), а местообитания, где размещается все множество популяций видов сообщества, обладают признаками *специфичности*. Матрица естественной среды (Brown, Southwood, 1987) – классическая форма представления взаимосвязей между сообществами и местообитаниями, а также способ характеризовать виды вдоль градиента адаптационных стратегий (см. рис. 1.1).

В гидробиологической литературе под термином сообщество также понимается «совокупность совместно обитающих и взаимодействующих организмов» (Шкорбатов, 1985, с. 73). Нередко сообщество отождествляется с термином *биоценоз*, и тогда под *биоценозом* понимается «совокупность растений, животных и микроорганизмов, которые населяют участок суши или акваторию и характеризуются определенными отношениями как между собой, так и с абиотическими факторами среды. *Биоценоз* – совокупность организмов, входящих в состав экосистемы, сформировавшаяся вследствие борьбы за существова-

ние путем естественного отбора и других факторов эволюции» (Романенко, 2004). Резюмируя определения, представленные в гидроэкологической литературе (Зернов, 1949; Поддубный, 1971; Козлов, Абрамович, 1982; Романенко, 2004 и мн. др.), термин *биоценоз* означает подвижную систему отдельных организмов, популяций и *сообществ*, связанных между собой на определенной территории водоема (местообитании) общностью требований к условиям среды. В то же время состав биоценозов во многом определяется стохастическими тенденциями миграции и трансформации видов, а также особенностями эволюционно-исторического процесса в изучаемом регионе.



Рис. 1.1. **Матрица естественной среды (по: Brown, Southwood, 1987)**

В приведенных определениях идея сообщества включает в себя две составляющие. Во-первых, речь идет о группах организмов, которые, однако, в каждом конкретном случае выделяются, исходя из субъективных задач локального исследования. Эти группы могут быть ограничены одной гильдией или включать в себя все формы «от бактерий до буйволов» (Мэгарран, 1992). Безусловно, сообщества лучше всего представлять как каскад взаимодействующих уровней биологической организации, но удачные примеры комплексного изучения всех иерархических составляющих водных экосистем крайне немногочисленны. Во-вторых, сообщество представляется в некоторых пространственных границах, которые на практике чаще всего оказываются размытыми. Еще более «тонкой материей» представляются рамки существования сообщества в биологическом времени. Таким образом, понятие «сообщество» остается неопределенным потому, что сами сообщества по своей природе являются неопределенными и нечеткими образованиями. Поэтому в работах, визитной карточкой которых является строгость в формулировках, вместо термина «сообщество» стараются использовать термины «композиция» (Chao et al., 2005) или «ассамблея видов» (Lambshead et al., 1983). Тем не менее, целесообразность в рабочем определении «сообщества» существует, а необходимость установления его границ возникает в большинстве экологических исследований.

Самой интересной с экологической точки зрения особенностью сообщества является прямое или косвенное взаимодействие видов друг с другом. Их интенсивность «определяет тип того или иного сообщества, образуя своего рода шкалу межвидовых взаимодействий. На одном полюсе шкалы находятся сообщества, состоящие из коэволюционирующих и в высшей степени взаимодействующих между собой видов, на другом – сооб-

щества почти никак не связанных между собой видов, которые просто обитают в одной и той же среде, имея сходные требования к ресурсам» (Митчелл, 2001, с. 154). Несмотря на то, что количественная оценка степени и частоты межвидовых связей является пока нерешенной научной проблемой, многие экологи придерживаются мнения, что большинство сообществ ближе по своей организации к «индивидуалистическому типу». Возможно, это связано с тем, что чаще всего изучаются не взаимосвязи в сообществах в форме трофических сетей, а гильдии, члены которых взаимодействуют между собой лишь опосредованно.

Выделяется (Southwood, 1966) 12 типов местообитаний: постоянные, сезонные, непредсказуемые, эфемерные, непрерывные, пятнистые, изолированные и другие комбинации фрагментов пространства и времени, где совместно встречаются виды. Границы отдельных классов представляются весьма размытыми: как, например, можно количественно представить степень эфемерности? Также не всегда понятно, в каких единицах можно оценить «специфичность» местообитания, которая определяется сложным комплексом факторов среды.

В функционировании сообществ, составляющих водные и наземные экосистемы, существуют определенные специфические особенности. Так, например, в водной среде формируются благоприятные условия для тесных биохимических взаимоотношений гидробионтов, тогда как наземные сообщества растений и животных пространственно разделены наземной и воздушной средами и метаболические связи в этих условиях проявляются менее четко, чем у водных организмов (Романенко, 2004).

В последние годы формируется принципиально иной взгляд на структуру экосистем и разрабатываются новые формализованные подходы к их изучению, не рассматривающие сообщества как абстрактную «паутину» взаимоотношений между видами (Митчелл, 2001). Основное внимание уделяется динамике изменения видового состава, которая определяется процессами, происходящими на региональном уровне.

Выявление закономерностей распределения гидробионтов и их сообществ по продольному профилю водотоков относится к фундаментальным задачам экологии и гидробиологии. Нет сомнения, что основную роль в этих процессах играют абиотические факторы среды, среди которых ведущее положение занимает скорость течения, которая закономерно изменяется по руслу рек и во многом определяет турбулентный режим перемешивания, аккумуляцию размытого материала и многие другие физико-химические характеристики. Жизненные пространства верховий и среднего течения (ритраль) биоценотически отделены от мест обитания рыб в долинном течении (потомаль), в то время как сообщества устьевых участков (эстуарий) находятся под сильным влиянием экосистемы водоема-приемника и зависят от его типологии и специфики.

Кроме общих гидрологических, ландшафтно-топографических и морфометрических особенностей, пространственная неоднородность водотока (а соответственно, и всей экосистемы) определяется многими локальными биотопическими факторами: характером распространения водных растений, условиями освещенности, температурными градиентами, наличием питательных веществ, содержанием кислорода и др. Мозаичность («пятнистость») размещения популяций бентосных организмов сильно зависит от мелкомасштабной изменчивости грунтов (включения песка, илистых отложений, гравия). Для планктонных сообществ пространственная изменчивость популяционной плотности имеет трехмерный характер, т. е. возникает необходимость учитывать вертикальную стратификацию по глубине. Линейное разграничение гидробиоценозов затруднительно из-за подвижности животных, способности к дрейфу или миграции гидробионтов при малейшем изменении условий существования, в результате чего «базовые» виды совместно обживают общие переходные территории. Наконец, кроме морфологических особенностей водотока необходим детальный учет внешних факторов: распределение по береговой линии интенсивности поверхностного стока, выделение мест локализации сброса загрязненных вод антропогенного характера, областей рекреационного использования и т. д.

Закономерности размещения организмов в пространстве (или *спейс-структура популяций*) обуславливаются реакцией особей на воздействие внешних факторов и особенностями их собственных взаимоотношений. Изучение пространственной структуры осуществляется на разных масштабных уровнях, из которых можно выделить три:

- глобальный (субконтинентальный) уровень, изучающий крупные реки и цепочки водохранилищ, размещенные на протяжении тысяч километров в разных ландшафтных зонах; здесь важен комплексный характер биосферных исследований с привлечением геоморфологии, климатологии и др. (Розенберг и др., 1995);
- уровень отдельного водотока или региональной сети малых и средних рек в линейном масштабе от 10 до 300 км (пример многофункциональных исследований такого рода см.: Биоиндикация экологического.., 2007);
- уровень отдельных участков или створов реки, когда детально изучается мелкомасштабное (менее 1 км) распределение сообществ гидробионтов в соответствии с биотопической неоднородностью в продольном, поперечном и вертикальном профилях русла.

Основной концепцией изучения водотоков, принятой до 80-х годов прошлого века, было выделение изолированных внутренне гомогенных зон, последовательно расположенных вдоль продольного профиля русла реки (Hawkes, 1975). Части водоема выделялись как по функциональному, так и по хорологическому принципу с учетом морфологических особенностей водного объекта и локализации внешних воздействий. Такое районирование, отражающее в той или иной мере объективно существующую гетерогенность водотока, теоретически должно осуществляться с использованием многофакторных математических методов выделения множества непересекающихся целостных кластеров (Баканов, 1997), однако на практике чаще всего осуществлялось приблизительно на основе экспертных оценок. Эта концепция нашла отражение в моделях экосистем, которые имели ярко выраженный камерный (блочный) характер (Страшкраба, Гнаук, 1989; Меншуткин, 1993). Деление водотока на дискретную последовательность изолированных участков нашло отражение в стратегии и методике проведения гидробиологической съемки, в результате чего отбор серий проб проводился не в случайных точках (random sampling), а согласно заранее разработанной субъективной схемы районирования (passive sampling).

В качестве альтернативы *теории дискретных изолированных зон Хоукса* (Hawkes, 1975), Р. Вэннот с соавторами (Vannote et al., 1980) рассматривали водоток как целостную экосистему и сформулировали концепцию *речного континуума* (river continuum concept – RCC). Основные положения RCC сводятся к следующему:

1. Для всех рек существует непрерывный градиент важнейших гидрологических условий, ориентированный по продольному профилю русла. К таким факторам, регулирующим биотические процессы и динамику поступления питательных веществ, относятся ширина и глубина водотока, скорость течения, турбулентность перемешивания, расход воды и др.
2. В соответствии с градиентом абиотических факторов закономерно и непрерывно изменяются структурно-функциональные особенности сообществ гидробионтов, реализуя тем самым стремление экосистемы минимизировать диссипативные потери вещества и энергии в каждой точке пространства и достичь равновесного энергетического оптимума всего водотока в целом.
3. Связь между физической и биотической составляющими экосистемы состоит в том, что функциональная организация сообществ определяется стратегией баланса между тенденцией максимально использовать энергию, поступающую с необходимыми для жизнедеятельности ресурсами (пища, субстрат и т. д.), с одной стороны, и стремлением равномерно распределить эту энергию в течение года и по руслу реки, с другой. Биоценозы структурно приспосабливаются к градиенту внешних условий, чтобы наиболее эффективным образом аккумулировать и перераспределить кинетическую энергию физической системы. Например, сообщества в нижнем течении сконструированы так,

чтобы в максимальной мере утилизировать те остатки ресурсов, которые были недоиспользованы гидробионтами в верховьях реки. В то же время принципиально более низкий КПД использования ресурсов для сообществ, обитающих выше по течению, представляется вполне предсказуемым.

4. Различия в стратегии функционирования находят отражение в закономерном изменении видовой структуры сообществ. Этот процесс представляется как формирование непрерывной последовательности синхронизированной смены таксономического состава комплекса доминантов. Если тот или иной вид испытывает затруднения в адаптации к изменяющимся факторам микросреды, он заменяется другими видами, выполняющими по существу ту же самую функцию, но более приспособленными к новым условиям.
5. Понятие смены видовых ассоциаций подразумевает статистический подтекст. Разумеется, полное исчезновение вида в естественных водотоках является маловероятным событием. Если рассматривать кривую распределения обилия видов в речной системе, то все виды, как правило, присутствуют в любой точке на пространственной оси, но численность их популяций на различных участках может существенно отличаться.

Динамику основных экосистемных показателей авторы концепции (Vannote et al., 1980) представили на графиках (рис. 1.2).

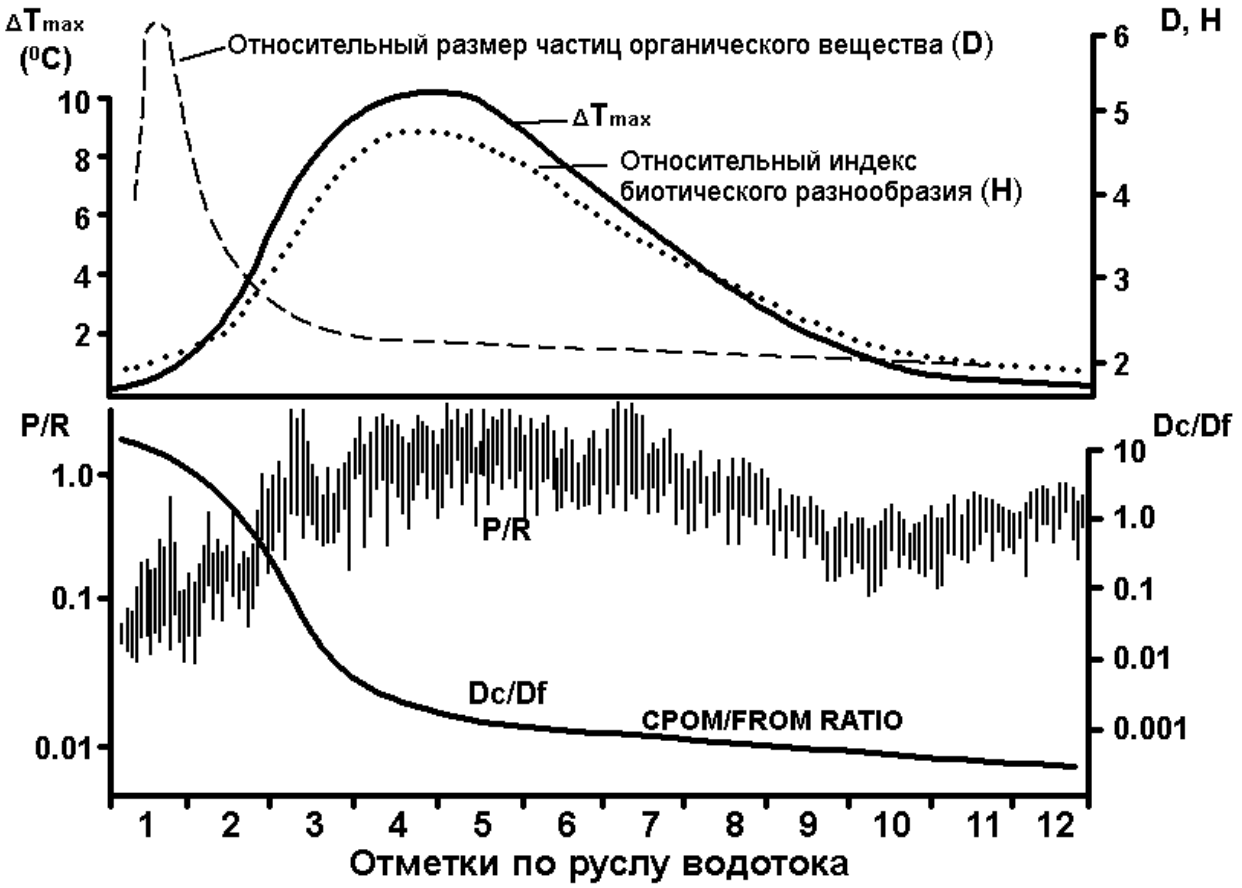


Рис. 1.2. **Гипотетическое распределение основных параметров экосистемы в пределах речного континуума по 12 участкам от истока к устью:**

ΔT – перепады температуры воды, H – относительный индекс видового разнообразия, D – относительный средний размер частиц органического вещества, Dc/Df – соотношение крупной и мелкодисперсных фракций органического вещества, P/R – отношение продукции к тратам на обмен (по: Vannote et al., 1980).

Наибольший перепад температурного режима ΔT_{max} они отнесли к среднему течению рек, полагая, что в верховьях температурная стабилизация осуществляется за счет питающих водных источников (например, родников), тогда как в устьевых участках это происходит вследствие аккумуляции водных масс и подпора со стороны водоема-приемника. Стрессовые воздействия температурного дрейфа сопровождаются повышением видового разнообразия сообществ на средних участках, как естественная реакция экосистемы по поддержанию своей стабильности и устойчивости.

В верховьях рек, протекающих в лесистой местности, преобладают крупные частицы аллохтонного органического вещества (листовой опад) и в сообществах гидробионтов преобладают первичные консументы, относящиеся к «механическим разрушителям». Отношение продукции (P) к тратам на обменные процессы (R) в этой части экосистемы значительно меньше 1, что указывает на её гетеротрофность. Ближе к ритрали (среднее течение) увеличивается поступление энергии с солнечной радиацией, а в потоке начинает преобладать тонко измельченное органическое вещество. Первичную продукцию на этих участках обеспечивают водоросли и водные макрофиты, а в сообществе консументов – фильтраторы, собиратели и хищники. Считается, что система здесь автотрофна, а отношение P/R обычно выше 1 (рис. 1.2). На нижних равнинных участках рек течение замедляется, вода становится мутной, что затрудняет фотосинтез. На большинстве трофических уровней видовое разнообразие снижается, и сообщество вновь становится гетеротрофным.

После создания теории речного континуума появился целый ряд работ, подтверждающих либо опровергающих описанные закономерности (Minshall, 1988; Цимдинь, Лиепа, 1989; Крылов, 2005; Thorp et al., 2006; Экосистема малой.., 2007 и др.). Например, было показано (Statzner, Higler, 1986), что вместо непрерывных градиентов по течению реки часто встречаются геоморфологические разрывы (излучины, поймы), которые приводят к резким изменениям в размере субстрата и, следовательно, к скачкообразным сменам видовых ассоциаций сообществ. Д. Перри и Д. Шеффер (Perry, Schaeffer, 1987), не найдя в своих водных объектах отчетливой континуальности, характеризовали распределения видов в реке как «пунктирные» градиенты (punctuated gradients), а не экологические континуумы. Д. Монтгомери (Montgomery, 1999) заключил, что RCC справедлива только для равнинных рек с относительно постоянным климатом и простой геоморфологией. Наконец, А.Ф. Алимов (2001, с. 10) совершенно справедливо пишет, что «в любых (!) реках условия существования в верхнем, среднем и нижнем течении совершенно различны». Следует также согласиться с авторами, которые указывают, что для полного понимания экологических процессов в реках необходимо исследовать всю область бассейна, из которого реки получают воду, химические и питательные вещества (Митчелл, 2001; Романенко, 2004).

Нет сомнения, что основную роль в распределении сообществ гидробионтов играют абиотические факторы среды, среди которых ведущее положение занимают гидрологические параметры, закономерно изменяющиеся по продольному профилю водотоков. Однако формирование речных сообществ гидробионтов в значительной мере зависит и от нарушений, имеющих разный временной интервал существования (Resh et al., 1988; Townsend, 1989). Нарушение определено как любое относительно дискретное во времени событие, которое разрушает экосистему, сообщество или популяционную структуру и изменяет физические условия среды, доступность субстрата и пищи (Resh et al., 1988). Различного рода нарушения, в зависимости от изначальных физико-химических характеристик участков, а также глубины и мощности преобразований, способствуют возникновению по продольному профилю водотока естественных *рефугиумов*[3] – участков, резко отличающихся скоростью течения, глубиной, площадью затопленной поймы, содержанием органических и био-

[3] См. аналогичные термины – *биотические лакуны* и *сгущения жизни* (Залетаев, 1997; Крылов, 2005)

генных веществ и др.

Очевидно, что при возникновении в речной среде экстремальных природных ситуаций биологическая система включает адаптационные механизмы управления популяциями. В результате в реке формируются слабо и сильно заселенные участки, характеризующиеся либо коренными сменами комплексов видов-доминантов вследствие массового развития отдельных таксономических групп, либо пропорциональным изменением численности (биомассы) организмов.

Традиционная концепция экологии основывалась на убеждении, восходящим к идеям Лотки–Вольтерра, что конкуренция за ресурсы является главной силой структурирования сообщества, а идеи о сосуществовании видов формулировались в контексте их конкурентного исключения. Эта простая детерминированная модель оказалась непригодной (Yodzis, 1986) для описания последовательностей рефугиумов и акцент был сделан на другом важном классе моделей, которые рассматривают экосистему как состоящую из многих обособленных ячеек, внутренне однородных по условиям среды и ресурсам. Такая «сотовая» структура сообществ предполагает, что количество видов и численность особей в отдельных ячейках изначально были сформированы случайным образом. В дальнейшем конкуренция в пределах ячеек может проходить согласно математическим моделям Лотки–Вольтерра, но только как частный и локальный процесс, в то время как стохастические механизмы модели представляют каждому новому поколению гидробионтов возможность свободно и случайно рассеиваться (мигрировать) между ячейками.

Процессы дифференцирования ниши перестают быть определяющим фактором регулирования сообществ с большим видовым богатством. Новые неравновесные представления подчеркивают динамический характер явлений и сосредотачиваются на значимости стохастических экологических флуктуаций (Yodzis, 1986; Townsend, 1989). Рост числа видов и обилия особей в пределах ячейки стимулирует миграцию и вызывает постоянные перестройки видовой структуры. Активный обмен видами между ячейками и степень приспособления видов-колонизаторов имеют сложный характер (см. рис. 1.3), зависящий от уровня местной конкуренции, локальных условий среды, экологических стратегий, жизненных циклов и др. Важен также и исторический фактор, связанный с заселением ячеек начальной совокупностью видов и их популяций.

Как и предусмотрено ячеечными моделями, структура реальных речных сообществ формируется одновременно под влиянием как пространственных закономерностей распределения популяций, так и динамики стохастических возмущений, обеспечивая множество вариантов, по которым изменяется вероятность появления видов и общее их разнообразие. При этом К. Таунсенд (Townsend, 1989, p. 38) определяет *возмущение* как «любое относительно дискретное событие, во время которого исчезают организмы или открывается пространство, которое может быть колонизировано особями тех же самых или иных видов».

В соответствие с концепцией *динамики пятен* (Townsend, 1989), считается, что поскольку «рефугиумы-пятна» расположены в пределах водотока случайным образом, то и видовой состав речных гидробионтов формируется в значительной мере случайно, хотя и на фоне некоторой детерминированной закономерности. В различных водотоках в зависимости от климатических, геологических, ландшафтных условий, могут наблюдаться самые различные комбинации «сгущений жизни» и «биотических лакун».

Без сомнения, что в любой случайности можно обнаружить «зашумленные» закономерности. В.В. Богатов (1995) считает, что в большинстве речных систем по мере продвижения от истоков к ритрали закономерно увеличивается количество и разнообразие таких рефугиумов и, таким образом, увеличивается вероятность появления новых групп гидробионтов. Тем самым он видит причину высокой степени видового разнообразия в среднем течении (рис. 1.2) в более многочисленных рефугиумах по сравнению, например, с зоной кренали.

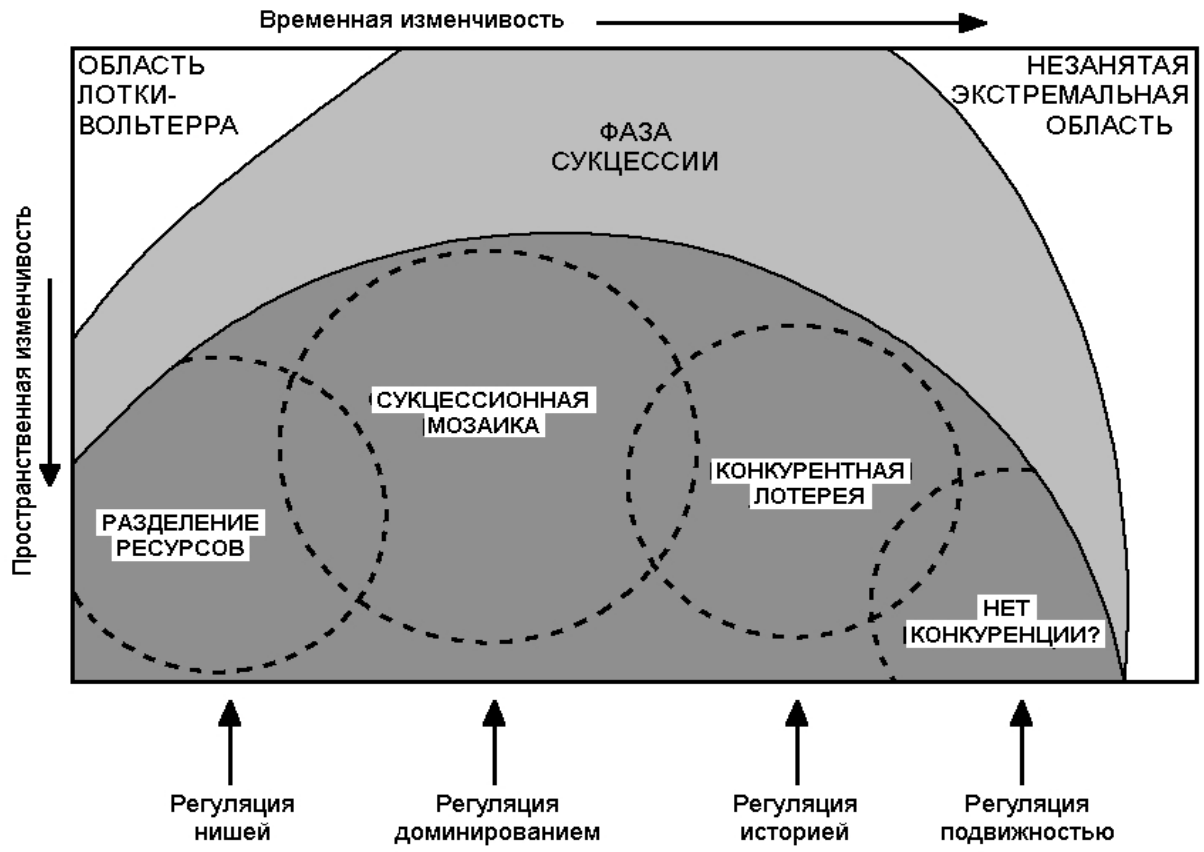


Рис. 1.3. **Влияние факторов пространственной и временной изменчивости на структурную организацию сообществ.**

Более темная штриховка соответствует большему богатству видов. Наименьшее количество видов развивается либо в гомогенной стационарной среде, где значительное большинство видов исключено в ходе конкуренции (область Лотки–Вольтерра), либо в экстремальной среде, где возмущения настолько часты, что большинство видов не выживает. Увеличение пространственной изменчивости обеспечивает появление разных микросред, дающих возможность сосуществовать большему числу видов. При увеличении временной изменчивости конкурентные отношения теряют свое значение и сообщества формируются за счет миграции организмов между ячейками, соответствующим непрерывно изменяющимся ландшафтным компонентам (Townsend, 1989).

К. Таунсенд (Townsend, 1989) определяет три метода, доступных для изучения возмущений в водотоках: 1) феноменологическая регистрация специфических явлений; 2) активный эксперимент и 3) сравнительный анализ рек, у которых есть контрастирующие режимы возмущений. Описание уникальных событий не пользуется популярностью в экологии, поскольку они производят впечатление анекдотических артефактов. Однако феноменологические исследования весьма существенны, когда познание сущности явлений находится на начальной стадии, а изучаемая экосистема обладает экстраординарными свойствами (Зинченко и др., 2010). Что касается активного эксперимента, то тип возмущения, создаваемого в моделируемом потоке, зависит во многом от изобретательности экспериментатора и варьируется от простого переворачивания камней до использования бетономешалки.

Сила сравнительного подхода состоит в возможности обобщения больших массивов данных, в результате чего анализируется весь комплекс факторов и условий, в которых происходят возмущения, и ищутся специфические закономерности в организации сообществ (Биоиндикация экологического.., 2007; Шитиков, Зинченко, 2008). К сожалению,

формальная сторона этого подхода, которая состоит в описании конкретной методики учета и сопоставления ключевых показателей, не была представлена ни Таунсендом, ни последующими исследованиями.

Научная реальность такова, что часто прекрасная теоретическая концепция несколько теряет свой блеск при столкновении с необходимостью ее подтверждения конкретным фактическим материалом. В этом плане интересна одна из немногочисленных попыток (Townsend et al., 1997) проверки гипотезы промежуточных возмущений путем обобщения достаточно репрезентативных экспериментальных данных. Ставилась задача оценить зависимость основных показателей сообществ макрозообентоса (таксономическое богатство, численность особей, разнообразие, выравненность) от интенсивности и частоты возмущений. Согласно рис. 1.3, максимум видового и популяционного богатства должен совпадать с областью умеренных возмущений, в то время как стационарная зона Лотки–Вольтерра и область экстремальных флуктуаций соответствуют бедным сообществам. Однако попытка изобразить колоколообразную кривую на множестве экспериментальных точек показалась нам весьма надуманной (см. рис. 1.4), поскольку левая ветвь параболы опирается всего на один образец – водоток 15.

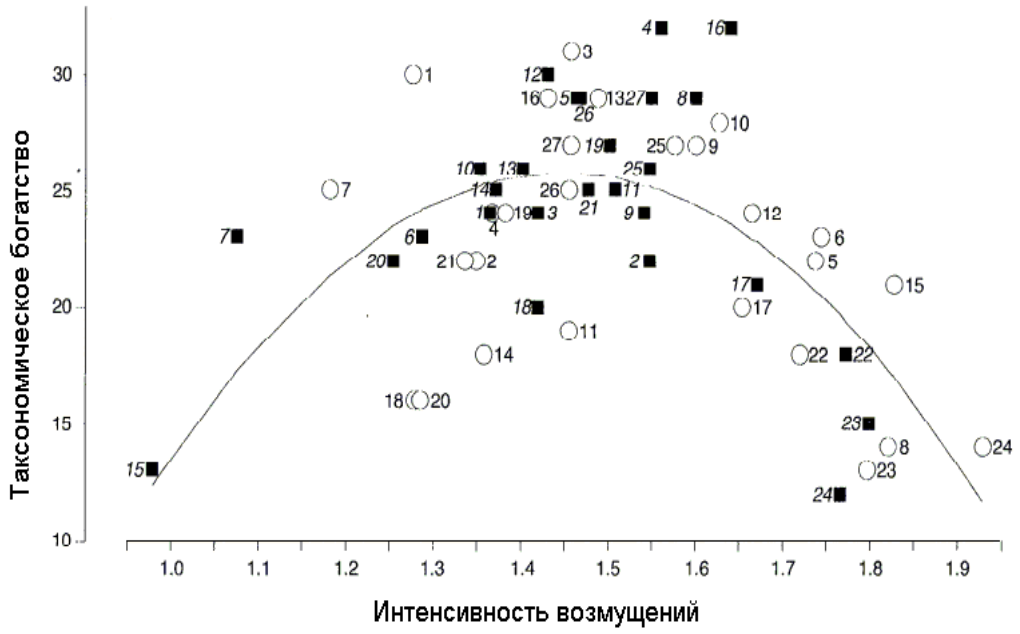


Рис. 1.4. **Зависимость числа таксонов макрозообентоса на 54 участках р. Тайери (Taieri, Новая Зеландия) от логарифма интенсивности возмущений (по: Townsend** et al., **1997)**

В своей достаточно убедительной попытке объединения концепций речного континуума и динамики пятен В.В. Богатов (1995) идет дальше. Он считает, что в крупных реках существуют разные экосистемы, совокупность которых позволяет говорить об особой, надэкосистемной форме организации природных комплексов – континууме речных экосистем, или реобиоме («...реобиом будет представляться совокупностью речных экосистем в речном континууме»; Богатов, 1995, с. 54).

Концепция *функциональных зон* (functional process zones), декларируемая Дж. Торпом с соавторами (Thorp et al., 2006), возвращает нас к тривиальной идее, что в речной экосистеме по совокупности гидрологических, ландшафтных, гидрохимических и иных факторов могут быть выделены достаточно изолированные гидрогеоморфологические области (hydrogeomorphic patches), которые и определяют видовой и функциональный состав сообществ гидробионтов. Эта теория опирается на данные исследования (Thoms, Parsons,

2003), где проводилось пространственное зонирование рек вдоль продольных градиентов с использованием 230 гидрологических и геоморфологических параметров, и является развитием сходных взглядов Д. Монтгомери (Montgomery, 1999), обозначившего в водотоках «процессные области» (process domains). Круг «колодок» замкнулся и, плавно пройдя через обе контрастирующие идеи (континуальности и дискретности), научная мысль вернулась к исходной точке – выделению изолированных относительно гомогенных зон, последовательно расположенных вдоль продольного профиля русла реки (Hawkes, 1975).

Концепции речного континуума и динамики пятен не противоречат существующим представлениям о таком трудно формализуемом явлении, как *сукцессия*, определяемом как «постепенные необратимые... изменения состава и структуры сообщества, вызываемые внутренними (автогенная сукцессия) или внешними (аллогенная сукцессия) причинами» (Миркин и др., 1989, с. 178). Неоднозначность трактовки экологами *постепенности*, *обратимости* и характера происходящих *изменений* в ходе сукцессий сообществ подробно проанализирована Г.С. Розенбергом с соавторами (1999).

Продольная сукцессия получила достаточно вольное определение (Fisher et al., 1983) как стационарная (т. е. не зависящая от времени) последовательность смены сообществ в проточных водах от истоков до устья. Такое расширенное толкование сукцессии как последовательности сообществ, где каждая составляющая испытывает влияние своих предшественников, здесь достаточно обоснована, если учесть, что проточные воды являются единственными экосистемами, в которых причинность в достаточной степени ориентирована в пространстве.

Продольные сукцессионные смены определяются, с одной стороны, естественными ландшафтно-гидрологическими характеристиками водотока, т. е. общим биоценотическим градиентом от истока к устью, в результате чего наблюдается детерминированная последовательность отличающихся по видовому составу, но функционально достаточно близких экоморф. С другой стороны, в условиях динамики стохастических возмущений естественного и антропогенного характера определяющими становятся процессы, когда локальные смены гидроценозов происходят благодаря периодическому или постоянному воздействию какого-либо внешнего фактора и целиком определяются его интенсивностью. Если внешний фактор оказывает постоянное воздействие слабой интенсивности, то может происходить *реверсионная* последовательность смен ассоциаций видов, которая противоположна той, что наблюдается при нормальном течении экогенеза. Если же возмущение коренным образом изменяет условия среды, возникает необратимая *трансверзальная* смена или, другими словами, полная смена одного экогенетического ряда на другой. При этом К. Таунсенд различает (рис. 1.3) *фазовую сукцессию*, отмечаемую в пространственно стационарных водоемах, и *сукцессионную мозаику* при появлении рефугиумов.

Л. Бенда с соавторами (Benda et al., 2004) обращает внимание на важную проблему при изучении малых и средних рек – пространственную и гидрологическую неоднородность, обусловленную типом организации сети водотоков. В местах слияния рек происходят резкие перепады удельного расхода воды и наносные явления, которые вызывают соответствующие изменения в морфологии поймы и русла. Иерархический и ветвящийся характер речных сетей сочетается с динамическими возмущениями, описанными Таунсендом, и в значительной мере определяет неоднородность распределения природных условий приречной местности. Это вызывает адекватные изменения биологического разнообразия, функциональных особенностей и видовой структуры сообществ гидробионтов. Немаловажным фактором является также «перемешивание» двух сообществ в зоне слияния рек.

Гипотеза *динамики сети* (network dynamics hypothesis; Benda et al., 2004) определяет развитие общих представлений и конкретных моделей, оценивающих факторы геометрии речной сети, рассматриваемой в контексте совокупности водотоков, взаимодействующих в местах их слияния. При этом показано (рис. 1.5), что при разветвленном (сердце-

видном и грушевидном) характере конфигурации сети притоков, последние аккумулируют большую поверхность водосбора и в точке слияния с главным руслом оказывают значительное влияние по сравнению с линейной регулярной структурой.

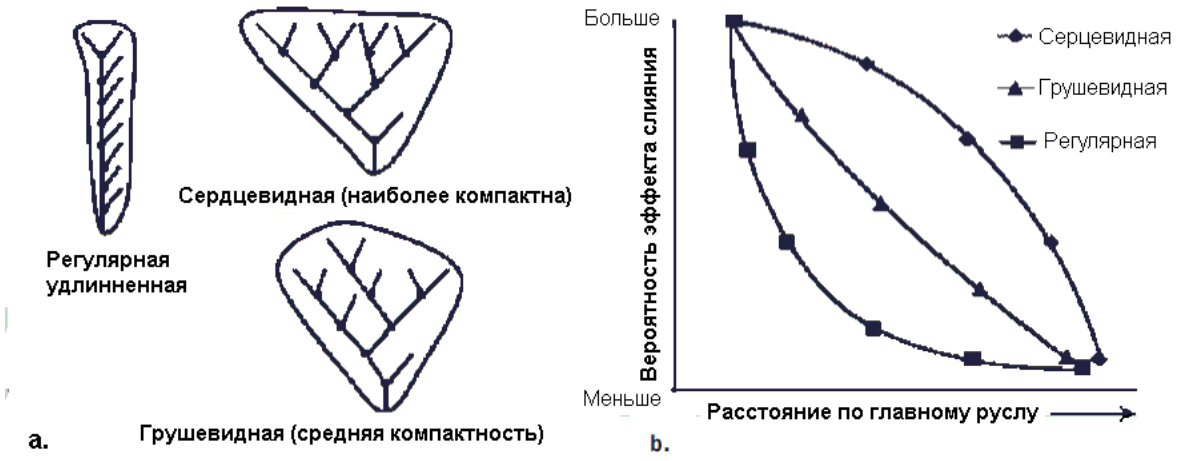


Рис. 1.5. **Различные схемы конфигурации сети водотоков (a) и предполагаемая динамика развития эффекта слияния водотоков по главному руслу реки (b) (по: Benda et al., 2004)**

Дальнейшая апология здравого смысла в построении концепций функционирования речных экосистем основывается на признании (Thorp et al., 2006), что все зависит от масштаба изучаемых явлений. Х. Фриссел с соавторами (Frissell et al., 1986) предложили (рис. 1.6) географически интегрированную иерархическую структуру, состоящую из совокупности вложенных друг в друга пространственных и временных масштабных уровней. Такой иерархический подход предусматривает: 1) интегрирование данных из различных источников, 2) выбор различной степени детализации данных, необходимой для работы на каждом уровне, и 3) активный обмен данными и их агрегирование при переходе от нижних уровней к высшим.

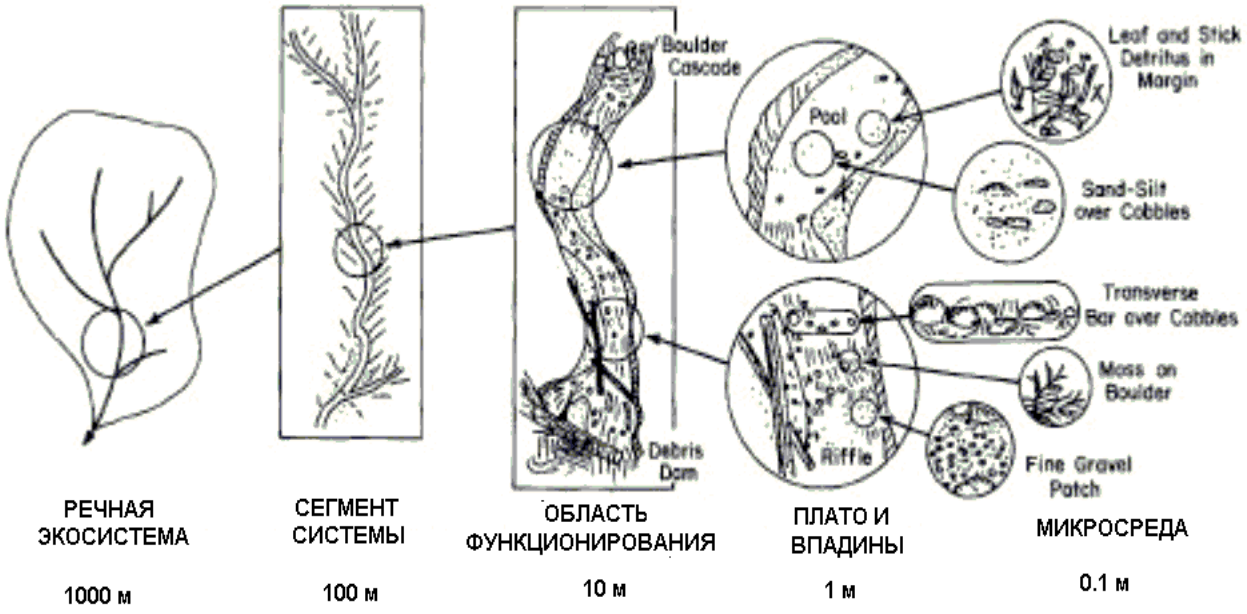


Рис. 1.6. **Пространственно вложенная иерархическая организация речных экосистем и их структурных компонентов (по: Frissell et al., 1986)**

В принципе, каждая выделенная функциональная зона водотока обладает как внешней, так и внутренней биотической изменчивостью, компоненты которой распределяются по уровням пространственно вложенной масштабной иерархии естественной среды. Но, в конечном счете, изменчивость между функциональными зонами реки должна быть больше, чем изменчивость во внутренних пределах каждой отдельно взятой зоны.

Попытка количественно оценить компоненты этой изменчивости будет одной из центральных тем нашего последующего изложения. Кроме того, мы собираемся установить, регулируется ли структура наших модельных сообществ детерминированными взаимодействиями между видами, такими как конкуренция и хищничество, или она определяется стохастическими факторами, такими как пространственные и временные возмущения.

1.3. Донные сообщества лотических систем как объект и материал исследования

Методы и модели анализа видовой структуры пресноводных сообществ рассматриваются нами с использованием данных гидробиологических наблюдений на равнинных реках бассейна Нижней Волги (Зинченко, Шитиков, 1999), протекающих в степной и лесостепной зонах с различными ландшафтно-географическими характеристиками. Мониторинговые наблюдения проводились начиная с 1985 г. по настоящее время на 263 станциях 43 малых и средних рек, часть которых представлена на картосхеме (рис. 1.7а).

В качестве модельного водотока была выбрана речная система Байтуган–Сок (рис. 1.7б), достаточно протяженная и включающая в себя значительное разнообразие биотопических и ландшафтных элементов. Территорию бассейна можно условно разделить на предгорную и равнинную части с типичными физико-географическими характеристиками.

Самый высокий участок находится на границе Самарской и Оренбургской областей, куда заходят отроги Бугульминско-Белебеевской возвышенности. Здесь берет начало река Сок и её основные притоки: Байтуган, Камышла, Б. Черемшан, Кондурча и другие. Основная равнинная часть водосбора расположена в приподнятой части Высокого Заволжья и находится в двух природных зонах – лесостепной и степной.

Река Сок – типичная средняя не зарегулированная равнинная река первого порядка, впадающая в Саратовское водохранилище на северной окраине г. Самара. Протекает по широкой долине с возвышенным правым берегом на территории Сокских гор. Среднее и нижнее течение р. Сок находится в зоне умеренной сельскохозяйственной и техногенной нагрузки. В целом по уровню загрязнения водных масс относится к II-IV классу качества. Основные гидрологические и гидрохимические показатели, а также итоговые результаты гидробиологической съемки представлены в табл. 1.1.

Река Байтуган является характерным предгорным водотоком, имеющим родниковое питание и небольшую протяженность (22 км) при среднем уклоне около 0,9%. Скорость течения в летнюю межень в верховьях реки изменяется от 1,2 м/сек (на перекатах) до 0,5 м/сек (на плесах), снижаясь в низовьях до 0,3-0,2 м/сек. Для нее характерны прозрачность воды до дна, высокое содержание растворенного кислорода, низкое содержание биогенных веществ (Зинченко, 2002). По качеству вод р. Байтуган относится к эталонным водотокам региона.

Нам представилось интересным выполнить анализ видовой структуры сообществ гидробионтов в нескольких вариантах. Сначала р. Сок рассматривалась как единая целостная экосистема, но в дальнейшем (в главах 2-5) мы, на основании анализа гидрологических особенностей, выделили два отдельных однородных участка: верхнюю часть до с. Исаклы и область ниже по течению до устья. В некоторых примерах главы 2 представленные водотоки сравниваются с аналогичными по типу реками Б. Кинель и Кондурча, а также с р. Чапаевка в нижнем её течении, где отмечен высокий уровень загрязнения вод. В главах 6-7

масштабная детализация была углублена и речная система Байтуган–Сок делилась на 13 небольших участков в зоне базовых станций наблюдения.

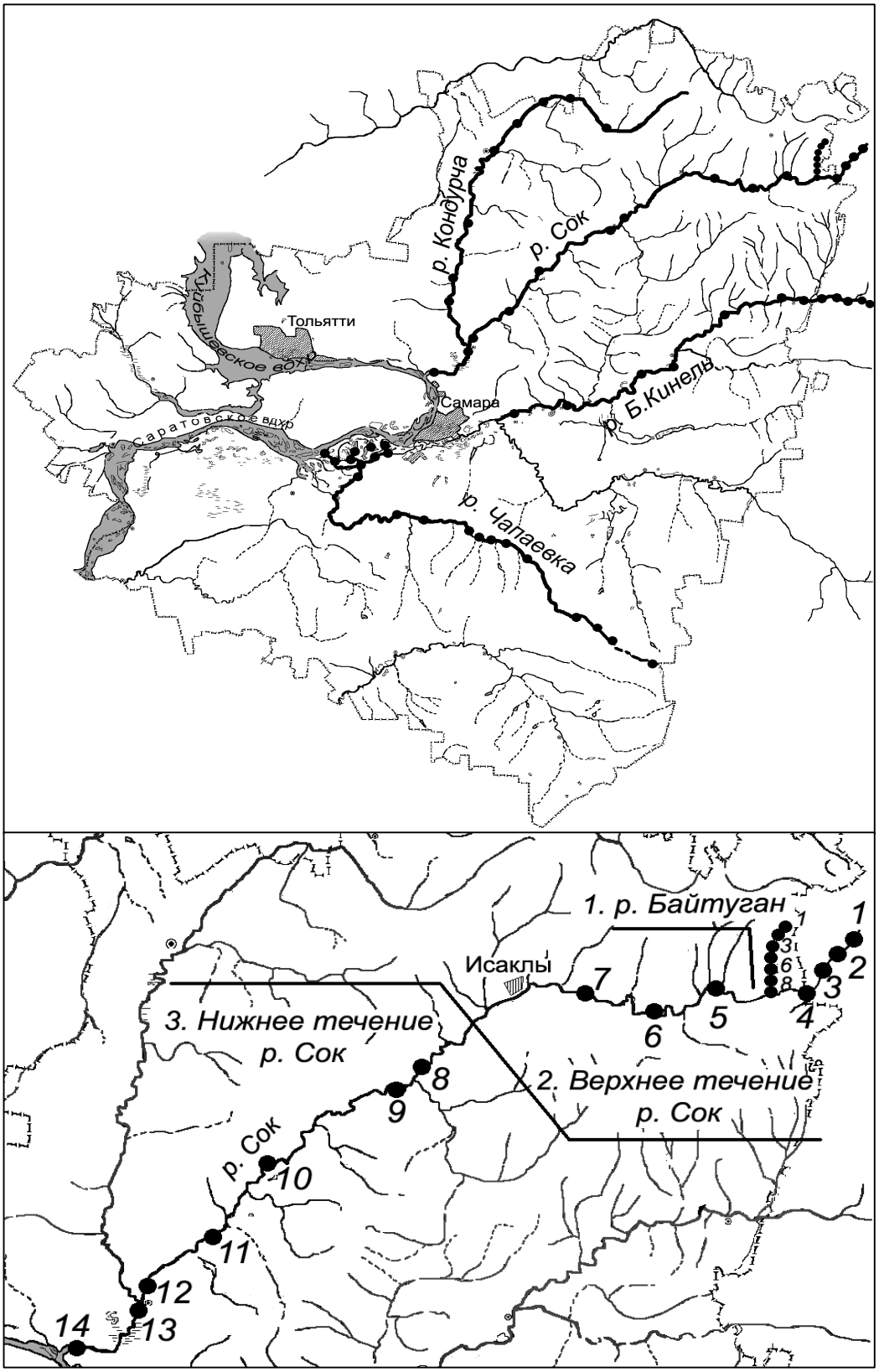


Рис.1.7. **Картосхемы малых рек: *а* - верхняя часть бассейна Саратовского водохранилища; *б* - участки речной системы Байтуган–Сок**

Таблица 1.1

Основные показатели по участкам изучаемых водоемов

Показатели	Река Сок			Река Байтуган
	В целом	Верхнее течение	Нижнее течение	
Гидрологические характеристики				
Длина, км	375			22
Площадь водосбора, тыс. км^2	11,87			0,14
Средний уклон, %	0,07			0,9
Скорость течения, м/сек	0,48	0,63	0,28	0,53
Качество воды		чистая – умеренно загрязненная	умеренно загрязненная – загрязненная	чистая
Класс качества		II-III	III-IV	II
Индекс ИЗВ		0,9-1,8	1,9-4,0	0,65-0,9
Определяющие компоненты (в скобках – максимально зарегистрированное превышение ПДК)		органические вещества	азот NH_4^+ (11,5); фенолы (3); нефтепродукты (2); медь (7)	органические вещества
Гидробиологическая съемка				
Число станций наблюдения	14	7	7	8
Количество проб	99	53	46	31
Всего видов и таксонов бентоса	281	206	176	144

В качестве объектов биологического мониторинга использовались сообщества макрозообентоса, которые обладают признанными биоиндикационными свойствами и являются наиболее стабильным и разнообразным компонентом лотических экосистем. Отбор проб бентоса и их последующая обработка осуществлялись с использованием общепринятых методов гидробиологических исследований, подробно изложенных нами ранее (Зинченко, 2002). Популяционная плотность донных организмов оценивалась в экз./м^2.

Детальное описание видовой структуры сообществ донных беспозвоночных организмов малых рек Нижнего Поволжья представлено в серии наших предыдущих публикациях (Экологическое состояние.., 1997; Зинченко, Головатюк, 2000; Зинченко, 2002; Биоиндикация экологического.., 2007). Поэтому в дальнейшем изложении мы делаем акцент на статистические свойства выборок, полученных на основе гидробиологических проб, а обзор фаунистического состава макрозообентоса рассматриваемых водотоков ограничим списком таксономических групп в табл. 1.2.

Структура сообществ макрозообентоса определяется преимущественно комплексом редких видов, и практически нельзя выделить какие-либо таксоны, которые могли бы быть характеризованы как базовые или доминирующие. Из 334 видов, встретившихся в 130 пробах, 194 вида (или 58%) было обнаружено только в одной или двух пробах (см. гистограмму на рис. 1.8). И лишь 22 вида встретились более чем в 10% проб (максимальная встречаемость для *Tanytarsus* sp. составила 54%).

Характер детальных таблиц гидробиологических данных, содержащих результаты камеральной обработки отдельных проб макрозообентоса, предопределяет значительные трудности статистического анализа, игнорирование которых может, в конечном счете, сильно повлиять на корректность итоговых выводов и заключений. Биогеографические матрицы, включающие сотни видов, заполнены измеренными показателями менее чем на 10%, а статистическое распределение выборочных значений численности особей имеет яр-

ко выраженный негауссовый характер (Шитиков и др., 2005). В этих условиях использование даже простейших традиционных статистик для оценки математического ожидания и дисперсии обилия видов приводит к абсурдным результатам.

Таблица 1.2

Таксономические группы макрозообентоса и их встречаемость в речной экосистеме Байтуган–Сок (T_S – средняя встречаемость видов группы в одной гидробиологической пробе, T_N – средняя популяционная плотность особей группы на 1 м^2)

№ пп	Таксономическая группа	Число видов	T_S	T_N	№ пп	Таксономическая группа	Число видов	T_S	T_N
1	Chironomidae	157	7,33	2702	20	Tipulidae	2	0,02	0,15
2	Bivalvia	22	0,67	83,63	21	Ceratopogonidae	2	0,36	42,70
3	Trichoptera	19	0,39	27,48	22	Stratiomyidae	2	0,03	0,96
4	Ephemeroptera	19	0,83	169	23	Dolichopodidae	2	0,02	2,62
5	Coleoptera	15	0,35	15,94	24	Ephydridae	2	0,02	0,54
6	Oligochaeta	10	0,77	769	25	Psychodidae	2	0,10	3,95
7	Hirudinea	9	0,17	26,16	26	Nematocera	2	0,02	0,38
8	Hemiptera	8	0,18	23,25	27	Dixidae	2	0,04	1,23
9	Unionidae	8	0,15	11,09	28	Hydrida	1	0,03	1,50
10	Gastropoda	6	0,05	3,54	29	Diplura	1	0,02	0,12
11	Plecoptera	6	0,11	4,53	30	Ptychopteridae	1	0,01	0,31
12	Odonata	5	0,08	2,86	31	Diptera (прочие)	1	0,05	1,19
13	Hidracarina	5	0,18	8,38	32	Cylindrotomidae	1	0,02	0,42
14	Crustacea	4	0,05	32,02	33	Megaloptera	1	0,03	0,62
15	Limoniidae	4	0,35	35,18	34	Muscidae	1	0,01	0,31
16	Tabanidae	3	0,05	1,15	35	Nematoda	1	0,09	11,54
17	Culicidae	3	0,02	0,24	36	Polychaeta	1	0,01	0,15
18	Amphipoda	3	0,04	2,31	37	Arachnoidae	1	0,04	0,88
19	Simuliidae	2	0,29	170	ИТОГО		**334**	**13**	**4157**

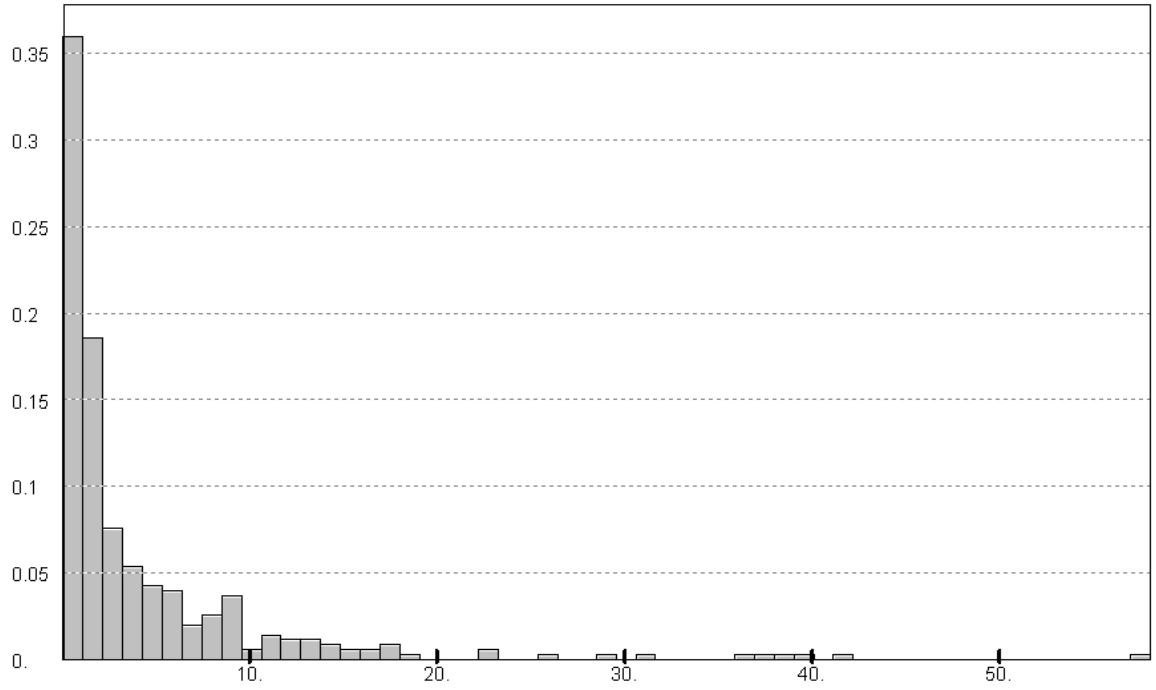


Рис. 1.8. **Гистограмма частот встречаемости 334 видов макрозообентоса в 130 пробах гидробиологической съемки речной системы Байтуган–Сок, 1990-2007 гг.**

Ось абсцисс – градации числа проб, в которых встретился тот или иной вид (т. е. в 1, 2, 3 и т. д. до 70); ось ординат – доля видов от их общего числа, принадлежащих к каждой из этих градаций

Рассмотрим в качестве примера произвольный и достаточно типичный фрагмент базы данных, включающий восемь проб, отобранных на ст. 9, и две пробы на ст. 10 р. Сок (табл. 1.3). Сразу можно отметить, что медиана и мода численности *Chironomus obtusidens* (ChChi_o) на ст. 9 равны нулю, а среднее арифметическое (520+20)/8 = 67 экз./м^2 не является интерпретируемой величиной из-за сильной асимметрии распределения.

Таблица 1.3

Фрагмент базы данных гидробиологической съемки зообентоса р. Сок на станциях 9 (7,5 км ниже пос. Сургут) и 10 (у с. Б. Чесноковка)

Коды таксонов в базе данных	Станция 9 (месяц и год отбора проб)								Станция 10	
	VI 1995	VII 1995	VII 1995	IX 1995	IX 1995	VII 1998	VII 1999	VI 2007	VII 1998	VII 1998
…										
BiAms_so							40			
BiPis_a		40	10	40	3160					
BiPis_i									40	
Ce								300		
CeCul_sp	10		10		40	140				10
ChAbl_m			10							
ChAbl_ph						100				
ChChi_ms			10							
ChChi_n										
ChChi_o					520	20				
ChChi_pi					1120					
ChCld_m			30			720				
ChCld_sp										490
…										
Число видов	8	5	17	18	20	35	6	19	8	16
	70								21	
Общая численность (экз./м^2)	150	560	390	290	6280	7650	1280	3900	2080	1440
Индекс Шеннона	2,82	1,81	3,74	3,88	2,58	4,04	2,02	3,60	0,95	2,73

Широко распространенной ошибкой анализа гидробиологических данных является используемый способ расчета обобщенных показателей путем механического осреднения проб без коррекции на относительное выборочное усилие. Например, представляется сомнительным сравнение индексов видового разнообразие по Шеннону[4], рассчитанных путем суммирования численностей видов на каждой станции по всем выполненным пробам. Что может означать в табл. 1.3 увеличение числа видов (70 против 21) или «объединенного» индекса Шеннона H (3,06 против 1,84) на ст. 9 по сравнению с ст. 10? Объясняется ли этот вывод экологическими причинами, вызванными, например, миграцией дополнительных видов с водными массами левого притока р. Сургут у ст. 9, или определяется статистическим эффектом того, что на ст. 10 просто было взято существенно меньше проб?

Представленный характер распределения численностей видов макрозообентоса в гидробиологических пробах делает мало пригодным использование при обработке данных мониторинга любых параметрических статистических методов, основанных на гауссовой

[4] Подчеркнем, что данный показатель, который в разных работах называется по-разному, правильно следует называть именно «индекс Шеннона» (Розенберг, 2010).

парадигме распределения вариационных рядов. Не только традиционные показатели, такие как средние значения численностей, дисперсии, медианы, но и все виды многомерного статистического анализа, использующие эти величины, являются принципиально некорректными. Например, серьезного сомнения заслуживает методика (Webb et al., 2009) оценки популяционной плотности каждого вида на участке как средняя геометрическая численностей во всех выборках, где этот вид был зарегистрирован (т. е. исключая обширные подмножества нуль-данных).

Отступление от темы (*проблемно-статистическое*). При обработке и интерпретации данных экологического мониторинга приходится преодолевать ряд стереотипов, которые выработались у исследователей, успешно применявших гауссовую математическую статистику в ряде отраслей естественных наук и неоправданно переносящих их на негауссовые явления (Хайтун, 1983).

Первый стереотип – «хвосты» распределений несущественны и все, что выходит за пределы трех сигм, можно отбросить без информационных потерь. Для нормального распределения и других гауссовых распределений эта практика себя оправдывает, поскольку там содержится пренебрежимо малая часть выборки. Для негауссовых распределений этот стереотип не срабатывает, так как их хвосты, медленно убывающие с ростом значений переменных, содержат существенную часть выборки, поэтому их отбрасывание ведет к значительным ошибкам.

Второй стереотип – близость медианы к среднему арифметическому. Когда речь идет о простых внутрипопуляционных показателях, имеющих нормальное распределение (например, индивидуальный вес особей), то это недалеко от истины. Однако при переходе к экосистемным показателям, учитывающим видовую структуру, необходимо принимать во внимание большую асимметрию негауссовых распределений, например, 80% числа всех особей сообщества принадлежит к 20% видов.

Третий стереотип – независимость полученных результатов от объема выборки. При вычислении среднего, дисперсии и других моментов распределения считается, что они незначительно изменяются по случайному закону при добавлении новых порций данных, поэтому мало кто принимает во внимание объем эмпирической совокупности, который фигурирует только при оценке ошибки выборочного значения оцениваемой величины. При работе с негауссовыми распределениями этот стереотип подводит самым серьезным образом, потому что на таких генеральных совокупностях моменты существенно и направленно растут с объемом выборки. При сравнении параметров разных экосистем объемы выборок, как правило, существенно отличаются, вследствие чего результаты не только не показывают должной воспроизводимости, но и вносится систематическая ошибка.

Распределения негауссового типа характерны, прежде всего, для сложных систем[5] с большим видовым разнообразием, которые рассматриваются как статистические ансамбли, состоящие из множества разнотипных компонентов. Работая с такими выборками, целесообразно обратиться к новой статистической идеологии, принимающей во внимание при анализе всю генеральную совокупность. Практическая важность этой концепции в том, что в случае неустойчивых частот почти все устойчивые плотности распределений не выражаются в элементарных функциях (через обычные параметрические формулы).

Одной из задач нашего последующего изложения является анализ работоспособности новых имитационных методов (на примере обработки данных мониторинга сообществ речного макрозообентоса) при решении следующих задач:

[5] Похожая ситуация – в социологии и других науках; например, неверно было бы полагать, что средний доход делит людей пополам – 50% имеют зарплату выше средней, а 50% – ниже.

- оценка полного видового богатства сообществ методом экстраполяции кумулятивных кривых накопления видов макрозообентоса с использованием *конечного* числа гидробиологических проб;
- сравнение параметров биоразнообразия сообществ с учетом выборочного усилия на основе непараметрических моделей разрежения видов;
- оценка сходства видовой структуры участков водотока на основе вероятностных мер подобия, учитывающих частотное распределение видов по гидробиологическим пробам;
- декомпозиция видового разнообразия на α- и β-компоненты с использованием аддитивной иерархической схемы масштабного разделения;
- построение функций рангового распределения обилия видов на основе различных моделей пропорционального или случайного распределения ресурсов;
- проверка гипотез о способах структурной организации композиций видов: являются ли они следствием межвидовых взаимодействий (например, конкурентного исключения) или их формирование произошло случайно под действием стохастических факторов;
- выделение комплексов видов (ценотических единиц и их ассоциаций), характерных для каждого местообитания.

Методика сравнительного анализа сообществ макрозообентоса основывалась на следующих основных принципах:

- отказ от многих традиционных приемов «объединения» или «усреднения» серий гидробиологических проб для оценки видового богатства, биоразнообразия и таксономического сходства изучаемых водотоков: ключевым анализируемым показателем явилась матрица частот встречаемости i-го вида $p_{ik} = n_{ik} / n_k$, где n_k – общее число гидробиологических проб, взятых на на k-ом участке водотока, $k = 1, 2, \ldots, m$; n_{ik} – число проб из n_k, в которых были обнаружены особи i-го вида, $i = 1, 2, \ldots, s$;
- оценка основных показателей видовой структуры сопровождалась расчетом доверительных интервалов, что позволило выполнять проверку достоверности различий сравниваемых параметров;
- вычисление оценок статистических параметров распределения величин проводилось с использованием процедур имитационного моделирования Монте-Карло и непараметрических алгоритмов генерации псевдовыборок (bootstrap, jackknife, permutation – Эфрон, 1988).

Сравнительный анализ сообществ может быть осуществлен на разных уровнях с использованием различных комплексов показателей или подмножеств данных (табл. 1.4). Методы оценки величины и статистической значимости рассчитываемых параметров, а также способы их интерпретации обсуждаются нами в дальнейшем по мере последовательного усложнения уровня описания данных: в главах 2 и 3 – отдельные количественные показатели и индексы (низкий уровень в табл. 1.4), в главе 5 – векторы (средний уровень) и в главах 4, 6 и 7 – матрицы (высокий уровень).

Основная концепция сравнительного анализа экосистем малых и средних водотоков часто основывается на предположении, что каждая взятая проба является случайной выборкой из одной и той же генеральной совокупности, что соответствует моделям локального α-разнообразия Р. Уиттекера (Whittaker, 1972). Иными словами, предполагается, что видовой состав всех проб принадлежит одному и тому же локальному сообществу, а имеющиеся флуктуации встречаемости видов обусловлены временной динамикой или пространственными координатами, в пределах которых выполнялась гидробиологическая съемка. Но, имея дело с разветвленной сетью водотоков, мы неизбежно сталкиваемся с неопределенностью: «Как оценить степень автономности двух смежных сообществ и где провести пространственную границу между ними?». Не следует ли, например, всю р. Сок вместе с её притоками р. Кондурча и р. Байтуган трактовать как единую экосистему? Или

наоборот, следует в нижнем течении р. Сок выделить несколько самостоятельных зон, учитывая специфику устьевого участка в зоне подпора водами водохранилища?

Таблица 1.4

Три способа описания структуры сообществ по увеличению степени точности или уменьшению простоты интерпретации (McGill et al., 2007)

Уровень	Характер данных	Достоинства	Недостатки
Низкий	Одномерные показатели (число видов S и особей N, индексы разнообразия).	1. Можно сравнивать сообщества, у которых есть немного (или нет совсем) общих видов. 2. Легкое понимание и однозначное толкование.	1. Высокая потеря информации. 2. Результат зависит от объема выборки, который не учитывается при расчете показателей.
Средний	Вектор – распределение численности видов (отсортированный вектор чисел с отсутствующими нулевыми значениями и метками принадлежности к видам); показывается двухмерным графиком или гистограммой.	1. Можно сравнивать сообщества, у которых есть немного (или нет совсем) общих видов. 2. Легкое понимание и возможность визуального сравнения. 3. Показывает относительное соотношение редких, промежуточных и доминантных видов. 4. Информация об объеме выборки включена в состав данных.	1. Умеренная потеря информации. 2. Сравнение двух сообществ, не столь легко, как сравнение одномерных показателей.
Высокий	Матрица с видами в строках, местообитаниями в столбцах и численностями в ячейках; сохранены принадлежность к видам и нулевая численность.	1. Полное использование информации. 2. Способность проследить за изменениями в численности особей и видовой структуре. 3. Доступность для обработки многомерными методами, например, для ординации.	1. Трудно сравнить сообщества, у которых есть немного (или нет совсем) общих видов. 2. Результат представлен как таблица чисел или как результат многомерной обработки (например, классификации), поэтому трудно интерпретируем.

Разделение сети водотоков на оптимально очерченные экологические зоны с учетом особенностей видовой структуры сообществ обычно связывают с формализмом построения компактных классификаций. Однако остается неясным, какие критерии оптимизации при этом использовать: индексы сходства, параметры β-разнообразия, изменение доминирующих видов или что-то иное. В главе 7 мы обсуждаем методы разрешения этой неопределенности.

1.4. Методы макроэкологии и анализ видовой структуры сообществ

Экология долгое время выступала как средство накопления фактографических данных о биосферных компонентах, полуколичественный анализ которых давал пищу для предположений, гипотез и теорий разной степени достоверности о том, как могут «работать природные механизмы». Выделение объективных закономерностей серьезно ослож-

нялось многопараметричностью экологических объектов, сложностью внутрисистемных связей, большой изменчивостью и динамичностью наблюдаемых показателей. Идентификация большинства экологических процессов при этом возможна только при сопоставлении с другими природными аналогами или в широком временном масштабе, что делает невозможным их экспериментальное исследование. Однако за последние два десятилетия развитие экологии стремительно проходило в направлении создания новой методологии изучения крупномасштабных объектов, широкого внедрения количественных методов и исследования процессов, которые носят универсальный характер. Среди таких общих закономерностей можно назвать широтный градиент разнообразия, зависимость количества видов от размера территории, а также связи между размерами тела, численностью популяции и площадью области распространения.

Анализ видовой структуры экологических сообществ заключается в поиске ответа на следующие ключевые вопросы:

- Как много видов объективно присутствует в определенном местообитании и сколько из них мы можем обнаружить, выполнив m измерений (проб)?
- Как видовой состав меняется во времени, и каков верхний предел числа видов, который обеспечивается доступными ресурсами?
- Каковы основные структурные особенности изучаемого сообщества (характер доминирования, доля и роль редких видов, наиболее подходящая модель рангового распределения, трофические взаимосвязи)?
- Какова взаимосвязь между видовым составом и моделями функционирования экосистем (продукцией, деструкцией, составляющими энергетического баланса)?
- Какие факторы определяют признаки совместного существования видов?
- Является ли сообщество устойчивым и стабильным?
- Каковы причины миграции, вымирания или образования популяций видов?

При попытке решить эти проблемы углубилось естественное понимание того, что:

1) таксономическая структура водных сообществ не является строго зафиксированной композицией видов и, в зависимости от свойств области исследования, характеризуется различной степенью пространственной и временной изменчивости;
2) основные структурные параметры сообществ (видовое богатство, разнообразие, частотные распределения численности особей и др.) и показатели сходства видовых ассоциаций между собой коренным образом зависят от масштабного фактора локализации экосистем и величины выборочного усилия (т. е. количества проб, взятых при проведении мониторинговых исследований).

Что же определяет специфику комбинаций видов, представленных в биологических сообществах, число этих видов и их относительное обилие? Эти вопросы представляют многолетний интерес для экологов, которые часто видят в замеченных флуктуациях результат межвидовой конкуренции, влияние хищничества, климатических условий, обеспеченность пищей, возможность случайного рассеивания или перемещения особей и т. д. Один из способов проверить эти гипотезы о роли различных процессов состоит в том, чтобы найти статистические закономерности (т. е. общие черты или различия) при сопоставлении отдельных сообществ, находящихся в разных условиях среды. Они могут отражать: 1) определенную регулярность в относительном обилии видов, представленных в сообществе; 2) пространственный градиент числа видов; 3) повторяемость некоторых специфических таксономических композиций и 4) различия в поведении или функциях некоторых видов, которые можно объяснить, например, конкуренцией или мутуализмом. Эти зависимости обычно представляются в форме графо-аналитических образов (или «паттернов» [patterns]), примерами которых могут служить многочисленные модели биоразнообразия (*геометрические ряды Мотомуры*, *разломанного стержня Мак-Артура*, *экспоненциально разломанного стержня Фёдорова*, *логарифмические серии Фишера* и т. д.; см.: [Шитиков, Розенберг, 2005]), результаты градиентного анализа – различного рода эко-, хроно-, ценок-

лины и топо(эко-, хроно-, цено)клины (Миркин, 1971; Миркин и др., 1975; Миркин, Розенберг, 1978а), развернутые графы биотических взаимодействий или направлений воздействия абиотических факторов и многое другое.

Принимая во внимание мультипараметрическую природу процессов, лежащих в основе генерации и развития множества компонентов биосферы, *макроэкология* позиционируется как наука, которая стремится построить объяснение этой сложности, изучая статистические выборки больших множеств популяций видов, распределенных в географическом пространстве и экологическом времени. Однако если рассматривать иерархию масштабных уровней организации экосистем, представленную на рис. 1.9, то область интересов макроэкологии, вопреки мнениям о ее глобальном характере, простирается в пространственной шкале от 1000 м^2 (что соответствует протяженности участков малых рек около 200 м) до континентальных размеров.



Рис. 1.9. **Масштабы в экологии: характер процессов при различных шкалах разрешения**

Макроэкология (Brown, 1995; Gaston, Blackburn, 2000; Blackburn, Gaston, 2003), оперируя экологическими «паттернами» и моделями процессов в различных масштабах разрешения пространственно распределенных биосферных компонент, стремится *идентифицировать и объяснить региональную изменчивость видового разнообразия и динамику развития популяций*. При этом, сохраняя свою специфику в изучении сообществ, она находится в тесной взаимосвязи с другими областями экологии (рис. 1.10).

Справедливости ради стоит отметить, что некоторые исследователи (McGill, 2003a) сводят макроэкологию лишь к инструменту исследования формы трех кривых:

- функции зависимости числа видов от площади (SAR или SPAR – Species Area Relationship), которой чаще всего приписывают степенной характер;
- кривой распределения обилия видов (SAD – Species Abundances Distribution);
- функции распределения размеров тела животных (BSD – Body Size Distribution), имеющей, например, смещенную логнормальную закономерность.

Действительно, большое количество публикаций конца прошлого века было посвящено малодоказательному обсуждению «правильной» формы этих частных кривых, что выглядело, как минимум, контрпродуктивно (McGill, 2003a).

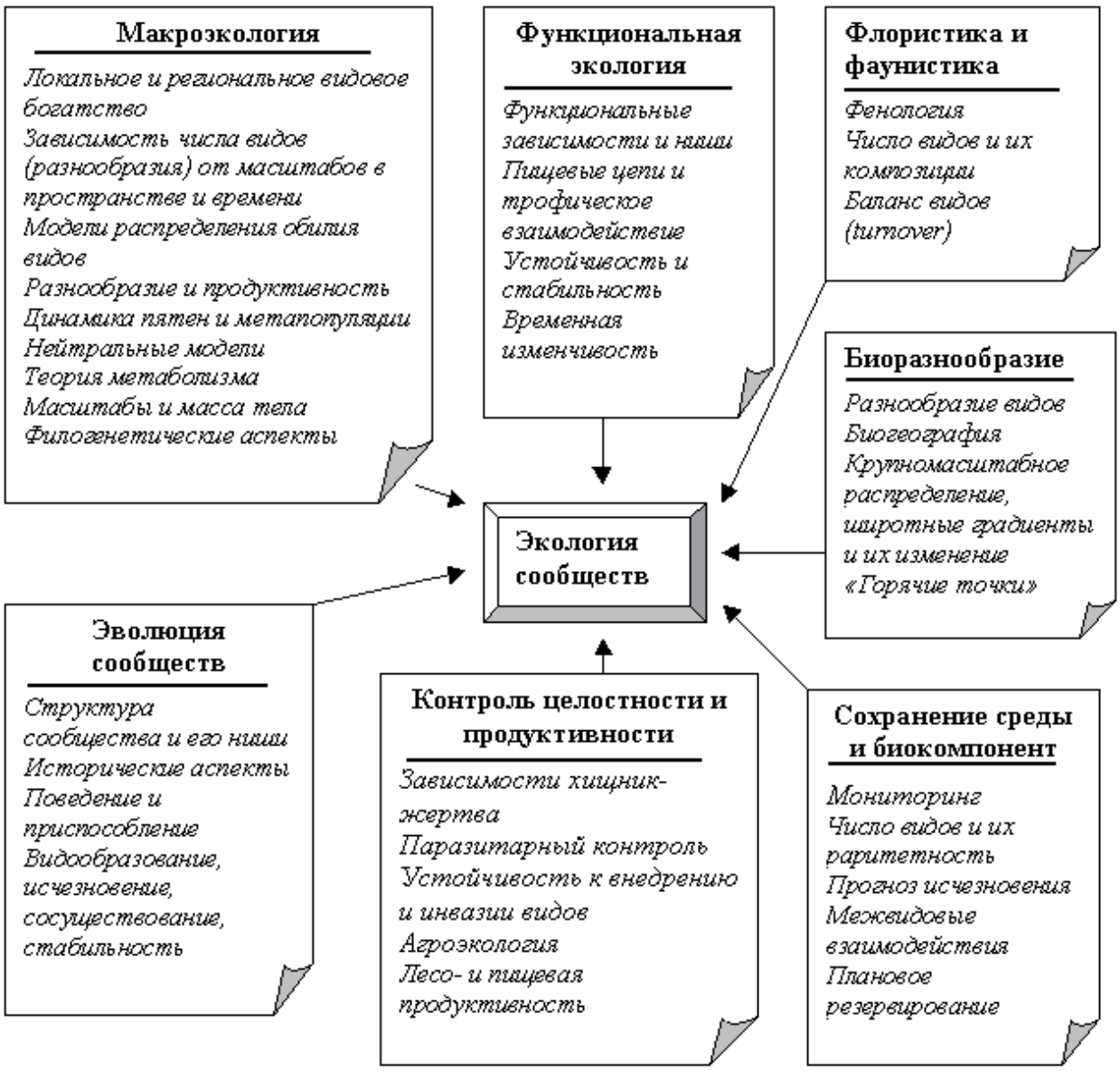


Рис. 1.10. **Основные модели («паттерны») и процессы, используемые при изучении экологии сообществ**

Многочисленные обобщения показали, что большинство экологических переменных с той или иной степенью погрешности связаны между собой простыми аллометрическими закономерностями (Enquist et al., 2002):

Зависимость от обследованной площади A	Зависимость от числа видов S
$B \propto A^z$	$B \propto S^z$
$D \propto A^z$	$D \propto S^z$
$S \propto A^z$	$V \propto S^z$
Зависимость от продуктивности P	Зависимость от массы тела W
$B \propto \mathrm{f}(P)^z$	$B \propto W^z$
$D \propto \mathrm{f}(P)^z$	$D \propto W^z$
$HR \propto P^z$	$HR \propto W^z$

где $\propto$ – математический символ пропорциональности; B – биомасса; D – популяционная плотность (экз./м^2); HR – зона обитания (Home Range); V – неоднородность распределения. Это означает, что при изменении разрешения шкалы масштабов одной и той же изучаемой области имеют место степенные соотношения типа $S_1/S_2 = (A_1/A_2)^z$, что привело к развитию идей о самоподобном иерархическом характере экологических сообществ (Allen, Starr, 1982; Азовский, 2001; Гелашвили и др., 2007а, 2008а).

Естественно, отношения простой степенной пропорциональности не могли удовлетворить ни экологов, ни математиков. Поэтому последние 20 лет ознаменовались бурным развитием макроэкологических методов на базе имитационных моделей Монте-Карло и алгоритмов генерации псевдовыборок (Gotelli, Colwell, 2001; Chao, 2005), стохастических моделей (Tokeshi, 1993), мультифрактального анализа (Гелашвили и др., 2007б, 2008б), универсальной нейтральной теории (Hubbell, 2001) и др.

Один из «священных» принципов в экологии – идея, согласно которой биологическое разнообразие видов обеспечивает стабильность сообщества и развитие экосистем (Elton, 1946, 1958). Правда остается неясным, какой аспект разнообразия (видовая насыщенность, функциональная сбалансированность, трофическая обусловленность или что-то иное) здесь наиболее важен и в какой мере сама окружающая среда формирует механизмы взаимосвязи разнообразия и стабильности (Sankaran, McNaughton, 1999). Вопрос запутывается окончательно, если попытаться задуматься над смыслом самого термина «стабильность», так как насчитано (Grimm, Wissell, 1997) 163 его определения и 70 версий различных толкований. Несмотря на эти методологические сомнения, в последние десятилетия активно продолжались исследования, основанные на идеях макроэкологии и связанные с изучением влияния на динамику видовой структуры сообществ всего комплекса движущих сил: воздействия внешних факторов (Brown, Lomolino, 1998), исторических процессов или характерных биологических циклов (Cornell, Lawton, 1992), особенностей морфологии ландшафта или региона (Turner, 1989; Schlosser, 1991; Алимов, 2001) и др.

Важнейшим вопросом экологии, имеющим длинную историю, является выявление и анализ механизмов распределения биотических компонентов в пространстве и времени. В истории развития моделей распределения видов (species distribution models – SDMs) отмечены три фазы:

- непространственный статичный анализ взаимосвязи флоры и фауны с окружающей средой, основанный на локальных массивах эмпирических данных;
- пространственное моделирование распределения видов в континентальных масштабах, основанное на представлениях экспертов и не использующее статистическую обработку данных мониторинга;
- пространственное статистическое и эмпирическое моделирование распределения видов на всех масштабных уровнях, использующее распределенные базы эмпирических данных.

В последние годы прогнозы, основанные на моделях SDMs, стали важным инструментом решения различных проблем в экологии, биогеографии, эволюции, сохранении биоразнообразия и исследованиях по климатическим изменениям (Guisan, Thuillier, 2005).

Процедуры, моделирующие распределения видов, столь же разнообразны, как многолики проблемы самой экологии. Однако практикой выработаны основные принципы идеального их построения, включающие шесть последовательных этапов (см. табл. 1.5): 1) осмысление задачи, 2) подготовка данных, 3) выбор алгоритма решения, 4) расчеты по моделям, 5) обобщение результатов и 6) оценка применимости модели. Подробно представленная схема обсуждается в обзорах (Guisan, Zimmermann, 2000; Guisan, Thuillier, 2005).

Таблица 1.5

Важнейшие особенности, связанные с построением моделей распределения видов (SDM), критерии для их потенциального применения и некоторые предложенные решения (по: Guisan, Thuiller, 2005, с сокращениями)

Рассматриваемые аспекты	Решаемая проблема	Метод индикации	Примеры предложенных решений
1. *Концептуализация задачи*			
Тип изучаемых организмов и их популяций	Мобильные виды, изучаемые вне их ареалов	Непрерывный мониторинг (радиодатчики)	Экстраполяция функций распределения; выбор размеров ячеек
	Сессильные виды, изучаемые вне их ареалов	Недостаточная выживаемость и низкое размножение	Использование критериев соответствия при настройке модели
	Виды, наблюдаемые в их естественной среде	Знание жизненной стратегии видов (рассеивание и т. д.)	Интерпретация различных типов отклонений
	Редкие и трудно обнаруживаемые виды	Полевые исследования, литературные источники	Уточнение тестов на вероятность обнаружения
	Родственные виды одного экотипа	Генетические исследования	Дифференцирование ниш вдоль градиентов
	Инвазионные виды	Непредвиденное обнаружение	Использование аналогий из регионов их происхождения
Тип факторов	Прямые и косвенные факторы	Знание экофизиологии	По возможности, избегать косвенных факторов
Планирование исследований	Разработка схемы отбора проб	Имитационные испытания	Рандомизация выборок в пространстве
Соответствие экологической теории при оценке характера зависимости вида от факторов	Линейный или унимодальный отклик	Частные графики, анализ корреляций	Модели GAM или квадратичные члены GLM
	Полимодальные кривые отклика	Частные графики, сглаживающие кривые	GAM, полиномы 3-й степени или β-функции в GLMs
	Асимметрия унимодальных кривых	Оценка асимметрии	HOF, β-функции в GLM, GAM, модели нечетких множеств
2. *Подготовка данных для построения моделей*			
Данные по видам	Подтверждение в исторических коллекциях (NHC)	Картографическое и статистическое исследование	Отбор наблюдений в пределах заданной точности
Масштаб (степень детализации или размер пространственных ячеек)	Дифференциация размера ячеек для каждого фактора	ГИС-исследования	Объединение всех слоев ГИС и установление масштабов разрешения
	Оценка степени усечения градиентов	Предварительное исследование кривых отклика видов	Увеличение области исследования для формирования полных градиентов

3. ***Идентификация моделей***			
Мульти-колинеарность	Корреляция факторов	Коэффициент возрастания дисперсии (VIF)	Удаление скоррелированных факторов; ортогонализация
Пространственная автокорреляция (SAC)	Взаимная зависимость наблюдений	Индексы автокорреляции	Изменение порядки отбора проб, чтобы избежать SAC; компенсация SAC в моделях
Тип статистической модели	Повышенная дисперсия (overdispersion)	Отклонение для остатков и остаточные степени свободы	Квази-распределение в GLMs и GAM; шкалирование отклонений
Выбор модели	Тип критерия и степень приближения	–	Оптимизация размеров модели на основе AIC-критерия; кросс-валидация
4. ***Оценка параметров модели***			
Тип данных	Бинарные (вид найден или нет)	Обычные методы не применимы	Использование моделей частот встречаемости
Структура оценки	Отсутствие независимости исходных выборок	–	Процедуры многократного формирования псевдовыборок (ресамплинг)
Метрики зависимости	Выбор порога для бинарных данных	–	Пороговые независимые меры (AUC)
	Стоимости ошибок	–	Взвешивание ошибок
Неопределенность модели	Недостаточное доверие к данным и моделям на их основе	Графики распределений остатков модели	Байесовский подход; пространственное взвешивание
Неопределенность выбора модели		Сравнение алгоритмов расчета и конкурирующих моделей	Объединение моделей в коллектив
5. ***Прогноз пространственной динамики***			
Проекция модели на новые регионы	Диапазон факторов находится вне изученной области	Статистические расчеты	Оценка формы кривых отклика и ограничение области экстраполяции
6. ***Оценка области применимости модели***			
Область действия и применимость	Модель не применима для другого региона	Нереалистичные кривые ответа	Снижение размерности модели
	Модель не поддается объяснению	Похоже на модель «черного ящика»	Выбор моделей, легких для объяснения и интерпретации (GLMs, GAM, CART)

Примечание: GLM – обобщенные линейные модели; GAM – обобщенные аддитивные модели; HOF – модели Хаусмана–Олфа–Фреско; CART – деревья классификации и регрессии; AIC – информационный критерий Акаике.

Интенсивное развитие макроэкологических исследований в последнее десятилетие проходило, прежде всего, при изучении наземных сообществ. Обширный потенциал водных экосистем в значительной степени остается недоиспользованным. Причиной тому являются большая трудоемкость сбора материала, скромное финансирование, большое число мало обследованных регионов, а также трудность оценки таких переменных, крайне важных в макроэкологическом анализе, как «размер и численность популяции», особенно для очень мобильных пелагических таксонов. Несмотря на эти трудности, существует несколько превосходных макроэкологических исследований морской среды (например, Foggo et al., 2003; Fisher, Frank, 2004; Blackburn et al., 2004). Определенный прогресс в этом направлении наметился в связи с разработкой в рамках Европейской программы MarBEF

(Marine Biodiversity and Ecosystem Functioning; www.marbef.org) общедоступной базы данных MacroBen по морской донной фауне (Webb et al., 2009).

Проведенные исследования морских экосистем основывались на канонической форме основных макроэкологических моделей и имели тенденцию поддерживать общие выводы, сделанные наземными макроэкологами. Например, было установлено, что кривые распределения обилия видов имеют одинаково типичную правостороннюю асимметрию и для морских и для земных таксонов (Gaston, 2003), как и основные закономерности пространственного размещения особей (Foggo et al., 2003). В то же время существуют серьезные различия в численных параметрах моделей, которые определяются большим таксономическим разнообразием и высокой долей редких видов, характерных для водных сообществ. Очень велика также межвидовая вариация численности особей и их массы тела, что делает натуральные показатели (экз./м^2, г/м^2 и т. д.) практически некорректными для статистической обработки. Приходится применять различные функциональные преобразования (степенные, логарифмические, тригонометрические [например, arcsin(N)] и др.), что часто лишает результаты сопоставимости и стандартизации.

Важно отметить, что макроэкология – это больше, чем просто модели (Blackburn, Gaston, 2006), и её основные усилия постепенно перемещаются к поиску механизмов процессов, которые описывают эти «паттерны». Например, модели популяций, которые рассматривают межвидовые взаимодействия как результат внутривидовой динамики, обеспечили эффективный подход для прогнозирования роли экологических факторов и параметров биологического цикла в формировании морских сообществ (Freckleton et al., 2005).

1.5. Нуль-модели в экологии

Стандартным инструментом макроэкологических исследований стали «нуль-модели», т. е. *теоретические конструкции, которые моделируют изучаемый объект с минимально необходимым количеством исходных экологических предположений, а во всем остальном предполагается стохастический характер процессов.*

Рассмотрим смысл нуль-моделей на простом примере. Рассчитаем *индекс*[6] *видового сходства Q Чекановского–Съеренсена* двух сообществ макрозообентоса: в р. Байтуган и в верхнем течении р. Сок. Вычисленное значение видового сходства по результатам гидробиологических проб равно 0,58. Однако, насколько оно велико и статистически значимо? Поскольку выбранный индекс Q (как и многие другие) – просто число, это не обеспечивает даже намека на подобную интерпретацию.

Для построения нуль-модели перемешаем все виды и будем многократно и в случайном порядке перераспределять их между сравниваемыми сообществами, каждый раз рассчитывая новое псевдозначение индекса Чекановского–Съеренсена. Равновероятная нуль-модель RA3 **(**Winemiller, Pianka, 1990) после 500 итераций дает среднее значение 0,543, причем 95,2% модельных индексов находится в диапазоне от 0,541 до 0,549. Поскольку индекс, полученный по результатам наблюдения, находится за пределами 95% доверительного интервала сравнений случайных композиций видов, мы можем предположить, что сходство между двумя участками водотока носит закономерный характер, определяемый экологическими причинами. Однако, если бы эмпирическое значение Q оказалось менее 0,55, то такой уровень сходства считался бы статистически незначимым и основанным на случайной вариации видового состава. В этом примере нуль-модель предоставила нам новые возможности, чтобы оценить степень подобия и его дивергенцию.

---

[6] Он называется также *индексом перекрывания ниш Чекановского* $Q = 1 - 0{,}5\sum_{j=1}^{n} | p_{1j} - p_{2j} |$, где p_{1j} и p_{1j} – доли использования j-го ресурса (Feinsinger et al., 1981).

Основной вопрос экологии, имеющий давнюю историю, – не является ли сложившаяся структура сообщества чистой случайностью, т. е. стохастической выборкой из регионального пула хаотически перемешанных видов и их особей? Первой «колодкой», вступившей в противоречие с принципом конкурентного исключения как основной детерминанты сообществ, явилось предположение П. Сейла (Sale, 1982) о «лотерее» видов, захватывающих ограниченное пространство. Он продемонстрировал на примере популяций рыб, что борьба за освоение участков территории коралловых рифов новыми «рекрутами» определяется системой случайных приоритетов, а не тонкими различиями в потребностях или конкурентных способностях отдельных видов. Взрослые особи производят огромное количество гамет, чтобы часть их потомства могла выиграть «лотерею» и первой достичь какого-то свободного участка. Прибывающие личинки завоевывают свое положение в иерархии видов формируемого сообщества и оказываются в состоянии защитить участок, не допуская опоздавших особей. Лотерея продолжается, пока не разыграны все свободные участки. Эта «лотерейная» гипотеза может объяснить высокое локальное разнообразие в пределах определенного местообитания.

В дальнейшем Д. Стронг и Д. Симберлофф (Strong, Simberloff, 1981; Бигон и др., 1989) подвергли критике то, что они назвали тенденцией объяснять «простые различия» действием межвидовой конкуренции. Действительно, каким образом изменения в видовой структуре сообществ можно счесть статистически значимыми и в какой степени эти сдвиги являются следствием биологических процессов, которыми их пытаются объяснить? Не обеспечивает ли адекватное объяснение наблюдаемых данных нулевая гипотеза H_0, которая исходит из предположения об отсутствии межвидовых взаимодействий?

Например, теория конкуренции предсказывает, что виды с очень близкими нишами должны взаимно исключаться, а следовательно, они должны расходиться в пространстве ниш сильнее, чем можно было бы ожидать на основании чистой случайности (т. е. эти ниши оказываются как бы раздвинуты – так называемая «сверхдисперсия» ниш). Однако при более доказательном исследовании приходится искать ответ на следующий вопрос: значимо ли отличается наблюдаемое распределение видов, которому приписывается эффект межвидовой конкуренции, от того, которое возникло бы в сообществе в отсутствие каких бы то ни было взаимодействий между видами?

Вопросы подобного типа стимулировали ряд работ, посвященных сравнению реальных сообществ с так называемыми «нейтральными или нуль-моделями». Последние представляют собой модели реальных сообществ, отражающие основные средообразующие характеристики своих природных прототипов, но объединяющие видовые компоненты случайным образом с сознательным исключением возможных последствий внутрибиологических взаимодействий.

П. Харви с соавторами (Harvey et al., 1983) и другие исследователи подробно обсудили построение и использование нуль-моделей в различных областях экологии сообществ.

Распределение относительного видового обилия. До 1970-х годов почти все биологические интерпретации предложенных статистических моделей распределения обилия видов не имели строгой теоретической базы. Основное «предсказание» этих гипотез сводилось к тому, что в случае отсутствия сильных межвидовых взаимодействий сообщество должно состоять из большего числа видов с равномерно распределенным относительным обилием. Г. Касвеллу (Caswell, 1976) пришло в голову, что проверить это утверждение можно единственным способом – смоделировать кривую обилия видов для сообщества без межвидовых взаимодействий вообще. Привлекая алгоритмы, используемые в генетике, Касвелл разработал модели, в которых вероятности рождения и смертности особей равны, а число новых видов, вносимых в сообщество, регулировалось параметром статистического распределения Пуассона, который не зависит от видовой насыщенности.

Частично как противовес моделям Касвелла, Дж. Сугихара (Sugihara, 1980) разра-

ботал количественные модели распределений обилия видов для сообществ с сильными взаимодействиями (конкуренцией), основанные на определенном алгоритме случайного, но последовательного деления многомерного пространства ниш среди конкурирующих видов. Поскольку кривые Сугихары хорошо аппроксимировались каноническим лог-нормальным распределением, это вызвало длительную дискуссию. С одной стороны, естественные сообщества слишком хорошо «укладывались» в рамки лог-нормального распределения, чтобы это было простой случайностью. Это привело к искушению объявить этот тип распределения неотъемлемым свойством сообществ, связанным с механизмом распределения используемых ресурсов (Sugihara, 1980).

С другой стороны, хорошая корреляция с эмпирическими данными сама по себе вряд ли может гарантировать корректность моделей Сугихары – многие массивы наблюдений являются, по сути дела, не статистическими повторностями, а беспорядочно составленной «смесью» выборок (Lambshead, Platt, 1985). Лог-нормальное распределение может быть реакцией на статистические свойства любых сложных систем, обусловленные центральной предельной теоремой (May, 1975). Эта теорема гласит, что когда величина некоторой переменной определяется большим числом факторов, то их случайное варьирование приводит к нормальному распределению значений этой переменной. В случае лог-нормального распределения обилий видов, такой переменной является число особей, стандартизованное логарифмическим преобразованием, а определяющими факторами – экологические процессы, управляющие сообществом.

Нуль-модели и разделение ресурсов. Если конкуренция является или была в прошлом важным фактором, определяющим структуру реальных сообществ, то ниши входящих в них видов должны заметно различаться, а перекрывание в использовании ресурсов должно быть меньшим, чем предсказывают нейтральные модели. На проверке этой гипотезы построил свое исследование Л. Лолор (Lawlor, 1980), проанализировав десять сообществ ящериц в Северной Америке.

Для каждого вида i (i = 1, 2, … N, N = 9) по результатам наблюдений были получены доли K_{ij} ($0 \geq K_{ij} \geq 1$) каждого j-го ресурса (j = 1, 2, … M, M = 20) в его пищевом рационе. По значениям K_{ij} рассчитывали индекс перекрывания используемых ресурсов для каждой пары видов, который также колебался в пределах от 0 (отсутствие перекрывания) до 1 (полное перекрывание). В итоге каждое сообщество было охарактеризовано одной единственной величиной: средним перекрыванием S'_j в использовании ресурса для всех имеющихся пар видов. Нейтральные модели были четырех типов, поскольку строились с помощью четырех «алгоритмов перегруппировки» (АП1 – АП4). В каждом варианте сохранялась какая-либо одна черта структуры исходного сообщества, в то время как остальные особенности использования ресурсов были рандомизированы. Например, при алгоритме АП1 первоначальную структуру сообщества изменяли в наибольшей степени: учитывали только исходное число видов N и число категорий корма M, а наблюдаемые значения индексов избирательности заменяли в каждом случае случайными величинами в пределах от 0 до 1. Каждый из четырех алгоритмов применяли к каждому из десяти сообществ, и было получено 4000 «нуль-модельных сообществ» и вычислены соответствующие им средние величины перекрывания ресурсов S^o. Если бы конкуренция играла важную роль в сообществах, то значения S^o оказались бы статистически больше наблюдаемых для реальных сообществ S'. Считалось, что среднее перекрывание в нейтральной модели значимо превышает отмечаемое в природе ($p < 0{,}05$), если не более 5 из 100 вариантов моделирования дадут средние величины S^o, меньшие, чем S_{obs}, характеризующая реальное сообщество. Было установлено, что почти во всех случаях модельные значения S^o значимо превышают S_{obs}. Таким образом, низкие показатели перекрывания используемых ресурсов в рассмотренных сообществах ящериц предполагают разделение ниш и значительную роль межвидовой конкуренции.

Нейтральные модели и морфологические различия. Разделение ниш часто проявляется в морфологической дифференциации. Например, Дж. Хатчинсон (Hutchinson, 1959) высказал предположение, что если отдельные виды гильдии четко расходятся вдоль оси, соответствующей какому-то ресурсу, то соседние по этому градиенту виды закономерно различаются размерами тела или размерами других структур, связанных с питанием.

Применение нуль-моделей для анализа этой кажущейся закономерности в морфологических различиях связано с гипотезой о том, что совместно обитающие потенциальные конкуренты иногда сильнее различаются по морфологии, чем ожидалось бы при действии одних лишь случайных факторов. Были исследованы, например (Bowers, Brown, 1982), соотношение размеров тела у пустынных грызунов в двух масштабах: в пределах небольших (несколько гектаров) участков местности и в пределах обширных географических областей. Подсчитывалась ожидаемая частота совместной встречаемости каждой возможной пары видов при допущении, что распределение каждого из них независимо от распределения остальных. Модельные значения сравнивали с реальными частотами совместной встречаемости, предварительно отметив, составляет ли отношение масс тела больше или меньше 1,5, и рассчитывали вероятность случайной сочетаемости.

Наличие межвидовой конкуренции в этом и в других подобных исследованиях признавалось, если в реальных сообществах отношения размеров тела были в целом значимо больше, чем в «нулевых». Оказалось, что во многих случаях такой разницы не наблюдается, поэтому отбросить нуль-гипотезу и доказать наличие межвидовой конкуренции не удалось.

Нуль-модели и различия ареалов. Отсутствие сосуществования морфологически сходных видов представляет собой частный случай невозможности совместного обитания потенциальных конкурентов в целом. На её основе Дж. Даймонд (Diamond, 1975), например, предложил механизм оценки конкурентного исключения среди птиц архипелага Бисмарка. Однако Э. Коннор и Д. Симберлофф (Connor, Simberloff, 1979), использовав нейтральные модели, случайным способом составили сообщества из исходной совокупности видов и пришли к заключению, что реальные сообщества не имеют значимых отклонений от случайно подобранных. Роль конкуренции была четко установлена лишь в немногих случаях, хотя признавалось, что она остается одним из основных факторов организации структуры сообщества.

В большинстве описанных случаев завязывались дискуссии и высказывались критические замечания в отношении статистических неточностей и методических промахов нейтральных моделей. Их оппоненты обнаруживали дополнительные свидетельства межвидовой конкуренции, хотя и не смогли доказать ее влияние в каждом отдельном случае. Вероятно, самый важный вывод из этих дискуссий состоял в том, что критики расходились в деталях анализа и в способах построения нейтральных моделей, однако в их пользе не сомневался никто.

Конечно, надо бороться с искушением видеть какие-то механизмы в развитии сообществ лишь только потому, что мы их там ищем, и даже те, кто критиковал первые нейтральные модели, старались предложить взамен улучшенные их версии. Однако особо подчеркивалось, что невозможность опровержения нейтральной модели еще не говорит об отсутствии конкурентных взаимодействий. «Общепризнано, что гораздо надежнее отвергнуть (или опровергнуть) отсутствие какого-либо эффекта, чем подтвердить его наличие, поскольку существуют хорошо разработанные статистические методы, позволяющие определить, имеются ли между объектами достоверные различия, а вот методов для определения "значимого сходства", к сожалению, нет» (Бигон и др., 1989, с. 235). В лучшем случае такие модели могут указать на нежелательность привлечения межвидовых взаимодействий для объяснения наблюдаемых фактов.

Безусловно, описанный метод правомерен только при исключении из самой нуль-модели всех возможных эффектов конкуренции. Это положение очень четко подтверждено

еще 25 лет назад компьютерными исследованиями Р. Колуэлла и Д. Уинклера (Colwell, Winkler, 1984). Использовав программу GOD, они построили модель сообщества, филогенетические связи которого подчиняются определенным законам. Интересным и несколько обескураживающим результатом было то, что некоторые нейтральные модели не позволяли обнаружить каких-либо последствий конкуренции, даже если она задавалась программно. Главная из причин, затрудняющих выявление «призрака конкурентного прошлого», состоит в том, что наиболее уязвимые для конкурентного исключения виды, по всей видимости, уже давно исчезли из массива данных, взятых для обработки.

Анализ данных с помощью нуль-моделей представляет собой олицетворение общего «доказательного» подхода к экологическим исследованиям, заключающегося в построении и проверке нуль-гипотез. Этот подход, давно утвердившийся в научной методологии, состоит в том, что имеющиеся экспериментальные данные перегруппировываются в форме, имитирующей отсутствие исследуемого явления, после чего они сопоставляются с реальными измерениями. Если имеются статистически значимые отличия, то нулевая гипотеза отвергается и делается вывод о высокой вероятности исследуемого явления.

Современные методы разработки и исследования нуль-моделей основываются на четырех сходных процедурах (Crowley, 1992; Manly, 2007; Шитиков и др., 2008), известных в статистической и биологической литературе как рандомизация (randomization), алгоритмы Монте-Карло, бутстреп (bootstrapping) и «складной нож» (jackknife). Все вместе они позиционируются как различные схемы ресамплинга (resampling), генерирующего большое число псевдовыборок с использованием исходного набора данных или иного стохастического процесса, который, как предполагается, мог бы сформировать эти данные.

Рандомизация, или *перестановочный тест* (permutation), заключается в многократном случайном перемешивании относительно уровней изучаемых факторов воздействия строк или столбцов таблицы наблюдений. При каждой итерации на основе сгенерированной псевдовыборки рассчитываются имитируемые значения Q_{sim} анализируемого показателя или статистики, которые сравниваются с аналогичной величиной Q_{obs}, найденной по эмпирическим данным. Обычно в ходе перестановок не меняется ни состав исходной таблицы, ни численность групп с разными уровнями воздействия, а только происходит беспорядочный обмен элементами данных между этими группами.

Если, например, нулевая гипотеза H_0 состоит в том, наблюдаемое значение Q_{obs} не превышает величину, характерную для случайной совокупности данных (односторонняя гипотеза), то вероятность ошибочного отклонения H_0, если она верна, будет соответствовать количеству зарегистрированных ситуаций, в которых $Q_{sim} \geq Q_{obs}$. Соответственно, уровень значимости p, с которым отвергается гипотеза, будет равен доле этих ситуаций из общего числа итераций. Например, в приведенном выше примере после 500 циклов перемешивания данных вычисляемый индекс перекрытия ниш Чекановского ни разу не превысил наблюдаемое значение 0,58, следовательно, гипотеза о случайном характере этого показателя отклоняется с вероятностью ошибки 1-го рода $p = 0{,}002$ (т. е. 1/499).

В ходе рандомизации часто оцениваются стандартные отклонения, доверительные интервалы и традиционные статистические критерии (например, t, F), с помощью которых вычисляются p-вероятности значимости проверяемых гипотез, не связанные никакими предположениями о характере распределения статистических величин.

В *методах Монте-Карло* задается некоторый многократно повторяемый вероятностный процесс различной степени сложности (от простого биномиального распределения подбрасываемой монеты до комплексной стохастической имитационной модели), который, как предполагается, лежит в основе данных наблюдений. Сгенерированные псевдовыборки могут быть использованы для оценки доверительных интервалов или проверки нулевой гипотезы. Тестирование значимости отличий эмпирических наблюдений от предсказаний модели происходит по той же схеме, что и при рандомизации.

Существуют мнения (Manly, 2007), что рандомизация вообще является частным случаем испытаний Монте-Карло, в которых специфический стохастический процесс просто осуществляет равновероятные перестановки данных среди уровней воздействия. Однако различия этих методов в основных предположениях и ограничениях становятся весьма существенными в более типичных для методов Монте-Карло исследованиях, когда данные наблюдений вообще не используются, чтобы смоделировать вероятностный процесс.

«*Складной нож*» (jackknife), как и его универсальный тезка, является наиболее простым, но весьма эффективным методом расчета, доступном на простом калькуляторе. Прямые результаты этих вычислений – «улучшенные» (т. е. менее обусловленные выборочными эффектами) оценки типичных статистических параметров: среднего, эксцесса, дисперсии, доверительного интервала, скорости экспоненциального роста и др.

Обычно используется «складной нож» первого порядка, хотя версии высшего порядка также могут быть иногда полезными (примеры см.: Manly, 2007). Предположим, что нас интересует параметр $\hat{\mathbf{S}}$, например, истинная стандартная ошибка популяционной плотности особей (Crowley, 1992). Пусть его значение, оцененное по всей выборке наблюдений m равно S. Если удалить из выборки одно i-е наблюдение, то можно рассчитать псевдооценку стандартной ошибки популяционной плотности $s_j = S - (s_{-i} - S)(m - 1)$, имеющую некоторое смещение относительно S, выраженное правой частью этого уравнения. Среднее из всех возможных псевдозначений s_j, $i = 1, 2, \ldots m$, является оценкой S_{-1} «складного ножа» первого порядка. Различие $S - S_{-1}$ определяет полное смещение первоначальной выборочной оценки S, т. е. отражает возможное искажение оценки параметра $\hat{\mathbf{S}}$, если, например, особи образуют локальные скопления в пространстве.

Бутстреп (bootstrap – ремешок ботинка) – быстро развивающийся набор алгоритмов, привлекающий большое внимание в статистической литературе. Основная идея метода – моделирование плотности многомерного распределения обрабатываемых переменных с использованием эмпирических выборок. Таким образом, бутстреп использует некоторую метафору исходных данных, формируя наилучшее доступное отображение их частотного распределения. На этой основе запускается итерационный процесс генерации псевдовыборок, в ходе которого в исходной матрице многократно осуществляются случайные извлечения с заменой отдельных фрагментов данных.

Можно отметить важное различие между рандомизацией и бутстрепом. Применение рандомизационного теста оправдано, если ставится задача оценить степень упорядоченности массива данных или взаимосвязи между отдельными его фрагментами. Если при подсчете статистики (например, среднего) порядок следования элементов выборки не имеет значения, то каждая итерация рандомизации возвращает одну и ту же величину, поскольку сами по себе данные не изменяются. В ходе бутстрепа, наоборот, на каждом шаге итерации «замены с возвращением» отдельные значения в первоначальной выборке подвергаются изменениям, что слегка модифицирует частотное распределение реализаций исходных данных. Следовательно, можно ожидать, что каждая следующая бутстреп-выборка будет возвращать значение среднего, немного отличающееся от вычисленного для первоначальной выборки. Образующийся разброс показателя дает возможность построения доверительных интервалов анализируемой статистики (Manly, 2007).

Возможность гибкой настройки и использование идей самоорганизации отличает бутстреп от метода «складного ножа» с его параметризованным и менее интенсивным вычислительным подходом. Вместе с тем, существование самостоятельных выборочных процессов для разных уровней воздействия внешних факторов делает бутстреп менее зависимым от выборочных эффектов, чем более ограниченная в этом смысле рандомизация.

Пути дальнейшей модернизации нуль-моделей, широко обсуждаемые на протяжении двух последних десятилетий, и методы ресамплинга, используемые при их конструировании, будут подробно рассмотрены нами на конкретных примерах, представленных в главах 2-7.

ГЛАВА 2. МЕТОДЫ ОЦЕНКИ ВИДОВОГО БОГАТСТВА СООБЩЕСТВ

Нет истины, а есть лишь ряд
асимптотических приближений к ней…
Дмитрий Быков "ЖД" (2006 г.),
приписывается В. Набокову

2.1. Взаимосвязь между локальным видовым богатством и региональным пулом видов

Оценка видового богатства в различных пространственных масштабах является краеугольным камнем изучения биологического разнообразия (Ricklefs, Schluter, 1993) и наиболее существенным параметром контроля экологических изменений (Karr, 1991; Rosenberg, Resh, 1993). Точное определение числа видов сообщества в пространственно-временных координатах позволяет количественно определить интенсивность процессов возникновения, исчезновения и миграции видов, а также проверить различные экологические модели и конструкции, такие как теория островной биогеографии (MacArthur, 1965), концепции режимов возмущений (disturbance regimes – Connell, 1978) и видового насыщения (species saturation – Griffiths, 1997; Caley, Schluter, 1997).

В соответствии с глобальной моделью видового разнообразия (Риклефс, 1979), число видов для организмов любых таксономических групп на больших материковых территориях определяется соотношением скоростей двух процессов: появлением новых видов, преимущественно вследствие колонизации (инвазии), и исчезновением видов в результате воздействия факторов среды на локальном или региональном уровне. Скорость видообразования возрастает прямо пропорционально числу видов, уже присутствующих в данной области, но по достижению некоторого уровня видового богатства, она замедляется и затем (это положение спорно: Василевич, 1979) стабилизируется на одном уровне из-за сокращения экологических возможностей для появления новых видов. Так как с возрастанием видового богатства сообщества S увеличивается доля специализированных и малочисленных таксонов, то риск исчезновения видов повышается, особенно в непредсказуемо варьирующейся среде (Hanski, 1982). Есть предположения (Ricklefs, Schluter, 1993), что зависимость скорости вымирания от S носит экспоненциальный характер (как в островной *модели Мак-Артура–Уилсона*; MacArthur, Wilson, 1963). В случае, когда в результате эволюции биоты скорости процессов возникновения и исчезновения видов уравновешиваются и достигается точка эволюционного равновесия, можно говорить об эволюционной *полночленности* сообщества (Работнов, 1983), т. е. соответствию видового богатства условиям (емкости) среды на занимаемой территории.

Если эволюционное равновесие в принципе существует, то теоретически оно может удерживаться неопределенно долгое время. Разумеется, по мнению В.А. Красилова (1973), ритмические колебания климата, геологические и другие факторы могут приводить (и неоднократно приводили в истории Земли) к изменению емкости экологического пространства. В результате этого при снижении емкости биоценозы оказывались перенасыщенными видами, а при увеличении – недонасыщенными (неполночленными). Соответственно, нельзя исключить, что видовое богатство многих хорошо известных современных пространственных структур не является прямым следствием различий физических условий и связано (по крайней мере, частично) со стохастической природой процессов видообразования и уникальными событиями истории (Ricklefs, 1987; Ricklefs, Schluter, 1993; Акатов и др., 2003; Акатов, Чефранов, 2007). Однако оценка полночленности и выявление причин сдвигов в эволюционном равновесии современных сообществ связаны со значительными методическими трудностями. Традиционно применяют три подхода к решению этой проблемы:

- анализ характера исторических изменений видового богатства сообществ по палеонтологическим данным (Бигон и др., 1989);
- сравнение видового состава и богатства сообществ биологически равноценных местообитаний, расположенных в разных регионах или даже на разных континентах (Уиттекер, 1980);
- анализ характера и силы связи между видовым богатством биоты и климатическими или физико-географическими условиями среды (Малышев, 1969; Пианка, 1981).

Ряд исследователей, цитируемых в настоящем разделе, считают, что эта задача имеет перспективу решения, если рассматривать её в аспекте анализа соотношений между локальными процессами и региональным видовым богатством.

Под видовым фондом (*пулом видов*) сообществ понимается группа видов, обитающих в пределах определенной области и потенциально способных существовать в этих сообществах (MacArthur, Wilson, 1963; Taylor et al., 1990). М. Партел с соавторами (Partel et al., 1996) детализировали понятие видового фонда и предложили различать три его типа, дифференцированных в соответствии с пространственным масштабом:

- *региональный* фонд (regional species pool) – группа видов, встречающихся в некотором регионе, способных существовать в некотором сообществе или достигать его при распространении;
- *локальный* фонд (local pool) – группа видов, встречающихся в окрестностях сообщества, способных существовать в нем или мигрировать в него относительно быстро;
- *актуальный* или *ценотический* фонд (actual pool) – группа видов, характерных для региона и являющихся диаспорой для любого его участка.

Пространственные границы регионального фонда определяются границами распространения сообществ определенного типа в регионе, а локального фонда – границами ландшафтных образований, включающих эти сообщества (например, водоток или бассейн водотока). Локальное богатство измеряется в пространственном масштабе, достаточно небольшом, чтобы все виды могли взаимодействовать друг с другом в течение относительно непродолжительного периода времени.

В пределах региона для конкретного изучаемого сообщества различают (Butaye et al., 2002) *полное* (total species richness – TSR), *локальное* (local species richness – LSR) и *наблюдаемое* видовое богатство (observed community richness – OSR). В принципе локальный пул видов можно считать «размытым множеством» (Roberts, 1986), к которому теоретически может принадлежать любой из видов, образующих общее таксономическое богатство TSR. При этом, однако, каждый вид имеет разную вероятность вхождения в локальное сообщество, которая определяется главным образом степенью соответствия между оптимумом экологических условий, характерных для данного вида, и специфическими экологическими параметрами для биотопа в целом. Важны также факторы, определяющие способность особей вида перемещаться в пространстве, в частности, способность к адаптации, условия видового распространения и рассеивания.

Подчеркивается (Eriksson, 1993; Brown et al., 2000), что формирование сообществ происходит в результате синхронных действий факторов среды, охватывающих целые ландшафтные зоны, и стохастических процессов, определяющих нарушение баланса миграции и вымирания локальных популяций. Поэтому, независимо от масштаба действия, факторы среды как абиотического, так и биотического характера, оказывают одновременное влияние и на локальную и на региональную встречаемость видов, а соответственно, и на размерность TSR и LSR. Процессы, определяющие размер каждого из формируемых видовых списков (или, в иной терминологии, фондов, пулов или богатств) для изучаемых сообществ, представлены на рис 2.1.

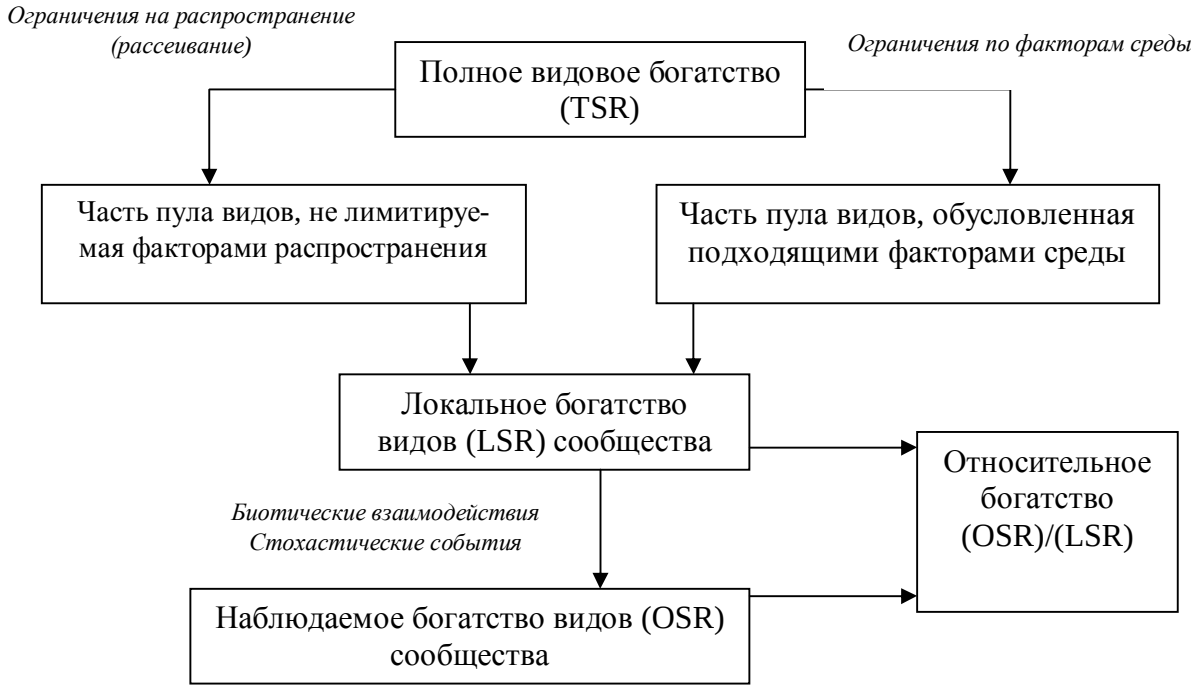


Рис. 2.1. **Процессы, определяющие размеры богатства видов (по: Butaye et al., 2002).**

Рассмотрим основные концепции и спектр мнений, которые сопутствуют гипотезе видовых фондов. Существует относительно небольшое число публикаций, в которых анализируется соотношение между локальным богатством и размерами фондов видов. В частности, для водных экосистем объектами исследований становились рыбы (Hugueny, Paugy, 1995), насекомые и паразиты амфибий (Cornell, Lawton, 1992).

Г. Корнелл (Cornell, 1985) предложил две возможные формы зависимости локального богатства от размера видового фонда. Первая – криволинейная зависимость, когда с увеличением размеров видового фонда локальное богатство сначала растет, а затем достигает верхнего предела, обусловленного уровнем сопротивления условий среды, и стабилизируется (это соответствует *насыщенным* сообществам – рис. 2.2a). Вторая – положительная линейная зависимость, которую на ряде примеров продемонстрировал Р. Риклефс (Ricklefs, 1987), свидетельствующая о значительной роли размеров видового фонда в формировании локального видового богатства в широком диапазоне его масштаба.

Таким образом, для *ненасыщенных* сообществ характерно фиксированное соотношение регионального и локального пулов видов. Однако практически в одних случаях эмпирический материал давал возможность предполагать линейную зависимость, в других – кривую с насыщением, но в большинстве случаев результаты следовало бы признать как неопределенные.

Д. Сривастава (Srivastava, 1999) обратила внимание на два интересных обстоятельства. Во-первых, зависимость между локальным S_{loc} и региональными S_{reg} пулами видов идентично выражению 1/(1 – β-разнообразие), если использовать для β-разнообразия в дефиниции Р. Уиттекера (Whittaker, 1972) соотношение (S_{reg} / S_{loc}) – 1. Предполагается (Cornell, Lawton, 1992), что β-разнообразие не будет изменяться в регионах с ненасыщенными сообществами (рис. 2.2b), поскольку у этих областей одинаковое соотношение локального и регионального богатства видов. В сообществах, насыщенных видами, увеличение регионального богатства будет сопровождаться увеличением β-разнообразия.

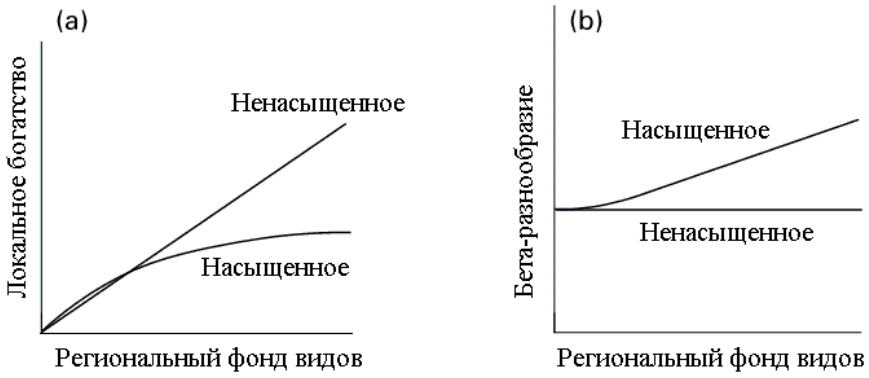


Рис. 2.2. **Зависимости локального богатства (a) и β-разнообразия (b) от размеров регионального фонда видов для насыщенных и ненасыщенных сообществ (по: Srivastava, 1999)**

И локальные, и региональные экосистемы отличаются по своим пространственным масштабам. При увеличении пространственных размеров от локального A_{loc} к региональному A_{reg} масштабам в пределах одной и той же территории число видов может возрастать согласно степенной зависимости $S = cA^z$ между площадью описания и количеством выявленных видов. Следовательно, для произвольной области i логарифм отношения локальных и региональных видовых богатств может быть связан с логарифмом отношения их площадей линейной зависимостью (рис. 2.3а)

$$\log(S_{reg\text{-}i} / S_{loc\text{-}i}) = z_i \log(A_{reg\text{-}i}/A_{loc\text{-}i}), \quad \text{где } z_i \approx 0{,}25.$$

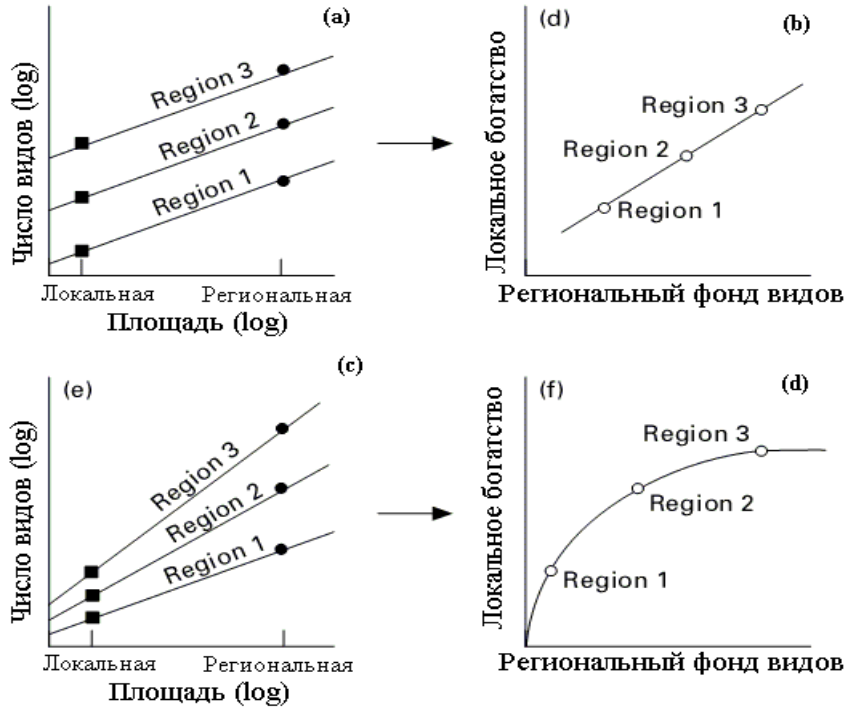


Рис. 2.3. **В ненасыщенном сообществе параллельные линии соединяют локальное (квадраты) и региональное богатство (круги) каждой территории (Region 1-3) на графике логарифмической зависимости числа видов от площади (a).**
Эти же территории показаны на прямолинейной зависимости между локальным и региональным богатством, характерным для ненасыщенного сообщества (b). Аналогичные графики (c) и (d) представлены для насыщенных сообществ (Srivastava, 1999).

Если предположить, что размеры $A_{reg\text{-}i}$ и $A_{loc\text{-}i}$ – одни и те же для каждой i-й территории, то соотношение S_{reg} / S_{loc} будет одинаковым при постоянном значении показателя z_i для каждой области. Другими словами, у ненасыщенных сообществ должны быть параллельные прямые зависимости между величиной площади описания и количеством выявленных видов в логарифмической шкале (рис. 2.3a, b), но эти линии могут отличаться по отрезку, отсекаемому на оси ординат, что соответствует коэффициенту c в выражении $S = cA^z$. Насыщенные сообщества будут отличаться по величине z между территориями (z увеличивается с региональным богатством), но имеют минимальные различия в значениях c (рис. 2.3c, d).

Точная оценка размера полного видового богатства TSR и его локальных подмножеств LSR для отдельных изучаемых сообществ является важнейшей отправной точкой для определения относительной значимости различных экологических механизмов, таких как колонизация, рассеивание, интенсивность обмена видами, влияние условий окружающей среды, биотические взаимодействия и стохастические события.

Г. Корнелл и Д. Лотон (Cornell, Lawton, 1992) сделали предположение, что влияние размера видового фонда сообщества на его локальное видовое богатство различно у разных типов сообществ. По этому признаку они выделяют типы сообществ с интенсивными межвидовыми взаимодействиями (interactive community) и неинтенсивными взаимодействиями (non-interactive community). В формировании первых значительную роль играют межпопуляционные взаимодействия, а скорость иммиграции новых видов зависит от числа уже присутствующих в них. У вторых влияние этого фактора ограниченно и динамика численности популяций каждого вида не зависит от других видов. Колонизация этих сообществ новыми видами является случайным процессом, а её скорость не зависит от числа присутствующих видов.

В общем случае вероятность $p(CR)$ принадлежности того или иного вида к рассматриваемому сообществу TC (target community) может быть представлено функцией (Zobel et al., 1998)

$$p(CR) = \mathrm{f}\,(ES, SD, FE, FA, DC, DM, GE, DB),$$

где ES – экологическое (средовое, ценотическое и функциональное) подобие вида и TC; SD – обратная величина расстояния между его ближайшими популяциями и TC; FE – фертильность среды; FA – частота встречаемости или обилие; DC – способность к рассеиванию; DM – фактическое наличие механизмов рассеивания (объектов транспортировки); GE – уровень адаптации вновь прибывающих видов; DB – жизнеспособность особей (семян, куколок и проч.) в «банке диаспор» TC. Не вполне ясно, как составляющие этой функции соотносятся друг с другом, но любой член уравнения с малой вероятностью приведет к низкой полной вероятности $p(CR)$ принадлежности того или иного вида к сообществу TC (*правило лимитирующего фактора*).

Ж. Бутей с соавторами (Butaye et al., 2002) на примере лесных сообществ выполнил сравнительный анализ двух основных ограничений, обуславливающих формирование богатства видов и связанных как с процессами миграции, так и с экологическими факторами – (см. рис. 2.1). Для того, чтобы рассчитать долю видов, которую можно было бы признать потенциальными мигрантами, использовалась имитационная модель Монте-Карло. При этом проверялась гипотеза о том, что видовой состав в недавно сформированных делянках определен видовым составом более старших по возрасту соседних фрагментов леса, против альтернативы, что он является случайной выборкой видов, характерных для изучаемого региона в целом. Для этого подсчитывалось число совпадений видов в каждом анализируемом участке и фрагментах, расположенных в некоторой его δ-окрестности ($\delta \le 1000$ м). В ходе имитации фрагменты мета-сообществ переставлялись в пространстве случайным образом 5000 раз, чтобы воспроизвести плотность распределения вероятности наличия каждого вида на каждой делянке. Если наблюдаемая средняя встречаемость находилась за

пределами правого хвоста распределения (одностороннее испытание при $\alpha = 0{,}05$), полученного в соответствии с процедурой Монте-Карло, то предполагалось, что появление вида на участке обусловлено не только случайными факторами (т. е. достоверна возможность иммиграции вида в тестируемое сообщество из сопряженных территорий). Показано, что изучаемые делянки содержат в среднем 91% видов, присутствующих в смежных фрагментах леса, находящихся в радиусе 1000 м. Это число уменьшается до 72% и 47% соответственно, если рассматривать δ-окрестности 500 и 100 м. С экологической точки зрения проведенное имитационное исследование ясно показывает, что видовой состав на тестируемых участках не является случайной выборкой таксонов, присущих исследуемому региону, а во многом определяется географической конфигурацией смежных пространственных фрагментов в пределах критического расстояния и возможностями видовой миграции (т. е. не столько региональным, сколько локальным пулом видов).

В этой же работе другая составляющая локального богатства, представленная на рис. 2.1 как «обусловленная подходящими факторами среды», оценивалась по следующим четырем параметрам: освещенность, влажность почвы, pH и наличие питательных веществ. Если эти четыре показателя для участка входили в четырехмерный гиперкуб, ограниченный минимальными и максимальными индикаторными значениями из таблиц Г. Элленберга с соавторами (Ellenberg et al., 1991), то вид-претендент считался принадлежностью тестируемого сообщества.

О. Эриксон (Eriksson, 1993) с учетом модели метапопуляционной динамики Р. Левинса (Levins, 1969) предложил простое уравнение для скорости изменения числа видов в локальном сообществе (S), отражающее результат соотношения двух процессов:

- скорости колонизации $c(S_{reg} - S)$, которая пропорциональна числу видов в видовом фонде S_{reg} за вычетом числа видов, уже присутствующих в сообществе (т. е. числу видов – потенциальных колонизаторов),
- и локальной скорости вымирания, пропорциональной только числу видов, присутствующих в сообществе.

Тогда

$$dS / dt = c(S_{reg} - S) - eS,$$

где c и e – константы колонизации и вымирания видов соответственно, зависящие от особенностей сообщества. В соответствии с этим

$S^* = c\,S_{reg} / (c + e)$, где S^* – равновесное число видов в сообществе.

Отсюда возможны следующие предположения:

- если скорость локального вымирания очень низка, то S^* будет приближаться к S_{reg};
- если $c = e$, то $S^* = S_{reg}/2$;
- если $e > c$, то $S^* < S_{reg}/2$.

Однако главным следствием гипотезы видового фонда, по мнению Эриксона, является вывод, что варьирование локального видового богатства сообществ полностью или частично определяется варьированием размера их видовых фондов.

Некоторые исследователи считают, что по ряду причин однозначная интерпретация результатов полевых наблюдений в пользу или против гипотезы о существенном влиянии видового фонда сообществ на локальное богатство является проблематичной. В первую очередь это связано с трудностями оценки размера самого видового фонда. Чаще всего в качестве его оценки используют общее число видов, зарегистрированных в сообществах определенного типа. Однако критический аспект уровня таких работ обычно ссылается на небольшую выборочную достаточность экспериментальных исследований, от которых напрямую и зависит обоснованность научных выводов.

Ряд авторов (Eriksson, 1993; Zobel et al., 1998) предлагает использовать дополнительную информацию о среде и видах (об их морфологии, физиологии, жизненных стратегиях и т. д.), свидетельствующую об их способности сосуществования в некоторых сообществах. По их мнению, можно использовать и результаты экспериментов по ис-

кусственному внедрению видов в сообщества, однако данный подход в полной мере практически не реализуем, а отдельные опыты дали неопределенные результаты (Zobel et al., 1998). По общему признанию все эти способы дают весьма грубую оценку размера видового фонда. В частности, Г. Корнелл (Cornell, 1985) считает, что правильное определение значения этого параметра для отдельных местообитаний является досадной проблемой, которая не имеет иного простого решения, кроме как путем дефаунизации или интродукции.

Другая причина, почему необходимо с осторожностью интерпретировать зависимость между локальным видовым богатством и размером видового фонда, заключается в сложности определения зависимого и независимого параметров (Cornell, Lawton, 1992; Caley, Schluter, 1997). Кроме того, поскольку факторы среды могут оказывать одновременное влияние как на размер пула, так и на локальное богатство, нельзя исключить предположение, что корреляция между значениями этих параметров может быть обусловлена их зависимостью от третьей переменной – условий или емкости среды (Srivastava, 1999; Herben, 2000; Akatov et al., 2005).

Результаты ряда исследований показывают, что видовое богатство сообществ на меньших по площади участках при увеличении размера видового фонда достигает верхнего предела (видовой емкости) быстрее, чем на более крупных (Terborgh, Faaborg, 1980; He et al., 2005). Соответственно, относительная роль исторических и региональных (островной эффект) процессов в формировании видового богатства должна зависеть от масштаба оценки феномена: чем меньше рассматриваемый участок сообщества или масштаб пула видов (региональный, локальный, актуальный), тем в меньшей степени их видовое богатство определяется этими процессами и в большей – условиями среды. Поэтому, по мнению В.В. Акатова, характер соотношения между S_{loc} и S_{reg} можно интерпретировать не только с позиции полночленности сообществ на локальных участках, но и полночленности их видовых пулов: чем выше значения соотношения $S_{\text{reg}}/S_{\text{loc}}$, тем выше уровень насыщенности S_{reg} (Акатов и др., 2002, 2003). Что касается видовой полночленности самих сообществ, то её предложено оценивать на основе того же подхода – через соотношение локального богатства на участках большего и меньшего размеров. На примере субальпийских фитоценозов изолированных высокогорных массивов Западного Кавказа были получены достоверные корреляционные зависимости данного показателя от площади и видовой емкости этих сообществ, что было интерпретировано, как признак проявления островного эффекта (Акатов, 1999).

Н. Муке с соавторами (Mouquet et al., 2003) обратил внимание на то, что соотношение между локальным и региональным богатством изменяется последовательно по мере эволюции сообществ в результате межвидовых взаимодействий. Сообщество является перенасыщенным на ранних стадиях его формирования, потому что идет активный процесс колонизации видами регионального пула. На промежуточных этапах формирования сообщества происходит существенное конкурентное исключение видов. Наконец, когда сообщества достигают равновесия, его насыщенность ограничивается результатом соревнования видов за лимитирующие ресурсы (если в принципе существует ресурс, влияющий на ограниченность видовой насыщенности). Авторами (Mouquet et al., 2003) была разработана простая теоретическая модель соревнования видов, позволяющая оценить размер регионального пула и обилие видов во времени с момента начала колонизации.

Пусть S_{reg} – число видов регионального пула, а S_{loc} – количество локально сосуществующих видов, обилие которых равно N_{i}, $i = 1, 2, \ldots, S_{\text{loc}}$. Вероятность P_{i} присутствия i-го вида в сообществе удовлетворяет условиям распределения Пуассона и является функцией обилия N_{i} и фертильности среды F (одинаковой для всех видов):

$$P_i = 1 - \exp(-FN_i / \sum_{k=1}^{S_{loc}} N_k) .$$

Вероятность того, что особь i-го вида выиграет соревнование за единицу ресурса, зависит от P_i и от вероятности C_{ij} того, что данный вид является лучшим конкурентом для единицы ресурса H: $C_{ij} = 1 - |H_j - R_i|$, где j – градация ресурса; R_i – параметр соревнования, выбранный случайно на равномерной шкале между 0 и 1. Описаны результаты реализации модели с использованием имитационных методов. На рис. 2.4 представлен график динамики видовой насыщенности локальных сообществ при различных размерах регионального пула. Горизонтальная линия обозначает максимальное число видов (S_{reg} = 20), которое может сосуществовать в равновесии весь период развития сообщества (т. е. без конкуренции за ресурсы). Вертикальная линия обозначает число лет, за которые в экосистеме достигается максимально возможное количество локально сосуществующих видов.

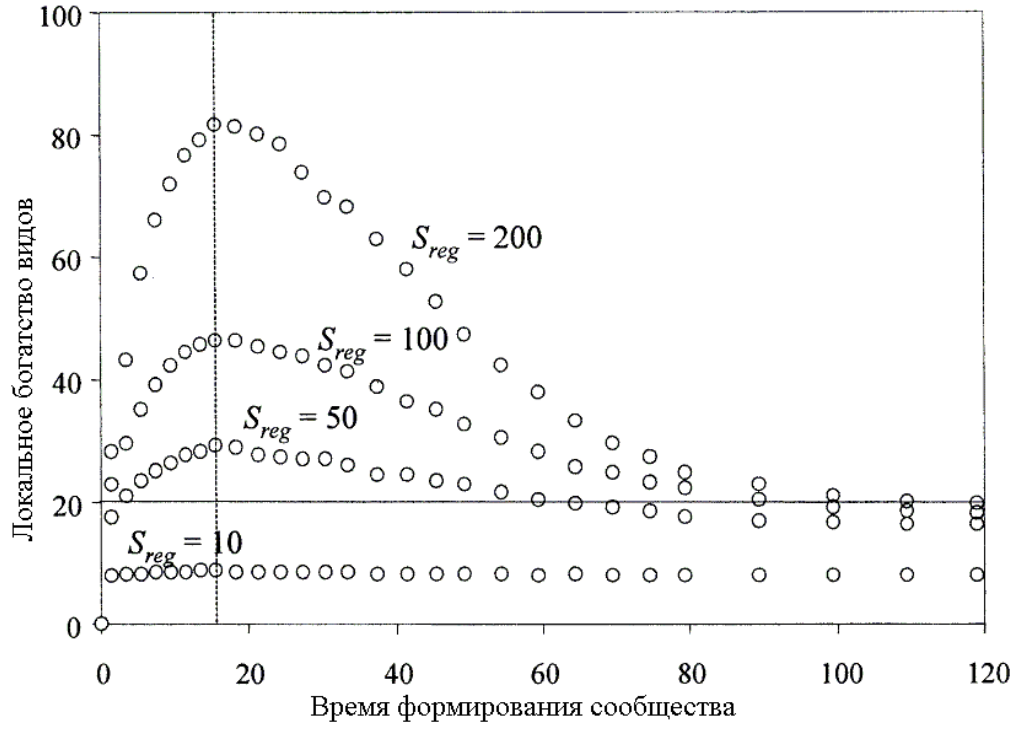


Рис. 2.4. **Число видов, сосуществующих локально в разные периоды формирования сообществ, в зависимости от размера регионального фонда S_{reg} (по: Mouquet et al., 2003)**

Подводя итог изложенному, мы хотим обратить внимание, что понятие пула видов используется главным образом как теоретическая конструкция для того, чтобы объяснить богатство видов в различных пространственных масштабах. Основная идея заключается в том, что количество видов в сообществе или пуле меньшего масштаба определяется прежде всего числом доступных видов в пуле большего масштаба, следующего по порядку (Zobel, 1997). Несмотря на неопределенность результатов тестирования этой гипотезы, перечисленные методы анализа пулов видов – потенциально полезный инструмент для оценки степени видовой насыщенности сообществ (Srivastava, 1999).

Однако работы, описанные выше, убедительно показывают, что эти методы должны использоваться осторожно и параллельно с учетом других факторов, таких как присутствие или отсутствие конкурентного исключения, ограничения на ресурсы, компенсация неоднородности популяционной плотности, оценка инвазии видов из внешней среды и др.

2.2. Оценка видового богатства: исходные данные, задачи, подходы, решения

Изучение видового состава сообществ проводится главным образом для решения трех задач:

- выявление пространственно-временной неоднородности распределения популяционной плотности и изменений структуры сообщества относительно экологических факторов;
- оценка биологической целостности (Karr, 1991);
- разработка мер по сохранению биоразнообразия (Allan, Flecker, 1993).

Массовые явления антропогенного преобразования естественной среды, тенденции глобального изменения климата, угроза исчезновения многих биологических видов сделали экстраполяцию и предсказание динамики экосистем решающим компонентом многих планов научных исследований. Асимптотические подходы к анализу богатства видов могут быть особенно важными для оценки биотических запасов или восполнения материальных ресурсов, а также в случаях, когда даже интенсивные исследования не дают возможности оценить *истинное* биоразнообразие (пример – тропические растительные и беспозвоночные сообщества; сообщества мирового океана и др.).

Список видов и численность их популяций в некоторых редких случаях могут быть определены путем полного пересчета (Palmer, 1990), однако, как правило, локальное (и, тем более, региональное) видовое богатство оценивается на основе результатов мониторинга в формате последовательности повторяемых выборок, начиная с того момента, как эти данные становятся доступными для обработки (Miller, White, 1986). Выборочный подход предоставляет наиболее ценную и достоверную информацию о региональном разнообразии и его долгосрочных тенденциях, однако достоверность выводов существенно зависит от репрезентативности данных, получаемых в ходе полевых исследований. Идеальным результатом изученности видовой структуры сообществ является её детализация до уровня функций плотности распределения обилия видов (Gauch, 1982) с учетом их пространственно-временной неоднородности. Этот подход практичен, если видовое богатство относительно невелико, а природная среда является относительно гомогенной. Однако территориальное распределение плотности популяций, как правило, весьма неоднородно и поверхности аппроксимирующих функций часто не могут быть восстановлены на основе практически разумного числа выполненных наблюдений. Поэтому наиболее распространенным альтернативным подходом является разработка теоретических моделей или установление эмпирических отношений на основе экстенсивного изучения ограниченного числа участков, после чего наблюдаемые соотношения экстраполируются на значительно более обширные территории (Lyons, 1992). Правда здесь всегда возникает вопрос, насколько выводы из полученной эмпирическим путем частной информации способны отражать общие закономерности экосистемы (Stohlgren et al., 1995). К сожалению, на практике этот аспект часто выпадает из поля зрения, как на этапе планирования полевых работ, так и на этапе обработки результатов.

Перед началом мониторинговых исследований в некоторой ограниченной области площадью A (или водотоке длиной A) количество видов в сообществе S и характер распределения совокупности организмов N являются фактически неизвестными параметрами, которые должны быть оценены. Эффективные схемы планирования выборочных исследований в экологии значительно отличаются от традиционных в демографической статистике, и их выбор связан, главным образом, с методическими и практическими рекомендациями, основанными на природе изучаемых сообществ (например, *планы Биттерлиха* или *выборки по линейным трансектам*). Однако в любом случае корректная статистическая оценка обилия видов и других показателей разнообразия предполагает условие, что особи были отобраны из сообщества путем того или иного рандомизированного плана с

повторностями (random sampling with replacement). Ряд авторов (например, De Vries, 1986; Tompson, 1992) попытались соединить некоторые из этих планов с основной теорией выборочного метода или оценить практические преимущества, которые могут быть получены при использовании традиционных схем статистики и планирования эксперимента. Не вдаваясь в подробности каждого из планов взятия проб, принятых экологами, отметим, что все они являются схемами столкновений ES (encounter schemes), т. е. наблюдаемые особи – это те, с которыми исследователь сталкивается в точках, линиях или участках, случайным или целесообразным образом выделенных в области исследования A.

Отступление от темы (*туманно-статистическое*). Общая постановка задачи оценки видового богатства с использованием конкретного плана ES может быть сделана следующим образом (Fattorini, Marcheselli, 2002). Пусть каждая особь сообщества $y_1, \ldots, y_N$ соотнесена со значениями $e_1, \ldots, e_s$ единичного вектора $\mathbf{E}^s$, т. е. $y_j = e_i$, если особь j принадлежит виду i. В этой системе обозначений вектор обилия сообщества может быть выражен как $\mathbf{H} = \sum_{j=1}^{N} y_j$, что позволяет свести расчет N к стандартной задаче оценки объема конечной выборки. Обозначим как $\pi_1, \ldots, \pi_N$ вероятности того, что j-я особь при данном варианте обследования будет выбрана первой. Если план взятия проб позволяет выполнить оценку этих вероятностей (по крайней мере, для уже выбранных особей), то несмещенную оценку вектора $\mathbf{N}$ с ковариационной матрицей $\mathbf{\Sigma}$ можно найти по формуле Горвитца–Томпсона $\hat{\mathbf{N}} = \sum_{j \subset S} y_j / \pi_j$ (Thompson, 1992). Значения $\hat{\mathbf{N}}$ и $\mathbf{\Sigma}$ в свою очередь зависят от вероятностей того, что каждая j-я особь будет выбрана второй, третьей и т. д. для всей последовательности π_{jh} ($h > j = 1, \ldots, N$). Вероятности π_{jh} определяются характеристиками экологического сообщества (в частности, территориальным распределением особей по области исследования).

Если рассмотреть саму возможность получения несмещенной оценки $\mathbf{\Sigma}$, то известно, что она достижима только при $\pi_{jh} > 0$ для любых $h > j$. Таким образом, любая оценка вектора $\mathbf{N}$ будет смещенной, если имеется единственная выборка ES, либо особи разных видов размещены очень далеко друг от друга и не могут столкнуться совместно на одной и той же площадке, точке или линии. Как подчеркнуто (Barabesi, Fattorini, 1998), эти проблемы разрешаются путем многократного повторения планов ES. Если выборки наблюдений $R_1, \ldots, R_m$ независимо повторяются во времени и рандомизируются в пространстве, тогда оценки $\hat{\mathbf{N}}_1, \ldots, \hat{\mathbf{N}}_m$ являются реализацией случайного вектора с математическим ожиданием S и ковариационной матрицей $\mathbf{\Sigma}$, а их средние являются наилучшими оценками обилия сообщества $\hat{\mathbf{N}}$.

Точная, несмещенная и адекватная оценка перечисленных показателей по мониторинговым данным связана с непростыми самостоятельными проблемами, активно обсуждаемыми в конце XX в. (Thompson, Seber, 1996; Brown, Manly, 1998). В любом случае, очевидно, что значения N_i и размерность вектора S напрямую зависят от интенсивности выборочных усилий и являются сложной функцией принятого масштаба, пространственных координат, ограничивающих сообщество, периода наблюдений, способа подсчета численностей и т. д.

Моделирование видовой структуры сообществ обычно требует объединенного использования двух типов моделей. Первый тип описывает, как особи распределены по видам (species-abundance distribution). Наиболее известными моделями распределения обилия видов являются логарифмические серии и лог-нормальное распределение (Fisher et al., 1943; Preston, 1948; Hubbell, 2001). Второй тип моделей связан с тем, как особи различных

видов распределены в пределах их естественной среды (т. е. по пространственным участкам или выборкам иного характера). Традиционными моделями размещения видов (species-occurrence distribution) является распределение Пуассона и отрицательное биномиальное распределение (Gibson et al., 1999; Williams et al., 2001). Однако любые модели видовой структуры, а также большинство индексов и метрик разнообразия напрямую опираются на достоверность такого фундаментального (но трудно оцениваемого) выборочного параметра, как число видов S, «без которого они не имеют твердой статистической основы» (Magurran, 2004).

Как много видов может быть обнаружено в некоторой ограниченной области или водном объекте? Поиск ответа на этот вопрос и обсуждение наилучших зависимостей числа выявленных видов от масштаба исследования продолжается уже полтора столетия (de Candolle, 1855; Jaccard, 1902, 1908; цит. по: He, Legendre, 1996). Сформулированное общее правило – *наблюдаемое видовое богатство сначала монотонно возрастает при увеличении объема использованного экспериментального материала, а потом постепенно стремится к некоторому порогу насыщения* – было названо «одним из немногих подлинных законов экологии сообществ» (Schoener, 1976, p. 629).

Э. Тьорв (Tjørve, 2003) выполнил всестороннее обобщение используемых 14 моделей SAR (Species Area Relationship) зависимости числа видов от площади, включая их типологию и основные характеристики. Он различал выпуклые модели, не стремящиеся к насыщению (2.1 и 2.2. в табл. 2.1), и сигмоидальные модели, приближающиеся к верхней асимптоте (2.5 и 2.6). Ю. Денглер (Dengler, 2008) расширил этот список до 23 моделей, хотя отметил, что большинство из них имеет исключительно теоретический интерес.

Таблица 2.1

Основные модели зависимости числа видов от площади SAR

№№ формул	Название функции	Дифференциальная форма зависимости (He, Legendre, 1996)	Модель *)	Источник
2.1	Логарифмическая **)	Постоянная $dS / dA = \alpha / A$	$S = c + z \lg(A)$	Gleason, 1922
2.2	Степенная	Линейная $dS / dA = (\alpha + \beta S) / A$	$S = cA^z$ ($\lg S = c + z \lg A$)	Уотсон, 1859 ***) Arrhenius, 1921; Preston, 1962
2.3	Линейная		$S = c + zA$	Connor, McCoy, 1979
2.4	Экспоненциальная		$S = ce^{zA}$ ($\lg S = c + z A$)	Connor, McCoy, 1979
2.5	Отрицательная экспоненциальная		$S = c(1 - e^{-zA})$	Holdridge et al., 1971
2.6	Логистическая	Параболическая $dS / dA = (\beta S + \gamma S^2) / A$	$S = \frac{b}{c + A^z}$	Archibald, 1949
2.7	Гиперболическая		$S = \frac{bA}{c + A}$	Monod, 1950; de Caprariis et al., 1976

*) z, c и b – настраиваемые коэффициенты моделей, связанные с плотностью распределения организмов по рассматриваемой территории;

**) Ф. Хи и П. Лежандр ошибочно назвали её экспоненциальной, что было отмечено Ю. Денглером (Dengler, 2008);

***) подробнее см. (Гелашвили и др., 2007а).

Большинство приведенных моделей SAR являются частным случаем обобщенного дифференциального уравнения $dS / dA = f(S) / A$, т. е. интенсивность обнаружения новых видов (species detectabilities) зависит от числа уже найденных видов S на единицу площади A. Ф. Хи и П. Лежандр (He, Legendre, 1996) показали, что, используя разные предположения относительно вида функции $f(S)$, можно придти к двум самым популярным формам интегральных моделей: если принять $f(S)$, равной постоянной α, то легко получить логарифмическую зависимость (2.1), а линейная функция для $f(S)$ приводит к степенному уравнению (2.2).

Последние 80 лет не прекращались дискуссии о том, имеет ли степенная модель Уотсона–Аррениуса (2.2) характер универсального закона. Здесь есть два важных аспекта. Во-первых, все выражения (2.1-2.4) представляют собой некоторое семейство математически трансформируемых друг в друга функций: настройка параметров моделей с использованием какой-то пары координат $\{S, \lg S\} – \{A, \lg A\}$ предполагает взаимно однозначное соответствие в пространстве любой другой пары размерностей (Dengler, 2008, 2009). Во-вторых, специальный теоретический подтекст обычно приписывается нелинейной форме степенной зависимости (2.2), но для удобства вычислений ее приводят к модели, линейной в логарифмах. Однако, хотя обе эти зависимости математически эквивалентны, они дают различные оценки коэффициентов модели и детерминации R^2, поскольку при аппроксимации степенная функция использует аддитивную сумму квадратов отклонений S, а линеаризованная версия – произведение квадратов отклонений, т. е. мультипликативную форму. Иными словами, *с точки зрения регрессионного анализа – это два разных уравнения, коэффициенты которых могут серьезно отличаться*.

Поэтому налицо очевидное противоречие между теоретическими изысками и методами практических вычислений. Особые свойства, связанные с масштабными характеристиками области исследования и представленные в многочисленных публикациях, приписываются параметру модели z, повсеместно рассчитываемому по линейной в логарифмах форме. В то же время его связь с параметрами канонического лог-нормального распределения обилия видов (Preston, 1948, 1962) интерпретируется уже на основе степенной функции. Впоследствии ряд исследователей (Harte et al., 1999; Гелашвили и др., 2007б) обратили внимание на фрактальный характер структуры биологических сообществ и теоретически обосновали коэффициент z как фрактальную размерность, снова обратившись к степенной зависимости. Думается, необходимо все же окончательно определиться с тем, какая из двух форм модели (2.2) имеется в виду.

Выпуклые кривые (2.1-2.2) достаточно хорошо описывают интенсивный рост числа видов для малых и средних масштабов. В частности, Э. Коннор и Э. МакКой (Connor, McCoy, 1979) показали, что из 100 ценотических исследований формула (2.2) обеспечивает разумно пригодное приближение данных в 75 случаях, по сравнению с 47, 38 и 22 случаями для формул (2.3), (2.1) и (2.4). Однако в случае оценки числа древесных пород в зависимости от большого числа обследованных участков (Holdridge et al., 1971) предпочтительней оказалась отрицательная экспоненциальная модель (2.5). Формально говоря, *степень приближения к эмпирическим данным нельзя считать единственным определяющим условием адекватности математической модели*. Для достижения этого можно, например, прибавить к степенной функции дополнительные члены более высокого порядка $S = c A^z e^{P(A)}$, где $P(A)$ – бесконечный многочлен, «усекаемый» при достижении заданного качества аппроксимации (Plotkin et al., 2000).

Ряд недавних работ показал, что для больших территорий теоретически более обоснованы сигмоидальные модели, которые точнее отражают асимптотический характер кривой накопления видов, стремящейся к порогу насыщения. Например, логистическая модель (2.6) предполагает параболический характер функции $f(S)$, и, следовательно, интенсивность увеличения числа видов после достижения некоторого $A_{крит}$, начинает падать,

приближаясь к 0. Близка к логистической модели и двухпараметрическая гипербола (2.7), широко известная в статистической и биохимической литературе как *уравнение ферментативной кинетики Михаэлиса–Ментен*. Этим, видимо, объясняется тот факт, что различными авторами были предложены, по крайней мере, шесть различных модификаций этой модели для оценки численности видов (Raaijmakers, 1987).

Важный методологический вопрос связан с дискуссией, следует ли различать процесс накопления видового богатства при увеличении изученной площади A (модели SAR) и аналогичный эффект при росте любого другого выборочного усилия, т. е. общего количества проб разного характера, взятых в пространстве и времени (модели SSR – species sampling relationships). С. Шейнер (Scheiner, 2003) выделил четыре главных типа структурной организации территории, относящихся к SAR:

I. Непрерывные экосистемы с вложенными участками (т. е. участки меньшего размера являются заранее предопределенной частью более крупных участков), причем каждая точка пространства вносит самостоятельный и независимый вклад при подсчете S.

II. Экосистемы с вложенными участками, но объединение данных при их обобщении осуществляется с использованием некоторого алгоритма (например, расчет среднего богатства видов на каждом участке).

III. Фрагментированные экосистемы с невложенными участками индивидуального размера, прилегающими друг к другу (например, страны).

IV. Изолированные острова различной формы и размера.

Появление этой классификации вызвало многочисленные исправления и альтернативные терминологические конструкции, однако в итоге участники дискуссии оказались в затруднении четко разграничить модели SAR и SSR. Например, Ю. Денглер (Dengler, 2008) считает, что для SSR эмпирическая выборка может быть выполнена на участках из произвольных областей, но только в SAR эти участки и ограничивающие их области должны быть смежными (т. е. четко проявляются биологические процессы рассеивания или миграции видов). В другой работе (Gray et al., 2004), напротив, утверждается, что к SAR относятся только модели, относящиеся к территориям типа IV.

К моделям SSR участники дискуссии единодушно отнесли «кривые аккумуляции или разрежения видов», определив их существенное «математическое отличие» от моделей SAR. Поскольку никакой конкретной аргументации приведено не было, нам это дистанцирование, как и необходимость разграничивать модели SAR и SSR, представляется сомнительным.

Модели SAR и SSR является достаточно грубым приближением, не учитывающим популяционную или пространственную неоднородность сообществ, область допустимых значений A, динамику экосистем и многие другие важные факторы. Последовательное объединение подмножества A выборок однократных наблюдений (например, серии гидробиологических проб) дает эмпирическую кривую $S(A)$ последовательного *накопления* числа видов (species accumulation curve), которая не является гладкой из-за пространственно-временной неоднородности сообществ или воздействия случайных факторов.

На рис. 2.5а представлена такая кривая накопления для сообществ макрозообентоса 43 малых рек, протекающих по территории Самарской области, в хронологическом порядке по результатам 628 гидробиологических проб за период с июля 1985 г. по июнь 2007 г. Можно отметить, что, несмотря на столь большой экспериментальный материал, предполагаемый процесс асимптотического видового насыщения пока не обнаруживается. Возможным объяснением этому может явиться эффект видовой неоднородности за счет разнообразия ландшафтных особенностей территории или различий в масштабной размерности и гидрологических характеристиках исследованных водотоков, т. е. что соответствует β- и γ-разнообразию крупномасштабных экосистем.

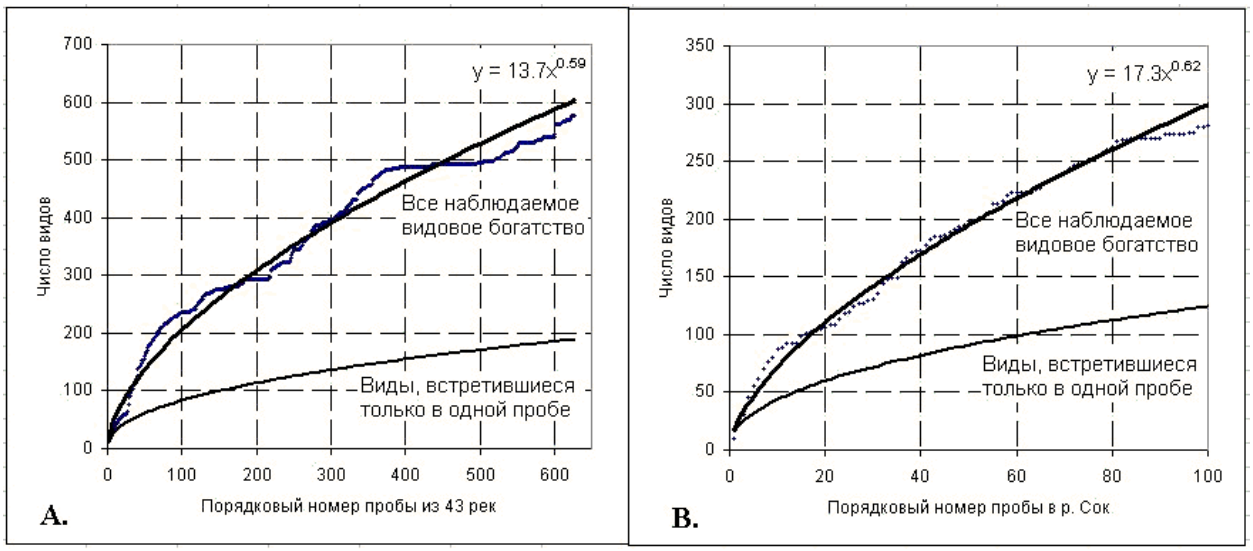


Рис. 2.5. **Кривые накопления видов макрозообентоса по результатам гидробиологических исследований малых рек Самарской области за 1985-2007 гг.**
А. – данные исследований 43 водотоков, В. – р. Сок

Однако если рассматривать функцию накопления видов макрозообентоса от выборочного усилия только в одном из водотоков – р. Сок, то общий характер кривой (рис. 2.5b) практически не меняется. Можно также отметить высокую долю в наблюдаемом богатстве редких видов, которые встретились только в одной из проб, хотя при росте выборочного усилия эта доля закономерно уменьшается. Заметим, что аппроксимация данных степенной зависимостью не вполне точно описывает тенденцию увеличения числа видов: в обоих случаях простой полином 2-й степени дает лучшие результаты приближения к экспериментальным данным.

Отсутствие асимптотического приближения кумулятивных кривых на рис. 2.5 к какому-то порогу видовой насыщенности является интересным экологическим феноменом и нуждается в объяснении. Можно, например, высказать общее предположение (нуждающееся, разумеется, во всесторонней проверке), что, по сравнению с растительными или другими наземными сообществами, водные экосистемы обладают более высокой таксономической емкостью. Косвенно это можно аргументировать интенсивностью факторов популяционной динамики: периодической суточной или сезонной сменой ассоциаций видов-гидробионтов, широким спектром питания, разветвленностью и поливариантностью трофических связей, миграционной активностью, высокой адаптационной способностью к различным экологическим факторам и т.д. (Шитиков и др., 2010).

В этом контексте достойна упоминания работа (Muneepeerakul et al., 2007), выполненная на примере сообществ водной растительности. Авторами сделан вывод, что показатели видового разнообразия речных систем лишь косвенно связаны с общей площадью изученного региона (в двух формальных измерениях). Здесь более существенное значение имеют градиенты направленности путей миграции и размножения, обусловленные наличием течения, а также общая топологическая структура пространства водотоков. В частности, речные системы характеризуются более высокими показателями β- и γ-разнообразия водной растительности в силу «эффекта локализации», возникающего в результате пространственной разобщенности речных притоков. Таким образом, *в разном пространственном масштабе и в разных сообществах могут действовать различные механизмы*, влияющие на показатели видового разнообразия.

Практическим приложением моделей SSR является оценка числа потенциально «скрытых» (т. е. еще не обнаруженных) видов на основе анализа кумулятивной кривой. При этом эмпирические данные о числе уже найденных видов при некотором произвольном выборочном усилии A сглаживаются той или иной параметрической модельной зависимостью с последующей экстраполяцией к некоторой асимптоте «насыщения». Использование методов прикладной статистики произвело большое количество непараметрических подходов к прогнозированию различных показателей видового разнообразия (см. обзор: Bunge, Fitzpatrick, 1993), где богатство видов – один из примеров. Как одну из первых работ в этом направлении можно упомянуть «кривую случайного размещения» (random placement curve) Б. Коулмана (Coleman, 1981), при построении которой предполагается, что любая особь из общего количества $\overline{Y}\,m$ (где $\overline{Y}$ – среднее число особей в одной эмпирической выборке, m – число площадок) может быть случайным образом размещена на любой из площадок.

Н. Готелли и Р. Колуэлл (Gotelli, Colwell, 2001) выделяют четыре типа подходов к моделированию кумулятивных кривых, которые отличаются между собой двумя характерными признаками, или «дихотомиями». Первая «дихотомия» основана на концептуальных отличиях самих задач интерполяции и экстраполяции. При интерполяции строится *кривая разрежения* (Sanders, 1968; Hurlbert 1971; Simberloff, 1978), которая формируется путем многократной перестановки данных внутри выборок из N_{obs} особей, обнаруженных в результате мониторинга, и является функцией математического ожидания видовой насыщенности $S(N)$ при увеличении численности сообщества. Разрежение дает возможность найти предполагаемое число видов S_{est} для любой промежуточной совокупности из N особей $(N_{\text{obs}} > N > 1)$, считая её случайной и независимой выборкой из всей генеральной совокупности. Когда выборочное усилие становится достаточно велико и объем располагаемых данных позволяет говорить о разумно устойчивом приближении к истинным параметрам разнообразия экосистемы, *экстраполяция* становится логической возможностью и предметом статистического анализа. При экстраполяции кривой накопления подсчитывается максимально возможное (т. е. потенциальное) видовое богатство S, лимитируемое, например, емкостью природной среды и превышающее наблюдаемое множество видов S_{obs}.

Вторая «дихотомия» касается выборочной схемы представления и обработки мониторинговых данных. Один подход, основанный на особях (individual-based), заключается в том, что наугад в пределах территории изучаемого сообщества отбирается случайное множество особей, которое последовательно идентифицируется на принадлежность к тому или иному виду. Другой подход (sample-based) предполагает выделение ряда «квадратов», адекватно отражающих всю пространственную совокупность. На этих выборочных площадках устанавливается численность и видовая принадлежность всех особей, после чего оценивается полное видовое богатство. В зависимости от протокола мониторинга, «квадратами», в широком диапазоне исследований, могут быть геоботанические площадки, трансекты, ловушки, приманки и загородки, отдельные гидробиологические пробы, взятые тралами, батометрами, дночерпателями или любыми другими устройствами, препарированные срезы и др. Ключевое различие представленных двух подходов заключается в том, что в первом случае целостной экспериментальной единицей, оперируемой в вычислительном процессе, является особь, а во втором случае – пробы или иные выборки данных.

На рис. 2.6 можно отметить, что две кривые (эмпирическая и сглаженная), основанные на выборках (пробах), лежат ниже двух кривых, основанных на численности особей. Причина этого почти универсального правила в том, что протоколы, основанные на выборках, объединяют в пределах каждой пробы особи, которые являются близко расположенными в пространстве или во временной последовательности. Любая пространственная или временная автокорреляция в размещении видов (т. е. взаимозависимость скоплений групп одинаковых особей) приводит к некоторым закономерностям появления таксонов в

выборках. Иными словами, в результате рандомизированной процедуры «отлова» будет выявлено меньше уникальных видов, если особи объединены в группы, по сравнению со случайной и независимой схемой распределения особей в той же естественной среде.

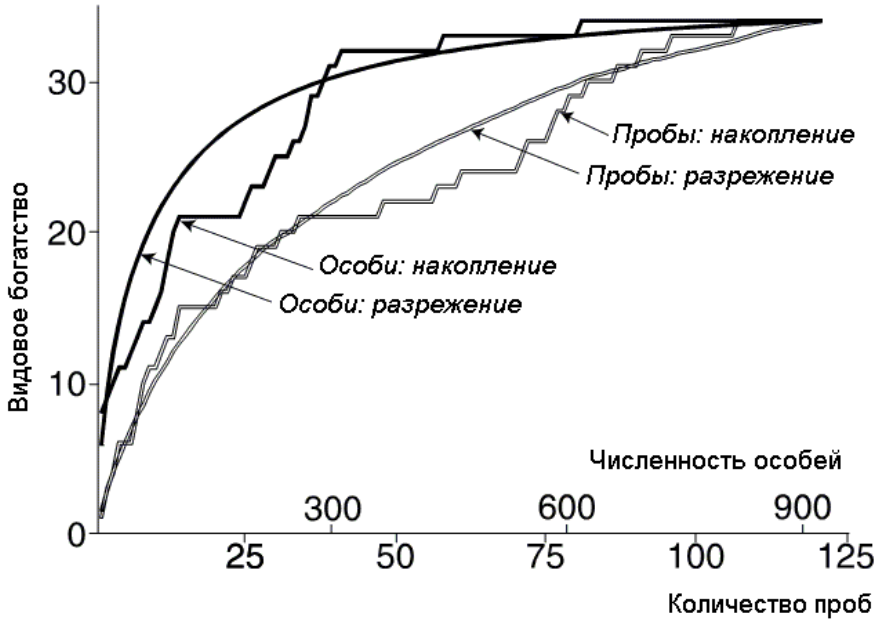


Рис. 2.6. **Эмпирические кривые накопления (ломаные линии) и сглаженные кривые разрежения, являющиеся математическим ожиданием соответствующих кривых накопления (по: Gotelli, Colwell, 2001).**

Кривая разрежения, основанная на численностях особей, получена с использованием *метода Коулмана* (Coleman, 1981). Аппроксимация кривой, основанной на выборках (пробах), осуществлена *методом Колуэла–Мао* (Colwell et al., 2004). Все четыре кривые основаны на данных о биоразнообразии тропической растительности в Коста Рике (Butler, Chazdon, 1998)

В ряде областей экологических исследований тип протокола достаточно понятен (вылов рыбы удочкой или геоботанические описания, с использованием «квадратов»). Однако при изучении водных объектов все не так однозначно. Можно согласиться с мнением Сандерса (Sanders, 1968), что каждая отдельная проба бентоса, взятая дночерпателем, должна рассматриваться как независимая выборка особей, поскольку является как бы массовым захватом хаотично расположенных животных из гомогенной естественной среды. В то же время повторности таких проб, взятых из одного местообитания, могут быть рассмотрены как выборки "квадратов". Поэтому одни и те же наборы данных в зависимости от масштаба исследований или смысла проверяемой гипотезы иногда могут рассматриваться одновременно в разных вариантах статистического анализа: и как подмножества особей, и как совокупности выборок.

Отступление от темы (статистика отловов). Теоретическая проработка конкретных методов оценки популяционного состава сообществ исторически осуществлялась в рамках схемы «повторного отлова» (capture-recapture), широко используемой экологами для исследования динамики биологических популяций, развивающихся в естественной среде. Основная идея этой схемы применительно к полевым исследованиям выражается в следующем:

- во время первой экспедиции исследователь отлавливает (например, с помощью западни) группу особей, отмечает их уникальным идентификатором (кольцом или клеймом), после чего возвращает в естественную среду;
- предоставляется возможность, чтобы отмеченные особи за достаточный период времени могли перераспределить себя среди немаркированной совокупности;
- во время второй экспедиции повторно отлавливается выборка особей, часть которых уже была помечена во время начального посещения и теперь только регистрируются как возвращенные; помечаются еще немаркированные особи и вся выборка снова возвращается в естественную среду;
- размер популяции может быть оценен по результатам двух отловов, но предпочтительнее многократные обследования; данные «истории захвата» для каждой особи анализируются математически, чтобы оценить популяционные параметры, выживание или перемещение.

Другими модификациями этого семейства сходных методик являются: «маркировка с освобождением и возвращением» (mark-release-recapture), «смена вида» (sight-resight), «восстановление полосы» (band recovery) и другие.

Одним из первых модель «отлова с возвращением» использовал известный французский математик П. Лаплас (Amstrup et al., 2005). По записям регистрационных книг общая рождаемость в 1802 г. во Франции составила M = 1000 тыс. младенцев. Лаплас детально обработал данные о численности населения и рождаемости в некоторой небольшой выборке департаментов (где были «рьяные и интеллигентные мэры») и нашел, что на 2037 тыс. жителей за год там приходится 71,9 тыс. рождений (R/C = 0,0353). Тогда оценка общей численности населения Франции будет равна $N = MC/R$ = 1000/0,0353 = 28 330 тыс. человек.

Действия Лапласа предопределили суть модели Линкольна–Петерсена (Lincoln, 1930; Petersen, 1896; цит. по: Otis et al., 1978), используемой для оценки общего размера популяции N по результатам двух отловов:

$$N = MC/R \quad \text{или} \quad N/C = M/R,$$

где M и C – общее число экземпляров, отловленных и промаркированных в первой и во второй выборках соответственно; R – число экземпляров второй выборки, которые были промаркированы и возвращены в первый раз. *Метод Линкольна–Петерсена* соответствует модели закрытой популяции, которая предполагает, что за период обследования вероятности отловов не изменяются, а общая и помеченная численности особей остаются постоянными, т. е. не учитываются факторы смертности, рождения или миграции.

В естественных биологических совокупностях размер популяции динамичен, а неоднородность вероятности отлова может возникнуть разными способами. В этой связи разрабатываются различные модели открытых популяций, учитывающие характер распределения (Chao, 1987; Maunder, 2004) и использующие различные уравнения регрессии (Tanaka, 1967; Cormack, 1989) или функции максимального правдоподобия (Cormack, 1979; Seber, 1982).

Механизмы учета животных в полевых условиях не являются темой нашего обзора, однако формальная методология «отлова с возвращением» нашла широкое распространение в моделях имитационного типа, осуществляющих многократные манипуляции с компьютерными выборками (Chao, 1987). Она предвосхитила появление семейства непараметрических методов экстраполяции полного видового богатства и других показателей биоразнообразия.

2.3. Непараметрические методы экстраполяции полного богатства видов

Как было отмечено выше, в большинстве случаев точная оценка *полной* видовой насыщенности изучаемых сообществ ввиду биологической и пространственно-временной неоднородности природной среды неизбежно потребует дорогостоящих и масштабных исследований. *Наблюдаемое* видовое богатство OSR по мере увеличения выборочного усилия первоначально быстро растет, но затем этот рост замедляется, в той или иной мере асимптотически приближаясь к полному видовому богатству TSR. Однако даже при достаточно большом объеме полученных мониторинговых выборок в числе зарегистрированных видов, вероятно, не будут присутствовать многие редкие виды, которые бы встретились при дальнейшей интенсификации обследований.

В этой связи привлекательной альтернативой является статистическая оценка регионального богатства видов с использованием ограниченного числа единиц повторности наблюдений. Поскольку каждая точка на кривой насыщения располагается ниже линии полного богатства TSR, можно, используя экстраполяционную процедуру, оценить число проб (или выборок иного характера), требуемое для того, чтобы получить ожидаемое число видов, составляющее, например, 90%, 95% или 99% от объема TSR, в зависимости от желаемой выборочной достаточности. Поскольку простые модели кривых накопления (степенная или экспоненциальная зависимости, модель Михаэлиса–Ментен и др.) часто приводят к смещению оценок в сторону уменьшения, основой современной методологии является использование семейства методов генерации псевдовыборок (resampling).

Я. Цао с соавторами (Cao et al., 2001, 2004, 2007), чтобы оценить уровень неоднородности природной среды и смоделировать выборочную достаточность, обосновали и использовали показатель J_c среднего подобия между видовым составом повторных проб, случайным образом извлеченных из массива экспериментальных данных (т. е. автоподобие). Пусть имеется m повторностей мониторинга, содержащих данные видового состава изучаемого сообщества, причем эти пробы могут быть разделены на два приблизительно равных подмножества, достаточно больших, чтобы сходство или различия между ними носили случайный характер. Коэффициент подобия Жаккара между видовым составом этих двух подмножеств проб может быть вычислен как

$$J = c/(a + b + c),$$

где c – количество видов, общих для обеих серий; a и b – количество видов, уникальных для серии 1 и 2 соответственно. Тогда наблюдаемое видовое богатство TSR_{obs} и среднее число видов SR из двух серий проб будет, соответственно, равно:

$$TSR_{obs} = (a + b + c) \quad \text{и} \quad SR = 0{,}5(a + b) + c.$$

Предположим, что из массива наблюдений многократно (например, 1000 раз) случайным образом формируется по две серии наблюдений и получены средние значения числа видов SR_c и коэффициента Жаккара J_c. На основании этого показатель выборочной достаточности наблюдаемого числа видов может быть записан как

$$\%TSR_{obs} = SR_c / TSR_{obs} = [0{,}5(a + b) + c] / (a + b + c) = 0{,}5 + 0{,}5\, J_c.$$

Если любая случайная выборка содержит полный список видов и для каждой повторности $J = 1$, то авторы (Cao el al., 2001) считают, что наблюдаемое видовое богатство TSR_{obs} отражает неизвестное истинное богатство TSR_{tru} (т. е. $TSR_{obs} = TSR_{tru}$). Теперь предположим, что для каждой пары серий проб $J = 0$. В этом случае каждая выборка таксономически уникальна и $\%TSR_{obs} = 0{,}5$, а $TSR_{tru} > TSR_{obs}$.

Если принять, что $\%TSR_{tru} = SR_c/TSR_{tru}$, то при $J_c < 1$ в реальных условиях $\%TSR_{obs}$ будет превышать $\%TSR_{tru}$, а возможный размах значений $(\%TSR_{obsr} – \%TSR_{tru})$ составляет $0{,}5(1 – J_c)$. Используя предположение о линейности зависимости разности $(\%TSR_{obs} – \%TSR_{tru})$ от J_c, можно рассчитать показатель выборочной достаточности ис-

тинного полного числа видов: $\%TSR_{tru} = \%TSR_{obs} - 0{,}5(1 - J_c) = J_c$, а значение $\%TSR_{obs}$ интерпретировать как несмещенную оценку $\%TSR_{tru}$.

Для проверки своей методологии Я. Цао с соавторами (Cao el al., 2001) провели имитационный эксперимент с использованием данных о речных рыбных сообществах в ряде штатов США. Из общего множества выборок случайным образом извлекались серии по две, четыре, шесть и т. д. проб, после чего каждая серия делилась пополам и рассчитывался коэффициент видового сходства J между этими «половинками». Этот процесс повторялся 1000 раз для каждого объема выборки, рассчитывались средние значения J_c, на основании чего были построены кривые накопления видов в зависимости от выборочного усилия и осуществлена калибровка значений $\%TSR_{obs}$.

В последующей публикации (Cao el al., 2004) авторы проводят аналогию предлагаемого ими алгоритма, получившего название CY-1, с моделью «повторного отлова» Линкольна–Петерсена (Otis el al., 1978), поскольку $TSR = SR_1 \cdot SR_2/c$, где SR_1 и SR_2 – число видов в выборках 1 и 2, а c – число общих видов. Однако алгоритм CY-1 предусматривает разную вероятность обнаружения (detectability) для видов-доминантов и редких видов, чем он и отличается от модели Линкольна–Петерсена, принимающей равную вероятность. Цао с соавторами приводят интересные данные о верификации предлагаемого метода на двух принципиально отличающихся сообществах (птицы и рыбы) двух крупных регионов (Восток и Запад США). Выполнен анализ работоспособности алгоритма CY-1 по сравнению с другими аналогичными методами (bootstrap, jackknife, ICE, ACE), описываемыми ниже.

С.Г. Чефранов (2002) использовал процедуры имитационного моделирования для оценки параметров статистического распределения видового богатства растительных сообществ субальпийских лугов Западного Кавказа. Пусть $S(j)$ – множество видов, обнаруженных на j-й площадке, где были выполнены геоботанические описания, $j = 1, ..., m$, а S_m – общее число видов, зафиксированное на всех площадках. Очевидно, что величина S_m зависит от конкретной выборки, носит вероятностный характер и, следовательно, может характеризоваться некоторым законом (плотностью) распределения. Его основные параметры определялись следующим образом.

Пусть x_i – событие, заключающееся в наличии i-го вида на произвольной площадке (в этом случае $x_i = 1$, при отсутствии вида $x_i = 0$), а p_i – вероятность данного события, $i = 1, ..., S$. Аналогично можно обозначить $X_i(m)$ как событие обнаружения вида хотя бы на одном из m участков. Если предположить равенство вероятностей встречаемости видов p_i, то вероятность этого события равна $P_i(m) = 1 - (1 - p_i)^m$.

Общее число видов, обнаруженных на m участках, равно сумме случайных величин $X_i(m)$, т. е.

$$S_m = \sum_{i=1}^{m} X_i(m),$$

математическое ожидание данной величины:

$$M(S_m) = \sum_{i=1}^{m} P_i(m), \qquad (2.8)$$

а дисперсия:

$$D(S_m) = \sum_{i=1}^{m} P_i(m) \cdot (1 - P_i(m)). \qquad (2.9)$$

Последнее равенство справедливо лишь при условии независимости вероятностей колонизации участков разными видами (более жестким является условие отсутствия взаимодействия видов).

Схема определения величины S_m соответствует опыту Бернулли. Выполнение условия $D(S_m) \geq m$ дает основания для применения формулы Муавра–Лапласа в качестве функции плотности распределения величины S_m:

$$f(S_m) = \frac{1}{\sqrt{2\pi \cdot D(S_m)}} e^{\frac{(S_j - M(S_m))^2}{D(S_m)}}, \qquad (2.10)$$

где $M(S_m)$ и $D(S_m)$ – параметры, определяемые по формулам (2.8) и (2.9).

Для проверки представленных соотношений, допущений и ограничений проводился имитационный эксперимент по следующей схеме. По результатам полевых исследований определялись встречаемости p_i, $i = 1, \ldots, S$ для 292 видов сообществ субальпийского пояса Западного Кавказа. Далее генерировалось достаточно большое количество (около 10^5) виртуальных площадок со средней встречаемостью видов, равной p_i. Последовательно определялось число видов по $m = 10$ площадкам и строилась гистограмма распределения числа обнаруженных видов. Подобранные в результате имитации параметры плотности распределения $f(S_m)$ по формуле (2.10) имели значения оценки общего числа видов $M(S_{10}) = 216{,}7$ и его дисперсии $D(S_{10}) = 33$.

Если функция плотности распределения величин регионального видового фонда S_{reg} и локального видового богатства S_{loc} определяется выражением (2.10) с параметрами m_{reg}, σ_{reg} и m_{loc}, σ_{loc}, вычисляемым, соответственно, по формулам (2.8) и (2.9), то плотность распределения соотношения этих двух значений задается законом Коши:

$$f(z) = \frac{\sigma'\sigma''}{\pi\left[(\sigma')^2 + (\sigma'')^2 (z - \frac{m'}{m''})^2\right]}.$$

Последняя формула позволяет находить вероятности отклонения соотношения между видовым фондом (S_{reg}) и локальным видовым богатством (S_{loc}) от среднего значения, определяя тем самым достоверность получаемых оценок.

Наиболее многообещающими методами экстраполяции числа видов являются непараметрические алгоритмы, основанные на генерации псевдовыборок (или ресамплинге – resampling) и позволяющие оценить доверительные интервалы сделанного прогноза. Наиболее известными из них являются бутстреп (bootstrap) и процедуры «складного ножа» (jackknife), успешно развивающиеся последние два десятилетия (Efron, 1979; Эфрон, 1988; Шитиков и др., 2008; см. также выше раздел 1.5).

Впервые алгоритм «складного ножа» для экстраполяции видового богатства был использован рядом авторов (Efron, Thisted, 1976; Burnham, Overton, 1978; Heltshe, Forrester, 1983), которые рассматривали в процедуре имитации прямоугольные площадки (квадраты), являющиеся независимыми выборками из одной и той же генеральной совокупности. Алгоритмы bootstrap и jackknife, осуществляющие генерацию большого числа псевдовыборок, игнорируют взаимоотношения между видами в пределах площадки и не делают никаких предположений о закономерностях распределения плотности вероятности их обнаружения. Непараметрические формулы оценки числа видов в изучаемом сообществе используют информацию о количестве малочисленных (т. е. представленных одиночными или несколькими особями) или уникальных (т. е. встретившихся только в одной пробе – unique) видов: чем больше их число в обрабатываемой совокупности данных, тем больше вероятность обнаружения новых видов, еще не представленных в выполненных исследованиях (Colwell, Coddington, 1994; Colwell et al., 2004; Mao et al., 2005).

Основная идея экстраполяции числа видов методом «складного ножа» (jackknife) заключается в расчете потерь числа видов при удалении одной из произвольных проб (Burnham, Overton, 1978; Smith, van Belle, 1984). Пусть m – количество независимых гидробиологических проб, взятых в изучаемом водотоке, а $\hat{S} = S_{obs}$ – число видов, наблюдаемое в этих пробах. Если случайным образом исключить одну из проб, то оставшееся видовое богатство на основе $m - 1$ измерений будет равно $\hat{S}_{-i} = S_{obs} - q_{1i}$, где q_{1i} – число уникальных видов, встретившихся только в i-й пробе. Тогда путем перебора всех значений

$i = 1, \ldots, m$ можно найти значение суммарных потерь Q_{1*}, после чего рассчитать статистическую оценку видового богатства –

$$\hat{S}_m^1 = S_{\text{obs}} + Q_{1*}(m-1)/m$$

и её дисперсию –

$$\text{var}(\hat{S}_m^1) = \frac{1}{m}\sum_{i=1}^{m}(q_{1i} - Q_{1*}/m)^2 .$$

Эта оценка известна как «складной нож первого уровня» (Jackknife-1) и используется для компенсации статистического смещения оцениваемого параметра S_{obs} порядка $1/m$.

Если отбросить, например, две произвольных пробы i и j, то оставшееся число видов $\hat{S}_{-ij} = S_{\text{obs}} - q_{1\text{i}} - q_{1\text{j}} - q_{2\text{ij}}$, где $q_{2\text{ij}}$ – количество видов, встретившихся только в паре проб i и j. Путем перебора значений i и j можно найти значения суммарных потерь Q_{1*} и Q_{2*}, а также статистическую оценку «складного ножа» второго уровня (Jackknife-2):

$$\hat{S}_m^2 = S_{\text{obs}} + 0.5\{[Q_{1*}(2m-3)]/m - [Q_{2*}(m-2)^2]/[m(m-1)]\}.$$

Глубину перебора можно продолжить и далее – формула для расчета оценок «складного ножа» k-го уровня приведена в (Smith, van Belle, 1984).

Основываясь на работе Б. Харриса (Harris, 1959), А. Чао (Chao, 1984, 1987) использовала в процедурах имитации по алгоритму Chao2 несколько видоизмененные формулы с теми же обозначениями:

$$\hat{S}_m^{Chao2} = S_{\text{obs}} + 0{,}5\,Q_{1*}^2/Q_{2*}; \quad \text{var}(\hat{S}_m^{Chao2}) = Q_{2*}[0{,}5(Q_{1*}/Q_{2*})^2 + (Q_{1*}/Q_{2*})^3 + 0{,}25(Q_{1*}/Q_{2*})^4],$$

а в алгоритме Chao1 учитывала также численность особей каждого вида:

$$\hat{S}_m^{Chao1} = S_{\text{obs}} + 0{,}5\,F_{1*}^2/F_{2*}; \quad \text{var}(\hat{S}_m^{Chao1}) = F_{2*}[0{,}5(F_{1*}/F_{2*})^2 + (F_{1*}/F_{2*})^3 + 0{,}25(F_{1*}/F_{2*})^4],$$

где F_{1*} и F_{2*} – число видов, представленных одиночными (singleton) и двумя (doubleton) особями соответственно.

Э. Смит и Г. ван Белле (Smith, van Belle, 1984) разработали алгоритм экстраполяции оценок видового богатства бутстреп-методом (bootstrap), основанный на подсчете относительной доли проб ρ_j, содержащих каждый j-й вид:

$$\hat{S}_m^{boot} = S_{obs} + \sum_{j=1}^{S_{obs}}(1-\rho_j)^m .$$

В работах Р. Колуэлла и А. Чао (Colwell, Coddington, 1994; Chao, 2005) представлены формулы, основанные на алгоритмах охвата по частоте встречаемости видов ICE (incidence-based coverage estimator) или их численности ACE (abundance-base coverage estimator), где «выборочный охват» определен как сумма разрядов λ, связанных с обнаруженными видами, а λ – параметр смешанной модели распределения Пуассона.

По приведенным формулам достаточно легко рассчитать для произвольного числа проб m превышение полного видового богатства над наблюдаемым и количественно оценить выборочную достаточность проведенного исследования. Нетрудно также отметить, что в описанной схеме имитации генерируется классическая «оценка, основанная на выборках» (sample-based), поскольку манипулируемыми единицами являются целостные площадки, а не отдельные случайные особи.

Исследования по непараметрическим методам экстраполяции богатства видов продолжают развиваться в нескольких направлениях, представленных в основных литературных обзорах (Colwell, Coddington, 1994; Chao, 2005):

- модификация алгоритмов для учета смещения и 95%-ых доверительных интервалов в случае логнормального распределения;
- непараметрическая реализация метода максимального правдоподобия;
- разработка бустреп-процедуры с использованием формулы Горвитца–Томпсона и т. д.

Перечисленные методы (и различные их варианты) широко представлены в общедоступных компьютерных программах. Наибольшей полнотой и удобством отличаются программы **EstimateS** 8.0 (Colwell, 1994-2005) и **SPADE** (species prediction and diversity estimation; Chao, Shen, 2003-2005).

Сопоставление результатов из разных литературных источников показало, что в общем случае характер кривых накопления видового богатства $S(a)$ зависит от трех факторов: изучаемой систематической группы (растения, птицы, насекомые, микроорганизмы и т. д.), экосистемных характеристик конкретного сообщества и интенсивности выборочного усилия. Опубликованные обзоры (Colwell et al., 2004) свидетельствуют о том, что опыт верификации непараметрических моделей оценки видового богатства основывается преимущественно на сообществах тропической растительности (Butler, Chazdon, 1998) или орнитофауны. Аналогичные работы применительно к водным сообществам относительно немногочисленны (например, Foggo et al., 2003). Поэтому представляет интерес проверка работоспособности описанных процедур на конкретных примерах донных сообществ пресноводных водоемов, у которых общепринятые шкалы учета численности популяций во многом отличаются от показателей обилия, например, растительных сообществ.

На рис. 2.7 представлено семейство кривых, экстраполирующих видовое богатство макрозообентоса равнинной реки Сок (см. также рис. 2.5) с использованием описанных методов. Расчет кривой Михаэлиса–Ментен (ММ) осуществлялся с использованием робастного алгоритма, асимптотически оценивающего параметры гиперболы для различных уровней объединения выборок $i = 1, \ldots, m$.

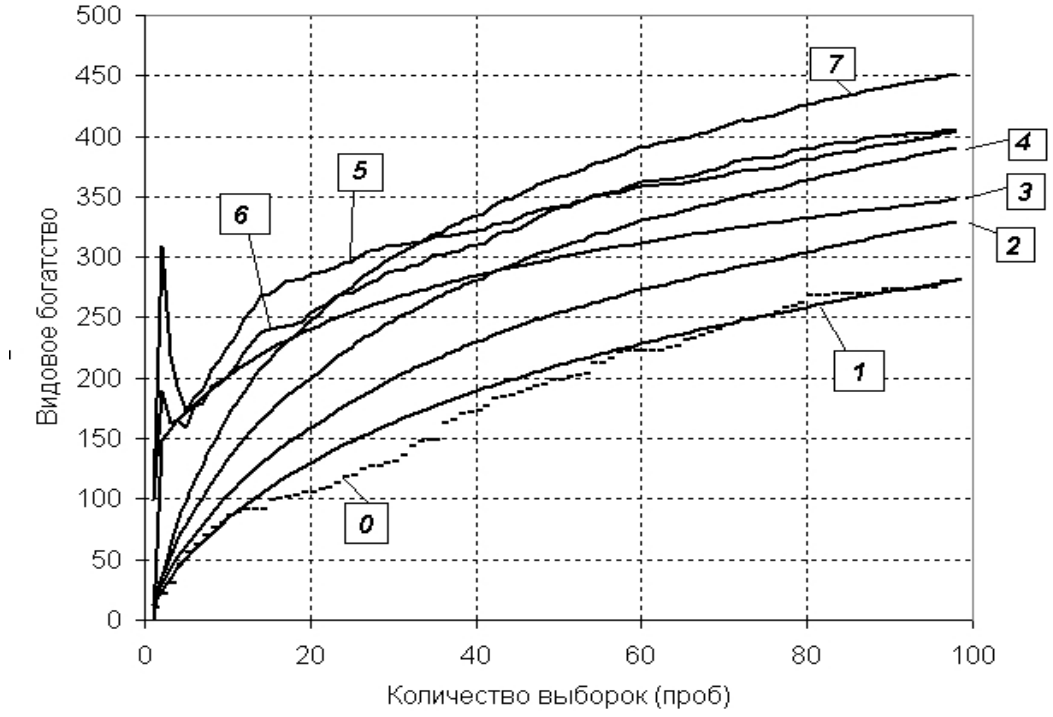


Рис. 2.7. **Экстраполяция видового богатства *S* сообщества макрозообентоса р. Сок:** цифрами отмечены асимптотические кривые, построенные в зависимости от выборочного усилия *m* различными непараметрическими методами: *1* – разрежение Колуэлла–Мао, алгоритм Chao1, алгоритм ACE; *2* – бутстреп-приближение; *3* – гипербола Михаэлиса–Ментен; *4* – оценки «складного ножа» 1-го уровня; *5* – алгоритм охвата ICE; *6* – алгоритм Chao2; *7* – оценки «складного ножа» 2-го уровня; *0* – эмпирическое накопления числа видов в процессе гидробиологической съемки

В табл. 2.2 приведено сопоставление наблюдаемого S_0 и прогнозируемого S_{est} количества видов сообществ макрозообентоса для отдельных участков р. Сок, а также для некоторых других малых рек бассейна Нижней Волги. Результаты расчетов приведены для трех методов экстраполяции, являющихся олицетворением трех возможных стратегий в любом подходе принятия решений: оптимистической (Jackknife-2), пессимистической (бутстреп) и сбалансированной (алгоритм Chao2).

Таблица 2.2

Сравнение наблюдаемого и экстраполированного видового богатства сообществ макрозообентоса для некоторых малых рек бассейна Нижней Волги

Река	Наблюдаемое богатство			Экстраполированное богатство			Отношение S_{pac}/S_0
	m*)	Число видов S_0	Из них, редких**)	Алгоритм	Число видов S_{est}	95% доверительные интервалы	
Чапаевка (нижнее течение)	104	88	31	Bootstrap Chao2 Jackknife-2	102 116 132	$99 \le S_{est} \le 156$	1,16 1,32 1,50
Сок (верхнее течение)	53	205	90	Bootstrap Chao2 Jackknife-2	244 304 341	$263 \le S_{est} \le 374$	1,19 1,48 1,66
Сок (среднее и нижнее течение)	46	176	84	Bootstrap Chao2 Jackknife-2	211 293 310	$242 \le S_{est} \le 384$	1,20 1,67 1,76
Байтуган (приток р. Сок)	31	144	70	Bootstrap Chao2 Jackknife-2	173 231 251	$191 \le S_{est} \le 306$	1,21 1,61 1,75
Б. Кинель	22	121	53	Bootstrap Chao2 Jackknife-2	143 188 201	$154 \le S_{est} \le 256$	1,18 1,55 1,67
Кондурча (приток р. Сок)	9	76	40	Bootstrap Chao2 Jackknife-2	92 123 131	$97 \le S_{est} \le 180$	1,21 1,62 1,73

*) m – количество проб; **) под «редкими» видами понимались виды, встретившиеся только в одной пробе.

В целом рассмотренные водотоки, за исключением нижнего течения р. Чапаевка, характеризуются достаточно близкой между собой величиной общего видового богатства макрозообентоса с учетом найденных доверительных интервалов (табл. 2.2). Относительно небольшое значение S_{est} для достаточно чистой р. Кондурча, по сравнении с реками Сок и Байтуган, можно объяснить, в первую очередь, малым числом отобранных проб. Весьма интересно сопоставление результатов экстраполяции числа видов для каждого из двух выделенных участков р. Сок (около 300 видов на основе 50 проб) и в целом по всему водотоку (340 видов при том же выборочном усилии). Величина этой разности позволяет оценить гетерогенность экосистемы реки как весьма умеренную.

Хотя экстраполяция всегда представляет менее устойчивый и сложнее аргументируемый процесс, чем интерполяция, некоторые из приведенных асимптотических формул показали достаточно хорошую эффективность прогноза общего видового богатства S_{est} макрозообентоса. Целая группа алгоритмов (Jackknife-1, Chao2 и ICE) дали между собой весьма близкие результаты, которые согласуются с нашими общими представлениями о составе донных сообществ сравниваемых водотоков (Биоиндикация экологического..,

2007). Несколько меньшее число видов было получено при аппроксимации функцией Михаэлиса–Ментен, тогда как самые скромные прогнозы принадлежат бутстрепу. В противоположность этому, алгоритм Jackknife-2, по сравнению с другими моделями, дал несколько завышенные результаты.

Следует отметить, что методы, основанные на анализе численности особей (Chao1 и ACE), не дали позитивного результата, и кривые экстраполяции почти в точности совпали с кривой разрежения (кривая 1 на рис. 2.8). Вполне понятно, что используемая в алгоритме Chao1 встречаемость F_1 видов, представленных одиночными экземплярами, для донных сообществ является практически маловероятным событием. Это замечание имеет отношение к одной из важнейших проблем экологии – разработке сопоставимых шкал оценки обилия и популяционной плотности для сообществ разных систематических групп (рыб, водорослей, беспозвоночных насекомых, бактерий и т. д.), что необходимо для сравнительной оценки их объединенного биоразнообразия. Становится все более ясным, что существующие шкалы измерения популяционной плотности гидробионтов (терминологически некорректно называемые «численностью» и «биомассой» видов – sic!) являются метрически неоптимальными и неудобными для системной обработки.

В работе А. Фоджо с соавторами (Foggo et al., 2003) также было отмечено, что наилучшие оценки обеспечиваются алгоритмами Chao2 и ICE, когда как бутстреп дает систематически заниженные оценки при небольшом выборочном усилии. Д. Хеллманн и Д. Фаулер (Hellmann, Fowler, 1999) проанализировали смещение, точность и правильность[1] прогноза видового богатства различными способами. Используя выборку из 160 описаний делянок леса, они с помощью генератора случайных чисел получили 5000 псевдонаборов данных в пределах от 2,5% до 70% исходной совокупности. На основе анализа кривых зависимости прогнозируемых значений S_{est} от выборочного усилия методом Монте-Карло оценивались вариационные статистики для каждого использованного алгоритма и, в частности, отмечалась стабильность бутстрепа и неустойчивость Jackknife-2.

В целом наши результаты согласуются с этими выводами, хотя вопрос о том, какая модель «недооценивает», а какая «переоценивает» истинное видовое богатство нуждается в дальнейшем тщательном изучении на широком экспериментальном материале. Однако нельзя не согласиться с основным выводом работы (Hellmann, Fowler, 1999), что к изучению каждого сообщества *надо подходить индивидуально, принимая во внимание его особенности и уровень видового богатства*.

2.4. Модели разрежения и сравнение биоразнообразия сообществ с учетом выборочного усилия

Последовательное объединение мониторинговых данных на основе подмножества однократных выборок (например, гидробиологических проб из одного биотопа) дает кривую накопления видов, которая не является гладкой из-за пространственной (или временной) неоднородности или различных случайных факторов. Классическая схема разрежения (Sanders, 1968; Hurlbert 1971; Simberloff, 1978) основана на предположении о случайном смешивании особей и используется, чтобы воспроизвести гладкую кривую, оценивающую

[1] Под смещением (bias) понимается различие между рассчитанным по формуле прогнозом (например, среднее из оценок всех возможных выборок, которые могут быть взяты из совокупности) и истинным неизвестным значением богатства видов. Точность (accuracy) – различие между оценками богатства видов, основанного на выборочных данных и истинным числом видов сообщества. Надежность (precision) – различие между выборочной оценкой и средним из оценок всех возможных выборок, которые могут быть взяты из совокупности. Точность соответствует дисперсии, а надежность – среднеквадратической ошибке метода оценки. Когда смещение равно нулю, надежность и точность идентичны.

видовую насыщенность, наблюдаемую для любого меньшего числа индивидуумов. Как отмечалось выше, схема разрежения, основанная на выборках (sample-based rarefaction; Gotelli, Colwell, 2001), оценивает число видов для любого меньшего числа проб из некоторой повторности наблюдений (replicated sets of samples) и использует предположение о случайном порядке выборок, полученных с использованием различных устройств, реализующих мониторинговые повторности (пробные площадки, западни, приманки, неводы, батометры, дночерпатели и др.).

Предварительно необходимо отметить, что наблюдаемое богатство видов локального сообщества в пределах естественной среды его обитания зависит как от выборочного усилия (α-разнообразие), так и от риска последовательного объединения видов из нескольких сообществ, принадлежащих разным биотопам (прирост числа видов, обусловленный β-разнообразием). Практически же в моделях локального (α) разнообразия предполагается, что видовой состав выборок в незначительной степени зависит от времени или от точки пространства, в пределах которых выполнялось взятие проб. В статистических терминах это означает, что каждая выполненная проба считается случайной выборкой из одной и той же генеральной совокупности.

Процесс разрежения видов на основе случайных повторностей отбираемых проб может интерпретироваться как сглаживание кривой накопления и служит для строгой оценки ожидаемого числа видов с доверительными интервалами для некоторого «стандартного» выборочного усилия. При необходимости возможна также экстраполяция наблюдаемого видового богатства сообщества на неограниченное множество выборок. Описанный ниже метод (Colwell et al., 2004; Mao, Colwell, 2005) основан на статистически строгой смешанной биномиальной модели (mixture model) встречаемости в многовидовых ассоциациях, строгое обоснование которой представлено в статье (Mao, Lindsey, 2004), где можно найти основные теоремы и доказательства.

Пусть из сообщества видов с неизвестным истинным богатством S взята серия независимых проб, которые отражают эмпирическую функцию распределения встречаемости видов: $z_{ij} = 1$, если i-й вид найден в j-й пробе, и $z_{ij} = 0$ в противном случае. Делается два статистических предположения: (1) вероятность ϕ_i обнаружения i-го вида одинакова для каждой пробы и (2) значения z_{ij}, определяемые ϕ_i, являются независимыми для всех i и j. Тогда функция накопления видов, которая задает ожидаемое их количество для произвольного числа m взятых проб, выражается через сумму вероятностей, определяющих наличие каждого из S видов хотя бы в одной из m выборок:

$$\tau(m) = \sum_{i=1}^{S}[1-(1-\phi_i)^m]\cdot$$

Виды по вероятности ϕ_i их встречаемости разделяются на G гомогенных групп инцидентности (некоторые группы могут содержать только один вид). Пусть для k-й группы (k = 1, 2, …, G), ψ_k – общая вероятность присутствия (мера редкости или тривиальности видов) и π_k – относительный размер группы (т. е. число видов k-й группы, деленное на общее число найденных видов S). Тогда

$$\tau(m) = S\sum_{k=1}^{G}\pi_k[1-(1-\psi_k)^m] \quad \text{или} \quad \tau(m) = S\sum_{k=1}^{G}\pi_k(1-e^{C_k m}), \quad \text{где} \quad C_k = -\ln(1-\psi_k).$$

Полученная модель является непараметрическим обобщением экспоненциальной модели Л. Холдриджа с соавторами (Holdridge et al., 1971), предполагающей, что все виды имеют одинаковую вероятность присутствия и сформирована единственная группа инцидентности. Действительно, если установить $G = 1$, $\pi_k = 1$ и $C_k = C$, то можно прийти к классическому экспоненциальному уравнению $\tau(m) = S(1-e^{-Ca})$.

Пусть s_j – количество видов, найденных точно в j пробах: например, s_1 – число видов, которые встретились только в одной пробе из m, s_2 – в двух пробах одновременно и т. д. Тогда число видов, зарегистрированное во всей исходной выборочной совокупно-

сти, будет равно $S_{obs} = \sum_{j=1}^{M} s_j$, где M – число выделенных градаций. По аналогии принимается, что s_0 – число видов, присутствующих в сообществе, но не встретившихся ни в одной из проб. С учетом этого, общее богатство видов будет равно $S = S_{obs} + s_0$.

Эмпирическая последовательность частот $s_1, s_2, ..., s_M$ содержит всю информацию, необходимую для несмещенной оценки числа видов как функции выборочного усилия $\tau(m)$, которая может быть получена по формуле, основанной на моментах:

$$\text{€}(m) = \sum_{j=1}^{M} (1 - \alpha_{jm}) s_j = S_{obs} + \sum_{j=1}^{M} \alpha_{jm} s_j, \quad m = 1, 2, ..., M,$$

где комбинаторные коэффициенты $\alpha_{jm} = \dfrac{(M-m)!(M-j)!}{(M-m-j)!M!}$ для всех $(j + m + 1) < M$.

Для этой функции можно аппроксимировать 95%-ые доверительные интервалы $\text{€}(m) \pm 1{,}96\text{€}_\tau(m)$, где $\text{€}_t(m) = \sum_{j=1}^{M} (1 - \alpha_{jm}) s_j - \text{€}(M)^2 / \overline{S}$ – оценка стандартной ошибки, а $\overline{S}$ – общее видовое богатство, которое можно найти любым описанным выше непараметрическим методом экстраполяции (jackknife, bootstrap и др.).

С использованием описанной модели разрежения Колуэлла–Мао и программы EstimateS, была выполнена сравнительная оценка видового богатства и биоразнообразия сообществ макрозообентоса малых рек Самарской области. На рис. 2.8 показаны кривые разрежения, построенные для изучаемых водотоков.

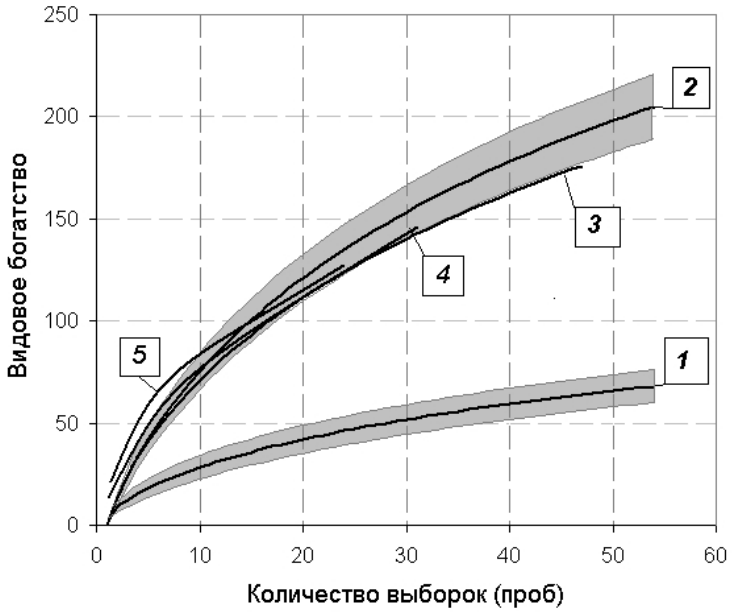


Рис. 2.8. **Кривые Колдуэлла–Мао разрежения видового богатства для сообществ макрозообентоса некоторых участков малых рек Самарской области:**

1 – нижнее течение р. Чапаевка; 2 – верхнее течение р. Сок; 3 – нижнее течение р. Сок; 4 – р. Байтуган; 5 – р. Б. Кинель; серым цветом показаны 95% доверительные интервалы для участков рек 1 и 2

Можно отметить, что видовое богатство в верхнем течении р. Сок (ст. 1-7) несколько больше, чем в нижнем (ст. 8-13), однако это различие статистически незначимо, поскольку доверительные интервалы среднего числа видов перекрываются, в то время как сравнительная таксономическая бедность донных сообществ р. Чапаевка в её нижнем течении – ниже г. Чапаевска (ст. 13-24) очевидна и визуально, и статистически. Полученные

результаты в целом не противоречат выводам о сравнительном биоразнообразии отдельных равнинных рек бассейна Нижней Волги, которые были опубликованы нами ранее (Биоиндикация экологического.., 2007).

Основная роль кривых разрежения заключается в сравнении видового богатства при одном и том же выборочном усилии. Например, в нижнем течении р. Чапаевка было зарегистрировано 88 таксонов макрозообентоса в 104 пробах, в то же время в р. Кондурча установлено 76 видов, но в 9 пробах. Однако если учесть, что расчетное значение S_{est} в р. Чапаевка при m = 9 выборок составляет лишь 26 видов (в р. Байтуган – 71 видов, в р. Сок – 79, в р. Б. Кинель – 80), то р. Кондурча займет законное место в рейтинге рек с высоким видовым богатством макрозообентоса.

Видовое богатство является важнейшим, но не единственным показателем «благополучия» сообществ. Другим подходом, используемым экологами, является оценка видового разнообразия с помощью индексов (Dennis et al., 1979; Мэгарран, 1992; Bunge, Fitzpatrick, 1993), которые принимают во внимание несколько таких важных аспектов, как равномерность распределения, доминирование и уникальность видов. И в этом случае недостаточный учет зависимости основных индексов видового разнообразия от выборочного усилия часто приводит к ошибочным выводам при сравнении показателей для разных водотоков.

Методика статистической оценки математического ожидания индексов видового разнообразия и их доверительных интервалов была разработана С. Цаалем (Zahl, 1977), использовавшем на примере индекса Симпсона процедуру «складного ножа». В последующих работах эти идеи получили теоретическое обоснование и были обобщены для энтропийной формулы Шеннона (Heltshe, Forrester, 1983; Chao, Shen, 2003). Л. Барабеси и Л. Фатторини (Barabesi, Fattorini, 1998) обосновали применение идеологии генерации псевдовыборок для расчета оценок $\hat{\Delta}$ большого класса индексов видового разнообразия Δ, если они удовлетворяют выражению $\hat{\Delta} = \Delta(p_1, p_2, \cdots, p_k)$, где p_i – вероятности встречаемости особей разных видов, а общие результаты исследований рассматриваются как независимые реплики «схемы столкновений». Показано, что с использованием имитационных методов можно не только асимптотически оценить дисперсию σ, но также скомпенсировать смещение математического ожидания и сгенерировать наилучшим образом центрированные доверительные интервалы для любого индекса из семейства Δ.

Рассмотрим два индекса оценки биоразнообразия – информационный индекс Шеннона H и параметр α модели распределения численности видов логарифмическими рядами Фишера (Мэгарран, 1992), которые, согласно проведенным исследованиям (Taylor et al., 1976), обладают наилучшей дискриминирующей способностью. Рассчитаем с использованием алгоритма интерполяции Колуэлла–Мао (программа EstimateS) кривые зависимости величины этих индексов от выборочного усилия (см. рис. 2.9). Очевидно, что корректное сравнение индексов видового разнообразия для различных сообществ также возможно лишь при одинаковом выборочном усилии. Например, индекс Шеннона H для нижнего течения р. Сок (кривая 2 на рис. 2.9а) при количестве проб до 25 превышает аналогичный показатель для верхнего течения, однако при дальнейшем росте выборочного усилия значения H выравниваются. Это подтверждает полученные данные наличия в верхнем течении р. Сок значительного комплекса относительно редких видов, вероятность обнаружения которых увеличивается при углублении мониторинговых исследований.

Полученные непараметрическими методами границы доверительных интервалов индексов биоразнообразия дают возможность проверить гипотезу о статистической значимости различий между двумя сравниваемыми оценками при одинаковом выборочном усилии (см. табл. 2.3).

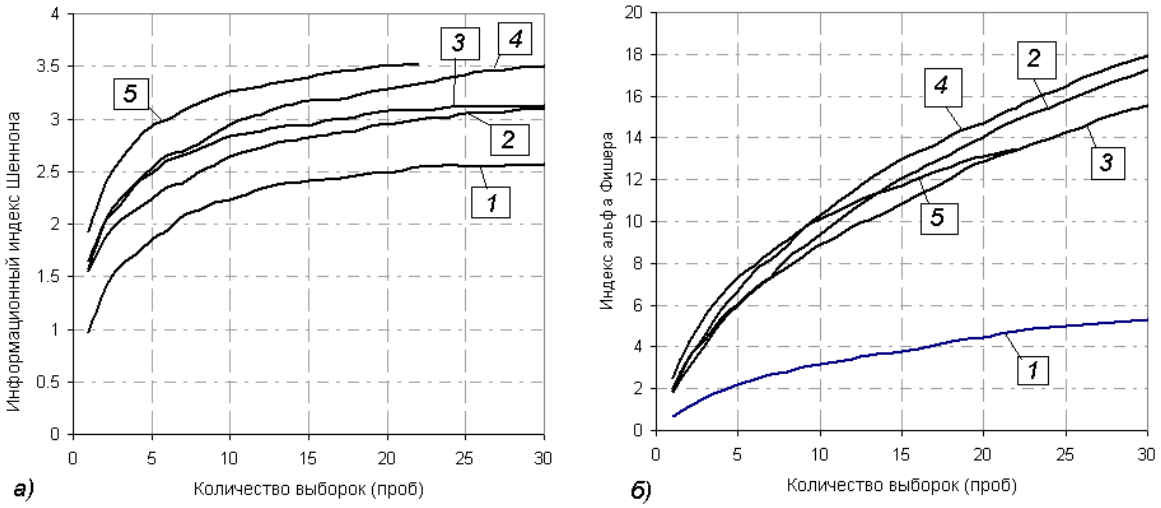


Рис. 2.9. **Кривые зависимости информационного индекса Шеннона (а) и параметра α логарифмического ряда Фишера (б) от выборочного усилия для сообществ макрозообентоса некоторых участков малых рек Самарской области (обозначения аналогичны рис. 2.8)**

Таблица 2.3

Интерполированное число видов (числитель) и его доверительные границы при *p* = 0,95 (знаменатель) для сообществ макрозообентоса

Река	Показатель разнообразия	Количество проб *m*					
		9	**22**	**31**	**46**	**53**	**104**
Чапаевка (нижнее течение)	Число видов	26,5	44,2	52,6	63,4	67,5	**88,0**
		21 ÷ 32	37 ÷ 51	45 ÷ 60	56 ÷ 71	60 ÷ 75	79 ÷ 97
	H Шеннона	2,2	2,55	2,58	2,7	2,73	**2,85**
	α Фишера	3,02	4,75	5,36	6,26	6,64	**8,27**
Сок (верхнее течение)	Число видов	75,3	131,8	158,7	192,5	**205,0**	
		66 ÷ 84	120 ÷ 144	145 ÷ 172	177 ÷ 208	189 ÷ 220	
	H Шеннона	2,56	2,99	3,14	3,25	**3,27**	
	α Фишера	8,8	14,86	17,55	20,8	**21,86**	
Сок (среднее и нижнее течение)	Число видов	70,9	121,0	145,0	**176,0**		
		62 ÷ 80	108 ÷ 134	131 ÷ 159	160 ÷ 192		
	H Шеннона	2,77	3,08	3,14	**3,24**		
	α Фишера	8,37	13,44	15,85	**18,66**		
Байтуган	Число видов	70,8	121,1	**144,0**			
		72 ÷ 80	108 ÷ 134	130 ÷ 158			
	H Шеннона	2,86	3,33	**3,52**			
	α Фишера	9,67	15,43	**18,18**			
Б. Кинель	Число видов	79,8	**121,0**				
		70 ÷ 90	108 ÷ 134				
	H Шеннона	3,21	**3,53**				
	α Фишера	9,67	**13,5**				
Кондурча	Число видов	**76**					
		66 ÷ 86					
	H Шеннона	**2,17**					
	α Фишера	**8,95**					

Предположим, необходимо сравнить две кривые на рис. 2.9а, соответствующие рекам Байтуган (4) и Б. Кинель (5). Для р. Байтуган при сопоставимом выборочном усилии $m = 22$ оценки индекса Шеннона $\overline{H}$ и его стандартной ошибки s_H, полученные методом максимума правдоподобия, соответственно, принимают значения $\overline{H} = 3.33$ и $s_H = 0{,}092$. Тогда нижняя $H_н$ и верхняя $H_в$ границы доверительного интервала будут равны

$$H_н = \overline{H} - t_\alpha s_H = 3{,}33 - 1{,}96 \cdot 0{,}092 = 3{,}15;$$

$$H_в = \overline{H} + t_\alpha s_H = 3{,}33 + 1{,}96 \cdot 0{,}092 = 3{,}51.$$

Аналогично для р. Б. Кинель границы доверительного интервала при том же выборочном усилии: $H_н = 3{,}527 \le \overline{H} \le H_в = 3{,}542$, что опровергает гипотезу о равенстве индекса Шеннона в этих двух реках.

ГЛАВА 3. СХОДСТВО И РАЗНООБРАЗИЕ ВИДОВОЙ СТРУКТУРЫ СООБЩЕСТВ

Идея – это трещина в темных очках, через
которую ты видишь мир: чем мощнее идея,
тем сильнее искажается восприятие мира.
"Пятничный manager" (Сборник юмора)

3.1. Развитие представлений о видовом разнообразии: от энтропии к фракталам

Экологические последствия снижения биологического разнообразия вызывают повышенный интерес и оживленные дискуссии экологов в течение последних нескольких десятилетий. Главные направления актуализации исследований состоят в разработке методов идентификации функционально важных таксономических групп, описании зависимостей между видовым разнообразием и продукционными процессами и раскрытии основных механизмов устойчивого развития экосистем. К. МакКанн (McCann, 2000) выделяет четыре основных этапа развития этих представлений:

- до 1970-х годов большинство экологов полагало, что разнообразие сообществ закономерно увеличивает устойчивость экосистем (Elton, 1958);
- с этими ранними интуитивными представлениями не согласился Р. Май (May, 1973), который, используя статистический анализ, основанный на вероятностных процессах, показал, что разнообразие имеет тенденцию дестабилизировать динамику сообществ;
- позднее П. Йодзис (Yodzis, 1981) подверг сомнению эти результаты и его модели, построенные на биологически обоснованной вероятности межвидовых взаимодействий, продемонстрировали положительную корреляцию между разнообразием и стабильностью;
- дальнейшие исследования развивались по двум основным направлениям – «механистическому», изучающему конкретные зависимости между структурой трофических цепей, продуктивностью экологических систем и их стабильностью (Kassen et al., 2000), и статистическому, обосновывающему общие отношения устойчивости и разнообразия (Tilman, 1996).

Исследования, основанные на концепции «обобщение через изучение конкретных экосистем» (Loreau et al., 2001), находят все новые аргументы, противоречащие устоявшимся представлениям, что самые производительные экосистемы характеризуются низким видовым разнообразием (впрочем, это, действительно, вполне типичная ситуация, часто наблюдаемая в природе). Анализ репрезентативной совокупности случайно выбранных растительных сообществ из разных стран Европы и штата Калифорнии показал, что первичная продукция (прирост биомассы на 1 $м^2$ в единицу времени) имеет положительную корреляцию с общим числом видов растений и разнообразием функциональных групп.

Последователи статистической линии исследований (Doak et al., 1998; Yachi, Loreau, 1999) считают, что биологическое разнообразие обеспечивает сообществу «страховочный полис» или буфер, компенсирующий экологические флуктуации, поскольку всегда найдется совокупность видов, толерантных или быстро адаптирующихся в изменившихся условиях. Согласно «страховочной» гипотезы виды, которые функционально избыточны на данный момент времени при существующих экосистемных процессах, перестают быть таковыми в долговременном аспекте. Другая гипотеза «эффекта портфолио» (Tilman, 1996; Lhommel, Winkel, 2002) отражает тот факт, что стабильность (т. е. постоянство во времени) таких агрегированных свойств сообщества, как производство биомассы, в целом повышается с ростом видового разнообразия, однако это происходит из-за эффекта статистического усреднения («statistical averaging effect») продукционных свойств видов.

Перечисленные здесь фундаментальные исследования ограничивались «простыми» представлениями о разнообразии, как о числе функционирующих в экосистеме видов S. Однако, как отметил С. Хелберт (Hurlbert, 1971), сравнение сообществ по одному индексу является не вполне корректным, поскольку различные показатели оценивают разные аспекты видового разнообразия и, в частности, обладают *специфической чувствительностью к видам с определенной частотой встречаемости*. Можно заметить (рис. 2.8, табл. 2.3), что, рассчитав только три показателя, мы получим три различно ранжированных по биоразнообразию списка водотоков. Например, при количестве проб $m = 22$ максимум числа обнаруженных видов наблюдается в верховьях р. Сок, при расчете индекса Шеннона – в р. Б. Кинель, а показатели индекса α Фишера свидетельствуют о наибольшем видовом разнообразии в р. Байтуган.

Большое количество используемых показателей разнообразия и противоречивость их поведения привело к заключению некоторых авторов (Hurlbert, 1971), что понятие разнообразия вообще бессмысленно. Научная и практическая ценность результатов расчета разнообразия и его связи с адаптивной устойчивостью экосистем (Артюхов, 2009) во многом определяется адекватностью выбранной модели и корректностью ее интерпретации. Разнообразие, безусловно, имеет смысл, поскольку с точки зрения ОТС[1] любой системе присущи *элементы* разнообразия $\{\boldsymbol{m}\}$, *отношения* между элементами $\{\boldsymbol{r}\}$ и ограничивающие эти отношения условия или *законы композиции* $\{\boldsymbol{z}\}$. Наиболее широко используемая информационная трактовка разнообразия как «числа различных элементов» является предельно редуцированным вариантом этого определения, в котором все потенциальное многообразие отношений между элементами ограничено простейшим требованием пересчитать их подряд. Исходя из этих принципов формирования, разнообразие, оцениваемое как число обнаруженных в экосистеме видов S, можно называть *счетным* или *нумерическим*.

Однако, кроме простого числа элементов, представляется целесообразным учесть при оценке разнообразия и законы композиции этих элементов (например, выравненность), и их иерархические отношения между собой. Для качественного сопоставления видового разнообразия используются выработанные практикой разумные критерии и показатели, которые, тем не менее, являются субъективными и отражающими лишь некоторые частные стороны этого свойства. Например, радиус сферы – индекс её объема, но не является непосредственно объемом, и использование радиуса вместо объема в технических уравнениях даст опасно вводящие в заблуждение результаты. То же самое – с показателями разнообразия в биологии.

В течение последних 50 лет не прекращаются дискуссии, отмечающие теоретическую противоречивость и практическую нестабильность индекса видового разнообразия Шеннона (Шитиков и др., 2005; Розенберг, 2010). Вместе с тем это, вероятно, самый глубокий и полезный из всех показателей разнообразия, оценивающей энтропию (т. е. уровень неопределенности) в результате выборочного процесса. Если рассчитать его, используя логарифмы с основанием 2, то он будет соответствовать среднему минимальному числу испытаний, в результате которых из сообщества будет извлечена особь самого многочисленного вида. Энтропия – разумный показатель разнообразия, но это не причина утверждать, что энтропия – и есть «разнообразие».

Меры разнообразия часто используются, чтобы оценить изменения в окружающей среде вследствие загрязнения, изменения климата, или других факторов. Предполагается, что большое снижение видового разнообразия экосистемы должно вызвать адекватное уменьшение значения оценивающего его показателя. Пусть (Jost, 2008) на континенте присутствует миллион видов с одинаковой численностью и произвольное катастрофическое воздействие (например, падение метеорита) убивает 999 900 видов, оставляя только 100 из них. Любой биолог, если его спросить, сказал бы, что воздействие метеорита вызвало

[1] ОТС – Общая Теория Систем (Урманцев, 1978).

большое абсолютное и относительное понижение разнообразия. Однако популярный индекс Симпсона–Джини уменьшается при этом только от 0,999999 до 0,99 (снижение менее чем на 1%). Очевидно, что поведение этой меры не соответствует интуитивному понятию разнообразия, и экологи, полагающиеся на индекс Симпсона, будут часто недооценивать величину изменений в экосистеме. В этих же условиях индекс Шеннона уменьшается с 19,9 до 6,6, т. е. уменьшение видового богатства в 10 000 раз сопровождается снижением меры разнообразия только в 3 раза. Это не означает, что шенноновская энтропия – неудачная мера, но её значение оценивает *неопределенность*, а не разнообразие.

Другим спорным моментом оценки разнообразия является иерархическая организация биологических сообществ. В наиболее общем смысле под биологическим разнообразием подразумевают «разнообразие жизни». Это понятие охватывает генетическое разнообразие различных видов и более высоких таксономических единиц (семейств, классов, типов и т. д.), а также разнообразие сред обитания и экосистем. Вследствие этого, вычислительные схемы разнообразия на некотором фиксированном уровне организации экосистем должны быть статистически (и с учетом филетических отношений) связаны с аналогичными «паттернами» на верхних и нижних уровнях (Enquist et al., 2002). В любом случае индексы разнообразия тесно связаны с таксономическими шкалами, в которых представлены данные мониторинга сообществ (Павлинов, 2008).

Поиск адекватных метрик оценки разнообразия не прекращается и до настоящего времени. К. Рао (Rao, 1982) предлагает аксиоматизацию мер разнообразия на основе расщепления их значений на две составляющие: «между» и «в пределах» таксономических групп, определенных в соответствии с иерархической классификацией исследуемого сообщества. Впоследствии К.-С. Лоу (Lau, 1985) показал, что любая точная мера разнообразия должна иметь форму $\mathbf{p}^{\mathrm{T}}\mathbf{D}\mathbf{p}$, где $\mathbf{p}$ – вектор вероятностей встречаемости видов, $\mathbf{D}$ – матрица размерностью $S \times S$, элементы которой d_{ij} отражают различие (обычно в биологическом смысле) между видами i и j ($i, j = 1, \ldots, S$).

Этот индекс вяло упоминался в работах как *квадратичная энтропия Рао* до появления статьи С. Чампли и Д. Чесселя (Champely, Chessel, 2002), которая стимулировала оживленное обсуждение. Сторонники практического использования квадратичной энтропии подчеркивали: «*Если в сообществе какой-то вид заменить другим с тем же обилием, но с уникальными биологическими характеристиками, то традиционные показатели разнообразия останутся незатронутыми, в то время как интуиция подскажет нам увеличение разнообразия*». Нет никаких сомнений, что и гидробиологи зачастую сталкивались с этим тезисом, имея в своем арсенале достаточный материал многолетних исследований, используемый для сравнения видового разнообразия разнотипных водоемов. В формуле квадратичной энтропии (Champely, Chessel, 2002) предложили использовать евклидову метрику для определения меры дистанции между видами, определив тем самым матрицу разнообразия, которая дает возможность геометрических интерпретаций и графических представлений иерархических деревьев разнообразия (tree species diversity and dominance). Однако их оппоненты (Izsak, Szeidl, 2002) указывают на некоторые аномалии квадратичной энтропии, обоснованно полагая, что проблема оценки разнообразия все еще остается открытой.

Р. Руссо с соавторами (Rousseau et al., 1999) отмечают, что необходимо принимать во внимание оба аспекта разнообразия: и равномерность распределения и богатство видов (т. е. для равного богатства видов его величину определяет равномерность распределения и наоборот), поскольку ни один из этих двух компонентов не определяет исчерпывающе свойства экосистемы. Нами (Шитиков, Розенберг, 2005) также было дано формальное определение разнообразия, как произвольного функционала от двух характеристик разбиения компонентов экосистемы на классы – числа таксонов S и выборочной функции распределения $F_s(p)$ обилия особей по классам: $D = \Psi\,[S, \Omega(F_s(p))]$, где Ω – некоторая наиболее подходящая выборочная статистика (дисперсия, мера информации или др.).

Одним из подходов, обеспечивающих появление практически бесконечного множества показателей видового разнообразия, является использование параметрических семейств индексов. А. Реньи (Rényi, 1961) расширил концепцию информации (энтропии) Шеннона и предложил семейство обобщенных энтропий, которое дает целый континуум индексов разнообразия, различающихся своей чувствительностью к обильным и редким видам, что позволяет получить наиболее полное представление о структуре сообщества:

$$H_\alpha = \frac{1}{1-\alpha} \ln \sum_{i=1}^{S} p_i^\alpha \text{ , где } -\infty \le \alpha \le \infty .$$

М. Хилл (Hill, 1973a) впервые использовал семейство *индексов Реньи* как обобщенную формулировку биологического разнообразия. Однако чтобы избавиться от логарифма в формуле Реньи, он предложил её экспонентный вариант, получивший известность как «*числа Хилла*»: $D_\alpha = \left(\sum_{i=1}^{S} p_i^\alpha \right)^{1/1-\alpha}$. Очевидно, что оба семейства индексов дают полностью эквивалентную между собой интерпретацию структуры сообщества.

В области небольших и отрицательных значений параметра α Реньи наибольший вклад в оцениваемый показатель энтропии вносят редкие виды, а при увеличении α возрастает вклад видов с высокой представленностью (доминантов). Например, более высокое разнообразие верхней части р. Сок, по сравнению с низовьями, определяется комплексом малочисленных видов, когда как основной фоновый видовой состав имеет одинаковые параметры разнообразия (рис. 3.1). Можно показать, что многие из предложенных индексов разнообразия являются частными случаями обобщенной энтропии Реньи: для $\alpha = 0$, $H_0 = \ln S$; для $\alpha = 1$, $H_1 = H$, т.е. индекс Шеннона, и т. д.

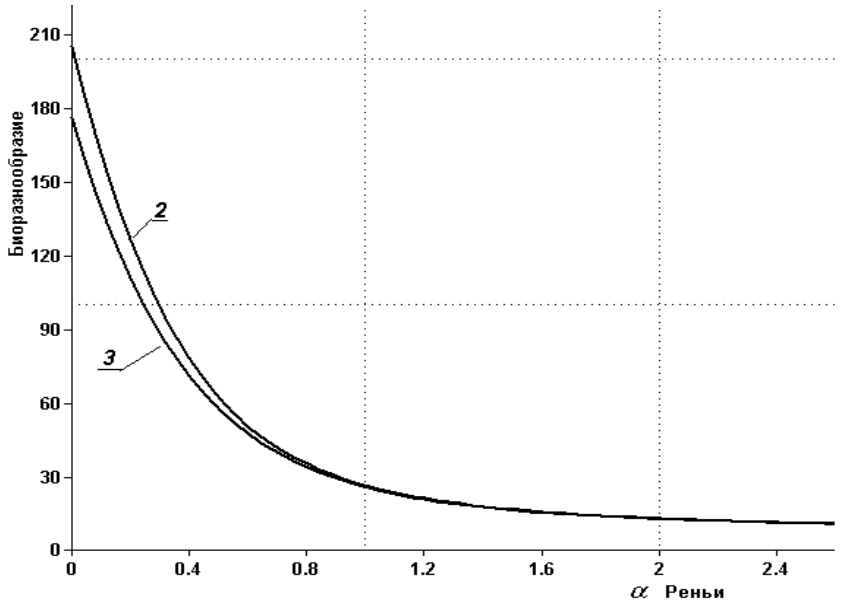


Рис. 3.1. **Кривые семейства чисел видового разнообразия Хилла для верхнего (2) и нижнего (3) течения р. Сок**

Спектр значений видового разнообразия, по Реньи или Хиллу, ограничивается только заданным параметрическим семейством индексов. Чтобы избежать конфликта при сравнении более широкого круга показателей, Г. Пэтил и К. Тейли (Patil, Taille, 1982) обсуждают концепцию «ординации внутреннего разнообразия» (concept of intrinsic diversity ordering), определяющую, что сообщество C_2 имеет большее внутреннее разнообразие,

чем сообщество C_1 ($C_1 < C_2$), если C_2 может быть получено из C_1 через конечную последовательность следующих операций: а) перераспределение численностей от более к менее обильным видам, не меняющее порядок их ранжирования; б) передача части численностей новым видам; в) перемаркировка видов. Исходя из представлений о «нейтральных видах»[2], эта концепция была расширена и был сформулирован (Ricotta, 2003; Jost, 2008) перечень стандартизованных требований к мерам разнообразия:

1) симметричность (функция разнообразия должна быть монотонно симметрична относительно своих аргументов);
2) расширяемость и независимость результата от нуля (добавление видов с нулевым обилием не должно изменять полного разнообразия);
3) принцип «выравниваемости» (перемещение единицы обилия от массового к более редкому виду не должно уменьшать разнообразие);
4) однородность (разнообразие зависит только от относительных частот видов, а не от их абсолютного обилия);
5) принцип репликации;
6) нормализация (если в сообществе присутствует S видов с одинаковым обилием, то мера его разнообразия должна быть равна S).

Свойства 1–4 являются достаточно бесспорными и им удовлетворяют многие известные меры композиционной сложности или энтропии. Свойства 5–6 нуждаются в дополнительном обосновании и составляют смысл новой концепции «истинного» разнообразия («true» diversity), отстаиваемой Л. Джостом (Jost, 2006).

Предположим, что каждое из m сообществ состоит из S видов с одинаковым распределением обилия p, однако ни один из видов не является общим для любой пары сообществ. У всех m сообществ обязательно должно быть одинаковое разнообразие $D(p)$. Теперь объединим все m сообществ. Если мера разнообразия $D(p)$ удовлетворяет условиям *принципа репликации*, то общее разнообразие m объединенных, одинаково равновеликих, но полностью отличных по видовому составу сообществ должно быть $m \cdot D(p)$. Очевидно, что индекс Шеннона H не удовлетворяет принципу репликации (при объединении двух сообществ из 8 уникальных видов H возрастет с 3 до 4).

Нормализацию разнообразия Джост (Jost, 2006) сводит к задаче поиска эквивалентного сообщества, составленного из видов с одинаковой встречаемостью p, у которого численно то же самое значение показателя разнообразия, что и у реально рассматриваемого сообщества. Количество видов D в эквивалентном сообществе будет оценкой видового разнообразия изучаемого сообщества. Р. Мак-Артур (MacArthur, 1965) назвал число D «эффективной видовой насыщенностью»; в физике это – набор состояний, связанных с расчетной энтропией, в экономике – «эквивалентные числа» мер разнообразия (Patil, Taille, 1982).

Большинство непараметрических показателей разнообразия, используемых в науке (включая все обобщенные энтропии, используемые в биологии), является монотонными функциями $\sum_{i=1}^{S} p_i^{\alpha}$, или пределами этих функций при неотрицательной постоянной α. Разнообразие нулевого порядка ($\alpha = 0$) ${}^0D = \sum_{i=1}^{S} p_i^0 = S$ полностью нечувствительно к частотам p и больше известно как видовое богатство. Все значения α, которые меньше единицы, дают оценки разнообразия, которое преднамеренно подчеркивают роль редких видов, в то время как значения $\alpha > 1$ непропорционально ориентированы на виды-доминанты (Tsallis et al., 1998; Keylock, 2005).

---

[2] Разнообразие «нейтральных видов» зависит только от частот их встречаемости, исключая любые другие характеристики, такие как филогенетические отношения или функциональная роль в экосистеме.

Критическая точка при $\alpha = 1$ обнаруживается для функции энтропии Шеннона ${}^1D = \exp\left(-\sum_{i=1}^{S} p_i \ln p_i\right) = e^H$. «Центральная роль этой формулы определяется не модой или предубеждением, а, скорее, как следствие её единственной способности взвесить элементы точно соизмеримо с их частотой, не ориентируясь ни на редкие, ни на доминирующие виды. *Даже если бы теории информации не существовало, биологи обнаружили бы это свойство и использовали H как главный показатель разнообразия* (*выделено нами. – Авторы*)» (Jost, 2006, p. 365). Показатель степени α (и верхний индекс ${}^{\alpha}D$) называется «порядком» разнообразия (Jost, 2006) и в общем случае

$$ {}^{\alpha}D = \left(\sum_{i=1}^{S} p_i^{\alpha}\right)^{1/(1-\alpha)} . \qquad (3.1)$$

Эта формула идентична выражению для чисел Хилла, но Л. Джост претендует на более общий её смысл. Для всех индексов, которые являются функциями $\sum_{i=1}^{S} p_i^{\alpha}$, разнообразие порядка α ${}^{\alpha}D$ зависит только от значения α и видовых частот p, а не от функциональной формы индекса. Это означает, что, оценивая разнообразие исследуемого сообщества, не имеет значения, используется ли индекс Симпсона, его обратное значение, индекс Джини, энтропия Реньи при $\alpha = 2$ или индекс Хелберта–Смита–Грасле (Hurlbert–Smith–Grassle): все дают одно и то же разнообразие второго порядка ${}^2D = 1/\left(\sum_{i=1}^{S} p_1^2\right)$. Сопоставление значений некоторых индексов видового разнообразия и эффективного числа видов по Джосту для речной системы Байтуган–Сок представлено в табл. 3.1.

Таблица 3.1

Конвертирование популярных индексов видового разнообразия x в значения «истинного разнообразия» Джоста D_x (на примере сообществ макрозообентоса системы Байтуган–Сок)

Используемый индекс	Формулы для расчета индекса x и перевода его в D_x	α Реньи	Участки рек и водотоки		
			р. Байтуган	р. Сок (верховья)	р. Сок (нижн. теч.)
Число видов	$x = \sum_{i=1}^{S} p_i$		144	205	176
	${}^0D_x = x$				
Энтропия Шеннона	$x = -\sum_{i=1}^{S} p_i \ln p_i$		3,52	3,27	3,24
	${}^1D_x = e^x$		33,7	26,3	25,5
Индекс Симпсона–Джини	$x = 1 - \sum_{i=1}^{S} p_i^2$		0,94	0,92	0,92
	${}^2D_x = 1/(1-x)$		16,3	12,8	12,7
HCDT-энтропия (Tsallis, 1988)	$x = (1 - \sum_{i=1}^{S} p_i^{\alpha})/(\alpha - 1)$	0,5	14,1	13,8	13,2
		1,5	1,57	1,51	1,50
	${}^{\alpha}D_x = [1-(\alpha-1)x]^{1/(1-\alpha)}$	0,5	64,7	62,3	57,7
		1,5	21,7	16,6	16,2
Энтропия Реньи	$x = (-\ln \sum_{i=1}^{S} p_i^{\alpha})/(\alpha - 1)$	0,5	4,17	4,13	4,06
		1,5	3,08	2,81	2,78
	${}^{\alpha}D_x = e^x$	0,5	64,7	62,3	57,7
		1,5	21,7	16,6	16,2

Уравнению (3.1) интуитивно приписываются свойства «истинного» разнообразия. Во-первых, оно всегда точно для всех значений α дает число видов S «эквивалентного сообщества» с равномерно распределенным обилием. Во-вторых, для всех значений α оно также обладает свойством «удвоения», представленного М. Хиллом (Hill, 1973): если численность каждого вида разделить пополам на два новых таксона, то разнообразие удвоенного сообщества всегда окажется $2 \times {}^{\alpha}D$ независимо от характера распределения p и порядка разнообразия α.

В последнее десятилетие в физических и математических исследованиях для описания сложных систем стал использоваться фрактальный анализ (Мандельброт, 2002), изучающий самоподобные структуры, отдельные компоненты которых инвариантны целому. В ряде работ (Иудин и др., 2003; Гелашвили и др., 2004, 2008а; Якимов и др., 2005) рассмотрены приложения мультифрактального анализа для описания распределения видов в сообществах и оценки структурного разнообразия экологических систем.

Рассмотрим произвольный фрактальный объект в евклидовом пространстве с размерностью d, который представляет собой множество из $\mathbf{N}_{\text{total}} >> 1$ точек, по какому-то закону распределенных в некоторой ограниченной области $\mathfrak{I}$ размером L. Разобьем всю область $\mathfrak{I}$ на ячейки со стороной $\varepsilon << L$ и объемом ε^d, в результате чего общее количество занятых ячеек, в которых содержится хотя бы одна точка, будет равно $N(\varepsilon)$. Пусть величины p_i, $i = 1, 2, \ldots, N(\varepsilon)$ характеризуют относительную заселенность ячеек, т. е. представляют собой вероятности того, что наугад взятая точка из фрактального множества находится в ячейке i.

Спектр обобщенных фрактальных размерностей D_q характеризует текущее распределение точек в области $\mathfrak{I}$ и определяется с помощью соотношений

$$D_q = \frac{\tau(q)}{q-1} \quad \text{и} \quad \tau(q) = \lim_{\varepsilon \to 0} \frac{\ln\left(\sum_{i=1}^{N(\varepsilon)} p_i^q(\varepsilon)\right)}{\ln \varepsilon},$$

где q – показатель степени (момент), который может принимать любые значения в интервале $-\infty < q < +\infty$. Если $D_q = D = \text{const}$, т. е. не зависит от q, то данное множество точек представляет собой регулярный монофрактал, который характеризуется одинаковой относительной заселенностью всех ячеек ($p_i \approx \varepsilon^D$). Напротив, если функция D_q как-то меняется с q, то рассматриваемое множество точек является мультифракталом и, вследствие его неоднородности, вероятности заполнения ячеек p_i неодинаковы, т. е. показатель степени для разных ячеек принимает различные значения ($p_i \approx \varepsilon^{\alpha i}$). Пусть заданы преобразования Лежандра, такие, что $\alpha = d\tau/dq$ и $f(\alpha) = q\, d\tau/dq + \tau$. Тогда для характеристики мультифрактального множества используется функция мультифрактального спектра, т. е. набор различных значений функции $f(\alpha)$ при разных α, соответствующих размерностям однородных фрактальных подмножеств $\mathfrak{I}_\alpha$, на которое можно разбить исходное множество $\mathfrak{I}$.

Фундаментальные теоретические законы экологии свидетельствуют о том, что биотические сообщества обладают удивительной, хотя и специфической структурной универсальностью. В частности, они демонстрируют самоподобие при изменении численности сообщества, но в то же время структурные особенности внутренних энергетических каналов и иерархия связей биотических сообществ остаются скрытыми от глаз: они не имеют непосредственных геометрических образов, инвариантных относительно преобразований масштаба. Поэтому принимаемые ниже аналогии между фрактальным анализом геометрических объектов и анализом структуры биологических сообществ носят, в определенном смысле, характер не формализованных пока допущений.

С другой стороны, эта связь хорошо «просматривается» на множестве объектов, для которых удалось получить фрактальные спектры видовой структуры:

- гидробиоценозов (сообщества субантарктических морских макроводорослей у острова South Georgia в южной Атлантике [Davenpor et al., 1996], сообщества зоопланктона Чебоксарского водохранилища [Гелашвили и др., 2004, 2006; Шурганова и др., 2005], зообентоса городских озер Нижнего Новгорода [Гелашвили и др., 2007б] и литорали [Азовский, Чертопруд, 1997]);
- наземных экосистем (теоретические аспекты [Scheuring, 1991; Scheuring, Riedi, 1994], растительные сообщества на дюнах (Yakimov et al. 2008), растительность в Калифорнийском заповеднике на серпентинной почве [Green et al., 2003], лесные сообщества горы Донглиншан [Donglingshan; западный Тянь-Шань; Zhang et al., 2006], сообщества орибатидных клещей [Ермилов, 2004; Иудин, 2006], сообщества мелких млекопитающих Нижегородского Поволжья [Гелашвили и др., 2008б; Солнцев, 2009] и крупных млекопитающих Самарской Луки [Гелашвили и др., 2009а]);
- ископаемых сообществ (видовая структура ископаемых сообществ мелких млекопитающих Северного Прикаспия в плейстоцене – голоцене [Гелашвили и др., 2008в; Солнцев, 2009]);
- сообществ симбионтов (сообщества гельминтов мелких млекопитающих Самарской Луки [Гелашвили и др., 2009б; Snegireva, 2010]); сообщества паразитических насекомых на примере наездников-ихневмоид Висимского заповедника [Снегирева и др., 2010]);
- структуры таксономического разнообразия (теоретические аспекты [Burlando, 1990, 1993], виды семейства Proteaceae (Протейные) во флористической области мыса Доброй Надежды [Южная Африка; эта область характеризуется самым высоким процентом эндемиков; Bossuyt, Hermy, 2004], таксономическая структура классов Aves и Mammalia [Ю. Пузаченко, А. Пузаченко, 1996], таксономическое разнообразие сосудистых растений национального парка «Самарская Лука» [Гелашвили и др., 2010]).

В случае биологического сообщества разбиение фрактального пространства на ячейки происходит стационарно, так как для элементов сообщества оно уже изначально разбито на таксономические категории (чаще всего в качестве таковых выступают виды). Биологическое сообщество, как и геометрический мультифрактал, обладает гетерогенностью: например, выделяются доминантные и редкие виды. Подсчет числа видов, составляющих сообщество (т. е. определение видового богатства сообщества), аналогичен подсчету занятых ячеек $N(\varepsilon)$ при определении размерности фрактала. Отсюда для сообщества оценки вероятностей p_i – это доли особей i-го вида. Менее очевидна аналогия в отношении размера ячейки ε. Поскольку мы интересуемся асимптотическим поведением моментов в области больших численностей, формулу для $\tau(q)$ следует трактовать как функцию от общей численности N:

$$\tau(q) = \lim_{\varepsilon \to 0} \frac{\ln(\sum_{i=1}^{N(\varepsilon)} p_i^q(N))}{\ln N}.$$

Являясь обобщенным геометрическим образом видовой структуры, мультифрактальные спектры (см. рис. 3.2) позволяют сделать детальный анализ таксономической «выравненности» (эквитабельности) отдельных изолированных сообществ, выделить степень доминирования отдельных групп, уточнить удельный вес редких видов и т. д. Их использование при оценке биоразнообразия является непротиворечивым, если анализируются закономерности распределения численностей по видам при увеличении количества проб из некоторого однородного биотопа; в этом случае суммарная численность N соответствует приписываемой ей дефиниции «размер фрактальной ячейки». Однако при сравнении двух или нескольких разнородных экологических объектов с использованием мультифрактального анализа остаются открытыми вопросы об уровне гетерогенности сообществ и необходимой эквивалентности выборочных усилий.

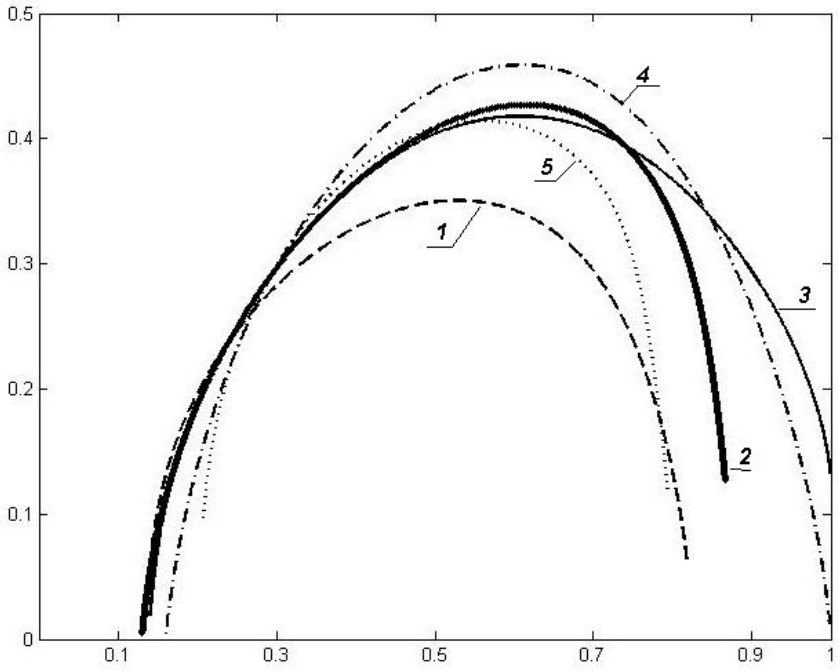


Рис. 3.2. **Мультифрактальные спектры видовой структуры макрозообентоценозов малых рек бассейна Нижней Волги:** 1 – нижнее течение р. Чапаевка; 2 – верхнее течение р. Сок; 3 – нижнее течение р. Сок; 4 – р. Байтуган; 5 – р. Б. Кинель

Например, если макрозообентос р. Сок на всем его протяжении является единым сообществом, то мультифрактальные спектры видовой структуры отдельных участков должны совпадать вне зависимости от способа деления водотока на фрагменты. Что это не вполне так, легко убедиться, анализируя рис. 3.2, где отчетливо видно расхождение кривых (2) и (3) в правой части графика, отражающих долю видов с небольшой численностью. С другой стороны, все четыре представленных сообщества достаточно сходны по степени доминирования, поскольку спектры в левой части графика практически совпадают. Можно также наглядно отметить относительно высокое (в смысле выравненности вероятностей p_i) разнообразие макрозообентоса р. Байтуган, спектральная кривая которой (4) в центральной части графика располагается над остальными.

В то же время при изучении графиков на рис. 3.2 возникает ряд вопросов. Например, можно ли с позиций формальной статистики принять гипотезу о ничтожности расхождений между кривыми 2, 3, 4 (т. е. об однородности видовой структуры рек Байтуган–Сок), или стоит полагаться на свойства человеческого глаза, улавливающего даже небольшой «зазор»? А как поведут себя спектральные кривые, если построить их по данным гидробиологических проб сначала для каждой из 22 станций наблюдения, а потом последовательно объединять их в группы, пытаясь обнаружить границы компактных характерных зон этой речной экосистемы?

Исследование приложений фрактальной геометрии к анализу видовой структуры находится в своей начальной фазе. Методика интерпретации спектров вызывает впечатление некоторой незавершенности, поскольку основана на субъективных рассуждениях и визуальных ощущениях исследователя о степени совпадения (или несовпадения) сравниваемых кривых, конфигурации спектра на отдельных участках, его предположительной связи с различными аспектами разнообразия и т. д. Метод, несомненно, приобрел бы новых единомышленников, если бы удалось разработать набор типовых шаблонов для чет-

кой классификации сообществ и объективные критерии оценки близости двух спектральных кривых. Желательна также строгая статистическая процедура проверки адекватности моно- и мультифрактальной гипотез, позволяющая оценить степень гетерогенности и неравновесности изучаемых экосистем.

«При применении фрактального подхода в теоретической экологии важно помнить и о его ограничениях. Наиболее принципиальное обстоятельство связано с локализацией самоподобия в определенном интервале масштабов. Это имеет отношение как к пространственной, так и видовой структуре сообщества» (Гелашвили и др., 2008а, с. 32). Поэтому насущна необходимость разработки динамических фрактальных моделей нового поколения, позволяющих обобщить последовательность статических состояний экосистемы (например, в виде пространственных или временных её срезов) и показать структурно-топологическую динамику мультифрактальных спектров.

3.2. Оценка сходства видовой структуры сообществ

В различных областях экологии (таксономический анализ, классификация сообществ, оценка видового соответствия фауны и флоры, дифференциация биоразнообразия) используется большое количество показателей сходства, имеющих разнообразную математическую природу. Многочисленные литературные источники (Bray, Curtis, 1957; Sokal, Sneath, 1963; Василевич, 1969; Goodall, 1973; Миркин, Розенберг, 1978а; Сёмкин, 1979; Песенко, 1982; Миркин и др., 1989; Шитиков и др., 2005) приводят длинные списки наиболее известных индексов и предпринимают попытки сделать хотя бы приблизительную их классификацию. Можно отметить следующую возможную типологию мер сходства:

- основанные на данных о присутствии/отсутствии видов (в англоязычной литературе – «*инцидентности видов*» – species incidence) или использующие значения численности особей (обилие видов);
- имеющие различное математическое обоснование (подсчет количества элементов подмножеств, оценка расстояния в многомерном метрическом пространстве, использование теоретико-информационных или статистических методов и т. д.);
- абсолютные меры или относительные коэффициенты, связанные с фиксированными диапазонами варьирования.

«Мера сходства/расстояния» является типичным искусственно сконструированным собирательным понятием, отражающим латентные (т. е. скрытые, принципиально не измеряемые инструментально) свойства реальных объектов и включающим некоторые, выработанные практикой разумные критерии для качественного сопоставления этих объектов (P. Legendre, L. Legendre, 1998). «*Подобие* – человеческая конструкция; у него нет и не может быть никакого точного математического определения» (Chao et al., 2005, p. 157). В связи с этим любые шкалы сходства являются по определению субъективными и нерепрезентативными (т. е. нельзя выстроить четкую систему «эталонов», по которой предполагается смысловая оценка анализируемых объектов на основании интенсивности измеряемых свойств). Относительные шкалы, нормированные в пределах некоторого диапазона, как правило, поверхностны и менее информативны, чем абсолютные: например, *коэффициент Жаккара* равен 0,5 как при совпадении 10 видов из 20, так и при 200 общих видов из 400, что для существа дела иногда имеет немаловажное значение.

В ряде работ (Миркин, Розенберг, 1978а) делаются попытки оценить, какие меры «завышают» или «занижают» сходство между сообществами и каким коэффициентам следует отдать предпочтение в работе. Однако вряд ли имеет смысл акцентировать эту сторону проблемы, так как единственным критерием адекватности оценок является последовательность агрегирования объектов на основании меры сходства в более крупные таксоны, иерархические деревья и пр.

Одним из важных шагов по упорядочению используемых оценок явилось формулировка понятий «*эквивалентности*» и «*коэквивалентности*» мер сходства. Согласно теореме Б.И. Сёмкина и В.И. Двойченкова (1973), две меры r_1 и r_2 эквивалентны, если они связаны монотонно возрастающей зависимостью φ, т. е. $r_1 = \varphi(r_2)$. Понятие «эквивалентности» имеет важное следствие: если два индекса сходства эквивалентны, то они приводят к одной и той же последовательности объектов, упорядоченных по степени их сходства – близкие объекты остаются близкими и наоборот. Например, эквивалентные *коэффициенты Жаккара* J и *Съеренсена* L, где $L = 2J/(J + 1)$, приводят к качественно идентичным иерархическим деревьям сходства. То же можно сказать и об использовании более «сложных» формул, которые часто создают только иллюзию объективности и точности классификации. В связи с этим споры о том, какой коэффициент из некоторого семейства эквивалентных мер является лучшим, можно считать беспредметными.

Все индексы, основанные на регистрации факта отсутствия/присутствия вида, используют только три составляющих простого подсчета инцидентности: c – количество видов, общих для двух сообществ; a и b – число видов, уникальных для каждого из сравниваемых множеств. В своей работе Б.И. Сёмкин (1979) предложил целый континуум мер сходства, обладающих свойством эквивалентности и объединенных формулой

$$K = \frac{2c}{(1+u)(a+b+2c) - 2u \cdot c}, \qquad (3.2)$$

где параметр u может принимать любые значения от -1 до ∞. Нетрудно заметить, что при $u = 0$ мы имеем хорошо известный *коэффициент Съеренсена*; мера при $u = 1$ численно совпадает с *коэффициентом Жаккара*, при $u = 3$ – с *коэффициентом Сокала–Снита* и т. д.

Индексы, основанные на инцидентности, нашли широкое и вполне оправданное применение в экологических исследованиях, особенно в тех областях, где нет понятия «особи» и обилие видов оценивается в баллах (геоботаника, фитопланктонология и др.). Однако когда популяционная численность в пробах может быть установлена, то часто возникает необходимость провести анализ с учетом экологической значимости («важности») видов, принимая во внимание относительное их обилие. Поскольку для семейства индексов (3.2) виды с высокой и низкой численностями особей рассматриваются одинаково, интерпретация результатов сравнения двух или более сообществ становится весьма проблематичной, особенно, если данные мониторинга ограничены недостаточно репрезентативными выборками, содержащими многочисленные редкие виды.

Действительно, почти все существующие подходы к вычислению мер сходства основаны на предположении (обычно неправильном), что по выборочным данным можно составить истинное и полное представление о видовых композициях сообществ, поскольку фактически невозможно обнаружить все таксоны и оценить их относительное обилие на основе разумно ограниченного числа наблюдений. Выборочные ограничения препятствуют точным оценкам α-разнообразия в пределах локальных, приблизительно гомогенных местообитаний, особенно для сообществ с высоким богатством и большой фракцией редких видов. Всегда существует некоторое подмножество «невидимых» видов, находящихся за «*линией занавеса*» *Престона*[3], за счет которых при росте выборочного усилия (т. е. числа сделанных проб) увеличивается α-разнообразие сообщества. При этом, как показывают исследования (Fisher, 1999; Chao et al., 2005), повышается вероятность того, что вновь регистрируемые виды будут принадлежать к общим видам двух сравниваемых сообществ.

[3] Логнормальное распределение описывается симметричной «нормальной», т. е. колоколообразной кривой. Однако если данные, которым она соответствует, получены из ограниченной выборки, то левая часть кривой (т. е. редкие, неучтенные виды) будет выражена нечетко. Ф. Престон (F. Preston) назвал такую точку усечения кривой слева «линией занавеса». «Линия занавеса» может сдвигаться влево при увеличении объема выборки (География и мониторинг.., 2002).

Очевидно также, что в случае, когда интенсивность взятия проб в сравниваемых местообитаниях существенно различается, то многие виды будут «*поддельно*» отсутствовать в одной из композиций с минимальным выборочным усилием.

Любые индексы (3.2), основанные на инцидентости, весьма чувствительны к подобным ситуациям и, как результат, почти всегда серьезно недооценивают уровень сходства сообществ (Morisita, 1959; Ricklefs, Lau, 1980). В принципе, может также произойти переоценка подобия, если сравниваются сообщества с высоким уровнем доминирования, в которых фоновые виды широко распространены, а редкие имеют тенденцию быть локально эндемичными (в работе [Ruokolainen, Tuomisto, 2002] обсуждается такой уникальный случай).

А. Чао с соавторами (Chao et al., 2005) обозначили вероятностный подход для индексов Жаккара и Съеренсена, основанных на инцидентности видов. Предположим, что мы случайным образом выбираем по одному виду из двух сообществ a и b, а затем проверяем каждую такую пару на совпадение (т. е. был ли при этом идентифицирован один и тот же вид). Если в результате таких испытаний выбрано s_a видов сообщества a и s_b видов сообщества b, из которых s_{ab} видов оказалось общими, то мощности этих подмножеств являются составляющими формулы Съеренсена: $L = \frac{2c}{(a+b+2c)} = \frac{2s_{ab}}{s_a + s_b}$. Отсюда легко перейти к оценкам вероятности u и v отнесения видов к подмножествам

$$L = \frac{2[(s_{ab}/s_a)(s_{ab}/s_b)]}{[(s_{ab}/s_a)(1-(s_{ab}/s_b)]+[(s_{ab}/s_b)(1-(s_{ab}/s_a)]+2[(s_{ab}/s_a)(s_{ab}/s_b)]} = \frac{2uv}{(u+v)}. \qquad (3.3)$$

Хотя в результате преобразований была получена знакомая классическая формула для подсчета индекса Съеренсена, вероятностный подход создает единую основу для разработки семейства обобщенных мер, использующих оценки вероятности встречаемости, как видов, так и отдельных особей.

Выше отмечалось, что некоторое количество дополнительных видов потенциально являются общими для сравниваемых сообществ, однако из-за принципиальной неполноты данных мониторинга не смогли встретиться в одном из подмножеств проб a или b, либо в обоих подмножеств вместе. Чао с соавторами (Chao et al., 2005) полагают, что число таких *невидимых* общих видов (unseen shared species) тем больше, чем выше частота наличия среди *наблюдаемых* общих видов уникальных таксонов, встретившихся только в одной или двух выборках. Иными словами, для прогноза числа видов, находящихся за «линией занавеса» Престона, используется выборочная информация о числе уже найденных редких видов.

Введем следующие обозначения:
M_a и M_b – количество проб (выборок), сделанных в сообществах a и b соответственно; $P_a(i)$ и $P_b(i)$ – частоты встречаемости вида i в этих пробах; $f_{a1} = \sum_{i=1}^{s_{ab}} I[P_a(i)=1, P_b(i)\geq 1]$ – число наблюдаемых общих видов, которые встретились только в одной пробе из сообщества a и в любом ненулевом числе проб из сообщества b; $f_{a2} = \sum_{i=1}^{s_{ab}} I[P_a(i)=2, P_b(i)\geq 1]$ – число наблюдаемых общих видов, которые встретились в двух пробах из сообщества a и в любом (ненулевом) числе проб из сообщества b; $f_{b1} = \sum_{i=1}^{s_{ab}} I[P_b(i)=1, P_a(i)\geq 1]$ и $f_{b2} = \sum_{i=1}^{s_{ab}} I[P_b(i)=2, P_a(i)\geq 1]$ – аналогично для сообщества b.

В этих формулах I['выражение'] – индикаторная функция, принимающая значение 1, если верно выражение, являющееся её параметром, и 0, если 'выражение' ложно. Тогда оценку $\hat{C}$ вероятности того, что случайно взятый вид из сообщества a встретится

также и в сообществе b можно рассчитать по формуле

$$w = \sum_{i=1}^{s_{ab}} P_a(i)/M_a + \frac{(M_b-1)}{M_b}\frac{f_{b1}}{2f_{b2}}\sum_{i=1}^{s_{ab}}[P_a(i)/M_a]\cdot I[P_b(i)=1].$$

Аналогично для сообщества b: $v = \sum_{i=1}^{s_{ab}} P_b(i)/M_b + \frac{(M_a-1)}{M_a}\frac{f_{a1}}{2f_{a2}}\sum_{i=1}^{s_{ab}}[P_b(i)/M_b]\cdot I[P_a(i)=1]$.

Полученные оценки вероятностей подставляются в общую формулу (3.3) вероятностной интерпретации меры Съеренсена: $L_I = \frac{2uv}{(u+v)}$.

Большинство индексов сходства, основанных на численности видов, используют те же идеи, что и меры, основанные на инцидентности. Например, *индекс Ренконена* (Renkonen, 1938), использующий сумму минимумов обилия, варьируется в диапазоне от 0 до 1 и определяется как

$$\xi_R(a,b) = \frac{2\sum_{i=1}^{s_{ab}}\min[N_a(i),N_b(i)]}{\sum_{i=1}^{s_a}N_a(i)+\sum_{i=1}^{s_b}N_b(i)}, \qquad (3.4)$$

где $N_a(i)$ и $N_b(i)$ – численности вида i в сообществах a и b соответственно; s_{ab} – общее число видов. Если индекс Съеренсена, по существу, является долей общих видов в средней видовой насыщенности сравниваемых выборок, то индекс Ренконена определяет долю числа совпадающих особей относительно средней численности особей рассматриваемых сообществ. Эта мера сходства также упоминалась (цит. по: Plotkin, Muller-Landau, 2002) как «*индекс Штейнгауза*» (Steinhaus index; Motyka, 1947), «*индекс Съеренсена с покрытием*», «*процентное подобие*» и «*коэффициент общности*». С небольшими изменениями индекс Ренконена часто называют *коэффициентом Чекановского* (Czekanowcki, 1911), *мерой Брея–Кёртиса* (Bray, Curtis, 1957) или «*процентным различием Одума*» (Odum, 1950). Нетрудно заметить его аналогию и с «*коэффициентом общности удельного обилия*», предложенным А.А. Шорыгиным (1939, 1946), для которого немного изменены условия нормировки:

$$\xi_S(a,b) = \sum_{i=1}^{s}\min[N_a(i)/\sum_{i=1}^{s}N_a(i), N_b(i)/\sum_{i=1}^{s}N_b(i)].$$

Широкое распространение получили также два *индекса Мориситы* (Morisita, 1959), в основе которых лежат вероятности обнаружения видов (Krebs, 1989):

- оригинальный $\xi_{OM}(a,b) = \frac{2\sum_{i=1}^{s_{ab}}[N_a(i)/T_a]\cdot[N_b(i)/T_b]}{\sum_{i=1}^{s_a}[N_a(i)/T_a]\cdot\{[N_a(i)-1]/[T_a-1]\}+\sum_{i=1}^{s_b}[N_b(i)/T_b]\cdot\{[N_b(i)-1]/[T_b-1]\}}$,

- и Мориситы–Горна (Horn, 1966) $\xi_{MH}(a,b) = \frac{2\sum_{i=1}^{s_{ab}}[N_a(i)/T_a]\cdot[N_b(i)/T_b]}{\sum_{i=1}^{s_a}[N_a(i)/T_a]^2+\sum_{i=1}^{s_b}[N_b(i)/T_b]^2}$, (3.5)

где T_a и T_b – суммарная численность особей сообществ a и b соответственно. Если предположить, что особи выбираются из каждого сообщества случайно, то числитель приведенных формул определяет вероятность того, что две произвольных особи принадлежат одному и тому же виду из подмножества общих видов, а знаменатель представляет нормирующую постоянную, которая является средним значением вероятностей обнаружения вида для этих двух сообществ. Смыл поправки Г. Горна к индексу М. Мориситы заключается только в том, что ξ_{MH} варьируется строго в диапазоне от 0 до 1.

Ряд исследователей (Horner-Devine et al., 2004; Chao et al., 2005) считают, что индекс Ренконена и ему подобные имеют смысл, если выборочные фракции (т. е. число сделанных проб и количество обнаруженных при этом видов) обоих сравниваемых сообществ являются приблизительно эквивалентными. В противном случае эти индексы ведут себя непредсказуемо, что подтверждается результатами имитационного моделирования.

Следует отметить, что отличительной чертой всех индексов, основанных на обилии, является очень высокая чувствительность к конспецифичным (относящимся к одному и тому же виду) скоплениям особей (Wolda, 1981; Magurran, 2004). Это интуитивно понятно из их вероятностной интерпретации: немногие виды с высокой численностью особей вносят главную часть вероятности того, что две случайно выбранных особи принадлежат к одинаковому виду. В результате значения этих индексов определяются, в основном, популяционной плотностью одного-двух доминирующих таксонов и практически игнорируют подмножество видов с низкой численностью, даже если их количество велико. Являясь относительно нечувствительными к композиционным различиям во фракциях редких видов, индексы типа Ренконена и Мориситы, вероятно, будут более стойкими к недостаточной выборочной репрезентативности (undersampling), поскольку доминирующие виды с высокой численностью особей, как правило, появляются уже в первых пробах.

Степень расхождения оценок сходства по индексам, рассчитанным на основе обилия или присутствия/отсутствия видов, зависит от разнообразия сообщества: при равномерном распределении численностей видов и большом их числе количественные индексы могут дать ту же картину взаимной упорядоченности местообитаний, что и индексы присутствия/отсутствия. В целом вряд ли разумно требовать одинаковых результатов классификации сообществ по их сходству, поскольку оба класса индексов основаны на разных теоретических предпосылках и отражают различные аспекты отличий видовой структуры:

- индексы на основе обилия отражают подобие сообществ по доминирующему комплексу видов и для их успешного использования целесообразно учитывать характер статистического распределения численности популяций, как в сравниваемых местообитаниях, так и в изучаемом регионе в целом;
- индексы, основанные на инцидентности, выделяют композиционные отличия с учетом всего списка видового богатства, делая акцент на редких видах, поэтому их корректное использование возможно, если выборочные данные достаточно полны.

С учетом этих рекомендаций целесообразно тщательно анализировать, что каждый раз понимается под «подобием сообществ» или «наложением видов», и соответственно выбирать тип индекса, принимая во внимание природу и ограничения исходных данных.

Рассмотрим основные идеи теоретико-вероятностного подхода применительно к оценке мер сходства, использующих обилие видов. Для аналитического вычисления (Plotkin, Muller-Landau, 2002) предполагаемой доли совпадения популяционной численности двух участков, удобно представлять суммарное обилие сообщества в форме непрерывной функции плотности вероятности $\phi(N)$. Выражение $\phi(N)\,dN$ означает вероятность того, что численность вида, случайно выбранного из сообщества, будет находиться в интервале между N и $N + dN$. Непрерывная плотность распределения нормирована к полной единичной вероятности: $\int_0^\infty \phi(N)dN = 1$. В действительности распределение численностей видов дискретно, однако показано, что непрерывные распределения обилия приводят почти точно к тем же результатам, что и их дискретные аналоги (May, 1975).

Д. Плоткин и Е. Мюллер-Ландау (Plotkin, Muller-Landau, 2002) осуществили вероятностную интерпретацию формулы Съеренсена, основанную на функциях плотности распределения численностей особей:

$$\chi(a,b) = \frac{2\int_0^\infty \phi(N)\psi(a,N)\psi(b,N)dN}{\int_0^\infty \phi(N)\psi(a,N)dN + \int_0^\infty \phi(N)\psi(b,N)dN}. \qquad (3.6)$$

Здесь $\phi(N)$ отражает распределение обилия видов в большой области (регионе или совокупности водных объектов), а $\psi(a, N)$ соответствует вероятности обнаружения вида в частном местообитании a при условии, что этот вид имеет численность N в большей области, частью которой является местообитание a. Тот же смысл имеет и функция $\psi(b, N)$ для местообитания b, являющегося частью того же региона. Следует отметить, что приведенное уравнение использует плотности распределения вероятности только тех видов, которые являются общими для двух небольших подобластей a и b, поэтому нет необходимости знать полное богатство видов, характерное для всей изучаемой области.

В работе (Plotkin, Muller-Landau, 2002) приведены формулы для расчета значений $\chi(a, b)$ согласно различным предположениям о форме непрерывных и дискретных распределений обилия видов. Например, для *модели «разламываемого» стержня Мак-Артура* (вернее, однопараметрического экспоненциального распределения $\phi(N) = \lambda e^{-\lambda N}$, являющегося её непрерывным аналогом) в результате интегрирования уравнения (3.6) можно получить следующее выражение: $\chi(a,b) = \dfrac{2ab(a+b+2\lambda)}{(a+b+\lambda)(2ab+a\lambda+b\lambda)}$. Аналогично для другого дискретного распределения лог-сериями Фишера $\phi(N) = \alpha x^N / N$ можно записать:

$$\chi(a,b) = \frac{2[\ln(1-xe^{-a}) + \ln(1-xe^{-b}) - \ln(1-x) - \ln(1-xe^{-a-b})]}{\ln(1-xe^{-a}) + \ln(1-xe^{-a}) - 2\ln(1-x)}.$$

Интегрирование уравнения (3.6) для остальных популярных распределений обилия видов непрерывного типа осуществлялось с использованием гамма-функций и приводит к достаточно громоздкой форме записи.

А. Чао с соавторами (Chao et al., 2005), в отличие от «параметрических» формул Д. Плоткина и Е. Миллер-Ландау, предлагают свой подход, не требующий никаких предварительных гипотез о законе распределения популяционных плотностей. Предположим, что все виды сообществ имеют одинаковую способность обнаруживать себя, следовательно, вероятности случайного извлечения особи каждого вида из гомогенной смеси зависят, прежде всего, от их относительного обилия. Пусть для сообществ a и b с богатством видов s_a и s_b, соответственно, можно оценить значения этих вероятностей: $(p_1, p_2, \ldots, p_{Sa})$ и $(\pi_1, \pi_2, \ldots, \pi_{Sb})$, где $p_i > 0$, $\pi_i > 0$ и $\sum_{i=1}^{s_a} p_i = \sum_{i=1}^{s_b} \pi_i = 1$, а s_{ab} – число общих видов, обнаруженных в результате попарного случайного извлечения по одной особи из двух сообществ одновременно. Тогда полное относительное обилие особей, принадлежащих подмножеству общих видов, будет равно $u = \sum_{i=1}^{s_{ab}} p_i$ для сообщества a и $v = \sum_{i=1}^{s_{ab}} \pi_i$ для сообщества b. С учетом этих обозначений формула (3.3) для индекса Съеренсена, учитывающая обилие видов, может быть записана как

$$L_0 = \frac{2uv}{(u+v)} = \frac{2\sum_{i=1}^{s_{ab}} p_i \cdot \sum_{i=1}^{s_{ab}} \pi_i}{\sum_{i=1}^{s_{ab}} p_i + \sum_{i=1}^{s_{ab}} \pi_i}. \qquad (3.7)$$

Используя суммарные вероятности u и v, можно записать аналогичные выражения и для других индексов, основанных на инцидентности (например, для *коэффициента Жаккара*). Отметим, что, в отличие от других вышеприведенных выражений, знаменатель формулы (3.7) представляет собой сумму вероятностей только для совместно встретившихся видов.

По аналогии с приведенной выше схемой расчета по формуле (3.3), Чао с соавторами (Chao et al., 2005) учитывают также количество *невидимых* общих особей, которые

потенциально могли бы встретиться при углублении мониторинговых исследований. Их число тем больше, чем выше частота наличия среди *наблюдаемых* общих видов малочисленных таксонов, представленных одной или двумя особями (т. е. синглетонов и дублетонов).

Если $N_a(i)$ и $N_b(i)$ – наблюдаемая численность вида i в сообществах a и b соответственно, то можно рассчитать следующие выборочные частоты:

$f_{a1} = \sum_{i=1}^{S_{ab}} I[N_a(i) = 1, N_b(i) \geq 1]$ – число наблюдаемых общих видов, которые представлены одной особью в сообществе a, и любым числом особей, большем нуля, в сообществе b;

$f_{a2} = \sum_{i=1}^{S_{ab}} I[N_a(i) = 2, N_b(i) \geq 1]$ – число наблюдаемых общих видов, которые представлены двумя особями в сообществе a.

Другие частоты f_{b1} и f_{b2} также приобретают соответствующий смысл, а $I[\ldots]$ – индикаторная функция, принимающая значение 1 и 0 в зависимости от истинности выражения.

С учетом этого скорректированные оценки вероятностей ¹ и º, подставляемые в формулу (3.7), будут являться суммой двух составляющих: первая обозначает сумму частот, связанных с наблюдаемыми общими видами, а вторая определяет предполагаемый эффект невидимых общих видов:

$$¹ = \sum_{i=1}^{S_{ab}} N_a(i)/T_a + \frac{(T_b - 1)}{T_b} \frac{f_{b1}}{2 f_{b2}} \sum_{i=1}^{S_{ab}} [N_a(i)/T_a] \cdot I[N_b(i) = 1];$$

$$º = \sum_{i=1}^{S_{ab}} N_b(i)/T_b + \frac{(T_a - 1)}{T_a} \frac{f_{a1}}{2 f_{a2}} \sum_{i=1}^{S_{ab}} [N_b(i)/T_b] \cdot I[N_a(i) = 1],$$

где T_a и T_b – суммарная численность особей сообществ a и b соответственно.

Эмпирическая оценка вероятностей u и v по выборочным данным в формуле для представленной меры сходства, основанной на обилии, связаны с двумя проблемами.

Во-первых, понятие «малой численности» особей весьма относительно. Например, при изучении популяций носорогов единичная особь в сообществе – вполне обыденное явление, в то время как для штаммов бактерий несколько тысяч экземпляров могут соответствовать редкому виду. Чтобы корректно оценить частоты синглетонов $f_{_1}$ и дублетонов $f_{_2}$ для каждого типа сообщества, необходимо вводить масштабирующие коэффициенты, учитывающие статистический разброс наблюдаемого числа экземпляров, которые, в той или иной степени, всегда будут носить субъективный характер.

Во-вторых, значения «численностей особей $N_a(i)$ различных видов сообщества», как правило, неизвестны экологам. Их можно оценить только статистически по данным экологического мониторинга, проанализировав характер распределения популяционной плотности во времени и пространстве. В любом случае, для большинства гидробиологических сообществ оценивание математических ожиданий численностей популяций не является тривиальным. Например, в 11 пробах макрозообентоса на станции 5 р. Сок за 1992-1999 гг. вид *Paracladius conversus* встретился следующим числом особей {0, 0, 960, 0, 10, 0, 0, 0, 0, 0, 10}. Большинство элементарных статистик этого вариационного ряда вряд ли удовлетворят эколога-аналитика: медиана и мода равны 0, среднее геометрическое – 2, среднее арифметическое – 89. Следует признать, что единственный правдоподобный вывод, который можно сделать в представленном, весьма типичном примере: вид *Paracladius conversus* действительно встретился в 3 пробах из 11. Это обстоятельство заставляет предположить, что для донных сообществ более предпочтительно использование формулы (3.3), предложенной А. Чао с соавторами (Chao et al., 2005).

В качестве примера зададимся целью установить, имеются ли отличия в видовом составе трех сообществ донных организмов: в р. Байтуган (31 выполненная проба, 144 найденных видов) и на двух участках р. Сок в верхнем (53 пробы, 206 видов) и нижнем её

течении (46 проб, 176 видов). Был выполнен расчет индексов сходства по различным формулам, учитывающим показатели обилия, в качестве которых использовались частоты встречаемости видов $P_a(i)$ и $P_b(i)$ в гидробиологических пробах (табл. 3.2).

Таблица 3.2

Анализ сходства видового состава сообществ макрозообентоса некоторых малых рек Самарской области с использованием индексов, вычисленных по различным алгоритмам

Показатели сходства видового состава сообществ	Сравниваемые композиции видов		
	р. Байтуган – р. Сок (ВТ)	р. Сок (ВТ) – р. Сок (НТ)	р. Байтуган – р. Сок (НТ)
Число общих видов	79	101	57
Индекс Жаккара (3.2)	0,291 ± 0,0179	0,359 ± 0,0179	0,217 ± 0,0159
Индекс Съеренсена, основанный на инцидентности видов (3.2)	0,451 ± 0,0223	0,529 ± 0,0207	0,356 ± 0,0227
Индекс Леннона с соавторами (Lennon et al., 2001)	0,549 ± 0,0300	0,574 ± 0,0238	0,396 ± 0,0268
Вероятностный индекс Съеренсена (3.3)	0,687 ±0,0210	0,705 ± 0,0203	0,489 ±0,0232
Вероятностный индекс Съеренсена с поправкой на скрытые виды (3.3)	0,775 ± 0,0349	0,833 ± 0,0345	0,676 ± 0,042
Индекс Ренконена–Брея–Кёртиса (3.4)	0,473 ± 0,0203	0,461 ± 0,0182	0,294 ± 0,0166
Индекс Мориситы–Горна по встречаемости в пробах (3.5)	0,804 ± 0,0232	0,658 ± 0,0290	0,496 ± 0,0329
То же – оригинальный индекс Мориситы	0,929 ± 0,0257	0,728 ± 0,0316	0,566 ±0,0369
Индекс Мориситы по среднему обилию видов	0,415	0,582	0,216

Примечание: ВТ – верхнее течение, НТ – нижнее течение.

Легко установить по всему комплексу показателей, что между р. Сок в её нижнем течении и р. Байтуган имеются очевидные различия в видовой структуре, в то время как участок р. Сок в верховьях занимает промежуточное положение между ними. С гидробиологической точки зрения эта ситуация легко объяснима, если вспомнить, что р. Байтуган является притоком р. Сок в её верхнем течении, что определяет активный обмен видами между смежными участками. В то же время приведенный простейший пример показывает некоторую противоречивость выводов, основанных на индексах различной природы: меры, использующие «суммы минимумов» (3.4), сочли, что верховья Сока очень близки по видовому составу р. Байтуган, тогда как различные варианты формул, основанных на инцидентности (3.2), обозначили более серьезные отличия.

Можно также отметить, что индекс Мориситы, рассчитанный по средней численности особей в пробах, взятых в разных водотоках, показал существенно более низкий уровень подобия, чем аналогичный показатель, использующий частоту встречаемости видов в пробах. Поскольку первая версия индекса подчеркивает отношения между подмножествами видов с высоким обилием, можно предположить, что отличия между сообществами макрозообентоса рек Байтуган и Сок в его верхнем течении обусловлены не общим видовым составом, а отдельными доминирующими таксонами.

Расчеты были выполнены по программе **SPADE** (Chao, Shen), которая включает дополнительную процедуру оценки стандартных ошибок для рассчитываемых индексов сходства. Оценка дисперсий осуществлялась бутстреп-методом (bootstrap; Chao et al., 1992; Шитиков и др., 2008), который выполняет генерацию большого количества псевдо-

выборок с использованием случайных комбинаций видов, входящих в состав сравниваемых сообществ. Из этих выборок в процессе имитации многократно формируются подмножества M_a и M_b искусственных «проб», между которыми и рассчитываются псевдозначения индекса сходства. Полученное статистическое распределение отражает изменчивость показателя и позволяет оценить его стандартную ошибку $\hat{\sigma}_L$. Значения $\hat{\sigma}_L$, представленные в табл. 3.2, вычислены в ходе бутстреп-процедуры, повторенной 200 раз.

Значения стандартных ошибок, полученные непараметрическими методами, дают возможность оценивать статистическую значимость индексов сходства и проверять гипотезы о соотношении видового состава сообществ. Доверительный интервал для истинного значения λ индекса сходства задается выражением $P(\bar{L} - t_\alpha \hat{\sigma}_L < \lambda < \bar{L} + t_\alpha \hat{\sigma}_L) > (1 - \alpha)$, где t_α – критерий Стьюдента для принятого уровня значимости α, P – вероятность того, что λ не выйдет за пределы полуразмаха $t_\alpha \hat{\sigma}_L$ относительно оцениваемого значения $\bar{L}$. На рис. 3.3 показано, что при уровне значимости $\alpha = 0{,}05$ значения индексов сходства Съеренсена (3.3) и Мориситы–Горна (3.5) с вероятностью 95% будут находиться в обозначенных диапазонах (можно отметить, что они не выходят за рамки отведенных им значений от 0 до 1).

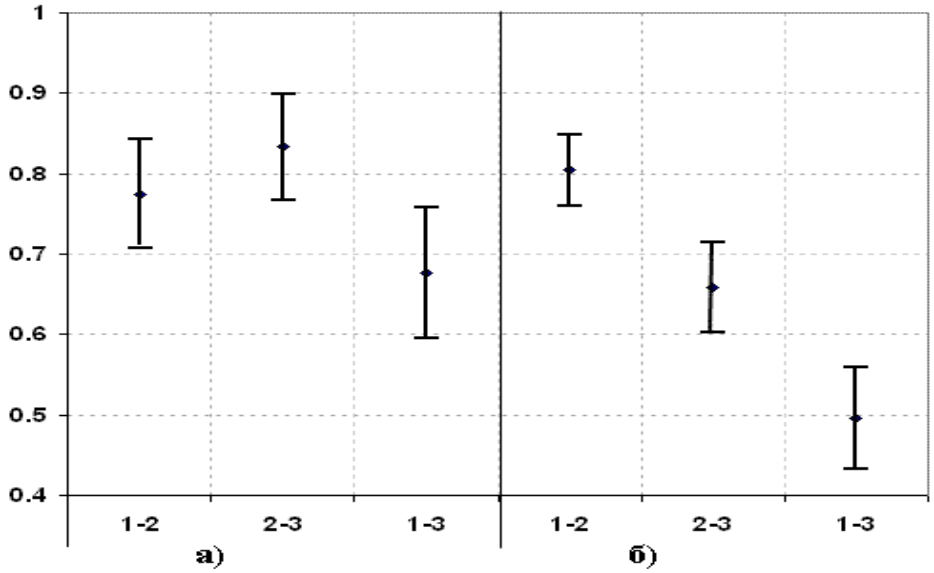


Рис. 3.3. **Доверительные интервалы индексов сходства Съеренсена с поправкой на скрытые виды (а) и Мориситы–Горна (б):**
по оси абсцисс – сравниваемые пары водотоков, обозначенные порядковыми номерами 1 – р. Байтуган, 2 и 3 – р. Сок в верхнем и нижнем течении соответственно

Используя найденные доверительные границы, можно проверить обоснованность предположения о том, какая пара сравниваемых сообществ является более сходной. Например, нельзя определенно сказать, что верхнее течение р. Сок по своему видовому составу ближе к низовьям этой реки, чем к р. Байтуган, поскольку доверительные интервалы коэффициентов сходства $L_{1\text{-}2}$ и $L_{2\text{-}3}$ по формуле Съеренсена пересекаются (рис. 3.3). В соответствии с этим нетрудно предположить, что *многочисленные дендрограммы, представленные в экологической литературе, носят приблизительный и вариативный характер, поскольку последовательность разбиений на кластеры, как правило, четко не подтверждена на основе общепринятых статистических процедур проверки гипотез.*

Оценка подобия видовой структуры двух произвольных локальных сообществ выглядит достаточно корректно, если ставится частная задача парного сравнения. Однако, если рассматривать отдельные фрагменты экосистемы как случайную выборку из большей области, то необходим переход к обобщенным мерам сходства для множества участков больше двух.

Если следовать обычной системе обозначений для индекса сходства Съеренсена, то мера подобия трех участков A, B и C будет иметь вид $L_3 = \frac{3}{2}\frac{ab + ac + bc - abc}{(a + b + c)}$, где a, b и c – совокупности видов, представленных в каждом из участках, ab – подмножество видов, общих для участков A и B, и т. д. до abc, которая определяет виды, найденные одновременно во всех трех местообитаниях. В работе (Diserud, Ødegaard, 2007) этот подход обобщен для случая m участков и новая версия «множественного» индекса сходства Съеренсена, отражающего ковариационные отношения внутри региона, будет иметь вид

$$L_m = \frac{m}{m-1}\left(1 - \frac{S_m}{\sum_i a_i}\right), \text{ где } S_m = \sum_i a_i - \sum_{i<j} a_{ij} + \sum_{i<j<k} a_{ijk} - \sum_{i<j<k<l} a_{ijkl} + \ldots, \qquad (3.8)$$

где a_i видовой состав на участке A_i, $i = 1, \ldots, m$; a_{ij} – число видов, общих для участков A_i и A_j; a_{ijk} – общая видовая насыщенность участков A_i, A_j и A_k и т. д. Легко заметить, что при $m = 2$ формула для индекса Съеренсена приобретает свой обычный вид.

Аналогично может быть получена обобщенная формула, например, индекса Мориситы, использующая вероятности встречаемости особей и оценивающая множественный уровень сходства более чем двух сообществ. Для того чтобы сравнить видовой состав m сообществ, предложен профиль из $(m - 1)$ индексов $C_{2m}, C_{3m}, \ldots, C_{mm}$, где C_{jm} – обобщенный показатель подобия, основанный на информации об общих видах для различных комбинаций j сообществ из m (Chao et al., 2008). Например, обобщенный индекс Мориситы для трех сообществ состоит из профиля двух индексов:

$$C_{23} = \frac{\sum_{i=1}^{s}\left[\frac{N_{i1}N_{i2}}{T_1T_2} + \frac{N_{i1}N_{i3}}{T_1T_3} + \frac{N_{i2}N_{i3}}{T_2T_3}\right]}{\sum_{i=1}^{s}\left[\frac{N_{i1}^{(2)}}{T_1^{(2)}} + \frac{N_{i2}^{(2)}}{T_2^{(2)}} + \frac{N_{i3}^{(2)}}{T_3^{(2)}}\right]} \text{ и } C_{33} = \frac{\sum_{i=1}^{s}\left[\frac{3N_{i1}^{(2)}N_{i2}}{T_1^{(2)}T_2} + \frac{3N_{i1}N_{i2}^{(2)}}{T_1T_2^{(2)}} + \ldots + \frac{6N_{i1}N_{i2}N_{i3}}{T_1T_2T_3}\right]}{8\sum_{i=1}^{s}\left[\frac{N_{i1}^{(3)}}{T_1^{(3)}} + \frac{N_{i2}^{(3)}}{T_2^{(3)}} + \frac{N_{i3}^{(3)}}{T_3^{(3)}}\right]},$$

где N_{ij} – численность i-го вида в j-м сообществе, T_j – суммарная численность, S – количество встретившихся видов, $x^{(k)} = x(x - 1)\ldots(x - k + 1)$.

Используя данные табл. 3.2, можно вычислить:

- среднее значение парного индекса Мориситы (0,415 + 0,582 + 0,216)/3 = 0,404;
- профиль $C_{23} = 0{,}412$, который интуитивно можно расценить как меру сравнения трех сообществ, основанную на общности видов между любыми двумя сообществами;
- глобальную меру $C_{33} = 0{,}384$, принимающую во внимание полную комбинаторику соотношений особей, принадлежащих общим видам, для любых двух и трех сообществ.

Снижение индекса C_{33} по сравнению с C_{23} означает, что имеется, по крайней мере, одно сообщество, которое отличается от других. Формулы для расчета профилей при $m > 3$, а также их интерпретация с использованием условных вероятностей приведены в (Chao et al., 2008).

Традиционным методом, осуществляющим оценку статистической значимости воздействия одного или нескольких факторов на совокупность параметров состояния экосистемы, является *многомерный дисперсионный анализ*. Предположим, что в ходе мониторинга природного объекта было сделано m проб, которые можно разделить на r групп, например, в соответствии со схемой зонирования водотока, где выполнялись гидробиологические пробы. Тогда общую изменчивость численности i-го вида можно разложить на

компоненты: Var N_i = Var τ + Var ε, где Var τ – вариация, обусловленная влиянием группирующего фактора, Var ε – изменчивость, связанная с воздействием прочих не учитываемых в эксперименте факторов (в том числе случайных). Значимость средних квадратов, соответствующих этому разложению, можно проверить по критерию Фишера.

Если количество зарегистрированных видов велико (300-400 в нашем примере), то одномерный дисперсионный анализ не дает возможности оценить совокупную изменчивость всего таксономического комплекса изучаемого сообщества под воздействием группирующего фактора. Один из приемов избежать «проклятия размерности» – выполнить дисперсионный анализ симметричной матрицы **D** размерностью $m \times m$, элементами которой являются коэффициенты расстояния между каждой парой выполненных проб. Однако евклидово расстояние или отношение Ренконена–Брея–Кёртиса $d = 1 - \xi_R$, принятые в качестве модели отношений между точками данных, не вписываются в общую концепцию классического дисперсионного анализа, оценивающего вариацию случайных измерений относительно групповых средних (McCune et al., 2002). Кроме того, традиционное вычисление значений p-вероятности ошибки 1 рода основано на предположении о многомерном нормальном законе распределения, которое вряд ли справедливо для данных экологического мониторинга.

М. Андерсон (Anderson, 2001) предложил метод, известный как непараметрический многомерный дисперсионный анализ (npMANOVA), использующий матрицу расстояний **D**. При этом «средние точки» (centroid; Anderson, 2001), т. е. центры распределения значений d в пределах каждой группы, находятся не по обычным формулам, а с использованием методов Монте-Карло. Проверка нулевой гипотезы об отсутствии влияния группирующего фактора также осуществляется не на основе принятых предположений о характере распределения, а в ходе перестановочной процедуры: пробы многократно случайным образом перемешиваются относительно выделенных групп (Anderson, ter Braak, 2003).

Поскольку процедура npMANOVA ведет обработку сбалансированных данных с одинаковым числом измерений для уровней фактора, каждую из трех выделенных групп (р. Байтуган, верхнее и нижнее течение р. Сок) ограничим до 40 проб. С использованием прологарифмированных значений численности 363 видов макрозообентоса рассчитаем матрицу **D** расстояний Ренконена–Брея–Кёртиса между всеми 120 гидробиологическими пробами. Результаты непараметрического дисперсионного анализа, представленные в табл. 3.3, свидетельствуют о ярко выраженной статистической достоверности различий видового состава макрозообентоса на трех участках водотока.

Вероятность p оценивалась путем сравнения эмпирической величины F-критерия с набором из 5000 модельных F-значений, полученных в ходе многократного случайного перемешивания проб между участками. Поскольку все значения F, полученные рандомизационной процедурой, не превышали 5,9 и находились в диапазоне от 0,99 до 2,05, вероятность ошибочного отклонения верной нулевой гипотезы составила $p = 1/4999 \cong 0{,}0002$.

Таблица 3.3

Дисперсионный анализ отличий видового состава макрозообентоса по участкам рек Байтуган–Сок

Факторы	Степени свободы *df*	Сумма квадратов *SS*	Средние квадраты *MS*	*F*-критерий	*p*-значение
Группирующий фактор τ	2	4,3	2,15	5,9	0,0002
Случайные воздействия ε	117	42,6	0,364		
Итого	119	46,9			

Вывод о достоверности различий видового состава макрозообентоса на отдельных участках водотока подтверждается также вычислением значений *t*-критерия (числитель) и соответствующих им *p*-вероятностей (знаменатель) для множественных попарных post-hoc-сравнений (или контрастов) между отдельными уровнями группировочного фактора:

	р. Байтуган	р. Сок (верховья)	р. Сок (нижнее течение)
р. Байтуган	0 / 1	1,84 / 0,0002	3,05 / 0,0002
р. Сок (верховья)	1,84 / 0,0002	0 / 1	2,24 / 0,0002
р. Сок (нижнее течение)	3,05 / 0,0002	2,24 / 0,0002	0 / 1

Нами был представлен только один, самый обширный класс коэффициентов сходства, использующих общую формулу Ренконена–Съеренсена–Мориситы. В экологической практике встречаются различные версии других типов индексов, основанных на теории информации, семействе мер расстояния Минковского, корреляционных матрицах и т. д. (Ricklefs, Lau, 1980; Шитиков и др., 2005). Процесс уточнения старых и «изобретения» новых индексов не прекращается и по сей день (см., например: [Lennon et al., 2002; Green, Plotkin, 2007] и другие современные публикации, цитируемые в настоящем разделе). Однако такие проблемы, как учет характера распределения мониторинговых данных, уровень выборочной достаточности исследования и оценка статистической достоверности полученного решения, остаются неизменными во всех случаях. Что касается личных авторских предпочтений, то мы продолжаем обращать внимание на использование вероятностных мер отличий между комплексами измерений, относящимся к различным сообществам, таким как обобщенное расстояние Махаланобиса. Эти идеи были изложены в недостаточно, на наш взгляд, оцененных статьях В.Л. Андреева и Б.И. Сёмкина (1979) и подробно продемонстрированы нами (Шитиков и др., 2005) на конкретных примерах.

3.3. Иерархия уровней альфа- и бета-составляющих видового разнообразия

Структурированность оценок видового разнообразия тесно связана не только с филетическими отношениями, но и с пространственными масштабами изучаемых экосистем. Пусть мы имеем иерархическую схему с тремя масштабными уровнями, на которых производились экологические наблюдения в отношении определенной категории организмов. В пределах самого высокого уровня, объединяющего изучаемую территорию или совокупность водных объектов (назовем его «ландшафт»), можно выделить r_1 специфических местообитаний, соответствующих сообществам более низкого уровня. В каждом из местообитаний в ходе мониторинга было произведено по r_2 выборок (будем считать число выборок везде одинаковым). Все выборки характеризуются набором относительных плотностей популяций $\{p_i\}$, $i = 1, 2, \ldots, s$, где s – размерность регионального фонда видов.

Согласно предложенной Р. Уиттекером (Whittaker, 1972; Уиттекер, 1980) концепции *инвентаризационного* биоразнообразия, каждому масштабному уровню соответствует свой уровень разнообразия. Самый низкий уровень соответствует *точечному разнообразию* микроместообитания в данный момент времени, т. е. оценивается по результатам одной выборки. Разнообразие по совокупности проб внутри однородного местообитания было названо *разнообразием сообщества* или α-*разнообразием*, а общее разнообразие ландшафта – γ-*разнообразием*. Уровням инвентаризационного разнообразия соответствуют уровни *дифференцирующего* разнообразия, оценивающие «дисперсию» видового состава между элементами каждого уровня, т. е. *мозаичное разнообразие* между выборками из одного местообитания и сравнительное β-*разнообразие* между отдельными сообществами.

Компоненты инвентаризационного разнообразия (α и γ) обычно измеряются путем прямого подсчета числа видов или с использованием традиционных индексов видового разнообразия (Симпсона, Шеннона и т. д.). Для оценки β-разнообразия Уиттекером и его последователями на основе изучения различных градиентов в растительных сообществах было предложено несколько метрик, учитывающих так называемые «полусмены», либо появление и исчезновение видов вдоль градиента (см., например, Мэгарран, 1993). Однако, поскольку β-разнообразие характеризует различие видового состава между участками α-разнообразия, нет причин, ограничивающих применение этой меры только трансектами или градиентами факторов среды. Альтернативный подход к измерению β-разнообразия – изучение степени ассоциированности участков или выборок при помощи стандартных методов экологической ординации и классификации.

Косвенно β-разнообразие может быть оценено с использованием индексов сходства (см. раздел 3.2) и полученных на их основе иерархических *дендрограмм*, *плеяд Терентьева*, *графоаналитических отображений*, *результатов многомерного шкалирования* и т. д. (Шитиков и др., 2005). В то же время такие производные показатели, как средний индекс сходства или разброс относительно «характерной комбинации видов» (Pignatti, Mengarda, 1962), не вполне корректно отражают средний уровень гомогенности территории, поскольку игнорируют общие черты экосистемы более высокого порядка.

Попарное сравнение смежных местообитаний является удовлетворительным, если ставится задача изучить, как видовой состав изменяется вдоль физического или экологического градиента, однако если рассматривать отдельные фрагменты экосистемы как случайную выборку из большей области, то необходим переход к мерам дифференциации для множества участков. Основываясь на работе (Koleff et al., 2003), показывающей связь между индексом Съеренсена для двух участков и видовым разнообразием по Уиттекеру β_W, О. Дисеруд и Ф. Одегард (Diserud, Ødegaard, 2007) без особенных сомнений от формулы (3.8) приходят к простому выражению: $L_m = \dfrac{m - \beta_W}{m - 1}$.

Имеют ли, однако, приведенные обобщения мер сходства прямое отношение к β-разнообразию? Описанные подходы часто оценивают β-разнообразие, не учитывая его скрытую зависимость от α-разнообразия (и наоборот). Чтобы избежать неверных выводов при сравнении регионов, которые отличаются по α-разнообразию, необходимо точно оценить статус обоих компонентов и разработать алгоритмы их независимой оценки.

Уиттекер на основе изучения различных градиентов в растительных сообществах предложил мультипликативную модель видового разнообразия региона γ в виде произведения двух показателей α- и β-разнообразия: $\gamma = \alpha \cdot \beta$. Недостатки такой схемы очевидны:

- α- и β-разнообразие имеют разные размерности и не могут использоваться для сравнения (Wagner et al., 2000);
- для трех и более масштабных уровней эти составляющие не будут обладать одинаковым весом (Gering et al., 2003).

В связи с этим в качестве альтернативы мультипликативной модели всё большее распространение получает аддитивная схема масштабного разделения разнообразия, определяющая γ-разнообразие как сумму компонентов α- и β-разнообразия: $\gamma = \alpha + \beta$ (Lande, 1996). Исторически эта схема обосновывалась как развитие *теории экологической ниши* (MacArthur, Wilson, 1963) и статистической *теории декомпозиции видового разнообразия* (Rao, 1982) – см. подробный обзор в (Veech et al., 2002).

Удобство использования аддитивной схемы Ланде состоит в том, что она допускает разложение общего разнообразия более чем на два компонента. Например, для трех масштабных уровней «проба (выборка)» – «местообитание» – «регион» имеем:

$$\gamma = \alpha_2 + \beta_2 = \alpha_1 + \beta_1 + \beta_2,$$

т. е. получаем разложение общего разнообразия (γ) в виде суммы среднего разнообразия, содержащегося в выборках (α_1), разнообразия, привносимого различиями между выборками в пределах «местообитания» (β_1) и различиями между «местообитаниями» в пределах «ландшафта» (β_2). Принципиально такая схема может иметь произвольное число уровней больше двух.

Абстрагируясь от конкретного способа расчета γ, Л. Джост (Jost, 2007) сформулировал общее соглашение о свойствах α- и β-разнообразия в виде совокупности следующих аксиом:

1. α- и β-разнообразия представляют собой две ортогональные компоненты, оценивающие полностью различные аспекты регионального разнообразия. Поэтому они должны свободно и независимо варьировать; в частности, высокое значение α-компоненты не должно вынуждать быть высоким (или низким) β-компоненту, и наоборот. Только математическая независимость β- от α-компоненты дает возможность сравнивать β-разнообразие регионов, α-разнообразие которых отличается.
2. Каждый компонент α- и β-разнообразия, вне зависимости от его природы, должен вносить строго соразмерное количество разнообразия (или неопределенности) в общий показатель γ-разнообразия, позволяющее обоснованно выделить составляющие «в пределах местообитания» и «между местообитаниями».
3. α-компонента – некоторым произвольным образом найденный центр статистического распределения показателей разнообразия «местообитаний» или «выборок», которые образуют «регион»; чтобы не подходить к этой проблеме предвзято, целесообразно сделать только одно минимальное предположение: если показатель разнообразия для всех «местообитаний» в регионе принимает одно и то же значение H_0, то α-компонента также должна равняться H_0.
4. γ-компонента должна быть полностью определена α- и β-компонентой; нет необходимости в конкретных математических предположениях о том, как α- и β- определяют γ-компоненту.
5. Компоненты должны обладать свойством строгой вогнутости (strict concavity; Lande, 1996), т. е. α-компонента никогда не может быть больше чем γ-компонента. Это свойство означает, что при объединении нескольких выборок разнообразие в любой суммарной выборке всегда должно быть больше (либо равно) среднему разнообразию по всем выборкам. С точки зрения теории информации эта особенность разумна. Если вспомнить, что большинство показателей разнообразия можно считать обобщенными мерами *неопределенности* (Taneja, 1989), то α-компоненту можно считать неопределенностью при оценке видовой принадлежности особей, случайно извлекаемых из «местообитания», при условии, что нам известно хотя бы место, откуда выбираются особи. γ-компонента – неопределенность в идентификации видовой принадлежности, когда мы точно не знаем, в каком месте осуществлялась выборка. Знание никогда не может увеличивать неопределенность, таким образом, α- компонента никогда не может быть больше, чем γ-компонента.

Чтобы разработать новую картину α- и β-разнообразия с учетом представленных (относительно бесспорных) свойств 1–5, необходимо создать инструментарий оценки показателей разнообразия более общим способом, чем это являлось общепринятым. Тем самым увеличивается наше понимание того, как видовая неоднородность сообществ изменяется при увеличении пространственных масштабов.

Пусть в результате экологического мониторинга была получена совокупность гидробиологических проб (в иных исследованиях – оценки видового обилия по квадратам, трансектам, загородкам и т. д.). Каждая элементарная выборка характеризуется индивидуальным α-разнообразием и, следовательно, по всей имеющейся совокупности проб может

быть рассчитано среднее разнообразие α_1, которое отражает самый детальный масштабный уровень – параметр в точке (sample grain). Подмножество проб, полученных из одного местообитания (например, на локальном участке реки), может быть объединено, чтобы сформировать выборку второго масштабного уровня и найти соответствующее ему значение α-разнообразия. Если было изучено несколько местообитаний, то можно найти среднее разнообразие α_2 «внутри» местообитаний, представляющих оценку параметра на небольшом ограниченном участке. Продолжая иерархическое объединение проб и укрупняя пространственные масштабы, можно рассчитать среднее α-разнообразие на каждом из последующих уровней, например, α_3 – для отдельных рек или крупных характерных их участков, α_4 – для отдельных бассейнов или регионов, α_5 – для природно-климатических зон или континентов и т. д. Таким образом, полное видовое разнообразие γ, полученное на всей располагаемой совокупности выборок, таково, что $\gamma \geq \alpha_r$ (Lande, 1996), причем в ряду $\alpha_1, \alpha_2, ..., \alpha_r$ значение α_1 определяет точечное разнообразие, а последующие члены зависят от способа пространственного объединения выборок.

В общем случае для r масштабных уровней $\gamma = \alpha_1 + \sum_i^r \beta_i$, поскольку на каждом i-ом уровне $\beta_i = \alpha_{i+1} - \alpha_I$ (Lande, 1996). Значения альф вычисляются по формуле $\alpha_i = \sum_j^{m_i} D_{ij} q_{ij}$, где D_{ij} – принятая метрика разнообразия, рассчитанная для каждой выборки, $j = 1, 2, ..., m_i$; m_i – число выборок, взятых на i-ом уровне иерархической схемы; q_{ij} – выборочные веса (Wagner et al., 2000), определяемые соотношением общего числа особей, найденных в каждой выборке j.

В качестве примера рассмотрим разложение общего видового разнообразия макрозообентоса речной системы Байтуган–Сок (см. рис. 1.7) на α- и β-составляющие при различных уровнях выборочного агрегирования. В качестве тестируемых показателей разнообразия используем индексы Шеннона H и Симпсона D, а также видовое богатство s (численные значения представлены в табл. 3.4).

Таблица 3.4

Разложение видового разнообразия сообществ макрозообентоса речной системы Байтуган–Сок на α- и β-составляющие

Уровень	Описание составляющих	Компоненты α-β	Число видов s	Индекс Шеннона H	Индекс Симпсона D
0	Общее разнообразие	γ	325	3,77	0,943
1	Среднее α-разнообразие 123 проб и β-разнообразие между пробами	α_1	13	1,62	0,671
		$\beta_1 + \beta_2 + \beta_3$	312	2,15	0,273
2	Среднее α-разнообразие 22 станций наблюдения и β-разнообразие между станциями	$\alpha_2 = \alpha_1 + \beta_1$	44,6	2,34	0,800
		$\beta_2 + \beta_3$	280,4	1,43	0,143
3	Среднее α-разнообразие трех участков водотока (р. Байтуган, верхнее и нижнее течение р. Сок) и β-разнообразие между ними	$\alpha_3 = \alpha_1 + \beta_1 + \beta_2$	168,0	3,33	0,925
		β_3	157,0	0,44	0,018

Заслуживает упоминания различие между долевым участием рассчитанных компонент видового α- и β-разнообразия по моделям, основанным на выборках (столбец «Число видов» в табл. 3.4) и на особях (индексы *H* и *D*). Если средняя видовая насыщенность одного участка составляет около половины полного богатства (168 из 325), то видовое разнообразие всей гидроэкосистемы несущественно отличается от его показателей на отдельных участках (на 12% и 2% *H* и *D* соответственно). Отмеченное обстоятельство объясняется достаточно тонкими методическими причинами. С одной стороны, эти различия можно объяснить природой самих индексов: соотношение характера их внутренней нелинейности и линейной природы аддитивной модели подробно обсуждается в работе Л. Джоста (Jost, 2006). С другой стороны, декомпозиция разнообразия, основанная на анализе особей, исходит из общей численности сообщества и расчетное значение α_{exp} учитывает среднее обилие экземпляров в одной пробе n_{ij}. Поскольку величина n_{ij} в существенной мере определяется подмножеством выборок с аномально высокой численностью отдельных видов, то это приводит, как правило, к более высокому значению α_{exp}, чем это можно получить по кривым разрежения, основанным на накоплении проб.

С использованием аддитивной иерархической схемы масштабного разделения можно рассчитать совокупность значений α- или β-разнообразия на каждом уровне, однако возникает вопрос: насколько их величина определяется экологическими причинами? Д. Геринг с соавторами (Gering et al., 2003) предложили статистический критерий оценки значимости отличий эмпирических значений α_{obs} и β_{obs} от ожидаемых величин при случайном характере распределения особей по выборкам. Ими была разработана имитационная процедура рандомизации в виде *перестановочного теста* (permutation), заключающаяся в следующем. Все выборки нижнего уровня объединяются в одну, в результате чего получается общее распределение численности видов. Затем из этого «общего фонда» случайным образом извлекаются особи (без возвращения) и распределяются по выборкам, причем исходный размер выборок по числу организмов сохраняется. Сгенерированные псевдовыборки используются для процедуры аддитивного разложения разнообразия и рассчитываются псевдозначения α_{sim} и β_{sim}. Процедура рандомизации повторяется достаточно большое число раз (обычно не менее 100) и формируется статистическое распределение оценок α- и β-разнообразия, которое и используется для поверки статистических гипотез. Например, в качестве гипотезы может быть выбрано следующее утверждение: $H_0 : \beta = \beta_{sim}$, $H_1 : \beta > \beta_{sim}$. Тогда статистическая значимость оценивается как доля значений из случайного распределения β_{sim}, которые превосходят реальное значение β_{obs}.

Для всех компонентов видового α- и β-разнообразия, представленных в табл. 3.4, были рассчитаны *p*-значения, основанные на рандомизационной процедуре перемешивания особей в ходе 100 итераций. Все эти значения оказались очень низкими ($p < 0.01$) и, следовательно, предположение, что данные мониторинга рек Байтуган–Сок являются результатом работы «нуль-модели полного хаоса», очень маловероятно.

К. Рикотта (Ricotta, 2003) предложил обобщить аддитивную схему масштабного разделения разнообразия и перейти от отдельных индексов к некоторому их параметрическому семейству. Наиболее распространенные семейства – обобщенные *энтропии Реньи* и «*числа Хилла*» – не подходят для этой цели, поскольку удовлетворяют свойству строгой вогнутости лишь на интервале параметра $0 \le \alpha \le 1$. Удобнее оказалось воспользоваться так называемой *обобщенной энтропией типа* α[4], определяемой как

$$_{\alpha}H = \sum_{i=1}^{S} \frac{1}{2^{1-\alpha}-1}\left(p_i^{\alpha} - p_i\right),$$

[4] Здесь возникает определенное неудобство с терминологией, поскольку одним и тем же символом α обозначаются параметр семейства и уровень разнообразия.

где $0 \le \alpha \le \infty$. Эти показатели также можно рассматривать как семейство индексов разнообразия. Энтропии типа α обладают свойством строгой вогнутости для любых положительных α, чем можно воспользоваться для аддитивного разделения разнообразия. При этом для каждого значения параметра α можно рассчитать значения α- и γ-разнообразия, а значение β-разнообразия вычисляется как их разность. В итоге полученное аддитивное разделение разнообразия можно представить в виде профилей ${}_{\alpha}H(\alpha)$.

Особый смысл имеет масштабная декомпозиция такого показателя, как число видов *S*, обнаруженных в экосистеме. Разделение полного видового богатства на α- и β-компоненты может быть использовано, чтобы проанализировать, как меняется β-разнообразие в зависимости от выборочного усилия. Пусть s_{ij} – число видов, зарегистрированных в каждой выборке $j = 1, 2, ..., m_i$, где m_i – число выборок на i-ом уровне иерархического плана объединения проб. Тогда, согласно (Lande, 1996), выражение для среднего числа видов (α_i), на каждом уровне i будет равно $\alpha_i = 1/m_i \sum_j^{m_i} s_{ij}$. В этой формуле для каждой выборки веса $q_{ij} = 1$. При аддитивном распределении по иерархической выборочной схеме значения β-разнообразия могут быть рассчитаны как $\beta_i = 1/m_i \sum_j^{m_i} (\gamma - s_{ij})$, где β_i имеет смысл числа видов, которое необходимо добавить к среднему видовому α-разнообразию на i-ом уровне, чтобы получить полное видовое богатство.

Аналитическую зависимость, обеспечивающую аддитивное разложение видового богатства на α- и β-компоненты, можно также получить из формулы, представленной в работе (Ugland et al., 2003): $E[S(a)] = S_{obs} - \sum_{k=1}^{S_{obs}} q_k(a)$, где $E[S(a)]$ – оценка числа видов при объединении любого подмножества из a выборок, S_{obs} – полное число видов во всех выборках и $q(a)$ – вероятность того, что вид k не будет найден в a выборках. Очевидно (Crist, Veech, 2006), что при $a = 1$ справедливы соотношения: S_{obs} = γ-разнообразию, $E[S(a)] = \alpha_1$, $\beta_1 = \sum_{k=1}^{S_{obs}} q_k(1)$. Другая версия математического обоснования выполнена в работе (Arita, Rodriguez, 2004), предлагающей последовательность вытекающих друг из друга формул: $\gamma = \alpha_i m_i / \bar{u}_i$, $\sum_i^m \beta_i = (1 - \bar{u}_i / m_i)\gamma$ и $\beta_i = (\bar{u}_{i+1}/m_{i+1} - \bar{u}_i / m_i)\gamma$, где $\bar{u}_i$ – среднее число выборок (проб), которые могли бы включить всё видовое богатство γ.

В ряде публикаций (Ugland et al., 2003; Olszewski, 2004; Crist, Veech, 2006) предпринимается попытка установить связь аддитивной модели разделения разнообразия с известными методами анализа и интерпретации видового богатства. Например, была показана тесная зависимость между компонентами иерархического α-β разделения и кривыми разрежения, подробно описанными выше в разделе 2.4.

Кривые разрежения строятся по мониторинговым данным, полученным в пределах естественных сообществ, чтобы оценить степень выборочной достаточности для достоверного определения истинного видового богатства или сравнить число видов между сообществами на основе равного выборочного усилия (Gotelli, Colwell, 2001). Используя различную группировку проб в соответствии с иерархическим планом объединения масштабных уровней, можно построить (Crist, Veech, 2006) семейство кривых разрежения, которые имеют разный масштаб по оси абсцисс, но все они асимптотически приближаются к одному и тому же полному богатству видов.

Графический способ разложения видового богатства макрозообентоса речной системы Байтуган–Сок по семействам кривых разрежения приводит к тем же соотношениям между компонентами разнообразия, что и в табл. 3.4. На рис. 3.4 представлены все эти составляющие, полученные на основе трех кривых для отдельных гидробиологических проб, станций наблюдения и отдельных участков рек соответственно:

- среднее видовое богатство в одной пробе соответствует α_1-разнообразию и является первой точкой на соответствующей кривой накопления (α_1 = 13);

- различие в числе видов между последней и первой точками на кривой разрежения для проб есть полное β-разнообразие между пробами;
- полное число видов в объединенном множестве проб соответствует γ-разнообразию;
- аналогично этому, богатство видов, соответствующее первым точками каждой последующей кривой накопления (для станций и участков водотоков) есть ни что иное, как α_i-разнообразие, а разница между ординатами первых точек любой пары смежных кривых – β_i-разнообразие между объектами *i*-го уровня.

Поясним, что абсцисса первой точки на второй кривой накопления равна 5,6; т. е. среднему числу проб на станции (123/22), а «масштабная единица» на третей кривой равна 41, т. е. среднему числу проб на участке водотока (123/3).

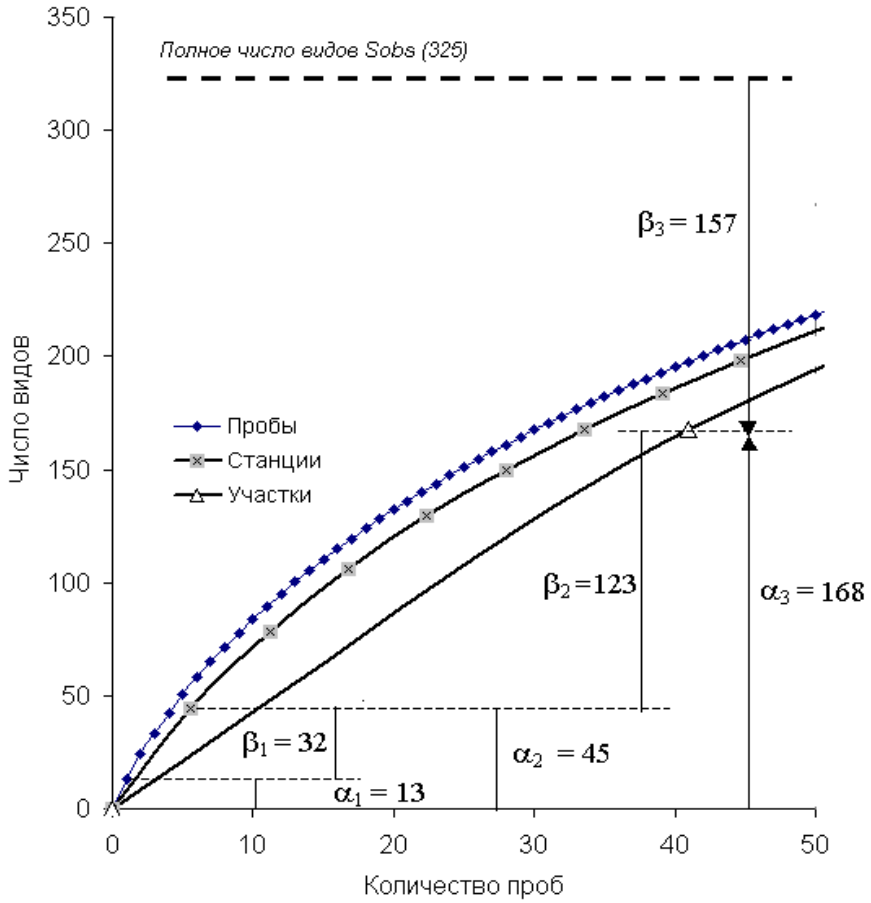


Рис. 3.4. **Декомпозиция видового разнообразия макрозообентоса рек Байтуган–Сок на α- и β-составляющие по трем уровням: 1 – гидробиологическая проба, 2 – станция наблюдения, 3 – участок водотока**

Разложение общего видового разнообразия по масштабным уровням позволяет подойти к проблеме оценки экологического статуса сообществ с учетом факторов биотопической и региональной неоднородности плотности популяций. Например, полученное соотношение α-компонентов видового богатства:

13 (в одной пробе) : **45** (на станции) : **168** (на участке) : **325** (для всего водотока)

свидетельствует как о возможном существовании продольных градиентов и локальных рефугиумов по течению изучаемых рек, так и о специфической особенности, характерной для всех макрозообентоценозов и связанной с большим числом редких видов. В частности, для рассматриваемых сообществ средней статистической «нормой» является вероятность обнаружения вида только в одной гидробиологической пробе из 12.

ГЛАВА 4. АНАЛИЗ СТАТИСТИЧЕСКИХ ЗАКОНОМЕРНОСТЕЙ ОРГАНИЗАЦИИ ВИДОВОЙ СТРУКТУРЫ

Охота на слона:
Математики отвергают все, что не-Слон
и ловят одного из того, что осталось…
Экологи ловят первое попавшееся серое
животное и называют его Слоном…
"Пятничный manager" (Сборник юмора)

4.1. Взаимодействия видов и их влияние на таксономическую структуру сообществ

Теоретическая экология выделяет девять различных типов взаимодействия двух видов в природных сообществах (Одум, 1975, с. 274):

- протокооперация и мутуализм (+ +),
- хищничество и паразитизм (+ –),
- прямая и ресурсная конкуренции (– –),
- аменсализм (– **0**), комменсализм (+ **0**) и нейтрализм (**0 0**).

Здесь в скобках использованы обозначения: (+) – благоприятное воздействие на рост, выживание и другие характеристики популяций; (–) – ингибирующее действие на эти показатели; (**0**) – отсутствие воздействия одной популяции на другую. Механизмы вертикальных взаимодействий традиционно описываются как пищевые цепи в трофических сетях. Для видов одного или смежных трофических уровней характерны горизонтальные взаимодействия, обусловленные конкуренцией, различными локальными и региональными экологическими причинами или эволюционной историей (Bascompte, Jordano, 2007).

В процессе эволюции экосистем общая (правильнее сказать – статистическая) картина взаимодействия видов изменяется: в недавно сформировавшихся ассоциациях велика вероятность сильных отрицательных взаимодействий, однако в ходе развития наблюдается тенденция к росту положительных связей, способствующих выживанию видов. Все эти механизмы находят свое отражение в динамике метаструктурной конфигурации сообществ, т. е. в пространственном или временном изменении видового состава за счет появления новых видов и исчезновения старых. В общем случае можно выделить (Leibold, Mikkelson, 2002) три важнейших структурных закономерности, связанные с взаимодействием видов: *взаимная согласованность* (coherence), *заменяемость* или *взаимное исключение* (turnover) и *формирование обособленных групп* (boundary clumping). Показано (Krishna et al., 2008), что многие описанные экологические свойства трофических сетей и типичные структурные артефакты так или иначе являются неизбежными следствиями этих трех процессов. Наиболее характерные типы структурной организации представлены на рис. 4.1.

Количественная оценка степени взаимодействия видов имеет давнюю традицию в экологии (Weiher, Keddy, 1999). Особенно длительная и все еще продолжающееся дискуссия ведется относительно правил ассоциирования видов в сообществе и *концепции* «*минимального сообщества*» (Diamond, 1975; Connor, Simberloff, 1979; Джиллер, 1988) которая инициировала развитие статистических инструментов выделения фрагментов видовой структуры, которые не могут быть интерпретированы как случайные.



Рис. 4.1. **Наиболее характерные типы структурной организации сообществ (на примере биогеографических матриц)**

Исходными данными для анализа межвидовых связей служат два типа матриц: биогеографические матрицы и экологические сети. Традиционная бинарная матрицы биогеографических описаний, в *s* строках которой представлены виды, а в *m* столбцах – местообитания (или пробы, взятые на одном участке в разное время), является непосредственной формой представления мониторинговых данных. Элементы такой матрицы указывают на присутствие (1) или отсутствие (0) вида на том или ином конкретном участке. Взаимодействия видов в этой форме анализа оцениваются не напрямую, а опосредованно, т. е. как некоторая функция частот их одновременной встречаемости в группах сходных местообитаний (очевидна аналогия с непрямым градиентным анализом).

Другим способом представления исходных данных являются двусторонние экологические сети (Sugihara, Ye, 2009), описывающие два множества потенциально взаимодействующих между собой видов (например, хищники и их жертвы). Эта версия анализа особенно распространена применительно к мутуалистическим сетям, описывающим, например, симбиотические комплексы растений и насекомых-опылителей. В матрицах $m \times n$, описывающих такие структуры, виды сравниваемых подсообществ представлены m строками и n столбцами, а значения 0 или 1 соответствуют градациям вероятности возможного совместного обнаружения каждой пары видов. Двусторонние экологические сети непосредственно отображают межвидовые связи в экосистеме, однако нельзя не считаться с риском субъективного разделения видов на две группы (например, один и тот же таксон в разных обстоятельствах может классифицироваться и как хищник, и как жертва).

Несмотря на принципиальные отличия биогеографических матриц и экологических сетей, у них есть очень много общего как в методах выделения неслучайных структурных фрагментов, так и в последующей статистической обработке. Анализ внутренней взаимообусловленности видовой структуры предполагает три этапа:

- вычисление метрики, оценивающей степень взаимосвязи видов в исходной матрице;

- сравнение параметров изучаемой экосистемы с соответствующей нуль-моделью и оценка статистической значимости метрики с помощью рандомизационной процедуры;
- формулировка гипотез об экологических механизмах, объясняющих полученные результаты.

По всем трем позициям отсутствует единая методика формализации и до сих пор не прекращается дискуссия экологов относительно причин и следствий различных структурных закономерностей, а также многочисленных нюансов количественного измерения их выраженности (Almeida-Neto et al., 2008; Timi, Poulin, 2008).

В большинстве случаев механизмы, объясняющие взаимодействие видов, нельзя выделить, основываясь только на исходной матрице отсутствия-присутствия видов и статистических индексах, полученных на её основе. Уточнение определяющих экологических процессов обычно будет требовать большой совокупности дополнительных данных, например, сведений о пространственном распределении основных абиотических факторов. Предпочтительно, чтобы изучение особенностей видовой структуры проводилось в комбинации с различными вариантами градиентного анализа, кластеризацией и другими многомерными методами (Leibold, Mikkelson, 2002). Это позволяет осуществить идентификацию самого сильного градиента, который определяет механизмы межвидовых взаимодействий и закономерности структурной организации сообществ.

Стандартный анализ параметров видовой структуры предполагает, что бинарные матрицы отсутствия/присутствия не содержат ошибок. Однако результаты и выводы на их основе в значительной мере зависят не только от характера и интенсивности экологических процессов, но и от методик мониторинга. В условиях мета-сообществ с неоднородным и фрагментарным распределением популяционных плотностей следует принимать во внимание эффект «пассивного планирования мест для взятия проб» (passive sampling).

Большое влияние на полученные результаты может оказать и «выборочный эффект», вследствие которого часть видов в местообитании не обнаруживается из-за недостаточного количества взятых проб. Например, низка вероятность обнаружить редкие виды-специалисты на небольших, относительно изолированных участках территории или водотока, тогда как для видов с высокой численностью особей есть больше возможностей рассеивания и колонизации. Добавление к матрицам инцидентности редких видов, которые наиболее чувствительны к выборочному усилию, может сильно изменить оценку вероятности межвидового взаимодействия (Greve, Chown, 2006). В то же время другая часть исследователей (Nielsen, Bascompte, 2007) считает, что зависимость результатов анализа от объема выборки не столь велика по сравнению с богатством видов и заполнением матриц.

У проблемы выборочной достаточности есть также теоретическая перспектива. С одной стороны, редкие и эндемичные виды усиливают внутреннюю обусловленность сообществ, поскольку их появление соответствует моделям избирательной колонизации и исчезновения (Patterson, Atmar, 2000). С другой стороны, этот процесс может позиционироваться как случайный, вследствие чего снижается оценка интенсивности инкорпорации видов. Поэтому проблема выборочной неполноты и редкости видов заслуживает большего внимания и увеличения количества исследований.

4.2. Меры взаимосвязи видов с использованием структурных элементов

За последние два десятилетия было предложено много метрик[1], количественно оценивающих интенсивность взаимодействия видов и характер структурной организации сообществ (Berlow et al., 2004). Однако, несмотря на широкое практическое использование,

[1] По этому поводу есть образная фраза Н. Готелли [Nicholas J. Gotelli]: *However, this freedom also gives you enough rope to effectively hang yourself* («Однако свобода выбора [мер и процедур] предоставляет Вам прекрасную веревку, чтобы эффектно повеситься»).

репрезентативные условия применения этих мер и процедур все еще слабо определены, а статистические свойства стали серьезно изучаться лишь в последнее время (Gotelli, 2000; Ulrich et al., 2009).

Взаимная согласованность видов (coherence) может быть оценена как площадь прямоугольных фрагментов бинарной матрицы, полностью заполненных единицами. Простая мера согласованности – $C = \frac{1}{sm}\sum \max(1\cdots 1)$, т. е. суммированное по всем s строкам и m столбцам максимальное количество смежных инцидентностей, нормированное на размерность матрицы $s{\cdot}m$. Однако для расчета меры C нужно предварительно преобразовать исходную матрицу таким образом, чтобы в ней появилось максимальное число прилегающих «единиц», что является само по себе нетривиальной задачей (Leibold, Mikkelson, 2002). Еще один из индексов этой серии – *COMBO* (D. Pielou, E. Pielou, 1968) – подсчитывает число уникальных комбинаций (combinations) совместной встречаемости видов и изменяется от 1 до 2^s. Чем больше таких комбинаций, тем значительнее композиция видов отличается от нейтральной модели сообщества без каких-либо взаимодействий.

В ряде работ для оценки зависимостей между всеми возможными парами видов использовался корреляционный анализ, однако ввиду того, что статистические испытания не являются независимыми, здесь бывает трудно оценить вклад в полученный результат опосредованных и других множественных взаимодействий. Анализ комбинаторных эффектов с использованием таблиц статистической сопряженности 2^s (Pielou, 1972) в какой-то мере решает эту проблему для данных присутствия/отсутствия видов, но этот метод быстро становится непрактичным при увеличении числа видов больше 5.

Взаимное исключение видов (turnover) соответствует закономерности, с которой два вида заменяют друг друга в пределах столбцов или строк, что определяется существованием подматриц: $\left[\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array}\right]$ и $\left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right]$. Например, можно предположить, что пара видов конкурирует друг с другом и никогда не встречается на одном и том же изолированном участке. Дж. Даймонд (Diamond, 1975) назвал это «распределением шахматной доски» и утверждал, что присутствие многих таких пар в сообществе является доказательством детерминированных правил его организации. Отметим, однако, что если сообщество представляет собой полную шахматную доску, то один из видов обязательно формирует с третьим видом идеально сокоррелированную пару (рис. 4.1в). Следовательно, парную разобщенность видов всегда можно связать с некоторой мерой их согласованности с другими видами.

Степень взаимного исключения в сообществе из m видов традиционно оценивается с использованием нескольких метрик (Gotelli, 2000):

- индекс «шахматной доски» *CHECKER* (Diamond, 1975), который просто подсчитывает число подматриц $\left[\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array}\right]$ и $\left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right]$ размерностью 2×2 и теоретически изменяется от 0 до $s(s-1)/2$;
- индекс C-заполнения (Checkerboard score – Stone, Roberts, 1990), оценивающий среднее число таких фрагментов для произвольной пары видов $CS = \sum_{ij}(N_i - N_{ij})(N_j - N_{ij})$, где N_{i} и N_{j} – встречаемости видов i и j, N_{ij} – одновременная встречаемость этой пары. Этот индекс изменяется в диапазоне от 0 до $\frac{2}{s(s-1)}\sum_{ij} N_i N_j)$;

- мера близости (togetherness score – Stone, Roberts, 1992) $t = \frac{2}{s(s-1)} \sum_{ij}^{s(s-1)/2} \frac{4 p_{ij} a_{ij}}{m^2}$, где p_{ij} и a_{ij} – число комбинаций пар присутствия и отсутствия видов соответственно;
- метрика Съеренсена $C_{soe} = \frac{2}{s(s-1)} \sum_{ij}^{s(s-1)/2} \frac{2n_{ij}}{n_i + n_j}$.

С увеличением большинства этих индексов также растет степень организации сообщества.

Важным структурным элементом матриц являются ассоциации (группы) видов, имеющие более или менее отчетливые границы (boundary clumping). Их существование объясняется двумя фундаментальными представлениями об экологических сообществах. Традиционный фитоценологический подход восходит к работам Ф. Клементса (Clements, 1916) и связан с выделением обособленных прямоугольных блоков матрицы, объединяющих небольшие группы близких по экологии видов, встречающихся в сходных местообитаниях. Индивидуалистическая концепция Л.Г. Раменского (1924) и Г. Глизона (Gleason, 1926) также не отрицает процессов ассоциирования видов, если они распределены вдоль какого-либо экологического градиента (Leibold, Mikkelson, 2002). В пределах трофических сетей это представление связывается с подсетями немногих взаимодействующих видов, которые имеют очень мало связей с другими аналогичными подсетями.

Степень обособленности групп (т. е. число и размеры относительно изолированных блоков матриц) традиционно оценивается с использованием *индекса инцидентности Мориситы* (Morisita, 1971), хотя алгоритмы, основанные на фрактальной размерности (Bloom, 1981), кажутся теоретически более обоснованными и удовлетворительными на практике. Отметим, что сами по себе понятия «агрегированности» и «разделяемости» (aggregation and comparatmentization) не вполне ясно определены, что приводит к существенным разногласиям о способах их измерения (Plotkin, Muller-Landau, 2002).

В литературе представлены и другие метрики, основанные на анализе группового распределения видов (Stone, Roberts, 1992), различных версиях оценки наложения ниш (niche overlap – Sohn, 2001) и др. Однако у этих методов не видно ощутимых преимуществ по сравнению с интуитивно ясными и простыми матричными мерами.

4.3. Инкорпорация видов и оценка «вложенности» (nestedness) структурных композиций

В пределах целостных континентальных сред (наземных ландшафтов, сетей водотоков и т. д.) таксономическая структура экологических сообществ формируется под влиянием двух конкурирующих процессов, определяющих, соответственно, степень её вложенности и изолированной фрагментарности (Patterson, Atmar, 2000). В полностью вложенных структурах древовидной формы видовой состав небольших композиций является строго детерминированной частью подмножеств более высокого уровня (рис. 4.1б). Степень приближения к таким структурам определяет меру порядка в экосистеме, обусловленную следующими факторами:

- общая биогеографическая история и соответствие текущим параметрам экологических условий для всех видов и их местообитаний;
- иерархические отношения ниши и детерминированный характер распределения ресурсов, определяющие для каждого местообитания различное число видов, принадлежащих, однако, к одному и тому же региональному пулу;
- закономерное изменение видового богатства и популяционной плотности в соответствии с градиентом параметров среды или пространственными координатами.

В то же время неоднородность природной среды определяет стохастичность видовой структуры (рис. 4.1а), а фрагментарность ландшафта (вызываемая частично и антропоген-

ными преобразованиями – дороги, плотины, цепочки поселений) приводит к появлению относительно изолированных субпопуляций, представляющих собой локальные ассоциации специфичных видов (рис. 4.1г).

Понятие «вложенности» (nestedness[2]) видовой структуры впервые было предложено Э. Хюлтеном (Hulten, 1937; цит. по: Hausdorf, Hennig, 2007), однако количественный nestedness-анализ стал возможным после серии работ (Patterson, Atmar, 1986; Patterson, 1987; Atmar, Patterson, 1993; Patterson, Atmar, 2000), где на большом экспериментальном материале были изучены процессы инкорпорирования[3] видов в ландшафтных фрагментах и предложены метрики, оценивающие степень вложенности матриц. Анализ вложенности видовой структуры сообществ стал после этого широко используемым экологическим инструментом для изучения основных механизмов возникновения и сосуществования популяций: поиск в библиографической системе Scopus дает более 150 статей, опубликованных с ключевым словом «*nestedness*». Этот интерес сопровождается быстро растущим числом прикладных и теоретических исследований по статистическим свойствам nestedness-метрик, формам нуль-моделей и методам формирования продуктивных гипотез (список литературных источников можно найти, например, в обзоре: Ulrich et al., 2009).

Nestedness-эффект был первоначально описан для замкнутых фаунистических сообществ (Darlington, 1957) и позиционировался как мера интенсивности исчезновения и колонизации видов. Представим себе островные сообщества на различных расстояниях от материка. Предположим, что острова, расположенные вдали от материка, будут иметь низкое видовое богатство, а те острова, которые расположены вблизи, будут иметь полный комплект материковых видов. В этом случае матрица «виды–острова» имеет высокую вложенность (т. е. низкую «nestedness-температуру» – см. далее). Если же на всех островах и на материке популяционная плотность имеющихся видов распределена приблизительно равномерно, то островные сообщества являются просто случайным фрагментом материковой фауны, и матрица «виды–острова» не будет показывать существенной организованности.

Разумеется, кроме фактора «островной изоляции», существует множество других экологических процессов (как детерминированных, так и стохастических), приводящих к различной степени инкорпорированности видовой структуры сообществ. В статье В. Ульриха с соавторами (Ulrich et al., 2009) приводится систематизированный обзор возможных объяснений эффекта вложенности субпопуляций. Для мета-сообществ эти предположения связаны, в основном, с различными вариантами процессов появления/исчезновения видов, интенсивность которых упорядочивается вдоль экологических или биологических градиентов. Таковыми могут быть ландшафтная изменчивость, различия в качестве окружающей среды или обеспеченности ресурсами, имеющиеся ограничения для миграции особей, распределение популяций видов-антагонистов и т. д.

Анализ вложенности видовой структуры начинается с ответов на два вопроса, которые определяют выбор nestedness-метрики и форму нуль-модели:

- Какой аспект вложенности представляет интерес для исследователя: взаимосвязи между видами, характер их распределения по местообитаниям или оба фактора вместе?
- Какой критерий следует использовать, чтобы оценить степень упорядоченности строк и/или столбцов исходной матрицы?

---

[2] Англоязычный термин «*nestedness*» не имеет однозначного перевода. Например, в немецкой литературе был принят эквивалент «*Verschachtelung*» (инкорпорация), а в польской – «*zagnieźdzenie*» (загнездование).

[3] *Инкорпорация* (лат. *incorporatio* – включение в состав) – объединение в одно морфологическое целое двух и более элементов, представляющих собой подвижные компоненты с обособленными свойствами (материал из Википедии).

Нами будут рассматриваться различные nestedness-метрики и алгоритмы рандомизации на примере матриц биогеографических описаний, которые являются более распространенной формой представления. Однако математическая сущность и основные принципы обработки данных вполне применимы и для анализа двусторонних сетей (основная разница имеется лишь на этапе окончательной интерпретации результатов).

Как правило, матрицы сортируются по суммам в строках и столбцах, после чего в верхних строках оказываются виды с наибольшей встречаемостью, а в столбцах слева – местообитания с наибольшим видовым богатством. С формальной точки зрения величина nestedness-эффекта оценивается двумя главными свойствами отсортированной таким образом матрицы: 1) насколько отличается максимальное и минимальное заполнение в крайних столбцах и строках и 2) совпадает ли встречаемость (т. е. наличие единиц) в менее заполненных столбцах и строках по сравнению с более заполненными столбцами и строками соответственно. Визуально степень вложенности сообщества представляется как концентрация «единиц» в верхнем левом треугольнике отсортированной матрицы (рис. 4.2) и концентрация «нулей» в нижнем правом углу.

Классической единицей измерения структурной вложенности матриц является *nestedness-температура сообщества* T (Atmar, Patterson, 1993), которая теоретически оценивает меру упорядоченности, согласно которой в экосистеме произошло бы исчезновение видов (или по другой версии – колонизация экосистемы видами). В «холодных» системах этот порядок детерминирован, когда как в «теплых» системах последовательность исчезновения приобретает все более и более произвольный характер. Значения T варьируют от 0^o (для строго вложенных структур) до 100^o (для абсолютно стохастических сообществ).

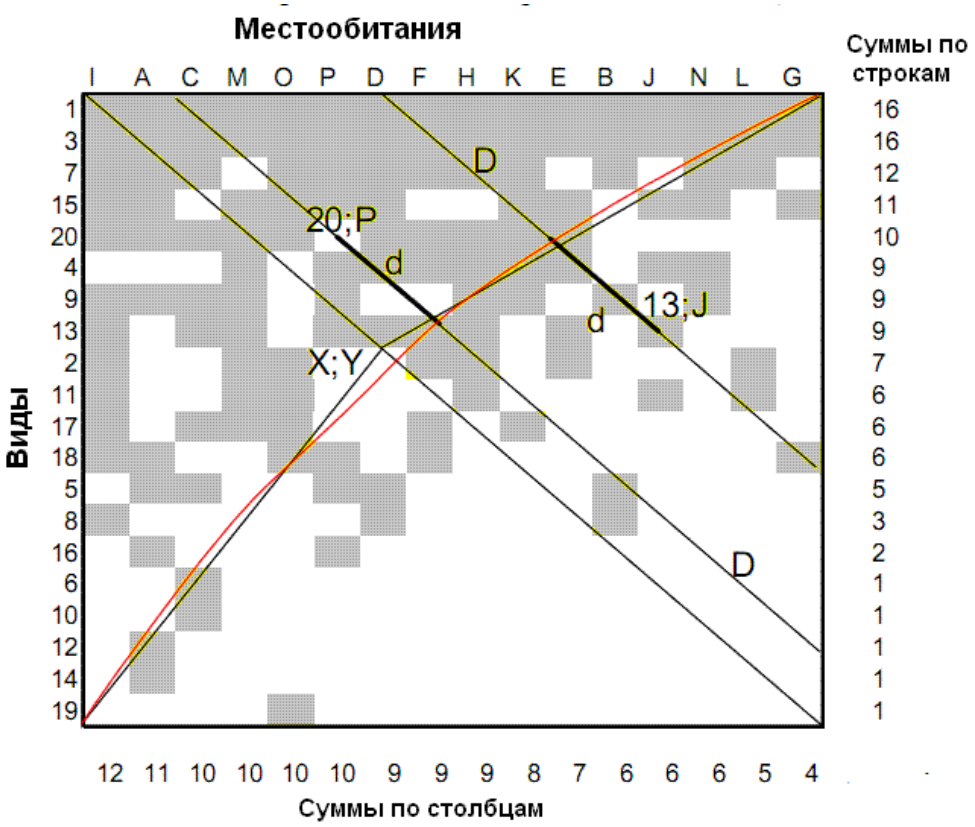


Рис. 4.2. **Пример отсортированной биогеографической матрицы данных**
(после ранжирования строк и столбцов идентификаторы видов
и местообитаний изменили свой натуральный порядок)

При вычислении матричной температуры T учитывается взаимное расположение пустых и заполненных ячеек и подсчитывается суммарное евклидово расстояние ошибочных элементов в двух частях исходной отсортированной матрицы, разделенных граничной кривой «истинной вложенности» (т. е. диагональной изоклиналью). Эта метрика в равной мере отражает как взаимосвязи между видами, так и характер их распределения по местообитаниям. Рассмотрим процедуру вычисления T на примере.

Точка [X, Y] на главной диагонали (рис. 4.2) выбирается пропорционально доле заполнения матрицы F (число занятых ячеек, деленное на общее количество ячеек), т. е. $X = F{\cdot}S$ и $Y = F{\cdot}M$, где S – число видов и M – количество местообитаний. От точки [X, Y] к углам матрицы проводятся прямые линии, которые интерпретируются как «границы вложенности». Далее строится гладкая кривая изоклина таким образом, чтобы суммы площадей между граничными линиями и изоклиналью выше и ниже последней были бы равны. Кривая изоклина делит всю матрицу на две части: преимущественно заполненную область истинной вложенности емкостью $F{\cdot}S{\cdot}M$ и нижнюю разреженную часть объемом $(1 - F{\cdot}S{\cdot}M)$. Появление пустой ячейки выше изоклинали («ошибочное отсутствие» – точка [20, P] на рис. 4.2), либо занятой ячейки ниже её («ошибочное присутствие» – точка [13, J]) являются отклонениями от строгой вложенности и проявлением «структурного беспорядка» в экосистеме.

Расстояние d_{ij} от каждой «ошибочной» ячейки [i; j] до разделяющей кривой отсчитывается по диагонали матрицы, проходящей через эту ячейку и имеющей полную длину D_{ij}. С учетом этого nestedness-температура сообщества T рассчитывается как

$$T = \frac{100}{0{,}04145 \cdot S \cdot M} \sum_{i=1}^{S} \sum_{j=1}^{M} \left(\frac{d_{ij}}{D_{ij}} \right)^2 ,$$

где значение $U_{max} = 0{,}04145$ является приблизительно максимально возможной мерой хаоса для любой исходной матрицы.

Чтобы оценить, являются ли статистически значимыми отличия найденной величины T для эмпирических данных от температуры в этой же матрице при случайном размещении видов, в разработанном программном продукте **NTC** (Nestedness Temperature Calculator – Atmar, Patterson, 1995) предусмотрено сопоставление исходной матрицы с нуль-моделью, разработанной по определенной схеме рандомизации.

Хотя на практике чаще всего используется индекс матричной температуры, были предложены и другие меры вложенности видовой структуры, которые, как правило, уже не учитывают взаимное расположение пустых и заполненных ячеек и основаны на механическом подсчете ситуаций, нарушающих порядок в упакованной матрице:

- *BR* (*мера несоответствия* – discrepancy measure) – подсчет числа несоответствий (т. е. непредусмотренного отсутствия или присутствия), которые должны быть устранены, чтобы получить максимально вложенную матрицу (Brualdi, Sanderson, 1999);
- *UT* (*число непредвиденных трансформаций* – unexpected transformations) – сумма отклонений от полной вложенности; UT = UA + UP, где UA – количество ошибочных ситуаций отсутствия слева от столбца с минимальным количеством ошибок; UP – число непредусмотренного наличия видов справа от этого столбца (Cutler, 1991);
- *NC* (nestedness index) – количество видов, общих для всех комбинаций пар местообитаний (Wright, Reeves, 1992);
- *N*0 (*число отсутствий* – absences) – количество видов, которые отсутствуют в местообитании с самым большим богатством, по сравнению с участком, где число видов минимально (Patterson, Atmar, 1986).

Недавно (Almeida-Neto et al., 2008) была представлена новая *метрика* «*уменьшения заполнения и парного наложения*» *NODF* (nestedness based on overlap and decreasing fill):

$NODF = \frac{\sum NP_{paired}}{s(s-1)/2 + m(m-1)/2}$, где s – число видов, m – число участков, а NP_{paired} – парная степень вложенности, т. е. подсчет, сколько взаимосвязанных пар видов и сколько взаимосвязанных пар участков существует в определенных подмножествах элементов матрицы. Использование *NODF* позволяет разложить полную вложенность на компоненты, обусловленные взаимосвязью видов ($NODF_{строк}$) и взаимосвязью местообитаний ($NODF_{столбцов}$). Это представляется важным, так как часто локальной задачей исследования сообществ является либо только оценка взаимосвязей внутри видовых композиций, либо выделение специфичных местообитаний.

Одна из целей nestedness-анализа состоит в том, чтобы идентифицировать виды, которые в наибольшей мере уменьшают оценку вложенности матриц. Существование таких видов, названных «*особенными*» (или *идиосинкразивными* – idiosyncratic species), можно объяснить (Atmar, Patterson, 1993) миграцией видов с последующей изоляцией, географическими барьерами и конкурентным исключением. В общем случае идиосинкразивные виды позиционируются как группа, в наименьшей мере подверженная действию факторов, сопряженных с экологическими и географическими градиентами. Поскольку у особенных видов есть больше потенциальных комбинаций для непредвиденных отсутствий и присутствий, они должны иметь более широкий диапазон встречаемости. Например, для сообществ диатомовых водорослей показано (Soininen, 2008), что у особенных видов наблюдаются более широкие ареалы, чем у «обычных» видов.

С точки зрения планирования охраняемых территорий представляет интерес идентификация в ходе nestedness-анализа «особенных» местообитаний – участков, видовой состав которых в наибольшей мере специфичен. Впрочем, у экологов нет единого мнения о том, территориям с каким соотношением вложенности и особенности следует отдать приоритет при сохранении (см. дискуссию относительно проблемы SLOSS «отдельный большой регион или несколько небольших участков – single-large-or-several-small; Patterson, Atmar, 2000).

4.4. Статистические методы оценки структурированности матриц и согласованной встречаемости видов

Большинство перечисленных выше мер основано на подсчете «особых» структурных фрагментов в исходной матрице. Д. Шлютер (Schluter, 1984) предложил использовать для тестирования значимости множественных зависимостей между видами оценку дисперсии числа видов для различных местообитаний.

Если использовать известное положение теории вероятности о том, что общая дисперсия суммы случайных величин равна сумме дисперсий каждой из них плюс удвоенная сумма ковариаций, то на основе данных о встречаемости видов X_i (i = 1, 2, …, s) в m выборках можно записать $D(X) = \sum_{i=1}^{s} D(X_i) + 2\sum_{i<l}^{s} D(X_i, X_l)$. Компоненты разложения дисперсии в общем случае неизвестны, но могут быть рассчитаны по данным мониторинга.

Пусть выборочная оценка $D(X_i)$ равна $\sigma_i^2 = t_i(1-t_i)$, где t_i – средняя встречаемость i-го вида в имеющихся выборках, а оценка $D(X)$ общей дисперсии числа видов – $\sigma_X^2 = \left(\sum_{i=1}^{S}(t_i - t)\right)/s$, где t – наблюдаемое среднее число видов в одной пробе.

Нулевая гипотеза (H_0) об отсутствии сопряженности между видами справедлива, если сумма ковариаций $D(X_i, X_j)$ равна нулю. Это будет верно, когда виды независимо распределены среди выборок, но также может иметь место, если положительные и отрица-

тельные ковариации уравновешивают друг друга. Если проверять H_0 против альтернативной гипотезы, что есть чисто положительная или чисто отрицательная зависимость между видами, то, при справедливости H_0, имеет место отношение $E(\sigma_X^2 \mid t_1, t_2, ..., t_s) = \sum_{i=1}^{S} \sigma_i^2$. Следовательно, выражение $V = \sigma_X^2 / \sum_{i=1}^{S} \sigma_i^2$ может служить обобщенным «*индексом взаимозависимости*» видов в выборках.

Если H_0 верна, то ожидаемое значение $V = 1$. Значение V, большее или меньшее 1, указывает, что между видами в сообществе есть статистическое положительное или отрицательное взаимодействие. Термин «статистическое» употреблен нами, чтобы подчеркнуть частный характер таких связей: например, на фоне «нейтральных» взаимоотношений между большинством видов может оказаться одна или несколько пар «видов-антагонистов», редко встречающихся совместно.

Д. Шлютер подчеркивал, что само по себе значение V при отклонении гипотезы H_0 не может быть свидетельством каких-либо конкретных экологических механизмов взаимодействия между видами (например, конкуренции). В самом общем виде дисперсионный тест оценивает не более чем меру внутриматричных различий распределения видов по рангам обилия (measure of matrix wide differences of rank order distributions). И эти различия могут быть объяснены самыми различными экологическими причинами (табл. 4.1).

Таблица 4.1

Экологические процессы с акцентом на межвидовые отношения, которые могут привести к положительной или отрицательной зависимости среди видов (по: Schluter, 1984)

Зависимость	Взаимодействие	Пример процесса
Отрицательная	Конкуренция	Недоступность или недостаток лимитирующего ресурса, что приводит к возможности исчезновения вида или снижению его популяционной плотности.
Отрицательная	Нет	Различия в распределении ресурса или климатические особенности, приводящие к локальному снижению популяционной плотности.
Отрицательная	Мутуализм	В условиях локального переизбытка ресурса, за который идет соревнование между видами, плотность некоторых популяций увеличивается более интенсивно, чем остальных.
Отрицательная	Хищничество	Высокий удельный вес хищников приводит к локальному снижению популяционной плотности жертв.
Положительная	Конкуренция	Плотность конкурирующих видов колеблется в унисон в ответ на флуктуацию запасов совместно используемых лимитирующих ресурсов.
Положительная	Нет	Совместная реакция на изменения в климате или в росте нелимитированных ресурсов.
Положительная	Мутуализм	Виды синхронно увеличивают вероятность выживания друг друга и т. д.
Положительная	Хищничество	Обилие хищников увеличивается при росте популяций жертв и наоборот.

Если предположить, что для данных мониторинга справедлива центральная предельная теорема, то при достаточно больших значениях m и s последовательности t_i можно интерпретировать как независимые случайных величины, приблизительно распределенные по нормальному закону. Тогда индекс взаимозависимости видов V при справедливости H_0 будет иметь χ^2-распределение с m степенями свободы, а критические

значения для того, чтобы отклонить нулевую гипотезу определяются пределами табличных значений, например: $\chi^2_{m,\,0,05} \le mV \le \chi^2_{m,\,0,95}$.

Предположения о нормальности распределения вероятностей обнаружения видов привели к тому, что *дисперсионный тест Шлютера* в его оригинальной версии оказался чрезвычайно чувствителен к конспецифичным скоплениям особей и склонен к гипердиагностике взаимодействий. Для решения этих проблем были разработаны имитационные процедуры рандомизации путем многократного случайного перемешивания видов (Gotelli, 2000) или с использованием алгоритма минимизации сумм квадратов (Mikloš, Podani, 2004), что дало возможность восстанавливать характер эмпирического распределения стандартных отклонений σ_X, σ_i и, соответственно, связанного с ними V-значения.

Проведенное тестирование (Gotelli, 2000) показало, что этот индекс имеет весьма интересные свойства и может быть полезен, например, для проверки *гипотезы ограниченности ниши* (Wilson, 1987), утверждающей, что дисперсия числа видов для различных местообитаний может быть достаточно мала, если изучаемая гильдия лимитируется конкурентностью. Также этот тест имеет существенно более низкую ошибку первого рода (ложное отклонение верной нулевой гипотезы) по сравнению, например, с метриками «шахматной доски».

Приведенные выше метрики дают обобщенную оценку степени структурной организации сообществ, детально не рассматривая при этом, какие конкретно виды показывают наибольшую согласованность встречаемости по отношению к другим видам. Поэтому другой проблемой статистического анализа является идентификация в матрицах наблюдений таких пар видов или их комбинаций, вероятность совместного появления которых определяется экологическими взаимодействиями (сегрегацией или кооперацией) и не может быть интерпретирована как случайность. Если принять во внимание, что общее число парных комбинаций в сообществе из s видов равно $s(s-1)/2$, то эмпирические данные могут часто содержать сотни и тысячи потенциально взаимодействующих пар видов. Подобная задача имеет большую актуальность также, например, в геномике и протеомике, где необходимо статистически проанализировать тысячи сочетаний генов в микрообразцах. Как исследователь может оценить, какие из этих парных комбинаций экологически достаточно интересны, чтобы обеспечить «полезность» дальнейшего анализа?

Н. Готелли и В. Ульрих (Gotelli, Ulrich, 2010) детально проанализировали четыре возможных подхода к выделению статистически значимых пар видов в биогеографической матрице «отсутствия/присутствия»: использование границ обычного 95%-го доверительного интервала, последовательный тест Бонферрони, а также два испытания, основанных на среднем и доверительном интервале в рамках эмпирического байесового метода (Efron, 2005). Для этого может быть использована любая из перечисленных в разделе 4.2 мер согласованности или взаимного исключения (индекс C-заполнения, мера близости t, метрика Съеренсена или иная). Для удобства значение произвольной метрики C нормируется в диапазоне от 0 (полное совпадение появления двух видов) до 1 (полная сегрегация).

Простейший способ проверить, случайно ли связаны два вида в своем совместном возникновении, состоит в том, чтобы сравнить эмпирическое значение метрики C с её теоретическим или восстановленным методами Монте-Карло частотным распределением. Согласно методу доверительных границ (*CL-критерий*), для дальнейшего рассмотрения отбираются все пары видов, для которых индекс C выше или ниже границ 95%-го доверительного интервала тестового распределения. Но здесь проблема состоит в высокой ошибочной степени обнаружения взаимодействий FDER (false detection error rate; Sanderson, 2004), поскольку многие комбинации пар видов показывают статистически значимые результаты, даже если исходная матрица была сформирована полностью случайным образом. Например, в такой матрице из 30 видов при общем числе пар $30 \cdot 29/2 = 435$ оказались признанными неслучайными комбинации из 22 пар, тогда как теоретическая вероятность такого исхода составляет только 3 пары (Gotelli, Ulrich, 2010).

Коррекция Бонферрони (Carlo Emilio Bonferroni, 1892-1960) – простой способ сократить FDER, деля уровень значимости на общее количество испытаний. Однако в случае парных зависимостей видов эта процедура будет приводить к слишком консервативным выводам из-за большого числа комбинаций видов в типичной матрице «отсутствия/присутствия». К тому же метод Бонферрони предполагает, что проводимые испытания независимы, что далеко не очевидно для биогеографических матриц, полученных по схеме «пассивной выборки». Заслуживает внимания метод взвешенной коррекции принимаемого уровня значимости нулевой гипотезы H_0 (Benjamini, Yekutieli, 2001), основанный на последовательном уменьшении весов вероятностей P_k в ходе проведения испытаний. Однако есть веские философские и статистические причины для отказа от любых исправлений или регулирований P-значений при проверке гипотез (Gotelli, Ellison, 2004).

Полезным статистическим инструментом, обеспечивающим разумное снижение ошибки FDER, является эмпирический байесовский подход (Efron, 2005; Gotelli, Ulrich, 2010). При его использовании сначала формируется ожидаемое частотное распределение используемой метрики C согласованности видов, вычисляемое в процессе многократного случайного перемешивания строк и столбцов исходной матрицы (см. раздел 4.5). Весь интервал нормирования от 0,0 до 1,0 делится на диапазоны (например, на 22 равномерных отрезка: 0, 0-0,05, 0,05-0,1, … , 0,9-0,95, 0,95-1), и по результатам имитации рассчитывается средняя частота появления случайной величины C в каждом диапазоне и стандартное отклонение этой частоты. Полученная функция распределения дает представление о том, как часто определенные значения метрики в условиях выполненного эксперимента могут оказаться случайными, независимо от того, являются ли эти величины статистически достоверными при анализе.

На втором этапе значения метрики C, рассчитанные непосредственно по исходной матрице наблюдений для каждой пары видов, сравниваются с полученным частотным распределением этой совокупности видов в рандомизированных матрицах. Пусть, например, в диапазоне значений C между 0,60 и 0,65 в эмпирических данных наблюдалось 24 парных комбинации видов, в то время как математическое ожидание этой частоты в той же матрице при случайном формировании сообщества составляет 19,85. Если все 24 пары будут признаны статистически значимыми, то это автоматически приведет к ошибке гипердиагностики взаимодействия FDER = 19,85/24 = 83%. В результате сравнения эмпирического и нуль-модельного распределений частот для дальнейшего анализа в каждом диапазоне отбирается строго ограниченное количество потенциально взаимодействующих пар видов с максимальными значениями Z-критерия (см. раздел 4.6). Это число пар в каждом случае определяется как положительное превышение эмпирической частоты над средней частотой (M-критерий Байеса) или её 95%-го доверительным интервалом (CL-критерий Байеса), полученными из моделируемого распределения.

На третьем этапе это множество отобранных пар взаимодействующих видов сокращается за счет того, что сохраняются только комбинации, признанные статистически значимыми на основе простого CL-критерия. Такое комплексное сочетание группового эмпирического байесовского подхода с индивидуальными испытаниями отдельных пар-претендентов обеспечивает значительное сокращение объема аналитической работы и ошибки гипердиагностики FDER. Следует отметить, что байесовский M-критерий по уровню консервативности занимает промежуточное положение между простым CL-критерием и коррекцией Бонферрони, тогда как CL-критерий Байеса приводит к весьма жестким правилам отбора небольшого количества потенциально взаимодействующих пар.

Наконец, некоторые терминологические вопросы возникают по поводу того, почему эти методы названы «байесовскими», в то время как они основаны на обычном частотном ограничении ошибки 1-го рода. Однако Б. Эфрон (Efron, 2005) развернуто показал, что у оценки вероятности ошибочной диагностики взаимодействий есть хорошее объяснение с позиций байесовской идеологии.

4.5. Нуль-модели для проверки статистической значимости мер согласованности видов

После статьи Э. Коннора и Д. Симберлоффа (Connor, Simberloff, 1979), всесторонняя проверка гипотез о достоверности взаимодействий между видами в ассоциациях стала одной из ключевых проблем изучения сообществ (Simberloff, Martin, 1991). Для оценки статистической значимости широкого круга индексов и метрик, оценивающих межвидовые связи, был разработан набор вычислительных алгоритмов, осуществляющих генерирование нуль-моделей с различными свойствами и их сравнение с матрицами наблюдений.

Процедуры использования любой нуль-модели основаны на последовательности действий классического рандомизационного теста (Edgington, 1987).

1. Имеющиеся у исследователя теоретические предположения об экологической сущности межвидовых взаимодействий сводятся к более простым гипотезам о существовании «особых» структурных фрагментов в матрице отсутствия/присутствия видов, полученной в ходе мониторинга. Исходя из классической теории Р. Мак-Артура (MacArthur, 1965, 1972), обычно считается, что совместное появление видов подавляется конкуренцией и его вероятность меньше, чем это ожидалось бы в сообществе, где виды распределяются по участкам случайно и независимо друг от друга.
2. Выбирается тип индекса $\boldsymbol{E}$, отражающего уровень взаимосвязи видов, и вычисляется его значение E_{obs} для эмпирической матрицы инцидентности.
3. Исходная матрица с данными наблюдений рандомизируется в соответствии с принятым типом нуль-модели (т. е. виды тем или иным способом случайно перемешиваются по участкам) и вычисляется значение E_{sim} для преобразованной матрицы.
4. Шаг (3) повторяется много раз (например, 1000), чтобы воспроизвести гистограмму частот распределения E_{sim}. Если эта гистограмма включает диапазон значений, ожидаемых для E_{obs}, то полученная нуль-модель верна. График распределения плотности оценок вероятности $\boldsymbol{E}$ может быть использован для проверки нулевой гипотезы, утверждающей, что E_{obs} является случайной реализацией стохастического процесса появления видов.
5. Используются стандартные методы проверки статистических гипотез, чтобы интерпретировать E_{obs}. Для этого на графике плотности распределения $f(E_{\text{sim}})$ определяется частота, соответствующая E_{obs}, и вычисляется вероятность P того, что случайная величина $\boldsymbol{E}$ примет значение, меньшее, чем E_{obs}. По традиции, если $0{,}05 < P < 0{,}95$, то принимается нулевая гипотеза H_0 об отсутствии значимых отличий исходной и рандомизированной матриц по индексу $\boldsymbol{E}$, а если P выходит за эти границы, то H_0 отклоняется в пользу альтернативы.
6. Результаты статистической проверки используются, чтобы оценить первоначальные гипотезы о значимости метрики E (а, следовательно, о достоверности взаимодействий).

Последний пункт является одним из спорных аспектов использования нуль-моделей (Gotelli, Graves, 1996; Gotelli, McGill, 2006), вызывающим возражения экологов и математиков. Как известно, любой статистический анализ нуждается в двух конкурирующих гипотезах. Формально оцениваемые метрики E соответствуют H_1, т. е. гипотезе о *существовании* определенных структурных фрагментов в пределах исходной матрицы, в то время как статистической проверке подвергается совсем другая, уже нулевая гипотеза H_0 о случайном характере распределения видов по участкам. Насколько эти гипотезы эквивалентны, остается недоказанным моментом.

Существует достаточно большой набор взглядов на способы конструирования алгоритмов рандомизации, поэтому существенной проблемой статистического анализа является выбор вычислительной процедуры генерации нуль-модели. Для бинарных матриц отсутствия/присутствия видов эти процедуры осуществляют перестановку значений в преде-

лах матрицы с теми или иными ограничениями на перебор. «Равновероятные» **EE** (Equiprobable) алгоритмы ограничены только постоянным значением средней встречаемости видов и сохраняют минимальное количество информации, содержащейся в исходных данных. Фиксированные алгоритмы дополнительно лимитируют число видов для некоторых подмножеств местообитаний. Например, наименее «либеральная» дважды фиксированная модель **FF** (Fixed-Fixed) требует, чтобы общая встречаемость видов в строках и столбцах нуль-матрицы соответствовала бы наблюдаемым значениям в эмпирической матрице. Комбинированная модель **EF** (Equiprobable-Fixed) сохраняет неизменным видовое богатство каждого участка, но позволяет частотам возникновения видов (т. е. общему количеству единиц в строках) изменяться беспорядочно и равновероятно. Модель **FE** (Fixed-Equiprobable) делает то же самое в отношении столбцов (т. е. местообитаний).

Разработана также коллекция пропорциональных (Proportional) моделей **P**, которые в процессе перебора отдают предпочтение тем или иным видам согласно вероятности их встречаемости на участках (Patterson, Atmar, 1986), а также пропорционально общей численности особей, либо иным популяционным параметрам.

Проверка статистической значимости эффекта вложенности видов с использованием нуль-моделей в идеале должна обладать, как хорошей мощностью идентифицировать структурированность изучаемого сообщества, если взаимосвязи между видами имеют неслучайный характер (низкая ошибка второго рода), так и не обнаруживать вложенности в матрицах, в которых распределение видов сгенерировано стохастическим процессом (низкая ошибка первого рода). В теоретическом плане становится все более ясным, что статистические тесты, которые используют полностью равновероятные нуль-модели **EE**, не вполне соответствуют экологическому определению стохастичности (Gotelli, McGill, 2006). Необходима ориентация на практически работоспособные нуль-модели, которые в состоянии отличать экологические и эволюционные механизмы от чисто статистических и выборочных артефактов (Krishna et al., 2008).

Все представленные выше имитационные процедуры с различными схемами генерации нуль-моделей можно ранжировать в ряд по мере снижения ограничений на рандомизацию: **P > FF > (EF, FE) > EE**. В этом же направлении увеличивается уровень ошибки 1-го рода (т. е. анализ становится менее консервативным) и нулевая гипотеза будет отклоняться, даже если наличие межвидовых взаимодействий выглядит весьма сомнительно. Однако в этом же ряду уменьшается ошибка 2-го рода (принятие ложной нулевой гипотезы), поэтому предпочтение часто отдается моделям типа **EF**. Согласно другим рекомендациям в условиях пассивного формирования выборок более надежные результаты дают модели **FF**.

Любой анализ структурированности матрицы должен быть инвариантен к простым алгебраическим преобразованиям, которые не затрагивают биологической сущности проверки гипотез. Результаты статистической обработки также не должны зависеть от размера матриц или их формы. Несмотря на то, что **NTC** (Температурный Калькулятор) стал стандартным инструментом при анализе матриц отсутствия-присутствия, он по различным причинам не всегда удовлетворяет экологов. В работах (Wright et al., 1998; Greve, Chown, 2006) была показана зависимость значения T от размера матриц и увеличение ошибки 1-го рода в случае большого числа видов-мутуалистов. В частности, другая программа BINMATNEST (Rodrigués-Gironés, Santamaría, 2006) решает проблемы большого размера матриц и генерирует целую совокупность различных нуль-моделей, чтобы точнее оценить значимость эффекта вложенности видов.

Н. Готелли и В. Ульрих (Gotelli, 2000; Ulrich, Gotelli, 2007; Ulrich et al., 2009) провели широкое тестирование восьми метрик структурированности и девяти алгоритмов генерации нуль-моделей на примере 200 эталонных рандомизированных матриц и 200 эмпирических матриц, полученных в ходе пассивной выборки. Среди анализируемых метрик наилучшую оценивающую способность продемонстрировали индекс несоответствия *BR* (Bru-

aldi, Sanderson, 1999) и метрика наложения/заполнения *NODF* (Almeida-Neto et al., 2008), тогда как многие другие популярные индексы не показали хороших статистических свойств.

4.6. Обсуждение результатов на примере сообществ макрозообентоса малых рек

Анализ структурированности и вложенности сообществ макрозообентоса выполнялся на примере речной системы Байтуган–Сок (см. раздел 1.3). Первая серия расчетов была проведена по непосредственным данным гидробиологической съемки отдельно для трех выделенных участков:

- р. Байтуган (выполнено 48 проб на 8 станциях наблюдения, в результате чего обнаружено 198 видов и таксонов);
- верхнее течение р. Сок (53 пробы на 7 станциях, 190 видов);
- нижнее течение р. Сок (46 проб на 7 станциях, 176 видов).

Сформировано три бинарных биогеографических матрицы с гидробиологическими пробами в столбцах и таксономическим составом водотоков по строкам. Степень заполнения матриц составила от 6% до 9%. Расчеты выполнялись с использованием комплекса программ **CoOccurrence** и **Nestedness** (W. Ulrich).

Статистическая оценка значений индексов структурной взаимосвязи и вложенности проводилась с использованием четырех нуль-моделей: **EE, EF, FF** и **P** (пропорционирование проводилось по частотам встречаемости видов). На рис. 4.3 представлены 95%-ные доверительные интервалы, полученные по всем четырем схемам рандомизации для индекса «шахматной доски» *CHECKERS* и температуры вложенности матриц *T*, которые сравнивались с этими же показателями, найденными по эмпирическим данным для трех участков водотока.

Представленные результаты свидетельствуют, что проверка гипотезы о стохастическом характере распределения видов макрозообентоса по гидробиологическим пробам в значительной степени зависит от выбранной схемы рандомизации. У нас нет формальных оснований для предпочтения той или иной нуль-модели, однако на нашем материале считаем возможным согласиться с рекомендациями проведенных исследований (Gotelli, 2000; Ulrich, Gotelli, 2007) и выбрать для дальнейшего анализа в качестве базовой модель **EF** (Equiprobable-Fixed), которая сохраняет неизменным число видов в каждой пробе, но случайно и с равной вероятностью переставляет виды между ними.

Кроме проблемы «*Какую нуль-модель предпочесть?*», при оценке степени взаимосвязи видов не менее важной и неоднозначной представляется другая проблема: «*Какая метрика является самой верной?*» (Gotelli, 2000). Несмотря на широкое применение большинства из них, однозначного ответа на этот вопрос нет, поскольку велика зависимость индексов от конкретных условий: размеров, степени заполнения и структурных особенностей эмпирических матриц. С использованием трех блоков данных по макрозообентосу речной системы Байтуган–Сок нами были рассчитаны 10 показателей межвидового взаимодействия (табл. 4.2) и выполнена проверка гипотезы H_0 о стохастическом характере распределения видов на основе **EF**-схемы рандомизации.

Нулевая гипотеза принималась, если эмпирически найденное значение индекса E_{obs} не выходило за пределы 95%-го доверительного интервала. Положение E_{obs} в пределах смоделированного ходе имитации распределения $\boldsymbol{E}$ оценивалось с помощью стандартизованного Z-критерия: $Z = \dfrac{E_{obs} - E_{sim}}{SD_{sim}}$, где E_{Sim} и SD_{sim} – соответственно, центр распределения и стандартное отклонение. Это преобразование основывается на концепции «величины эффекта», широко используемой в мета-анализе (Gurevitch et al., 1992; Шитиков и

др., 2008), и часто дает приблизительно нормально распределенную статистику со средним, равным 0 и стандартным отклонением, равным 1.

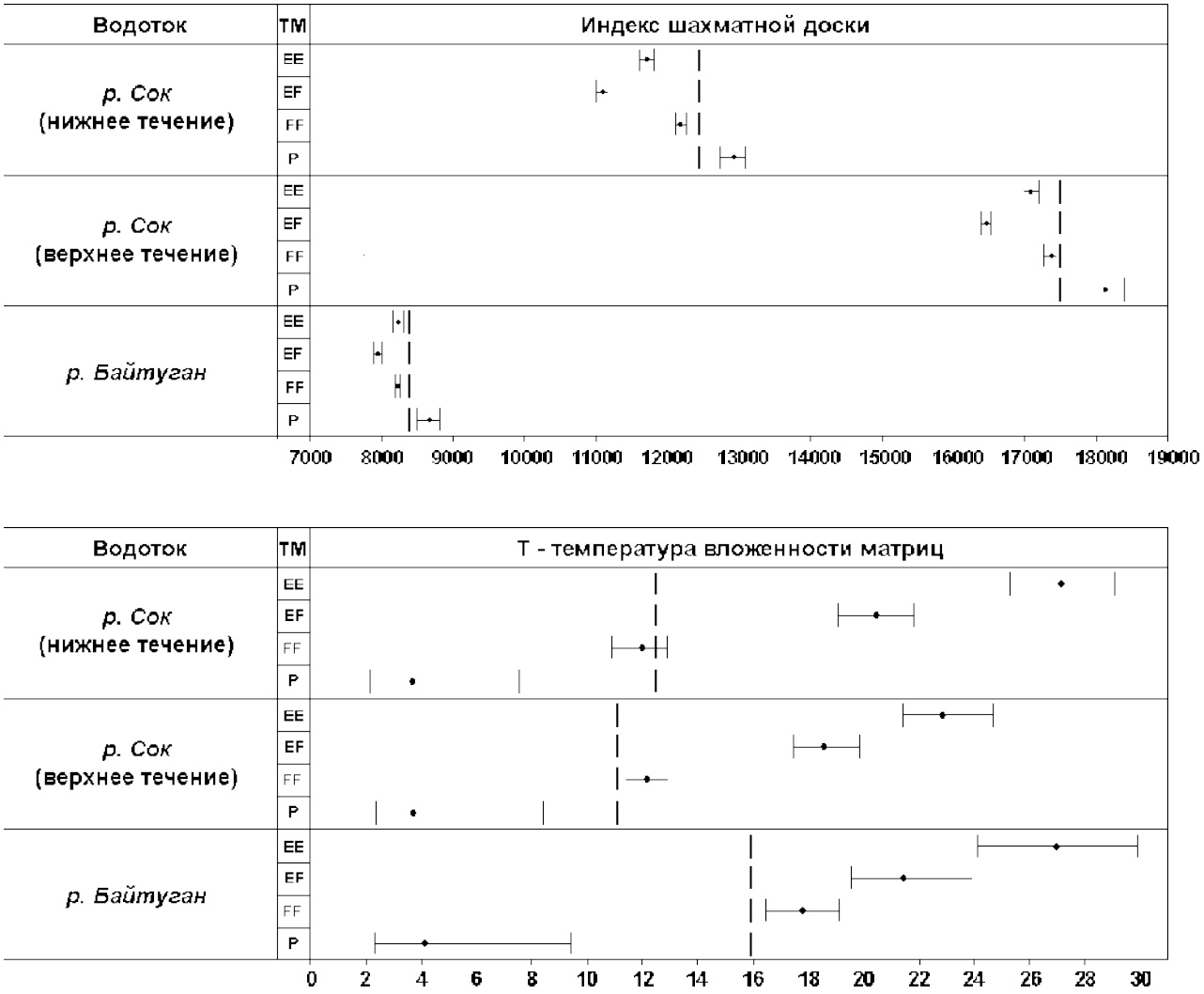


Рис. 4.3. **Индексы структурированности и вложенности эмпирических матриц для трех участков малых рек, в сравнении с доверительными интервалами, полученными по четырем нуль-моделям (обозначения моделей – по тексту раздела 4.5).**

В целом, для всех вычисленных метрик нулевая гипотеза была отклонена, что свидетельствует о высокой вероятности детерминированного характера распределения видов по пробам. Ориентируясь на абсолютную величину Z-критерия, взаимная обусловленность встречаемости видов в большей степени характерна для верхнего участка р. Сок и в наименьшей степени – для р. Байтуган.

С другой стороны, можно не согласиться с предлагаемой в публикациях (например, Ulrich, Gotelli, 2007) схемой статистического анализа. Используемые метрики E_{obs} оценивают, насколько велика детерминированная составляющая структурной организации изучаемого сообщества по сравнению со случайной композицией видов. Следовательно, отклонение H_0 целесообразнее проводить, если $E_{obs} > E_{Sim}$ (или $E_{obs} < E_{Sim}$ в зависимости от содержательного смысла применяемого индекса), и использовать для этого одностороннюю доверительную границу. Исключение составляет дисперсионный тест Шлютера (V),

схема использования которого обсуждалась выше. В этой связи, «странными» кажутся результаты, полученные для числа уникальных комбинаций (*COMBO*) и С-заполненности (*C-Score*). По величине этих метрик можно предположить, что стохастический характер эмпирических данных более выражен по сравнению с матрицами, подвергнутыми рандомизации.

Таблица 4.2

Статистический анализ десяти индексов структурированности и вложенности сообществ макрозообентоса на трех участках малых рек

Наименование индекса *E* и условие структурированности сообщества		**Участки рек**	**E_{obs} по эмпирическим данным**	**E_{Sim} – среднее из 100 итераций**	**Z-критерий**	**Нижний уровень ДИ при 95% ДВ**	**Верхний уровень ДИ при 95% ДВ**	**Итог: 0 – принятие H_o, + – отклонение**
Число уникальных комбинаций (*COMBO*)	E_{obs} < E_{Sim}	Байтуган	142	56,83	3,91	25	109	0
		Сок (верховья)	125	24,53	7,42	8	58	0
		Сок (нижн. теч.)	143	13,55	13,93	3	41	0
С-заполненность (*C-Score*)	E_{obs} > E_{Sim}	Байтуган	4,49	5,47	-16,16	5,33	5,61	0
		Сок (верховья)	7,29	9,4	-45,08	9,3	9,49	0
		Сок (нижн. теч.)	6,83	10,15	-50,27	10,02	10,28	+
Индекс «шахматной доски» (*CHECKERS*)	E_{obs} > E_{Sim}	Байтуган	8290	7926	11,89	7875	7992	+
		Сок (верховья)	17465	16449	26,08	16376	16526	+
		Сок (нижн. теч.)	12402	11070	31,15	10987	11154	+
Дисперсионный тест Шлютера (*V*)	см. таб. 5.1	Байтуган	7,92	31,17	-5,62	24,21	40,85	Отриц.
		Сок (верховья)	7,96	54,05	-7,75	44,17	67,81	Отриц.
		Сок (нижн. теч.)	5,67	46,57	-6,98	9,42	10,57	Отриц.
Средний коэффициент корреляции Спирмена	E_{obs} > E_{Sim}	Байтуган	0,11	0	26,04	-0,01	0,01	+
		Сок (верховья)	0,11	0	57,26	-0,01	0,01	+
		Сок (нижн. теч.)	0,16	0	60,77	-0,01	0,01	+
Температура матричной вложенности (*T*)	E_{obs} < E_{Sim}	р. Байтуган	15,87	21,38	-4,91	19,52	23,82	+
		р. Сок (верховья)	11,17	18,52	-12,91	17,40	19,76	+
		р. Сок (нижн. теч.)	12,46	20,41	-10,73	19,04	21,74	+
Метрика наложения/заполнения (*NODF*)	E_{obs} > E_{Sim}	Байтуган	13,59	10,14	9,51	9,59	10,99	+
		Сок (верховья)	11,53	8,01	16,9	7,63	8,39	+
		Сок (нижн. теч.)	13,92	10,03	13,61	9,39	10,52	+
Число несоответствий (*BR*)	E_{obs} < E_{Sim}	Байтуган	246	298,96	-11,52	288	308	+
		Сок (верховья)	446	573,8	-29,57	564	583	+
		Сок (нижн. теч.)	351	508,18	-28,42	497	518	+
Сумма отклонений от вложенности (*UT*)	E_{obs} < E_{Sim}	Байтуган	302	337,63	-11,16	330	342	+
		Сок (верховья)	591	642,09	-15,35	635	649	+
		Сок (нижн. теч.)	475	571,83	-25,58	564	580	+
Число видов, общих для комбинаций пар (*NC*)	E_{obs} > E_{Sim}	Байтуган	981	483	28,67	451	527	+
		Сок (верховья)	2928	1124	66,94	1072	1172	+
		Сок (нижн. теч.)	3091	1086	68,40	1043	1144	+

Примечание: ДИ – доверительные интервалы, ДВ – доверительная вероятность; прочие обозначения – по тексту.

Для каждого вида в биогеографической матрице могут быть рассчитаны число несоответствий BR_S и частная температура вложенности T_S (т. е. среднее евклидово расстояние ошибочных ячеек каждого вида от изоклинали – рис. 4.2). Представляет интерес рассмотреть список идиосинкразивных видов (табл. 4.3), вносящих наибольший «диссонанс» в структурные закономерности сообщества, которые можно выделить по максимальным значениям BR_S и T_S.

Анализ экологии видов показал, что индикаторная ценность отбора характерных комплексов по частной температуре T_S представляется неоднозначной. С одной стороны, максимальные значения T_S приобретают относительно часто встречаемые виды, обладающие повышенной сопряженностью с локальными условиями среды на фоне основного массива неспецифических видов. Так, в списке табл. 4.3 можно выделить структурообразующие таксоны *Simulium* sp., *Baetis* sp., являющиеся характерными для реофильного комплекса донных сообществ эталонной речной системы р. Байтуган. Аналогично высокие значения T_S свидетельствуют, например, о связи ценоза пелофильных и лимнофильных видов, таких как *Chironomus plumosus, Cladotanytarsus mancus, Micropsectra* sp., *Paratendipes albimanus,* с воздействием факторов пониженной аэрации и заиленности грунтов в нижнем течении р. Сок. С другой стороны, высокие значения T_S и меры несоответствия BR_S могут получить и случайные таксоны, например, мигрировавшие из смежных зон, но являющиеся в анализируемом сообществе «чужеродными» структурными элементами. В частности, вид *P. ferrugineus*, в силу своих экологических особенностей, не является характерным для р. Байтуган.

Таблица 4.3

Идиосинкразивные (особенные) виды макрозообентоса на различных участках

Вид	Код	*P*	T_S	BR_S	Участок реки
Amesoda solida	BiAms_so	4	16,52	5	Сок (нижн. теч.)
Pisidium inflatum	BiPis_i_	4	18,23	4	Сок (нижн. теч.)
Brillia modesta	ChBri_m_	7	17,25	10	Сок (верховья)
Chironomus agilis	ChChi_ag	5	13,96	8	Сок (нижн. теч.)
Chironomus obtusidens	ChChi_o_	5	18,94	8	Сок (верховья)
Chironomus plumosus	ChChi_p_	4 4	24,48 20	7 4	Сок (нижн. теч.) Сок (верховья)
Cladotanytarsus mancus	ChCld_m_	23	21,15	11	Сок (нижн. теч.)
Micropsectra sp.	ChMir_sp	8	20,88	11	Сок (верховья)
Monodiamesa bathyphila	ChMnd_ba	5 12	20,58 14,62	5 15	Байтуган Сок (верховья)
Paracladopelma gr. *Camptolabis*	ChPar_c	6	29,4	9	Байтуган
Paratendipes albimanus	ChPat_a_	4	21,92	6	Сок (нижн. теч.)
Pseudodiamesa branickii	ChPdd_br	4	21,89	6	Байтуган
P. convictum	ChPol_c_	5	17,3	8	Сок (верховья)
P. ferrugineus	ChPrc_f	5 10	27,9 14,83	7 11	Байтуган Сок (верховья)
Parametriocnemus sp.	ChPrm_sp	4	21,16	4	Байтуган
Paratanytarsus sp.	ChPtt_sp	5	17,77	8	Сок (нижн. теч.)
Baetis sp.	EpBst_sp	6	47,14	6	Байтуган
Cloeon dipterum	EpClo_di	6	14,47	9	Сок (верховья)
Caenis macrura	EpCsn_m	16	14,71	12	Сок (нижн. теч.)
Dicranota bimaculata	LiDir_b_	22	17,35	24	Сок (верховья)
Simulium sp.	SiSim_sp	8	32,29	18	Байтуган

Примечание: *P* – встречаемость, T_S – частная температура вложенности, BR_S – мера несоответствия для каждого вида.

Следует отметить, что в этой серии расчетов, основанных на гидробиологических пробах, не удалось проследить четкую пространственную упорядоченность отсортированных матриц: столбцы, относящихся к одному местообитанию, вследствие большой гетерогенности видового состава проб, оказались разобщены.

Это обстоятельство приводит нас к выводу о том, что важнейшей проблемой анализа закономерностей изменения видовой структуры речных систем на основе мониторинговых данных является учет масштабного фактора при переходе от уровня отдельной гидробиологической пробы к описанию всего биотопа и далее: участка водотока, всей экосистемы в целом, гидрографической сети крупного региона. В частном случае это связано с задачей статистически корректной «свертки» подматрицы гидробиологических проб, сделанных в разное время на одной и той же станции наблюдения, в вектор обилия видов, характерных для этого местообитания.

В рамках настоящего изложения представляет интерес оценить, как изменится характер структурной связанности видов, если таксономическое богатство каждой станции представить как объединение подмножеств видов во всех сделанных пробах. В результате усреднения проб, выполненных на станциях наблюдения, получим объединенную бинарную матрицу размерностью 13×375, где 13 – число биотопов, выделенных на всем протяжении водотока Байтуган–Сок, а 375 – общее количество обнаруженных видов макрозообентоса. В отсортированном виде эта матрица (см. фрагмент на рис. 4.4) в значительной мере отличается от обработанных ранее мелкомасштабных фрагментов по своей конфигурации и, в первую очередь, по более высокой степени заполнения $F = 0{,}21$. Сходную форму имеют и частные матрицы по отдельным участкам водотока: р. Байтуган (4×198) и р. Сок в её верхнем (5×206) и нижнем течении (4×176).

Вычисленные для агрегированных матриц значения основных мер вложенности и оценки их статистической значимости относительно базовой нуль-модели **EF** (Equiprobable-Fixed) представлены в табл. 4.4. В этих условиях матричная температура возросла до $T_{\text{obs}} \cong 33$ и намного превысила значения, рассчитанные для матриц, основанных на гидробиологических пробах (см. табл. 4.2), что свидетельствует о существенном искажении оценок структурной упорядоченности одних и тех же сообществ после агрегирования данных. Можно также отметить совпадение значений T_{obs} для отдельных рек и всего водотока в целом, т. е. никак не учитывается более высокая видовая и биотопическая неоднородность всей речной экосистемы по сравнению с каждым из её отдельных участков. Это является подтверждением широко обсуждаемой в литературе проблемы сильной зависимости температуры вложенности от заполнения и конфигурации исходных бинарных матриц.

Сходный вывод можно сделать и относительно двух других метрик вложенности. Мера несоответствия *BR*, ищущая число непредусмотренного отсутствия или присутствия «единиц», непосредственно связана с размерами матриц, не является стандартизованной величиной и не может использоваться для сравнения структурной детерминированности разных сообществ. Метрика «уменьшения заполнения и парного наложения» *NODF*, подсчитывающая число разнородных и иногда взаимоисключающих структурных артефактов отсортированной матрицы, также не является монотонной и нормированной оценкой вложенности. В то же время индекс *NODF* – единственный показатель, который дал совпадающий вывод о статистической значимости явления вложенности до и после агрегирования гидробиологических данных.

Представляется интересным проследить динамику изменения температуры вложенности для последовательности станций, выделенных по продольному руслу изучаемой речной системы. В целом для обоих участков р. Сок можно сделать вывод о существовании определенного «островного эффекта»: по мере приближения к устью происходит последовательное накопление видового богатства, частично аккумулирующего виды, оби-

тающие выше по течению. Эта тенденция нарушается лишь на ст. С_05 и С_13 (табл. 4.4), где рост частной температуры свидетельствует о существенной перестройке видового состава. Экосистема р. Байтуган, напротив, характеризуется значительной структурной гетерогенностью и β-разнообразием отдельные её участков.

График на рис. 4.5, обобщающий данные для всего водотока в целом, также показывает, что сообщества макрозообентоса в верхней части р. Байтуган (ст. Б_01 – Б_03), и, в особенности, на устьевых станциях р. Сок (СН_13) имеют специфический видовой состав, отличающийся от достаточно однородной и более детерминированной структуры остальной части экосистемы (СВ_3 – СН_12). При этом сообщества в верховьях р. Сок (СВ_01 – СВ_02) с более бедным видовым составом являются структурной «копией» сообществ на смежных участках, что подтверждается низкими значениями температуры вложенности.

	Б_03	Б_06	СВ_04	Б_01	СН_13	СВ_05	СВ_03	СН_09	СН_12	СН_08	Б_08	СВ_02	СВ_01	Итого
ChCri.b._	1	1	1	1	1	1	1	1	1	1	1	1	1	13
ChTar.sp_	1	1	1	1	1	1	1	1	1	1	1	1	1	13
Ol_	1	1	1	1	1	1	1	1	1	1	1	1	1	13
CeCul.sp_	1	1	1	1	1	1	1	1	1		1	1	1	12
Ep_	1		1	1	1	1	1	1	1	1	1	1	1	12
ChPlc.co_	1		1	1		1	1	1	1	1	1	1	1	11
LiDir.b._	1	1	1	1		1	1	1	1		1	1	1	11
ChPrc.f._			1	1	1	1	1	1	1	1	1	1		10
SiSim.sp_	1	1	1	1		1	1	1		1	1	1		10
ChCld.m._		1	1		1	1	1	1	1	1		1		9
ChMnd.ba_	1	1	1	1		1	1			1	1	1		9
ChPar.c._	1	1	1	1			1	1			1	1	1	9
ChPat.a._			1	1		1	1		1	1	1	1	1	9
ChPro.o._	1	1	1	1		1	1				1	1	1	9
Hc_	1	1	1			1		1	1	1	1		1	9
ChEuk.gr_	1	1	1	1		1	1				1	1		8
ChOrc.sp_	1	1	1	1		1		1			1	1		8
ChPol.n._			1		1	1	1	1	1	1		1		8
ChPrm.lu_	1	1	1	1		1	1				1		1	8
Co_		1	1	1	1		1		1	1			1	8
Si_	1	1	1			1	1	1			1		1	8
Tr_		1	1	1	1	1		1	1				1	8
BiEug.sp_	1	1	1		1	1		1		1				7
ChCri.s._			1	1	1	1	1	1					1	7
ChCry.d._			1		1	1	1	1	1	1				7
ChEuk.c._	1	1	1	1		1	1				1			7
ChMir.p._	1	1		1	1		1					1	1	7
ChNan.bi_		1	1	1	1		1	1	1					7
ChOrc.o._			1	1			1	1			1	1	1	7
ChPrc.c._	1			1	1				1	1		1	1	7
ChPtt.co_			1	1	1	1	1		1				1	7
Итого	102	99	98	97	91	89	89	83	63	59	55	49	49	

Рис. 4.4. **Фрагмент отсортированной матрицы «виды – станции наблюдений»**

Таблица 4.4

Оценка статистической значимости температуры вложенности *T*, меры несоответствия BR и метрики NODF по агрегированным матрицам

Водный объект	Последовательность станций в упакованной матрице, число видов (*S*) и частная температура вложенности (T_m) на каждой из них						Оценка статистической значимости метрик вложенности *E*				
							Метрика *E*	E_{obs}	E_{Sim}	LC 95%	UC 95%
р. Байтуган	№№ ст.	Б_03	Б_06	Б_01	Б_08		*NODF*	**46,60**	39,04	36,37	41,87
	S	102	99	97	55		*BR*	107	115	106	124
	T_m	34,3	27,6	38,3	9,58		*T*	34,4	31,2	28,3	35,0
р. Сок (верховья)	№№ ст.	С_04	С_03	С_05	С_02	С_01	*NODF*	**46,08**	38,12	35,53	40,59
	S	98	89	89	49	49	*BR*	118	139	128	148
	T_m	36,6	18,7	34,5	15,2	6,17	*T*	32,5	34,7	30,4	39,0
р. Сок (нижн. теч.)	№№ ст.	С_13	С_09	С_12	С_08		*NODF*	**43,20**	36,05	33,35	38,41
	S	91	83	63	59		*BR*	99	101	91	110
	T_m	35,4	23,8	21,8	15,5		*T*	33,9	32,7	28,9	38,0
Водоток в целом	№№ ст.	См. рис. 4.4 и 4.5					*NODF*	**28,61**	22,56	21,70	23,51
	S						*BR*	**487**	629	614	647
	T_m						*T*	**33,4**	43,4	41,1	45,7

Примечание: E_{obs} – значение метрики по эмпирическим данным, E_{Sim} – среднее по нуль-модели после 100 итераций, LC и UC – соответственно, нижний и верхний уровень 95%-го доверительного интервала.

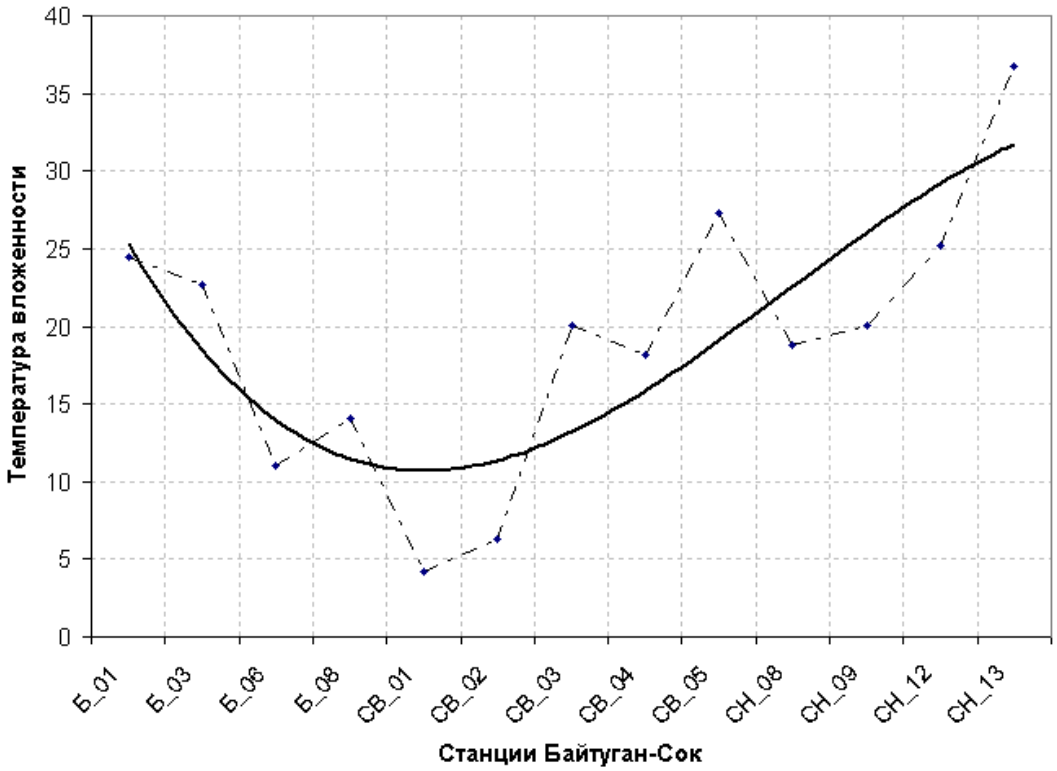


Рис. 4.5. **Распределение температуры вложенности станций по продольному профилю** (на оси абсцисс – порядковые номера станций наблюдений по участкам речной системы)

Для идентификации пар взаимно исключающих видов используем агрегированную матрицу суммарной встречаемости на 13 станциях наблюдений, причем для сокращения перебора выделим 129 видов, обнаруженных более чем в трех местообитаниях. Общее количество возможных комбинаций пар – 129 (129 – 1)/2 = 8256. Эффект «шахматной доски» для каждой комбинации будем оценивать с использованием нормированного индекса *C*-заполнения, значения которого C_{obs}, вычисленные по исходной таблице, будем сравнивать с расчетными величинами C_{sim}, полученными по наиболее консервативной нуль-модели **FF** (Fixed-Fixed), сохраняющей степень заполненности строк и столбцов.

В диапазоне значений *C* от 0,975 до 1 по эмпирической таблице было зафиксировано 1164 комбинаций пар видов, тогда как после 100 итераций рандомизации по нуль-модели – только 824, причем лимит отбора по байесовскому CL-критерию составил 61 комбинацию. Из этого числа 36 комбинаций пар оказалась статистически значимыми по простому *Z*-критерию, $Z = (C_{obs} - C_{sim}) / SD_{sim}$, а для 34 пар нулевая гипотеза о случайном характере взаимного исключения была отклонена и с учетом модифицированной поправки Бонферрони (Benjamini, Yekutieli, 2001). Некоторые парные комбинации таксонов макрозообентоса, прошедшие все перечисленные статистические фильтры и обнаруженные на 4-8 станциях из 13, но ни разу не встретившихся одновременно хотя бы в одном местообитании (C_{obs} = 1), представлены в табл. 4.5.

Таблица 4.5

Парные комбинации таксонов макрозообентоса, обнаруженные на T_1 и T_2 станциях из 13, но ни разу не встретившихся в одном местообитании

Таксон 1	T_1	Таксон 2	T_2	C_{sim}	*Z*	*p*	p_{BJ}
Parametriocnemus lundbecki	8	*Ablabesmyia monilis*	5	0.22	5.38	≅ 0	≅ 0
Stempellina bausei	6	*Brillia modesta*	5	0.34	4.47	0.000007	0.00007
Microchironomus tener	6	*Odontomesa fulva*	6	0.29	4.33	0.000014	0.00013
Odontomesa fulva	6	*Polypedilum scalaenum*	6	0.29	4.04	0.000051	0.00049
Eukiefferiella claripennis	7	*Ablabesmyia monilis*	5	0.3	4	0.000059	0.00056
Polypedilum nubeculosum	8	*Nemoura cinerea*	4	0.24	3.89	0.000096	0.00092
Ephemerella ignita	6	*Ablabesmyia monilis*	5	0.31	3.88	0.000099	0.00095
Chironomus obtusidens	6	*Nemoura cinerea*	4	0.37	3.7	0.000205	0.00197
Procladius choreus	7	*Cheumatopsyche lepida*	4	0.34	3.53	0.000403	0.00387
Cryptochironomus gr. *defectus*	7	*Pristina rosea*	4	0.31	3.52	0.000408	0.00392
Cryptochironomus gr. *defectus*	7	*Elmis aenea*	4	0.34	3.32	0.000856	0.00821
Odontomesa fulva	6	*Pisidium amnicum*	5	0.33	3.27	0.001	0.01
Ephemerella ignita	6	*Caenis macrura*	5	0.36	3.2	0.00132	0.0127
Ablabesmyia monilis	5	*Brillia modesta*	5	0.37	3.15	0.00158	0.0151
Ephemerella ignita	6	*Paralauterborniella nigrochalteralis*	4	0.35	3.08	0.00197	0.0189
Caenis macrura	5	*Elmis aenea*	4	0.41	3.01	0.00253	0.0242
Brillia modesta	5	*Paralauterborniella nigrochalteralis*	4	0.4	2.99	0.00272	0.0261
Caenis macrura	5	*Nemoura cinerea*	4	0.42	2.96	0.00298	0.0286

Примечание: C_{Sim} – среднее значение нормированного индекса *C*-заполнения по нуль-модели FF после 100 итераций; *Z* – нормированная разность наблюдаемых и расчетных значений; *p* и p_{BJ} – традиционная и скорректированная по методу Бонферрони (Benjamini, Yekutieli, 2001) вероятности, соответствующие *Z* – статистике.

Если взаимное исключение видов можно рассчитать по агрегированным матрицам, попарно сравнивая общий видовой состав на участках, то оценки взаимной согласованности видов могут быть корректно получены только при детализации информации на уровне гидробиологической пробы, т.е. с учетом пространственно-временных координат встречаемости.

Отбор пар взаимно согласованных видов осуществим на основе полной матрицы гидробиологической съемки, включающей 147 столбцов. Для сокращения перебора выделим 76 таксонов макрозообентоса, найденных более чем в 5 пробах. В диапазоне значений нормированного индекса *C*-заполнения от 0 до 0,3 в исходной матрице было найдено 129 сопряженных пар, тогда как в 100 модельных матрицах после рандомизации таких комбинаций в среднем оказалось только 27, что свидетельствует об отчетливой инкорпорированности сообщества. Дальнейший статистический анализ проводился по той же схеме, что и для взаимного исключения, и некоторые значимые парные комбинации взаимно согласованных таксонов представлены в табл. 4.6.

Таблица 4.6

Парные комбинации таксонов макрозообентоса, встретившиеся в T_1 и T_2 пробах, из которых T_{12} - совместно

Таксон 1	T_1	Таксон 2	T_2	T_{12}	C_{obs}	C_{sim}	Z	p_{BJ}
Ephemerella ignita	11	*Brachycentrus subnubilis*	6	5	0.091	0.867	-5.2	≅ 0
Lipiniella araenicola	13	*Dicrotendipes nervosus*	10	7	0.138	0.829	-4.93	≅ 0
Chironomus nudiventris	13	*Lipiniella araenicola*	13	10	0.053	0.799	-4.91	≅ 0
Chironomus nudiventris	13	*Dicrotendipes nervosus*	10	7	0.138	0.807	-4.74	≅ 0
Caenis macrura	17	*Henslowiana dupuiana*	6	5	0.118	0.817	-4.41	0.00008
Baetis rhodani	14	*Elmis aenea*	7	6	0.082	0.836	-4.37	0.0001
Herpobdella octoculata	7	*Brachycentrus subnubilis*	6	3	0.286	0.913	-4.32	0.00012
Henslowiana acuminata	6	*Henslowiana dupuiana*	6	3	0.25	0.903	-4.31	0.00013
Cladotanytarsus mancus	34	*Chironomus nudiventris*	13	12	0.05	0.649	-4.3	0.00013
Caenis macrura	17	*Chironomus anthracinus*	8	6	0.162	0.817	-4.16	0.00026
Cryptochironomus gr. *defectus*	21	*Unio pictorum*	7	6	0.102	0.772	-4.14	0.00028
Cladotanytarsus mancus	34	*Microchironomus tener*	11	10	0.064	0.661	-4.09	0.00035
Cricotopus (Isocladius) trifasciatus	10	*Rheocricotopus fuscipes*	7	4	0.257	0.875	-4.09	0.00035
Dicrotendipes nervosus	10	*Sphaeriastrum (=Rivicoliana) rivicola*	9	5	0.222	0.844	-3.99	0.00052
Cryptochironomus gr. *defectus*	21	*Microchironomus tener*	11	8	0.169	0.747	-3.87	0.00086
Procladius ferrugineus	38	*Microchironomus tener*	11	10	0.067	0.626	-3.67	0.00197
Simulium sp.	29	*Cricotopus* gr. *tremulus*	9	7	0.169	0.719	-3.67	0.00195
Polypedilum nubeculosum	33	*Dicrotendipes nervosus*	10	8	0.152	0.684	-3.66	0.00204
Baetis rhodani	14	*Pristina rosea*	6	4	0.238	0.837	-3.66	0.00205
Procladius ferrugineus	38	*Chironomus nudiventris*	13	11	0.109	0.617	-3.59	0.00271
Cladotanytarsus mancus	34	*Stempellina bausei*	7	6	0.118	0.669	-2.92	0.0287

Примечание: в шапке таблицы C_{obs} – эмпирическое значение нормированного индекса *C*-заполнения; остальные обозначения аналогичны табл. 4.5.

Необходимо подчеркнуть, что найденные закономерные положительные или отрицательные связи между отдельными таксонами в сообществе носят исключительно статистический характер и могут объясняться различными экологическими причинами. В частности, взаимное исключение в нашем случае является в первую очередь следствием про-

странственной неоднородности биотопических условий водотока на всем его протяжении, а не конкурентной борьбы за ресурсы. Например, в выделенной паре обитателей прибрежной зоны рек (табл. 4.5) реофильный вид с фильтрационным типом питания *Odontomesa fulva* встретился на шести станциях р. Байтуган, в то время как район обитания пело- и псаммофильного вида *Polypedilum scalaenum* находится в нижнем течении р. Сок (т.е. на расстоянии 300 км ниже по водотоку). Аналогичным образом оказались составлены и многие другие пары. В качестве примера возможной конкуренции или даже взаимного исключения при достижении определенной численности личинок можно привести обитание на сходных биотопах пар хищник-жертва: личинок хирономид активных хватателей *Ablabesmyia monilis* и фитодетритофагов-собирателей *Eukiefferiella claripennis* (другой пример *Ablabesmyia monilis - Brillia modesta*).

Более экологически обоснованный характер имеют пары видов с согласованной встречаемостью на одной станции (в одной пробе). В табл. 4.6 приведены примеры высокой вероятности совместного обнаружения, определяемой биологическими и экологическими особенностями выделенных таксонов, занимающих специфические микробиотопы, в частности пара пелофильных видов хирономид *Chironomus nudiventris + Lipiniella araenicola*, обитающих в заиленных биотопах или слабо заиленных песках нижнего течения р. Сок. Закономерна также положительная связь оксифильных поденок *Baetis rhodani* и жуков *Elmis aenea*, пело- и реофильных моллюсков *Henslowiana acuminata + Henslowiana dupuiana* или пар эврибионтных видов, характерных для эвтрофных участков системы равнинная река – водохранилище, таких, как *Cladotanytarsus mancus + Chironomus nudiventris, Polypedilum nubeculosum + Dicrotendipes nervosus, Procladius ferrugineus + Chironomus nudiventris.*

Описанная методика статистической «фильтрации» при отборе пар видов с взаимно согласованной встречаемостью является действенным и эффективным инструментом при изучении внутренних закономерностей организации сообществ. Комплексное сочетание группового эмпирического байесовского подхода с индивидуальными испытаниями отдельных пар-претендентов обеспечило значительное сокращение объема аналитической работы и ошибки гипердиагностики. Вместе с тем включение комбинации видов в список является лишь «сигналом» о возможно неслучайном характере связи между ними, а содержательная интерпретация, является эта связь следствием определенных межвидовых взаимодействий или разным уровнем толерантности к внешним факторам, остается за специалистом-экологом.

Расселение гидробионтов, связанное с их подвижностью, – один из важных факторов не только для лотических систем, но и вообще в экологии. В донных сообществах даже малоподвижные организмы распространяются с использованием планктонной стадии (некоторые виды моллюсков). Нельзя забывать и о гетеротопности хирономид, наиболее массовых и разнообразных таксонов водотоков. Расселение организмов в реках в нестабильных местообитаниях или подверженных сукцессиям вполне объяснимо. В незарегулированной речной системе Байтуган–Сок, подверженной эрозионным процессам, а значит, и появлению новых мест обитания (например, весенние паводки, влияние подпорных вод водохранилища в нижнем течении реки), более слабые конкуренты выживают благодаря быстрому расселению, опережая более приспособленных конкурентов (своеобразный, «рудеральный» тип жизненных стратегий; Грайм, 2005; Grime, 2007). Полагаем, что расселение, как один из ключевых вопросов в экологии, приобретает особое значение в качестве реакции видов на «фрагментирование» мест обитания.

ГЛАВА 5. МОДЕЛИ РАНГОВОГО РАСПРЕДЕЛЕНИЯ ОБИЛИЯ ВИДОВ

Тверже всего верят в то, о чем меньше всего знают…
Мишель де Монтень (1533-1592 гг.)

5.1. Классические модели распределения обилия и оценка роли редких видов

Структуру сообщества определяют два фундаментальных параметра – *список видов* и *функция неоднородности их обилия*. Если истинная общая численность особей в сообществе равна $\boldsymbol{N}$, то она может быть разложена на компоненты вектора таксономического разнообразия $\boldsymbol{N} \rightarrow \{N_1, N_2, \ldots, N_s\}$, где S – количество видов; N_i – численность экземпляров i-го вида, которая статистически значимо отличается от нуля. Значения N_i и размерность вектора S являются сложной функцией пространственных координат, ограничивающих сообщество, периода наблюдений, интенсивности выборочных усилий, способа подсчета численностей, принятого масштаба агрегирования и т. д. Точная, несмещенная и адекватная оценка этих показателей по мониторинговым данным связана с самостоятельными непростыми проблемами. Будем, однако, считать, что эти проблемы были успешно преодолены.

В историческом аспекте самые значительные усилия в реализации потребности понять структуру сообщества сводились к моделированию достаточно примитивной кривой распределения численности особей по видам, которая может быть представлена в трех основных формах:

- *кривая рангового распределения*, т. е. диаграмма численностей N_i, где по оси абсцисс – порядковый номер вида в ранжированном ряду $N_i \geq N_{i+1}$ (она же – *кривая доминирования-разнообразия Уиттекера* или *кривая важности видов*);
- *гистограмма распределения общего количества видов S по диапазонам численностей особей* (как правило, эти диапазоны выбираются в соответствии с правилом удвоения, или октав Престона: 2, 4, 8, 16, 64 и т. д.);
- *кумулятивная кривая накопленных долей численностей видов*, которую во многих предметных областях называют *кривой Лоренца* (Lorenz, 1905).

Характерной чертой всех этих распределений является хорошо подтверждаемая детерминированность в области видов-доминантов и таксонов со средней численностью особей, тогда как в области видов, представленных единичными экземплярами, многие модели становятся неопределенными. По сути, качество любой функции, аппроксимирующей распределение обилия видов, зависит, в первую очередь, от того, насколько адекватно она отражает реальные соотношения в области редких видов.

Понятие редкости видов в экологии и биогеографии принимает много форм в зависимости от пространственных и временных масштабов исследования. Редкость в контексте α-разнообразия (Whittaker, 1972) на уровне естественной среды обычно понимается как фракция видов, которые даже после интенсивного обследования представлены одиночными особями преимущественно в одной или двух пробах (уникальные виды и дубликаты). Часто увеличение повторности наблюдений приводит к обнаружению дополнительных индивидуумов этих редких видов, перемещая их в классы более высокой встречаемости, но одновременно появляются новые виды, квалифицируемые как редкие. Это – результат работы «демона Престона» (Preston, 1948) или передвигающаяся «линия завесы» между обнаруженными и необнаруженными видами при увеличении выборочного усилия.

Биологи долго подозревали, что редкие виды во многих сообществах – это смесь подлинно редких (т. е. «постоянных» или «резидентных») видов и случайных (или «транзитных») видов, которые могут быть широко распространены в других местах (Magurran, Henderson, 2003). Само по себе понятие «редкость вида» опирается на анализ статистической структуры сообществ, т. е. на исследование законов распределения каких-либо характеристик экосистем. Этот подход казался перспективным всю вторую половину прошлого века, и был предложен широкий круг моделей, описывающих неоднородность популяционной плотности таксономических единиц, однако ни одна из них не оказалась универсальной.

Полвека поисков привели, наконец, к осторожному признанию, что *характер распределения обилия видов лучше всего* (но далеко не всегда) *описывается*, например, *логнормальной моделью*. Дальнейшие поиски хорошей универсальной закономерности, пригодной для описания любых эмпирических последовательностей подмножеств видов, стали считаться мало результативными, – слишком неустойчива ситуация, которая определяется многообразием, изменчивостью и мозаичностью природных биоценозов (Левич, 1980). Так, распределение Пуассона, которое, в целом, является признанным эталоном, нетипично, в частности, для объектов учета сорной растительности и многих сообществ насекомых, которым свойственно не только предпочтение отдельных частей учетной единицы (поля, делянки, специфического биотопа), но и тяготение друг к другу, т. е. «стадность».

Остановимся на основных подходах к моделированию кривых распределения видов. Еще в начале XX века К. Раункиер (цит. по: McIntosh, 1967a) обнаружил, что обилие видов растений на пробных площадках следует некоторому правилу, названному им *законом частоты*. Дальнейший поиск количественных закономерностей видового состава сообществ протекал по двум сценариям: индуктивному и дедуктивному (Hubbel, 2001).

Вехой *индуктивного подхода*, доминировавшего вначале, стали работы энтомологов А. Корбета и К. Вильямса (1939-1941 гг.), которые заметили (Corbet, Williams, 1973; цит. по: Василевич, 1972), что частотный ряд численности видов может быть представлен относительно гладкой гиперболической прогрессией со многими редкими и немногими фоновыми видами. Они обратились к Р. Фишеру (Sir Ronald Aylmer Fisher, 1890-1962), который подобрал составное распределение, имеющее вид отрицательного бинома. Было, однако, две проблемы:

- как найти значение нулевого члена, оценивающего число видов, слишком редких, чтобы быть представленными в эмпирической выборке, и
- как найти, сколько частотных градаций должно быть в сообществе.

Соавторы (Fisher et al., 1943) поступили интуитивно просто: нулевой член отсекли, а видовую насыщенность сообщества представили эффективно бесконечной, разложив отрицательный бином в сходящийся *логарифмический ряд Маклорена*: αx, $\alpha x^2/2$, …, $\alpha x^n/n$, где $\alpha x^i/i$ – количество видов в группе с i-й численностью особей, α и x – параметры лог-серии. Оценка α и x осуществляется, исходя из условий: общее число особей в сообщества $\boldsymbol{N} = \alpha x/(1 - x)$, $x < 1$, общее число видов или сумма всех членов ряда $S = -\alpha \ln(1 - x)$. Параметр α, известный как *α Фишера*, является широко используемой мерой видового разнообразия, потому что теоретически независим от объема выборки (Fisher et al., 1943). Впрочем, мы убедились на практике, что это далеко не так (см. рис. 2.9б).

Несколько лет спустя Ф. Престон (Preston, 1948), не удовлетворенный аппроксимацией данных лог-сериями Фишера, нашел, что относительное распределение обилия видов имеет колоколообразный характер, т. е. виды, имеющие промежуточное обилие, встречаются чаще, чем редкие или доминирующие виды. Поскольку предложенное им логарифмическое нормальное распределение, в отличие от лог-серий, является непрерывным, Престон разработал метод «дискретной подгонки», основанный на интервалах удвоения численности особей, названных им «октавами» (Мэгарран, 1992):

$$S_R = S_0 e^{-a^2 R^2}, \quad a = 1/\sqrt{2\sigma},$$

где S_0 – число видов в медианной октаве; S_R – число видов в октаве, отстоящей на R влево или вправо от медианной, R = 0, 1, 2, …; σ – дисперсия логнормального распределения. Подбор кривой распределения может быть сделан приблизительно, рассчитав уравнение линейной регрессии $\ln(S_R)$ по R^2 и получив в качестве его параметров a^2 и $\ln(S_0)$. Более точную настройку усеченного непрерывного логарифмически нормального распределения к эмпирическим данным можно выполнить с использованием метода наибольшего правдоподобия (Krebs, 1989).

Как и Фишер с соавторами, которые попросту отбросили нулевой член лог-серии, Престон столкнулся с проблемой редких видов. Если данные, которым соответствует гауссиана, получены из ограниченной выборки, то левая часть кривой (т. е. редкие и потенциально неучтенные виды) будет выражена нечетко. Престон назвал такую точку усечения кривой слева «*линией занавеса*». Эта линия может сдвигаться влево при увеличении объема выборки, и, как считал Р. Мак-Артур (Robert Helmer MacArthur; 1930-1972), число скрытых (или «невидимых») видов, которые легко вычислить по уравнению гауссианы, является вполне приемлемой мерой видового разнообразия.

Все авторы, разработавшие индуктивные модели, использовали их лишь как удобный способ описания экспериментальных данных, не считая необходимым обосновать теоретически экосистемный смысл их применения (и, возможно, даже не проверяя адекватность выбранного закона с помощью статистических критериев). Но их работы вдохновили сторонников теоретических «дедуктивных» подходов, основанных на интуитивных гипотезах о том, как должны быть организованы экологические сообщества.

В 1957 г. Р. Мак-Артур опубликовал статью (MacArthur, 1957), описывающую модель «разбиваемой стеклянной палочки»[1], основанную на схеме монотонного изменения обилий N_i в ранжированном ряду: $N_i = \sqrt{S} \cdot \sum_{l=i}^{S} (1/l)/S$. Теоретически это можно представить как разделение пространства ниши в пределах сообщества по аналогии со случайным и одновременным разламыванием стержня на S кусков.

Тем временем Р. Уиттекер (Whittaker, 1965) заново обнародовал намного более ранний дедуктивный подход в работе японского эколога И. Мотомуры (Motomura, 1932), изданной на японском языке. Мотомура на примере растительных сообществ, для которых характерен высокий уровень доминирования, предложил геометрическое распределение, представленное как модель «выгрузки ниши»:

$$N_i = \frac{Nk(1-k)^{i-1}}{1-(1-k)^S}; \qquad N = \sum_{i=1}^{S} N_i, \quad k = \text{параметр ряда.}$$

Модель геометрического ряда концептуально подобна гипотезе «разбитой палочки», за исключением того, что разделение лимитируемых ресурсов идет последовательно.

Наконец, снова возродился интерес к логнормальному распределению, когда Дж. Сугихара (Sugihara, 1980) разъяснил, что эта модель – не просто способ описания относительного обилия видов, но и специальный класс так называемых канонических отношений в сообществах. При этом особое, почти мистическое внимание было уделено сдвигу медианы распределения численности особей относительно правой границы кривой распределения видов (если эти точки совпадают, то такая модель называется канонической). Значимость канонических отношений Сугихара связал с биологическими представлениями о последовательном разделении пространства многомерной ниши между существующими в экосистеме видами (здесь также проводится аналогия с неоднократно и случайно разбиваемой палочкой).

---

[1] Как обратил внимание проф. В.Н. Максимов, устоявшийся перевод *broken stick* как «разломанный стержень» не вполне точен, поскольку не отражает случайный характер процесса.

Подбор и последующая интерпретация моделей распределения численности видов (если серьезно учитывать структурные особенности изучаемых сообществ и анализировать шкалы измерения показателей обилия) представляет собой трудно формализуемый процесс, который сродни искусству хиромантии. Например, для сообщества макрозообентоса р. Сок (по результатам 97 проб, в которых найдено 277 видов и других таксономических единиц) на рис. 5.1 можно усмотреть, что геометрический ряд Мотомуры (прямая 1) почти идеально подходит для описания длинного линейного участка с умеренной численностью видов, тогда как нелинейные участки на концах эмпирической кривой с очень высоким и очень низким обилием оказались ему недоступными. Лог-серии Фишера (кривая 2) хорошо описывают область видов с высоким и средним обилием, но совершенно беспомощны для аппроксимации малочисленного «правого хвоста». Модель «разбитой палочки» Мак-Артура (кривая 3), в принципе, верно угадала теоретическую S-образность наблюдаемых данных, но низкая параметричность не дает возможности использовать эту модель для реальных расчетов.

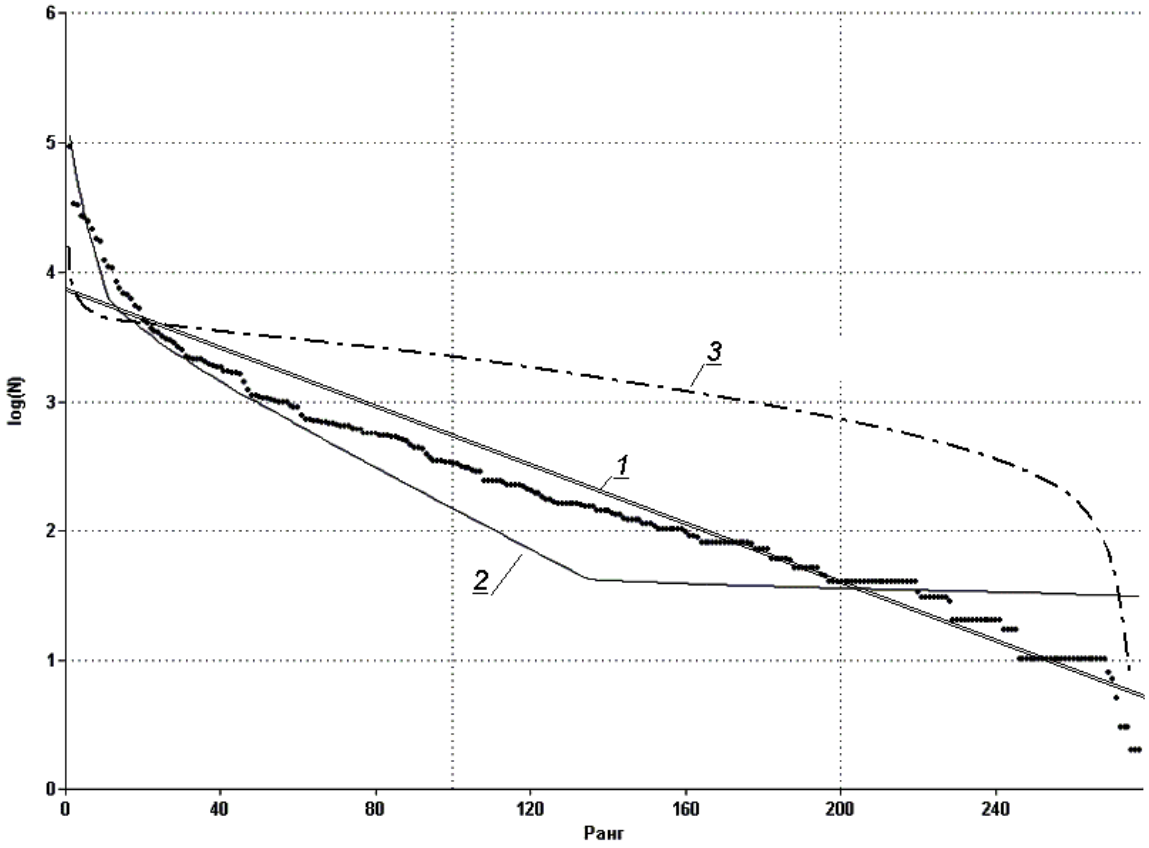


Рис. 5.1. **Ранговые модели численности видов сообщества макрозообентоса р. Сок:** на оси абсцисс – порядковый номер вида в убывающем ранговом ряду; на оси ординат – логарифм его численности; 1 – геометрический ряд Мотомуры ($\chi^2 = 2{,}1 \cdot 10^6$, $p \approx 0$); 2 – лог-серии Фишера ($\chi^2 = 7{,}9 \cdot 10^4$, $p \approx 0$); 3 – модель «разбитой палочки» ($\chi^2 = 1{,}1 \cdot 10^6$, $p \approx 0$)

Можно найти и много других изъянов в представленных попытках разработать общую теорию относительного обилия видов, исходя из основных концепций и законов экологии (см. дискуссионную часть работ: May, 1975; Pielou, 1975; Мэгарран, 1992). Однако историческую роль четырех основных моделей в теоретическом осмыслении экологической структуры сообществ нельзя преуменьшить.

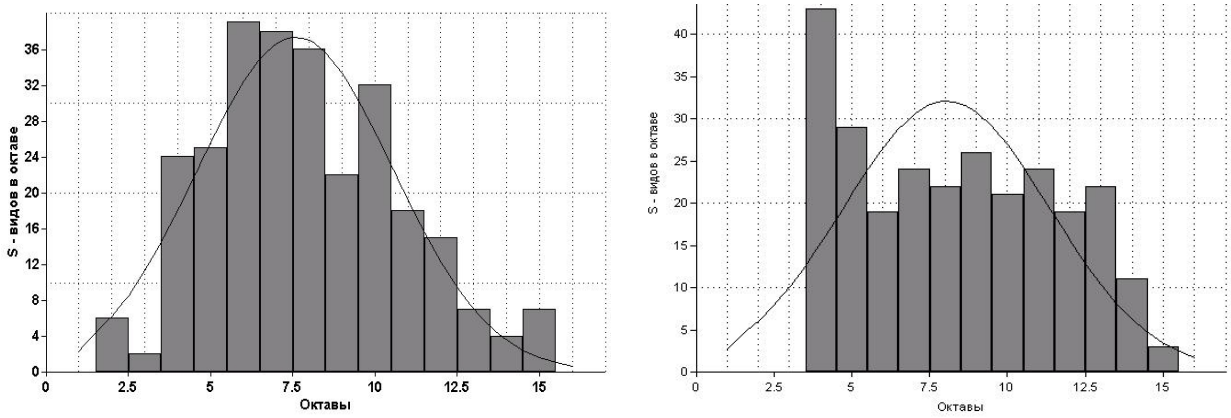
К сожалению, практическое значение этих моделей и всех их усовершенствованных вариантов оказалось весьма ничтожным. «Многие массивы данных одинаково хорошо описываются как лог-рядом, так и логнормальной моделью, и для экологов бывает трудно решить, какое распределение лучше подходит к конкретному случаю» (Мэгарран, 1992, с. 30). Какие конкретные результаты по сравнительной оценке разнообразия можно извлечь из результатов аппроксимаций распределения той или иной кривой, кроме естественного чувства удовлетворения от «модельных игр»? Имеющиеся публикации не дают на это четкого ответа, а лишь смутные предположения их авторов, что тот или иной параметр модели может быть как-то связан с разнообразием.

При отсутствии каких-либо универсальных законов распределения численности сообществ по видам, исследовать и отстаивать одну какую-то конкретную модель достаточно бессмысленно. В общем случае должен использоваться либо коллектив моделей (Брусиловский, 1987; Розенберг и др., 1994) с учетом уровня компетентности каждого из его «членов», либо переборный процесс, когда из всего множества моделей-претендентов выбирается одна, наилучшим образом аппроксимирующая имеющиеся экспериментальные данные.

Отклонение от темы (*о шкалах измерения*). Здесь можно отметить еще одну грань проблемы правильного использования статистических критериев согласия, оценивающих степень расхождения двух кривых: соотношение эмпирики и теории. Например, для проверки гипотезы о случайном характере различий между эмпирическими и теоретических частотами широко используется критерий согласия χ^2 (см. подписи к рис. 5.1). Однако, как и у всех иных статистик, использующих частотные характеристики, значение χ^2 сильно зависит от выбранной шкалы единиц измерения. Например, масштаб шкалы оценки эмпирических численностей различных сообществ составляет: для уссурийских тигров – единицы особей, лесных угодий – десятки стволов деревьев, бабочек – сотни, макрозообентоса – десятки тысяч, сообществ бактерий – сотни миллионов экземпляров. Поэтому можно по разному расценивать высокие значения критерия χ^2: и как объективную реальность плохой адекватности теоретических моделей распределения, и как результат нелинейной деформированности учетной шкалы численности макрозообентоса по сравнению, например, со шкалой учета бабочек. Если, в частности, предварительно прологарифмировать значения численности бентоса, то итог оценки моделей по критерию χ^2 будет совсем иным.

В практике современных гидробиологов, исповедующих балансовый принцип энерго- и массообмена, считается, что единственная цель учета состоит в оценке популяционной плотности на единицу площади или объема (числа особей [1/м^2], [1/м^3] или биомассы [г/м^2], [г/м^3]). Эти показатели не очень пригодны для статистических расчетов. А.И. Кудрин (1998, с. 10) обратил внимание на ограниченность прямолинейного подхода к выбору шкал измерения: «… измеряемая величина удачно выбрана, если она хорошо распределена, легка в учете и имеет отношение к сути дела». Говоря только о сути дела, следует признать, что количество донных организмов, найденных при удачном «скребке» драгой, имеет весьма косвенное отношение к оценке численности популяции на достаточно большой площади.

В нашем случае оба показателя обилия не имеют прямого отношения к «полному числу экземпляров *N*» (в том смысле, в каком понимается, например, численность уссурийского тигра в Амурской области) Но, поскольку нет точной спецификации *N* и для SAD, можно построить модели рангового распределения обилия как для численности, так и для биомассы видов, выразив её размерность, например, в мг/м^2 (рис. 5.2).



а) $\chi^2 = 32{,}9,\ \ p = 0{,}0018$ б) $\chi^2 = 105,\ \ p \approx 0$

Рис. 5.2. **Гистограмма распределения количества видов сообщества макрозообентоса р. Сок по октавам численности (а) и биомассы (б).**
Для сравнения приведены соответствующие кривые логнормального распределения и расчетные значения критерия согласия χ^2

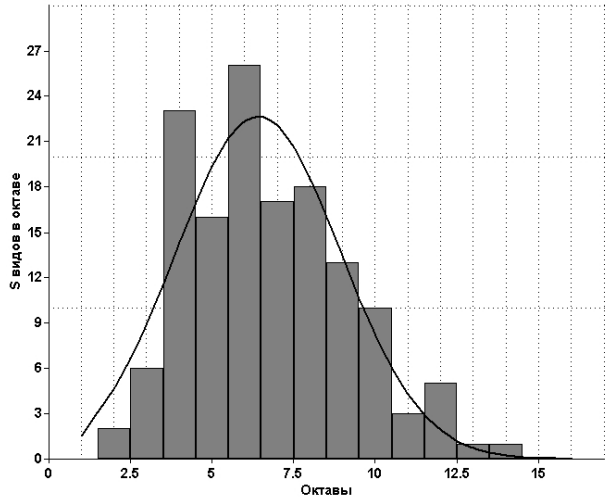
В последние десятилетия, когда оказались доступными для комплексной обработки большие объемы выборок, а экология редких видов стала объектом пристального внимания, отчетливая асимметрия распределения относительного обилия видов в модели Престона стала очевидной. Фактически, эмпирическая кривая хорошо аппроксимируется логнормальным распределением справа от медианы, т. е. в области, представляющей фоновые виды. Но слева от медианы ситуация почти всегда нестационарна: при движении «линии занавеса» влево в результате длительного и интенсивного мониторинга экосистемы до стадии видового насыщения появляется значительно большее число редких и чрезвычайно редких видов, чем это предсказывается теоретическим логнормальным распределением.

Сразу две нейтральных теории претендуют на объяснение и учет отрицательной асимметрии логнормальных распределений, наблюдаемой в эмпирических наборах данных. «*Унифицированная нейтральная теория биоразнообразия*», или UNTB (Hubbel, 2001) представляет новую модель распределения обилия видов в форме многочлена нулевой суммы (zero-sum multinomial), в котором отрицательная асимметрия учитывается как функция размера сообщества и интенсивности миграции. *Модели самоподобия* (Harte et al., 1999) предсказывают, что изменение богатства видов в некоторой области является постоянным для всех пространственных масштабов. Оба подхода принимают предположение об экологической эквивалентности всех видов в сообществе. Упомянутая выше процедура разделения многомерной ниши (Sugihara, 1980) также призвана для объяснения отрицательной асимметрии, но с несколько иных теоретических позиций.

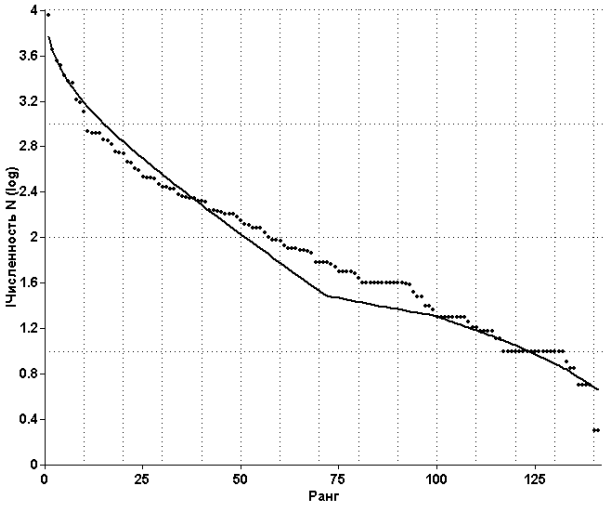
Э. Мэгарран и П. Хендерсон (Magurran, Henderson, 2003) предлагают более «бережливый» учет отрицательной асимметрии распределения эмпирических данных, не привлекая фракталы или концепцию нейтральности, а просто разделив все множество видов на две фракции: *постоянные* компоненты сообществ (или *резидентные виды*) и *случайные* виды (или *мигранты*). Появление новых малочисленных видов интерпретируется авторами как стационарный случайный процесс Пуассона, и на этом основании в качестве частной модели распределения численности подмножества редких видов используются логарифмические ряды Фишера. Оставшееся подмножество резидентных видов на втором этапе аппроксимируется логнормальным *распределением Престона*, и при этом отрицательная асимметрия должна исчезнуть. Работа (Magurran, Henderson, 2003) убедительно демонстрирует эффективность этого подхода на примере 21-летнего мониторинга рыбных сообществ в Бристольском проливе.

Идея «раздельного» учета резидентных и случайных видов распространяется и на аппроксимацию кривой накопления видов. Например, в описании к прекрасной программе **SDR-IV** (Seaby, Henderson, 2007) предлагается линейно-гиперболическая модификация *модели Михаэлиса–Ментен*: $S(a) = \frac{S_{\max} a}{K_1 + a} + K_2 a$, где $S(a)$ – оценка числа видов в зависимости от выборочного усилия a; $S_{\max}$ – общее количество резидентных видов; K_1 – параметр, связанный со скоростью накопления резидентных видов (выборочное усилие, которое соответствует половине общего количества резидентов); K_2 – параметр, который описывает скорость накопления случайных видов.

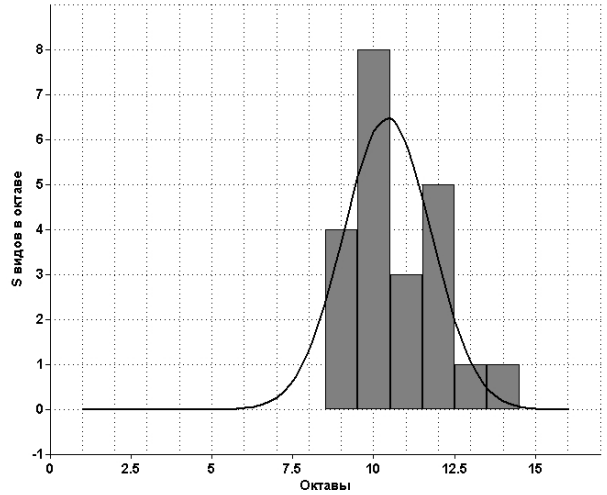
В изученных нами сообществах макрозообентоса отчетливой асимметрии логнормальных распределений численности особей отмечено не было (некоторая асимметрия наблюдалась в распределении биомассы – рис. 5.2б). Например, в р. Байтуган была найдена 141 таксономическая единица, из которых 119 видов в соответствии с *индексом доминирования Палия–Ковнацки* ($D < 0{,}1$) были отнесена к редким видам, а остальные 22 вида – к резидентным. При раздельном моделировании этих двух подмножеств видов *моделями Фишера* и *Престона* (рис. 5.3а–г) никакого позитивного результата не было получено.



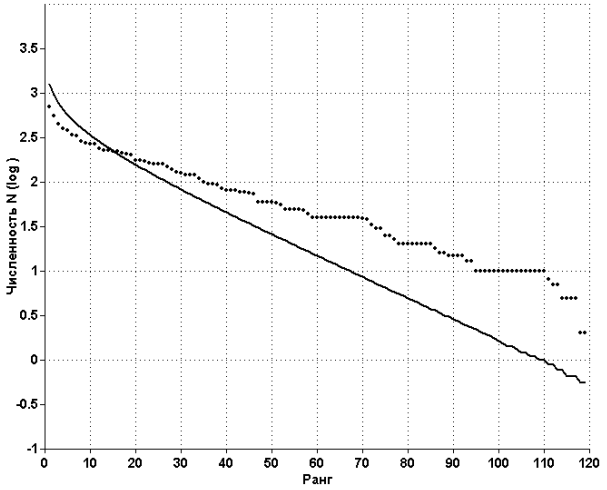
а) $\chi^2 = 2.3, p = 0.21$



б) $\chi^2 = 3727, \; p \approx 0$



в) $\chi^2 = 15{,}6, \; p = 0.076$



г) $\chi^2 = 4420, p \approx 0$

Рис. 5.3. **Модели распределения численности сообщества макрозообентоса р. Байтуган:**
а) и б) логнормальное распределение и лог-серии Фишера, соответственно, для полного списка видов; в) логнормальное распределение численности резидентных видов;
г) лог-серии Фишера для редких видов

И визуально, и статистически по критерию χ^2 обе кривые на усеченных списках видов явно проигрывают тем же моделям на полных множествах. Это лишний раз свидетельствует о том, что не существует общих рецептов моделирования «на все случаи жизни», не учитывающих экологические и статистические особенности изучаемых сообществ.

Третья «знаковая» кривая распределения численности особей по видам – *кривая Лоренца* – представляет собой график, на котором по оси абсцисс отложены накопленные относительные численности особей видов, а по оси ординат – соответствующие им доли числа видов. Пусть экологическое сообщество состоит из N особей, относящихся к S видам. Тогда ранжированный по убыванию вектор относительного обилия имеет вид $\mathbf{p} = [p_1, \ldots, p_s]^T$, $p_i = N_i/N$ ($i = 1, \ldots, S$), где N_i – обилие вида i, $p_i \geq p_{i+1}$. Классическая кривая Лоренца строится как полигональный путь, соединяющий последовательные точки $\pi_0 = (0,0)$, $\pi_1 = (p_1^*, 1/S)$, $\pi_2 = (p_1^* + p_2^*, 2/S)$, … , $\pi_S = (p_1^* + p_2^* + \ldots + p_S^*, S/S) \equiv (1, 1)$ (Ricotta, Avena, 2002; Ricotta, 2003). В прямоугольной системе координат она располагается выпуклостью вниз и проходит под диагональю прямоугольника с координатами (0; 0); (0; 1); (S; 1) и (S; 0). Чем больше отклонение эмпирической кривой от диагонали, которая совпадает с линией равномерного распределения долей, тем больше степень неравномерности обилия видов. Мере этого отклонения соответствует *индекс Джини* (Тейл, 1971; Weiner, Solbrig, 1984), равный половине площади фигуры между кривой Лоренца и главной диагональю:

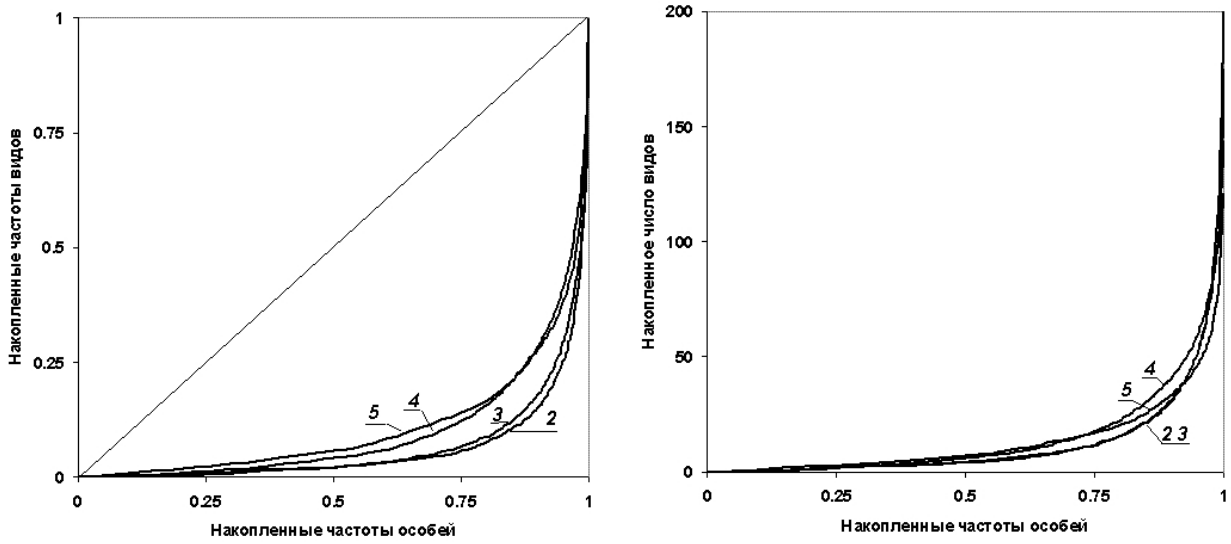
$$G = 1 - 2\int_0^1 L dx \quad \cong \quad \frac{1}{2S-1}\sum_{i=1}^{S}\sum_{j=1}^{S} | p_i - p_j | \cdot$$

Чем выше индекс Джини, тем выше неравномерность распределения обилия S таксонов, и, следовательно, при прочих равных условиях, выше гетерогенность сообщества.

Кривая Лоренца не несет в явном виде информации о втором компоненте разнообразия – видовом богатстве, поэтому был предложен её несколько модифицированный вариант под названием «внутренний профиль разнообразия» (intrinsic diversity profile), который отличается тем, что вместо накопленных долей видов используется накопленное число видов (Patil, Taillie, 1979). Этот профиль призван дать представление сразу и о выравненности, и о видовом богатстве, однако лишившись возможности оценивать неравномерность по индексу Джини, он мало что приобрел в «изобразительной убедительности» (см. рис. 5.4а, б). Предложение изменить порядок ранжирования видов по обилию на обратный, в результате чего кривые будут располагаться над главной диагональю выпуклостью вверх, мало что добавляет к существу дела.

Другая модификация кривых Лоренца – «внутренний профиль стабильности» (intrinsic stability profile; Rousseau et al., 1999) – отличается от кривой Лоренца тем, что значения по оси абсцисс нормированы на среднюю плотность. В конструкции этого профиля неявным образом заложено довольно спорное предположение о том, что суммарная (или средняя) плотность особей сообщества отражает видовое богатство.

Однако на основе профилей разнообразия разработан (Patil, Taille, 1979, 1982) достаточно эффективный и унифицированный подход к оценке разнообразия на основе интуитивного предположения, что разнообразие сообщества определяется в первую очередь количеством редких и малочисленных видов. Предлагается измерить «редкость» каждого вида и принять за показатель разнообразия среднюю редкость видов сообщества. Ими введена в качестве такого показателя сумма относительного обилия для *семейства видов правой части* внутреннего профиля разнообразия, т. е. для самых редких видов, накопленные частоты которых правее значения k (см. рис. 5.4б). Фактически, если профиль разнообразия C_2 всюду выше, чем C_1, то сообщество C_2 имеет большее разнообразие. С другой стороны, если два профиля пересекаются между собой один или более раз (как это имеет место в нашем случае), то ординация этих двух сообществ становится невозможной.



а) б)

Рис. 5.4. **Кривые Лоренца (а) и внутренние профили разнообразия (б) для сообществ макрозообентоса некоторых малых рек Самарской области:**
2 – верхнее течение р. Сок (индекс Джини G = 0,885); 3 – нижнее течение р. Сок (G = 0,871); 4 – р. Байтуган (G = 0,802); 5 – р. Б. Кинель (G = 0,786).

В этом ключе Р. Руссо с соавторами (Rousseau et al., 1999) доказывают, что частный порядок, полученный из внутренних профилей разнообразия, наилучшим образом соответствует балансу между числом видов и выравненностью обилия. Так как любой профиль разнообразия единственным образом определен функцией $\varphi(N)$ неизвестного распределения по видам множества особей N изучаемого сообщества, то, в конечном итоге, все сводится в очередной раз к статистической задаче восстановления функции φ по эмпирическим выборкам о наблюдаемом видовом обилии. Подробнее эта задача и принципы её решения с использованием процедур «складного ножа» (jackknife) рассматриваются в работах итальянских биологов (Barabesi, Fattorini, 1998; Fattorini, Marcheselli, 2002), однако нам не удалось найти публикации, где приводились бы конкретные результаты исследований в этом направлении.

5.2. Модель Ципфа–Мандельброта и *H*-распределения видов

Многочисленными социологическими и экологическими исследованиями установлено, что все устойчивые плотности распределений (кроме гауссовой) убывают при больших значениях аргумента *приблизительно как гиперболы*, имеющие длинные правосторонние «хвосты», что и составляет формальную сущность ценологического подхода (Кудрин, 2007). А.И. Яблонский констатирует (1986, с. 76), что *закон Ципфа–Парето*, описывающий гиперболический характер распределений, «играет в соответствующих областях (связанных, как правило, с информационными, биологическими, социально-экономическими системами) практически ту же универсальную роль, что и закон Гаусса в стохастических процессах с конечной дисперсией, связанных обычно с естественными процессами в неорганических, физико-химических системах».

Пусть имеется некоторая, хотя бы относительно замкнутая, система элементов N (в нашем случае – экологическое сообщество) и множество $S = \{x\}$ различных типов этих элементов (т. е. список видов, а в других предметных областях – каталог конст-

рукций, словарь словоформ и т. д.). С помощью рубрикатора S все объекты сообщества разбиваются на популяции (классы), т. е. каждому виду $x \in S$ соответствует подмножество $N(x)$ или частота всех вхождений этого вида в сообщество N. Перенумеруем элементы списка $S=\{x_1, x_2, ..., x_s\}$ так, чтобы частота встречаемости вида $N(x_i)$ была невозрастающей функцией его номера: $N(x_1) > N(x_2) > \ldots > N(x_S)$. Ранговым распределением (Арапов и др., 1975) называется функция $\Phi(r) = N(x)$, которая ставит в соответствие номеру или рангу $r(x)$ вида x, $x \in S$, оценку вероятности $N(x)/N$ встречаемости этого вида, т. е. Φ: $\mathcal{D} \rightarrow$ R, где $\mathcal{D}$ – отрезок натурального ряда, а R – множество положительных вещественных чисел.

Если каждому элементу сообщества x_i сопоставить некоторую характеристику его значимости $E(x_i)$, то наиболее вероятное состояние ансамбля совпадает с *распределением Гиббса* в статистической термодинамике: $N(x_i)/N = p(x_i) = \mu e^{-\lambda E(x_i)}$, где $p(x_i)$ – вероятность появления в данном сообществе вида x_i, μ и λ – постоянные, характеризующие данный ансамбль. Здесь переменной $E(x_i)$, являющейся аналогом энергии элемента x_i в термодинамике, разные авторы придают смысл сложности (Мандельброт, 1957; Шрейдер, 1967) или характеристики важности вида (Уиттекер, 1980). А.П. Левич (1978) трактует параметр λ как меру эволюционного совершенства системы, а $E(x_i)$ – как объем ресурса, доступный одной особи i-го вида, определяя тем самым объем всех используемых ресурсов $\sum \lambda E(x_i)$ как функцию общей энергии системы, являющуюся инвариантом её разнообразия.

В работах Б.Мандельброта (1957) и Ю.А. Шрейдера (1967) показано, что вполне естественные соображения позволяют перейти от формулы Гиббса к тому или иному закону рангового распределения. Очевидно, что при термостатистическом подходе к восприятию ансамбля видов функция рангового распределения $\Phi(r)$ несет определенную информацию о внутренней структуре сообщества, а именно о распределении ресурсов и своего рода «статусе» его членов. В зависимости от типа отношения к ресурсу, реализуемы различные модели распределений, аргументом которых являются ранги каждого вида i, а μ и β – постоянные параметры (Левич, 1978, 1980):

- *ранговое распределение Гиббса* в термостатике или *Мотомуры* в экологии $p(x_i) = \mu e^{-\beta i}$ при линейной зависимости состояния системы от ресурса;
- *ранговое распределение Ципфа* $p(x_i) = \mu e^{-\beta \log(i)} = \mu i^{-\beta}$ при логарифмической зависимости от ресурса;
- *распределение «разбитой палочки» Мак-Артура* $p(x_i) = \mu e^{-\beta \log(\log(i))} = \mu \log(i^{-\beta})$ при дважды логарифмической зависимости от ресурса;
- *распределение Ципфа–Мандельброта* $p(x_i) = \mu e^{-\beta \log(\alpha + i)} = \mu i^{-(\alpha+\beta)}$, где α – число «пустых» состояний с неиспользуемым ресурсом.

Укоренившаяся в экологии практика диагностики состояния экосистем по индексам видового разнообразия (энтропия Шеннона, показатель Симпсона и др.) имеет свое оправдание в том, что все эти индексы однозначно связаны с параметрами ранговых распределений μ и β (Левич, 1980). Поэтому вполне оправдана идея интерпретировать сами эти параметры непосредственно как обобщенные показатели экологического разнообразия.

Метод оценки экологического состояния природных объектов с помощью параметров ранговых распределений открывает путь к региональному нормированию воздействий, нарушающих экологическое благополучие. Показано (Левич, 1980), что в нормальном (т. е. ненарушенном, фоновом и т. п.) состоянии сообщества параметр рангового распределения варьируется во вполне определенном диапазоне значений. Например, для численностей планктонных организмов в случае гиперболической модели Ципфа $\beta \approx 1{,}5 – 2{,}5$.

В общем случае параметр рангового распределения специфичен для типа сообщества (например, для сообществ фитопланктона, зоопланктона или перифитона) и зависит от сложившегося комплекса условий среды, в которых адаптировано сообщество конкретной экосистемы. Н.Г. Булгаков с соавторами (2005) на весьма представительном множестве данных Роскомгидромета по биологическому мониторингу сообществ фитопланктона (от р. Ангара до Азовского гидрографического района) подробно рассмотрели закономерности изменения параметра β от региональных особенностей, объема выборок, времени и сезонности взятия проб. С использованием *детерминационного анализа* была выполнена оценка сопряженности между качественными градациями значений параметров ранговых распределений и шкалой качества вод с учетом пространственно-временного фактора. Однако значительный разброс полученных значений β на столь разнородном материале не позволил авторам однозначно выделить долю изменчивости характеристик распределений, которая бы объяснялась влиянием качества среды. Тем не менее, при сравнении сообществ, где региональные и сезонные факторы необходимым образом учтены при формировании выборок, отклонения от «нормативных» значений параметров распределений могут служить мерой патологии экосистем (Левич, 1980). Другими словами, в той степени, в какой справедлив выбранный закон рангового распределения, его параметры служат индикаторами стрессовых воздействий на сообщество и оценкой его стабильности. Ранговые распределения применялись для анализа процессов эвтрофирования вод, оценки влияния на гидробионтов загрязняющих веществ и теплового загрязнения, изучения сукцессий, сезонных изменений и многого другого (Inagaki, Lenoir, 1974; Левич, 1980; Lecordier, Lavielle, 1982; Сахаров, Ильяш, 1982; Максимов и др., 1997).

Общие выводы об адекватности различных моделей распределений таковы:

- при существующей точности экспериментальных экологических данных для их описания, как правило, достаточен выбор любой одно- или двухпараметрической модели;
- невысокая точность данных не позволяет отдавать предпочтение одной из модели;
- применение многопараметрических моделей для таких данных нецелесообразно.

Огромный опыт, основанный на моделировании реальных данных, показывает, что чаще всего ранговые распределения видов в сообществе описываются гиперболическим распределением Ципфа (Zipf, 1949). Формально исторический приоритет Д. Ципфа в этой области не однозначен, если вспомнить модель «оценки неравенства доходов», которая восходит к пионерским работам В. Парето (Pareto, 1897), закон «обратных квадратов» А. Лотки (Lotka, 1926), теорию устойчивых распределений П. Леви (1924; цит. по: Феллер, 2010), работы Р. Фишера, предшествовавшие описанию логарифмических рядов (Fisher et al, 1943), соображения Б. Мандельброта (1957) о смысле фрактальных размерностей и др. В связи с этим в разных источниках выводимые соотношения связывают с разными парными комбинациями имен Ципфа–Парето–Леви–Мандельброта. С использованием формализма гиперболических моделей, получивших название H-распределений, были получены прозрачные и элегантные доказательства многих фундаментальных теоретических законов экологии (Левич, 1978; Рябко и др., 1978; Абросов, Боголюбов, 1988; Кудрин и др., 1989; Левич и др., 1994; Ю. Пузаченко, А. Пузаченко, 1996), таких как:

- эмпирическое соотношение, известное как *формула Престона* зависимости числа видов от площади (или, при более общем подходе, от объема доступных ресурсов, «мощности» среды, эволюционного времени развития экосистемы и т. д.);
- взаимосвязь устойчивости развития сообщества с её энтропийными показателями, частотным спектром и пропускной способностью «каналов связи» с окружающей средой;
- зависимость между напряженностью межвидовой конкуренции и эволюцией морфологических признаков особей и т. д.

Рассмотрим обобщенную математическую постановку, называемую *первым* и *вторым законами Ципфа* (Кудрин, 1996, 2002, 2007). Пусть мы имеем сообщество из N осо-

бей, каждая из которых относится к какому-либо виду из списка **S**, объединяющего найденные экземпляры в группы (популяции). Наибольшей по численности популяции, вероятность которой $\lambda(1)$, присваивается первый ранг $r_1 = 1$, и далее по убывающей. Распределение $\Lambda(r)$ выстраивает все популяции, число которых равно числу видов S.

Первый закон Ципфа утверждает, что произведение вероятности $\lambda(r)$ обнаружения особи x и ранга r её частоты при некотором фиксированном показателе степени β дает константу b, а сам закон имеет вид $b = r\lambda(r)$ при $\beta = 1$. Иными словами, в первом законе Ципфа речь идет о вероятностях

$$\lambda(r) = \frac{b}{r^{\beta}}, \qquad (5.1)$$

где $\lambda(r) = N(r)/N$, а β и b – параметры модели рангового распределения.

Второй закон Ципфа («количество–частота») объединяет число элементов биоценоза с размерностью списка его видов: $N = \gamma S \ln S$, где γ – характеристический коэффициент.

Функция (5.1) получила название *ранговидового гиперболического H-распределения*. Если представить на оси ординат количество особей каждого вида, а по оси абсцисс – ранги видов, то полученный график имеет вид гиперболы, значение параметров которой различно для разных сообществ. В логарифмических координатах распределение Ципфа имеет вид прямой, наклон которой к осям координат определяется величиной β, равной тангенсу угла $\ln r / \ln[N(r)/N]$ между этой прямой и осью абсцисс, т. е. этот угол уменьшается с уменьшением β. Чем меньше β, тем длиннее хвост данного распределения и тем более оно отличается от нормального распределения и ему подобных гауссиан. Значительное увеличение β говорит о том, что в сообществе есть один или несколько доминирующих видов и видовое разнообразие этих сообществ минимально.

Остановимся на вычислительных аспектах и точности восстановления распределения Ципфа (5.1) по эмпирическим данным, когда задана выборка и необходимо оценить параметры модели β и b. Представления об этих процедурах выглядят неоднозначными, как и туманным представляется тезис о необходимости «хорошего приближения теоретической модели к результатам опытных наблюдений».

Если необходимо в классе функций $\mathbf{F}(\mathbf{x}, \mathbf{b})$ отыскать наилучшую функцию $\mathbf{F}(\mathbf{x}, \mathbf{b^*})$, наиболее близкую к результатам измерений $\mathbf{y} - \mathbf{x}$, то обычно выполняют минимизацию *функционала среднего риска* (Вапник, 1979):

$$\mathbf{I}(\mathbf{b}) = Q(\mathbf{x}, \mathbf{b^*})\, P(\mathbf{y}|\, \mathbf{x})\, P(\mathbf{x})\, d\mathbf{y}\, d\mathbf{x},$$

где $Q(\mathbf{x}, \mathbf{b^*})$ – некоторая функция потерь. Таким образом, построение модели распределения сводится к нахождению функций (алгоритмов), которые на выборках фиксированного объема гарантировали бы с заданной надежностью достижение значения риска, близкого к минимальному.

Следует отметить, что, упомянув термин «наилучшая функция», мы неизбежно сталкиваемся с рядом вопросов, например: «Что значит "наилучшее" приближение к восстанавливаемой зависимости?», «Как найти "наилучшую функцию", если искать её в некотором конструктивно заданном множестве (в нашем случае, в классе гипербол)?» и, наконец, «Приведет ли подмена задач к успеху, т. е. гарантирует ли успешная минимизация функционала $\mathbf{I}(\mathbf{b})$ нахождение "наилучшей функции"?»

Многими авторами считается допустимым прологарифмировать выражение (5.1):

$$\ln [N(r)/N] = \ln(b/N) - \beta \ln r \qquad (5.2)$$

и аппроксимировать ранжированный ряд величин относительного обилия видов уравнением прямой в двойных логарифмических координатах. Например, В.Г. Суховольский и Ю.Н. Баранчиков (1998) рассчитали по этой методике и выполнили сравнительный анализ показателя β для 5 семейств бабочек в разных регионах Сибири ($\beta = 0{,}478 \div 4{,}008$) и тропических птиц с разными пищевыми предпочтениями ($\beta = 1{,}25 \div 1{,}91$). В этой рабо-

те особо подчеркивается то обстоятельство, что стандартные ошибки параметров уравнения регрессии $b' = \ln(b/N)$ и β, линейного в логарифмах, можно считать ошибками оценок видового разнообразия, что дало авторам основание проверять статистическую значимость различий этих показателей для сравниваемых сообществ с использованием критерия Стьюдента.

В процедуре линеаризации модели (5.1) есть определенная методологическая неточность: согласно известным положениям регрессионного анализа нелинейная функция может быть приведена к линейной форме *только в том случае*, если ошибка обоих уравнений регрессии остается *аддитивной*, т. е. в обоих случаях зависимая переменная является суммой своего математического ожидания и ошибки. В нашем случае, при обратном преобразовании линейного в логарифмах уравнения (5.2) в уравнение (5.1) будет имеет место *мультипликативное* вхождение ошибок ε_i в нелинейное уравнение гиперболы. Иными словами, выражения (5.1) и (5.2) представляют собой два *принципиально разных* уравнения регрессии, поэтому интерпретация МНК-оценок (b', β) линеаризированного уравнения в качестве параметров распределения Ципфа неизбежно приведет к заведомо искаженным результатам. Вопрос заключается лишь в том, насколько велики эти искажения? Их вполне можно было бы избежать, используя, например, в пакете Excel вместо функции ЛИНЕЙН( ) процедуру нелинейной оптимизации ПОИСК РЕШЕНИЯ.

Другое возражение высказано В.Н. Максимовым (2004, с. 355): «… если логарифмирование численностей и может иметь какой-то биологический или хотя бы статистический смысл, то логарифмирование рангов неоправданно ни с точки зрения экологии, ни с точки зрения статистики». Нам не кажется это серьезным препятствием, если принять вполне разумное предположение об инвариантности двух шкал: порядковой шкалы рангов 1, 2, …, r_i и метрической шкалы натурального ряда 1, 2, …, x_i, такой, что $r_i = [x_i]$. Со значениями x_i уже можно оперировать как с обычными числами (складывать, умножать, логарифмировать и проч.), называя полученные значения «оценками статистик вариационного ряда, измеренного в порядковой шкале».

Модель распределения Ципфа (5.1) представлена уравнением нелинейной регрессии с двумя параметрами β и b. Во многих литературных источниках свободный член b принимается равным $N(1)$, т. е. численности особей наиболее часто встречающегося вида. Например, была показана связь уравнения (5.1) с выражением $\sum_{r=1}^{S} N(r) = N(1)\sum_{r=1}^{S} 1 \cdot r^{\beta} = 1$, где величина $\sum_{r=1}^{S} 1 \cdot r^{\beta}$ является неполной ζ-*функцией Римана*, что дает возможность перейти к оценке только одного параметра модели β (Суховольский, Баранчиков, 1998). Однако сомнительно, насколько обоснована принципиальная целесообразность перехода от двухпараметрической к однопараметрической модели, имеющей меньше степеней свободы для подбора функции распределения.

И, наконец, третий вопрос возникает в связи с тем, какой критерий минимизации отклонений между теоретическими и эмпирическими частотами следует применять? Для оценки близости двух функций – теоретической $f(\mathbf{x})$ и эмпирической $f^*(\mathbf{x})$ – можно использовать, по крайней мере, две метрики (Вапник, 1979):

- среднеквадратичной близости (L^2), определяемой как
$$\rho_L(f\text{-}f^*) = \{P(\mathbf{x})\, d\mathbf{x}\, [f(\mathbf{x}) - f^*(\mathbf{x})]^2\}^{0,5} \rightarrow \min,$$
- и равномерной близости (C) – $\rho_C(f\text{-}f^*) = \sup | f(\mathbf{x}) - f^*(\mathbf{x}) | \rightarrow \min.$

Заметим, что требование равномерной близости является более сильным, чем среднеквадратичной (из выполнения второго условия следует выполнение первого, в то время как обратное утверждение, вообще говоря, неверно). Можно высказать предположение, что при подборе распределения Ципфа следует минимизировать сумму абсолютных разностей ме-

жду теоретическими и эмпирическими частотами, которая, в отличие от суммы квадратов отклонений, не столь чувствительна к выбросам в области больших численностей особей.

В табл. 5.1 представлены результаты расчета параметра β распределения Ципфа для сообществ макрозообентоса некоторых водотоков Самарской области, полученные по четырем разным алгоритмам. Очевидно, что изменчивость показателей, рассчитанных для одной реки различными способами, существенно превышает вариацию параметров между сообществами разных рек.

Таблица 5.1

Сравнение показателя β распределения Ципфа, вычисленного по различным алгоритмам для сообществ макрозообентоса некоторых малых рек Самарской области

Река	Количество видов	Общая численность	Алгоритм расчета			
			ЛГЛ_2	МНК_1	МНК_2	МАР_2
Сок (выше с. Исаклы)	205	257796	2,000	0,948	0,973	1,214
Сок (ниже с. Исаклы)	176	232600	2,120	0,898	0,987	1,214
Байтуган	144	50046	1,772	0,917	0,930	1,023
Б. Кинель	121	105640	1,888	0,667	0,750	0,992

Примечание: ЛГЛ_2 – линейное уравнение в двойных логарифмических координатах; МНК_1 – подбор параметра β методом Ньютона–Рафсона с минимизацией суммы квадратов отклонений; МНК_2 – подбор двух параметров *b* и β тем же методом; МАР_2 – то же, но с минимизацией суммы абсолютных разностей.

На рис. 5.5 представлены графики найденных функций, соответствующие тестируемым методам аппроксимации, в двойных логарифмических координатах для участка р. Сок в её верхнем течении, где отчетливо проявляются различия в стратегиях приближения. Методы МНК стремятся минимизировать отклонения в области рангов видов с высокой численностью, чем они диаметрально отличаются от линейного тренда (ЛГЛ) в логарифмических координатах, тогда как использование критерия минимизации разностей (МАР) приводит к промежуточным результатам.

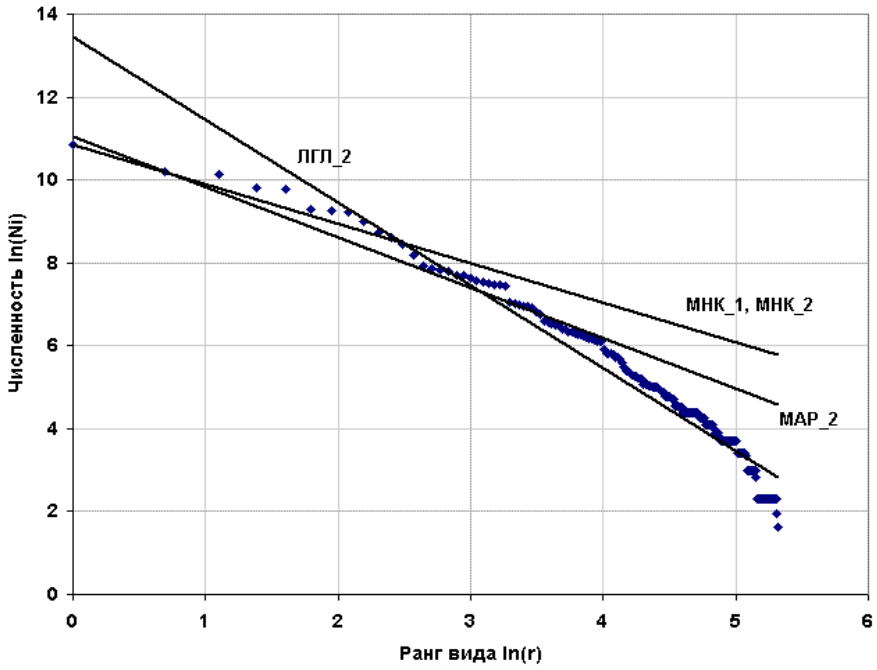


Рис. 5.5. **Подбор параметров ранговидового *H*-распределения эмпирических данных численности видов макрозообентоса р. Сок (верховья) с использованием различных алгоритмов:**

обозначения приведены в примечании к табл. 5.1

Мы не ставим своей задачей настоятельно рекомендовать к всеобщему применению какой-то конкретный метод аппроксимации (несмотря на отмеченные предпочтения), однако убеждены, что прежде чем устанавливать параметры распределения Ципфа в качестве некоторого «стандарта» видового разнообразия, было бы целесообразно определиться с конкретным способом их расчета.

Следующим этапом развития методов аппроксимации негауссовых (ранговых) распределений явилась «свертка» ранговидового распределения по степени повторяемости отдельных видов и моделирование частотно-видовых спектров численностей объектов. Техническим аналогом этого процесса является фракционирование смеси разнородных частиц разного диаметра путем просеивания через последовательность сит с уменьшающимся размером ячеек. Полученные модели были названы Б.И. Кудриным (2002, 2007) *видовым гиперболическим H-распределением*, которое позиционируется им как более общая форма законов Ципфа, Парето, Уиллиса, Фишера, Лотки, Бенфорда, Брэдфорда и др.

Для формирования видового распределения нужно выбрать вначале все виды, встретившиеся по одному разу (т. е. популяции, состоящие из одной особи $a_i = 1$), затем – все виды, представленные двумя особями, тремя и т. д. Тогда полученная группировка популяций в биоценозе может быть наложена на натуральный ряд чисел $k = 1, 2, \ldots,$ занимая, фактически, в нем лишь эмпирически найденные значения. Первая (как правило, самая многочисленная) фракция видов, наименьших по численности, составляет, по терминологии Б.И. Кудрина и его последователей, «*ноеву касту*». Количество видов в этой касте – w_1, а эмпирическая численность особей – $a_1 w_1$. Если все последующие зна́чимые фракции нумеровать по порядку, то общее число каст будет равно K, а w_K соответствует самой мощной (как правило, единственной) популяции доминирующего вида – «*саранчовой касте*», представленной наибольшим количеством особей. Тогда можно записать очевидные соотношения для списка всех встретившихся видов S и общего числа N встретившихся (идентифицируемых) экземпляров в исследуемом сообществе:

$$S = \sum_{k=1}^{K} w_i \; ; \qquad N = \sum_{i=1}^{K} a_i w_i \; .$$

Если $x \in [1,\infty]$ – непрерывный аналог порядкового номера i касты (фракции), $i = [x]$, то

$$\Omega(x) \; = \frac{W_0}{x^{1+\alpha}} , \tag{5.3}$$

где $\alpha > 0$ – характеристический показатель; $\gamma = 1 + \alpha$ и $W_0 = AS$ – постоянные распределения. Теоретические (не обязательно дискретные) значения W_0 или A можно найти из условий нормировки, хотя это может выглядеть ошибочным из-за отсутствия математического ожидания и бесконечности дисперсии. Часто из практических соображений принимают $[W_0] = W_1$, где W_1 – фактическое (экспериментальное) значение первой точки, соответствующее количеству видов ноевой касты.

Таким образом, модель видового распределения (5.3) объединяет популяции, представленные одинаковым количеством особей, в группу k ($k = 1, 2, \ldots, K$), называемую кастой, и утверждает, что между значениями $i = 1, 2, \ldots$ и частотой появления касты есть зависимость, отражаемая гиперболической кривой, параметры которой могут быть оценены для всех без исключения биоценозов. Нумерация каст в видовом распределении имеет физический смысл: это своеобразное ранжирование (по порядку без про́пусков) экспериментально полученных результатов наблюдений, т. е. в данном случае естественная классификация. Экспериментальные ошибки в определении численности редких видов обычно перемещают тот или иной вид в соседнюю касту (также популяционно малочисленную); ошибки в определении числа особей для многочисленных видов, как правило, даже не меняют номера касты.

Обе формы H-распределения – ранговидовое $\Lambda(r)$ и видовое $\Omega(x)$ – по определению являются взаимозаменяемыми: ранговое распределение «свертывается» в видовое, образуя обычно более короткую запись, и обратно. Тот факт, что ранговое распределение является промежуточным при построении видового, позволяет рассматривать только видовые распределения, имея в виду идентичность моделирования и для ранговых распределений. Предпочтение, отданное видовому распределению (5.3), объясняется неочевидностью того предположения, что ноева каста должна быть наиболее многочисленной. В этом случае мы не делаем насилия над фактическими данными, выделяя уникальные единичные виды, затем – встреченные дважды и т. д. Нет никаких оснований до опыта утверждать, что при аппроксимации данных должна образоваться гладкая гипербола.

Б.И. Кудрин (2002, 2007) обратил внимание, что гипербола $\Omega(x)$ делится точкой R на две ветви: слева $i = 1, 2, ..., R$ – неоднородные касты, состоящие из множества видов; справа $i = R + 1, R + 2, ..., K$ – однородные касты, в каждой из которых теоретически ровно один вид (здесь i связана с возрастающим числом особей каждого вида). Эта важная точка перегиба, статистически связанная с числом каст K, была названа *пойнтер-точкой*. Впоследствии было показано, что с ценотической точки зрения виды, группирующиеся вокруг $i = R$, есть виды-индикаторы, в наибольшей мере определяющие специфику сообщества. Математическая точка зрения формулируется кратко: *структура ценозов описывается числом каст K и пойнтер-точкой R.* Видовое H-распределение, выстраиваемое до точки R, имеет всплески и провалы (зубцы), которые не случайны, а закономерно обязательны вследствие нестационарности комплекса редких видов. Учитывая, что в модели принимается $W_0 = R^{1+\alpha}$, то аналитическое выражение для нахождения пойнтер-точки R будет иметь вид (Кудрин и др., 1989; Фуфаев, 1996):

$$\Omega(x) = \frac{W_o}{x^{1+\alpha}} = \frac{R^{1+\alpha}}{x^{1+\alpha}} = (R / x)^{1+\alpha} . \qquad (5.4)$$

Было бы ошибочно априори предполагать существование при заданной паре S и N единственных параметров закона видового распределения $\Omega(x)$, которые бы соответствовали некоторой оптимальной гиперболе Ципфа–Кудрина. Экология ценозов показывает, что из одного списка видов можно получить множество значений численности сообществ N и наоборот, а одинаковые по численности сообщества могут иметь различные комбинации доминирующих видов.

Нами была выполнена аппроксимация видового H-распределения по данным о численности макрозообентоса некоторых малых рек Самарской области и произведена оценка его параметров: числа каст K, параметра α и положения пойнтер-точки R. Предварительно шкала численностей была разделена на интервалы (градации), кратные 40 особям, и, чтобы не плодить множество малоинформативных каст, обилие каждого вида принималось равным среднему значению соответствующего интервала. Подбор параметров гипербол α и R (см. рис. 5.6) осуществлялся по формуле (5.4) с использованием процедуры минимизации суммы квадратов отклонений расчетных и эмпирических значений.

Для интерпретации полученных результатов обратимся к предложенной (Кудрин, 1993) классификации двух аномальных состояний: 1) увеличение эффекта концентрации – слишком многочисленная саранчовая каста и 2) увеличение эффекта рассеяния – слишком велика ноева каста. Ориентируясь на эти представления об объясняющей роли параметров видового H-распределения, можно по положению пойнтер-точки R на рис. 5.6 получить следующий сравнительный ряд экосистем рассматриваемых водотоков по убыванию доли доминирующих видов (или увеличению эквитабельности сообщества):

{р. Сок в нижнем течении} → {р. Сок в верхнем течении} → {р. Байтуган} → {р. Б. Кинель}.

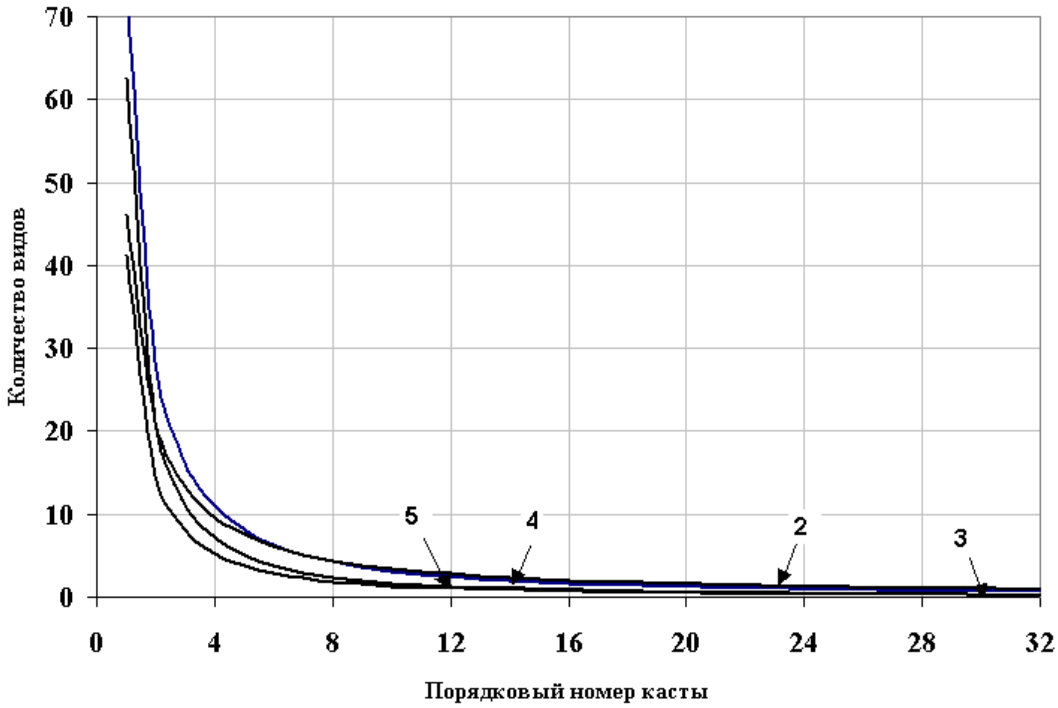


Рис. 5.6. **Графики видового *H*-распределения эмпирических данных по численности макрозообентоса малых рек Самарской области:**
стрелки указывают на расположение пойнтер-точек *R*; 2 – р. Сок в верхнем течении (число каст $K = 46$, видов в ноевой касте $W_1 = 73$, параметр $\alpha = 0{,}73$, $R = 23$); 3 – р. Сок в нижнем течении ($K = 43$, $W_1 = 46$, $\alpha = 0{,}89$, $R = 30$); 4 – р. Байтуган ($K = 27$, $W_1 = 63$, $\alpha = 0{,}64$, $R = 14$); 5 – р. Б. Кинель ($K = 40$, $W_1 = 42$, $\alpha = 0{,}67$, $R = 12$).

Сами по себе утверждения «слишком велико W_1» или «слишком мало $K - R$» без какой-нибудь стандартизации этих значений имеют неконкретный и субъективный характер. Поэтому при разработке механизмов статического сравнения двух сообществ выявилась необходимость построения некоторой универсальной (стандартной) модели видового *H*-распределения, позволяющей сопоставлять разнообразие ценозов различной природы. Первые предложения по моделированию эмпирических *H*-распределений частотными спектрами простых чисел относятся к 1974 г. (Кудрин, 1993). Идея метода заключается в том, что если рассмотреть факториал $S!$, то можно разложить каждый сомножитель ряда на простые сомножители:

$$S! = 2\times3\times(2\times2)\times5\times(3\times2)\times7\times(2\times2\times2)\times(3\times3)\times(5\times2)\times11\times(3\times2\times2)\times13\times\ldots\times S. \qquad (5.5)$$

Предположим, что каждый вид сопоставляется с абстрактным простым числом p_r, где r – порядковый номер простого числа в натуральном ряду: 2, 3, 5, 7, ..., 137, 139, 149, 151, ..., 509, 521, 523, 541, ..., а каждая особь – с появлением этого простого числа как сомножителя в любом из чисел натурального ряда. Тогда распределение численности групп одинаковых сомножителей представления (5.5) может служить канонической моделью распределения видов по повторяемости (т. е. видового *H*-распределения).

Разработка метода генерации канонического *H*-распределения на основе простых чисел (*генератор Кудрина*) позволила начать этап универсального моделирования структуры ценозов любой природы. Б.И. Кудрин (1993) считает, что имело место открытие аналога предельного закона распределения случайных величин для негауссовых генеральных совокупностей. Это дает возможность раскрывать объективные структурные закономерности экосистем посредством анализа чисто математических параметров.

Попытки учета динамики сообществ в виде канонических моделей растущего ценологического объема начались с предложения рассматривать ряд при $N \to \infty$ как динамический. Итоговым завершением данного этапа моделирования явилось обобщение (Кудрин и др., 1989), в котором приведен окончательный канонизированный облик модели в графическом виде и формулы расчета основных характеристик.

В ходе имитационных исследований было показано, что природа H-распределений поддерживает адаптационные механизмы: при сокращении доли малочисленных видов (за счет таксонов-доминантов) общая численность постепенно перераспределяется, причем с большей вероятностью будут заселяться касты, близкие к ноевой. В случае роста доли редких видов (представленных одной особью и близких к ним) процесс идет аналогичным образом в обратном направлении. Это соответствует высказанной гипотезе о стационарности заселения объемов системного пространства (этажей), занимаемых особями различных каст.

При более строгом взгляде на суть таких моделей, они представляют собой не динамику, а некоторый ряд статических распределений, упорядоченный по росту параметров объема моделируемой совокупности. К тому же, далеко не частым явлением в природе являются сообщества с быстрорастущим общим количеством особей ($N \to \infty$). Поэтому такая канонизированная форма представила определенный теоретический интерес (подобно изучаемым предельным теоремам теории вероятностей), но объективно не могла описывать реальные динамические процессы, происходящие в структуре экосистем.

Существование объективной реальности в сообществах любой природы в форме активных процессов рождения, размножения, старения и гибели особей, популяций или видов в целом предопределили разработку динамических моделей нового поколения. В.В. Фуфаев (1996; Фуфаев, Калашников, 2005) обобщил статику H-распределений и предложил структурно-топологическую динамику H-распределения, позволяющую следить за поведением каждого вида во времени и оценить относительное движение точек по поверхности H-распределения. На этой основе был разработан алгоритм и программный комплекс самоорганизующейся имитационной модели гиперболических H-распределений, основанный на разложении факториала натурального числа на простые сомножители. Модель включает настраиваемые алгоритмы, учитывающие два конкурирующих процесса «рождения–размножения» и «старения–гибели» рядов простых чисел, которые, самоорганизуясь в структурно-топологической динамике, дают при постоянном объеме выборки каноническую модель, позволяющую проводить в комплексе все этапы анализа, моделирования и прогнозирования ценотических структур сложных систем.

5.3. Стохастические модели распределения: идеология и идентификация параметров

В двух предыдущих разделах обсуждались детерминированные модели распределения относительного обилия видов. Их конкретное выражение ограничивается константой и одним или двумя фиксированными параметрами. В работах М. Токеши (Tokeshi, 1990, 1993, 1999) рассматривается серия последовательных *моделей пропорционального распределения ресурсов* (sequential apportionment models) и предложен критерий согласия, обеспечивающий подбор параметров, наилучшим образом аппроксимирующих данные полевых наблюдений. В максимально обобщенном виде стохастические модели имеют вид (Tokeshi, 1990; Ulrich, 2001):

$$D_i = D_{\min (i-1)} \rho^z, \qquad (5.6)$$

где D_i – относительное обилие вида i; ρ – случайная переменная с равномерным распределением; z – параметр, определяющий тип модели. Использование случайной переменной ρ обеспечивает максимальное имитирование нестационарности распределения ресурсов

между видами, наблюдаемое в природе (Tokeshi, 1990, 1993). Математическое ожидание такой стохастической модели представляет собой среднее из многих частных моделей, и для каждого ранга обилия вида может быть оценена плотность распределения вероятности (см. рис. 5.7).

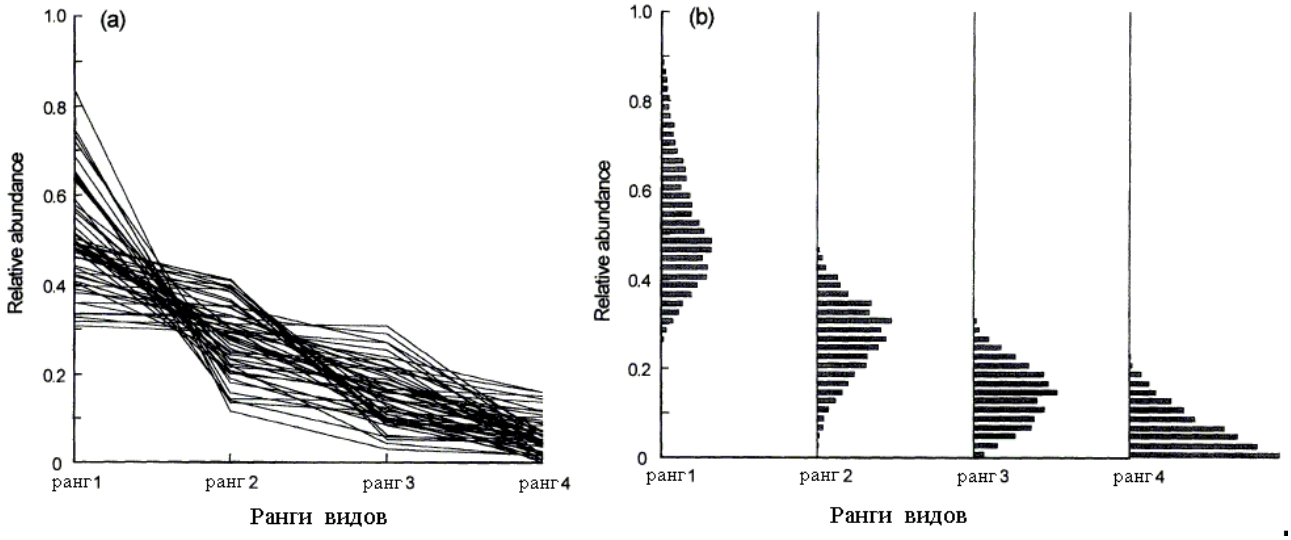


Рис. 5.7. **Пример, показывающий генерацию кривых рангового распределения по стохастической модели (a) и функцию плотности распределения, построенную для каждого ранга (b; по: Bersier, Sugihara, 1997)**

Теоретической основой моделей пропорционального распределения является условие, что «порции» ресурсов экологической ниши, выделяемые видам, пропорциональны их обилию (MacArthur, 1957; Pielou, 1972; Sugihara, 1980; Мэгарран, 1992). Пространство ниши последовательно распределяется каждому новому виду, входящему в сообщество, причем размер этой доли определяется случайным фактором ρ и параметром типа модели z.

Модели отличаются по правилам, используемым для последовательного разделения фракций ниши (breakage rules) – рис. 5.8. В моделях, поддерживающих доминантов (DP), для перераспределения ресурсов всегда выбирается наименьшая фракция, тогда как в моделях, противодействующих доминированию (DD), наоборот, первой делится наибольшая фракция. Если фракция, которая будет распределяться, выбирается наугад (независимо от её размера), то это соответствует модели случайного фракционирования (RF). В модели раздробления Мак-Артура (MF) вероятность распределения фракций пропорциональна их размерам. Модель MF является последовательной версией модели «разбитой стеклянной палочки» (MacArthur, 1957). В модели случайного выбора (RA) приоритет каждой фракции установлен равномерным случайным числом. В сложной модели (CM) сообщество разделяется на виды с высокой численностью особей и редкие виды, для которых установлены разные правила фракционирования.

Для каждого типа моделей используются разные диапазоны значений параметра z в выражении (5.6), который в общем случае может варьировать от $-\infty$ до ∞ (модель power fraction). В частности, для модели MF Мак-Артура $z = 1$, а диапазон значений $0{,}5 \leq z < 1$ соответствует модели обобщенного последовательного распределения Сугихары (Sugihara, 1980).

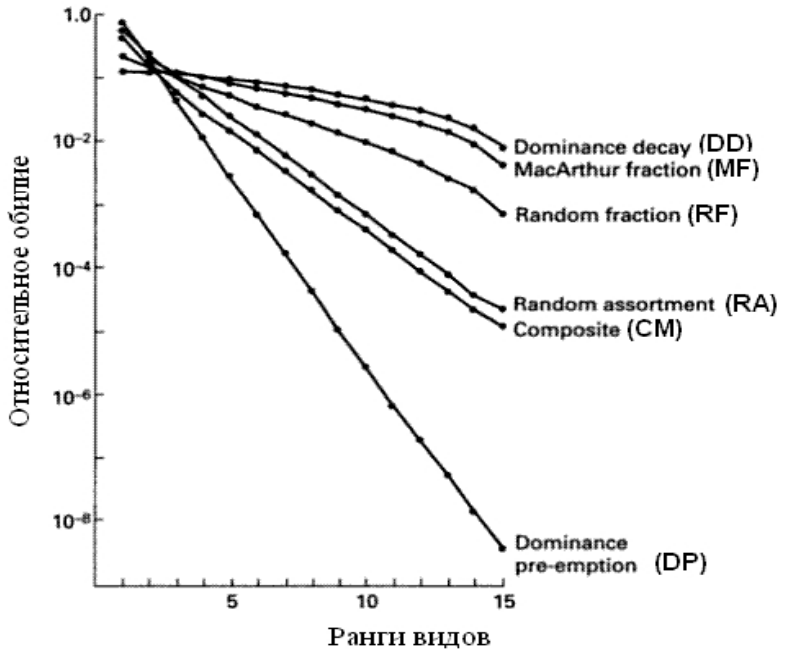


Рис. 5.8. **Кривые «ранг–обилие» для разных типов моделей пропорционального распределения ниши (по: Tokeshi, 1999)**

Обычный способ проверки согласия наблюдаемого и теоретического распределения видов состоит в использовании *критерия* χ^2 или *теста Колмогорова–Смирнова*. Однако, рассматривая распределение видов в каждой имеющейся полевой пробе как эквивалент отдельной выборочной реализации, такое тестирование приобретает множественный характер, а каждый акт сравнения имеет очень плохую дискриминационную мощность (Pielou, 1975), поскольку неоднородность видовой структуры сообщества определяет намного более высокий статистический разброс, чем это учитывается любой специфицированной моделью. Кроме того, никакое стандартное статистическое испытание не имеет смысла для коллекции теоретических распределений, которые произведены стохастическими моделями, рассмотренными выше. Аналитический метод, предложенный М. Токеши (Tokeshi, 1990), сравнивает для каждого ранга (от 1 до числа видов S в сообществе) среднее относительное обилие теоретического распределения с серией выборочных наблюдений. Недостатки такой процедуры привели к разработке алгоритмов статистической оценки, основанных на методе Монте-Карло (Bersier, Sugihara, 1997).

В. Ульрихом (Ulrich, 2001) была разработана расширенная процедура подбора параметров стохастических моделей, заключающаяся в следующем. На первом шаге в качестве минимизируемого статистического критерия суммы наименьших квадратов использовалось евклидово расстояние между точками предсказанных и наблюдаемых данных $r = \sum_{i=1}^{S} d_i^2$, где d_i – логарифмически преобразованная разность между модельной и эмпирической относительными обилиями (т. е. плотностями [densities] i-го вида). Для стохастической модели расстояния r вычисляются с использованием средних плотностей по n сгенерированным повторностям. Процесс настройки параметров моделей z и X (при пошаговой реализации) осуществляется путем их сканирования по достижению минимума r. Поскольку значение r весьма чувствительно к максимальным различиям плотностей $d_{\max}$ и зависит от общего числа видов S, дополнительно вводятся поправочные коэффициенты.

Другим критерием для подбора модели по эмпирическим данным является статистическая величина суммы квадратов отклонений OC_{test} относительного обилия по видам, распределенным по бинарным классам Престона (октавам). Как в случае r_{test}, OC_{test} также зависит от величины максимальных различий в плотностях и общего числа видов S, поэтому для каждого класса стохастических моделей (рис. 5.8) вводятся соответствующие поправочные коэффициенты.

Для практических расчетов рекомендуется выбрать по одной модели из каждого класса и проверить, насколько эти модели соответствуют эмпирическим данным. Анализируя вид модельных кривых, представленных на рис. 5.9, можно сделать следующие относительно бесспорные выводы:

- степень доминирования (т. е. неравномерность распределения обилия по видам) повышается в ряду «верхнее течение р. Сок (2)» → «р. Байтуган (4)» → «нижнее течение р. Сок»;
- это обстоятельство вполне очевидно при рассмотрении моделей случайного выбора (RA), имеющих вид прямой с монотонно увеличивающимся показателем z при снижении эквитабельности видов;
- графики моделей, поддерживающих распределение ресурсов малочисленным видам (DD), увеличивают свою кривизну при росте доминирования, что также может быть использовано для характеристики сообществ.

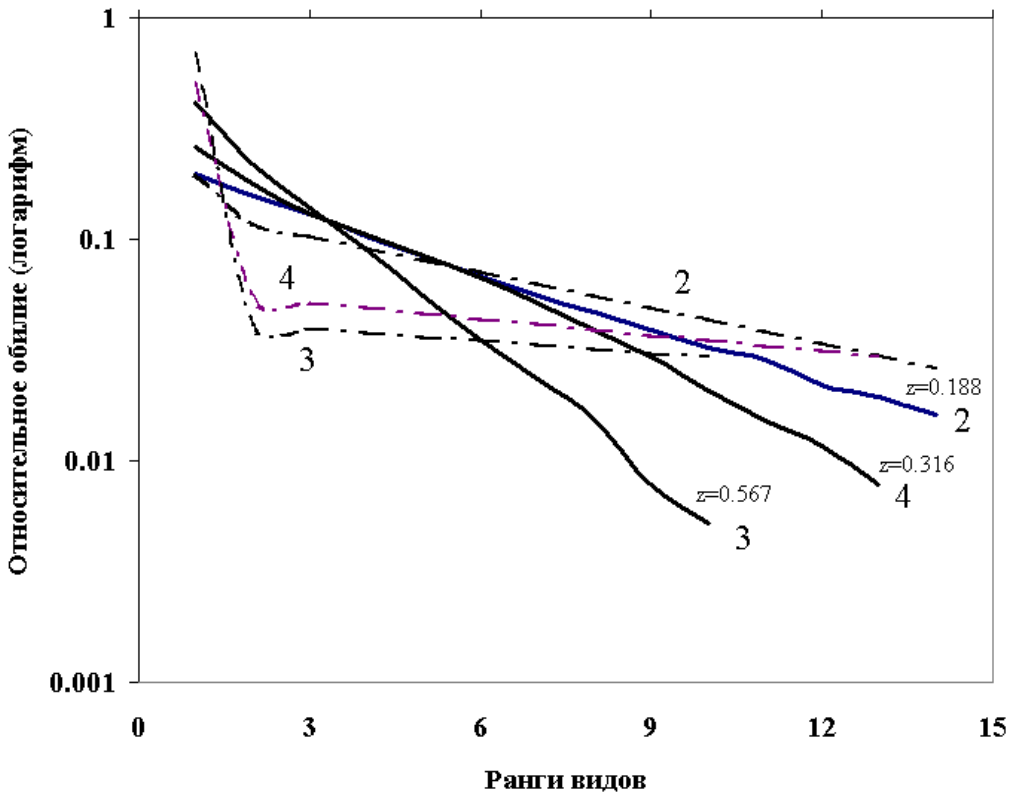


Рис. 5.9. **Графики модели случайного выбора RA (сплошная линия) и модели противодействия доминированию DD (штрих-пунктирная линия) для рассматриваемых сообществ макрозообентоса:**
2 – верхнее течение р. Сок; 3 – нижнее течение р. Сок; 4 – р. Байтуган

Всего 30 лет назад В.Д. Фёдоров (цит. по: 2004, с. 274), внося некоторые математические усовершенствования в модель «разбиваемой палочки» Мак-Артура, писал: «можно надеяться, что новая модель представляет собой шаг вперед на пути отыскания некой гипотетической идеально предсказывающей модели. Должно полагать, что сегодня мы находимся все еще в начале пути». Думается, что работы М. Токеши и его последователей еще в прошлом веке практически подвели черту под поиском математически изощренных моделей «ранг–обилие», перейдя к стохастической их форме. Манипулируя двумя-тремя численными параметрами и экологическими смыслами, можно получить бесконечное множество кривых, вполне «доброкачественно» интерполирующих мониторинговые данные. Однако, несмотря на математическую красоту и теоретическую гармоничность многих моделей распределения обилия, так и осталось неясным, на какую из них следует ориентироваться и как оценивать уровень «идеального предсказания». Другие сомнения состоят в том, в какой мере процесс моделирования предоставляет практическим экологам дополнительный материал для конкретных выводов по сравнению, например, с расчетом любого индекса выравненности сообществ.

Б. Макгилл и его 17 соавторов, авторитетных специалистов в области изучения структуры сообществ из 7 стран (McGill et al., 2007), сложившуюся ситуацию характеризуют как «коллективную неудачу» (collective failure). Выводы этого исследования «могут быть поучительными не только для неудачливых экологов, но и всех, кому приходится сталкиваться с отсутствием прогресса в решении той или иной научной проблемы» (Гиляров, 2008)[2].

Без сомнения, модель распределения «виды–обилие» (SAD) является одним из самых старых и самых универсальных законов экологии – каждое сообщество имеет на гистограмме вогнутую кривую гиперболической формы со многими редкими видами и несколькими базовыми видами. Полная доступность к объединению данных, имеющийся потенциал для сравнения сообществ и визуальная природа в сочетании с внутренней сложностью, сделали SAD очень популярными в экологических исследованиях.

Однако, начиная с 1970-х годов и до настоящего времени, модели, пытающиеся объяснить причины вогнутости SAD-кривой с позиций старых или новых теорий, распространились в экстраординарном количестве. Их далеко не полная классификация включает пять семейств с более чем 40 элементами (см. табл. 2 в работе [McGill et al., 2007]):

1. Статистические модели, которые просто основываются на известных теоретических распределениях (непрерывных гамма- и логнормальном распределениях, или дискретных биномиальных и пуассоновских распределениях). Если нет специальных соображений, рекомендуется использовать простое (неусеченное) логнормальное распределение.
2. Модели, ориентированные на ветвящиеся биологические процессы (branching process), в которых популяции формируются с учетом своей предыстории, а каждая особь имеет своих предков (это предполагает случайное ветвление процессов в модели).
3. Модели динамики популяции: большой спектр от абсолютно детерминированных до полностью стохастических моделей.

---

[2] Одним из возможных выходов из «кризиса по Макгиллу» может стать обращение к фрактальной парадигме устройства природы, заложенной Б. Мандельбротом (2002); см. раздел 3.1 и работы Д.Б. Гелашвили с соавторами (2004, 2007б, 2008а, 2010 и др.). Правда, восприятие фрактальной идеологии требует реальной смены «колодок мышления» (метафоры, предложенной В.Н. Тутубалиным и др. [1999, с. 68], под которыми авторы понимают «модель, не обязательно формализованную, неадекватность которой осознана»; см. также "*Общенаучные размышления*" в разделе 1.1). Иными словами, это означает необходимость приложения определенных усилий для выхода из прокрустова ложа абстрактных стереотипов евклидовых метрик пространства и перехода в мир фракталов, который является следующим приближением отражения реального мира.

4. Модели наложения и разделения ниши (the overlapping niche and the particulate niche model) – рис. 5.8.
5. Пространственные модели SAD, которые разрабатываются, если известно (или можно смоделировать статистически) как пространственное распределение особей каждого вида, так и распределение видов друг относительно друга.

В. Ульрих (Ulrich, 2001) все множество детерминированных и стохастических моделей объединяет в четыре класса по другому принципу – по характеру визуализации получаемых результатов. *Первый класс* моделей приводит к линейному виду зависимости «ранг – логарифм относительного обилия», что характерно для простой показательной функции. К этому классу относятся модели случайного выбора (RA) и поддерживающие доминантов (DP), а также негативное биномиальное распределение (Anscombe, 1950), логсерии и геометрические ряды Мотомуры. *Второй класс* включает модели случайного фракционирования (RF), макартуровских фракций (MF) и «разбитой палочки», а также логнормальное распределение Престона: все они приводят к S-образной функции для логарифма относительного обилия (S-образность кривой может быть в той или иной степени усечена). *Третий класс* содержит модели, основанные на аллометрических функциях, которые приводят к прямой в двойных логарифмических координатах «ранг–обилие». *Последний класс* объединяет модели, вогнутые вниз, но без верхнего изгиба. К ним относятся различные модели перекрывания ниш. Впрочем, считается (без приведения каких-либо доказательств), что классы 3 и 4 очень редко встречаются в природе.

Существует много эмпирических и априорных критериев, позволяющих оценить, какая модель является «лучшей»: учет «экологических» механизмов, приближение к мониторинговым данным и степень предсказуемости, «бережливость» в отношении используемых параметров, степень симметричности распределения обилия видов и т. д. Если неэмпирические аргументы из-за их субъективности являются принципиально неопровержимыми, то качество аппроксимации тестовых данных различными моделями часто составляет микроскопическую разницу. В частности, М. Уилсон с соавторами (Wilson et al., 1998) нашли, что вариация параметров одной и той же модели, полученной на разных подмножествах проб из одного сообщества, значительно превышает вариацию разных моделей на одном комплекте эмпирических данных.

Б. Макгилл и его соавторы видят возможный прогресс в соблюдении совокупности методических установок Д. Платта (Platt, 1964), называемых «строгим выводом» (strong inference). Трудно сказать, насколько можно разделять этот оптимизм авторского коллектива (McGill et al., 2007) и А.М. Гилярова (2008), подробно прокомментировавшего эту важную в концептуальном плане публикацию.

5.4. От распределения ресурсов ниши к концепции «нейтральности» видов в сообществе С. Хаббела

Выше было показано, как в своих попытках понять экологическую сложность сообществ, исследователи в течение всего XX в. старались разработать универсальные модели распределения относительной численности, популяционной плотности и других показателей видового обилия в пространстве и времени (Lotka, 1926; Fisher et al., 1943; Hutchinson, MacArthur, 1959; Preston, 1962, 1981; Pielou, 1975; Taylor et al., 1976; May, 1986). Параметры таких статистических моделей было трудно интерпретировать или приспособить для объяснения процессов динамики сообществ. Стремление построить универсальную концепцию на каком-то этапе уступило место более сфокусированным локальным исследованиям, которые попытались идентифицировать частные механизмы взаимодействия видов в пределах мелкомасштабных экспериментальных единиц (Maurer, McGill, 2004). И лишь некоторые работы, посвященные моделированию флуктуаций плотности

популяций в течение долгого времени (Dennis, Taper, 1994; Lande et al., 2003) или закономерностям пространственного распределения (Maurer, Taper, 2002), косвенно связывали динамику относительной видовой численности сообществ с аспектами биологического разнообразия.

Безусловно, сама структура «дерева» больших экосистем предполагает, что нижний уровень иерархии состоит из моделей, описывающих динамику множества особей и видов в пределах локальных сообществ (на территории ограниченных регионов, в отдельных створах водотока и т. д.). В то же время любая теория должна обслуживать «потребность в понимании и объяснении» того, какие движущие силы определяют биологическое разнообразие в экосистемах больших масштабов, принимая во внимание неоднородность популяционных плотностей в пространстве и времени.

Любая теоретическая попытка объяснить распределение обилия и разнообразия видов в биогеографическом контексте требует, по крайней мере, трех компонентов. *Во-первых*, теория должна учитывать стохастический характер динамики видов от их происхождения до их исчезновения. *Во-вторых*, она должна иметь (явно или неявно) тенденцию к пространственной интерпретации. *В-третьих*, она должна быть адекватно проверяемой, т. е. учитывать все аспекты выборочного процесса и позволять прямое испытание теории на основе эмпирических данных.

Фундаментальный прогресс в установлении прямой связи между динамикой локальных популяций и закономерностями биоразнообразия макросистем был сделан в результате описания, параметризации и эмпирической проверки моделей, основанных на *унифицированной нейтральной теории* (unified neutral theory of biodiversity или UNTB), разработанной С. Хаббелом и его коллегами (Hubbell, 2001, 2003; Volkov et al., 2003). Эта модель использует ту же математическую сущность, что и нейтральная теория генетических алгоритмов, разработанных для конечных популяций (Ewens, 1979), и основывается на ряде простых допущений (Alonso et al., 2006):

Первое предположение UNTB – принцип экологической эквивалентности: *виды могут сосуществовать, если в расчете на одну особь имеют примерно равные вероятности рождения, гибели, миграции и видообразования* (Hubbell, 2001; цит. по: Гиляров, 2007). В дополнение к ранее сформулированным принципам нейтральности (Caswell, 1976, 1988; Bell, 2000, 2001), С. Хаббел подчеркнул, что организмы всех видов обладают *экологической идентичностью* и в расчете на одну особь (т. е. «на душу населения» своей популяции) имеют равную вероятность выжить, размножиться и заселить новое местообитание. Если согласно *принципу конкурентного исключения Гаузе* «каждый вид должен быть ограничен *своей* пищей или *своим* хищником» (Гиляров, 2007), то, согласно UNTB, каждая особь представлена как совершенно *независимая повторность* конкретного сообщества (разумеется, с учетом экологических условий и факторов среды, в пределах которой происходит развитие экосистемы). Это предположение определяет *случайность* комбинаций сосуществующих видов (Lande et al., 2003; Chave, 2004), причем вероятность появления каждого из них определяется только его положением на кривой плотности распределения, которая характеризует относительное обилие видов изучаемого сообщества.

В экологии имеет место некоторое методологическое противоречие: все межвидовые взаимодействия, как очевидно, происходят на индивидуальном уровне отдельных особей, а регулярность, которую мы хотели бы объяснить, является свойством макрообъектов – сообществ или ансамблей видов. Физики успешно разрешили это противоречие, разработав статистические подходы в термодинамике. Нейтральная стохастическая концепция, подобно кинетической теории идеальных газов, также обеспечивает несколько основных идеализированных подходов, чтобы, отталкиваясь от них, можно было далее исследовать отклонения от теории и выдвигать более сложные комплексные предположения. Здесь UNTB работает как нуль-модель структуры сообщества и является хорошей первой ап-

проксимацией, когда эмпирические данные совместимы с нейтральностью (Gotelli, McGill, 2006; Chave et al., 2002).

Второе предположение, крайне важное в теории, – каждое локальное сообщество, *изолированное* в пространстве, является стабилизированным и равновесно насыщенным, т. е. никаких изменений в *общем* количестве особей в течение долгого времени не происходит. Динамикой популяций управляет непрерывная последовательность замены индивидуумов, часть которых в любой фиксированный период времени выводится из состава сообщества (умирает, меняет жизненную форму, мигрирует и пр.), но дефицит быстро восполняется новыми экземплярами. Так как особи всех видов экологически идентичны, то нет никаких преимуществ, которые бы имел любой организм при доступе к освобождаемым ресурсам, кроме априорной вероятности его рождения.

Не только UNTB, но и другие стохастические модели разнообразия, постулируют фиксированный конечный размер сообщества, равный параметру J (число особей, суммарная биомасса и т. д.). Есть неизбежные физические ограничения на общее количество особей, которые могут быть упакованы в данный объем *пространства*[3]. Предположим, что в сообществе существует некоторое конечное число «мест», которые могут быть заняты особями, а численность какого-то вида возрастает: это автоматически означает снижение численности других видов. Если рассматривать широкий диапазон видов (например, деревья гигантской секвойи и лианы; Maurer, McGill, 2004), то предположение о постоянном размере сообщества из этих двух видов было бы обескураживающим – лианы быстро стали бы монокультурой. Поэтому UNTB относится только к сообществам из трофически подобных конкурирующих видов (гильдий) и тогда маловероятно, что плотность всего сообщества сильно изменится от замены одного вида на другой. Другое исключение из принципа насыщенности составляют искусственные или деформированные экосистемы (например, нарушенные в результате антропогенного воздействия).

Таким образом, согласно допущениям UNTB, в моделируемом сообществе учитываются компенсируемые экологические взаимодействия (такие как сотрудничество, соревнование или конкуренция за ограниченный ресурс и т. д.), но при условии, что все особи удовлетворяют условиям одних и тех же правил. Асимметричные процессы, такие как паразитизм и хищничество, были исключены из сферы взаимодействий, поскольку они фактически протекают «между сообществами», т. е. на уровне экосистем, принципы организации которых считались уже совсем другими (Гиляров, 2007).

В-третьих, нейтральная теория тесно связана с теорией выборочного метода. Весьма серьезная методология, основанная на математическом аппарате, используемом в популяционной генетике, рассматривает особенности случайного процесса выборки из сообщества фиксированного числа N особей независимо от их видовой принадлежности и оценивает распределение вероятности нахождения S видов в этой выборке. Поскольку исследование взаимодействий в экосистеме всегда проходит исключительно через выборочный процесс, нейтральная теория Хаббела оформлена так, что имеются строгие отличия между отдельными выборками и всем целостным сообществом (которое как раз и является результатом обобщения этих имеющихся выборок).

Выборочная формула (sampling formula) представляет собой многомерное распределение вероятностей, которое основано на частотах извлечения n_1 особей вида 1, n_2 особей вида 2 и т. д., в соответствии с конкретным характером выборочного процесса. Выборочные *формулы Эвенса*, *Этьена* и т. д. (Ewens, 1972; Etienne, 2005) – это функции

[3] Здесь «пространство» понимается не обязательно в буквальном смысле; это – просто полезная переменная, олицетворяющая любой лимитирующий ресурс, который распределен в объеме среды: солнечный свет, хозяева для паразитов и т. д. Если количество ресурсов является постоянным, то это подразумевает также постоянную численность сообщества. Исходя из предположения *нулевой суммы* (zero-sum assumption), все доступные ресурсы жизнедеятельности стационарного изолированного сообщества полностью использованы в любой момент времени (Ricklefs, 2003).

максимального правдоподобия, которые используются как параметры при построении и сравнении моделей оценки этих вероятностей.

Это – важная составляющая UNTB, потому что обеспечивает настройку нейтральной модели к данным, полученным различными способами: нейтральная теория сконструирована «как на заказ» для ученого-экспериментатора (Alonso, McKane, 2004). Такой подход обеспечивает адекватную верификацию нейтральных моделей и возможность корректного сопоставления сравниваемых локальных сообществ.

Этим нейтральные модели отличаются от классических моделей обилия видов, которые почти полностью игнорируют вопрос о процедурах укомплектования эмпирических данных. Для подбора параметров, например, логнормальной модели Престона необходимо задать одномерный вектор оценок популяционных численностей $\{n_1, n_2, \ldots, n_s\}$ видов, которые, строго говоря, неизвестны. В частности, в приведенных примерах предыдущих разделов в качестве n_i нами использовалась суммарная численность i-го вида во всех пробах водотока. Если принять во внимание, что значительное большинство видов макрозообентоса встретилось только в 1-2 выборках, то эти модели скорее описывали «распределение проб», чем реальную кривую разнообразия–доминирования видов в сообществе.

Наконец, в UNTB объект исследований (т. е. конкретное локальное сообщество) рассматривается как часть *мета-сообщества* (metacommunity; Gilpin, Hanksi, 1991; Leibold et al., 2004), содержащего весь пул видов и служащего источником для восполнения тех из них, которые по каким-то причинам исчезли в локальных сообществах. Иначе говоря, мета-сообщество – это нечто вроде материка или большого архипелага островов по отношению к небольшому изучаемому острову. Другое образное клише для мета-сообщества – гигантский «термостат» (heat bath) или резервуар энергии бесконечной емкости, эффективно используемый множеством приемников.

Существенное место в нейтральной теории занимают модели обмена видами и индивидуумами между локальным сообществом и мета-средой (Etienne, Alonso, 2007). Выделяются *сообщества, формируемые рассеиванием* (dispersal-assembled communities), видовой состав которых определен главным образом процессами рассеивания (миграции). При ограничении на рассеивание (dispersal limitation) инвазионные виды не всегда могут свободно проникать в сообщество и вероятность того, что свободный участок будет занят в результате миграции, будет меньше 1. Мера, количественно оценивающая рассеивание в данном сообществе – *фундаментальное число рассеивания* (fundamental dispersal number). Это сложный параметр, включающий интенсивность миграции и локальный размер сообщества (число особей), который может быть оценен по эмпирическим данным обилия видов. В конечном итоге локальное нейтральное сообщество рассматривается как биномиальная выборка из бесконечного мета-сообщества или гипергеометрическая выборка из конечного мета-сообщества с теми или иными ограничениями на рассеивание.

Следствия из описанных предположений приводят к тому, что С Хаббел (Hubbell, 2001) назвал «нейтральным или экологическим дрейфом» (neutral drift), когда динамика обилия различных видов определяется только стохастическими событиями рождения, смерти и рассеивания. Дрейф характеризуется случайными флуктуациями численностей видов, причем ни у одного из них нет устойчивого преимущества в процессе эволюционного изменения сообществ. В результате, разные виды могут длительное время сосуществовать вместе, несмотря на конкуренцию за общие ресурсы.

Дальнейшее развитие теории организации сообществ без сомнения сделает менее категоричными предположения нейтральности, рассматривая их как предельный случай. Вполне возможно, что экологические сообщества часто не являются нейтральными, но они также и не являются сообществами строгой конкурентной иерархии. В пределах большого сообщества могут быть нейтральные подсообщества, связанные друг с другом через ненейтральные процессы (Alonso et al., 2006).

Выдвинув принцип нейтральности как основу сосуществования видов, Хаббел предложил многомерную модель, предсказывающую относительное обилие видов в сообществе, а И. Волков с соавторами (Volkov et al., 2003**)** разработал аналитическое решение этих уравнений. По их мнению (впрочем, вполне естественному), этой модели следует отдать предпочтение, поскольку она исходит из биологически обоснованных предположений, лучше описывает число редких видов и реально подразумевает определенный механизм формирования сообщества. В свою очередь, Б. Макгилл (McGill, 2003a, b), подробное изложение одной из статей которого мы приводим в «Заключении», не нашел сколько-нибудь серьезных отличий результатов тестовой аппроксимации, показанных моделью UNTB, от традиционного логнормального распределения Престона.

Выше подчеркивалось, что в отношении совокупностей данных, описывающих естественные сообщества, почти универсальными являются два утверждения:

- существенную (если не сказать, подавляющую) фракцию общей численности особей составляет небольшое количество базовых видов (common species);
- большую часть видов от их общего числа в сообществе составляют редкие виды (rare species), представленные малой популяционной плотностью или одиночной встречаемостью.

Подобные утверждения вызывают обычно большое количество вопросов. Почему виды с высокой численностью настолько превышают среднее обилие видов в сообществе? Почему редкие виды являются редкими? Это результат плохой адаптации к местным условиям, или редкий вид имеет совершенно иную жизненную стратегию, чем виды с высокой численностью? Нейтральная теория подразумевает, что такими рассуждениями можно пренебречь. Поскольку у массовых эврибионтных видов более высокий коэффициент рождаемости, следовательно, UNTB предсказывает, что формирование сообществ происходит с большей вероятностью базовыми видами, чем редкими.

А.М. Гиляров (2007) внимательно проанализировал этот обмен мнениями и подчеркнул, что речь идет о двух принципиально разных механизмах сосуществования видов. П. Чессон (Chesson, 2000) обозначил их как *выравнивание* и *стабилизация*. Нейтральная модель Хаббела основана на сближении видов по своим демографическим характеристикам, т. е. на выравнивании. Относительная скорость роста любой популяции остается при этом постоянной и не зависит от частоты встречаемости вида. Стабилизация предполагает усиление внутривидовой конкуренции по сравнению с межвидовой по мере того, как растет численное доминирование вида в сообществе. Поэтому, если какой-то редкий вид становится все более массовым, относительная скорость его популяционного роста будет закономерно снижаться.

П. Эдлер с соавторами (Adler et al., 2007), развивая представления Чессона, предложил гипотетическую графическую модель сосуществования видов, объединяющую механизмы выравнивания и стабилизации. Пока трудно судить, насколько «объединительная» модель Чессона–Эдлера поможет сближению разных точек зрения на механизм формирования сообществ. Однако Гиляров (2007) считает, что наблюдаемое оживление интереса к изучению сообществ есть не просто преходящая мода, а преддверие некого действительно нового важного этапа.

5.5. Динамика обилия видов с позиций *Унифицированной нейтральной теории* (UNTB)

Математические конструкции С. Хаббела (Hubbell, 2001; Alonso et al., 2006) были приняты по аналогии с генетическими приложениями: например, процесс образования новых видов соответствовал мутациям в аллелях. Основными являются четыре параметра:

- J_L – суммарная численность всех популяций локального сообщества;

- J_M – общая численность всех популяций мета-сообщества;
- M – интенсивность перемещения особей из локального сообщества и в него;
- ν – интенсивность образования новых видов в сообществе.

С использованием этих параметров можно рассчитать многие производные величины. Выше упоминалось фундаментальное число рассеивания $I \cong 2mJ_L$. Аналогично этому теория предсказывает существование фундаментальной постоянной *биологического разнообразия* (speciation rate), записываемой как $\theta = 2\nu J_M$, которой приписывается параметризация числа видов в широком диапазоне пространственных и временных масштабов. Тогда функция распределения численности видов будет выглядеть как

$$\Phi(x) = \theta\, x^{-1}\, (1-x)^{\theta-1}.$$

Видовой состав моделируемого сообщества стохастически изменяется с течением времени. Однако любому мгновенному срезу частной структуры может быть приписана совокупная вероятность (associate probability) её возникновения. Пусть сообщество, насчитывающее J особей, составлено из S различных видов и имеет показатели обилия n_1 для вида 1, n_2 – для вида 2 и так далее до n_s. Согласно *Унифицированной нейтральной теории* (UNTB), совокупная вероятность такого сообщества будет равна

$$\Pr(n_1, n_2, \cdots, n_S \mid \theta, J) = \frac{J!\theta^S}{1^{\varphi_1} 2^{\varphi_2} \cdots J^{\varphi_J}\ \varphi_1!\,\varphi_2!\cdots\varphi_J!\ \ \prod_{k=1}^{J}(\theta+k-1)}, \qquad (5.7)$$

где $\theta = 2\nu J_M$, а φ_i – число видов сообщества, состоящего из i особей. Это уравнение показывает, что UNTB подразумевает нетривиальное равновесие процессов доминирования и выравнивания в сообществе, обусловленное образованием и исчезновением видов.

Рассмотрим в качестве примера сообщество из 10 особей, принадлежащих трем видам с распределением численностей 3, 6 и 1 соответственно. При этом имеем $S = 3$ и $\varphi_1 = \varphi_3 = \varphi_6 = 1$, а остальные значения φ равны нулю. Тогда подстановка в формулу (5.7) позволяет записать выражение

$$\Pr(3, 6, 1 \mid \theta, 10) = \frac{10!\theta^3}{1^1 \cdot 3^1 \cdot 6^1 \cdot 1!\,1!\,1! \cdot \theta(\theta+1)(\theta+2)\cdots(\theta+9)},$$

максимум которого, найденный численными методами, дает возможность получить значение для θ. Оценка максимального правдоподобия для θ в нашем случае составляет приблизительно 1,148. Логика говорит нам также, что совокупные вероятности различных срезов сообщества будут одинаковыми, если просто переставлять значения численностей между видами:

$$\Pr(3; 3, 6, 1) = \Pr(3; 1, 3, 6) = \Pr(3; 3, 1, 6)\,.$$

Если использовать ранжированный ряд доминирования–разнообразия, то выражение для объединенной вероятности (5.7) можно записать как

$$\Pr(S; r_1, r_2, \ldots, r_s, 0, \ldots, 0),$$

где r_i – обилие i-го вида, $i = 1, 2, \ldots, S$, причем $r_i \geq r_{i+1}$, т. е. r_1 – относительное обилие самого доминирующего вида и так далее до r_s – доли самого малочисленного вида. Для удобства этот ряд обычно дополняется нулевым «хвостом», чтобы отразить тот факт, что в состав сообщества J может входить некоторое количество видов с обилием, близким к нулю. Тогда можно оценить *ожидаемое* обилие i видов с наиболее высокой численностью особей:

$$E(r_i) = \sum_{k=1}^{C} r_i(k) \cdot \Pr(S; r_1, r_2, \cdots, r_S, 0, \cdots, 0)\ ,$$

где C – общее количество конфигураций; $r_i\,(k)$ – обилие i ранжированных видов в k-й конфигурации; $\Pr(\ldots)$ – объединенная вероятность доминирования–разнообразия.

Представленные формулы отражают модель локального полностью закрытого сообщества. В общем случае биосфера представляет собой совокупность локальных сооб-

ществ, связанных между собой множеством миграционных потоков потенциально взаимодействующих видов, что все вместе выше было определено как мета-сообщество. Распределение видов в мета-сообществе определяется соотношением популяционных плотностей и общими механизмами динамического равновесия между появлением новых видов и исчезновением старых.

Движущие силы двухсторонних миграционных потоков особей между локальным сообществом и мета-сообществом могут быть смоделированы с использование пространственных соотношений между ядрами рассеивания. В общем случае, показатель интенсивности перемещения m зависит от соотношения вероятностей переноса особей (например, семян) из мета-сообщества в локальное сообщество с целью занятия свободного участка p_{in} и миграции особей из локального сообщества во внешнюю среду p_{out}. Тогда в терминах числа перемещаемых особей N_{in} и N_{out} можно записать соотношения (Clark et al., 1999; Etienne, Alonso, 2005):

$$m = \frac{N_{out}}{N_{in} + N_{out}}; \qquad N_{out} = 2\pi\rho P \int_R^{\infty} l p_{l,out} dl; \qquad N_{in} = 2\pi\rho P \int_0^{R} l p_{l,\text{int}} dl,$$

где R – радиус элементарной площадки локального сообщества, ρ и P – плотность и производительность источников перемещения. Эти выражения должны учитывать все элементарные площадки, по которым идет приход или расход особей как внутри локального сообщества, так и на смежной территории мета-сообщества.

С теоретико-вероятностных позиций в каждый момент времени происходит одно из двух событий: (*а*) с вероятностью $(1 - m)$ из локального общества наугад извлекается особь и заменяется другой случайной особью с возвращением первой; (б) с вероятностью m случайная особь заменяется иммигрантом, привлеченным из мета-сообщества. Считается, что изменения мета-сообщества происходят в значительно большем масштабе времени и его состав предполагается условно постоянным на период обмена особями с локальным сообществом.

Оценку параметра m можно осуществить с использованием известных в теории вероятностей *схем урны* (Феллер, 2010, с. 124). В соответствии с этими схемами из урны, наполненной шарами разного цвета (т. е. сообщества, включающего множества особей разных видов), последовательно извлекаются шары с добавлением новых. Вероятность извлечь шар определенного цвета зависит как от соотношения шаров в урне, так и от исхода предыдущих извлечений. Процесс извлечения шаров из урны, определяющий миграцию особей и, следовательно, динамику распределения обилия видов, в данном случае задается выборочной формулой У. Эванса (Ewens, 1972). Особый случай, который С. Хаббел назвал *точкой мутации*, соответствует появлению и извлечению из урны шара совершенно нового цвета (т. е. вида, не характерного для локального сообщества).

Если $m = 1$, то рассеивание неограниченно и локальное сообщество – только случайная выборка из мета-сообщества. Для $m = 0$ локальное сообщество полностью изолировано от окружающей среды, и все виды со временем исчезнут, кроме одного (Hubbell, 2001). Если $0 < m < 1$, то локальное сообщество – часть мета-сообщества, ограниченная рассеиванием, и меру этих ограничений следует учитывать в формулах. Тогда ожидаемые значения распределения видов в локальном сообществе зависят от всех четырех параметров: J_{L}, J_M, ν и m, причем J_M и ν интерпретируются как параметры урны. И. Волков с соавторами (Volkov et al., 2003) показали, что ожидаемое число видов с обилием n_{i} в условиях миграции может быть рассчитано на основе модели с использованием гамма-функций. В последствие эта модель была уточнена (Etienne, Alonso, 2007), а программные модули аппроксимации числа видов на основе принципа максимального правдоподобия представлены в Интернете.

Все модели биоразнообразия, описанные до UNTB, были связаны со значениями обилия видов, не зависящими от времени. Однако для анализа процессов эволюции сообществ и проблемы сохранения видов крайне важно сравнить динамику экосистем с моде-

лями биоразнообразия (Leigh, 2007). Нейтральная модель, которая была представлена в статье (Azaele et al., 2006), может аналитически предсказать временную динамику относительного обилия видов, а также рассчитать индекс оборота распределения видов (species turnover distribution – STD), определенный как вероятность $P(r, t)$, с который совокупность любых видов может изменить свою фракцию r за период времени t. В пределах этой модели совокупность любых видов представлена непрерывной (случайной) переменной x, динамика которой описывается следующем *уравнением Ланжевена*:

$$\dot{x} = b - x/\tau + \sqrt{Dx}\xi(t) ,$$

где b – показатель иммиграции особей из мета-сообщества; составляющая x/τ отражает уровень конкуренции за конечные ресурсы; член D связан с демографической стохастикой; $\xi(t)$ – гауссов белый шум. Решение может быть получено как непрерывное приближение основного уравнения при условии, что показатели рождения и смертности особей не зависят от вида, а в установившемся режиме распределение обилия может быть аппроксимировано гамма-распределением.

У каждого из представленных выше параметров J_L, J_M, ν и m есть достаточно четкая спецификация, которая в принципе позволила бы оценивать их непосредственно из результатов экологических наблюдений. Однако на практике это происходит чрезвычайно редко. Хотя размеры популяций локальных сообществ могут быть рассчитаны по выборочным данным, оценка размера мета-сообщества J_M проблематична, а интенсивности видообразования ν и перемещения m практически никогда не могут быть измерены непосредственно (Gotelli, McGill, 2006). Можно припомнить лишь одно уникальное исследование (Ricklefs, 2003), в котором некоторые из нейтральных модельных параметров были оценены опытным путем. Р. Риклефс рассчитал размер мета-сообщества (J_M) для тропических деревьев и сопоставил его с θ-оценкой Хаббела. Получающаяся нейтральная модель предсказала, что в среднем виды древесных пород будут исчезать после 9 поколений, что большинству специалистов показалось невероятно коротким.

В большинстве исследований нейтральных моделей один или несколько параметров оценивался косвенно с использованием алгоритмов подбора оптимальных функций, аппроксимирующих некоторые частные наборы данных. Такой подход не вполне корректен, потому что качество нейтральной модели будет зависеть от значения параметра (например, m), который представляет собой статистически «наилучшее приближение» к данным наблюдений, по которым впоследствии строится сама нейтральная модель. Это приводит к увеличению вероятности статистических ошибок при проверке гипотез (например, невозможность отклонить нейтральную модель, когда она неверна; McGill et al., 2006).

Выше, в разделе 4.4, мы обсуждали склонность к гипердиагностике нуль-моделей без ограничений на рандомизацию (Gotelli, 2000). Нейтральные модели, использующие принцип экологической эквивалентности, при оценке вероятности возникновения видов и перемещения особей могут испытывать ту же самую проблему, особенно в случаях с видами, которые очень отличаются по размеру тела или другим координатам осей ниши. Дальнейшее исследование риска статистических ошибок 1-го и 2-го рода представляется чрезвычайно важным для анализа нейтральных моделей (McGill et al., 2006; Gotelli, McGill, 2006).

ГЛАВА 6. МЕТОДЫ ОРДИНАЦИИ ВИДОВ И МЕСТООБИТАНИЙ

1. *Никто не знает, как всё было.*
2. *Все источники в той или оной мере сфальсифицированы.*
5. *Любая интерпретация истории [фактов] верна.*
6. *Если какое-либо утверждение является верным, то также верно и обратное.*

Дмитрий Быков "ЖД" (2006 г.), из "Памятки альтернативщику"

6.1. Актуальные проблемы современной фитоценологии

На протяжении XX в. при проведении фитоценологических исследований был выработан комплекс понятий, идей и концептуальных предположений, во многом предопределивших развитие теоретических представлений об экологии сообществ в целом. Одновременно создавался практический инструментарий обработки эмпирических данных, включая оценку сходства экологических объектов, выделение групп устойчивых ассоциаций видов, градиентный анализ, ординацию и т. д. Эти концепции и методы оказались столь универсальными, что постепенно стали приниматься на вооружение при изучении фаунистических и водных сообществ. В связи с этим целесообразно предварительно уточнить ключевые моменты разброса мнений, существующих в фитоценологии, применительно к выделению ценотических групп видов и классификации биотопов.

Как обычно бывает при изучении сложных систем, отдельные гипотезы о механизмах их функционирования принципиально односторонни, неполны и противоречивы. Длительные дискуссии, развернувшиеся относительно структурной изменчивости растительных сообществ, привели к тому, что некоторые исследователи начали расценивать итоги развития гносеологической базы фитоценологии (с точки зрения её целостности) как «коллективную неудачу». «В настоящее время эта дисциплина состоит из многочисленных, иногда противоборствующих школ, представители которых в течение длительного времени предпринимают отчаянные попытки выработать общепризнанную концепцию» (Прилуцкий, 2007, с. 6).

Соотношение континуальности и дискретности (Гиляров, 1988; Тимонин, 1989; Миркин, 1989, 1990; Кафанов, 2006; Миркин и др., 2007; Прилуцкий, 2007). Сторонники континуального подхода утверждают, что непрерывный характер изменения состава и структуры сообществ детерминируется постепенностью изменения экологических условий в пространстве и индивидуальностью экологии разных видов. Радикальное проявление континуальности – принципиальное отрицание целостности территориальных комплексов, а следовательно, порождаемых ими ассоциаций видов. С другой стороны, «организмисты» считают[1], что, даже в случае значительных градиентов факторов среды, последователь-

[1] При этом можно договориться, например, и до таких заключений: «компактная группа модальных растений представляет собой надорганизменную систему, достигшую стационарного состояния», что «фитоценозы как самостоятельные единицы растительного покрова существуют не в пространстве, а только во времени» или «растительные системы – это материальные структуры, способные извлекать из окружающей среды информацию в форме ритмических колебаний градиентов экологических факторов среды, освобождать ее от вещества и энергии и включать в свой наследственный информационный стереотип, осуществляя тем самым эволюцию системы популяции в направлении оптимизации стохастических и детерминированных связей» (Прилуцкий, 2007, с. 10, 12, 22).

ность местообитаний можно разбить на элементарные участки, внутри которых вариация условий среды не выходит за пределы нормы популяционных реакций. Биота внутри участков такого экологического ряда однородна, поэтому их сообщества искусственно наделяются свойством целостности.

И в том и другом случае логика рассуждений упрощена, а обе полярные концепции внутренне противоречивы (отметим большую аналогию с концепциями речного континуума и динамики пятен, представленными в разделе 1.2). Из всех многочисленных признаков организованности экосистем в этой дискуссии абсолютный приоритет отдан лишь одному из них – целостности, решающим критерием которой признается лишь наличие или отсутствие четких границ между различными сообществами. В то же время экосистемы принадлежат к классу динамических систем, границы и состав которых размыты и далеко не очевидны, а реальность существования фиксированных социумов носит ярко выраженный вероятностный характер.

Сегодня оппоненты уже не столь непримиримы, как это было в начале 60-х годов прошлого столетия, когда Р. Макинтош (McIntosh, 1967b) назвал их «геоботаническими Монтекки и Капулетти»; тем не менее, их видение природы растительности различно. Правда, «континуалисты» уже не так категоричны и их основной постулат – независимость распределения видов – «ныне воспринимается не как абсолют с обязательным соответствием этих распределений форме симметричной колоколовидной кривой» (Миркин, Наумова, 1998, с. 14): была показана и ценотическая обусловленность асимметричности распределения видов на градиентах, и описаны процессы формирования групп видов (нуклеации), противоречащие «классическому» взгляду на континуум, и обоснованы механизмы тонкой биотической дифференциации, как коадаптации видов друг к другу, и продемонстрированы случаи формирования резких границ между экологически различными сообществами за счет эффекта «переключения», и т. д. Таким образом, 100-летняя дискуссия[2], как нам представляется, пришла к равновесию – «и ты – прав, и ты – прав, и ты, Сарра, тоже права…» (см. название методологической статьи В.Д. Александровой [1966]).

Основываясь на работе В.С. Ипатова и Л.А. Кириковой (1997), еще раз проанализируем причины, приводящие к континууму и дискретности (см. табл. 6.1).

Таблица 6.1

Причины непрерывности и дискретности растительности

Причины непрерывности	Причины дискретности
1. Постепенность изменения среды и тесная зависимость от нее распределения видов в пространстве (непрерывность экотопа).	1. Наличие переломных пунктов в изменении прямодействующих факторов (дифференцированность экотопа).
2. Неспецифичность воздействия видов на среду.	2. Специфичность трансформации среды растениями.
3. Равномерность воздействия на среду природных факторов («растекание» видов в пространстве).	3. Катастрофическое воздействие на среду и растительный покров природных факторов.
4. Непрерывность воздействия на среду и растительность деятельности животных и человека.	4. Дискретность воздействия на среду и растительность деятельности животных и человека.
5. Отсутствие экологических групп видов.	5. Наличие экологических групп видов.

[2] Заметим, что первые публикации «континуальной направленности» появились в хронике XII съезда естествоиспытателей и врачей в начале 1910 г. (Раменский, 1910) и несколько позже (Gleason, 1910; цит. по: Трасс, 1976, с. 231).

Если принять эти пять пар причин за основу, то на конкретном участке пространства возможно возникновение $2^5 = 32$ различных ситуаций, проанализировать которые достаточно сложно в силу того, что каждая из причин может оказывать различное по силе влияние, зачастую «уравновешивая» или «заменяя» другие причины. Однако, учитывая, что причины 1, 3, 4 и 5, в известной мере, отражают разные стороны одной комплексной причины, связанной с характером воздействия среды на растения, эту схему можно редуцировать всего до двух причин: непрерывности - дискретности экотопа и специфичности - неспецифичности воздействия растений на среду; тогда число возможных ситуаций сокращается до четырех.

Две «крайние» ситуации легко задают необходимые условия существования непрерывности (непрерывность экотопа и неспецифичность воздействия видов на среду) и дискретности (дифференцированность экотопа и специфичность трансформации среды растениями). Правда, в последнем случае возможна ситуация «несовпадение границ» основных причин, что приведет к возникновению более «пестрой картины» в распределении растительности, которая может идентифицироваться как непрерывность. Ситуация постепенности изменения среды и специфичности воздействия на нее растений будет в значительной степени зависеть от числа эдификаторов: для одного эдификатора (бореальная растительность) в этом случае можно говорить о дискретности; в случае полидоминантности (например, луговая растительность) на фоне непрерывности воздействия среды происходит «наложение» полей воздействия эдификаторов, что приводит к возникновению непрерывности. Наконец, при неспецифичности воздействия видов на среду и наличии переломных пунктов в воздействии экологических факторов будет наблюдаться дискретность (влияние снежного покрова на растительность тундры – по данным В.Д. Александровой; см.: Ипатов, Кирикова, 1997, с. 229). С этих позиций можно говорить о превалировании непрерывности над дискретностью в растительности, однако само наличие «дилеммы дискретность – непрерывность» заставляет использовать различные методы изучения растительности и экосистем в целом – и классификацию, и ординацию, причем классифицируют не только дискретные, но и непрерывные сообщества (например, растительные сообщества лугов), так же как и ординируют существенно дискретные сообщества (например, «крест Сукачева» для типов еловых лесов).

Еще одна группа критериев (в какой-то степени «пересекающаяся» с критериями Ипатова и Кириковой) приведена в табл. 6.2 (по: Миркин и др., 1989, с. 133). Можно также подчеркнуть зависимость степени непрерывности растительного покрова от наличия доминантов и видового разнообразия (табл. 6.3; Миркин, Розенберг, 1983, с. 48).

С современных позиций континуум представляет собой сложное и иерархическое явление (Миркин и др., 1989; Миркин, Наумова, 1998). При этом различают разнообразные формы континуума: *топографический* (постепенное изменение растительности в пространстве), *временно́й* (постепенное изменение растительности в ходе экологической сукцессии), *синтаксономический* (наличие переходов между классификационными единицами фитоценозов) и *биоценотический* (отсутствие четких границ у каналов трансформации вещества и энергии в экосистеме). Для обоснования идеи иерархии континуумов полезны исследования И. Хэнски (Hanski, 1982; Миркин, Наумова, 1998): виды расклассифицированы по двум параметрам – широте амплитуды их распространения и обилию (см. табл. 6.4). При этом континуумы высших уровней иерархии (зональные смены, высотная поясность) формируются центральными видами с различным составом видов-спутников и видов-горожан (последние дают вспышки обилия в отсутствии центральных видов). Континуумы низших уровней более «мобильны» и в пространстве, и во времени; на фоне ограниченного числа центральных видов непрерывность формируют виды-селяне, а мозаичность (например, в сообществах высокотравной прерии) – виды-спутники.

Таблица 6.2

Сравнение основных элементов парадигм организмизма и континуализма

Элемент парадигмы	Парадигма	
	Организмизм	**Континуализм**
Понимание фитоценоза	Реальные, исторически обусловленные целостные совокупности популяций, связанные в первую очередь взаимоотношениями растений и формирующиеся под контролем эдификаторов.	Условно однородные части континуума, совокупности дифференцированных по нишам популяций, объединенных условиями среды. Вклад взаимоотношений различен в разных типах растительности.
Категории для оценки разнокачественности популяций в сообществе	Фитоценотипы – типы популяций по характеру отношений друг к другу.	Типы эколого-ценотических стратегий – типы популяций по реакции на биотические и абиотические условия.
Представления о синморфологии	Ценоэлементы различаются четко.	Ценоэлементы могут различаться и четко, и иметь континуумообразную структуру.
Представления о синдинамике	Сообщества изменяются как целостные единства, детерминировано с достижением ограниченного числа климаксовых состояний.	Популяции меняются более или менее независимо, процессы носят стохастический характер, сопровождаются дифференциацией экологических ниш и завершаются климакс-мозаикой.
Представления об эволюции	Коадаптация популяций	Сеткообразный характер эволюции.
Классификация экосистем	Построение естественной иерархической классификационной системы на основе сходства эдификаторов, как видов, определяющих внутреннюю сущность сообществ.	Естественная иерархическая классификация невозможна, любая классификация – приближение к естественной.

Таблица 6.3

Зависимость степени непрерывности растительного покрова от числа доминантов и видового разнообразия

Видовое разнообразие	Доминанты	
	Сильные	**Слабые**
Высокое	непрерывность (тропический лес)	непрерывность (травяные сообщества)
Низкое	дискретность (бореальный лес)	дискретность (пустынные сообщества); непрерывность (лишайниковый покров тундры)

Таблица 6.4

Классификация видов по И. Хэнски (Hanski, 1982)

Обилие	Диапазон условий среды	
	Широкий	**Узкий**
Высокое	виды ядра (англ. *core*)	горожане (англ. *urban*)
Низкое	селяне (англ. *rural*)	спутники (англ. *satellite*)

Наконец, приведем достаточно пространное высказывание Л.Г. Раменского (1924, с. 26-27), которое демонстрирует его глубокое понимание одного из коренных свойств растительности, а, в более широком плане, – существа экосистем (если не принимать несколько прямолинейную аналогию растительного сообщества с некоторыми однофазными и гетерогенными химическими равновесиями в силу системного видения проблемы; см. далее): «Аналогичными путями [*как в физике и физической химии. – Ремарка наша*] должна идти и работа синэколога: 1) выделение факторов, прямо и косвенно определяющих равновесия растительности...; 2) анализ состава и прочих особенностей равновесных ценозов, отвечающих различным местообитаниям; 3) их группировка в экологические ряды с построением кривых обилия (концентраций) растительных видов в этих рядах, по координатам условий; 4 и 5) впоследствии (в будущем или может отдаленном будущем) математический анализ координатных схем и их приведение к законам физики, химии и физиологии; 6) подобно физику и синэколог должен изучать динамику ценобиоза и установления равновесий (нарушенные и эволюционирующие ценозы и т. д.). Как и физико-химик, синэколог и биоценолог должен н е р а с ч л е н я т ь своего объекта (растительность) на иерархию условных единиц.., но с о е д и н я т ь , с п а и в а т ь все разрозненные а н а л и з ы ц е н о з о в в е д и н у ю к о о р д и н а т н у ю с х е м у (*разрядка автора*. – *Ремарка наша*). Работа в этом отношении не игнорирует возможных случаев скачков в сменах растительного покрова, подобных точкам инверсии физико-химических систем (вода – лёд – пар, два гидрата глауберовой соли в одном растворе и т. п.). Координатный метод универсален и в нем одинаково полно отразятся и непрерывность и возможные (но мне лично неизвестные) скачки, переломы экологических рядов».

Индивидуалистский и холистический подход. Еще один «краеугольный камень» современной теоретической фитоценологии – это соотношение «индивидуализма» («редукционизма») и «холистизма» (возможность привлечения методологии системного подхода). Об этот «камень» также сломано не мало копий.

Успехи современной физики повлекли не только проникновение в биологию физико-химических методов исследования объектов различных уровней биологической иерархии, но и определенное «навязывание» физического «образа мышления» при постановке и решении различных биологических задач. При этом забывалось, что физический научный метод – наблюдение, размышление и опыт – применялся к анализу *простых свойств* исследуемых систем. А. Эйнштейн (1965, с. 9) писал: «… физик вынужден сильнее ограничивать свой предмет, довольствуясь изображением наиболее простых, доступных нашему опыту явлений, тогда как все сложные явления не могут быть воссозданы человеческим умом с той точностью и последовательностью, которые необходимы физику-теоретику. Высшая аккуратность, ясность и уверенность – за счет полноты». А фитоценозы и экосистемы – объекты системологии (науки о сложных системах).

Несмотря на это, один из основных принципов физического научного метода – редукция – широко используется в биологии. А.А. Любищев (1977) различал три основных типа редукционизма в биологии:

- молекулярно-биологический (принцип исследования, состоящий к сведению сложного к совокупности или сумме его частей, при изучении которых получают сведения и о свойствах исходного целого);
- иерархический (возможность интерпретации явлений высших уровней биологической иерархии на языке молекулярных моделей);
- эволюционный (сведение всей социальной эволюции к биологической).

Оптимизм физиков в объяснении экологических (надорганизменных) феноменов связан с первым и вторым типом редукционизма. При этом сторонники редукционизма оказываются более воинствующими, чем системологи, и не приемлют иных точек зрения. Любищев говорит даже об «ультраредукционистском энтузиазме». Примером может слу-

жить статья А.Н. Прилуцкого [2007, с. 22], в которой автор утверждает, что «… существующие методы описания растительности не совместимы с системным подходом. Присутствие элементов софистики в теории фитоценологии свидетельствует о том, что её современная парадигма эклектична, составляющие её учения концептуально разнородны, а видимость единства всей научной дисциплины обеспечивается отклонениями в логике анализа от диалектического метода». Вот так!

Спокойный и трезвый обзор соотношения редукционизма и холистизма можно найти в целом ряде работ (например: Любищев, 1977; Мейен, 1978; Миркин, Розенберг, 1978б; Борзенков, 1982; Розенберг, 1984). При этом авторы сходятся в том, что необходимо «ясно сформулировать основные принципы такого подхода, который бы включал редукцию как рабочий метод, но не включал редукционизм как систему постулатов, не повторял увлечений редукционизма, но и не уступал ему по продуктивности» (Мейен, 1978, с. 164). Действительно, редукция в том или ином виде неизбежно присутствует в любом экологическом (хотя бы в принятии иерархической организации мира) и фитоценологическом исследовании (например, индивидуалистическая гипотеза Раменского–Глизона: «Видовая специфичность распределения, несомненно, является отражением факта физиологического своеобразия каждого вида организмов» [Раменский, 1938, с. 15]). И это справедливо подчеркивает Ю. Одум (1986, с. 18): «… и холистический, и редукционистский подходы следует использовать в равной мере, не противопоставляя их друг другу... Экология стремится к синтезу, а не к разделению».

В этой ситуации необходимо четко представлять, что **метод редукции применим для анализа простых свойств как простых, так и сложных систем** (Розенберг, 1988, 2005). Однако степень экстраполяции полученного при такой редукции знания будет определяться естественной типологией исследованных объектов (в экологии – экосистем), а это уже сложная и целостная характеристика.

Приоритетное значение функциональной роли различных компонентов биоценоза по сравнению с видовой структурой. Карл Линней (Carl Linnaeus, 1707-1778) в XVIII в. создал столь блестящую таксономическую парадигму, что экологи восприняли виды как элементарные составляющие сообществ. Появились представления, что достаточно разложить биоценоз на ценопопуляции и можно получить ключ к познанию механизмов формирования и локального сообщества, и экосистемы в целом (Галанин, 2005). Однако для поддержания стабильности экосистем и биогеохимической оптимизации биоценозов важно не видовое разнообразие само по себе, а *функциональная* или *биоморфологическая сбалансированность* (Любищев, 1982). Все чаще в декларативной форме признается, что классификация сообществ должна осуществляться не на таксономической основе, а на основе жизненных форм.

Однако существующую реальность отражает высказывание В.И. Василевича (1983, с. 130): «В отношении функциональной организованности фитоценотических систем вопрос еще менее ясен, так как никогда и не поднимался». Приходится признать, что современная экология не располагает практическим методологическим аппаратом массового количественного исследования функциональных систем[3]. Например, можно достаточно

[3] В частности, представления о «функциональной экологии» А.С. Керженцева (2006) весьма пу́танны (чего только стоят аналогии экосистем и технических механизмов – «метод» технико-биологических аналогий «позволяет решать не только философские, но и конкретные вопросы функциональной экологии, например, управления функциями природных систем… Основная трудность такого подхода… доказать сходство или хотя бы подобие между природными и техническими системами» [Керженцев, 2006. с. 65, 66]) и не конструктивны (если не считать вводимые представления о *гидротермическом акселераторе метаболизма экосистем* – достаточно повысить интенсивность «метаболизма», используя такой «акселератор», и экосистема «самонастроится» на измененный режим оптимального функционирования путем метаморфоза или эволюционных изменений…; критику этих представлений см.: [Краснощеков, Розенберг, 2008]).

легко с использованием того или иного апробированного выборочного метода оценить видовой состав участка растительного покрова или водоема, но корректно рассчитать баланс материальных потоков для отдельных функциональных групп сообщества (или хотя бы качественно идентифицировать эти группы) является пока трудно разрешаемой проблемой.

Выделение устойчивых ассоциаций видов. В экологии сообществ, о чем свидетельствуют многочисленные литературные источники, на протяжении последних десятилетий сохраняется постоянный интерес исследователей к изучению сопряженной встречаемости видов в биоценозах. Ширина диапазонов толерантности к воздействию специфических экологических факторов в сочетании с особенностями межвидовой конкуренции определяют совокупность потенциальных сред обитания каждого вида. Основываясь на знании экологических оптимумов на шкале градиента изменяющихся биотопических условий, можно найти распределение таксоценозов по принадлежности к отдельным синтаксонам (или группам синтаксонов), которые соответствуют определенным средам обитания с конкретными экологическими параметрами.

Следует отметить множественный характер выделения композиций видов. Представления, согласно которым виды образуют группировки – синтаксоны, зачастую носят размытый, множественный и отчасти также противоречивый характер. С одной стороны, в рамках флористических подходов к классификации растительности достаточно хорошо проработана одна из вертикальных иерархий синтаксонов, которая включает субварианты, варианты, субассоциации, ассоциации, подсоюзы, союзы, классы и дивизионы (Баркман и др., 1988; Weber et al., 2000). Однако отсутствуют четкие методологические рекомендации по выделению отдельных уровней иерархии разбиения сообщества (что, в общем-то, косвенно свидетельствует о превалировании в естественных экосистемах непрерывности над дискретностью).

Нет также четких рекомендаций, связанных с методологией априорного выделения классов местообитаний. Группы биотопов обычно определяются в соответствии с наиболее «характерными» параметрами окружающей среды, уровнями внешнего (чаще всего, антропогенного) воздействия, сложившимися представлениями о типологии сообществ на основании предыдущего опыта и т. д. Эта проблема не является тривиальной из-за многоуровневой иерархии данных, поэтому редко удается сразу построить группировку, разумную и удобную для последующей математической обработки.

Под эколого-ценотической группой (ЭЦГ) принято понимать устойчивое сочетание видов, сходных по отношению к совокупности экологических факторов, приуроченных к одинаковым местообитаниям и показавших высокую взаимную встречаемость (Юрцев, Петровский, 1971; Миркин и др., 2001). В отечественной литературе встречаются иные термины, обозначающие подобные сочетания: «синузии» (Алёхин, 1935), «свиты» (Ниценко, 1969), «биоэкогруппы» (Апаля-Шидлине, 1971), «эколого-ценотические элементы» (Королюк, Намзалов, 1999) и т. д. Подчеркнем следующие аспекты формирования ЭЦГ:

- с формальной точки зрения ЭЦГ представляет собой минимальный и далее неделимый набор видов, необходимый и достаточный для индикации хотя бы одного специфического местообитания; объединение «*единиц*» в «*группы*» представляет самостоятельную нетривиальную научную проблему;
- анализ «совокупности *экологических* факторов» (т. е. статистическая оценка достоверности и силы их влияния на встречаемость и обилие видов), прозвучавший в представленном определении, обычно является привычным литературным клише, не предполагающим серьезных математических исследований; мы далее покажем, что этот аспект может изучаться в рамках *канонического анализа соответствий* и *ординации нечетких множеств*, но чаще всего группировка видов проводится косвенными методами, не предполагающим непосредственного анализа всего комплекса внешних параметров;

- «*одинаковость*» местообитаний также обычно строго не анализируется, так как в условиях биотопического разнообразия водных экосистем трудно выбрать корректную количественную метрику оценки значимости многофакторных экологических различий между отдельными биотопами и комплексно оценить сходство условий среды;
- как правило, основным критерием группировки видов в ценотические единицы принимается высокая вероятность их совместной встречаемости (т. е. «верность» видов друг другу и объединяющему их биотопу).

«Верность» (или *сродство*, *индикаторная ценность*) – одно из самых важных понятий в экологии растительных сообществ (Braun-Blanquet, 1918; Szafer, Pawlowski, 1927), которое определяет, следует ли считать вид характерным, дифференцирующим, сопутствующим или случайным для некоторого набора однотипных местообитаний. *Характерные* виды концентрируются исключительно в одном синтаксоне, который они диагностируют (Миркин, Наумова, 1998, с. 49). *Дифференцирующие* (*диагностирующие*) виды могут рассматриваться, как частный случай характерных видов, но они диагностируют синтаксоны пограничной частью своего ареала и отражают специфику видовой структуры одновременно нескольких ценотических единиц (Barkman, 1989).

Мера «*верности*» (fidelity) – количественная переменная, определяющая степень статистической принадлежности вида к той или иной ценотической единице. Ж. Браун-Бланке еще в 1921 г. разработал возрастающий ряд[4] пятибальной шкалы категорий верности (Александрова, 1969; Westhoff, van der Maarel, 1973). Например, виды считаются верными в высшей степени (с баллом 4 и 5), если они имеют относительную частоту встречаемости от 61 до 100% в данной ценотической единице и не более чем 40% вне этой единицы. В настоящее время такая грубая шкала «верности видов» подвергается критике, так как представление о существовании в синтаксоне видов с баллами 4 и 5 чаще всего связано с недостаточной репрезентативностью эмпирического материала. По всей вероятности, верность 3 балла для характерных и диагностирующих видов является наиболее распространенной; более высокие баллы верности могут иметь только диагностические виды некоторых эндемичных сообществ (Миркин и др., 1989).

Обоснование взаимосвязи между параметрами жизнеспособности отдельных видов и характеристиками среды их обитания является одной из общих и приоритетных задач экологии при изучении всех сообществ. Например, задолго до введения понятия «верности» видов в фитоценологии Р. Кольквитцем и М. Марссоном (Kolkwitz, Marsson, 1902) была выполнена массовая классификация видов по принадлежности к 4 зонам сапробности водоемов. В этой связи было бы разумно отказаться от частных терминов вроде «верности», «сапробности» и т. д., поскольку все они соответствуют общеэкологическому понятию *индикаторной ценности* видов применительно к определенной группе биотопов. Однако утверждение терминологических нормативов не входит в задачу нашей книги.

Если изменения, вызываемые внешней средой, находят надежное отражение в выборочных элементах мониторинговых данных, то вычисление значений индикаторной ценности видов принимает интуитивно ясный вероятностный характер. Для оценки степени совместной встречаемости видов в одних и тех же местообитаниях может быть использована, например, непрямая ординация методом главных компонент (Миркин и др., 1979; P. Legendre, L. Legendre, 1998). Считается, что наибольшую индикаторную ценность имеют виды наиболее удаленные в пространстве ординации от центра координат. Напротив,

[4] «Верные» виды (нем. *treue*; виды связаны исключительно [или почти исключительно] с одним синтаксоном) – 5 баллов; «прочные» (нем. *feste*; виды, предпочитающие один синтаксон, но с меньшим постоянством встречающиеся и в других) – 4 балла; «благосклонные» (нем. *holde*; виды, встречающиеся более или менее постоянно в разных синтаксонах, но все же предпочитающие один из них) – 3 балла; «неопределенные» (нем. *vaga*; виды лишены явно выраженной связи с одним из синтаксонов) – 2 балла; «чуждые» (нем. *fremde*; редкие и случайные виды [адвентивные, реликтовые и пр.]) – 1 балл (Миркин и др., 1989).

положение видов вблизи облака сгущения точек свидетельствует об их индифферентном отношении к ведущим факторам среды (Маслов, 1990). Однако эти методы, подробно описываемые в последующих разделах, не дают прямой оценки «верности» видов. Другим способом оценить взаимную встречаемость видов является использование коэффициентов сопряженности в четырехпольных таблицах (Миркин, 1971; Миркин, Розенберг, 1978а).

Понятие «верности» связано с распространенным дефектом аргументации, построенной на основе замкнутой логики: мы выделяем каждую ценотическую единицу по списку характерных и дифференцирующих видов и в то же время количественно ранжируем дифференцирующие виды по степени предпочтения, оказываемого этой же ценотической единице. К сожалению, теоретическая фитоценология не дает конкретного рецепта разрешения этого логического противоречия.

Основу современных методов моделирования сложных систем составляют различные приемы аппроксимации, которые становятся одним из генеральных направлений научных исследований (Новик, 1975). Целесообразное приближение на базе логических выводов существенно упрощает разрабатываемую модель экосистемы, сохраняя при этом её адекватность оригиналу. В итоге, строгая логика аппроксимации приобретает способность вступать в конкуренцию с логикой сверхупрощенных описательных подходов и вытеснять их из познавательного процесса.

В частности, итеративные оптимизирующие алгоритмы являются единственным способом разрешить замкнутое логическое противоречие «курицы и яйца», т. е. что является первичным при группировке: дифференцирующие виды, образующие синтаксоны, или сами синтаксоны, для которых выделяются индикаторные виды. Используя базы данных биогеографических описаний для выделения характерных композиций видов, обычно начинают с некоторого произвольного исходного набора ценотических единиц, после чего запускают процесс взаимной оптимизации видового состава синтаксонов и совокупностей однотипных местообитаний, связанных с выделяемыми ассоциациями. Впоследствии эти группы видов могут быть объединены в иерархические конструкции тем или иным способом, наиболее подходящим для практического использования в отношении эмпирических наборов данных.

6.2. Статистические методы оценки «верности» видов и формирования ценотических групп

Чтобы осуществить итеративную группировку видов с использованием конкретной оптимизирующей процедуры, необходимо определить непрерывную меру «верности», позволяющую выполнить корректную сравнительную оценку информативности дифференцирующих видов. Очевидно, что пяти упомянутых выше классов для этого недостаточно: «Чтобы корректно определить степень верности видов, мы нуждаемся в не слишком сложном, логически последовательном критерии, наилучшим образом отражающем наши идеи относительно экологических оптимумов и роли видов в растительных сообществах» (Barkman, 1989, p. 110). В качестве такой меры Ж. Бэркман предложил полное отношение проективных покрытий (TCR, total cover ratio), т. е. отношение относительной доли обилия вида в данной ценотической единице к сумме проективных покрытий в других ценотических единицах.

Определенная проблема в обосновании формул оценок «верности» связана с трактовкой понятия «другие ценотические единицы». В принципе, для каждого синтаксона есть две возможности сравнения частот встречаемости *j*-го дифференцирующего вида:

- относительно его частоты в остатке базы данных, т. е. во всем доступном массиве местообитаний, не принадлежащих к рассматриваемому синтаксону (Barkman, 1989);
- с учетом только частоты в подмножестве самых близких ценотических единиц, общий список которых ранжирован по убыванию встречаемости *j*-го вида.

В любом случае предполагается, что все возможные ценотические единицы должны быть заранее известны. Г. Бриссе с соавторами (Brisse et al., 1995) решили эту задачу изящно, определив каждую единицу, как характеризуемую одним из всех возможных видов во французской базе данных "SOPHY". Они оценивали «дискриминирующую способность видов» (species discrimination capability) как отношение частот совместной встречаемости каждой такой единицы в местообитаниях, не вполне верно называя полученную величину «верностью».

Продолжает оставаться дискуссионным вопрос о том, какой показатель использовать в качестве базового для выделения ценотических групп: присутствие/отсутствие вида или его обилие/покрытие. Есть ряд аргументов за определение информативности только на основе одного присутствия/отсутствия вида:

- показатели обилия/покрытия имеют значительную сезонную динамику и многолетнюю флуктуацию в пределах большинства популяций;
- использование этих показателей подразумевает введение системы весовых коэффициентов для различных категорий видов по обилию/покрытию (Goodall, 1973);
- негауссовый характер распределения количественных параметров вызывает необходимость введения функциональных преобразований (например, логарифмирования), компенсирующих этот эффект (van der Maarel, 1979);
- крупномасштабные обзоры изучаемых сообществ используют материал различных авторов, которые не всегда однозначно оценивают количественные показатели; использование только признака инцидентности (присутствия/отсутствия) может минимизировать эти индивидуальные отклонения.

С теоретико-вероятностных позиций верность вида может быть интерпретирована как отклонение частоты его встречаемости в ценотической единице от частот, которые ожидались бы при случайном характере распределения видов по местообитаниям. Количественно это отклонение может оцениваться с использованием значений χ^2 (Goodall, 1953) или *F*-критерия при проведении дисперсионного анализа (Jancey, 1979). Сравнение эмпирических частот с ожидаемыми значениями в соответствии с тем или иным теоретическим распределением является основой расчета индекса индикаторной ценности *IndVal* (Dufrêne, Legendre, 1997) и меры верности *u*, предложенной Х. Брюльхейдом (Bruelheide, 2000).

Пусть виды распределены случайно в пределах биогеографических описаний базы данных. Тогда функция ожидаемых частот μ оценивает равновероятное число появлений каждого вида в той или иной ценотической единице: $\mu = n\, N_k / N$, где N – общее число описаний в базе данных; N_k – число описаний, соответствующих рассматриваемой единице; n – общая встречаемость вида во всех местообитаниях. Если известно теоретическое распределение встречаемости видов, то может быть найдена вероятность того, что наблюдаемое число n_k появления вида в k-й ценотической единице будет выше, чем ожидаемое число μ.

Если предположить, что случайные величины n_k и μ имеют гипергеометрическое распределение, то мера верности u_{hyp} может быть определена как отклонение наблюдаемой частоты появления вида в ценотической единице от ожидаемой частоты μ, нормированное на стандартное отклонение (Chytrý et al., 2002):

$$u_{hyp} = \frac{(n_k - \mu)}{\sigma_{hyp}} = \frac{(n_k - \mu)}{\sqrt{nN_k(N-n)(N-N_k)/[N^2(N-1)]}}. \qquad (6.1)$$

Максимум u_{hyp}, равный $\sqrt{(N-1)}$, достигается для характерных видов при $N_k = n = n_k$.

Х. Брюльхейд (Bruelheide, 2000) в своей первоначальной статье использовал в качестве эталона биномиальное распределение, которое, как он считал, обеспечивает независимость появления видов в выборках. М. Хитрый с соавторами (Chytrý et al., 2002) показали, что эта форма индекса имеет тенденцию существенно завышать роль редких видов.

Кроме того, если гипергеометрическая форма u_{hyp} является симметричной мерой взаимной верности видов к ценотической единице и наоборот, то биномиальная форма u_{bin} имеет отчетливую асимметрию. Впрочем, легко вывести простые формулы пересчета из одной формы в другую (Chytrý et al., 2002): $u_{bin} = u_{hyp}\sqrt{(N-n)/(N-1)}$.

Легко также может быть показана связь меры верности u_{hyp} с двумя известными статистическими критериями, которые, в свою очередь, сами могут использоваться как индексы верности между видами и ценотическими единицами. *Коэффициент ассоциации* (Sokal, Rohlf, 1981), он же *коэффициент корреляции Чупрова* или *коэффициент сопряженности Браве* $\Phi = \dfrac{ad-bc}{\sqrt{(a+b)(a+c)(b+d)(c+d)}} = \dfrac{u_{hyp}}{\sqrt{(N-1)}}$, где a, b, c, d – компоненты таблицы сопряженности 2 × 2, является мерой корреляционной связи и не может быть непосредственно использован для проверки статистических гипотез. Значение Ф варьируется от –1 до 1 и не зависит от объема выборки *N*. Практически эквивалентным ему является широко известный критерий $\chi^2 = u_{hyp}^2 N/(N-1)$, недостаток которого в данном случае заключается в его положительности. Отрицательное значение u_{hyp} удобно для интерпретации результатов, поскольку показывает низкий уровень встречаемости вида в ценотической единице, меньший, чем это было бы в случае равновероятного распределения.

Рассмотрим результаты оценки характерных и дифференцирующих видов макрозообентоса для трех участков речной экосистемы Байтуган–Сок на основе представленных индексов верности (табл. 6.5). Здесь N = 147 – общее число выполненных гидробиологических проб.

Для первого выделенного участка р. Байтуган, относящегося к эталонным водным объектам, характерные (или индикаторные) виды распределились в соответствие с их реофильностью и оксифильностью, с одной стороны, и биотопической приуроченностью видов к обитанию на специфических каменистых, растительных и песчанистых субстратах – с другой (*Nemoura cinerea, Elmis aenea, Pristina rosea, Macropelopia nebulosa, Pseudodiamesa nivosa* и др.).

В состав дифференцирующих видов, выделенных для этого же участка в правом столбце табл. 6.5, также вошла композиция из реобионтных и оксибионтных таксонов, характерных для «эталонно чистых» биотопов: *Monodiamesa bathyphila, Centroptilum luteolum, Baetis rhodani, Cricotopus* gr. *tremulus, Ephemerella ignita.* С другой стороны, сюда вошли достаточно эврибионтные виды, проявляющие как лимно-, так и реофильные свойства и ценотически ожидаемые в типологически сходных лотических системах с более широким диапазоном загрязнения, например: *Micropsectra* gr. *praecox, Dicranota bimaculata, Eukiefferiella* gr. *gracei, Paracladopelma* gr. *camptolabis, Prodiamesa olivacea, Psectrocladius* sp.

Второй участок верхнего и среднего течения р. Сок имеет иной набор характерных таксонов, где наряду с видами предгорных рек, в качестве индикаторных выделяются виды, толерантность которых к различным экологическим факторам (кислородным условиям, скорости течения, биотопической специфичности и др.) достаточно высока: пиявки *Herpobdella octoculata*, хирономиды *Pseudodiamesa branickii, Potthastia longimanus, Ablabesmyia* sp., олигохеты *Tubifex tubifex,* двукрылые *Dicranota bimaculata.* Сформированный комплекс дифференцирующих видов представлен различными по своим аутэкологическим характеристикам таксонами: эвритопными *Prodiamesa olivacea, Cloeon dipterum, Corynoneura lobata,* литореофильными и эврионсибионтными *Paratanytarsus confusus, Paracladius conversus, Paratendipes albimanus, Brillia modesta,* фитореофильными *Cricotopus bicinctus, Orthocladius* (*Orth.*) *oblidens* и др. (табл. 6.5).

Таблица 6.5

Характерные и дифференцирующие виды макрозообентоса для трех участков речной экосистемы Байтуган–Сок

Характерные виды			Дифференцирующие виды				
Наименование вида	n,n_k	u_{hyp}	**Наименование вида**	n	n_k	u_{hyp}	**Ф**
р. Байтуган – число гидробиологических проб N_1 = 55							
Nemoura cinerea	8	3,9	*Micropsectra* gr. *praecox*	20	16	4,3	34,92
Elmis aenea	7	3,6	*Dicranota bimaculata*	52	29	3,6	28,06
Pristina rosea	6	3,3	*Baetis rhodani*	14	11	3,4	27,59
Macropelopia nebulosa	6	3,3	*Eukiefferiella* gr. *gracei*	21	14	3,1	24,68
Pseudodiamesa nivosa	5	3,0	*Pericoma* sp.	13	9	2,6	20,48
Hydropsyche angustipennis	4	2,7	*Orthocladius* (*O.*) *oblidens*	8	6	2,3	18,63
Diamesa sp.	4	2,7	*Simulium* sp.	29	16	2,3	18,19
Tvetenia sp.	4	2,7	*Prodiamesa olivacea*	41	21	2,3	17,74
Rheocricotopus effusus	3	2,4	*Centroptilum luteolum*	5	4	1,9	16,51
Capnia bifrons	3	2,4	*Cricotopus* gr. *tremulus*	9	6	1,9	15,44
Pseudodiamesa nivosa	3	2,4	*Ephemerella ignita*	11	6	1,3	10,07
Elmis sp.	3	2,4	*Monodiamesa bathyphila*	23	11	1,3	9,27
Cricotopus tibialis	3	2,4	*Eukiefferiella* sp.	7	4	1,2	9,12
Limnodrilus profundicola	3	2,4	*Potamothrix hammoniensis*	5	3	1,1	8,76
Silo pallipes	3	2,4	*Parakiefferiella bathophila*	3	2	1,1	8,73
Nemurella picteti	3	2,4	*Ephemera lineata*	3	2	1,1	8,73
Amphineumura standfussi	3	2,4	*Leuctra fusca*	3	2	1,1	8,73
р. Сок выше с. Исаклы – число гидробиологических проб N_2 = 51							
Herpobdella octoculata	7	3,7	*Paratanytarsus confusus*	18	13	3,6	29,45
Pseudodiamesa branickii	3	2,4	*Cricotopus* (*C.*) gr. *trifascia*	10	8	3,1	25,72
Potthastia longimanus	3	2,4	*Paratendipes albimanus*	17	11	2,8	22,80
Centroptilum sp.	3	2,4	*Monodiamesa bathyphila*	11	8	2,8	22,72
Tubifex tubifex	3	2,4	*Paracladius conversus*	32	17	2,5	20,43
Beckidia zabolotskyi	3	2,4	*Cloeon dipterum*	8	6	2,5	20,31
Dicranota bimaculata	3	2,4	*Prodiamesa olivacea*	41	20	2,2	18,41
Stempellinella sp.	2	1,9	*Aphelocheirus aestivalis*	5	4	2,2	17,86
Ablabesmyia sp.	2	1,9	*Corynoneura lobata*	5	4	2,2	17,86
Lepidostoma hirtum	2	1,9	*Orthocladius* (*O.*) *oblidens*	13	8	2,1	17,57
Diamesa carpatica	2	1,9	*Brillia modesta*	11	7	2,1	17,29
Brillia longifurca	2	1,9	*Cricotopus bicinctus*	56	25	2,0	16,40
р. Сок (нижнее течение) – число гидробиологических проб N_3 = 41							
Chironomus nudiventris	13	5,9	*Caenis macrura*	17	15	5,8	48,66
Lipiniella araenicola	13	5,9	*Polypedilum nubeculosum*	33	22	5,7	46,52
Dicrotendipes nervosus	10	5,1	*Procladius ferrugineus*	38	24	5,6	46,43
Sphaeriastrum (=*Rivicoliana*) *rivicola*	9	4,8	*Cryptochironomus* gr. *defectus*	21	16	5,3	43,97
Unio pictorum	7	4,2	*Polypedilum* (*T.*) *scalaenum*	14	11	4,4	36,66
Gomphus vulgatissimus	5	3,5	*Chironomus anthracinus*	8	7	3,8	31,89
Amesoda scaldiana	3	2,7	*Cladotanytarsus mancus*	34	18	3,7	30,64
Chironomus agilis	3	2,7	*Microchironomus tener*	11	8	3,4	28,43
Dikerogammarus caspius	3	2,7	*Chironomus obtusidens*	12	8	3,2	25,78
Amesoda solida	3	2,7	*Henslowiana dupuiana*	6	5	3,0	25,50
Dicrotendipes notatus	2	1,9	*Ablabesmyia monilis*	6	5	3,0	25,50
Glyptotendipes gripekoveni	2	1,9	*Pisidium amnicum*	8	6	3,0	25,20
Caenis horaria	2	1,9	*Henslowiana henslowana*	8	6	3,0	25,20
Schizorhynchus bilamellatus	2	1,9	*Harnischia fuscimana*	5	4	2,6	21,80
Anodonta piscinalis	2	1,9	*Paralauterborniella nigrochalteralis*	5	4	2,6	21,80

Примечание: n – общая встречаемость вида в гидробиологических пробах; n_k – встречаемость вида на k-ом участке; u_{hyp} – мера верности; Ф – коэффициент сопряженности Браве.

Характерными видами нижнего течения р. Сок являются представители пелофильного ценоза, повсеместно распространенные в крупных равнинных реках и водохранилищах: *Chironomus nudiventris, Lipiniella araenicola, Dicrotendipes nervosus, Sphaeriastrum rivicola, Unio pictorum.* В составе дифференцирующего комплекса можно также дополнительно выделить практически эврибионтный ценоз пело-псаммофильных и фитофильных видов, широко распространенных в водоемах Палеарктики. Эта группировка видов, сходная по отношению к абиотическим факторам и связанная функционально трофическими и топическими взаимоотношениями, включает представителей двукрылых, ручейников и моллюсков *P. ferrugineus, P. nubeculosum, Cryptochironomus* gr. *defectus, Polypedilum scalaenum, Chironomus anthracinus, Cladotanytarsus mancus, Microchironomus tener, Chironomus obtusiden, Ablabesmyia monilis, Caenis macrura, Henslowiana dupuiana, Pisidium amnicum* и др. (табл. 6.5).

Важным этапом анализа полученных результатов является количественная оценка того, насколько обоснованно была выполнена исходная априорная классификация биотопов (в нашем случае – деление водотока Байтуган-Сок на 3 участка) и как велика степень компактности сформированных синтаксонов. Для этого используются различные индексы и критерии: *среднее значение константности* (average values of constancy), *индекс уникальности* (uniqueness of unit), *«хрупкость» классификации* (crispness of classification), *индексы* ISAMIC, PARTANA и *наложения ниш Мориситы*, *коэффициент бисериальной корреляции* и др. (Chytrý, Tichy, 2003; Aho et al., 2008).

В частности, индекс ISAMIC (indicator species analysis minimizing intermediate occurrences; пакет **labdsv** для **R,** Roberts, 2005) вычисляется по формуле $I = 1/s\sum_{i-1}^{s}\left(2\sum_{k}^{g}|C_{ik} - 0{,}5|\right)/g$, где s – число видов; g – число групп; C_{ik} – относительная частота встречаемости i-го вида в k-й группе. Значение индекса ISAMIC более 0,7 свидетельствует о высокой индикаторной способности видов, образующих группы. В случае сравнения 147 проб, взятых на трех участках системы Байтуган–Сок, отмечена высокая уникальность видового состава по индексу ISAMIC: $I = 0{,}934$.

Формула (6.1) и все остальные производные от нее статистики могут быть использованы не только для данных о наличии/отсутствии видов, но и для обработки количественных показателей обилия. Тогда, например, n – суммарная относительная численность вида во всех местообитаниях, а n_k – суммарная относительная численность вида в k-й ценотической единице.

Другой мерой верности, которая объединяет обе группы показателей – инцидентности и обилия вида в той или иной группе описаний, является индекс индикаторной ценности (*IndVal* Indicator Value Index; Dufrêne, Legendre, 1997):

$$IndVal_{jp} = 100 * RA_{jp} * RF_{jp}, \qquad (6.2)$$

где RA_{jp} (relative abundance) – относительное обилие или отношение среднего обилия вида j в группе описаний p к среднему обилию по всем описаниям; RF_{jp} (relative frequency) – относительная частота, т. е. отношение числа описаний, где встретился вид j в группе p к общему числу описаний в группе p. Если интерпретировать биогеографические данные как категориальные, то формула для расчета индекса индикаторной ценности будет иметь вид (Chytrý et al., 2002):

$$IndVal = \frac{n_p(N - N_p)}{nN_p - 2n_pN_p + n_pN} \cdot \frac{n_p}{N_p}.$$

Нетрудно заметить, формула (6.1) для меры верности Брюльхейда аналогична выражению для стандартного параметрического критерия $z = (x - \mu)/\sigma$, что дает формальное основание использовать u для непосредственной проверки односторонней нулевой гипотезы H_0: $n_k \le \mu$, формулируемой как «встречаемость вида в k-ом синтаксоне статистически значимо не превышает ожидаемой частоты при равномерном распределении вида по группам». Если значение u неотрицательно и больше $u_{крит}$, то нулевая гипотеза отверга-

ется, а вид может интерпретироваться как индикаторный для k-й ценотической единицы. В частности, если $u > 1{,}96$, то верность вида является значимой с вероятностью ошибки $p < 0{,}05$. Конкретное значение p-вероятности может быть найдено точным методом Фишера с помощью *таблиц Геннеса* (Гублер, Генкин, 1973), либо с использованием формулы аппроксимации (Chytrý et al., 2002), которая имеет следующий вид: $P(n_k \le \mu) = \sum \frac{n!N_k!(N-n)!(N-N_k)!}{i!N!(n-i)!(N_k-i)!(N-N_k-n+i)!}$, где суммирование ведется по всем $i > n_k$.

При использовании больших биогеографических наборов данных точный метод Фишера может привести к очень небольшим значениям p-вероятности ($10^{-3} – 10^{-100}$), которые трудно использовать в практической работе непосредственно в качестве одной из специальных версий меры верности. Поэтому предлагается (Chytrý et al., 2002) использовать величину $-\log_{10} P\,(n_k \le \mu)$, которая более удобна в работе.

Нам представляется не вполне обоснованным использование p-статистики в качестве обычного показателя оценки верности, например, для дальнейшего многомерного математического анализа. *U-метрика Брюльхейда*, основанная на гиперболическом распределении, и ряд других функционально связанных с ней показателей (*коэффициент ассоциации* Ф, χ^2, *G-статистики*, *модифицированный индекс индикаторной ценности Дюфрена–Лежандра* и др.), образуют целое семейство *коэквивалентных* мер. Использование любой из этих мер приводит практически к одному и тому же списку диагностических видов, упорядоченному по степени верности, но все они обладают несомненным преимуществом перед p-статистикой – более удобным интервалом варьирования и внятным законом распределения (см. максиму А.И. Кудрина [1998]: «… измеряемая величина удачно выбрана, если она хорошо распределена, легка в учете и имеет отношение к сути дела»). Кроме того, само p-значение нельзя интерпретировать как «силу» влияния фактора, поскольку оно не является мерой в математическом смысле, не обладает свойством аддитивности, и не удовлетворяет, как минимум, двум важнейшим принципам теории статистики – принципу правдоподобия и P-постулату (Хромов-Борисов, 2011).

Необходимо также отметить, что корректная проверка статистической гипотезы о значимости верности основана на предположении, что z-критерий представляет собой *независимую случайную величину*, распределенную по нормальному закону с нулевым математическим ожиданием и единичной дисперсией. Предположение о независимости и случайности величины u, даже из самых общих представлений, представляется весьма сомнительным. Эмпирические массивы биогеографических описаний, традиционно изучаемые в экологии, формируются выборочными методами, основанными, как правило, на субъективных планах обследований, недостаточно учитывающих схемы необходимой рандомизации. Для таксонов, входящих в состав описаний, характерны межвидовые взаимодействия, в результате чего образуются различные ассоциации совместно существующих видов, группы-антагонисты и т. д. Наконец, выделение синтаксонов (групп описаний и ценотических единиц) на этапе анализа также носит неслучайный характер. Все это в совокупности не позволяет принять необходимое статистическое предположение о независимости и случайности частот встречаемости видов.

Предположение о нормальном характере распределения разностей $(n_k – \mu)$ также не подтверждается серьезными статистическими исследованиями. Можно вычислить, например, на основе 136 проб макрозообентоса, взятых в реках Байтуган и Сок, распределение 7414 значений верности u для 337 видов по 22 местообитаниям (т. е. станциям наблюдения). Полученный вариационный ряд имеет приблизительно нулевую оценку среднего (0,007) и единичное среднеквадратическое отклонение (1,087). Однако сама форма гистограммы, представленной на рис. 6.1, резко отличается от стандартной гауссианы. Остальные статистики, такие как *критерии Шапиро–Уилка* и χ^2 ($p \approx 0$), асимметрия (3,89), эксцесс (20,96), смещение медианы относительно среднего (-0,266), также противоречат предположению о нормальности распределения u.

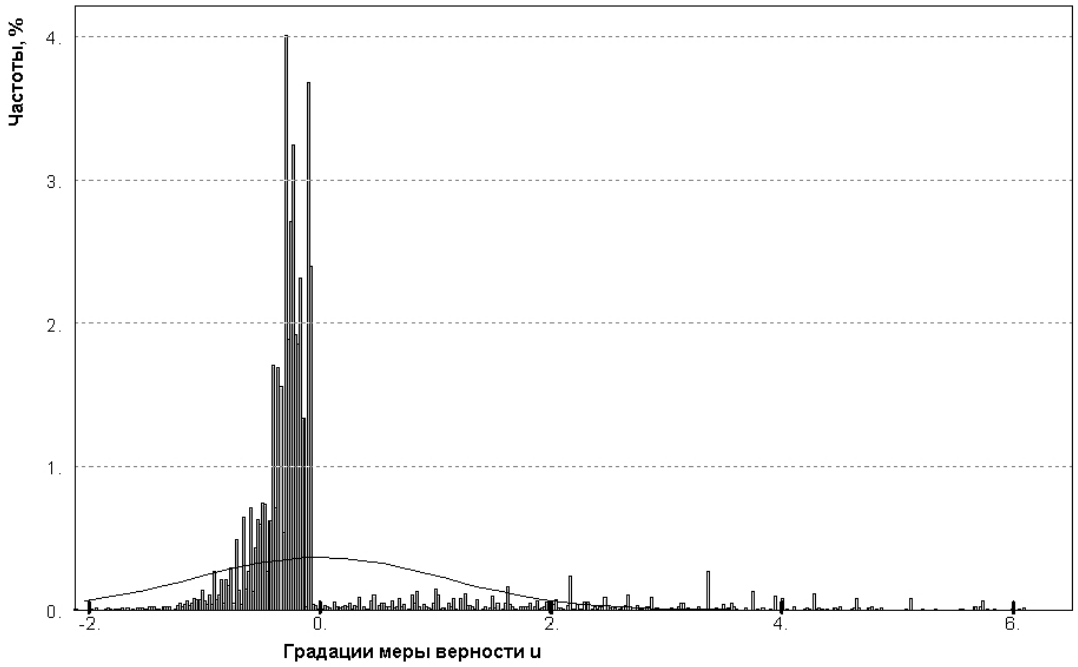


Рис. 6.1. **Гистограмма распределения метрики *u*-верности Брюльхейда для видов макрозобентоса рек Байтуган–Сок**

Представленные аргументы не следует понимать, как рекомендацию не применять при расчетах *p*-статистику верности, основанную на точном методе Фишера. Её использование в пределах каждого предварительно выделенного синтаксона позволяет провести достаточно обоснованное формирование подмножества видов, ранжированных по степени верности. Однако результаты, полученные нами, заставляют предположить, что проверка статистической гипотезы H_0: $n_k \leq \mu$ связана с большой вероятностью ошибки 1-го рода (ложное отклонение нулевой гипотезы). Это приводит к излишне «демократичным» результатам анализа, когда много видов признается верным *k*-й ценотической единице, когда на самом деле это не так.

Несомненно, дальнейшее изучение статистических свойств *u*-метрики верности приведет к учету формы её распределения, что позволит минимизировать ошибку 1-го рода и более обоснованно провести анализ. Сейчас же, чтобы компенсировать некоторую тенденцию к гипердиагностике и сделать анализ более консервативным, используется достаточно практичный прием – снижение уровня значимости ошибочного отклонения H_0 от традиционных для естественных наук $\alpha = 0{,}05$ до 0,001 или менее, что и вызвало не вполне оправданную критику Ю. Денглера (Dengler, 2009). Однако в этом случае принятие уровня значимости $\alpha = 0{,}001$ относится не к теоретико-статистической стороне дела, а является лишь эффективным инструментом, обеспечивающим стандартизированный способ выделения в ранжированном списке видов компактного подмножества наиболее диагностических таксонов, надежно отличая их от сопутствующих и случайных. Например, для проб макрозообентоса из рек Байтуган и Сок уменьшение порога доверия α с 0,05 до 0,001 вызывает уменьшение среднего числа статистически значимых верных видов, приходящихся на один синтаксон, с 22,6 до 12,4, что дает возможность существенно снизить объем последующей неавтоматизированной обработки.

6.3. Группировка видов и местообитаний в биогеографических описаниях

Данные экологического мониторинга, представленные подмножествами таксонов, обнаруженных в пробах из различных биотопов, могут быть геометрически интерпретированы как существенно «размытые» сгущения точек (местообитаний) в многомерном пространстве признаков (видов). Поэтому естественной задачей экологической *синтаксономии* является изучение путей упорядочивания или группировки этих объектов с использованием количественных критериев.

Здесь можно усмотреть дуализм, характерный для многих областей биологии (McCune et al., 2002). Тысячи страниц, написанных экологами, посвящены дебатам о существовании в природе дискретных типов сообществ. С одной стороны, биота демонстрирует бесконечную вариацию своих мелкомасштабных композиций во времени и пространстве, что предполагает большую долю условности любой классификации. С другой стороны, чтобы изучить любой экологический процесс или явление, необходимо:

- обобщить локальные структурные закономерности изменчивости сообществ для выделения границ между ними в среднем масштабе и
- оценить упорядоченность изменений основных характеристик всей экосистемы по градиентам в более крупных пространственных масштабах.

М. Майлз (Miles, 1979; цит. по McCune et al., 2002, p. 80) так обобщил стремление исследователей найти закономерности в геометрической организации данных: «… приходится признать, что любая классификация – есть только рабочая гипотеза, придуманная для конкретного случая (*ad hoc*), но необходимая, чтобы углубить наше научное понимание; её полноценность целиком ограничена спецификой исследования, для которого она была сформулирована».

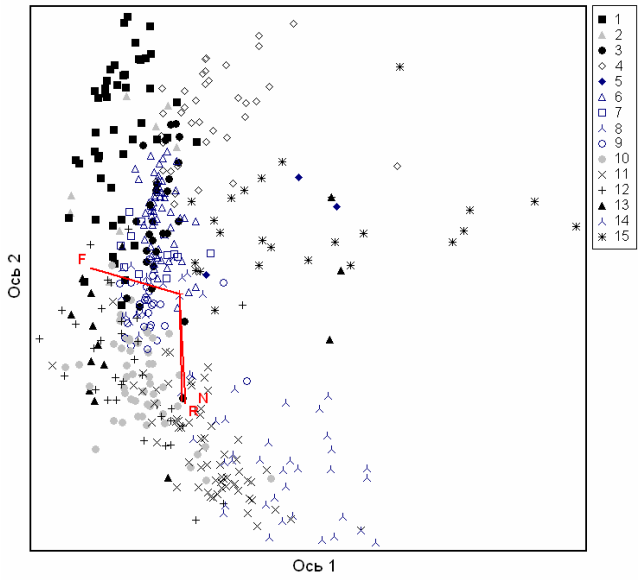
В общем случае, любая группировка объектов (*R-анализ* – объединение видов, *Q-анализ* – объединение сообществ; Миркин, Розенберг, 1978а, 1979) является своеобразной «сверткой» исходных информационных таблиц, поскольку число выделяемых классов всегда меньше, чем уникальных объектов. Таким образом, в итоге получается по возможности *лаконичное*, *наглядное* и *полезное* представление данных в пространстве существенно меньшей размерности. При обработке биогеографической информации используется несколько основных подходов – дискриминация, ординация и классификация.

В ***дискриминантом*** анализе предполагается, что классы заданы своими плотностями вероятностей или обучающими выборками. Задача состоит в том, чтобы на этой основе построить решающие правила, с помощью которых можно было бы отнести к одному из заранее обозначенных классов любой тестируемый объект.

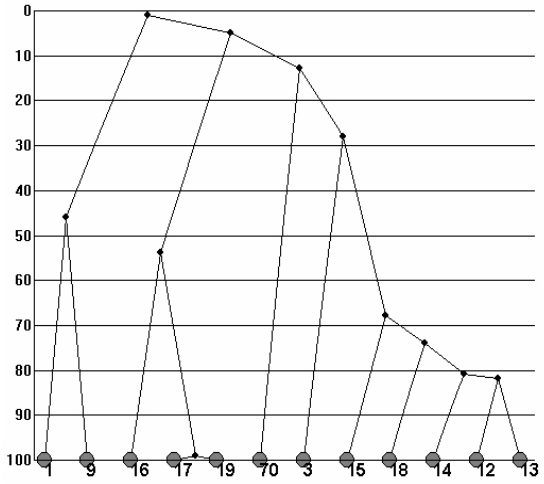
Принцип ***ординации*** заключается в анализе конфигурации и последовательной редукции облака точек многомерного пространства. Для этого используются различные методы оптимального целенаправленного проецирования (projecting pursuit) в пространства малой размерности (с 2 или 3 осями координат). Такой подход, основанный на минимально возможном искажении исходной взаимной упорядоченности точек, обеспечивает наглядное графическое представление геометрической метафоры исследуемых объектов (рис. 6.2а). Подробно эти методы рассматриваются в следующем разделе.

Под ***классификацией*** (*кластеризацией*) понимается задача разбиения всей исходной совокупности элементов на отдельные группы (классы) однородных объектов, сходных между собой, но имеющих отчетливые отличия этих групп друг от друга. Пусть $d(x, y)$ – некоторая мера близости между каждой парой классифицируемых объектов x и y. В качестве таковой может использоваться любая полезная функция: *коэффициенты сходства Жаккара*, *Съеренсена*, *Стугрена–Радулеску*, *евклидово расстояние*, *коэффициент корреляции Пирсона*, *расстояние* χ^2 и многое другое. Любой алгоритм кластеризации может считаться результативным, если при использовании метрики сходства $d(x, y)$ можно найти

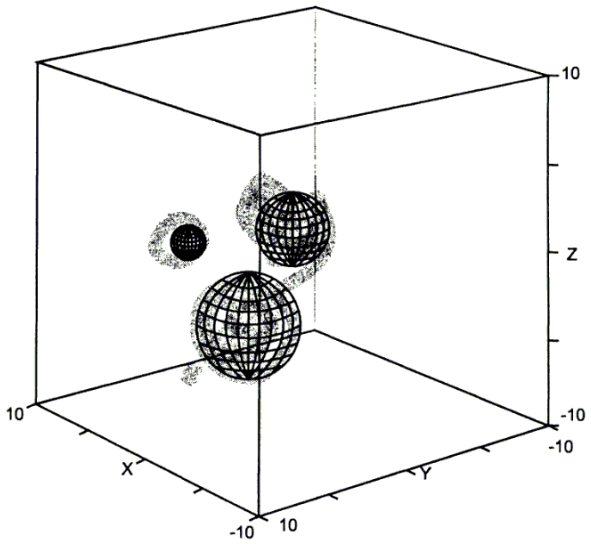
такое разбиение объектов на группы, что расстояния между объектами из одной группы будут меньше ε, а между объектами из разных групп больше ε, где ε > 0 – задаваемый уровень сходства.



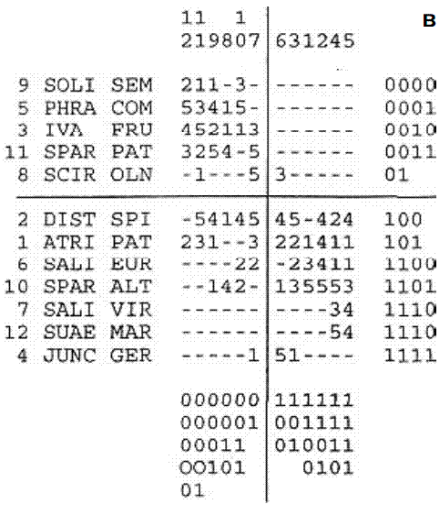
а)



б)



в)



г)

Рис. 6.2. **Различные способы обработки и визуализации биогеографических данных: ординация (а), иерархическая дендрограмма (б), кластеры fuzzy C-Means (в), блочно-диагональное представление исходной таблицы процедурой TWINSPAN (г)**

Наиболее часто применяется так называемый *агломеративный иерархический алгоритм «Дендрограмма»*, отдельные версии которого отличаются правилами вычисления расстояния между кластерами (рис. 6.2б). Дендрограммы просты в интерпретации и предоставляют полную информацию о соподчиненности всех классифицируемых объектов. Однако они имеют содержательный смысл при количестве исходных объектов не более 30-40 (см. обзор: [Миркин, Розенберг, 1979]).

Другим существенным недостатком этого метода является его детерминированность: объединение классов на более высоком уровне полностью определяется результатами агломерации на нижних уровнях. Если исходное множество объектов недостаточно гетерогенно и в определенном смысле «деклассированно», то иерархическая кластеризация дает очень неустойчивые и весьма неожиданные результаты. Деревья начинают резко

менять свой облик при изменении используемой метрики сходства или параметров алгоритма. Каждое из таких разбиений не являются «естественным», поскольку отражает прежде всего свойства вычислительного процесса, а не исходных данных.

Для того чтобы отличить такую ситуацию от противоположной, когда имеются реальные кластеры и алгоритм классификации более или менее точно их выделяет, используют *критерии устойчивости*. Если «реальное» разбиение на диагностические классы возможно, то оно находится с помощью любого алгоритма автоматической классификации. Следовательно, критерием «естественности» классификации может служить устойчивое совпадение результатов работы двух достаточно различающихся алгоритмов, например, «ближайшего соседа» и «дальнего соседа».

Кроме иерархических методов классификации большое распространение также получили различные итерационные процедуры, которые пытаются найти наилучшее разбиение, ориентируясь на заданный критерий оптимизации, не строя при этом полного дерева. Наиболее популярный неиерархический алгоритм – *метод k-средних Мак-Кина*, в котором сам пользователь должен задать искомое число конечных кластеров, обозначаемое как "*k*". Его главного преимущество – возможность обрабатывать очень большие массивы данных, поскольку нет необходимости хранить в памяти компьютера всю матрицу расстояний целиком. Другим современным подходом к кластеризации объектов являются *алгоритмы типа нечетких С-средних* и *Гюстафсона–Кесселя* (Барсегян и др., 2004), которые ищут кластеры в пространстве нечетких множеств в форме эллипсоидов, что делает его более гибким при решении различных задач.

В литературе также описывается множество других методов кластеризации, не использующих матрицы сходств и основанных на оценивании функций плотности статистического распределения, эвристических алгоритмах перебора, идеях математического программирования. Однако, поскольку современные методы дискриминации и классификации широко представлены в специальных руководствах и методиках неэкологического профиля, мы не будем здесь останавливаться на их более подробном описании.

В геоботанических исследованиях для выделения устойчивых сочетаний совместно встречающихся видов на протяжении длительного времени широко использовалась достаточно специфическая классификационная процедура, которая основана на двухсторонней сортировке биогеографической матрицы путем перестановки её строк и столбцов. Основой этих важнейших в практическом отношении методов обработки, в ходе которой размерность исходной матрицы не изменяется, явилась многомерная последовательная процедура, разработанная Ж. Браун-Бланке и его последователями (Braun-Blanquet, 1921).

Общая идея этого метода чрезвычайно проста – необходимо целенаправленно переставлять строки и столбцы исходной матрицы наблюдений до тех пор, пока таблица не приобретет, насколько это возможно, отчетливую блочную структуру (рис. 6.2г). Процесс такой *RQ*-диагонализации биогеографической матрицы приводит к тому, что по её главной диагонали начинают просматриваться прямоугольные блоки клеток, соответствующие встрече групп близких по экологии видов в группах биоценозов сходных местообитаний. Тем самым (в неявной форме) решается известная математическая задача поиска закономерностей плотности распределения объектов в нечетких множествах. Одновременно предлагается способ их оптимального проецирования: локализуются наиболее густые зоны скопления точек и выделяются области разрежения.

Ранние версии табличной обработки по Браун-Бланке были неформализованными и субъективными (Mueller-Dombois, Ellenberg, 1974; Миркин, Розенберг, 1978а), поэтому выделение групп дифференцирующих видов и типов сходных местообитаний в значительной мере зависело от индивидуального понимания экологами характера межвидовых отношений и их связи с экологическими факторами в пределах области исследования.

Однако начиная с 80-х годов прошлого столетия началось широкое внедрение в синтаксономию объективных автоматизированных процедур, основанных на общеприня-

тых многомерных статистических методах и алгоритмах распознавания образов (Gauch, 1982). Полная программа исследований стала включать в себя интегрированное сочетание методов двухсторонней сортировки исходных таблиц с использованием различных версий косвенного и прямого градиентного анализа, многомерного шкалирования и т. д., что дало возможность комплексно изучать и интерпретировать межвидовые взаимодействия и влияние экологических факторов среды. Поэтому формы современной реализации методики Браун-Бланке вряд ли могут быть названы «настоящим рудиментом естественной истории» (Гиляров, 2005, с. 215).

Значительная алгоритмическая завершенность аналитической части метода Браун-Бланке была достигнута уже в компьютерной программе TWINSPAN (Two-Way INdicator SPecies ANalysis) в постановке M. Хилла (Hill, 1979, 1989; Gauch, 1982). Двухсторонний анализ индикаторных видов основан на одновременной группировке строк (видов) и столбцов (местообитаний) и размещении их вдоль главного градиента изменчивости целостного набора данных (McCune et al., 2002).

Алгоритм преобразования таблицы и формирования групп основан на следующих процедурах. Проводится иерархическая группировка местообитаний, которые делятся сначала на 2 класса, потом на 4, 8, 16 и т. д. Каждая отдельная дихотомия основана на трех шагах последовательной процедуры:

- выполняется ординация видов с использованием анализа соответствий (CA – correspondence analysis); дифференцирующие виды располагаются при этом по обеим сторонам от первой оси CA, которая проходит посередине области максимального сгущения «облака» частот распределения видов;
- координаты дифференцирующих видов на главной оси CA используются для скорректированной группировки местообитаний, играя роль взвешивающих коэффициентов; т. е. на этом этапе классификации биотопы ранжируются в соответствии с суммой «весов», вносимых подмножеством видов, характерных для этих экологических условий, и строится дискриминирующая функция, разделяющая эти группы;
- полученная группировка местообитаний и выделенные при этом «виды-индикаторы» используются для следующей дихотомии;
- итеративный процесс заканчивается, если найдены минимальные подмножества видов, которые обеспечивают группировку местообитаний настолько точно, насколько это возможно.

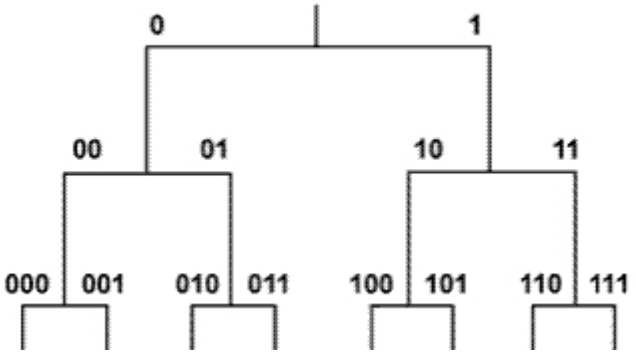
Процедуры TWINSPAN унаследовали многие недостатки метода его прародителя – анализа соответствий (CA). В частности, плохая обусловленность исходных матриц вызывает трудности расчета собственных значений, что приводит к неустойчивости классификаций (Oksanen, Minchin, 1997). TWINSPAN плохо работоспособен при наличии более чем одного главного градиента, что становится особенно очевидным, когда пытаются некорректно интерпретировать подобные таблицы. В отличие от кластерного или факторного анализа, у которых нет никаких ограничений на размерность исходных таблиц, TWINSPAN не всегда может эффективно обрабатывать и представлять иерархические структуры и многомерные гетерогенные наборы данных.

Другая проблема состояла в трудности подстройки алгоритма работы оригинальной FORTRAN-программы путем задания исходных параметров. О смысле многих произвольных, но потенциально важных решений при выборе вариантов расчета ясность наступала лишь после анализа машинного кода программы, поэтому прозрачность других методов кластерного анализа на этом фоне выглядела весьма привлекательной. Однако, после выхода версий программы в среде Windows (**WinTWINS,** Hill, Smilauer), эти трудности несколько потеряли свою остроту.

Двусторонний анализ индикаторных видов для относительно небольших таблиц просто, в традиционной манере и, насколько возможно, сжато выражает весь спектр синэкологических отношений между видами и местообитаниями, чем привлекает сейчас вни-

мание все большего числа сторонников. Модуль TWINSPAN включается в состав известных пакетов обработки биогеографических данных (PC-Ord, Juice), ведутся исследования по модернизации отдельных частей алгоритма. Например, Я. Ролечек с соавторами (Roleček et al., 2009) объединили классический алгоритм TWINSPAN с анализом неоднородности формируемых иерархических деревьев различными методами (с использованием *полной инерции*, *средней меры несходства по Съеренсену* и *Жаккару*, β-*разнообразия Уиттекера* и т. д.). Эта модификация дала возможность более гибко настраивать классификационные процедуры и обоснованно оценивать оптимальную глубину иерархии.

Для каждого вида справа в обработанной таблице приводится дихотомический ключ, определяющий уровень диагностирующей способности вида. Первый разряд этого ключа соответствует первой дихотомии, осуществляющей деление списка видов на две группы – 0 и 1. На втором уровне каждая из этих групп снова делится пополам, при этом формируются группы «безразличных» видов 01 и 10. Третий уровень определяет небольшой список наиболее верных дифференцирующих видов-индикаторов – 000 и 111.



По аналогичной схеме строятся дихотомические ключи и для каждого местообитания, которые помещаются внизу результативной таблицы.

TWINSPAN является полуколичественным методом, использующим уровни «псевдообилия» (pseudospecies cut levels) в порядковой шкале баллов. Если исходные данные представлены в относительных долях, то возможны, например, следующие градации шкалы Браун-Бланке: **r** – от 0 до 1%; **+** – от 1 до 2%; **1** – от 2 до 3%; **2** – от 3 до 13%; **3** – от 13 до 38%; **4** – от 38 до 63%; **5** – от 63 до 100%. Это обстоятельство вызвало негативную реакцию Б. Мак-Кана с соавторами (McCune et al., 2002, p. 97): «Если нам предлагают количественный метод обработки, то почему бы не иметь дело непосредственно с количественными данными, а не [лезть] через черный ход?» По нашему мнению, преобразование показателей обилия в единую систему баллов – прекрасный способ нормирования исходных данных, имеющих не всегда подходящий для статистической обработки характер распределения.

Для анализа продольной изменчивости видовой структуры сообществ макрозообентоса речной экосистемы Байтуган–Сок с помощью программы TWINSPAN был изменен формат обрабатываемых гидробиологических данных. По руслу водотока была выделена последовательность створов, соответствующих станциям наблюдения, а видовой состав объединялся по всем пробам, сделанным в зоне каждого створа. Чтобы обеспечить независимость результатов от выборочного усилия, усреднялось примерно одинаковое количество проб (от 9 до 14). Обилие каждого вида в агрегированных данных рассчитывалось как среднее от натуральных логарифмов численностей особей, что примерно соответствует численным градациям баллов в системе Браун-Бланке. Для дальнейшей обработки было использовано только 214 видов, т. е. из рассмотрения исключалось подмножество из 161 редких и характерных видов, которые встретились только в одном створе (см. фрагмент результатов классификации в табл. 6.6).

Таблица 6.6

Блочно-диагональное представление участков речной экосистемы Байтуган–Сок

Наименования видов	Байтуган				Сок 1		Сок 2			Сок 3				Классы видов
	01	03	06	08	01	02	03	04	05	08	09	12	13	
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Elmis sp.	2	3	3	-	-	-	-	-	-	-	-	-	-	1111
Tvetenia sp.	2	3	3	-	-	-	-	-	-	-	-	-	-	1111
Eukiefferiella claripennis	3	2	2	-	-	-	-	-	-	-	-	-	-	1111
Nemoura cinerea	1	2	2	2	-	-	-	-	-	-	-	-	-	1111
Pristina rosea	3	2	2	3	-	-	-	-	-	-	-	-	-	1111
Elmis aenea	2	3	2	2	-	-	-	-	-	-	-	-	-	1111
Euglesa casertana	-	-	3	2	-	-	-	-	-	-	2	-	-	1101
Centroptilum luteolum	2	-	3	3	-	-	-	-	-	3	-	-	-	1101
Corynoneura sp.	-	2	2	2	-	-	-	-	-	-	2	-	-	1101
Thienemanniella clavicornis	3	2	-	-	-	-	-	-	2	-	-	-	-	1101
Parakiefferiella bathophila	2	-	3	-	-	-	-	-	2	-	-	-	-	1101
Ephemerella ignita	3	2	3	3	-	-	3	2	-	-	-	-	-	1101
Orthocladius (*Orth.*) *clarki*	3	2	3	-	-	3	3	-	-	-	-	-	-	1101
Brillia modesta	2	2	2	-	1	-	2	-	-	-	-	-	-	1101
Pericoma sp.	1	2	3	-	2	-	-	2	2	-	-	-	-	1100
Psectrocladius sp.	-	1	2	2	-	3	2	2	-	-	-	-	-	1100
Orthocladius (*Pog.*) *oblidens*	3	-	-	2	3	2	2	2	-	-	2	-	-	1100
Parametriocnemus lundbecki	2	2	2	2	2	-	2	2	2	-	-	-	-	1100
Orthocladius sp.	2	2	2	3	-	3	-	2	3	-	2	-	-	1100
Eukiefferiella gr. *gracei*	2	2	2	3	-	3	2	3	2	-	-	-	-	1100
Micropsectra sp.	-	2	-	-	2	3	2	2	-	-	-	-	-	1100
Odontomesa fulva	-	1	2	2	1	3	2	-	-	-	-	-	-	1100
Heterotrissocladius gr. *marcidus*	2	2	-	-	2	3	-	2	-	-	2	-	-	1100
Paracladopelma gr. *camptolabis*	2	1	1	2	2	3	1	2	-	-	2	-	-	1100
Culicoides sp.	2	2	2	2	2	2	1	1	2	-	1	1	1	1100
Limnodrilus sp.	2	-	-	-	-	3	-	-	-	-	-	-	3	1100
Potamothrix hammoniensis	-	2	2	-	2	-	-	-	-	-	-	-	3	1100
Eukiefferiella sp.	2	2	2	-	-	-	2	2	-	-	-	-	-	1011
Eukiefferiella claripennis	2	2	2	2	-	-	2	2	2	-	-	-	-	1011
Procladius choreus	2	2	-	-	3	2	-	-	-	2	-	2	2	1010
Limnophilus flavicornis	-	2	2	-	-	2	-	2	-	-	-	-	2	1010
Micropsectra gr. *praecox*	1	1	2	-	2	2	2	-	-	-	-	-	3	1010
Parametriocnemus sp.	2	2	-	2	-	2	3	2	-	-	-	-	-	1001
Ephemera vulgata	2	3	-	2	-	2	2	-	2	-	-	-	2	1001
Prodiamesa olivacea	2	1	1	1	1	2	1	2	3	-	-	-	-	1001
Boophthora erythrocephala	-	2	2	2	2	-	2	2	3	-	2	-	-	1001
Monodiamesa bathyphila	2	1	2	1	-	2	1	2	2	3	-	-	-	1001
Simulium sp.	1	1	2	2	-	3	1	2	2	2	2	-	-	1001
Cricotopus gr. *tremulus*	1	2	3	-	-	-	3	3	-	2	-	-	-	1001
Baetis gr. *rhodani*	2	2	2	2	-	-	-	3	2	-	3	-	-	1001
Tanytarsus pallidicornis	-	-	-	-	2	3	-	2	-	-	2	-	2	1000
Rheocricotopus fuscipes	3	2	-	-	-	-	2	3	2	-	2	-	-	0111
Monodiamesa bathyphila	-	2	-	-	-	3	2	2	2	-	-	-	-	0110
Cricotopus. (*Isocladius*) *trifasciatus*	-	2	-	-	-	2	2	2	2	-	-	-	-	0110
Chironomus plumosus	3	-	-	-	2	-	2	2	-	-	-	-	2	0110
Culicoides grisescens	-	2	2	-	2	-	2	-	3	-	3	-	-	0110
Paratendipes albimanus	2	-	-	2	3	3	2	2	3	2	-	2	-	0110
Procladius ferrugineus	2	-	-	2	-	3	2	3	2	2	1	1	1	0100
Euglesa sp.	-	2	3	-	-	-	-	2	2	2	2	-	2	0100
Chironomus sp.	2	2	-	-	-	-	2	-	-	2	-	2	2	0100

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Stempellina bausei	-	-	-	2	-	2	-	-	-	2	2	1	2	0100
Dicranota sp.	-	-	-	-	2	-	2	3	-	-	-	-	-	0011
Centroptilum sp.	-	-	-	-	2	-	3	-	2	-	-	-	-	0011
Pisidium sp.	-	2	-	-	-	-	2	-	3	-	-	-	-	0011
Cloeon dipterum	2	-	-	-	-	-	2	2	3	-	-	2	-	0011
Haliplus ruficollis	-	2	2	-	-	-	2	3	2	2	2	-	-	0011
Cricotopus gr. *sylvestris*	3	-	-	-	2	-	3	3	3	-	2	-	2	0011
Sialis lutaria	-	-	-	-	2	-	2	-	2	2	-	-	-	0011
Psectrocladius sordidellus	-	-	-	-	2	-	-	-	2	-	-	-	3	0010
Popypedilum convictum	-	-	-	-	-	2	2	2	-	-	2	-	2	0010
Pisidium inflatum	-	2	-	-	-	-	-	-	2	-	2	-	1	0010
Stictochironomus crassiforceps	-	-	2	-	-	-	2	2	-	-	2	2	-	0010
Baetis sp.	2	2	-	-	-	-	2	2	2	2	2	-	-	0010
Nanocladius bicolor	1	-	2	-	-	-	2	3	-	-	2	1	2	0010
Micronecta sp.	3	-	-	-	-	-	-	3	-	2	2	-	2	0010
Paratanytarsus sp.	-	-	-	-	-	2	2	-	-	2	2	-	2	0010
Polypedilum nubeculosum	-	-	-	-	-	2	1	2	2	2	2	1	1	0010
Cladotanytarsus gr. *mancus*	-	-	2	-	-	2	2	1	2	3	2	1	1	0010
Cheumatopsyche lepida	-	-	-	-	-	-	2	2	2	-	2	-	-	0001
Stempellina sp.	-	-	-	-	-	-	2	2	-	-	2	-	3	0001
Pisidium amnicum	-	-	-	-	-	-	-	3	2	-	2	2	2	0001
Cryptochironomus gr. *defectus*	-	-	-	-	-	-	2	2	2	2	1	1	2	0001
Paratanytarsus intricatus	-	-	-	-	-	-	-	3	-	-	2	-	3	0000
Chironomus anthracinus	-	-	-	-	-	-	3	-	-	-	-	2	2	0000
Thienemannimyia sp.	-	-	-	-	-	-	2	-	1	-	2	3	-	0000
Piscicola geometra	-	-	-	-	-	-	2	-	2	-	2	2	2	0000
Polypedilum scalaenum	-	-	-	-	-	-	-	2	2	3	1	1	2	0000
Microchironomus tener	-	-	-	-	-	-	-	3	2	3	2	1	1	0000
Chironomus obtusidens	-	-	-	-	-	-	2	3	-	3	2	2	2	0000
Zavrelia pentatoma	-	-	-	-	-	-	-	-	-	2	3	-	3	0000
Paralauterborniella nigrohalteralis	-	-	-	-	-	-	-	-	2	3	3	2	-	0000
Caenis macrura	-	-	-	-	-	-	-	-	2	1	2	1	2	0000
Euglesa acuminata	-	-	2	-	-	-	-	-	-	1	2	-	2	0000
Ablabesmyia monilis	-	-	-	-	-	2	-	-	-	2	2	2	2	0000
Henslowiana henslowana	-	-	2	-	-	-	-	-	-	3	2	2	2	0000
Классы участков	**00**	**00**	**00**	**00**	**01**	**01**	**10**	**10**	**10**	**11**	**11**	**11**	**11**	

В результате синэкологической обработки была получена классификация створов на четыре группы 00, 10, 01 и 11, которая в целом совпадает с использованным нами априорным делением водотока на три участка (рис. 1.7). В то же время в верхнем течении р. Сок было выделено дополнительно две группы станций (1-2 и 3-5), что свидетельствует о биотопической неоднородности этой зоны экосистемы.

Группировка видового состава была осуществлена по 4 иерархическим уровням. Представленный анализ разбиения видового состава является достаточно корректным с содержательной точки зрения оценки продольного распределения донных сообществ.

Отчетливо выделяются группы видов 1111 и 1101, относящихся исключительно к таксоценозу реофильных и чистоводных видов, например: *Nemoura cinerea, Pristina rosea, Elmis aenea, Orthocladius* (*Orth.*) *clarki, Brillia modesta, Tvetenia* sp., *Eukiefferiella claripennis Centroptilum luteolum, Euglesa casertana*, характерных для предгорной р. Байтуган. В состав группы 0000 и 0001 входят таксоны разных систематических групп, обитание которых приурочено в основном к медленно текучим, умеренно загрязненным водам: это хирономиды *Zavrelia pentatoma Chironomus obtusidens, Ablabesmyia monilis, Polypedilum scalae-*

num, *Microchironomus tener*, *Chironomus anthracinus*, *Cryptochironomus* gr. *defectus*, моллюски *Henslowiana henslowana*, *Pisidium amnicum*, *Euglesa acuminata*, локализующиеся преимущественно в эвтрофных участках нижнего течения р. Сок.

В то же время комплекс видов, характерных для типологически сходных рек и относительно равномерно распределенных по всему продольному течению системы водотоков (рео- и эвриоксибионтные хирономиды *Monodiamesa bathyphila*, *Prodiamesa olivacea*; поденки *Ephemera vulgata Baetis* gr. *rhodani*, *Cloeon dipterum*; жуки *Haliplus ruficollis*, мошки *Boophthora erythrocephala* и моллюски *p. Pisidium*) оказался размещенным в промежуточных группах от 1001 до 0011.

Параметры, задаваемые TWINSPAN, позволяют регулировать состав и количество формируемых групп. Например, чтобы уменьшить число ценотических единиц, состоящих из небольшого числа редких видов могут быть заданы пороговые значения верности. Однако непосредственно обозначить априорные группы видов или местообитаний исследователь уже не может – все процедуры классифицирования осуществляются только автоматически.

Алгоритм TWINSPAN (Hill, 1979) и кластерный анализ относятся к методам неконтролируемой классификации, основной недостаток которых в том, что результат группировки всегда уникален для данного набора данных. Если исходная матрица была частично изменена (например, после добавления новых проб), или выбраны несколько иные настройки алгоритма, то состав выделенных групп может серьезно измениться.

Задать независимые от обрабатываемого набора данных априорные критерии того, на что должны быть похожи выделяемые группы видов, можно при использовании контролируемых классификаций. Так, в отличие от метода TWINSPAN, алгоритм COCKTAIL (Bruelheide, 2000; **JUICE,** Tichy, Holt) основан на предварительном выборе небольшой стартовой группы видов, что, в определенной степени, определяет общие итоги построения композиций видов. В процессе последующей шаговой статистической процедуры к стартовым группам добавляются новые виды на основе некоторого принятого критерия, например, по их *u*-значению верности. Следующая фаза позволяет регулировать состав формируемых синтаксонов путем фильтрации данных на основе формальных логических высказываний, включающих операторы И, ИЛИ, НЕТ. Эти «фильтры» определяют, какие комбинации видов должны присутствовать или отсутствовать в описании, чтобы отнести его к той или иной группе. Размеры формируемых групп определяются пороговым *u*-значением, выбираемым пользователем: чем ниже порог, тем больше видов включается в группы.

6.4. Многомерные методы ординации

Ординация в буквальном смысле – нахождение таких координатных осей, вдоль которых можно выполнить целенаправленное графическое упорядочивание (нем. *ordnung*; Goodall, 1953) анализируемых объектов, т. е. видов или местообитаний. «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).

Проблемы кластер-анализа и ординации связаны с поиском наиболее «естественных» экологических закономерностей. Классификация обычно является успешной, если местообитания образуют локальные обособленные группы вдоль некоторого главного градиента факторов, но приводит к неустойчивым результатам, если структура сообщества изменяется закономерно и непрерывно. Учитывая континуальную природу экосистем, ординацию можно было бы считать более «естественным» подходом. Методы ординации особенно эффективны при наличии комплекса взаимно скоррелированных влияющих фак-

торов, «которые являются кошмаром, например, для множественной регрессии» (McCune et al., 2002, p. 102).

Одно из критических замечаний в отношении методов ординации состоит в том, что её пользователи имеют тенденцию управлять получаемыми результатами, чтобы они лучше соответствовали их предвзятым мнениям. Другая, связанная с этим, критика – ординация выявляет только очевидные тенденции, которые были известны и без её применения. Эти замечания, безусловно, имеют свои резоны в отношении к большинству статистических инструментов экологии, однако ординация во многих случаях помогает исследователям более четко увидеть закономерности в анализируемых данных, например:

- выбрать самые важные влияющие факторы из всего множества предполагаемых или вероятных;
- выявить не только главные тенденции, но и идентифицировать возможные непредвиденные процессы в динамике изучаемых экосистем.

История развития методов ординации имеет следующие основные этапы их возникновения и начала применения в экологии (цит. по: Jongman et al., 1987, раздел 5.8):

- анализ главных компонент (PCA, principal component analysis; Pearson, 1901; Hotelling, 1933; Goodall, 1954);
- прямой однофакторный градиентный анализ (Раменский, 1956; Whittaker, 1967; Миркин, 1971; Миркин, Розенберг, 1978а);
- анализ соответствий и алгоритм встречного осреднения (CA, correspondence analysis, RA, reciprocal averaging; Hirshfield, 1935; Curtis, McIntosh, 1951; Roux, 1967);
- полярная ординация Брея–Кёртиса (Bray, Curtis, 1957);
- неметрическое многомерное шкалирование (NMS, nonmetric multidimensional scaling; Shepard, 1962; Anderson, 1972);
- бестрендовый анализ соответствия (DCA, detrended correspondence analysis; Hill, 1979);
- анализ избыточности или каноническая форма PCA (RDA, redundancy analysis; Rao, 1964; ter Braak, 1986);
- канонический анализ соответствия (CCA, canonical correspondence analysis; ter Braak, 1986);
- различные версии анализа полиномиального CCA/RDA, главных координат PCoA, методы, основанные на дистанциях dbRDA, и др. (P. Legendre, L. Legendre, 1998).

Для ординации используется значительное количество различных методов или их модификаций, основные из которых представлены на рис. 6.3. Формальная сущность ординации заключается в представлении исходной биогеографической матрицы **Y** в виде совокупности p (обычно $p = 2$) латентных переменных **F:** $Y_1, Y_2, \ldots, Y_m \Rightarrow F_1, F_2, \ldots, F_p$, которые и являются осями ординационной диаграммы (или биплота – biplot). Выбор этих осей осуществляется с использованием принципа оптимальности: стремления достичь минимума потерь содержательной информации, находящейся в исходных данных.

Как правило, первая ось F_1 проводится через центр сгущения значений y_{ij} и совпадает с направлением наибольшей по длине оси эллипсоида рассеяния. Вторая ось F_2, также проходящая через центр распределения, проводится перпендикулярно к первой и совпадает по направлению со второй из главных полуосей эллипсоида рассеяния. Эта операция обеспечивает формирование «картинки» данных с минимально возможными искажениями, а новые обобщающие переменные F_1–F_2 становятся ортогональными (т. е. взаимно некоррелированными, независимыми).

«Идеальный» метод ординации пока еще не разработан, но, если бы он существовал, он бы обладал следующими свойствами (McCune et al., 2002):

- близость объектов (т. е. композиций видов и местообитаний) в ординационном пространстве F_1–F_2 тесно связана с их экологическим подобием;
- главные градиенты ценоклина отображаются без искажений, а принятый масштаб осей определяется β-разнообразием сообществ;

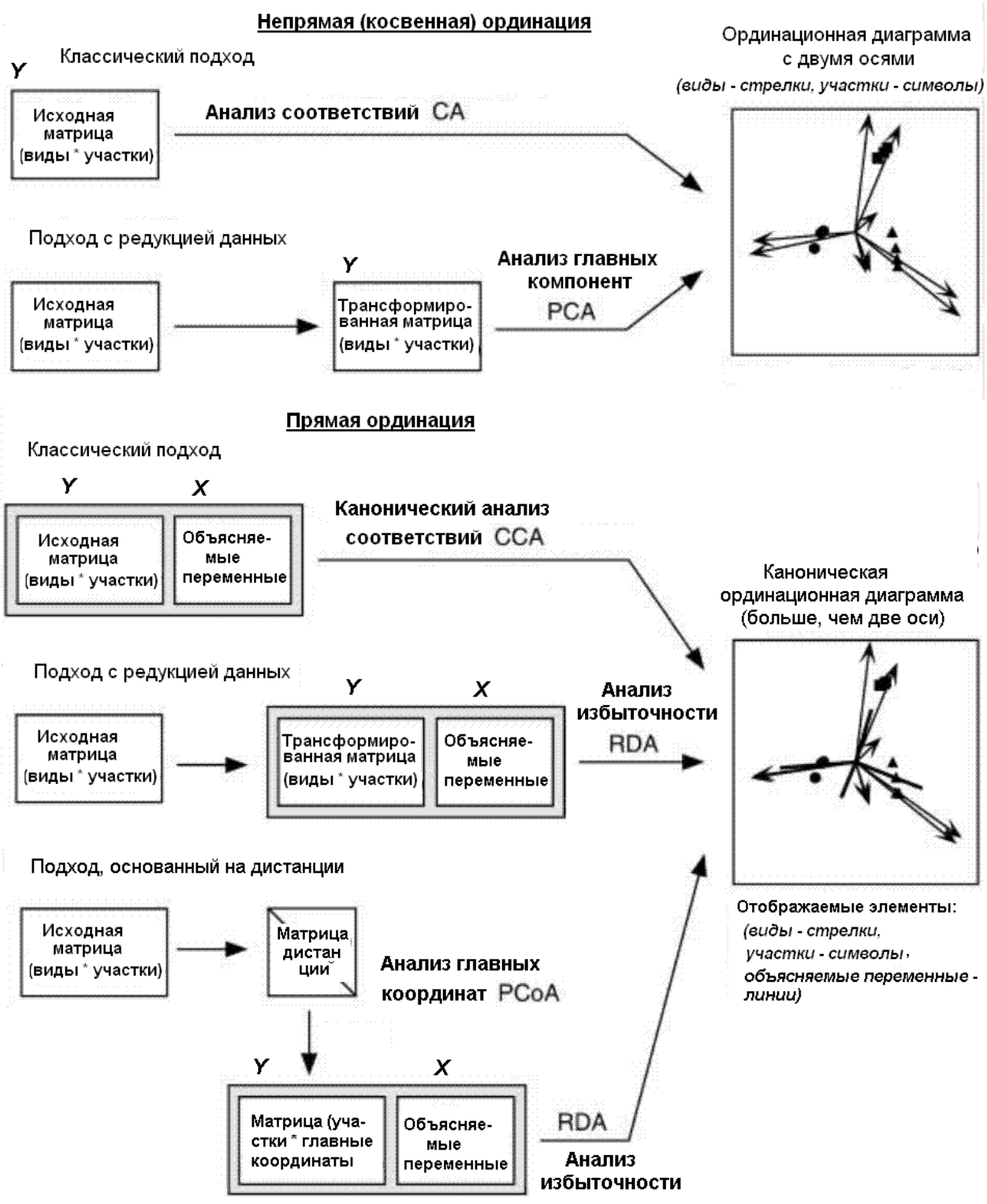


Рис. 6.3. **Схема сравнительного анализа основных методов ординации (Legendre, Gallagher, 2001)**

- если в изучаемой экосистеме объективно существует возможность группировки, то это находит отражение при ординации, однако не существующие псевдо-кластеры на диаграмме не обнаруживаются;
- метод не чувствителен к «шуму», однако «сигнал» и «шум» надежно отличаемы;

- алгоритм является результативным и воспроизводимым в различных условиях, т. е. одинаково работоспособен для:
  - «коротких» и «длинных» градиентов,
  - разреженных и заполненных матриц,
  - наборов данных различной размерности,
  - редких и массовых видов и т. д.;
- реализуем необходимый набор пользовательских настроек для задания числа осей ординации, взвешивания видов и местообитаний и т. д.;
- метод математически элегантен, а для экологов – доступен и легок в понимании.

Практически важным является разграничение методов *прямой* и *непрямой* (*косвенной*) ординации. Непрямая ординация использует только стандартные биогеографические матрицы **Y**, не учитывает внешние факторы среды и показывает изменение видового состава вдоль некоторой абстрактной оси, которая отражает максимальную изменчивость в структуре данных. При наличии информации о факторах окружающей среды **X** (на рис. 6.3 они названы «объясняемыми переменными»), проводится прямая ординация, которая чаще всего используется в дополнение к косвенной ординации, как некоторый инструмент интерпретации и объяснения.

Следует отметить, что упомянутый выше прямой однофакторный градиентный анализ формально не является методом ординации, а лишь обычным графиком зависимости обилия анализируемого вида от произвольно выбранного фактора среды, что часто не принимается во внимание. Поэтому высказываемые соображения, что «прямая ординация является довольно слабым методом, так как выбор факторов, служащих осями ординации, субъективен» (Миркин, 1971; Василевич, 1972; Галанин, 2005), никак нельзя отнести собственно к ординации, поскольку адресованы они, вероятно, прямому однофакторному градиентному анализу.

Можно также различать методы ординации,

- основанные на дистанции (т. е. матрице расстояний), и
- использующие вычисление собственных значений.

Это различие несколько произвольно, потому что, например, анализ главных координат (PCoA) может быть выполнен с использованием собственных векторов и, наоборот, методы, основанные на собственных значениях, могут интерпретироваться в контексте метрик расстояния: корреспондентный анализ (CA) часто описывают в терминах χ^2.

Кривые распределения обилия видов могут иметь *монотонный* характер, когда имеющиеся значения y_{ij} только увеличиваются или уменьшаются вдоль экологических градиентов. Монотонное распределение может быть линейным или иметь более сложную функциональность. В случае *унимодальных* кривых распределения, имеющих колоколообразную форму гауссиан, для видов характерен некоторый условный оптимум комбинации факторов окружающей среды. Это различие также довольно условно, потому что монотонное распределение можно интерпретировать как частный случай унимодального распределения, когда имеющаяся выборка данных охватывает лишь небольшую часть объективно существующего «длинного» градиента. В соответствие с этим, методы, использующие собственные значения, подразделяются на линейные (PCA, RDA) и унимодальные модели (DCA, CCA). Последние лучше использовать, если распределение большинства видов имеет унимодальный характер, хотя модели DCA, CCA достаточно хорошо работоспособны и с линейными данными (описание **CANONO**, ter Braak, Smilauer, 1998).

Анализ собственных значений и собственных векторов является важнейшим разделом линейной (матричной) алгебры, поэтому полное понимание методов ординации требует определенной математической подготовки. В рамках нашего изложения отметим только следующие основные положения:

- нахождение собственных значений и собственных векторов осуществляется в ходе математической операции линейного преобразования квадратной симметричной матрицы

R размерностью m; примерами такой матрицы, симметричной относительно главной диагонали, являются корреляционная и ковариационная матрица, а также любая другая матрица дистанций;

- собственными значениями квадратной матрицы **R** называются такие значения λ_k, при которых система m уравнений вида $\mathbf{RV}_k = \lambda_k\mathbf{V}_k$ имеет нетривиальное решение; здесь $\mathbf{V}_k$ – собственные векторы матрицы **R**, соответствующие λ_k; $k = 1, 2, \ldots, m$;
- собственные векторы $\mathbf{V}_k$ представляют собой информационные веса видов, если строки и столбцы матрицы **R** определяют ковариацию между видами, или веса местообитаний, если использовалась ***Q*** – техника свертки;
- каждое собственное значение λ_k соответствует величине дисперсии, объясняемой на k-м уровне; если **R** – ковариационная матрица, то сумма собственных значений будет равняться сумме дисперсий показателей обилия всех видов;
- ряд собственных значений обычно ранжируется от самого большого до минимального; первое собственное значение λ_1, объясняющее наибольшую долю вариации данных, часто называют «доминирующим» или «ведущим»; значения λ_1 и λ_2 определяют меру значимости осей F_1-F_2 ординационной диаграммы, т. е. концентрацию обилия видов вдоль каждой оси и, в итоге, выраженность экологического градиента;
- собственные значения λ_k находятся в ходе итерационной процедуры и точность их вычисления зависит от степени обусловленности исходной матрицы **R**.

Непрямая (косвенная) ординация

Анализ главных компонент (PCA, principal component analysis), как одна из форм факторного анализа, является классическим методом редукции данных, широко используемым в различных областях науки и техники и детально описанным в многочисленных руководствах (ter Braak, 1983; Айвазян и др., 1989). Оси ординационной диаграммы F_1–F_2 рассчитываются как линейные комбинации факторных нагрузок, основанных на первых двух собственных значениях и их собственных векторах. При этом гарантируется ортогональность системы координат (фактически, независимость главных компонент) и наименьшая вариация точек отображаемых объектов относительно осей главных факторов, т. е. при этом вносятся минимально возможные искажения в геометрию исходных данных. Разработанные алгоритмы вращения осей главных компонент позволяют получить удобную для интерпретации факторов ординационную диаграмму.

В то же время PCA является линейным методом, предполагающим монотонный характер изменения обилия видов по шкале градиента. Кроме того, вычисление собственных значений предполагает нормальный характер распределения элементов матрицы **R**, а её вырожденность приводит к неустойчивым и мало достоверным результатам. Это бывает, в частности, при большом количестве редких видов.

Интерпретация PCA-диаграмм основана на учете следующих закономерностей, характерных для большинства методов прямой и непрямой ординации, использующих двухмерное проецирование:

- точки, соответствующие каждому из видов, обозначают координаты их экологических оптимумов; следовательно, среднее обилие или вероятность встречаемости вида уменьшается с увеличением расстояния от его позиции на графике;
- точки, соответствующие местообитаниям, лежат на диаграмме ординации в центре тяжести точек тех видов, которые преимущественно встречаются на этих станциях (аналогично, если точка вида находится близко от точки местообитания, то вероятно высокое обилие этого вида, связанное с этим биотопом);
- точки видов на краю диаграммы часто принадлежат редким видам, либо предпочитающим экстремальные условия среды, либо имеющим низкую встречаемость, случайно совпадающую с местообитаниями, имеющими экстремальные условия.

В качестве примеров, иллюстрирующих различные методы ординации, используем

данные о встречаемости видов макрозообентоса речной экосистемы Байтуган–Сок в гидробиологических пробах, сделанных в зоне каждой из 13 станций наблюдений (см. табл. 6.6). На ординационной PCA-диаграмме (рис. 6.4) можно выделить группы станций и сопутствующие им виды, в целом соответствующие как априорному делению водотока на три участка (раздел 1.3), так и результатам блочной TWINSPAN-диагонализации (раздел 6.3).

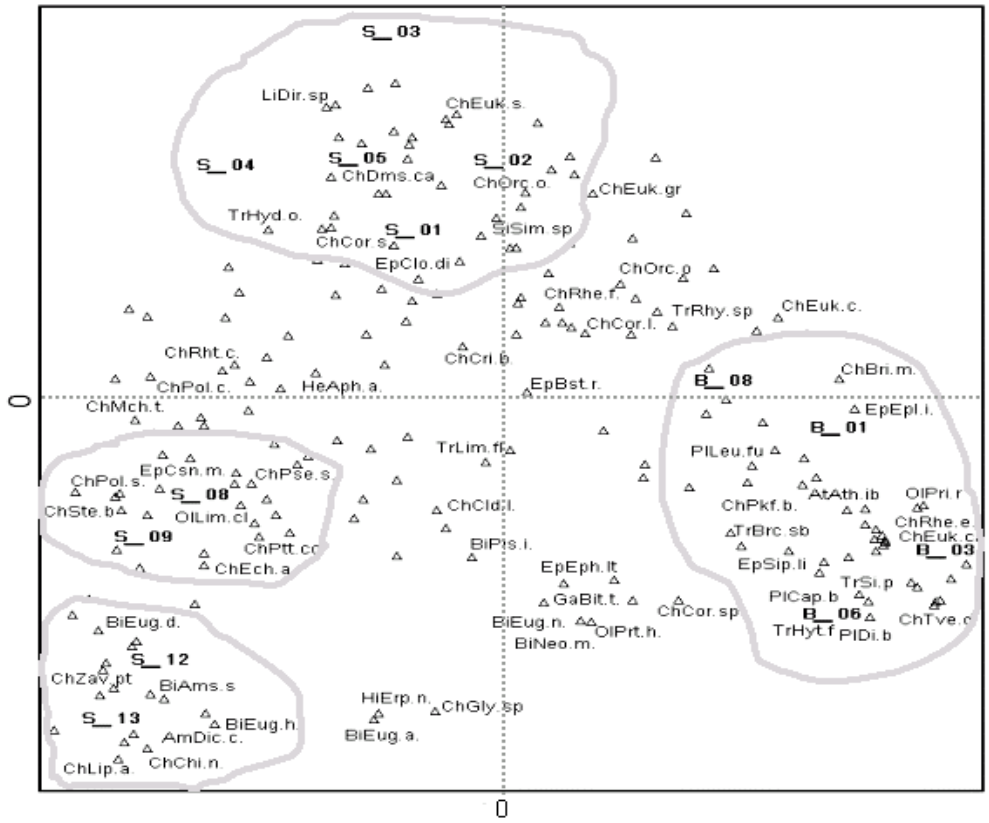


Рис. 6.4. **Пример ординационной диаграммы видовой структуры на станциях рек Байтуган–Сок (метод главных компонент – PCA)**

Анализ соответствий (CA, correspondence analysis), разработанный независимо друг от друга несколькими учеными-экологами (Curtis, McIntosh, 1951; Williams, 1952; Hill, 1973b), основан на процедуре встречного осреднения (или, в другой терминологии, реципрокного взвешивания RA), которая заключается в следующем. Алгоритм стартует с произвольного задания весов для всех местообитаний либо для всех видов. Пусть B_i – некоторые числа, заданные случайно и интерпретируемые как веса местообитаний (i = 1, 2, …, n). Далее для каждого j-го вида (j = 1, 2, …, m) рассчитывается его средний вес: $A_j = \sum_{i=1}^{n} B_i y_{ij} / \sum_{i=1}^{n} y_{ij}$, где y_{ij} – показатель обилия j-го вида в i-м описании, причем если вид отсутствует в i, то его обилие приравнивается нулю. Соответственно, зная веса всех видов, можно пересчитать веса для местообитаний: $B_i = \sum_{j=1}^{m} A_j y_{ij} / \sum_{j=1}^{m} y_{ij}$.

Такие итерации продолжаются до тех пор, пока веса B_i и A_j не стабилизируются. Можно математически строго показать, что итеративный процесс сходится (т. е. стабилизация всегда будет достигнута) и результат не зависит от выбранных начальных значений. В итоге, веса B_i и A_j будут максимально скоррелированны друг с другом (т. е. нельзя надеяться на лучший результат с учетом имеющихся данных), а коэффициент корреляции

между ними будет численно равен собственному значению λ_1. Эти величины и являются координатами местообитаний и видов для первой оси ординации F_1.

Вторая ось F_2 (и, при необходимости, все последующие) строится по аналогичному принципу, но в вычисления вводится дополнительный шаг, направленный на то, чтобы вторая ось была независимой от первой. Таким образом, смысл расчета второй оси – получение дополнительной информации из исходной биогеографической матрицы, которая не объясняется при формировании первой оси F_1.

Анализ соответствий изначально был ориентирован на учет унимодального распределения обилия видов. Однако он обладает и рядом существенных недостатков. Если первая полученная ось отражает влияние одного самого сильного фактора среды, то вторая и последующая оси часто являются искажениями той же самой первой оси, и зачастую не вскрывают влияние других факторов. Кроме того, этот метод сильно преувеличивает влияние редких видов и ординационная картина оказывается не всегда корректной с содержательной точки зрения. При использовании этого метода иногда возникает так называемый «эффект арки», обусловленный квадратичными связями между осями. При этом масштаб осей деформируется и сжимается в области минимальных и максимальных значений, а точки отдельных описаний выстраиваются на графике в дугу.

В нашем случае ординационная CA-диаграмма (рис. 6.5) не имеет принципиальных отличий от ранее полученной по методу PCA. Их сходство становится еще более очевидным, если один из графиков повернуть на 180° относительно центра координат, что не изменяет предупорядоченности точек. Можно отметить, в частности, только более акцентированное взаимное расположение устьевых станций р. Сок (S_12 – S_13).

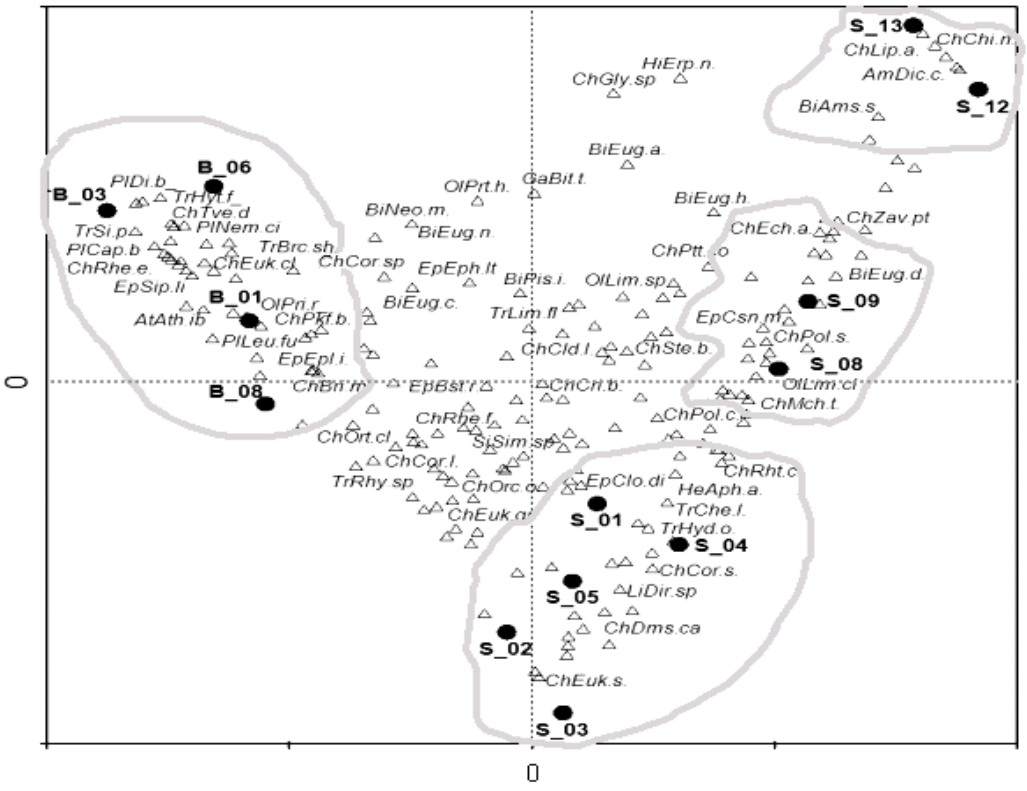


Рис. 6.5. **Пример ординационной диаграммы видовой структуры на станциях рек Байтуган–Сок (анализ соответствий CA)**

Для интерпретации скрытого смысла факторных осей обычно каждой из них приписывают некоторый содержательный смысл, связанный с предметной областью. Чтобы понять, какая гидробиологическая реальность скрыта в найденных факторах, проводится

анализ корреляций факторных нагрузок с исходными переменными. Из табл. 6.7 видно, что ось 1-й главной компоненты, объясняющая около 26% статистического разброса, связана с вариацией видов в верхней части р. Байтуган и нижнем течении р. Сок. Вторая факторная ось, объясняющая около 13% дисперсии, определяется изменчивостью видового состава в среднем течении р. Сок.

Таблица 6.7

Коэффициенты корреляции Пирсона (r) и ранговой корреляции Кендалла (τ) между встречаемостью 214 видов на станциях рек Байтуган–Сок и факторными нагрузками двух главных осей факторов (выделены значения $r > \pm 0{,}45$)

Станции Байтуган–Сок	Метод PCA (рис. 6.4)				Метод CA (рис. 6.5)			
	1 факторная ось		2 факторная ось		1 факторная ось		2 факторная ось	
	r	τ	r	τ	r	τ	r	τ
B_01	**0,508**	**0,390**	0,013	0,051	**-0,402**	-0,334	0,008	-0,009
B_03	**0,726**	**0,565**	-0,175	-0,078	**-0,665**	**-0,550**	0,187	0,127
B_06	**0,553**	**0,425**	-0,274	-0,190	**-0,487**	**-0,414**	0,220	0,193
B_08	0,334	0,261	0,073	0,123	-0,274	-0,240	-0,083	-0,104
S_01	-0,068	-0,026	0,300	0,259	0,063	0,033	-0,183	-0,189
S_02	0,001	0,012	0,436	0,337	-0,027	-0,034	-0,322	-0,291
S_03	-0,078	-0,049	**0,648**	**0,521**	0,044	0,019	**-0,590**	**-0,491**
S_04	-0,329	-0,234	**0,427**	**0,362**	0,228	0,187	-0,354	-0,332
S_05	-0,122	-0,073	**0,450**	**0,364**	0,056	0,038	-0,378	-0,345
S_08	-0,333	-0,260	-0,054	-0,011	0,268	0,233	-0,049	-0,043
S_09	**-0,495**	**-0,399**	-0,172	-0,098	0,394	0,342	0,036	0,025
S_12	**-0,493**	**-0,433**	-0,355	-0,182	**0,484**	**0,411**	0,267	0,165
S_13	**-0,504**	**-0,427**	-0,469	-0,303	**0,472**	**0,435**	0,388	0,299

Анализ соответствия с удаленным трендом (DCA, detrended correspondence analysis), разработанный М. Хиллом и Г. Гаучем (Hill, 1979; Hill, Gauch, 1980), призван устранить главный недостаток классического метода анализа соответствий – «эффект арки». Он отличается от своего предшественника тем, что после вычисления первой оси выполняется дополнительный шаг «детрендирования», в ходе которого ось разбивается на сегменты и в каждом сегменте срезаются отклонения от предыдущего. Такой подход имеет и оборотную сторону. DCA приводит к сглаживанию изменений, связанных с одним из скрытых градиентов, т. е. кроме артефактных «подков» разрушаются и реально существующие тренды в данных, если таковые имеются. Э. Пилоу (Pielou, 1984a, p. 197) отмечала, что корректировка дефектов методом DCA «может иногда привести к непреднамеренному уничтожению информации, имеющей экологическую ценность».

Анализ соответствия с удалением тренда при ординации видов макрозообентоса по продольному градиенту рек Байтуган и Сок (рис. 6.6) привел к более «смазанной» картине группировки, что является свидетельством определенной правоты замечания Пилоу.

Анализ главных координат (PCoA), или метод метрического многомерного шкалирования (Айвазян и др., 1989) в определенной степени абстрагируется от исходной таблицы данных мониторинга и оперирует только с произвольной матрицей расстояний **D**, где d_{ij} – мера дистанции между каждой парой местообитаний i и j, $i, j = 1, 2, \ldots, n$. Д. Робертс в лабораторной работе (http://ecology.msu.montana.edu/labdsv/R/labs/lab8/lab8.html) для студентов университета штата Монтана рекомендует использовать для разных типов данных одну из 16 метрик: χ^2-, евклидово, манхеттенское или хеммингово расстояние, индексы Съеренсена, Кульчицкого, Минковского, Брея–Куртиса, Робертса, Мориситы, Горна, Охайи (Ochiai), Ружички (Ruzicka), Маунтфорда (Mountford) и др.

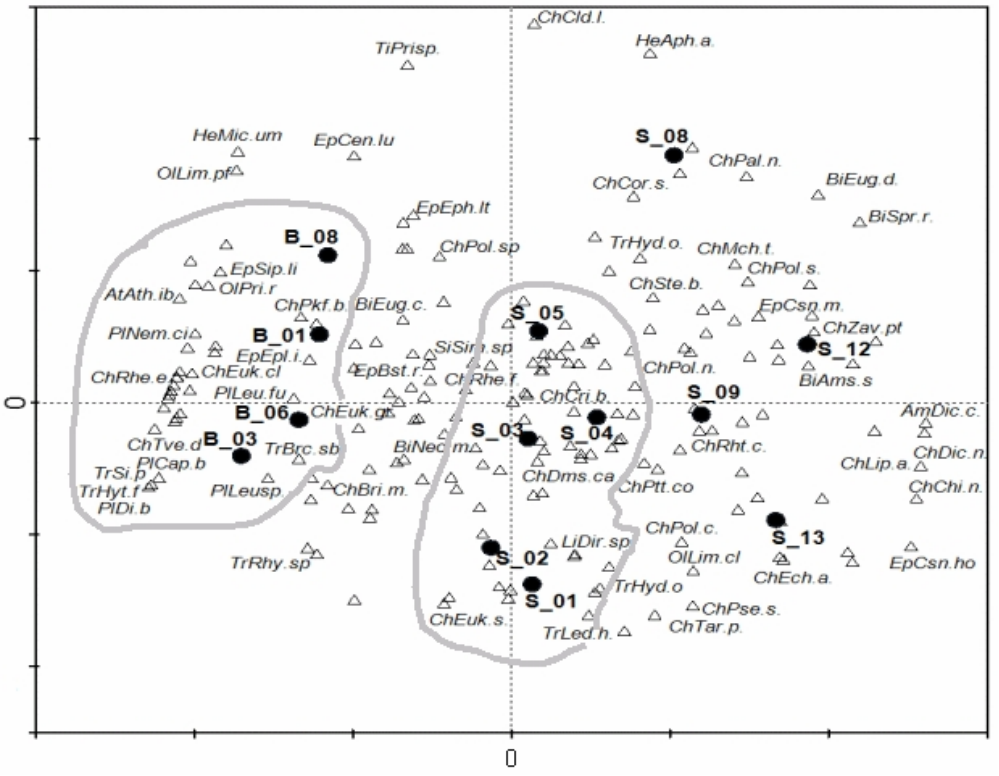


Рис. 6.6. **Пример ординационной диаграммы видовой структуры на станциях рек Байтуган–Сок (анализ соответствий с удаленным трендом DCA)**

Другим перспективным методом ординации, находящим все большее применение в экологии, является алгоритм неметрического многомерного шкалирования (NMS, nonmetric multidimensional scaling), также использующий произвольную матрицу дистанций **D** (Kenkel, 1987; Clarke, 1993; Шитиков и др., 2003). Считается, что этот метод дает наиболее адекватные результаты, особенно для больших биогеографических матриц с сильными шумами (Prentice, 1977; Minchin, 1987). Его главным преимуществом является то, что он не требует от исходных данных никаких априорных предположений о характере статистического распределения.

В общем случае, задача многомерного шкалирования состоит в том, чтобы создать такой *p*-мерный (p = 2 или 3) «образ» наших объектов (видов и местообитаний), в котором взаимные попарные расстояния оказались бы наименее искажены по сравнению с исходным состоянием **D**. Главные оси геометрической метафоры данных в пространстве меньшей размерности обычно находятся путем минимизации критерия «стресса»: $\Delta = \sum_{i,j=1}^{n} d_{ij}^{\alpha} \left| \hat{d}_{ij} - d_{ij} \right|^{\beta}$, где d_{ij} и $\hat{d}_{ij}$ – расстояния между объектами i и j в исходном и редуцированном пространствах, α и β – задаваемые коэффициенты. Отличие между метрической и неметрической модификациями заключается в том, что поиск решения PCoA осуществляется на множестве линейных функций (с точностью до ортогональных преобразований) и основан на операциях с собственными числами и собственными векторами (Айвазян и др., 1989, с. 439). В методе NMS выполняется последовательность итераций для минимизации критерия Δ, оценивающего степень сходства между исходной и моделируемой матрицами расстояний. Детальное описание обоих метододов широко представлено в доступной методической литературе.

Ординация донных сообществ рек Байтуган–Сок путем анализа главных координат матрицы *расстояний Брея–Куртиса*, представленная на рис. 6.7 (точки, обозначающие виды и таксоны макрозообентоса, были нами опущены), не выявила принципиальных отли-

чий от ранее использованных методов с учетом того, что допустимо произвольное вращение ординационных диаграмм относительно центра их координат.

В тех же условиях ординация донных сообществ рек Байтуган–Сок методом неметрического шкалирования (рис. 6.8) дала, на наш взгляд, наиболее реалистическую картину, которая в целом воссоздала (с использованием только видового состава!) геометрический контур самого водотока.

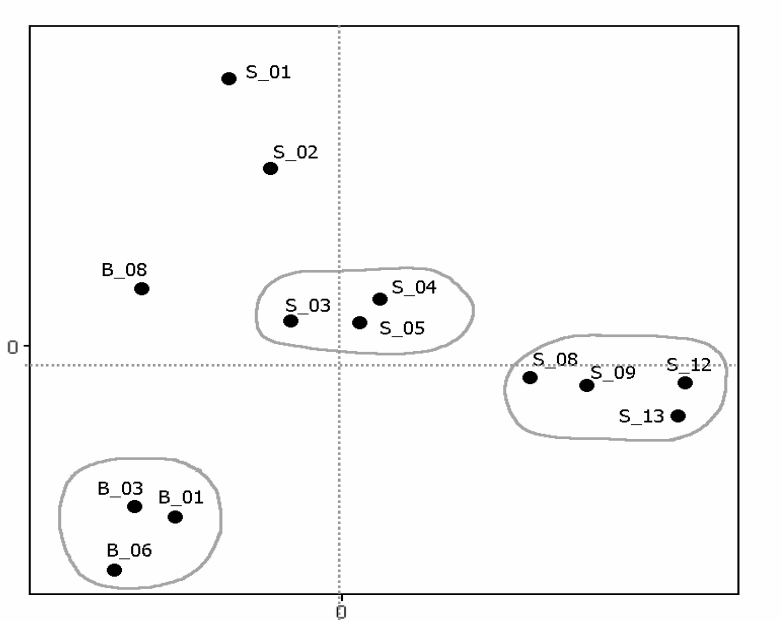


Рис. 6.7. **Пример ординационной диаграммы станций рек Байтуган–Сок (анализ главных координат PCoA)**

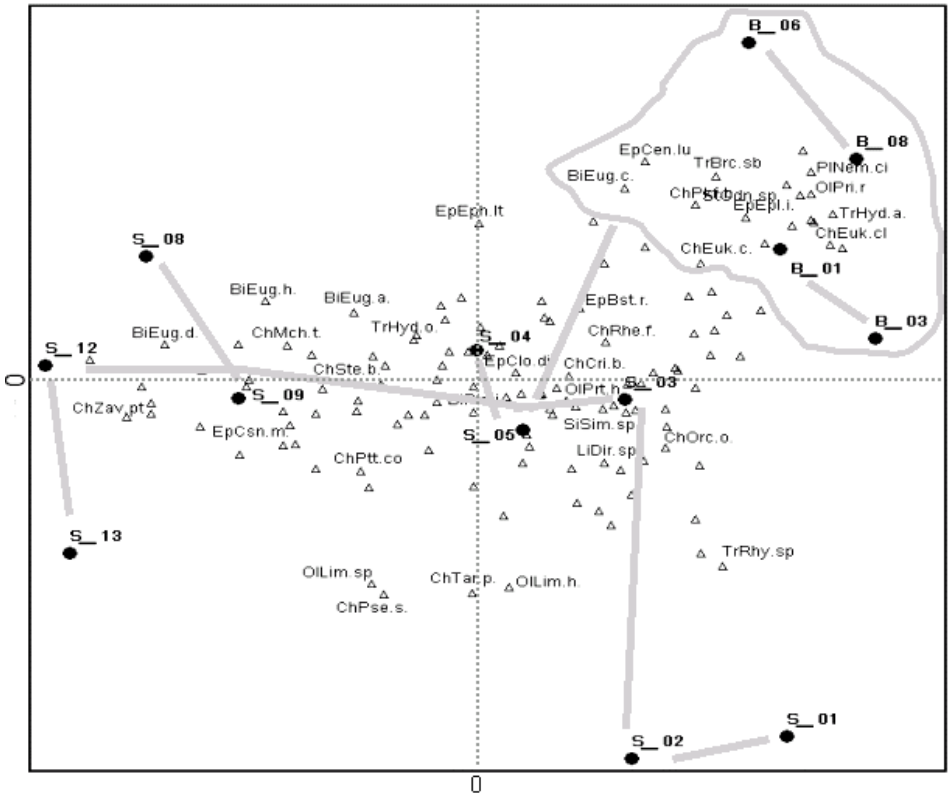


Рис. 6.8. **Пример ординационной диаграммы видовой структуры на станциях рек Байтуган–Сок (метод многомерного шкалирования NMS)**

Можно отметить некоторые специфические отличия ординаций, выполненных различными методами: например, с разной степенью отчетливости показывается расстояние между кластерами {S_1 – S_2} и {S_3 – S_5} или {S_8 – S_9} и {S_12 – S_13}. Несколько обособленное положение станций B_6 и B_8 в нижней части р. Байтуган, обозначенное на рис. 6.8, также объясняется специфическими биотопическими условиями на этом участке.

Найденные ординационные шкалы по отношению к водным экосистемам могут также рассматриваться как результат определенного типа взаимоотношений между совместно обитаемыми видами. Интерпретация осей биплота, как и в случае методов PCA (CA), осуществляется в соответствии с коэффициентами корреляции между натуральными показателями изменчивости видов и их координатами относительно выбранных шкал.

Например, все виды, представленные в табл. 6.8 и в наибольшей степени определяющие ординационную картину, можно разделить на несколько групп: связанные отрицательной или положительной связью только с одной шкалой 1 либо 2, а также определяющие положение обоих шкал одновременно. Например, отрицательные значения шкалы 1 связаны с эврибионтными таксонами, характерными для нижнего течения р. Сок, в то время как положительные значения определяются реофильной группой. Виды этой группы определяют также положительное направление шкалы 2.

Таблица 6.8

Коэффициенты корреляции Пирсона (r) и ранговой корреляции Кендалла (τ), связывающие распределение видов по станциям рек Байтуган–Сок с нагрузками ординационных шкал (приведены виды с r > ±0,45)

Таксоны, определяющие шкалу 1	Шкала 1		Шкала 2		Таксоны, определяющие шкалу 2	Шкала 2	
	r	τ	r	τ		r	τ
Paratanytarsus sp.	-0,802	-0,686			*Unio* sp.	-0,702	-0,600
Ablabesmyia monilis	-0,765	-0,677	-0,652	-0,512	*Piscicola geometra*	-0,611	-0,490
Stempellina bausei	-0,712	-0,626			*Paralauterborniella nigrochal*	-0,587	-0,496
Harnischia fuscimana	-0,638	-0,572	-0,560	-0,453	*Pisidium amnicum*	-0,564	-0,518
P. nubeculosum	-0,607	-0,556			*Thienemannimyia* sp.	-0,563	-0,419
Paralauterborniella nigrochal	-0,607	-0,537			*E. dupuiana*	-0,562	-0,493
Cladotanytarsus mancus	-0,602	-0,492			*Caenis macrura*	-0,542	-0,556
Zavrelia pentatoma	-0,585	-0,480	-0,591	-0,52	*Paratanytarsus confusus*	-0,517	-0,380
Polypedilum scalaenum	-0,537	-0,473	-0,597	-0,534	*Odontomesa fulva*	0,582	0,545
P. convictum	-0,529	-0,413			*Orthocladius* sp.	0,585	0,453
Chironomus obtusidens	-0,522	-0,439			*Elmis aenea*	0,586	0,548
Sphaeriastrum rivicola	-0,520	-0,480	-0,700	-0,560	*Hydropsyche angustipennis*	0,596	0,532
Microchironomus tener	-0,475	-0,439			*Pristina rosea*	0,597	0,594
P. ferrugineus	-0,458	-0,279			*Psectrocladius* sp.	0,603	0,545
Parakiefferiella bathophila	0,459	0,440			*Limnophila* sp.	0,604	0,560
Tvetenia sp.	0,479	0,414			*Eukiefferiella* gr. *gracei*	0,613	0,566
Odontomyia sp.	0,489	0,414	0,532	0,453	*Culicoides* sp.	0,618	0,478
Teleopelopia sp.	0,499	0,520			*Orthocladius* (*Orth.*) *clarki*	0,619	0,534
Elmis sp.	0,500	0,453	0,547	0,493	*Parametriocnemus* sp.	0,619	0,473
Brillia modesta	0,647	0,580			*Eukiefferiella claripennis*	0,696	0,616
Parametriocnemus lundbecki	0,649	0,412	0,581	0,443	*Nemoura cinerea*	0,713	0,654
Pericoma sp.	0,697	0,575			*Ephemerella ignita*	0,721	0,647

Прямая ординация отображает изменение видового состава вдоль некоторого выбранного исследователем физического градиента среды (уровень загрязнения, ландшафтно-климатические изменения, временная динамика и т. д.).

Канонический анализ соответствий. В общем случае, когда требуется рассмотреть влияние целого комплекса факторов, которые в общем случае могут быть зависимыми друг от друга, используется *канонический анализ соответствий* (CCA, canonical correspondence analysis), предложенный С. тер Брааком (ter Braak, 1986).

Сущность метода CCA состоит в том, что выбирается такая линейная комбинация факторов среды $c_k \cdot x_{kj}$, где c_k – специально подобранные взвешивающие коэффициенты, которая обеспечивает максимальные расстояния между видами (и местообитаниями) в редуцированном ординационном пространстве. Пусть $\mathbf{Y} - \{y_{ij}\}$ является матрицей $m \times n$, содержащей центрированные значения обилия по m видам (строки) и n местообитаниям (столбцы). Аналогично $\mathbf{X}$ – матрица $q \times n$, в которой j-я строка содержит центрированные значения фактора среды x_{kj}. Задачей канонического анализа соответствий CCA является определение таких коэффициентов для видов $\mathbf{a} = (a_i)$ $[i = 1, ..., m]$ и для факторов среды $\mathbf{c} = (c_k)$ $[k = 1, ..., q]$, которые делают максимальной корреляцию между $\mathbf{z}^* = \mathbf{Y'a}$ и $\mathbf{z} = \mathbf{X'c}$. Решениями для $\mathbf{a}$ и $\mathbf{c}$ являются первые собственные векторы соответствующих уравнений:

$$\mathbf{YX'}\,(\mathbf{XX'})^{-1}\,\mathbf{Y'X\,c} = \lambda\,\mathbf{YY'a} \qquad \text{и} \qquad \mathbf{Y'X}\,(\mathbf{YY'})^{-1}\,\mathbf{YX'\,a} = \lambda\,\mathbf{XX'c}\,,$$

где собственная величина λ равна квадрату канонической корреляции.

Таким образом, в каноническом анализе соответствий распределение обилия видов $\{y_{ki}\}$ дополнительно описывается с помощью гауссовой модели отклика, в которой объясняющая переменная является линейной комбинацией факторов $\{c_k \cdot x_{ji}\}$. При этом ординационные оси F_1-F_2 ищутся с учетом той доли изменчивости данных, которая может быть связана с физическими градиентами среды. Тем самым метод CCA в комплексе решает задачи как собственно ординации, так и регрессионного анализа, статистически оценивая упорядоченность видов и местообитаний под влиянием внешних факторов.

В соответствии с этим результирующая диаграмма ординации отражает не только изменчивость видового состава относительно двух абстрактных осей F_1-F_2, но и основные связи между видами и каждым из факторов $\{x_{ji}\}$. Для этого из центра координат диаграммы проводятся дополнительные оси физических градиентов, которые обычно изображаются стрелками в плоскости, совмещенной с базовым биплотом, и тем самым не вносят никаких искажений в ординационную картину. Проекция точки вида на каждую стрелку показывает экологический оптимум (точнее, центр тяжести распределения обилия) этого вида относительно анализируемого физического фактора.

Положения концов стрелок зависят от рассчитанных собственных значений и внутригрупповых корреляций. При нормировке Хилла координата начала стрелки для k-й переменной среды на оси l равна $r_{kl} = [\lambda_s(1 - \lambda_l)]^{0,5}$, где r_{kl} – коэффициент корреляции этого фактора с осью l, λ_l – собственная величина для оси l. Переменные среды, обозначаемые длинными стрелками, сильнее связаны с осями ординации, чем факторы, обозначаемые короткими стрелками, и, следовательно, более значимо определяют изменчивость видового состава. Таким образом, канонический анализ позволяет одновременно отобразить связь средневзвешенных характеристик каждого вида, как с конкретными местообитаниями, так и с градиентом каждого фактора среды.

Пример ординационной CCA-диаграммы с наложением дополнительных осей объясняющих переменных, представленный на рис. 6.9, содержит в концентрированной форме богатый материал для различных экологических выводов и умозаключений. Найденные собственные значения λ_1 и λ_2 объясняют 34,4% вариации зависимости между встречаемостью видов и переменными среды, а корреляция последних с факторными нагрузками (табл. 6.9) предоставляет дополнительную информацию для интерпретации главных ординационных осей диаграммы. Нетрудно заметить, например, что положительное направле-

ние оси 1 связано с комплексом сильно связанных между собой гидрохимических показателей, определяющих концентрацию биогенных элементов, а ось 2 во многом определяется скоростью течения и содержанием растворенного кислорода.

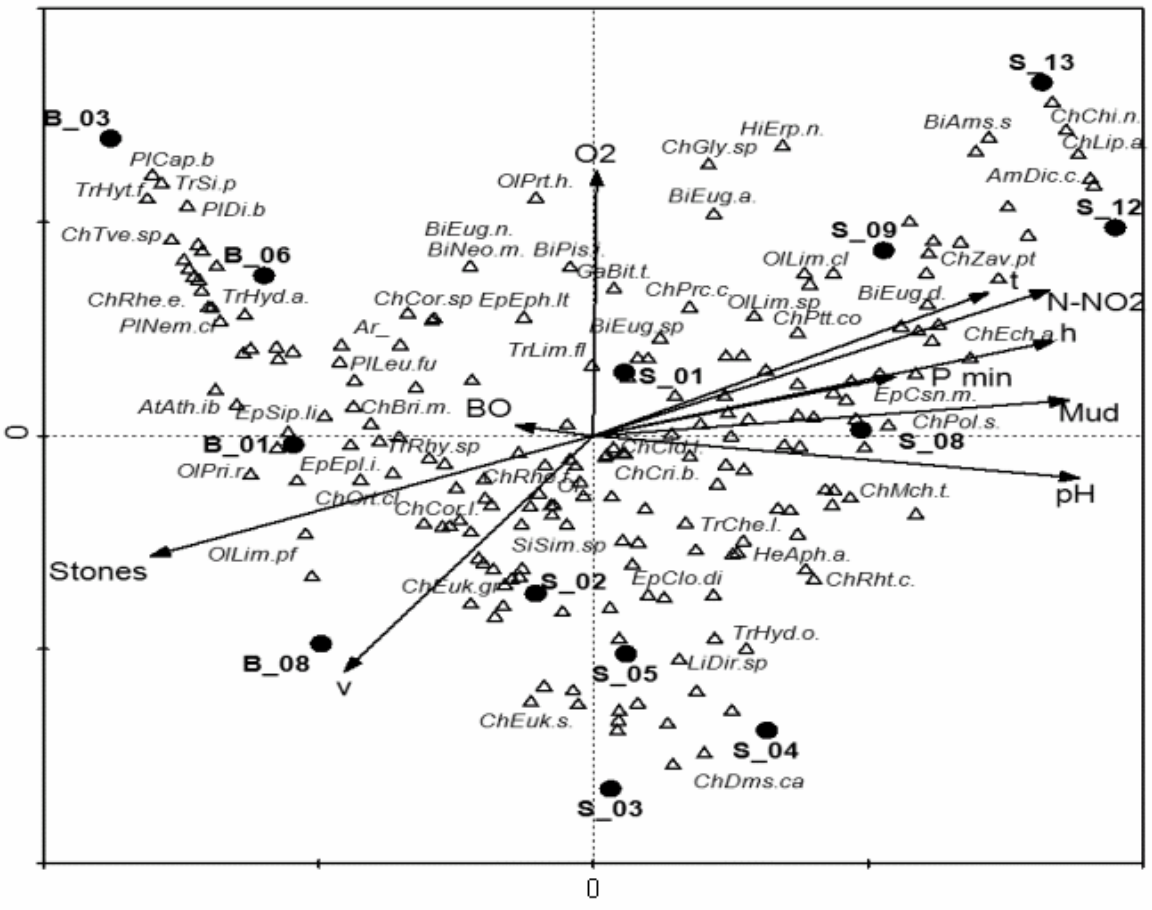


Рис. 6.9. **Пример ординационной диаграммы видовой структуры на станциях рек Байтуган–Сок (канонический анализ соответствий CCA):**

N-NO2 – содержание нитритного азота, **t** – температура, **BO** – бихроматная окисляемость, **h** – глубина в местах отбора проб, **v** – скорость течения, **O2** – содержание кислорода, **P min** – содержание минерального фосфора, **Stone** – каменистость грунта, **Mud** – заиленность грунта

Таблица 6.9

Коэффициенты корреляции Пирсона (r) и ранговой корреляции Кендалла (τ), связывающие переменные среды с факторными нагрузками (r > ±0,45)

Переменные среды	1 факторная ось		2 факторная ось	
	r	τ	r	τ
Каменистость грунта (**Stone**)	-0,805	-0,650		
Скорость течения (**v**)	-0,450	-0,395	-0,566	-0,369
Содержание кислорода (**O2**)			0,570	0,405
P min – содержание минерального фосфора	0,582	0,348		
Содержание нитритного азота (**N-NO2**)	0,744	0,763		
Глубина в местах отбора проб (**h**)	0,827	0,649		
Температура (**t**)	0,833	0,658		
Заиленность грунта (**Mud**)	0,835	0,714		
pH	0,864	0,722		

В целом местообитания по рассмотренному множеству переменных среды можно условно разделить на три категории:

- песчано-галечные субстраты с относительно чистой водой и высоким содержанием кислорода (ст. B_03, характерные виды *Capnia bifrons, Silo pallipes, Hydroptila femoralis, Diura bicaudata, Tvetenia* sp.) или его умеренной концентрацией (ст. B_01–B_08, *Atherix ibis, Ephemerella ignita, Hydropsyche angustipennis, Pristina rosea, Limnodrilus profundicola*);
- илистые грунты, с относительно высоким содержанием биогенных элементов, достаточной концентрацией кислорода, что характерно для устьевых станций р. Сок S_12 и S_13 (характерные виды *Chironomus nudiventis, Lipiniella araenicola, Dikerogammarus caspuis, Amesoda scaldiana)*;
- смешанные грунты с низким содержанием кислорода в среднем течении р. Сок (ст. S_08–S _09, виды *Limnodrilus* sp., *L. claparedeanus, Euglesa dupuiana, Procladius cho-reus, Tanypus vilipennis, Cricotopus* gr. *sylvestris, Dicrotendipes nervosus, Nanocladius bicolor*).

Анализ избыточности (RDA, redundancy analysis), предложенный К. Рао (Rao, 1964), является канонической формой PCA и считается линейной альтернативой алгоритму CCA с ограниченной изменчивостью весов местообитаний. Долгое время метод RDA не обращал на себя внимание экологов, однако в комбинации с анализом главных компонент его использование обладает рядом несомненных достоинств (Roberts, 2009).

Анализ избыточности, как и CCA, принимает во внимание не отдельные факторы среды, а линейные комбинации последних, которая дает наименьшую остаточную сумму квадратов. Коэффициенты канонической корреляции между видами и факторами среды рассчитывается таким же образом, как и в CCA, но в RDA – это корреляция между весами местообитаний, которые являются взвешенными суммами весов видов, и линейными комбинациями факторов среды, рассчитанными для каждого местообитания.

Диаграмма ординации RDA также может быть интерпретирована как совмещенный график, где точки видов и точки местообитаний отображают центры тяжести распределения обилия. Если виды представить стрелками, исходящими из начала координат, то косинус угла между стрелкой вида и стрелкой фактора среды приблизительно равен коэффициенту корреляции между ними. Кроме того, стрелки, имеющие примерно одно направление, говорят о значимой положительной корреляции, стрелки, пересекающиеся под прямым углом, говорят о корреляции, близкой к нулю, а стрелки, имеющие примерно противоположное направление, указывают на высокую отрицательную корреляцию.

Пример RDA-диаграммы для рассматриваемых донных сообществ, приведенный на рис. 6.10, содержит незначительные расхождения по сравнению с рассмотренной выше диаграммой CCA (после поворота осей на 180^{o}). Однако их идентичность является косвенной гарантией надежности получаемых результатов.

В каноническом корреляционном анализе (COR, canonical correlation analysis) проблема изучения взаимосвязи видов и факторов среды занимает центральное место, когда как в CCA и RDA она является вспомогательным средством. Идея анализа COR заключается в том, чтобы выбрать такие взвешивающие коэффициенты, которые бы максимизировали коэффициент корреляции между обилием видов и факторами среды.

Канонический линейный дискриминантный анализ (CVA, canonical variate analysis), часто называемый *анализом случайной переменной*, принадлежит к классическим линейным множественным методам (как PCA, RDA и COR). Формально CVA является особым случаем COR, в котором набор факторов среды состоит из единственного номинального фактора, определяющего классы. В отличие от CCA и RDA алгоритмы COR и CVA могут давать значительные погрешности при высоких уровнях взаимодействия между видами.

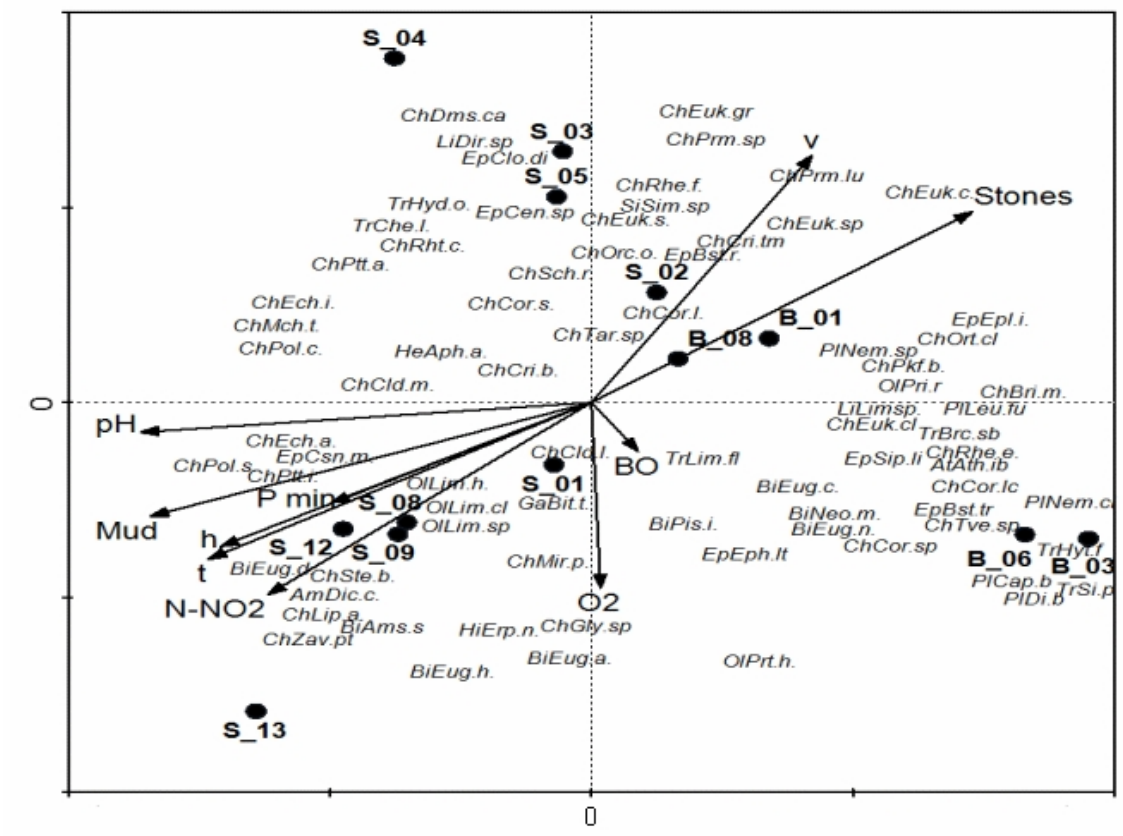


Рис. 6.10. **Пример ординационной диаграммы видовой структуры на станциях рек Байтуган–Сок (канонический анализ избыточности RDA).**
Условные обозначения см. на рис. 6.9

Дальнейшее расширение аналитических возможностей ординации связано с частной канонической ординацией. Предположим, что воздействие отдельных факторов среды должно быть выделено на фоне вариации признаков, вызываемых другими сопутствующими (латентными) факторами. Для этого матрица **X** объясняющих переменных заменяется остатками, полученными при построении регрессии воздействующих факторов среды на сопутствующие факторы. Детально теория частных версий RDA и CCA описана К. тер Брааком (ter Braak, 1988).

6.5. Ординация сообществ с использованием нечетких множеств

Традиционные принципы анализа структуры сообщества и формирования эколого-ценотических групп предполагают, что выделяемые классы видов или местообитаний представляют собой детерминированные совокупности: т. е. каждый объект может принадлежать только к одному таксону. Ограниченность такого подхода проявляется, например (Pillar, Orloci, 1991), в неразрешимости проблемы адекватного представления числа отсутствующих видов (т. е. нулей) при полевом учете, что приводит к аналитической неопределенности и множественности выводов при сравнении сообществ. Разумной альтернативой понятию абсолютной дискретности в классической таксономии является интерпретация компонентов экосистем как нечетких объектов в составе гибко настраиваемых ординационных структур.

Операции с нечеткими множествами появились, как эффективная практическая мера преодоления правила несовместимости Неймана–Заде[5]: «повышение точности описания сложной системы становится несовместимым со здравым смыслом, поскольку сложность модели становится соизмеримой со сложностью самого объекта». В конце 80-х годов, после бурного развития технических устройств на базе нечетких контроллеров и экспертных систем, теория нечетких множеств (fuzzy sets) и нечеткая логика (fuzzy logic) становятся важными обобщениями классических математических теорий и неотъемлемой составной частью современных систем искусственного интеллекта.

Основные понятия fuzzy-концепций, впервые предложенные американским ученым Лотфи Заде (Zadeh, 1965; Bezdek, 1987), сводятся к следующему:

- множество **C** является нечетким, если существует функция принадлежности (membership function) $\mu_C(x)$, принимающая на этом множестве значения в интервале [0, 1];
- функция $\mu_C(x)$ конструируется на основе экспертных заключений или любого подходящего формального метода и оценивает степень сродства (grade) анализируемого объекта x к произвольного множеству **C**: $\mu_C(x) = 0$ означает полную несовместимость, т. е. $x \notin$ **C**, а $\mu_C(x) = 1$ – полную принадлежность или $x \in$ **C**;
- нечеткое множество **C** задается множеством упорядоченных пар типа $C = \{x , \mu_C(x)\}$; в частном случае, если функция принадлежности принимает значение только 0 или 1, то **C** становится «четким» или обычным множеством.

Например, применительно к методам нечеткой таксономии при анализе компонент экосистемы, функция $\mu_C(x)$ задает в масштабе от 0 до 1 степень принадлежности некого объекта x (вида или местообитания) к выделяемому классу **C**.

Результаты использования алгоритмов нечетких множеств в биогеографии, экологии и численной систематике планомерно обобщались, анализировались и широко публиковались (Bezdek, 1974; Roberts, 1986; Marsili-Libelli, 1989; Zhang, 1994; Boyce, 1998; Feoli, Zerihun, 2000; Boyce, Ellison, 2001; Pedersen et al., 2006). Временами звучали "фанфары, предвещающие светлое будущее": «Новая концептуальная основа, объединяющая нечеткую математику и динамическую теорию систем, позволит в математически строгой и экологически глубокой манере идентифицировать природу взаимодействия между ключевыми компонентами экосистем в гиперпространстве значимых факторов» (Roberts, 1987, p. 28). Однако отсутствие стандартной методики конструирования нечетких экосистем зачастую не приводили к повышению точности вычислений по сравнению с существующими вероятностными методами. Хотя считается, что «нечеткая математика обеспечивает более интуитивный подход ко многим экологическим вопросам», эта интуиция в первую очередь основывается на необходимости априорно ранжировать параметры окружающей среды, которые могут представить потенциальный интерес для исследователя. С одной стороны, это – прекрасный шаг, активно включающий полевого эколога в процесс компьютерного моделирования, но вносящий, тем не менее, определенную долю субъективизма.

В наибольшей мере преимущества математики нечетких множеств были реализованы в методе ординации сообществ, известном под аббревиатурой FSO (Fuzzy Sets Ordination – Roberts, 1986, 2008). FSO, в отличие от непрямых (косвенных) методов ординации, использует данные о факторах окружающей среды непосредственно при проведении вычислений и, таким образом, может относиться к семейству прямых ординаций, таких как канонический анализ соответствия CCA или канонический анализ главных координат (CAPC; Anderson, Willis, 2003). Фактически FSO выглядит как скорректированная на

[5] На самом деле основанием для создания новой теории послужил спор проф. Л. Заде со своим другом о том, чья из жен привлекательнее. К единому мнению они, естественно, так и не пришли, но это вынудило Заде сформулировать концепцию, которая выражает нечеткие понятия типа «привлекательность» в числовой форме.

учет факторов среды версия *полярной ординации Брея–Куртиса*, предшественницы всех ординационных методов.

Метод FSO оперирует с данными в нечетком топологическом пространстве, т. е. в пространстве, где элементы связаны между собой нечеткими отношениями на интервале [0, 1], значения которых определяют меру близости между собой каждых двух элементов. Рефлексивная и симметричная матрица любых мер подобия является примером такого нечеткого отношения. Р. Бойс и П. Элисон (Boyce, Ellison, 2001) рассмотрели множество индексов сходства на предмет их использования в FSO и рекомендовали восемь индексов, включая меры Жаккара, Съеренсена, Горна, Юла и других, которые игнорируют составляющую совместного отсутствия видов. Одно из главных преимуществ FSO состоит в том, что, сохраняя формальную логику и математику «четких» (евклидовых) метрик, он не зависит от каких-либо предварительно задаваемых порогов, а основывается на нечетких функциях, оценивающих уровень принадлежности на множестве элементарных исходов.

В *одномерном* FSO исследователем задается один ключевой параметр среды *z*, который потенциально может оказаться ведущим градиентом. После этого реализуется последовательность операций на промежуточных множествах, позволяющих преобразовать конфигурацию точек из нечеткого топологического пространства в евклидово пространство, где становятся доступными средства параметрической статистики (Roberts, 1986). Результирующее нечеткое множество **E**, которое обычно используется для отображения на ординационной диаграмме, вычисляется с использованием оператора *антикоммутативной разности*, оценивающего контраст двух полярных нечетких множеств **C** и **D**, соответствующих высоким и низким значениям изучаемого фактора среды *z*:

$$\mu_E(x) = \{1 + [1 - \mu_D(x)]^2 - [1 - \mu_C(x)]^2\}/2;\ \mu_{\mathbf{C}}(x) = \frac{\sum_{y \neq x} d(xy)\mu_{\mathbf{A}}(x)}{\sum_{y \neq x} \mu_{\mathbf{A}}(y)};\ \mu_{\mathbf{D}}(x) = \frac{\sum_{y \neq x} d(xy)[1 - \mu_{\mathbf{A}}(x)]}{\sum_{y \neq x} [1 - \mu_{\mathbf{A}}(y)]},$$

где x и y – анализируемые экологические объекты (например, местообитания); $d(x, y)$ – сходство между этими объектами в соответствии с выбранной метрикой; $\mu_P(x)$, – функции принадлежности, задающие каждое из представленных нечетких множеств, $P = \{x, \mu_P(x)\}$, $P = \mathbf{E}, \mathbf{C}, \mathbf{D}, \mathbf{A}$.

Здесь функция принадлежности $\mu_A(x)$ является просто набором нормированных значений фактора среды z_x для всех классифицируемых объектов x: $\mu_{\mathbf{A}}(x) = \dfrac{z_x - \min(z)}{\max(z) - \min(z)}$.

Практически метод FSO моделирует распределение оптимумов относительной таксономической насыщенности разных местообитаний на шкале градиента фактора среды, т.е. оценивается роль, которую играет переменная *z* в формировании или ограничении видовых композиций сообщества. Эта аналитическая процедура была ранее описана К. тер Брааком (Jongman et al., 1987, глава 4) как «калибровка».

Выше (раздел 6.1) отмечалось, что практически трудно найти идеальные биологические индикаторы, т.е. виды, которые всегда встречаются при фиксированном множестве условий среды и не встречаются в другом месте. Безусловно, существуют относительно редкие виды с узкими экологическими амплитудами, но они не всегда могут быть обнаружены в предназначенном им ареале, а их отсутствие вряд ли может предоставлять полезную информацию об окружающей среде. Это – главная причина использования в целях калибровки и последующей экологической экспертизы целостных таксономических композиций сообществ на участках, включая виды с более широкими экологическими амплитудами. Теоретической основой разработки конкретных дефиниций такого многомерного отклика является модель совместного распределения видов (species packing model), основанная на концепциях максимального правдоподобия и учитывающая различное положение экологических оптимумов каждого вида на шкале градиента фактора среды. Практически же оценка отклика чаще всего сводится к обоснованию различных полуколичествен-

ных биотических индексов, использующих, как правило, алгоритм среднего взвешивания (Ellenberg et al., 1991) или обратную регрессию (Jongman et al., 1987). Функцию принадлежности $\mu_E(x)$ можно рассматривать как одну из версий количественного представления мультивидового отклика, оценивающего на интервале [0, 1] приведенную таксономическую изменчивость при различных значениях фактора среды *z*.

Выполним расчет компонентов нечетких множеств на примере данных о встречаемости 214 видов макрозообентоса на 13 участках речной экосистемы Байтуган–Сок. Для расчета матрицы подобия, согласно имеющимся рекомендациям (Boyce, Ellison, 2001), будем использовать индекс Горна, дающий наиболее стабильные результаты для количественных признаков:

$$a_{ij} = \frac{\sum_{k=1}^{S}(t_{ik}+t_{jk})\log(t_{ik}+t_{jk}) - \sum_{k=1}^{S} t_{ik}\log t_{ik} - \sum_{k=1}^{S} t_{jk}\log t_{jk}}{(\sum_{k=1}^{S} t_{ik} + \sum_{k=1}^{S} t_{jk})\log(\sum_{k=1}^{S} t_{ik} + \sum_{k=1}^{S} t_{jk}) - \sum_{k=1}^{S} t_{ik}\log\sum_{k=1}^{S} t_{ik} - \sum_{k=1}^{S} t_{jk}\log\sum_{k=1}^{s} t_{jk}}, \quad (6.4)$$

где t_{ik} – частота встречаемости вида k (k = 1, …, S) в гидробиологических пробах, взятых на i-м участке. Меры подобия были трансформированы в меры расстояния: $d_{ij} = (1 - a_{ij})$.

Простым критерием значимости полученных осей нечетких множеств может служить корреляция между функцией принадлежности $\mu_E(x)$ и значениями соответствующей экологической переменной *z*. Если в распоряжении исследователя оказывается целый набор потенциальных параметров среды, определяющих закономерность распределения видов по местообитаниям, то можно последовательно провести сравнительный анализ всех факторов (см. Pedersen et al., 2006): параметры окружающей среды с высокой корреляцией считаются наиболее влияющими.

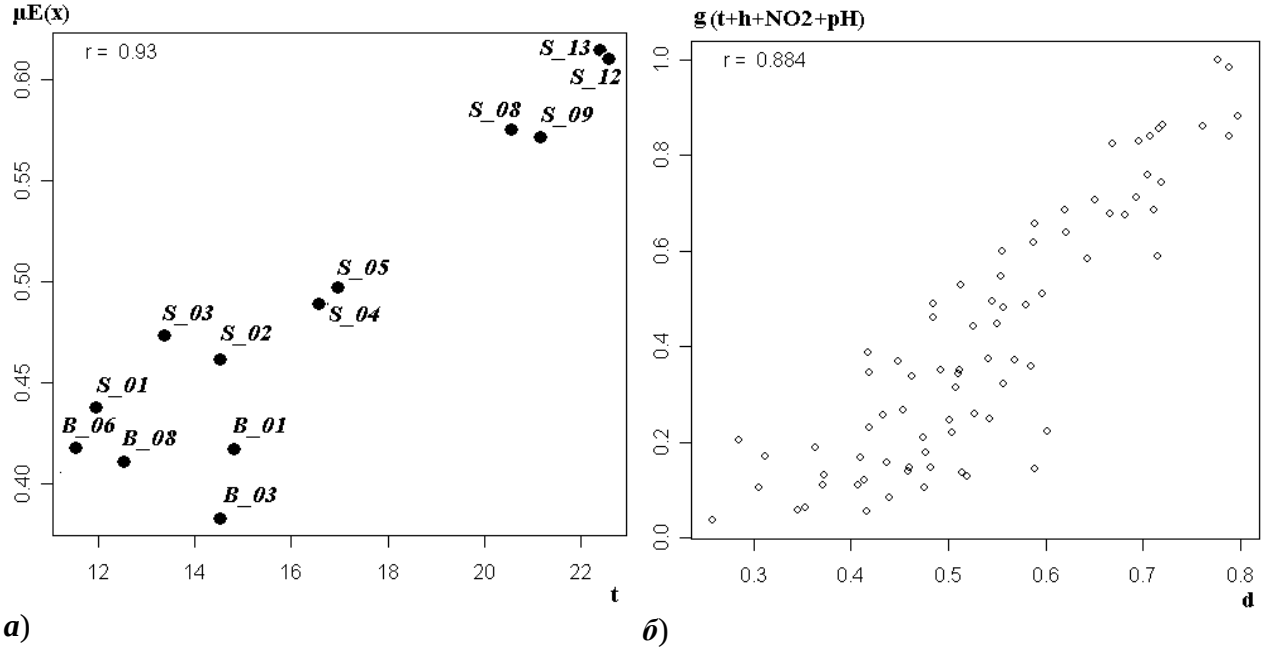
Оценка *p*-значений достоверности коэффициентов корреляции Пирсона *r* может быть вычислена обычным путем нормального приближения с соответствующими числами степеней свободы. Тем не менее, имеющийся опыт (Roberts, 1986) показывает при этом очевидную тенденцию высокой ошибки первого рода, поэтому как альтернатива нами был выполнен расчет статистических величин *p*, основанный на рандомизации. Для этого исходные описания участков *i* и *j* случайным образом перемешивались 1000 раз и рандомизированная оценка *p* соответствовала вероятности того, что имитируемый коэффициент корреляции превысит или будет равен эмпирическому значению *r*. Расчеты показали (табл. 6.10), что можно принять статистически значимым влияние 8 из 12 анализируемых факторов среды на таксономическую изменчивость донных сообществ.

Таблица 6.10

Коэффициенты корреляции Пирсона *r*, связывающие переменные среды с функциями принадлежности нечетких множеств $\mu_E(x)$; p_n и p_m – оценки статистической значимости *r*, найденные, соответственно, методами нормального приближения и рандомизации

Переменные среды	***r***	p_n	p_m
Температура воды (**t**)	0,929	$4{,}18\cdot10^{-6}$	**0,001**
Глубина в местах отбора проб (**h**)	0,921	$7{,}55\cdot10^{-6}$	**0,001**
Содержание нитритного азота (**N-NO2**)	0,828	$4{,}72\cdot10^{-4}$	**0,001**
pH	0,822	$5{,}61\cdot10^{-4}$	**0,001**
Каменистость грунта (**Stone)**	0,806	$8{,}67\cdot10^{-4}$	**0,001**
Высота над уровнем моря (**H**)	0,797	0,001	**0,001**
Заиленность грунта (**Mud**)	0,778	0,0017	**0,001**
P min – содержание минерального фосфора	0,641	0,018	**0,007**
Скорость течения (**v)**	0,412	0,16	0,077
Содержание кислорода (**O2**)	0,180	0,55	0,226
Площадь водосбора (***F***)	0,044	0,88	0,323
Бихроматная окисляемость (**BO**)	-0,526	0,064	0,784

Можно отметить тесную связь найденных коэффициентов корреляции с проекциями факторов среды на первую главную ось ординации, полученными по методу CCA (рис. 6.9). Однако методы нечеткой логики на современном этапе не предоставляют столь разносторонних способов графической интерпретации, как классические алгоритмы ординации. Обычно исследователю предлагается проанализировать корреляционные поля зависимостей $\mu_E(x)$ от z (рис. 6.11а), матрицы расстояний d от z (рис. 6.11б) или различные парные взаимодействия двух нечетких множеств.



а) *б*)

Рис. 6.11. **Корреляционная связь функции принадлежности нечеткого множества $\mu_E(x)$ с температурой воды t (*а*) и комплексного ординационного расстояния g(t+h+NO_2+pH) с компонентами матрицы таксономических расстояний d (*б*)**

Одномерный FSO позволяет, подобно одномерному регрессионному анализу, оценить статистическую зависимость функций принадлежности размытых множеств только от одного фактора среды. Для анализа всего комплекса переменных в целом можно воспользоваться *многомерной* версией ординации нечетких множеств MFSO (Roberts, 2008). В методе MFSO для определения нечетких множеств в многомерном пространстве сначала выполняется нахождение ординационной оси относительно экологической переменной, которая в наибольшей мере объясняет вариацию композиций видов (т.е. имеет наибольший коэффициент корреляции). Вторая ось, связанная со следующим фактором, рассчитывается как перпендикуляр к первой с использованием процедуры ортогонализации *Грамма-Шмидта***,** а ее значения соответствуют остаткам регрессии на первую ось. Таким образом, значения второго градиента нечеткого множества оценивают только ту долю влияния второй переменной, которая не может быть объяснена первым фактором. Оси третьего и высшего порядка рассчитываются аналогично и отражают долю таксономической вариации, не связанную со всеми предыдущими осями нечеткого множества.

При использовании MFSO вследствие применения ортогонализации Грамма-Шмидта предварительно должен быть проведен статистический анализ и тщательно исключены те переменные среды, которые сильно коррелированны с другими факторами. В общем случае алгоритмы FSO предполагают априорный неформальный выбор переменных, основанный на экологических представлениях. У исследователя должны быть одна

или несколько рабочих конкурирующих гипотез, которые могут быть проверены путем сравнения полученных результатов ординации на нечетких множествах.

Предварительный отбор факторов среды может быть проведен с использованием обычного корреляционного анализа, метода главных компонент или теста Мантеля, описываемого ниже. В нашем случае все 8 отобранных признаков образуют сильно взаимосвязанный комплекс переменных: коэффициенты корреляции между парами статистически значимых векторов $\mu_E(x)$ варьируют от 0,982 до 0,999. Это подтверждается при использовании шаговой процедуры FSO, которая может оказаться полезной для быстрой селекции значимых осей ординации. Алгоритм стартует с фактора, имеющего наибольший коэффициент корреляции $r(d, z)$, и последовательно включает в модель остальные переменные, оценивая их по приращению, вносимому ими в результирующее значение r. Представленные результаты (табл. 6.11 слева) позволяют исключить из дальнейшего рассмотрения три фактора среды, использование которых снижает общий коэффициент корреляции.

Таблица 6.11

Результаты шаговой процедуры селекции признаков и многомерного анализа нечетких множеств с выделением значимых осей ординации с использованием коэффициентов корреляции Пирсона *r*; *p* – оценки статистической значимости приращений *r*, найденные методом рандомизации

Переменные среды	Шаговая процедура			Многомерная процедура		
	***r*(*d*, *z*)**	**Приращение**	***p***	***r*(μ_E(*x*), z)**	**Приращение**	***p***
Температура воды (**t**)	0,879			0,929		**0,001**
pH	0,8824	0,0034	0,07	0,9328	0,0038	**0,017**
Содержание нитритного азота (**N-NO2**)	0,8855	0,0031	0,14	*0,9337*	*0,00095*	*0,055*
Высота над уровнем моря (**H**)	0,8857	0,00017	0,17		*-0,002*	
Глубина в местах отбора проб (**h**)	0,8859	0,00024	0,18	*0,9339*	*0,0002*	*0,071*
P min – содержание мин. фосфора	*0,8859*	*-0,00002*	*0,34*			
Каменистость грунта (**Stone)**	*0,8839*	*-0,002*	*0,23*			
Заиленность грунта (**Mud**)	*0,8831*	*-0,00079*	*0,25*			

Примечание: курсивом отмечены факторы, исключаемые из рассмотрения на каждом этапе

Применение многомерной процедуры анализа нечетких множеств MFSO к оставшимся признакам (табл. 6.11 справа) позволяет выделить две относительно некоррелированные и статистически значимые оси ординации донных сообществ, основанные на температуре воды и pH.

6.6. Оценка надежности и достоверности результатов классификации

Как отмечалось выше, кластерный анализ и ординация используются для поиска наиболее «естественных» экологических закономерностей. Существуют, впрочем, различные точки зрения на понятие «естественности», например:

- естественная закономерность описывает глубокие природные механизмы, тогда как искусственная закономерность – неглубокие;
- естественна та закономерность для конкретной предметной области, которая наиболее быстро вытекает из ее тезауруса или предыдущего опыта;
- естественная закономерность удовлетворяет многим целям, в то время как цель поиска искусственных закономерностей задает человек;
- естественная закономерность позволяет делать прогнозы;
- естественная закономерность наиболее устойчива.

Приведенные высказывания дают представление о столь больших расхождениях в трактовке этого тонкого предмета, что термин «естественность» следует признать таким же нечетким, как и многие другие экологические понятия.

Анализ результатов ординации. Задача ординации состоит в нахождении существенной, стабильной и нетривиальной геометрической метафоры данных, полезной для дальнейшей экологической интерпретации. И здесь существует, например, опасность «ложных выводов» и заблуждений, основанных на выделении случайных и непостоянных тенденций в экосистеме, которые вновь не проявились бы при анализе других выборок из той же генеральной совокупности. Другое обстоятельство связано с определенной субъективностью визуального восприятия исследователем сгущений точек (см. рис. 6.4–6.10).

Исходя из общих представлений, количественно оценить достоверность найденных ординаций можно различными известными способами (McCune et al., 2002):

- используя долю объясняемой вариации данных;
- выполняя ординацию на последовательности различных фрагментов из одних и тех же исходных данных;
- по найденным статистическим отличиям при сопоставлении с нуль-моделью;
- сравнивая найденную закономерность с некоторым эталонным образцом.

Если не принимать во внимание ошибки измерения и часто недостаточную выборочную репрезентативность исходных данных, то доля объясненной дисперсии, связанная с ведущими собственными значениями λ_1 и λ_2, может служить хорошей оценкой достоверности ординации. Ниже приведены частные и накопленные коэффициенты детерминации r^2, оценивающие корреляцию между отображаемыми объектами в исходном многомерном пространстве и проекциями на оси факторов в рамках расчетов, выполненных в разделе 6.3:

Оси	PCA (рис. 6.4)		CA (рис. 6.5)		DCA (рис. 6.6)		CCA (рис. 6.9)	
	прира-щение	накоп-лено	прира-щение	накоп-лено	прира-щение	накоп-лено	прира-щение	накоп-лено
1	0,259	0,259	0,269	0,269	0,269	0,269	0,268	0,268
2	0,128	0,387	0,109	0,378	0,109	0,378	0,140	0,408
3	0,058	0,445	0,090	0,468	0,090	0,468	0,091	0,499

Однако при этом учитывается только один аспект – информационные потери при редукции исходного многомерного пространства видов и местообитаний до 2-мерного представления. К тому же, имеющиеся выбросы значений обилия видов могут дать сильные искажения этих оценок. Существуют и другие параметрические методы проверки статистической значимости найденной корреляционной структуры многомерных данных, например, *тест Бартлетта* на сферичность.

М. Вильсон (Wilson, 1981; Wilson, Mohler, 1983) разработал метод оценки согласованности результатов ординации видов вдоль градиентов на основе многократного разделения одной и той же исходной биогеографической таблицы на непересекающиеся блоки (подмножества) данных. Пусть $X(w,w_0)$ – число рассогласованных пар в двух ранжированных рядах видов w и w_0, где w_0 – значения w, приводящие к минимуму $\sum_{i=1}^{q} X(w, \ddot{w})$. Два вида составляют рассогласованную пару, если их порядок следования в обоих ранжированных рядах полностью нарушен. В результате последовательных аппроксимаций ищется «истинная» ординация видов, соответствующая оценке максимального правдоподобия. Мера рассогласования результатов, рассчитанная как *коэффициент* τ *ранговой корреляции Кендалла*, колеблется от -1 (полное разногласие) до +1 (полная согласованность) и может использоваться как оценка точности ординации видов.

В то же время Вильсон (Wilson, 1981) отмечал, что отклонение нулевой гипотезы о статистической независимости полученных ранжированных рядов не является сильным утверждением о достоверности полученных закономерностей. Сделанный вывод о взаимо-

связи результатов ординации на частных подмножествах данных не указывает на степень выраженности общего решения и не слишком пригоден для интерпретации. Более соответствует поставленной задаче сравнение результатов ординации с нуль-моделью, в которой любое ранжирование видов имеет полностью случайный характер.

Непараметрические методы, использующие испытания Монте-Карло (чаще всего – многократное случайное перемешивание исходных таблиц или матриц расстояния), обладают рядом решающих преимуществ в оценке достоверности структурных закономерностей, выделяемых любыми многомерными методами. В отличие от классических тестов, основанных на критериях t, F и χ^2, перестановочные процедуры (permutation procedure) предоставляют возможность получить точное значение вероятности проверяемой гипотезы, не прибегая к каким-либо предположениям о характере распределения анализируемых данных. Точнее, единственным исходным предположением является независимость результатов проверки нулевой гипотезы от схемы обмена (exchangeability) перемешиваемых единиц (Hayes, 1996).

Вычисленная вероятность p оценивает меру достоверности найденной ординационной структуры по сравнению с набором данных, содержащим переменные с тем же самым характером распределения, но случайной зависимостью между ними. Если величина p меньше, чем задаваемый уровень значимости (например, $\alpha = 0{,}05$), то результат ординации можно считать нетривиальным и пригодным для интерпретации. В противном случае следует принимать во внимание, что полученная главная ось может оказаться неустойчивой, поскольку вариация относительно нее статистически значимо не отличается от случайных флуктуаций.

Одна из версий рандомизированной нуль-модель для описываемой задачи может быть получена (McCune et al., 2002), если для каждого столбца (местообитания) независимо и случайно переставлять значения обилия видов. В результате образуется последовательность матриц с одинаковыми суммами по столбцам и общим количеством нулевых элементов, что и в исходной таблице, т. е. рандомизация сохраняет многие из свойств эмпирических данных, но (насколько это возможно) разрушает взаимоотношения между видами. Например, в рамках расчетов по методу PCA (раздел 6.4) в ходе 1000 итераций перемешивания сравнивались эмпирические λ_{obs} и рандомизированные λ_{sim} значения собственных чисел и подсчитывалось количество случаев, когда $\lambda_{isim} > \lambda_{iobs}$:

Главные компоненты	**Эмпирические λ_{obs}**	**Собственные числа λ_{sim} при рандомизации**			
		Минимум	**Среднее**	**Максимум**	***p*-значение**
1	2,189	1,286	1,43	1,641	0,001
2	1,574	1,196	1,32	1,488	0,001
3	1,187	1,137	1,24	1,353	0,923

Очевидно, что третья ось ординации уже не имеет статистической значимости и нет необходимости в её применении, несмотря на достаточно ощутимое приращение доли объясняемой дисперсии.

В. Пиллар (Pillar, 1999a) предложил следующую уточненную схему тестирования метрических ординаций (CA, PCA или PCoA) с использованием бутстреп-метода:

- генерируется большое число B (в нашем случае, $B = 1000$) псевдовыборок на основе случайных комбинаций из столбцов исходной биогеографической матрицы, соответствующих отдельным участкам водотока;
- полученные бутстреп-матрицы подвергаются ординации по одинаковой схеме с последующей «прокрустовой стандартизацией» (Procrustean adjustment), т. е. путем вращения и корректировки размерностей результаты приводятся к сопоставимой форме;
- для каждой псевдовыборки рассчитывается средний коэффициент корреляции θ^*_1 факторных весов 1-й главной ординационной оси со значениями натуральных признаков;

- значения показателей, представленных в строках псевдоматрицы, многократно случайным образом перемешиваются, рассчитывается средний коэффициент корреляции θ^{o}_{1} для рандомизированной структуры и проверяется выполнение условия $\theta^{o}_{1} > \theta^{*}_{1}$;
- вычисляется вероятность ошибки 1-го рода $P(\theta^{o}_{1} > \theta^{*}_{1}) = (1 + F) / (1 + B)$, где F – число случаев, когда коэффициент корреляции для рандомизированных данных оказался больше, чем для эмпирических;
- аналогичные вычисления проводятся для остальных осей ординации.

Как и при других статистических тестах, небольшой объем выборки вызывает высокую вероятность ошибки второго рода. Невозможность отклонить нулевую гипотезу здесь просто означает, что для уверенного заключения необходимо продолжить исследования и получить новые данные.

Как получить оптимальную ординацию с высокой разрешающей способностью, отражающую «естественные» связи в сообществах? Наряду с многообразием методов ординации, на результаты расчетов могут оказывать существенное влияние формат и размерность обрабатываемых таблиц, состав и способ трансформирования исходных показателей, выбор выражения для меры сходства и другие факторы. Вероятность $P(\theta^{o}_{1} > \theta^{*}_{1})$, рассчитанная бутстреп-методом по схеме Пиллара, является таким индикатором устойчивости ординационной структуры изучаемой экосистемы по сравнению с ее нуль-моделью.

Нами было проанализировано на одном и том же исходном материале, как изменяется надежность результатов ординации в зависимости от таких ключевых параметров расчета как размерность признакового пространства, тип обрабатываемых данных и формулы для расчета меры сходства (табл. 6.12).

Представленные результаты показывают, что несомненным приоритетом получения качественной ординации обладает представление данных в виде частот встречаемости видов в гидробиологических пробах. Использование этого показателя оказалось единственным, приводящим к статистически значимой корреляции с главным градиентом. Существенно более зашумленные результаты были получены при использовании средней численности особей каждого вида при различных версиях нормировки (в том числе, проективного покрытия, широко употребляемого в геоботанике). Наименее стабильные результаты, как и ожидалось, дало представление данных в виде бинарных переменных («наличие/отсутствие» вида) и использование индекса сходства Съеренсена.

Применение различных метрик расстояния не оказывает столь большого влияния на итоги ординации. Можно лишь отметить несколько худшие результаты, полученные с использованием евклидова расстояния, и хорошую стабильность применения процентной разности Брея-Куртиса.

Несколько парадоксальные для нас выводы были получены при анализе влияния количества видов, представленных в исходной матрице. Как и принято считать (см. Jongman et al., 1987), включение комплекса редких видов немного снизило стабильность ординации за счет нарастающей ошибки при оценке собственных значений. Однако, вопреки ожиданию, еще более ухудшил её качество переход к тщательно укомплектованному списку из 69 надежных видов с наилучшими биоиндикационными свойствами (что является, видимо, проявлением закона Коммонера «*Природа знает лучше*»). Наконец, к еще менее стабильным результатам привело агрегирование видов в систематические группы более высокого уровня (семейства, подсемейства и трибы).

Анализ результатов классификации. Общепринятая методика оценки надежности классификаций основана на интерпретируемости и повторяемости. Если одна и та же закономерность проявляется при использовании различных вариантов или методов классификации, отличаясь лишь в некоторых деталях, то аналитик приходит к мнению, что основная тенденция изменчивости структуры экосистемы найдена. Строгость и стройность этому субъективному подходу могут придать количественные методы оценки надежности и статистической значимости найденных группировок.

Таблица 6.12

Результаты анализа устойчивости различных вариантов ординации методом главных координат (PCoA) на основе процедуры бутстреппинга (Pillar, 1999a)

Исходная матрица		Мера сходства[1)]	Главная ось 1[2)]			Главная ось 2[2)]		
Видов	Тип данных		d_1, %	θ^*_1	$P(\theta^o_1 > \theta^*_1)$	d_2, %	θ^*_2	$P(\theta^o_2 > \theta^*_2)$
214[3)]	Наличие/отсутствие	*Sr*	15,02	0,664	0,524	11,29	0,65	0,501
	Средний $ln(N+1)$,	*E*	19,35	0,789	0,258	12,22	0,720	0,421
	N – численность особей	*R*	18,39	0,808	0,181	11,58	0,762	0,461
	вида (экз/м^2)	*BC*	24,6	0,88	0,148	12,58	0,69	0,466
	Численность, нормированная на сумму по участкам (проективное покрытие)	*R*	16,41	0,621	0,341	12,23	0,644	0,412
	То же, по встречаемости видов	*R*	18,38	0,79	0,206	11,58	0,686	0,458
	Встречаемость вида в пробах	*E*	40,0	0,974	**0,04**	11,88	0,768	0,379
		R	33,61	0,971	**0,029**	11,29	0,672	0,519
		BC	34,81	0,964	**0,032**	12,44	0,679	0,463
375[4)]	Встречаемость	*BC*	31,1	0,95	0,059	12,2	0,678	0,454
69[5)]	Встречаемость	*BC*	37,24	0,965	0,074	17,8	0,813	0,348
45[6)]	Встречаемость	*BC*	47,1	0,949	0,195	17,55	0,919	0,3

Примечания:

1) *Sr* – индекс Съеренсена для бинарных признаков; *E* – евклидово расстояние; *R* – коэффициент корреляции Пирсона; *BC* – коэффициент различий Брея-Куртиса (процентная разность);
2) d_i – доля, объясненной дисперсии, соответствующая главной координате с собственным значением λ_i; θ^*_i – средний коэффициент корреляции бутстреппированных данных с факторными нагрузками *i*-й ординационной оси; $P(\theta^o_i > \theta^*_i)$ – вероятность того, что средний коэффициент корреляции θ^o_i для рандомизированной структуры окажется больше, чем θ^*_i (жирным шрифтом показаны случаи отклонения нулевой гипотезы);
3) число таксонов, встретившихся более двух раз на двух или более участках (т. е. без учета редких видов);
4) общее число обнаруженных таксонов макрозообентоса;
5) число таксонов с максимальным индексом верности (коэффициент ассоциации Ф > 0,185 – Chytrý et al., 2002);
6) число систематических групп макрозообентоса с уровнем выше вида.

Непараметрическое многомерное тестирование различий между двумя или более группами наблюдений обеспечивает *процедура перестановки с анализом множественного отклика* (MRPP, multi-response permutation procedure; Mielke, 1984; McCune et al., 2002), использующая матрицу расстояний. Метод дает весьма хорошие результаты и практически не накладывает никаких ограничений на анализируемые данные, однако применим только в случае простой схемы исследований, основанной на одноуровневой классификации.

Предположим, что множество n местообитаний разбито на κ групп так, что $\sum_{i=1}^{k} n_i = n, \quad n_i \geq 2$. Процедура MRPP предполагает априорный характер такого разбиения, основанный на тех или иных экологических предположениях. Например, необходимо проверить нулевую гипотезу, что видовой состав донных сообществ никак не связан с типом грунтов ложа водотока, классификация которых установлена заранее. Суть данного метода заключается в следующем:

- вычисляется матрицу расстояния **D** между всеми n местообитаниями; могут использоваться различные меры расстояния в пространстве видов: *евклидово расстояние*, его

квадрат, *манхэттенская мера* (т. е. ненормированный *индекс сходства Съеренсена*) и т. д.; если MRPP используется в сочетании с ординацией, то желательно выбрать одну и ту же меру расстояния для обоих методов;

- находится среднее расстояние d_i между объектами каждой выделенной группы $i = 1, ..., k$ и средневзвешенный индекс внутригруппового расстояния $\delta_{obs} = \sum_{i=1}^{k} \frac{n_i}{n} d_i$; для локальных групп, у которых $d_i > \delta_{obs}$, можно предположить ослабленную выраженность группировки;
- запускается процедура случайной перестановки векторов местообитаний между группами, и каждый раз вычисляются средние значения различий объектов δ_{sim} внутри «перемешанных» групп; если размерности n и $к$ относительно невелики, то возможен полный перебор (например, если 30 местообитаний разделены на 2 группы и $n_1 = n_2 = 15$, то число необходимых перестановок определяется по формуле $n!/(n_1!\cdot n_2!) = 1{,}55\cdot 10^8$), в противном случае процедура рандомизации повторяется достаточно большое число раз, чтобы можно было построить статистическое распределение величины δ;
- в ходе итерационной процедуры подсчитывается число случаев, когда значение δ_{sim} для псевдо-группировки не превышает эмпирическое значение δ_{obs}, и вероятность p этого события определяет вероятность ошибки 1-го рода при проверке нулевой гипотезы (это соответствует площади под кривой распределения δ левее значения δ_{obs}).

Другим способом оценки вероятности p, который может быть применим только к средним или большим наборам данных, является аппроксимация распределения δ непрерывным распределением Пирсона (тип III), которое, в случае справедливости нулевой гипотезы, имеет среднее m_δ, стандартное отклонение s_δ и асимметрию g_δ. Тогда тестовая статистика, аналогичная t-критерию Стьюдента, будет иметь вид $T = (\delta - m_\delta)/s_\delta$. Чем меньше значение T (при $T < 0$), тем сильнее отличаются средние значения сходства для выделенных и «перемешанных» групп и тем выше дискриминирующая ценность группировки. Если задаться вопросом, достоверно ли отличается эмпирическое значение δ_{obs} от полученного распределения всех возможных дельт, то вероятность того, что δ_{sim} превысит величину δ_{obs} будет соответствовать вероятности p, связанной с вычисленным T-критерием.

Кроме значения p, необходимого для проверки статистической значимости группировки, может быть вычислена также оценка эффективности априорной классификации. Индекс внутригрупповой однородности, независимый от объема выборки и показывающий превышение среднего уровня сходства объектов внутри классов по сравнению со случайным размещением, может быть вычислен как $A = 1 - \delta / m_\delta$. Если $A < 0$, то однородность внутри выделенных групп ниже, чем при случайном их формировании, а при $\delta = 0$ и $A = 1$ группы состоят из совершенно идентичных объектов. В качестве порогового значения принято использовать $A = 0{,}1$. Если значение индекса больше этого порога, то можно говорить о существенной ценности классификации, а при $A = 0{,}3$ – об очень высокой.

Сделаем сравнительный анализ эффективности различных вариантов группировки гидробиологических проб по участкам речной экосистемы Сок–Байтуган. В качестве исходной матрицы используем значения численностей 375 видов, найденных в 147 пробах. Матрицу расстояний **D** между пробами вычислим по формуле Съеренсена для количественных признаков $d_{ij} = \frac{\sum_{k=1}^{p} |a_{ik} - a_{jk}|}{\sum_{k=1}^{p} a_{ik} + \sum_{k=1}^{p} a_{ik}}$ [6], где a_{ik} и a_{jk} – прологарифмированные численности k-го вида в i-й и j-й пробах.

[6] Этот показатель известен также как модификация коэффициентов Ренконена–Брея–Кёртиса–Чекановского, относительного манхеттенского расстояния и др.

Расчеты, выполненные с использованием процедуры MRPP (см. табл. 6.13), не позволяют сделать вполне однозначных выводов. С одной стороны, бесспорно заключение о статистической значимости всех априорных классификаций по сравнению со случайным перемешиванием сделанных гидробиологических проб ($p \cong 0$), причем абсолютное значение T-критерия увеличивается с ростом агрегированности выделяемых групп. В то же время величина индекса внутригрупповой однородности A свидетельствует о существенной неоднородности выполненных группировок видов, причем степень «размытости» классификации возрастает с уменьшением числа групп.

Таблица 6.13

Оценка эффективности разбиения речной экосистемы Байтуган–Сок на участки с использованием процедуры MRPP

	Показатели и группы			Уровень 1 – 13 участков	Уровень 2 – 6 участков	Уровень 3 – 3 участка
	река	станция	проб			
Среднее внутригрупповое расстояние	Байтуган	01	15	0,860	0,856	0,890
		03	14	0,847		
		06	13	0,928	0,913	
		08	13	0,886		
	Сок	01	8	0,769	0,804	0,852
		02	8	0,783		
		03	11	0,854	0,857	
		04	11	0,822		
		05	11	0,881		
		08	8	0,861	0,814	0,812
		09	11	0,834		
		12	11	0,682	0,787	
		13	13	0,787		
Индекс расстояния δ наблюдаемый				0,8346	0,8469	0,8552
Индекс расстояния δ по нуль-модели				0,8901	0,8901	0,8901
Индекс однородности A				0,0624	0,0486	0,0392
T-критерий для проверки H_0				-18,72	-23,12	-29,93
Вероятность p, связанная с T				0,0	0,0	0,0

Весьма близок описанной процедуре MRPP по своим основным предпосылкам другой метод ANOSIM (analysis of similarities; Clarke, 1993), который использует немного модифицированную тестовую статистику R для оценки межгрупповых различий. Она основана на разности средних рангов различия объектов между группами b и различия внутри групп w: $R = \dfrac{4(b-w)}{n(n-1)/4}$, где n – общее число анализируемых объектов. Знаменатель дроби выбран таким образом, чтобы величина тестовой статистики R всегда находилось в интервале [-1; +1], причем чем ближе значение R к единице, тем больше различаются сформированные группы. Отрицательное значение R свидетельствуют о том, что внутренняя неоднородность сформированных групп настолько велика, что превышает межгрупповые отличия. Статистическая значимость эмпирической тестовой статистики R оценивается с помощью описанной выше перестановочной процедуры.

6.7. Анализ компонентов экосистемы с использованием теста Мантеля

При изучении структуры экосистемы исследователь часто задается вопросом: изменятся ли результаты ординации местообитаний, если использовать несколько иной список видов (например, ограничиться только одной систематической группой или взять данные за разные экспедиционные периоды)? Поскольку в основе всех ординационных методов лежит расчет матрицы сходства или корреляции, то в математической плоскости эта проблема сводится к анализу: статистически значимы ли различия между двумя произвольными симметричными матрицами **A** и **B**.

Для оценки связи между двумя матрицами расстояний широко используется *тест Мантеля* (Mantel, 1967), который проверяет гипотезу, что расстояния (или близости) между объектами в матрице **A** независимы от расстояний (или близостей) между теми же самыми объектами в матрице **B**. Этот тест является альтернативой регрессии при анализе зависимости первой матрицы расстояний от второй, вычисляя, по сути, коэффициент корреляции между ними, не нуждаясь при этом в каких-либо строгих статистических предположениях.

В отличие от процедуры MRPP, которая сравнивает несколько групп неодинакового размера, составленных из произвольных совокупностей местообитаний, тест Мантеля предполагает, что обе матрицы расстояний **A** и **B** относятся к одному и тому же фиксированному набору местообитаний. В то же время этот метод позволяет найти линейные отношения между матрицами, составленными на основе меры расстояния любой природы. С использованием теста Мантеля можно, например, оценить обоснованность следующих содержательных гипотез:

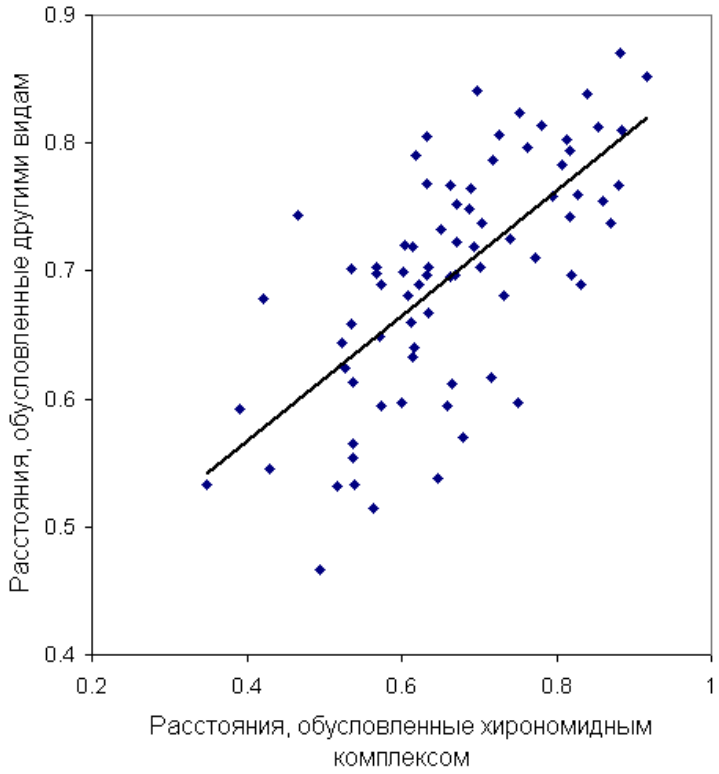
- имеется ли взаимосвязь между двумя группами организмов, обитающих в одних и тех же биотопах (например, включающих растения и беспозвоночные);
- имеются ли изменения в видовой структуре сообществ до и после нанесения возмущения;
- связаны ли статистически значимой зависимостью географическое расстояние между местообитаниями и экологическое сходство их видовых составов.

Если найти сумму элементов в матрице **Z = AB**, которая является произведением обоих сравниваемых матриц расстояний **A** и **B**, при этом исключая элементы на главной диагонали, то получим *Z*-статистику Мантеля, т. е. $Z = \sum_{i=1}^{n-1}\sum_{j=i-1}^{n} a_{ij}b_{ij}$. Стандартизованная статистика Мантеля r изменяется от +1 до –1 и соответствует *коэффициенту корреляции Пирсона* между матрицами расстояний **A** и **B**.

Оценим, имеется ли статистическая взаимосвязь между характером распределения 165 видов семейства хирономид (Chironomidae) и 210 остальных видов макрозообентоса по 13 створам рек Байтуган–Сок. Матрицы расстояний **A** и **B** между местообитаниями рассчитаем с использованием формулы Съеренсена:

$$a_{ij}(b_{ij}) = \frac{\sum_{k=1}^{p}\left|t_{ik} - t_{jk}\right|}{\sum_{k=1}^{p} t_{ik} + \sum_{k=1}^{p} t_{ik}}\ , \qquad (6.5)$$

где t_{ik} и t_{jk} – частота встречаемости k-го вида в пробах, взятых на i-й и j-й станциях. На рис. 6.12 представлен характер взаимосвязи между этими матрицами: каждая из точек биплота соответствует расстояниям между каждой парой местообитаний для двух композиций видов. Если эта зависимость имеет статистически значимый линейный характер, то все семейства макрозообентоса в одинаковой мере реагируют на биотопическую изменчивость по руслу водотока.



Эмпирические значения
Статистика Мантеля Z_{obs} = 19,6
Стандартизированная статистика Мантеля r = 0,657;

Аппроксимация Мантеля
- t-критерий = 5,84;
- p = 0,001

Перестановочный тест Монте-Карло
Параметры рандомизированного распределения:
- среднее Z_{sim} = 19,16:
- минимум Z_{sim} = 18,95;
- максимум Z_{sim} = 19,43;

Число перестановок N = 999;
Число значений с $Z_{sim} > Z_{obs}$ n = 0;
Число значений с $Z_{sim} < Z_{obs}$ – 999;
Вероятность p ошибки I рода = 0,001.

Рис. 6.12. **Характер взаимосвязи между значениями матриц расстояний, обусловленных двумя подмножествами видов макрозообентоса (реки Байтуган–Сок, 13 станций наблюдений)**

Для оценки нулевой гипотезы H_0 об отсутствии зависимости между матрицами **A** и **B** может быть использована *аппроксимация Мантеля* (Mantel, 1967), которая трансформирует Z-статистику в t-значение, распределение которого асимптотически приближается к нормальному. Метод асимптотической аппроксимации Мантеля дает хорошие результаты для больших наборов данных.

Если предполагается, что элементы матриц **A** и **B** рандомизированы (т. е. представлены случайным образом), то процессы, порождающие сравниваемые матрицы, являются независимыми. Поэтому оценка нулевой гипотезы H_0, не связанная с какими-либо априорными предположениями, может быть получена с помощью перестановочного теста: в обеих матрицах расстояний случайным образом меняют порядок рядов и строк. Если процедуру перестановки повторять многократно (например, L = 1000 раз), то можно смоделировать распределение значений статистики **Z***. Значимость связи между эмпирическими матрицами сходства оценивается как выход статистики Z за пределы правостороннего доверительного интервала распределения **Z***, а вероятность p ошибки I рода находится по формуле $p = (1 + l)/(1 + L)$, где l – число итераций, когда модельное значение Z_{sim} оказывалось больше эмпирического Z_{obs}.

Представленные на рис. 6.12 результаты теста Мантеля показывают статистически значимую связь между характером распределения двух комплексов видов макрозообентоса по течению водотока. Сходные выводы могут быть получены при сравнении матрицы расстояний **A**, рассчитанной по полному списку видового состава макрозообентоса, с двумя другими матрицами:

- **B´** по хирономидному комплексу – r = 0,756, t = 6,59, $p \cong 0{,}0$;
- **B´´** по списку видов, не включающих хирономид – r = 0,744, t = 6,50, $p \cong 0{,}0$.

Эти результаты свидетельствует о тесной экологической связи между обоими компонентами макробентоценоза.

Тест Мантеля стал основой процедуры *множественной матричной регрессии* (Quinn, Keough, 2002; Manly, 2007), приводящей к модели следующего вида: $\mathbf{Y} = \beta_0 + \sum_{i=1}^{m} \beta_i \mathbf{A}_i + \mathbf{E}$, где $\mathbf{Y}$ – зависимая матрица; $\mathbf{A}_i$ – матрицы, независимые от $\mathbf{Y}$; β_0 и β_i – свободный член и коэффициенты регрессионной модели; $\mathbf{E}$ – матрица ошибок (остатков). Метод специально предназначен для анализа матриц сходства или корреляции, у которых элементы, расположенные по главной диагонали, равны 1. Множественная процедура Мантеля реализует при этом обычную линейную регрессию, однако вследствие специфической структуры исходных данных для оценки параметров модели используются перестановочные методы Монте-Карло.

Оценим, какой вклад вносят отдельные таксономические группы макрозообентоса (на уровне семейств и подсемейств) в общую изменчивость донных сообществ по руслу рр. Байтуган–Сок. Рассчитаем частные матрицы сходства, составленные из коэффициентов Брея–Кёртиса, для видов семейств Ephemeroptera ($\mathbf{A}_1$), Oligochaeta без *Nais* sp. ($\mathbf{A}_2$), Trichoptera ($\mathbf{A}_3$) и Coleoptera ($\mathbf{A}_4$), а также для отдельных таксонов хирономид Orthocladiinae ($\mathbf{B}_1$), Tanypodinae ($\mathbf{B}_2$), Chironomini ($\mathbf{B}_3$) и Tanytarsini ($\mathbf{B}_4$). В качестве зависимых матриц будем использовать аналогичные матрицы сходства, рассчитанные по всему комплексу видов (отдельно для хирономид и остальных таксонов). Результаты множественного матричного регрессионного анализа представлены в табл. 6.14.

Таблица 6.14

Параметры моделей множественной матричной регрессии для оценки вклада отдельных групп макрозообентоса в изменчивость донных сообществ речной экосистемы Байтуган–Сок

	Коэффициенты β	**Стандартное отклонение**	***t*-критерий Стьюдента**	***p*-значение вероятности**	**Коэффициент детерминации (R^2) и критерий Фишера (F)**
Модель **V** – виды, не относящиеся к семейству Chironomidae					
Свободный член	0,127				$R^2 = 0{,}552$ $F = 22{,}5$ ($p = 0{,}001$)
Ephemeroptera	***0,334***	0,041	8,13	***0,001***	
Oligochaeta	0,085	0,046	1,81	0,121	
Trichoptera	0,083	0,046	1,80	0,092	
Coleoptera	-0,007	0,031	-0,22	0,840	
Модель **C** – виды, относящиеся к семейству Chironomidae					
Свободный член	-0,040				$R^2 = 0{,}888$ $F = 144{,}8$ ($p = 0{,}001$)
Orthocladiinae	***0,491***	0,040	12,30	***0,001***	
Tanypodinae	0,014	0,024	0,58	0,670	
Chironomini	***0,381***	0,033	11,55	***0,001***	
Tanytarsini	***0,201***	0,035	5,64	***0,001***	

На основании результатов анализа можно сделать вывод, что при общей значимости моделей регрессии по критерию Фишера далеко не все таксономические группы вносят одинаковый вклад в объяснение вариации видовой структуры по руслу рек. В частности, для модели **V** статистически значимо отличается от 0 коэффициент регрессии для семейства Ephemeroptera, тогда как остальные группы макрозообентоса не оцениваются как значимые для объяснения вариации коэффициентов сходства участков. Для более однородного хирономидного комплекса (модель **C** в табл. 6.14) основные группы видов, за исключением Tanypodinae, вносят приблизительно одинаковый вклад в пространственную изменчивость видовой структуры.

ГЛАВА 7. МЕТОДЫ СРАВНЕНИЯ КОМПОЗИЦИЙ ВИДОВ В ЭКОЛОГИЧЕСКИХ ПОСЛЕДОВАТЕЛЬНОСТЯХ МЕСТООБИТАНИЙ

Теория – это когда вы знаете всё,
но ничего не работает.
Практика – это когда всё работает,
но никто не знает, почему.
Научный фольклор

7.1. Модели распределения видов по экологическим градиентам

Любая экосистема, реально существующая в природе, может быть рассмотрена как часть пространства, ограниченная тем или способом в соответствии с целями исследования. Каждый биотический и абиотический компонент выделенного фрагмента и структура их взаимосвязей неоднородны относительно пространственных осей координат и постоянно изменяются во времени. Изменчивость компонент происходит под воздействием одного или нескольких экологических факторов (Фёдоров, Гильманов, 1980), под которыми понимается любая внешняя или внутренняя движущая сила, модифицирующая показатели жизнедеятельности экосистемы. Для системного анализа этих взаимосвязей введены важнейшие абстрактные понятия общей экологии: «*пространство экологических факторов*» и «*функция отклика экосистемы*».

Количественная оценка реакции различных видов живых организмов на внешние и внутренние возмущения является основной задачей моделирования окружающей среды (habitat models). В зависимости от природы воздействий, выделяют три основные группы экологических факторов (Austin, 2007):

- факторы, оказывающие прямое физиологическое воздействие на каждую конкретную особь и развитие всей популяции в целом (температура, минерализация, pH и др.);
- косвенные факторы, которые статистически тесно связаны с прямыми факторами, но непосредственного воздействия на биоту не оказывают (широта, высота, глубина водотока, расстояние от устья и т. д.);
- ресурсы жизнедеятельности, прямо или опосредованно потребляемые организмами и лимитирующие их численность (органическое вещество, биогенные вещества и пр.).

Реакция сообществ на изменение условий среды, выражающаяся в изменении обилия и видового состава популяций, зависит от всего комплекса воздействующих факторов. Однако круг анализируемых экологических данных обычно лимитируется возможностями массового мониторинга, как показателей окружающей среды, так и ключевых экофизиологических параметров биоты. Разумеется, качество модели повышается, если в роли переменных используются только прямые факторы и показатели обеспеченности ресурсами. Однако они не всегда доступны для измерения, и при крупномасштабных исследованиях все чаще используются косвенные факторы.

Распространенные в 70-е годы XX в. концепции *прямого градиентного анализа*, учитывающие только о д и н параметр среды, показали свою ограниченность (несостоятельность) в сколько-нибудь сложных региональных условиях (Василевич, 1972; Галанин, 2005). Традиционным методом многомерного анализа является редукция исходного множества переменных и сведение их к небольшому количеству латентных факторов, обладающих некоторым внутренним единством и соотносимых с определенным биологическим смыслом. Для этого используется, представленный в главе 6 *анализ главных компонент*, формирующий в пространстве признаков набор новых ортогональных осей, прохо-

дящих сквозь эллипсоид облака точек в оптимальных направлениях максимальной изменчивости.

Две главные компоненты (рис. 7.1), обобщающие 12 различных геофизических и гидрохимических показателей, измеренных на водотоке Байтуган–Сок, объясняют 68,3% общей изменчивости этих признаков, т. е. при переходе к двум осям мы теряем не более 32% исходной информации, представляющей, в основном, «стохастический шум», неизбежно присутствующий в реальных данных. Первый фактор с большой очевидностью может быть связан с пространственным градиентом условий и объединяет высоту над уровнем моря, глубину и температуру воды в месте отбора проб, изменчивость грунтов и т. д. Это убедительно представлено на диаграмме расположения станций в пространстве факторов. Если предположить, что в первой компоненте сосредоточена вся пространственная изменчивость, то в следующих компонентных осях она «снята», и второй фактор связывается нами с органическим загрязнением и насыщенностью кислородом. Большу́ю роль тут играет и площадь водосбора, как источник аккумуляции органического вещества.

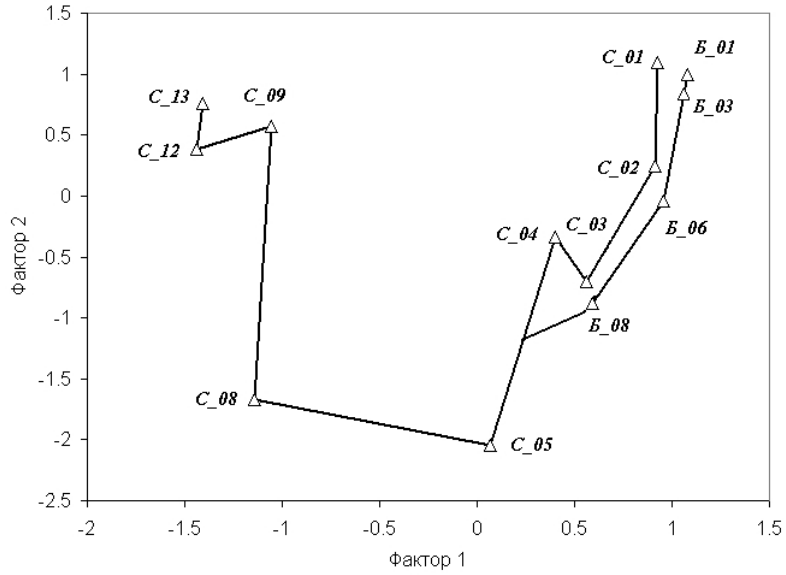


Рис. 7.1. **Диаграмма расположения участков рек Сок (префикс «С») и Байтуган (префикс «Б») в пространстве двух главных компонент и относительные нагрузки исходных переменных, вносимые в эти факторы**

Показатель	Фактор 1	Фактор 2
Доля каменисто-гравийных субстратов	**0,154**	0,027
Высота над уровнем моря	**0,149**	0,116
Азот нитритный	**-0,132**	0,046
Глубина отбора проб	**-0,148**	0,038
Доля ила и песка	**-0,149**	-0,034
Температура	**-0,157**	0,041
pH дна	-0,112	-0,018
Минеральный фосфор	-0,123	**-0,226**
Скорость течения	0,078	**-0,312**
Площадь водосбора	0,014	**-0,332**
Бихроматная окисляемость	-0,016	**-0,291**
Содержание кислорода	-0,050	**0,276**

Расчет главных компонент является основной процедурой PCNM-анализа (principal coordinates of neighbor matrices – главные координаты соседства матриц), разработанного Д. Боркардом и П. Лежандром (Borcard, Legendre, 2002). В соответствие с этими рекомендациями будем ниже использовать значения фактора 1 (рис. 7.1) как направление *комплексного пространственного градиента X*, обобщающего изменение условий среды по руслу водотока. В качестве первой задачи рассмотрим методы моделирования популяционной плотности отдельных видов по шкале градиента.

Значительная часть традиционно используемых статистических методов предполагает, что кривая зависимости обилия y от величины воздействующего фактора x имеет симметричную форму колоколообразной гауссианы с тремя интерпретируемыми параметрами: $y = he^{-(x-\mu)^2/2\sigma^2}$, где μ – локальный оптимум на оси x, которому соответствует мак-

симум обилия вида h; σ – стандартное отклонение на шкале градиента относительно этого оптимума. Эта форма отклика, теоретически основанная на представлениях о лимитирующем ресурсе и *законе Либиха*, позволяет определить ширину ниши $\pm 2\sigma$, т. е. диапазон толерантности и экологический оптимум каждого вида, а также идентифицировать его стратегию как «универсал» или «специалист».

Отступление от темы (*эмпирико-статистические рассуждения*). Многочисленные исследования показали, что в реальных экосистемах зависимость в виде гауссианы наблюдается весьма редко (например, трудно ожидать идентичную симметричную биологическую реакцию популяций на недостаток или переизбыток пищевого ресурса). Р. Уиттекер (1980) определяет пять основных типов характера кривой отклика на шкале градиента: *симметричный*, *асимметричный*, *монотонный*, (*поли*)*мультимодальный* и *платообразный*.

В рамках *прямого градиентного анализа* по о д н о м у фактору среды наиболее широко используется схема *однофакторного дисперсионного анализа* (Миркин, 1971; Миркин, Розенберг, 1978). При этом 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: *средневзвешенной напряженность фактора*, *средневзвешенной дисперсии*, *силы влияния* исследуемого фактора на данный вид, получаемой из однофакторного дисперсионного анализа. Однако в эксперименте, как справедливо подчеркивает В.И. Василевич (1969, с. 8), трудно добиться «изменения одного фактора при постоянстве других». Поэтому более корректным представляется использование 2-х, 3-х и т. д. факторных комплексов, для получения которых на практике следует прибегать к достаточно сложным схемам планирования экспериментов (см.: Проблемы экологического.., 2008). Более того, если эксперимент не планировался для многофакторного дисперсионного анализа, то отсутствуют исходные данные для построения дисперсионного комплекса, и, следовательно, отсутствует возможность воспользоваться результатами этого анализа. Поэтому некоторые авторы при оценке силы влияния исследуемого фактора на данный вид прибегают к ряду эвристических процедур. Один из способов состоит в сглаживании влияния неконтролируемых факторов на изучаемый путем многократного выравнивания распределения вида *методом скользящей средней* (Розенберг, 1984, 2007), что позволяет 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. Определим средние после выравнивания по следующей формуле:

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

для всех $1 < i < r$; r – число градаций исследуемого фактора. 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.

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. После преобразований общая дисперсия исходного комплекса C_y представляется в следующем виде (зависимость между дисперсиями для безграничных свя-

занных рядов в общем случае получена Е.Е. Слуцким [1927]):

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

где $n \cdot \overline{C}_x / 3$ и C_z – это дисперсии, соответственно, 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:

$$\eta^2_x = t\, C_x / C_y ,$$

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 Стъюдента. При количестве выравнивании $h > 1$ 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 экоклины) значительно усложняется процесс нахождения коэффициента t, так как он является функцией от числа наблюдений. В качестве оценки величины n можно использовать среднее значение числа наблюдений в каждой градации фактора. Как и любая эвристическая процедура, выравнивание методом скользящей средней статистически мало (или совсем не-) корректно, но позволяет избежать сложностей, связанных с необходимостью планирования эксперимента с учетом воздействия на распределение вида многих факторов. Рассматриваемые далее статистические методы моделирования различных форм асимметричных кривых отклика – это шаг в направлении от эвристических к более строгим и статистически достоверным методам.

Проявление конкретной функции распределения обилия вида в определенной мере зависит от типа изучаемого фактора. Например, для прямого фактора можно ожидать *унимодальную* форму с той или иной степенью *асимметрии*. Эффект влияния косвенных факторов, которые часто представляют собой сложный коррелированный комплекс трудно идентифицируемых воздействий, редко бывает «чистым» и имеет ярко выраженный *полимодальный* характер. Поэтому, сохранив традиционную симметричную гауссову модель только как образец для сравнения, были разработаны статистические методы моделирования различных форм асимметричных кривых отклика (Guisan, Thuillier, 2005):

- обобщенные линейные модели GLM (generalized linear model);
- обобщенные аддитивные модели GAM (generalized additive model);
- модели HOF Хаусмана–Олфа–Фреско (Huisman et al., 1993), являющиеся иерархическим объединением пяти моделей с увеличивающейся сложностью;
- обобщенные линейные модели пространственного распределения (GRASP; Lehman et al., 2002) и модели сверхниши (hyper niche; McCune, 2006).

Я. Оксанен и П. Минчин (Oksanen, Minchin, 2002) обосновали характер асимметричных моделей отклика с позиций теории континуума, а также разработали программные модули реализации большинства из них в статистической среде программирования **R** вместе с детальной технической спецификацией скриптов. Перечисленные методы аппрокси-

мации позволяют гибко осуществлять подбор наилучших моделей кривых отклика, при этом каждая из процедур имеет свои специфичные ограничения, преимущества и недостатки.

Модели GLM и GAM являются частным случаем обобщенного уравнения нелинейной регрессии

$$y = g^{-1}\left[\sum_{i=1}^{n} a_i q_i(x_i)\right], \qquad (7.1)$$

где $g(y^{-1})$ является *функцией связи* (link function), а $q_i(x_i)$ – произвольными функциями нелинейного преобразования независимых факторов. Если принимаются некоторые условия тождественности этих функций, то мы получаем следующие классы моделей:

- при $q_i(x_i) = x_i$ – обобщенную линейную модель $y = g^{-1}\left(\sum_{i=1}^{n} a_i x_i\right)$;
- при $g(y') = y$ – обобщенную аддитивную модель $y = \sum_{i=1}^{n} a_i q_i(x_i)$;
- при $g(y') = y$ и $q_i(x_i) = x_i$ – обычную линейную регрессию $y = \sum_{i=1}^{n} a_i x_i$.

Обобщенные линейные модели (McCullagh, Nelder, 1989) в зависимости от используемой функции связи могут задавать различный характер кривой изменения обилия вида y. Легко проследить, например, математическую идентичность гауссовой модели отклика и GLM в виде полинома 2-го порядка при логарифмической функции связи

$$\log(y) = a_0 + a_1 x + a_2 x^2 \quad \Leftrightarrow \quad y = h\exp[-(x-\mu)^2 / 2\sigma^2],$$

если удовлетворяются условия $\mu = -a_1/2a_2$; $\sigma = (-1/2a_2)^{0,5}$; $h = \exp(a_0 - a_1^2/4a_2)$. Использование иных коэффициентов или полиномов 3-го и более порядков будет подчеркивать отличия асимметрических отклонений функции отклика от гауссианы.

Следует также отметить, что при $g(y) = \log(y)$ предполагается распределение ошибок регрессии по *закону Пуассона*, поэтому использование для настройки модели метода наименьших квадратов не является корректным. Коэффициенты GLM-моделей a_0, a_1, a_2 и т. д. могут быть найдены с использованием итеративной нелинейной процедуры оценки наибольшего правдоподобия (Oksanen, Minchin, 2002). Другой проблемой является поиск набора предикторов, приводящего к модели оптимальной сложности. В частности, в случае одного фактора необходимо подобрать степень «наилучшего» аппроксимирующего полинома. Для этого предлагается использовать *критерий Акаике*, *информационный критерий Байеса* или иные методы (Burnham, Anderson, 2004).

Отклик y может выражаться дихотомической переменной, связанной только с фактом встречаемости вида. Если p – доля единичных значений y, а $(1 - p)$ – доля нулевых значений для того же вида, то функция связи обычно задается в виде логита (logit link): $g(y) = \log(\frac{y}{1-y}) = \log(\frac{p}{1-p})$. Соответствующая ей обобщенная логит-линейная модель $\log\frac{p(x)}{1-p(x)} = \sum_{i=1}^{n} a_i x_i$ характеризуется кривой S-образной формы и предполагает биномиальный закон распределения ошибок.

Представляет интерес для экологов также β-функция, предложенная для градиентного анализа (Austin, 1976) и используемая для оценки асимметрии и эксцесса распределения отклика (Oksanen, Minchin, 2002): $y = k(x-a)^{\alpha}(b-x)^{\gamma}$, где a и b – граничные значения, k – ширина полосы отклика, α и γ – параметры формы кривой. Эквивалентная запись в логарифмической форме $\log(y) = \log(k) + \alpha\log(x-a) + \gamma\log(b-x)$ позволяет рассматривать β-функцию, как модель GLM с пятью параметрами.

Для обобщенных аддитивных моделей GAM процесс настройки заключается в поиске непараметрических функций произвольного вида $g_i(x_i)$, которые связывают расчетные значения переменной y со значениями предикторов x_i. Уникальным аспектом GAM является использование различных общих методов сглаживания диаграмм рассеивания, к

которым относится аппроксимация кубическими сплайнами или полиномами более высоких порядков.

Обобщенные аддитивные модели являются очень гибким инструментом и могут обеспечить идеальную подгонку кривой даже при наличии сильных нелинейных связей и значительной зашумленности исходных факторов. Но здесь важным компонентом подбора моделей является выбор разумного (т. е. не слишком большого) числа узлов сглаживания, чтобы избежать вырождения моделируемой зависимости, которая часто приобретает нереалистичный характер. Кроме того, если модели GLM полностью описываются уравнениями с фиксированным и четко объясняемым числом параметров, то уравнения GAM являются непараметрическими, и поэтому их очень трудно содержательно интерпретировать.

Модель HOF (Huisman et al., 1993) может быть в общем виде представлена, как $y(a, b, c, d, x, M) = \frac{M}{[1+\exp(a+bx)][1+\exp(c-dx)]}$, т. е. отклик y зависит от значения градиента x, максимально возможного значения M и четырех параметров $\{a, b, c, d\}$. Каждый из параметров может быть свободно варьируемым или равным определенному фиксированному значению и, в соответствии с этим, модель может принимать иерархическое множество состояний из пяти возможных форм (I–V):

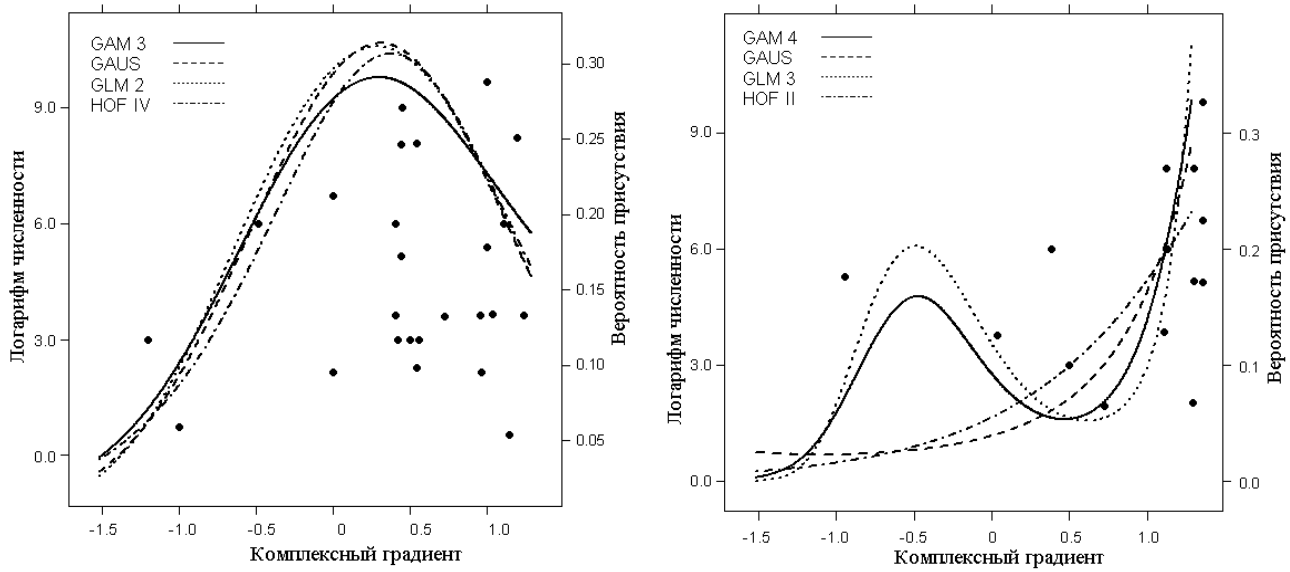
Формы модели HOF	Список параметров				Кол-во
• V – асимметричный унимодальный отклик	***a***	***b***	***c***	***d***	4
• IV – симметричный унимодальный отклик	***a***	***b***	***c***	*b*	3
• III – монотонный рост с «плато»	***a***	***b***	***c***	∞	3
• II – монотонный рост	***a***	***b***	0	0	2
• I – «плато» (отсутствие отклика)	***a***	0	0	0	1

Самая сложная форма модели HOF – асимметричная кривая, включающая полный комплект из всех четырех параметров. Иерархичность здесь заключается в том, что более простая модель может быть получена из более сложной модели путем фиксации значений одного из параметров (Oksanen, Minchin, 2002). Модели HOF позволяют гибко учитывать всю совокупность априорных исходных ограничений и теоретических предположений, традиционно связываемых с характером кривых отклика, и, вероятно, предоставляют наилучший результат с экологической точки зрения. Например, можно задать поиск только унимодальных зависимостей, которые необходимы для нахождения оптимума вида на градиенте. Модели GLM и GAM с низким числом степеней свободы в этих условиях могут легко привести к «полубимодальной» (semi-bimodal) зависимости с неясной содержательной интерпретацией.

Как и любая модель, кривые отклика – только упрощенное отображение действительности, и полученные результаты в значительной мере зависят как от принятых допущений, так и используемого набора данных. Одно из главных предположений градиентного анализа, традиционно ориентировавшегося на геоботанические представления, – унимодальный характер зависимости обилия вида от фактора среды. Для его проверки выполним с использованием различных моделей GAM, GML и HOF оценку распределения логарифма численности особей отдельных видов макрозообентоса рек Байтуган–Сок по шкале комплексного градиента X (см. фактор 1 на рис. 7.1). Как показывают примеры, представленные на рис. 7.2, иногда весьма трудно сделать однозначный вывод о характере зависимости. Если для *Baetis rhodani* все три модели практически совпали с гауссианой (GAUS), то для *Paracladius conversus* интерпретация осталась неясной, поскольку, в отличие от модели HOF, обобщенные модели GLM на основе полинома 3-й степени и GAM с 4-мя узлами сглаживания уверенно демонстрируют полубимодальную кривую.

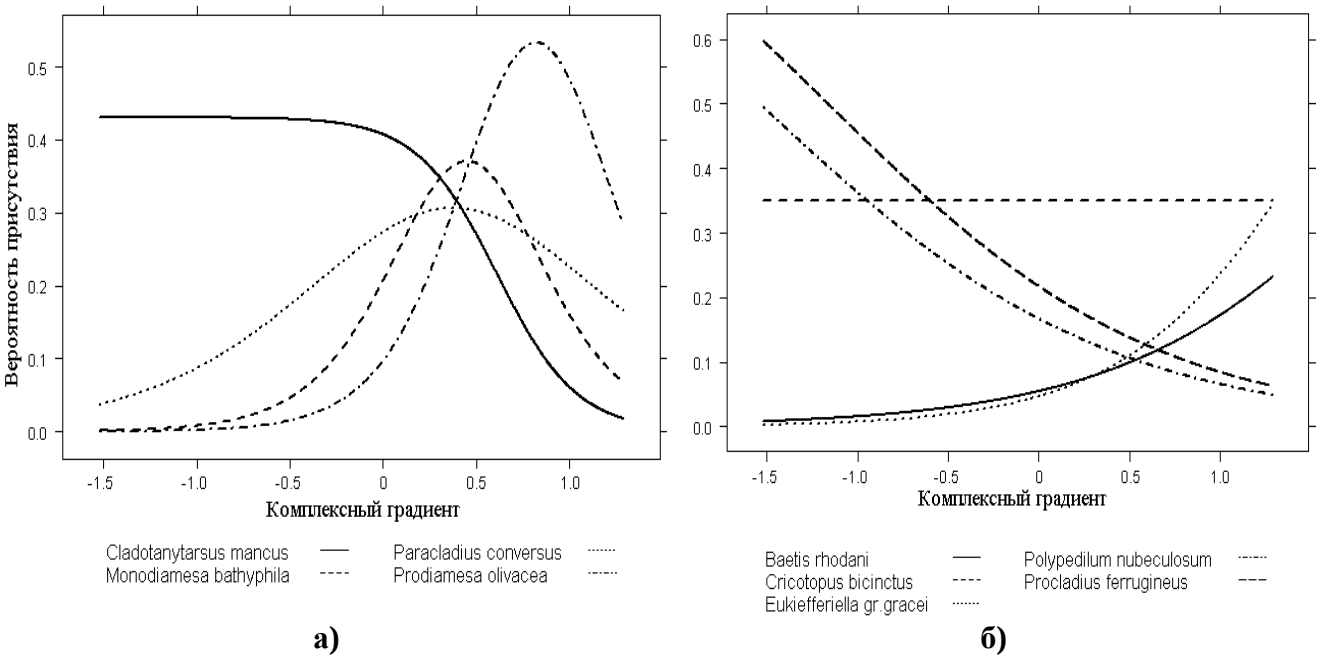
Классическая унимодальная форма кривой отклика с максимумом в середине шкалы продольного градиента X оказалась характерной для немногочисленного и мало интересного с точки зрения биоиндикации подмножества видов с широким диапазоном толерантности к абиотическим факторам (рис. 7.3а). Для некоторых видов (например, *Tany-*

tarsus sp.) это распределение имело неопределенную многовершинную форму, свойственную видам-эврибионтам. Для значительного числа редко встречающихся видов говорить о каком-то распределении по градиенту вообще было некорректно, так как их встречаемость носит «штрих-пунктирный» или «точечный» характер.



а) ***Baetis rhodani*** **б)** ***Paracladius conversus***

Рис. 7.2. **Оценка распределения обилия двух видов макрозообентоса на шкале комплексного градиента с использованием различных моделей отклика**



а) **б)**

Рис. 7.3 **Распределение обилия некоторых видов макрозообентоса в гидробиологических пробах по шкале комплексного градиента с использованием модели HOF**

Значительно более важными в биоиндикационном отношении оказались виды бентофауны с максимумом встречаемости на обоих концах шкалы градиента *X*, модели отклика которых носила специфический сигмоидальный характер. Сигмоиды с максимумов в

правой части шкалы X объединили структурообразующие таксоны (*Prodiamesa olivacea, Eukiefferiella* gr. *gracei, Parametriocnemus lundbecki, Eukiefferiella claripennis, Odontomesa fulva, Parametriocnemus* sp.), характерные для реофильного комплекса донных сообществ р. Байтуган и верховьев р. Сок. С другой стороны, S-образная кривая с максимумом в левой части шкалы X была характерна для пелофильных и лимнофильных видов (*Chironomus obtusidens, Pisidium amnicum, Microchironomus tener, Ablabesmyia monilis, Dicrotendipes nervosus, Euglesa henslowana, Dikerogammarus haemobaphes*), что связывается с воздействием факторов снижения течения и заиленности грунтов в нижнем течении р. Сок.

Пока трудно дать однозначное объяснение причинам этих отклонений от общеэкологической теории:

- связаны ли они с изучением всех речных экосистем,
- являются ли они общим свойством при изучении пространственных градиентов в шкалах среднего масштаба (10-50 км) или
- они характерны для сообществ, имеющих значительное число редких видов, какими являются макрозообентоценозы.

Но в любом случае очевидно, что градиентный анализ воспроизводит динамику отклика лишь для отдельных доминирующих или «важных» видов, но не дает возможности оценить совокупную реакцию на изменение факторов среды в целом для всего донного сообщества, образующего сложный комплекс из большого числа экологически взаимосвязанных компонент. Об общих закономерностях речного континуума можно корректно говорить, лишь изучая все сообщество как единое целое.

7.2. Оценка статистической значимости вариации экологических показателей по течению водотоков

Представленные выше модели распределения популяционной плотности по продольному градиенту абиотических условий в водотоках позволяют детально изучить характер этих зависимостей, однако требуют в каждом частном случае трудоемкого анализа и в силу своей многовариантности не всегда дают возможность однозначно оценить статистическую значимость экологических последствий воздействия факторов среды. С практической точки зрения более эффективен классический корреляционный тест, позволяющий проверить, имеют ли изменения обилия или встречаемости видов по изучаемому градиенту статистически значимый монотонный характер. Для этого достаточно вычислить коэффициенты ранговой корреляции τ Кендалла между экологическими показателями донных сообществ и связанным с ними комплексным пространственным градиентом параметров среды (фактор 1 на рис. 7.1).

Для оценки статистической значимости корреляционной связи проводилась рандомизационная процедура многократного перемешивания данных об обилии по шкале градиента. В ходе имитации для каждого вида макрозообентоса восстанавливалось распределение возможных значений коэффициента τ Кендалла и рассчитывались значения модельных среднего m_{sim} и стандартного отклонения δ_{sim}. Отклонить нулевую гипотезу об отсутствии монотонной зависимости можно при соблюдении двух условий:

- коэффициент τ_{obs}, найденный по эмпирическим данным, находится вне вычисленных доверительных интервалов;
- величина критерия $Z = (\tau_{obs} - m_{sim}) / \delta_{sim}$ превышает критическое значение.

Виды макрозообентоса, имеющие наиболее выраженную положительную и отрицательную монотонную зависимость встречаемости по шкале продольного градиента экосистемы Байтуган–Сок, представлены в табл. 7.1.

Как и в случае моделей градиентного анализа, статистические связи, найденные для отдельных видов, не дают возможность сделать какой-либо определенный вывод относи-

тельно всего сообщества в целом. В частности, оценка средней корреляции рангов в значительной мере осложняется наличием большого числа видов, которые встретились только на одной (161 вид, или 43%) или двух (85 видов, или 27%) станциях, в результате чего делать заключение о монотонном характере динамики видов не вполне корректно. На рис. 7.4 показано распределение относительной вероятности *Z*-критерия для всех 375 видов макрозообентоса, из которого следует, что только 63 вида (16,8%) имеют значимую линейную статистическую связь с продольным градиентом водотока, т.е. $|Z| > 1{,}96$. Если рассмотреть только 129 видов, которые встретились в трех и более местообитаниях, то доля таковых оказывается несколько больше – 33,3%.

Таблица 7.1

Виды макрозообентоса, имеющие выраженную линейную связь с продольным градиентом рек Байтуган–Сок, и оценка значимости коэффициентов ранговой корреляции

Виды и таксоны	*V*	τ_{obs}	m_{sim}	δ_{sim}	LCL	UCL	Z
Виды, выпадающие из донных сообществ от истока к устью							
Dicranota bimaculata	11	-0,9133	-0,0003	0,2878	-0,5407	0,5480	-3,1721
Prodiamesa olivacea	9	-0,8360	-0,0038	0,2927	-0,5574	0,5651	-2,8433
Eukiefferiella gr. *gracei*	8	-0,7791	-0,0051	0,2946	-0,5478	0,5965	-2,6267
Elmis aenea	4	-0,7520	-0,0080	0,2905	-0,5530	0,5751	-2,5614
Baetis rhodani	7	-0,7498	-0,0013	0,2907	-0,5453	0,5794	-2,5751
Micropsectra gr. *praecox*	7	-0,7457	-0,0057	0,2942	-0,5500	0,5500	-2,5153
Pristina rosea	4	-0,7443	0,0003	0,2952	-0,5513	0,5789	-2,5221
Eukiefferiella gr. *claripennis*	3	-0,7319	0	0,2955	-0,5855	0,5855	-2,4768
Elmis sp.	3	-0,7319	0	0,2955	-0,5855	0,5855	-2,4768
Odontomyia sp.	3	-0,7319	0	0,2955	-0,5855	0,5855	-2,4768
Paracladopelma gr. *Camptolabis*	9	-0,7244	-0,0028	0,2974	-0,5824	0,5824	-2,4263
Tvetenia sp.	3	-0,7128	-0,0057	0,2920	-0,5431	0,5431	-2,4216
Orthocladius (*Orth.*) *clarki*	5	-0,7127	-0,0031	0,2908	-0,5791	0,5345	-2,4400
Виды, появляющиеся в составе донных сообществ от истока к устью							
Unio sp.	3	0,7132	-0,0091	0,2850	-0,5706	0,5298	2,5350
E.henslowana	5	0,7135	-0,0060	0,2844	-0,5659	0,5659	2,5295
Piscicola geometra	5	0,7184	-0,0087	0,2934	-0,5916	0,5494	2,4782
Cladotanytarsus mancus	9	0,7403	0,0034	0,2848	-0,5535	0,5535	2,5878
Ablabesmyia monilis	5	0,7454	-0,0052	0,2874	-0,5510	0,5510	2,6113
P. nubeculosum	8	0,7879	-0,0073	0,2837	-0,5342	0,5142	2,8032
Cryptochironomus gr. *defectus*	7	0,8004	-0,0036	0,2868	-0,5593	0,5401	2,8034
P. ferrugineus	10	0,8037	-0,0123	0,2911	-0,5699	0,5334	2,8030
Microchironomus tener	6	0,8106	-0,0006	0,2918	-0,5639	0,5463	2,7797
Chironomus obtusidens	6	0,8108	-0,0101	0,2797	-0,5406	0,5225	2,9351
Caenis macrura	5	0,8420	-0,0059	0,2918	-0,5761	0,5650	2,9057

Примечание: *V* – число участков из 13, на которых встретился вид; LCL и UCL – нижняя и верхняя границы 95% доверительного интервала; остальные обозначения – по тексту.

Виды, для которых нулевая гипотеза об отсутствии линейной зависимости не отвергается, в свою очередь могут разделяться на виды, имеющие отчетливо выраженную унимодальность (или бимодальность) нелинейного тренда, и эврибионтные виды, равномерно или стохастически распределенные по градиенту. Для некоторых видов из этого числа была найдена значимая нелинейная зависимость с использованием моделей HOF (см. раздел 7.1), однако остается открытым вопрос о статистической достоверности влияния фактора среды для всего сообщества в целом.

Э. Пиелу (Pielou, 1984b) разработала непараметрический метод случайного зондирования (random skewers), который позволяет объективно судить, насколько существенна детерминированная тенденция в изменении структуры всего сообщества в целом вдоль изучаемого градиента. Данные мониторинга рассматриваются как «облако» (swarm) *m*

точек, соответствующих выполненным пробам и представленных в *s*-мерном пространстве обилия видов. Точки отбора проб выстраиваются в определенном порядке 1, 2, ..., *m* согласно их положению на водотоке или ином физическом градиенте.

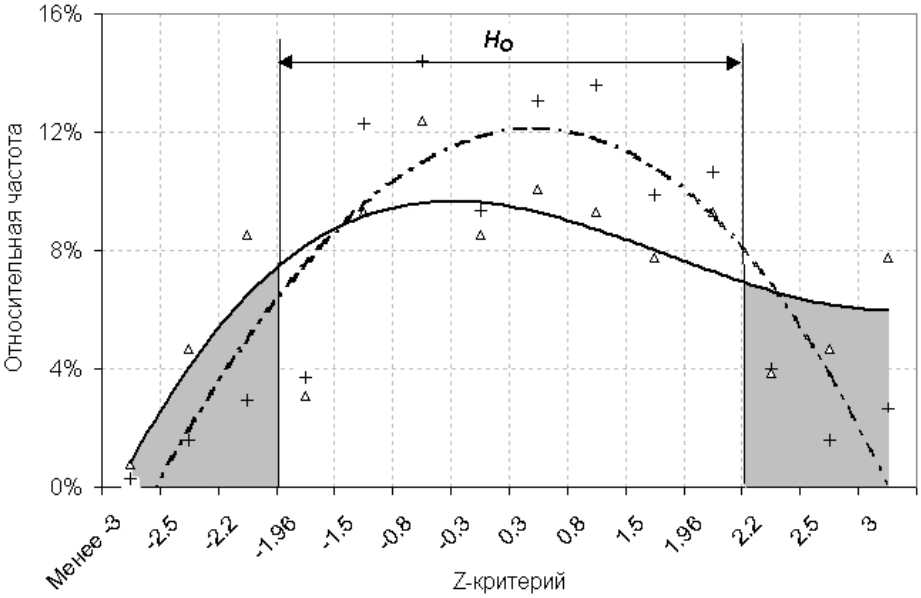


Рис. 7.4. **Распределение значений *Z*-критерия, оценивающего статистическую значимость связи встречаемости видов макрозообентоса с величиной продольного градиента**
(штрих-пунктирная линия – все виды, сплошная – виды, встретившиеся на трех или более станциях; показана область принятия нулевой гипотезы при 95% уровне значимости)

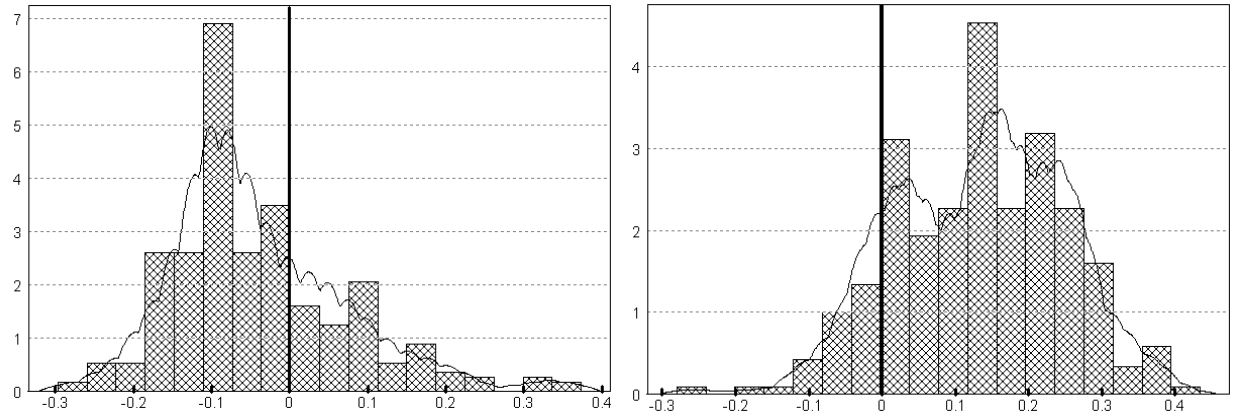
«Зонд», пронизывающий эту *s*-мерную структуру, представляет собой произвольно ориентированный вектор, координаты которого определяются *s* случайными числами, нормированными так, чтобы их сумма равнялась 1. Из каждой точки «облака» данных на зонд опускается перпендикуляр и фиксируется место его пересечения со случайным вектором. Далее вычисляется коэффициент ранговой корреляции τ Кендалла между порядком следования проекций *m* точек на зонд и натуральным рядом чисел от 1 до *m*. Если значение τ близко к 1 (или -1), то делается вывод, что точки мониторинговых данных имеют очевидную закономерную упорядоченность относительно некоторого направления в многомерном пространстве. При $\tau \cong 0$ можно предположить одну из двух причин:

- все экспериментальные точки беспорядочно расположены по всем направлениям или
- ориентация случайного зонда не совпадает ни с одной из возможных осей многомерного эллипсоида «облака» точек эмпирических данных.

Процедура «пронизывания данных зондом» проводится многократно (например, 300 раз), и формируется частотное распределение всех зарегистрированных значений коэффициента ранговой корреляции τ. Если в результате получается унимодальное распределение с математическим ожиданием $\tau \cong 0$, то в «облаке» точек нет ни одного направления, относительно которого эмпирические данные упорядочены вдоль изучаемого градиента, что соответствует нулевой гипотезе. Согласно альтернативной гипотезе, принимаемой, когда центр распределения $|\tau_{sim}| > 0$, точки наблюдений закономерно упорядочены относительно оси фактора. Бимодальное распределение с центрами, приближенными к +1 и -1, свидетельствует о существовании ярко выраженной детерминированной закономерности. Несколько вариантов проверки нулевой гипотезы об унимодальности эмпирическо-

го распределения τ, что означает отсутствие связи данных мониторинга с направлением по градиенту, представлено в работах (Pielou, 1984b; Perry, Schaeffer, 1987).

Тест случайного зондирования Пиелу был выполнен с использованием данных о встречаемости 129 видов макрозообентоса, обнаруженных на трех или более станциях (рис. 7.5а), и в их числе – на 66 таксонах, относящихся к семейству хирономид (рис. 7.5б). Для этих сообществ получены гистограммы эмпирического распределения значений τ, которые носят достаточно выраженный унимодальный характер. При этом доверительные интервалы для центров распределений не включают нулевое значение, что позволяет отклонить нулевую гипотезу. В качестве основного итога можно предположить в целом умеренную изменчивость тестируемых сообществ вдоль продольного градиента водотока.



а) для 129 видов макрозообентоса
(правая доверительная граница среднего
$m + t \cdot s_m$ = -0,04 + 1,96·0,0065 = -0,027)

б) для 66 таксонов Chironomidae
(левая доверительная граница среднего
$m - t \cdot s_m$ = 0,13 – 1,96·0,0067 = 0,119)

Рис. 7.5. **Гистограмма распределения коэффициента ранговой корреляции Кендалла, полученная в ходе 300 итераций по методу случайного зондирования**

Можно отметить, что сдвиг относительно нуля центра распределения τ для ценоза хирономид носит более акцентированный (и противоположный по знаку!) характер, чем для всего бентоценоза. Это вполне объяснимо, поскольку разные группы организмов имеют различный характер адаптационных реакций на изменяющиеся биотопические условия (а также различные стратегии жизненных циклов организмов) в направлении от истока к устью водотока. Чем разнообразнее состав сообщества, тем слабее выражена его совокупная экосистемная зависимость от градиента фактора.

Д. Перри и Д. Шеффер (Perry, Schaeffer, 1987) выполняли тест случайного зондирования на различных локальных группах бентосных организмов, комбинируя их по таксономической и функциональной принадлежности, и добивались при этом отчетливого «расщепления» пиков распределения значений τ. Эти исследования не нашли продолжения, поскольку на смену методу Пиелу скоро пришли более информативные и надежные методы многомерной ординации. Однако один из важнейших выводов был при этом сделан: *нельзя говорить о речном континууме вообще, а следует говорить о закономерностях изменчивости отдельных групп организмов с определенной экологической стратегией и адаптационными механизмами.*

Для иллюстрации теорий речного континуума и динамики пятен (см. главу 1), обычно используются обобщенные индексы, связанные со структурно-функциональными параметрами сообществ: *таксономическое богатство*, *коэффициент P/B*, *индексы разнообразия* и др. Прежде чем приступить к оценке стационарности рядов обобщенных показателей, введем понятие «*экологической последовательности*», под которой будем пони-

мать цепочку «срезов» состояния экосистемы вдоль характерного ландшафтного образования, например, по руслу течения водотока. На сети рек эти последовательности могут иметь сложную иерархическую и криволинейную структуру (рис. 1.5а), чем они отличаются от традиционных линейных ценотопоклинов, используемых в градиентном анализе. От обычных пространственных биогеографических моделей анализ последовательностей для водотоков отличается направленным характером выделяемых цепочек (например, при продвижении от истоков к устью) и неопределенностью значений экологических параметров вне трансекты наблюдений.

Рассмотрим динамику изменения следующих основных обобщенных показателей сообществ макрозообентоса в цепи экологических последовательностей, расположенных по течению рек Байтуган–Сок:

1) общее число зарегистрированных видов и среднее число видов в одной пробе;
2) индекс видового разнообразия Шеннона;
3) общее обилие бентоса по формуле $N_{сум}(r) = \sum_{i=1}^{s} \log(\sum_{j=1}^{p_r} N_{ij} / p_r)$, где p_r – число проб, выполненных на r-й станции, N_{ij} – численность особей i-го вида в j-й пробе;
4) доля обилия хищников-хватателей $N_h(r)/N_{сум}(r)$;
5) доля обилия хирономид групп Diamesinae+Orthocladiinae $N_{Di+Or}(r)/N_{сум}(r)$.

Показатели 1-3 представлены слева на рис. 7.6 в виде нормированных значений в диапазоне от 0 до 1, т. е. 0 соответствовал минимальной, а 1 – максимальной величине показателя. Доли обилия 4-5 в % представлены на рис. 7.6 справа.

Визуальный анализ графиков не всегда дает возможности выделить отчетливый детерминированный тренд: со всей очевидностью можно отметить лишь наличие гармонической составляющей синусоидального типа. Для более точных выводов о характере выборочных зависимостей выполним проверку трех статистических гипотез:

- наблюдается ли корреляция обобщенных показателей с продольным градиентом условий среды (фактор 1 на рис. 7.1)?
- являются ли ряды экологических последовательностей в целом стационарными?
- имеет ли место пространственная автокорреляция между отдельными участками водотока?

Абсолютные величины коэффициентов линейной корреляции Пирсона и ранговой корреляции τ Кендалла (табл. 7.2) имеют достаточно «пристойные» для экологических измерений значения, однако мы рекомендуем читателю обратить внимание в первую очередь на найденные бутстреп-методом p-значения, отражающие значимость этих коэффициентов. Очевидно, что говорить о наличии статистически достоверной детерминированной связи с продольным градиентом можно только в отношении обилия групп Diamesinae+Orthocladiinae, которое последовательно уменьшается от истоков к устью. Для других показателей принимается нулевая гипотеза о случайном характере распределения обобщенных показателей в пространстве факторов среды.

Следует также отметить низкую достоверность линейной связи между отдельными экологическими показателями: по существу, значимой оказалась лишь тривиальная зависимость между общим обилием бентоса и средним числом видов в пробе. Это свидетельствует о том, что оценка континуальности сообществ или выраженности сукцессионных изменений по различным обобщенным экологическим показателям может привести к неоднозначным, а часто и к диаметрально противоположным выводам.

Экологическую последовательность можно интерпретировать как случайный процесс, т. е. как ряд независимых величин $Y(r)$, принимающих случайные значения в дискретных точках пространства. Выборочные данные, содержащие векторы наблюдений над объясняемыми y_r и независимыми x_r переменными, а также вектор возмущений h_r являются реализациями этого случайного процесса и используются для анализа и моделирования. Важными характеристиками случайного процесса $Y(r)$ являются математическое ожидание $m_Y(r)$, дисперсия $D_Y(r)$ и автоковариация.

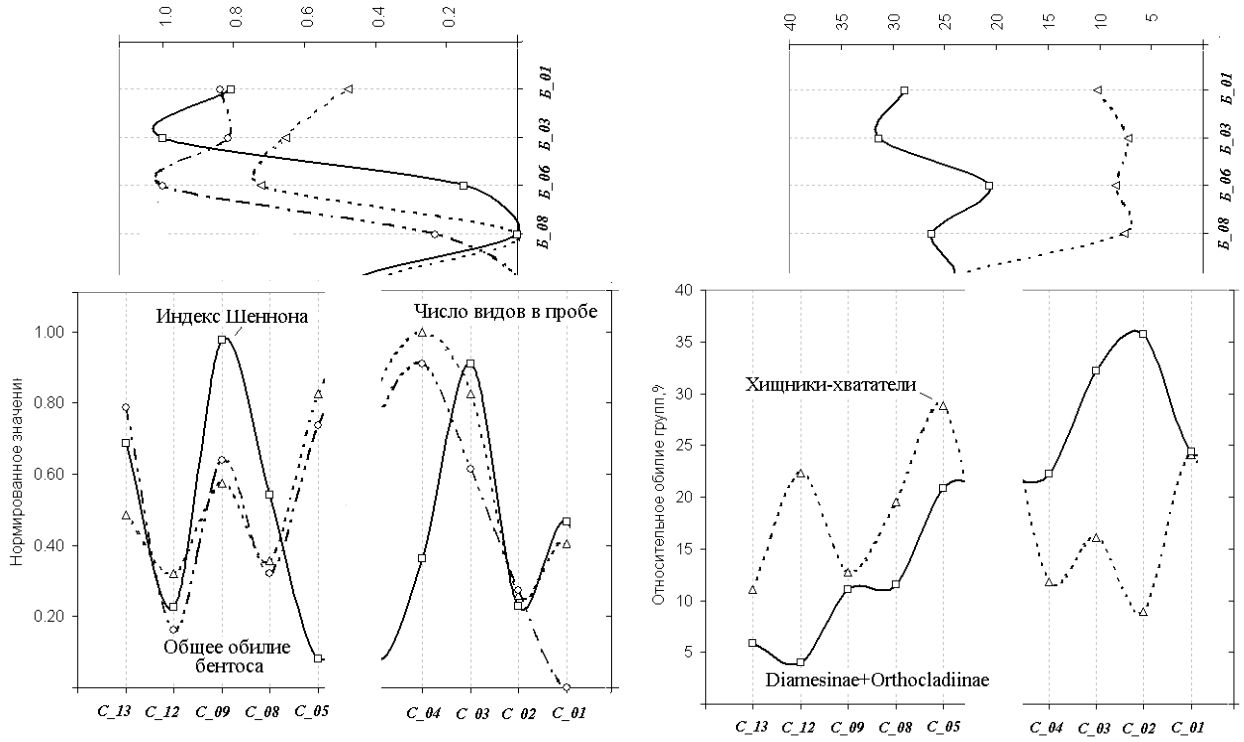


Рис. 7.6. **Изменение основных показателей сообществ макрозообентоса по руслу рек Байтуган (станции с префиксом «Б») и Сок («С»)**

Таблица 7.2

Коэффициенты корреляции Пирсона (ниже главной диагонали) и τ Кендалла (выше главной диагонали) между обобщенными показателями сообществ макрозообентоса

	0	1	2	3	4	5
0. Продольный градиент		0,09 (0,66)	0,08 (0,71)	0,18 (0,39)	-0,38 (0,07)	**0,59 (<0,01)**
1. Число видов в пробе	0,14 (0,63)		0,25 (0,24)	**0,48 (0,02)**	0,12 (0,58)	0,09 (0,66)
2. Индекс Шеннона	-0,08 (0,80)	0,27 (0,36)		0,18 (0,39)	-0,08 (0,71)	0,08 (0,71)
3. Общее обилие зообентоса	0,16 (0,59)	**0,73 (<0,01)**	0,29 (0,33)		-0,33 (0,11)	-0,03 (0,9)
4. Доля хищников хватателей	-0,33 (0,26)	0,13 (0,66)	-0,22 (0,46)	-0,40 (0,18)		-0,28 (0,17)
5. Диамезиды и ортокладины	**0,92 (≅0)**	0,09 (0,77)	0,05 (0,87)	0,07 (0,82)	-0,34 (0,25)	

Примечание: в скобках приведены вероятности ошибки 1-го рода при отклонении H_0.

В каждой точке последовательности r текущие значения параметров состояния сообществ **Y** определяются интенсивностью возмущений h_r в этой точке, которые включают в себя влияние всего комплекса факторов среды, биотопические условия, а также погрешности измерения и случайные флуктуации. В общем случае, в нестационарном динамическом ряду $y(r) = f(r) + g(r) + \varepsilon_r$ можно выделить три компоненты:

- $f(r)$ – детерминированную компоненту, представляющую собой некоторую аналитическую функцию, выражающую направленную тенденцию (тренд) в ряду динамики;
- $g(r)$ – стохастическую компоненту, моделирующую характер периодической и квазипериодической вариации исследуемого явления;
- ε_r – случайную компоненту типа «белого шума».

Основным предметом анализа последовательностей является проверка нуль-гипотезы о том, что изучаемый случайный процесс является *стационарным*, т. е. вероятностные характеристики процесса остаются неизменными на всей траектории (трансекте) **r** выделенных точек. Случайный процесс $y(r) = h_r$ является стационарным в широком смысле, если его математическое ожидание, дисперсия и автоковариация одинаковы для всех точек последовательности:

$$E(h_r) = \mu = const; \quad E(h_r - \mu) = \sigma^2 = const; \quad E((h_r - \mu)(h_r - k - \mu)) = \gamma_k = const.$$

Ответ на вопрос «определяется ли изменчивость параметров водных сообществ по течению водотока только чисто случайной нерегулярной составляющей, или их значения подчинены некоторой детерминированной или квазипериодической закономерности?» может дать комплекс критериев стационарности динамических рядов, которых разработано достаточно большое количество (Андерсон, 1976). Выбор наиболее подходящего из них связан с тем, какой смысл вкладывается в альтернативу проверяемой гипотезы: наличие тренда и какого именно типа, наличие периодичности или то и другое.

Наиболее распространенный тест для проверки динамического ряда на стационарность был разработан Д. Дики и В. Фуллером (Dickey, Fuller, 1979). При этом учитывается, что в экологических данных нестационарность чаще всего имеет характер случайного блуждания ($y_r = \rho y_{r-1} + \varepsilon_r$); случайного блуждания с дрейфом ($y_r = \alpha + \rho y_{r-1} + \varepsilon_r$) или комбинации случайного блуждания с линейным трендом ($y_r = \alpha + \rho y_{r-1} + \mu r + \varepsilon_r$). Нулевая гипотеза в *тесте Дики–Фуллера* состоит в том, что ряд нестационарен и имеет один единичный корень H_0: $\rho = 1$, $\mu=0$. При $\rho < 1$ нулевая гипотеза отвергается в пользу альтернативы и ряд считается стационарным. Для проверки H_0 оценивают статистику $t = (\rho - 1)/s_\varphi$, сравнивая её с табличными критическими значениями, полученными методом Монте-Карло. Если t-статистика незначима и соответствующее ей p-значение больше 0,05, то нулевая гипотеза принимается и ряд считается нестационарным.

Основные статистические критерии и тесты на стационарность шести основных обобщенных показателей макрозообентоса экосистемы Байтуган–Сок представлены табл. 7.3. По существующей традиции исследования временных рядов, пространственные координаты выборочных точек были приняты эквивалентными натуральному ряду чисел $r = (1, 2, ..., m)$, т. е. физические расстояния в цепочке экологических последовательностей не учитывались.

Анализ динамики представленных показателей на стационарность показал, что большинство рядов соответствует модели «случайного блуждания», что свидетельствует об отсутствии тех или иных пространственных экосистемных закономерностей. Стационарными относительно постоянного дрейфа являются только ряды значений индекса разнообразия Шеннона и доли хирономид групп Diamesinae+Orthocladiinae. В то же время отчетливая нестационарность ряда индекса Шеннона относительно функции линейного тренда говорит о том, что эта зависимость не имеет монотонного характера, т. е. изменчивость этого показателя связана с проявлениями периодичности (см. график на рис. 7.5 слева). И только для доли обилия групп Diamesinae+Orthocladiinae можно усмотреть статистически значимую гладкую закономерность монотонного типа, что ранее было подтверждено анализом коэффициентов корреляции. Тест Дики–Фуллера для рядов, состоящих из первых разностей, свидетельствует об их отчетливой стационарности, что отвергает гипотезу о существовании на водотоке резких перепадов и «динамики пятен» (по крайней мере, в отношении показателей сообществ макрозообентоса, представленных в табл. 7.3).

Таблица 7.3

Оценка стационарности и автокоррелируемости рядов показателей экологических последовательностей сообществ макрозообентоса в речной системы Байтуган–Сок

Параметры и тесты	Всего видов	Среднее число видов в пробе	Индекс разнообразия Шеннона	Общее обилие бентоса (по тексту)	Доля (%) хищников-хватателей	Доля (%) Diamesinae+ Orthocladiinae
Минимальное значение и номер станции	49 (ст. С_01)	4,23 (ст. Б_08)	3,07 (ст. Б_08)	127 (ст. С_01)	7,2 (ст. Б_03)	4,0 (ст. С_12)
Максимальное значение и номер станции	102 (ст. Б_03)	8,91 (ст. С_04)	4,80 (ст. Б_03)	343 (ст. Б_06)	28,8 (ст. С_05)	35,7 (ст. С_02)
Тест Дики–Фуллера (модель с дрейфом)	-2,33 [0,178]	-2,67 [0,108]	-3,51 [***0,028***]	-2,57 [0,124]	-0,70 [0,809]	-3,36 [***0,035***]
Тест Дики–Фуллера (линейный тренд)	-2,08 [0,498]	-2,52 [0,314]	-3,29 [0,119]	-2,32 [0,393]	-1,87 [0,602]	-4,25 [***0,032***]
Тест Дики–Фуллера (первые разности)	-3,26 [***0,044***]	-4,39 [***0,008***]	-4,55 [***0,006***]	-3,47 [***0,031***]	-6,34 [***0,0005***]	-4,26 [***0,009***]
Q-статистика Льюнг–Бокса	1,55 [0,21]	0,47 [0,494]	0,09 [0,764]	0,71 [0,398]	0,04 [0,835]	7,30 [***0,007***]

Примечание: в квадратных скобках представлены *p*-значения, соответствующие тестируемым статистикам t или Q; жирным курсивом выделены случаи отклонения нулевой гипотезы.

Одним из общих методов географического анализа является оценка пространственной автокорреляции (spatial autocorrelation), которая используется, чтобы оценить, как взаимосвязаны экологические показатели и географические расстояния, и выделить характерные особенности этой пространственной зависимости. Объективные тенденции динамики экосистемы и её внутренний детерминизм определяют инерционность пространственного изменения ключевых параметров, значения которые, в известной степени, стабилизированы на определенной δ-окрестности относительно произвольной точки r. Однако вследствие пассивного переноса или активной миграции видов под действием локальных факторов среды происходит пространственное перераспределение компонентов сообществ. Иными словами, существует поле рассеяния пар значений $y(r)$, $y(r + k)$ последовательности, где k – некоторый пространственный интервал (или лаг; англ. *lag* – смещение), характеризующий взаимозависимость реализаций процесса в точках r и $(r + k)$. Теснота этой взаимосвязи оценивается коэффициентом автоковариации

$$\gamma(k) = \mathrm{E}[y(r) - m_Y)(y(r + k) - m_Y)]$$

или автокорреляции

$$R(k) = \gamma(k)/ D_Y(r).$$

Для выявления наличия автокорреляции ε_r случайных величин может быть использован *метод Льюнг–Бокса* (Ljung, Box, 1978), при которым вычисляется выборочная статистика $Q(k) = n(n+2)\sum_{k=1}^{m} \frac{R(k)^2}{n-k}$, где $R(k)$ – коэффициент автокорреляции k-го порядка. Если расчетное значение статистики Льюнг–Бокса больше 95%-го квантиля $\chi^2(m)$ распределения, то признается наличие автокорреляции (до m-го порядка) в последовательности ε_r. На наших данных (табл. 7.3) анализ статистики Льюнг–Бокса выявил отсутствие автокорреляции для большинства рядов обобщенных экологических показателей.

Другим распространенным в биогеографии способом выделения доли пространственной ковариации в общей вариации исходных данных является использование *I-статистики Морана* (Moran, 1950), которая, как и коэффициенты корреляции Пирсона или Кендалла, варьирует в интервале от -1 до +1. Чем больше значение I, тем более сходны ме-

жду собой экологические объекты, находящиеся в относительной пространственной близости, причем при увеличении географического расстояния между ними уровень сходства уменьшается. При отрицательной пространственной автокорреляции максимально несходные объекты появляются в относительной близости друг от друга, что для экосистем является далеко не характерным.

В общем виде формула для расчета коэффициента Морана имеет следующий вид:

$$I = \left(\frac{n}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}} \right) \cdot \left(\frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} \cdot (x_i - \bar{x}) \cdot (x_j - \bar{x})}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \right),$$

где n – число точек или пространственных единиц (в нашем случае, станций на водотоке); x – значение интересующей нас переменной (в нашем случае, показатели сообщества макрозообентоса); $\bar{x}$ – среднее значение признака для всей совокупности значений; w_{ij} – «вес», который отражает степень близости (или удаленности) между точками i и j в пространстве. Стандартная ошибка выборочных значений I может быть найдена по формулам аппроксимации или методами Монте-Карло.

Пространственные кореллограммы обобщенных показателей сообществ макрозообентоса, показанные на рис. 7.7, представляют определенный исследовательский интерес.

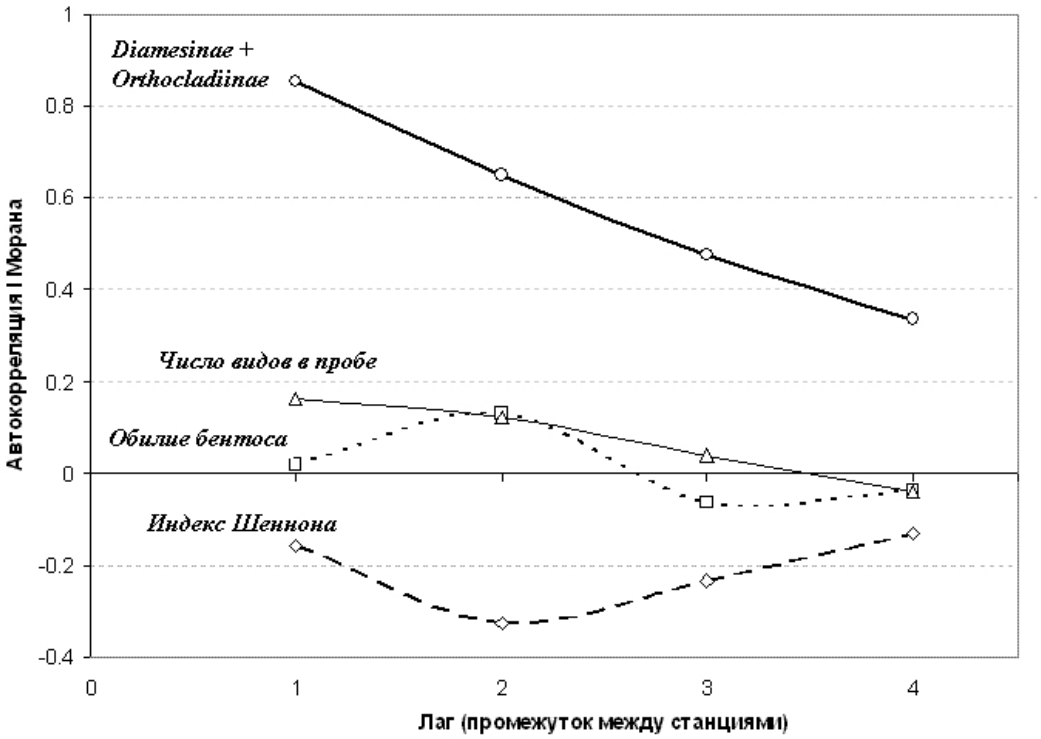


Рис. 7.7. **Автокорреляционная диаграмма индексов Морана для различных показателей сообществ макрозообентоса экосистемы Байтуган–Сок**

Например, максимум корреляционной связи для общего обилия бентоса приходится на лаг «через одну станцию», в то время как для индекса Шеннона H эта зависимость имеет противоположный характер. Это обычно свойственно рядам с отчетливо выраженной периодичностью. Однако мы здесь не будем выдвигать каких-либо конструктивных гипотез, поскольку значения индексов автокорреляции I для трех рядов статистически значимо не отличаются от нуля, т. е. пространственное распределение этих показателей по уча-

сткам является случайным. В отличие от них автокорреляционный спектр доли таксонов хирономид имеет высокую статистическую достоверность и характерен для монотонных зависимостей умеренной кривизны.

Современная биогеография и теория анализа временных рядов имеют немалые достижения при моделировании одномерных последовательностей. Однако разработка практически эффективных алгоритмов структурного анализа, оценки причинно-следственной коинтеграции и целенаправленной селекции моделей динамики зашумленных рядов большой размерности (в нашем случае – популяционной плотности одновременно 200-300 видов изучаемого сообщества) является для этих областей математики проблемой неопределенного будущего.

7.3. Оценка динамики видовой структуры с использованием индексов скорости и метрик сходства

Во многих экологических исследованиях установление количественной зависимости функции отклика от экологических факторов должно сочетаться с анализом скорости изменения функции отклика $y' = \mathrm{d}y / \mathrm{d}x$. Из соображений общности С.А. Дыренков и В.В. Акатов (1990) ввели формализм экологического пространства k-го порядка, координаты которого соответствуют $(k-1)$-ым производным от значений факторов и функций отклика. Практический интерес представляет анализ экологического пространства второго порядка, иначе говоря, «пространства скоростей». При этом, разумеется, имеется в виду не мгновенная скорость, а её средняя оценка $\Delta y / \Delta x$, где Δx – величина шага наблюдений в единицах длины, времени или произвольных факторов среды.

Оценка скорости изменения функции отклика связана с проблемой поиска способов определения величины Δy или, иными словами, *биологической длины градиента*, выраженной в единицах показателей, характеризующих состояние экосистемы. Она достаточно просто решается в отношении обобщенных показателей: биомассы, продуктивности, α-разнообразия и других, но гораздо сложнее – при оценке такого комплексного признака, как видовой состав биоценозов, где Δy соответствует β-разнообразию на градиенте. В литературе известно немало способов оценки скорости изменения ценотической структуры (Whittaker, 1967; Миркин и др., 1975; Diamond, 1975; Уиттекер, 1980 и др.), однако наибольший интерес для нас представляют те показатели, которые способны оценить скорость направленных изменений. При анализе пространственных градиентов под *ненаправленными изменениями* понимается весь комплекс нерегулярных волнообразных флуктуаций и случайных изменений («экологический шум»), а *направленные изменения* подразумевают необратимые сукцессионные смены видового состава, например, по продольному профилю течения водотока.

М. Вильсон и А. Шмида (Wilson, Shmida, 1984) использовали подход М. Коди (Cody, 1975), в основе которого лежит учет числа видов, появившихся $g(\mathbf{r})$ и исчезнувших $l(\mathbf{r})$ вдоль анализируемого интервала градиента $(\mathbf{r})$, и предложили индексы

- абсолютного $\beta c = | g(\mathbf{r}) + l(\mathbf{r}) | / 2$ и
- относительного $\beta\tau = \beta c / a_c$ оборота видов, где a_c – среднее число видов в пробах.

С.А. Дыренков и В.В. Акатов (1990) изящно показали, что индексы Вильсона–Шмиды можно вычислить гораздо проще: $\beta c = S - (a_1 + a_r)/2$; $\beta\tau = 2S/(a_1 + a_r) - 1$, где S – общее число видов; a_1 и a_r – видовое богатство на концах экологической последовательности. С учетом этого, для обеих ветвей последовательностей изучаемой экосистемы Байтуган–Сок указанные индексы могут быль легко рассчитаны, как:

- от участка С_01 до С_13: $\beta c = 276 - (49 + 91)/2 = 206$; $\beta\tau = 2 \cdot 276/(49 + 91) - 1 = 2{,}94$;
- от участка Б_01 до С_13: $\beta c = 334 - (97 + 91)/2 = 240$; $\beta\tau = 2 \cdot 334/(97 + 91) - 1 = 2{,}55$.

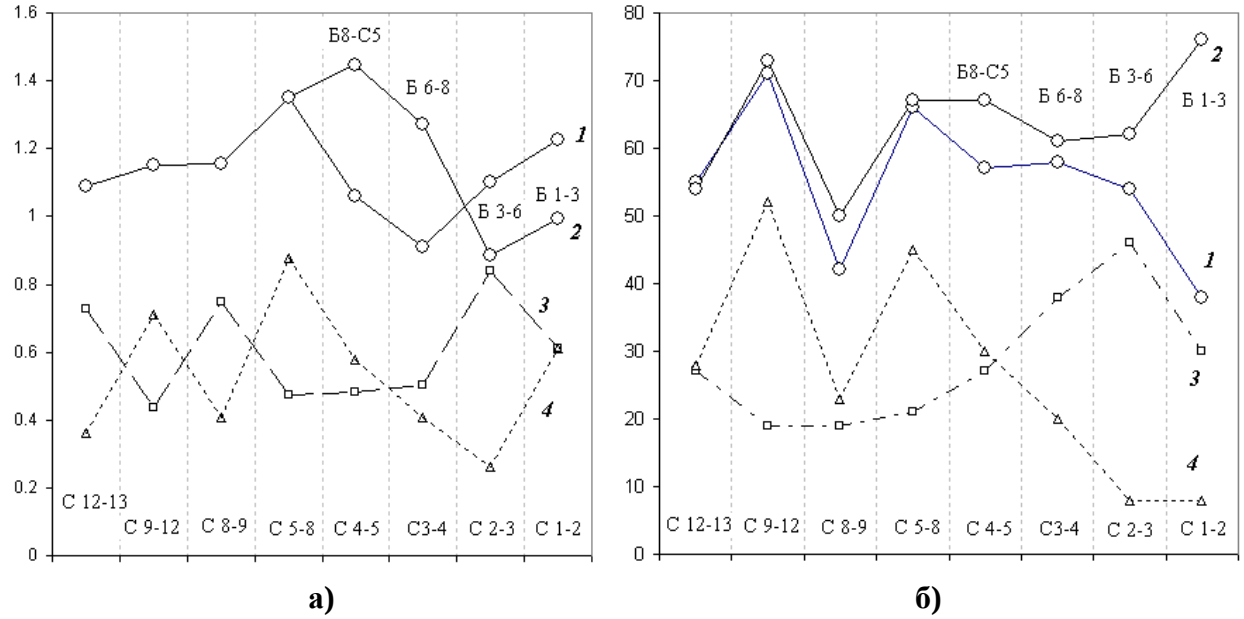
Можно также отметить, что приведенное выражение для $\beta\tau$ в точности совпадает с формулой *индекса β-разнообразия Уиттекера* (Whittaker, 1967); однако, несмотря на наличие коллектива столь авторитетных авторов, сами эти показатели не обладают ни законченным теоретическим смыслом, ни эффективностью практического использования.

Для решения задач оценки пространственно-временного континуума в большей степени важно не итоговое суммарное значение индекса, представленного «одним числом», а динамика частных значений скорости биоценотических изменений на градиенте. Если проанализировать $(n - 1)$ шагов, продвигаясь от первого узла экологической последовательности к последнему, то в качестве таких показателей могут быть использованы следующие величины (Дыренков, Акатов, 1990):

- число видов, появляющихся A_i и исчезающих B_i на каждом i-ом шаге последовательности;
- общая интенсивность оборота видов $I_{si} = (A_i + B_i)$ на i-ом шаге;
- число a_i впервые встретившихся видов;
- число видов b_i, встретившихся на предыдущих шагах и окончательно исчезнувших на i-ом шаге;
- интенсивность направленного изменения видового состава $I_{vi} = (a_i + b_i)$ на i-ом шаге.

При необходимости, можно также рассчитать для всей последовательности средние значения интенсивности оборота видов I_s и I_v, а также их дисперсии, что дает возможность статистически сравнивать между собой несколько анализируемых экологических объектов.

Рассмотрим пространственную динамику приведенных показателей скорости изменения таксономической структуры на примере донных сообществ экосистемы Байтуган–Сок: общую относительную интенсивность оборота видов I_{si} / S_i, где S_i – среднее число видов на двух смежных участках водотока (рис. 7.8а), и абсолютную интенсивность направленного изменения видового состава I_{vi} (рис. 7.8б). Расчеты проведены по двум веткам последовательности: на р. Сок от участка С_01 до С_13 и от истока р. Байтуган до устьевого участка речной системы Байтуган–Сок (Б_01 – С_13).



а) **б)**

Рис. 7.8. **Динамика относительной (а) и абсолютной (б) скоростей изменения видового состава донных сообществ экосистемы Байтуган–Сок:**

1 – указанные показатели для р. Сок; 2 – то же для ветки Байтуган–Сок;
3 – появление новых видов для р. Сок; 4 – исчезновение видов, найденных в р. Сок

Можно увидеть, что максимум общей смены видового состава наблюдается в месте впадения р. Байтуган, при переходе от участка Б_08 к С_05. Следует также отметить высокую скорость изменения таксономической структуры между участками С_5 и С_8 (рис. 7.8а), обусловленную массовым исчезновением видов реофильного комплекса, характерного для верхнего течения экосистемы.

Анализ пространственной динамики направленной скорости изменения видового состава (рис. 7.8б) свидетельствует об определенной специфичности комплекса таксонов макрозообентоса р. Байтуган: средняя степень направленности изменения видового состава для ветки Байтуган–Сок (I_v/I_s = 63,7) существенно превышает аналогичный показатель для основного водотока (I_v/I_s = 55,1). В частности, между участками Б_06 и Б_08 исчезло 47 видов, которые ниже по течению больше уже не появились. Можно также отметить сукцессионную смену видов между участками С_9 и С_12, связанную с трансформацией р. Сок водами Саратовского водохранилища.

Как уже упоминалось в разделе 1.2, Д. Перри и Д. Шеффер в своей статье с характерным названием "Продольное распределение речного бентоса: речной дисконтинуум?" (Perry, Schaeffer, 1987) феномен, представленный на рис. 7.8б, характеризовали, как «акцентированный градиент», когда постепенное изменение видовой структуры от истока к устью сопровождается локальными коренными сменами типов сообщества, обусловленными гидрологическими и гидрохимическими флуктуациями экстремального характера.

Если поставить перед собой более широкие цели, то можно задаться вопросом о статистической значимости пространственной структурированности сообществ по течению водотока. Проверяемой нулевой гипотезой будет утверждение о том, что изучаемое сообщество однородно и не имеет статистически значимых изолированных образований, т. е. значения разностей частот встречаемости комплекса видов в смежных участках реки зависят только от случайных флуктуаций. Альтернативная гипотеза, напротив, будет утверждать, что сообщество структурировано в пространстве, т. е. на границе выделяемых участков имеют место резкие смены видового состава.

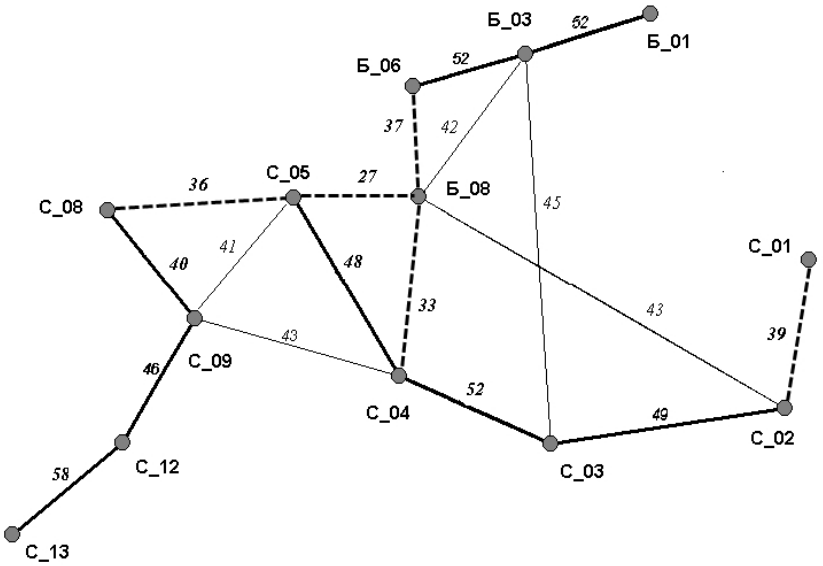
Перечисленные индексы скорости изменения таксономической структуры прямого ответа на эти вопросы не дают. Например, применяемый в таксономических исследованиях *индекс замены* (Allen, Starr, 1982) определяет, что если при переходе от одной точки последовательности к другой 50% видов оказались заменены другими видами, то было бы справедливо полагать, что мы имеем при этом новый тип сообщества. Почему именно 50% видов и следует ли этот порог применять для всех типов сообществ, – остается при этом неясным.

В биогеографии анализ пространственного распределения популяций был впервые формализован в модели изоляции расстоянием (isolation by distance IBD; Malecot, 1948), оценивающей линейную в логарифмах регрессию между соответствующими матрицами экологических и географических расстояний. В общем случае модель IBD можно проверить по уровню значимости коэффициента корреляции между соответствующими элементами двух квадратных матриц (видового сходства D и пространственной удаленности R между каждой парой анализируемых участков) с использованием *теста Мантеля* (раздел 6.7). Однако эта процедура не вполне корректна для сетей водотоков, поскольку практический интерес имеет соотношение видового состава на смежных участках.

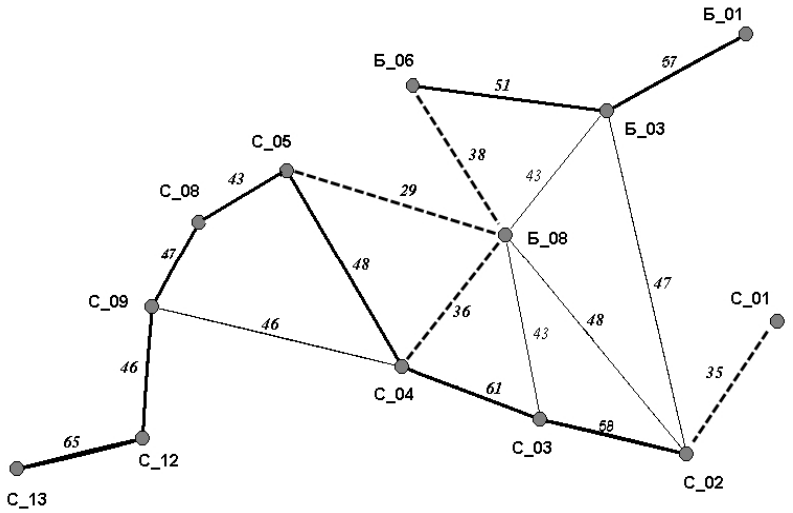
Рассмотрим динамику изменения таксономической структуры донных сообществ экосистемы Байтуган–Сок с использованием показателей видового сходства. Использованная нами вероятностная интерпретация формулы Съеренсена (раздел 3.2) в целом идентична индексу замены, но более обоснована, поскольку дополнительно учитывается частота встречаемости каждого анализируемого вида в пробах.

Одной из форм графического представления матрицы сходства D является дендрит, или дерево с максимальными весами, вершины которого соответствуют рассматриваемым объектам, а ребра – значениям метрики Съеренсена между этими объектами (Андреев,

1979). Графы на рис. 7.9, построенные на основе мер сходства полного видового состава макрозообентоса (а) и хирономидного комплекса (б), являются такими дендритами, которые дополнены некоторыми важными для содержательного анализа ребра с меньшим значением индекса Съеренсена. Учитывая общий характер сообществ макрозообентоса с их большим числом редких видов, что существенно снижает среднее сходство по Съеренсену, мы сочти целесообразным увеличить критический 50% порог Аллена смены видового состава для пары смежных станций до 60%. (т.е. значения метрики сходства менее 40%).



а) сходство по всему видовому составу макрозообентоса



б) сходство ценоза хирономид

Рис. 7.9. **Графы сходства по индексу Съеренсена видовой структуры сообществ макрозообентоса рек Байтуган (станции с префиксом «Б») и Сок («С»):**
жирные линии – ребра графа между каждой парой смежных станций наблюдения с уровнем сходства более 40%.; прерывистые линии – с меньшим уровнем сходства

Анализ графов позволяет проследить, в какой мере композиция видов в каждой точке последовательности связана с составом видов на станции наблюдения, находящейся выше по течению. Для большинства пар станций, смежных по течению водотока, композиции видов отличаются менее, чем на 60%, что является свидетельством континуальности

донного сообщества. Можно отметить лишь предполагаемую зону сукцессионных смен видового состава в районе ст. 8 р. Байтуган и относительно обособленное положение ст. 1 р. Сок. Изменение таксономической структуры хирономид на протяжении всего водотока имеет динамику, мало чем отличающуюся от всего сообщества в целом, за исключением участка р. Сок между станциями 5 и 8, где наблюдается формирование специфических видовых композиций за счет остальных групп макрозообентоса. Детальный характер видовой изменчивости сообществ был проанализирован нами ранее в разделах 6.3 и 6.4.

В разделе 3.2 был представлен метод непараметрического многомерного дисперсионного анализа (npMANOVA), предложенный М. Андерсоном (Anderson, 2001), и была выполнена проверка статистической гипотезы о значимости различий в видовом составе макробентоса при разделении всего водотока Байтуган–Сок на три участка.

Рассчитаем матрицу **D** расстояний Ренконена–Брея–Кёртиса между 120 гидробиологическими пробами на основе прологарифмированных значений численности 363 видов макрозообентоса. В качестве группировочных факторов представим иерархическую схему из двух уровней: каждый из трех крупных фрагментов водотока разделен на 4 участка в соответствии со схемой расположения станций наблюдения, на каждом из которых выполнено по 10 проб. Чтобы получить сбалансированный факторный план, пробы со станций С_01 и С_02 были объединены. С использованием алгоритма npMANOVA выполним проверку двух гипотез:

- гипотеза 1 – являются ли значимыми в целом отличия видового состава макрозообентоса между 12 станциями наблюдений?;
- гипотеза 2 – в какой мере вариация видовой структуры на трех укрупненных участках определяется изменчивостью проб на четырех станциях внутри каждого участка?

Для проверки второй гипотезы используем алгоритм вложенного (nested) дисперсионного анализа (McCune et al., 2002) с двумя факторами, из которых второй является иерархическим вложением первого. Общая линейная модель для любого отдельного измерения записывается, как $y = \mu + \tau + \beta(\tau) + \varepsilon$, где μ – математическое ожидание общего среднего; τ – влияние основного группировочного фактора 1; $\beta(\tau)$ – влияние вложенного фактора группировки 2 (т. е. изменчивость между повторностями в блоках данных с одинаковым уровнем воздействия фактора 1); ε – влияние случайных и не учитываемых в анализе факторов. Предполагается, что все факторы – τ, $\beta(\tau)$ и ε – независимы друг от друга, поэтому можно общую суммарную изменчивость разложить на три компоненты: $\mathrm{Var}\, y = \mathrm{Var}\, \tau + \mathrm{Var}\, \beta(\tau) + \mathrm{Var}\, \varepsilon$, статистическую значимость которых можно проверить по *F*-критерию. Результаты непараметрического дисперсионного анализа, выполненного с одним группировочным фактором (гипотеза 1) и двумя иерархически вложенными факторами (гипотеза 2), представлены в табл. 7.4.

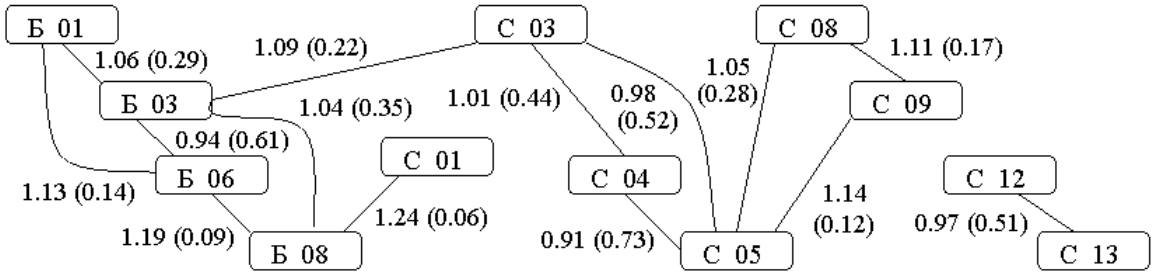
Таблица 7.4

Результаты одно- и двухфакторного иерархического анализа пространственной изменчивости сообществ макрозообентоса рек Байтуган–Сок

Факторы	Степени свободы *df*	Сумма квадратов *SS*	Средние квадраты *MS*	*F*-критерий	*p*-значение
Однофакторный непараметрический дисперсионный анализ – группировка по 12 станциям					
Группировка по станциям	11	9,22	0,838	2,40	0,0002
Остаток (ε)	108	37,70	0,349		
Итого	119	46,90			
Иерархический дисперсионный анализ – группировка на 3 участка, включающих по 4 станции					
Группировка на участки	2	4,30	2,150	3,93	0,0002
Внутри участков по станциям	9	4,92	0,547	1,56	0,0002
Остаток (ε)	108	37,70	0,349		
Итого	119	46,90			

Примечание: вероятности *p* отклонения нулевой гипотезы, когда она верна, оценивались в ходе рандомизационной процедуры, выполняющей 5000 итераций перемешивания данных.

Для схемы однофакторного анализа выполним множественные сравнения (межгрупповые контрасты) для каждой пары из 12 станций. Граф компонент связности (дендрит), вершинами которого являются станции наблюдения, а ребра соответствуют статистически значимым выводам об отсутствии различий в видовом составе, имеет следующий вид (рядом с линиями связи приведены значения *t*-критериев и в скобках соответствующие им *p*-вероятности):



Представленные результаты позволяют сделать следующий вывод: несмотря на то, что вариация видовой структуры между станциями одного участка носит в целом статистически значимый характер, она в значительной мере определяется изменчивостью таксономического состава для пар станций, наиболее удаленных друг от друга в пределах каждого участка. Два «разрыва» в непрерывном характере изменчивости в месте впадения р. Байтуган и для ст. 1 р. Сок, найденные по превышению 60% порога смены видов (рис. 7.9), были подтверждены и статистически с использованием множественных сравнений. Однако дисперсионный анализ показал дополнительно отсутствие статистически значимого сходства между парами станций {C_02 – C_03} и {C_09 – C_12}.

Описанные традиционные статистические методы анализа биогеографических матриц не всегда дают точную и адекватную оценку зависимостей между экологической структурой сообществ и пространственным распределением популяций. Более реалистичная картина может быть получена с использованием геометрических подходов, применяемых в биогеографии и фиксирующих изменения структуры в реальных географических координатах. Это в полной мере относится к идентификации «барьеров», т. е. выделению границ, относительно которых параметры изучаемых сообществ претерпевают резкие характерные изменения.

Одной из базовых структур вычислительной геометрии является *триангуляция Делоне*, которая отображает множество точек на плоскости сетью треугольников. Иными словами, решается задача деления географического пространства R на m подпространств C_m треугольной формы, удовлетворяющих ряду условий: каждое подпространство является изолированным, границы являются непересекающимися, а сумма длин всех ребер минимальна среди всех возможных триангуляций, построенных на тех же исходных точках (Скворцов, Костюк, 1998).

После того как получена сеть, объединяющая окрестности всех географических точек, каждое звено сети может быть связано со значением матрицы видового сходства **D**, полученной по данным экологического мониторинга. Чтобы найти границы областей, вдоль которых расстояния между смежными парами экологических объектов являются наибольшими, может быть использован *алгоритм максимума различий Монмоньера* (Monmonier, 1973). Перебор обычно начинается с краевого звена с максимальным значением расстояния и продолжается через прилегающие звенья, пока формируемая граница не достигнет пределов триангуляции или замкнется на себя, формируя контур вокруг совокупности точек. В целом алгоритм Монмоньера напоминает дихотомический процесс образования иерархических деревьев: как только формируемая граница подходит к краю треугольника, дальнейшее продвижение может проходить только по одному из двух ос-

тавшихся его ребер (правое или левое решение). Чтобы оценить статистическую значимость вычисленных барьеров, предлагается рандомизационный тест (Manni et al., 2004), который использует бутстреп-процедуру многократного преобразования исходной матрицы экологических расстояний.

Выполним поиск границ (барьеров) между отдельными донными сообществами экосистемы Байтуган–Сок с использованием географических координат отдельных станций. Матрицу экологических расстояний таксономической структуры рассчитаем по формуле Съеренсена, учитывающей встречаемость видов в пробах (раздел 3.2). Первая главная граница (рис. 7.10а) разделяет весь водоток на две части: выше станции С_5 и участки, расположенные ниже по течению р. Сок (расстояние Съеренсена d изменяется от 0,626 до 0,722). Вторая по значимости граница (рис. 7.10б) выделяет как самостоятельную область участки р. Байтуган (d = 0,567 ÷ 0,703). Третья граница выделяет в качестве характерного биотопа ст. 1 на р. Сок. И, наконец, четвертая граница проходит по р. Сок у с. Исаклы между станциями 5 и 8.

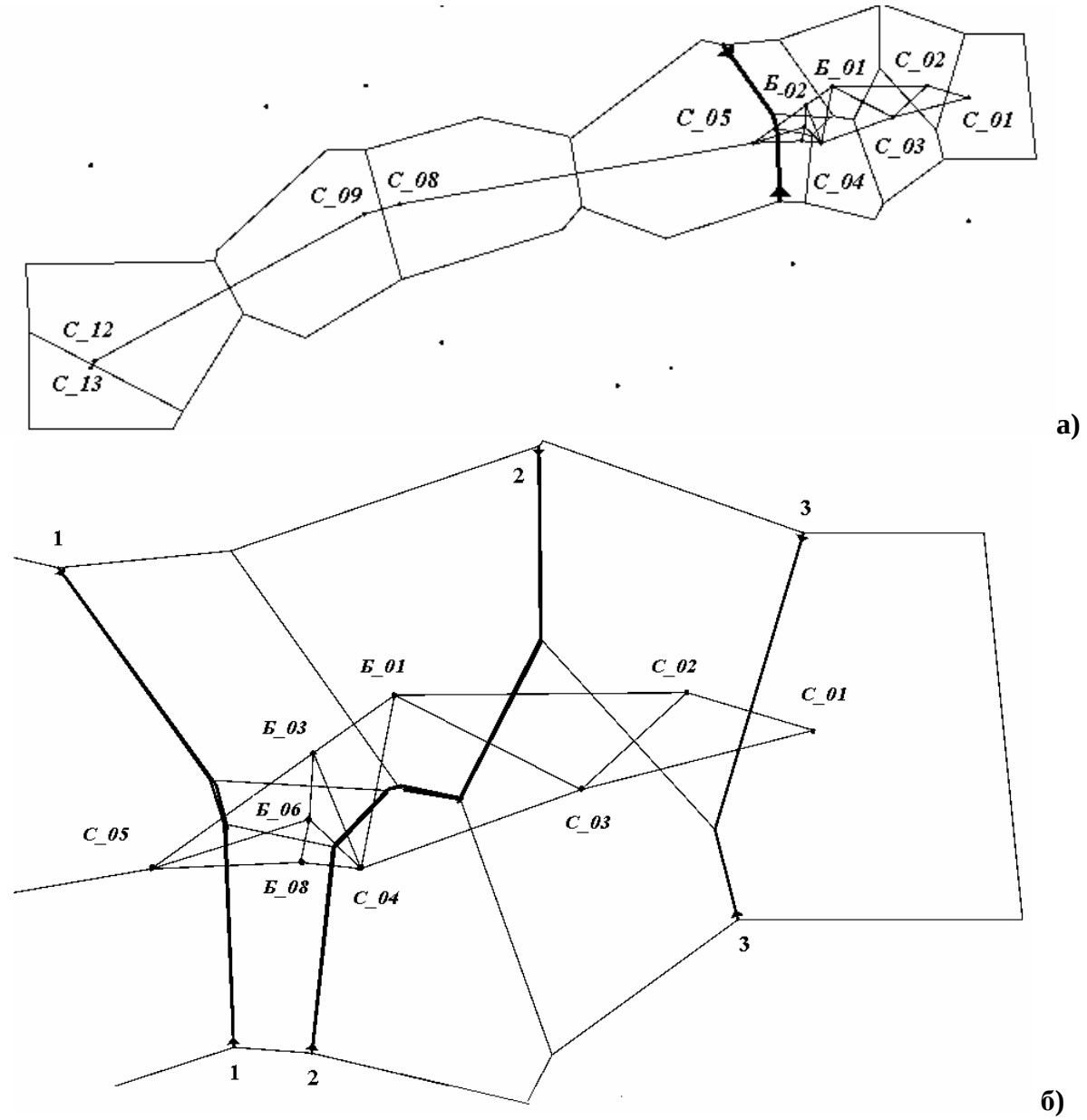


Рис. 7.10. **Выделение границ («барьеров») между отдельными сообществами макрозообентоса экосистемы Байтуган–Сок с использованием географических координат**

Проведенные статистические тесты на стационарность экологических показателей и продольную изменчивость видовой структуры показали, что речной континуум донных сообществ изучаемого водотока в целом следует оценивать как «*акцентированный*». Динамика изменчивости большинства обобщенных структурных или функционально-продукционных параметров сообществ по течению реки, за исключением некоторых групп хирономид, статистически значимо не отличается от чисто случайного процесса типа «белого шума». Носит ли эта изменчивость континуальный характер или является результатом стохастических возмущений естественного и антропогенного характера – статистика вряд ли может дать обоснованный ответ на этот вопрос.

В то же время, ориентируясь на характер полученных графических материалов (рис. 7.4-7.9) и результаты специальных гидробиологических исследований (Зинченко, Головатюк, 2000), можно выделить две характерных «зоны» изменения общей численности и видового состава донных сообществ:

- в зоне слияния двух рек (ст. 8 р. Байтуган) и в верховьях р. Сок (ст. 1-2), объясняемые наличием большого числа родников и каменистым субстратом участков;
- на р. Сок, в районе станций 8 и 12, связанные, возможно, с аккумуляцией загрязняющих веществ антропогенного происхождения.

Отмеченные разрывы континуума носят вполне объяснимый экологический смысл. Видовой состав макрозообентоса станций 1 и 2 р. Сок и на всем течении р. Байтуган является типичным для «ритрали» водотоков (Богатов, 1995), где основную долю фаунистического разнообразия составляют относительно немногочисленные таксоны литореофильных личинок амфибиотических насекомых (Diptera, Ephemeroptera, Trichoptera). На ст. 5 происходит сукцессионная смена экоформ и видовая структура донного сообщества существенно изменяется за счет развития привносимых эврибионтных таксонов.

Структурные особенности макрозообентоса приустьевого участка, начиная со ст. 12, во многом связаны с трансформацией р. Сок водами Саратовского водохранилища. Вторая сукцессионная смена донных сообществ обусловлена развитием лимнофильных моллюсков, пелофильных хирономид и ракообразных, характерных для донных сообществ водохранилища и в большинстве своем относящихся к инвазионным видам, распространенным в волжских водохранилищах.

ЗАКЛЮЧЕНИЕ

Много лет размышлял я над жизнью земной.
Непонятного нет для меня под луной.
Мне известно, что мне ничего не известно!
Вот последняя правда, открытая мной.

Омар Хайям (1048-1131 гг.)

У читателя может создаться определенное ощущение, что настоящая книга перегружена большим многообразием гипотез, подходов и методов статистического анализа при решении проблем, связанных с изучением экологии сообществ, использование которых не всегда имеет отчетливый практический смысл. Действительно, после расчета, например, предполагаемого количества видов или параметров стохастического распределения их численности в экосистеме возникает сакраментальный вопрос: ну и какой практический вывод из этого можно извлечь?

Авторы уверены, что такое впечатление обманчиво. В действительности при изучении природных водных объектов экологи стремятся создать «обобщенный портрет» исследуемой экосистемы с достаточной степенью детализации и в масштабе, адекватном поставленной задаче (т. е. «увидеть мир, стоя на глобусе»). Но только всесторонний анализ и тщательная проверка всех предположений, не всегда явно вытекающих из многообразия полученных результатов исследований, может дать достаточно полную картину.

Сложный многопараметрический и экосистемный характер донных сообществ требует оценки различных аспектов их функционирования и структурной изменчивости. Анализу подлежат оценка полного видового богатства и совокупного биоразнообразия, выделение основных функциональных компонент и оценка характера распределения популяционной плотности в пространстве и времени. Большое значение имеет также анализ таксономической структуры сообществ, оценка интенсивности межвидовых взаимодействий и вклада, вносимого в устойчивость и жизнеспособность сообщества различными группами гидробионтов – от доминантов до редких видов. Если продолжить аналогию, например с живописью, то весь этот комплекс многокомпонентных исследований эколога напоминает отдельные «мазки», составляющие общую гармонично скомпонованную картину. Проработка различных аспектов анализа требует использования большого арсенала современных проблемно-ориентированных математических методов, оснащенных процедурами проверки статистических гипотез.

Структурная организация водных сообществ носит иерархический полимасштабный характер. На каждом масштабном уровне и в зависимости от прикладных целей задачи изучения механизмов функционирования и пространственного распределения гидробионтов формулируются и решаются по-разному. Например, на мелкомасштабном уровне от 1 м до 100 м мозаичность («пятнистость») размещения популяций бентосных организмов во многом определяется точечной изменчивостью структуры биотопов: гранулометрическим составом и композиционным распределением песчаных, илистых, гравийных и прочих отложений. На этом уровне важен также учет стохастического (или направленного по течению) перемещения особей, сравнительный анализ способности различных видов осваивать новые субстраты и т. д. Для оценки мелкомасштабной изменчивости используют большое количество специальных моделей и критериев (модели распределения ансамблей случайных точек, модели дрифта и др.).

На среднемасштабном уровне (100 м – 100 км) анализируются более общие закономерности макроизменчивости сообществ гидробионтов. В водотоках, например, основным определяющим фактором становится биоценотический градиент от истока к устью, в на-

правлении которого наблюдается некоторая экологическая последовательность отличающихся по видовому составу, но функционально достаточно близких экоморф. Такие продольные сукцессионные смены определяются, с одной стороны, континуумом естественных ландшафтно-гидрологических характеристик водотока, а с другой – локальным воздействием комплекса различных внешних факторов высокой интенсивности, имеющих постоянный, периодический или даже «залповый» характер. Различные статистические подходы к количественной оценке этих явлений представлены в главах 6 и 7. На крупномасштабном (более 100 км) и континентальном уровнях состав показателей начинает носить укрупненный агрегированный характер, а исследования проводятся в русле приоритетных направлений развития ландшафтной экологии, природно-климатического зонирования и т. д.

Методология мультимасштабного сегментирования экосистем вплотную подошла к разработке методов решения проблемы, обозначенной еще в начале 1980-х годов Р. О'Нейлом, П. Делькуртом, Ю.Г. Пузаченко и другими: выбор оптимальных масштабов, адекватных для отображения экологических процессов в пространстве и во времени. Разработаны методы корректной экстраполяции точечных данных для широкого диапазона масштабов (Borcard, Legendre, 2002; Peterson, Parker, 1998), предложены схемы мультимасштабных профилей, внутри которых пространственные структуры самоподобны, т. е. не меняются или меняются монотонно при изменении масштаба (Zurlini et al., 2007; Хорошев, Алещенко, 2008), созданы экспертные системы и базы данных, предоставляющие научно обоснованную информацию для принятия решений в области природопользования (Розенберг, 2009). Все это можно расценить как целенаправленный переход от точечно-ориентированного анализа к объектно-ориентированному представлению территориальных сегментов, содержащих целостные взаимоувязанные сочетания биотических и ландшафтных компонентов.

Материал, изложенный в основных главах книги, – лишь вершина айсберга огромного числа англоязычных публикаций по использованию имитационных моделей и статистических методов при изучении экологических сообществ. Даже скромные обзоры литературы по этой тематике включают десятки фундаментальных монографий (Pielou, 1975; Sokal, Rohlf, 1981; Gauch, 1982; Seber, 1982; Jongman et al., 1987; Krebs, 1989; Gilpin, Hanksi, 1991; Renshaw, 1991; Thompson, 1992; Brown, 1995; Gotelli, Graves, 1996; Legendre, Legendre, 1998; Tokeshi, 1999; Gaston, Blackburn, 2000; Caswell, 2001; Hubbell, 2001; Quinn, Keough, 2002; McCune, Grace, 2002; Lande et al., 2003; Magurran, 2004; Manly, 2006 и др.) и многие сотни журнальных статей. К сожалению, российская библиография насчитывает весьма скудное количество оригинальных работ в этой области.

Нам представляется, что целому ряду теоретических положений макроэкологии целесообразно найти самое широкое применение в практике биосферных исследований. Например, «калибровка» показателей видового богатства и разнообразия по моделям разрежения в зависимости от величины выборочного усилия должна явиться обычным стандартом обработки данных при сравнении различных природных объектов. Экстраполяция кумулятивной кривой и расчет потенциально полной численности видов в сообществе также является объективной оценкой репрезентативности данных, получаемых в ходе полевых исследований (глава 2).

Оценка видового разнообразия продолжает оставаться наименее проработанной задачей теоретической экологии. Например, в планах РАН до 2025 г. предусмотрено всего две программы: «Выявление характера связей между биологическим разнообразием, устойчивостью сообществ, интенсивностью продукционных и деструкционных процессов в экосистемах для основных природных зон Северной Евразии» и «Анализ многолетних трендов и выявление факторов изменения биоразнообразия в районах сильных локальных антропогенных нарушений природных комплексов».

Построение профилей видового разнообразия, семейств обобщенных энтропий и мультифрактальных спектров могут в определенной степени прояснить возникающие при этом проблемы. Однако интерпретация этих зависимостей часто основана на субъективных представлениях из-за отсутствия методов «параметризации» и статистических процедур, позволяющих отнести эмпирические данные к определенному типу сообществ по их разнообразию. Важным элементом дифференциации видового разнообразия является аддитивная схема его иерархического разложения на α-, β- и γ-компоненты, позволившая перенести на формальную, легко интерпретируемую основу фундаментальные конструкции Р. Мак-Артура – Р. Уиттекера. На фоне всеобщей модной критики информационного индекса Шеннона (на наш взгляд, не всегда обоснованной) привлекает внимание его апология, сделанная Л. Джостом, как наиболее сбалансированной меры видового разнообразия (глава 3).

Для выявления структурных закономерностей формирования композиций таксонов и оценки статистической значимости взаимной согласованности встречаемости видов большую популярность приобретает nestedness-анализ степени «вложенности» эмпирических таблиц «виды-местообитания» (глава 4). Вместе с различными схемами ординации (глава 6) и непараметрическим иерархическим анализом матриц сходства, разработанным М. Андерсоном (главы 3 и 7), эти методы могут стать эффективным инструментом объективного выделения экологически сходных групп местообитаний и экоморф видов (синтаксонов).

Объектом нашего пристального внимания явились методы и процедуры, позволяющие оценить вариацию видовой структуры гидробионтов вдоль продольного течения водотока (т. е. по трансекте местообитаний), чтобы проверить статистическую гипотезу о закономерном континууме сообщества против альтернативы о стохастической «мозаике пятен». Достаточно убедительную информацию о таксономической изменчивости нам удалось получить с использованием методов ординации, широко используемых в фитоценологии (глава 6).

Этот анализ был продолжен в глава 7 с использованием математического аппарата анализа многомерных временных рядов, ГИС-технологий и других методов исследования пространственной структуры сообществ. В то же время многие интересные подходы и модели (например, использование фазовых переходов и неоднородных марковских процессов для моделирования сукцессии, представленные в работах А.С. Исаева и Д.О. Логофета) остались за рамками настоящей монографии, как требующие более обстоятельного изложения.

Далеко не все схемы и процедуры обработки данных могут удовлетворить непредвзятого эколога-аналитика. Например, значительная доля публикаций второй половины XX века была посвящена моделям SAR, SAD, BSD (раздел 1.4) и поиску теоретических кривых «правильной формы» (т. е. формы, предположительно наблюдаемой в природе). Б. МакГилл в своей статье "Строгие и слабые стороны теоретических моделей в макроэкологии" (McGill, 2003a, с. 679) выразился предельно категорично: «Сосредотачиваться на формах кривой бесперспективно и контрпроизводительно для макроэкологии».

Д. Гудолл (Goodall, 1952, p. 217) справедливо писал, что решение выбрать ту или иную модель «не может приниматься из априорных соображений, а должно опираться на данные наблюдений». Однако МакГилл (McGill, 2003a, b) показал, что удачное подтверждение модели частными эмпирическими данными – слишком слабое свидетельство правильности заключения о механизмах изучаемого явления. Этому есть, по крайней мере, четыре причины:

1. *Множественность объяснения*. Например, для большинства наборов данных можно считать «правильными» не менее двух дюжин моделей (и механизмов их расчета) распределения обилия видов по градиенту. И лишь очень немногие кривые будут решительно отклонены, как не соответствующие результатам наблюдений.

2. *Свобода выбора параметров*. Манипулируя числом переменных модели можно достичь любого, сколь угодно точного приближения модельной кривой к экспериментальным данным. МакГилл (McGill, 2003a) приводит цитату со ссылкой на Л. Эйлера: «Дайте мне три свободных параметра, и я нарисую Вам слона; дайте мне четыре, и он начнет качать хоботом».
3. *Неопределенность критерия для выбора «лучшей» модели*. МакГилл использует 9 различных методов оценки согласия модели с данными: *критерий Колмогорова–Смирнова*, *вероятность максимального правдоподобия*, *три меры*, *основанные на нелинейных формах корреляции* r^2 *фактических и расчетных значений*, *четыре схемы расчета* χ^2*-критерия*, основанные на различных законах распределения испытаний. Все методы вполне работоспособны, хотя некоторые оценивают согласие данных по всему имеющемуся диапазону, а другие подчеркивают уровень адекватности в зоне специального интереса (например, где сосредоточена наибольшая часть плотности вероятности). В зависимости от применяемого метода, представления о «наилучшей» модели могут существенно меняться (см. также раздел 5.3).
4. *Особенности центральной предельной теоремы в экологии*. Основные процессы в экологии носят характер взаимодействия мультипликативных случайных величин. Примерами этому являются «преодоление последовательности барьеров» (sequence of hurdles; Shockley, 1957) при выживании $W = p_1 \cdot p_2 \cdot p_3 \cdot p_4 \cdot \ldots$ и составляющие роста (compound growth; Caswell, 2001) популяций $N_0 = N_i \lambda_1 \cdot \lambda_2 \cdot \lambda_3 \cdot \lambda_4 \ldots$ Легко можно показать (MacArthur, 1960; May, 1975), что любой процесс, объединяющий множество мультипликативно взаимодействующих переменных, так или иначе, сходится к логарифмическому нормальному распределению. С ним близко связана показательная функция или распределение Парето–Ципфа–Мандельброта (см. раздел 5.2), что дало возможность МакГиллу говорить об основополагающей единой концепции моделирования POLO (power-lognormal), как новой форме центральной предельной теоремы. Он считает, что нет нужды в многообразии подходов, поскольку любая экологическая модель будет сходиться в некотором пределе к логнормальному распределению, либо к показательной функции.

Если скомбинировать пункты 1 и 2, то можно предположить, что «почти любая теория будет соответствовать почти любым данным». МакГилл сформулировал «уровни строгости» проверки моделей, сводящиеся в конечном итоге к статистическому анализу различий с нуль-моделью, тщательной отбраковке лишних параметров, использованию методов рандомизации и Монте-Карло.

Время покажет, являются ли действительно *самоподобие структур*, *функция Ципфа–Мандельброта* или *логнормальное распределение* фундаментальной основой законов и моделей экологии, либо они останутся в памяти потомков как быстро преходящая мода. Мы же закончим наш обзор аналитических исследований другой цитатой из Б. МакГилла (McGill, 2003a, p. 684): «Макроэкология – сегодня жизненно важная область экологии. Её развитие обещает как большие надежды на распутывание клубка скороспелых теоретических конструкций, так и может стать полезным инструментом в практической деятельности. Будет позором, если эту полезность сочтут ненадежной, поскольку мы не предоставили строгие процедуры проверки макроэкологических моделей и предположений».

СПИСОК ЛИТЕРАТУРНЫХ ИСТОЧНИКОВ

Абросов Н.С., Боголюбов А.Г. Экологические и генетические закономерности сосуществования и коэволюции видов. Новосибирск: Наука, 1988. 327 с.

Азовский А.И. Соотношение пространственно-временных диапазонов в экологических иерархиях различной природы // Журн. общ. биол. 2001. Т. 62, № 6. С. 451-459.

Азовский А.И., Чертопруд М.В. Анализ пространственной организации сообществ и фрактальная структура литорального бентоса // Докл. АН. 1997. Т. 356, № 5. С. 713-715.

Айвазян С.А., Буштабер В.М., Енюков И.С., Мешалкин Л.Д. Прикладная статистика. Классификация и снижение размерностей. М.: Финансы и статистика, 1989. 607 с.

Акатов В.В. Анализ степени флористической неполночленности фитоценозов как способ тестирования островного эффекта // Бот. журн. 1995. Т. 80, № 4. С. 49-64.

Акатов В.В. Островной эффект как фактор формирования высокогорных фитоценозов Западного Кавказа. Майкоп: Изд-во МГТИ, 1999. 114 с.

Акатов В.В., Чефранов С.Г. Локальное видовое богатство древесного яруса лесов о. Мадагаскар и Западного Кавказа: опыт тестирования исторической гипотезы путем анализа структуры распределения обилия видов // Бюл. МОИП. Отд. биол. 2007. Т. 112, вып. 1. С. 65-71.

Акатов В.В., Чефранов С.Г., Акатова Т.В. Гипотеза видового фонда: необходимость смены акцента // Журн. общ. биол. 2002. Т. 63, № 2. С. 112-121.

Акатов В.В., Чефранов С.Г., Акатова Т.В. Об эволюционной полночленности видовых фондов современных растительных сообществ высокогорной зоны Западного Кавказа // Журн. общ. биол. 2003. Т. 64, № 4. С. 308-317.

Акимова Т.А., Хаскин В.В., Сидоренко С.Н., Зыков В.Н. Макроэкология и основы экоразвития: Учеб. пособие. М.: Изд-во РУДН, 2005. 368 с.

Александрова В.Д. О единстве непрерывности и дискретности в растительном покрове // Философские проблемы современной биологии. Л.: Наука, 1966. С. 191-204.

Александрова В.Д. Классификация растительности: Обзор принципов классификации и классификационных систем в разных геоботанических школах. Л.: Наука, 1969. 275 с.

Алёхин В.В. Основные понятие и основные единицы в фитоценологии // Сов. ботаника. 1935. № 5. С. 21-34.

Алимов А.Ф. Элементы теории функционирования водных экосистем. СПб.: Наука, 2001. 147 с.

Андерсон Т. Статистический анализ временных рядов. М.: Мир, 1976. 523 с.

Андреев В.Л. Статистические методы классификационных построений в биогеографии и систематике // Иерархические классификационные построения в географической экологии и систематике. Владивосток: ДВНЦ АН СССР, 1979. С. 60-96.

Апаля-Шидлине Дз. Об индикаторных биоэкогруппах // Теоретические вопросы фитоиндикации. Л.: Наука, 1971. С. 37-43.

Арапов М.В., Ефимова Е.Н., Шрейдер Ю.А. О смысле ранговых распределений // Науч.-техн. информ. 1975. Сер. 2. № 1. С. 9-20.

Артюхов В.В. Общая теория систем: Самоорганизация, устойчивость, разнообразие, кризисы. М.: URSS, 2009. 224 с.

Аруцев А.А., Ермолаев Б.В., Кутателадзе И.О., Слуцкий М.С. Концепции современного естествознания. Учебное пособие. М.: МГОУ, 2006. [http://nrc.edu.ru/est/pos].

Баканов А.И. Теоретические основы экологического районирования водохранилищ // Водные ресурсы. 1997. Т. 24, № 3. С. 336-343.

Баркман Я., Моравец Я., Раушерт С. Кодекс фитосоциологической номенклатуры // Бюл. МОИП. Отд. биол. 1988. Т. 93, вып. 6. С. 112-130.

Барсегян А.А., Куприянов М.С., Степаненко В.В. Методы и модели анализа данных: OLAP и Data Mining. СПб.: БХВ-Петербург, 2004. 336 с.

Бигон М., Харпер Дж., Таунсенд К. Экология. Особи, популяции и сообщества. В 2-х т. М.: Мир, 1989. Т. 2. 477 с.

Биоиндикация экологического состояния равнинных рек / Отв. ред. О.В. Бухарин, Г.С. Розенберг. М.: Наука, 2007. 404 с.

Богатов В.В. Комбинированная концепция функционирования речных экосистем // Вестн. ДВО РАН. 1995. № 3. С. 51-61.

Борзенков В.Г. Биология и физика. М.: Знание, 1982. 64 с.

Брусиловский П.М. Коллективы предикторов в экологическом прогнозировании. Саратов: Изд-во Сарат. ун-та, 1987. 104 с.

Булгаков Н.Г., Абакумов В.А., Максимов В.Н. и др. Методические вопросы применения ранговых распределений численности фитопланктона к анализу массовых данных экологического мониторинга пресных вод. Приближенные расчеты // Изв. РАН. Сер. биол. 2005. № 6. С. 718-725.

Вапник В.Н. Восстановление зависимостей по эмпирическим данным. М.: Наука, 1979. 415 с.

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

Василевич В.И. Количественные методы изучения структуры растительности // Итоги науки и техники. Ботаника. Т. 1. М.: Изд-во ВИНИТИ, 1972. С. 7-83.

Василевич В.И. Почему существуют многовидовые растительные сообщества // Бот. журн. 1979. Т. 64, № 3. С. 341-351.

Василевич В.И. Очерки теоретической фитоценологии. Л.: Наука, 1983. 248 с.

Галанин А.В. Флора и ландшафтно-экологическая структура растительного покрова. Владивосток: Дальнаука, 2005. 272 с.

Гелашвили Д.Б., Дмитриев А.И., Иудин Д.И. и др. Мультифрактальный анализ видовой структуры сообществ мелких млекопитающих Волго-Уральского палеоценоза // Докл. АН. 2008в. Т. 421, № 4. С. 562-566.

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

Гелашвили Д.Б., Иудин Д.И., Розенберг Г.С. и др. Основы мультифрактального анализа видовой структуры сообщества // Успехи соврем. биол. 2008а. Т. 128, № 1. С. 21-34.

Гелашвили Д.Б., Иудин Д.И., Розенберг Г.С., Якимов В.Н. Элементы фрактальной теории видовой структуры гидробиоценозов // Изв. Самар. НЦ РАН. 2006. Т. 8, № 1. С. 70-79.

Гелашвили Д.Б., Иудин Д.И., Розенберг Г.С., Якимов В.Н. Степенной характер накопления видового богатства как проявление фрактальной структуры биоценоза // Журн. общ. биол. 2007а. Т. 68, № 2. С. 115-124.

Гелашвили Д.Б., Иудин Д.И., Солнцев Л.А. и др. Мультифрактальный анализ видовой структуры сообществ гельминтов мелких млекопитающих Самарской Луки // Докл. АН. 2009б. Т. 427, № 5. С. 703-706.

Гелашвили Д.Б., Пухнаревич Д.А., Иудин Д.И. Структурная организация и фрактальная природа макрозообентоса малых городских водоемов // Поволж. экол. журн. 2007б. № 1. С. 3-15.

Гелашвили Д.Б., Розенберг Г.С., Саксонов С.В. и др. Видовая структура сообществ крупных млекопитающих Самарской Луки в связи с проблемой изменения климата // Изв. Самар. НЦ РАН. 2009а. Т. 11, № 1(2). С. 28-41.

Гелашвили Д.Б., Якимов В.Н., Иудин Д.И. и др. Мультифрактальный анализ видовой структуры сообществ мелких млекопитающих Нижегородского Поволжья // Экология. 2008б. № 6. С. 456-461.

Гелашвили Д.Б., Якимов В.Н., Иудин Д.И. и др. Фрактальные аспекты таксономического разнообразия // Журн. общ. биол. 2010. Т. 71, № 2. С. 115-130.

География и мониторинг биоразнообразия / Н.В. Лебедева, Д.А. Криволуцкий, Ю.Г. Пузаченко и др. М.: Изд-во науч. и учеб.-метод. центра, 2002. 432 с.

Гиляров А.М. Соотношение органицизма и редукционизма как основных методологических подходов в экологии // Журн. общ. биол. 1988. Т. 49, № 2. С. 202-217.

Гиляров А.М. Перестройка в экологии: от описания видимого к пониманию скрытого // Вестн. РАН. 2005. Т. 75, № 3. С. 214-223.

Гиляров А.М. От ниш к нейтральности в биологическом сообществе // Природа. 2007. № 11. С. 29-37. [http://rogov.zwz.ru/Macroevolution/ghilarov_neutral.pdf].

Гиляров А.М. В поисках универсального закона устройства биологических сообществ, или Почему экологи потерпели неудачу? // Элементы большой науки. 2008. [http://elementy.ru/news/430675].

Грайм Д.П. Классификация растительности по соотношению стратегий / Пер. с англ. Г.С. Розенберга // Теоретические проблемы экологии и эволюции (Четвертые Любищевские чтения). Тольятти: ИЭВБ РАН, 2005. С. 48-58.

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

Гублер Е.В., Генкин А.А. Применение непараметрических критериев статистики в медико-биологических исследованиях. Л.: Медицина, 1973. 132 с.

Джиллер П. Структура сообществ и экологическая ниша. М.: Мир, 1988. 184 с.

Динамические модели в экологии. Реестр моделей (Моск. ун-т им. М.В. Ломоносова, каф. биофизики; рук. проф. Г.Ю. Ризниченко). [http://www.dmb.biophys.msu.ru/models].

Дыренков С.А., Акатов В.В. Экологическое пространство второго порядка: скорости изменений функций отклика на градиенте экологических факторов // Экология. 1990. Т. 21, № 3. С. 13-21.

Ермилов С.Г. Особенности населения орибатидных клещей крупного промышленного центра (город Нижний Новгород): Автореф. дис. … канд. биол. наук. Н. Новгород, 2004. 21 с.

Залетаев В.С. Структурная организация экотонов в контексте управления // Экотоны в биосфере / Отв. ред. В.С. Залетаев. М.: РАСХН, 1997. С. 11-29.

Зернов С.А. Общая гидробиология. М.; Л.: Изд-во АН СССР, 1949. 588 с.

Зинченко Т.Д. Хирономиды поверхностных вод бассейна Средней и Нижней Волги (Самарская область). Эколого-фаунистический обзор. Самара: ИЭВБ РАН, 2002. 174 с.

Зинченко Т.Д., Головатюк Л.В. Изменение состояния бентоса малых рек бассейна Средней Волги. 1. Состав и распределение бентофауны притоков реки Сок (Камышла, Сосновка, Черновка, Хорошенькая) // Изв. Самар. НЦ РАН. 2000. Т. 2, № 2. С. 257-267.

Зинченко Т.Д., Головатюк Л.В., Абросимова Э.В. Экологическая характеристика *Cricotopus salinophilus* (Diptera, Chironomidae) из соленых рек бассейна оз. Эльтон // Изв. Самар. НЦ РАН. 2010. Т. 12, № 1. С. 196-200.

Зинченко Т.Д., Шитиков В.К. Гидробиологический мониторинг как основа типологии малых рек Самарской области // Изв. Самар. НЦ РАН. 1999. № 1. С. 118-127.

Ипатов В.С., Кирикова Л.А. Фитоценология: Учебник. СПб.: Изд-во СПб. ун-та, 1997. 316 с.

Иудин Д.И. Методология принципа самоподобия в исследовании видовой структуры биотических сообществ: Дис. … докт. биол. наук. Н. Новгород, 2006. 273 с.

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

Кафанов А.И. Континуальность и дискретность живого покрова: проблема масштаба // Журн. общ. биол. 2006. Т. 67, № 4. С. 311-313.

Керженцев А.С. Функциональная экология. М.: Наука, 2006. 259 с.

Козлов В.И., Абрамович Л.С. Краткий словарь рыбовода. М.: Россельхозиздат, 1982. 160 с.

Королюк А.Ю., Намзалов Б.Б. Эколого-ценотические элементы степной флоры гор Южной Сибири // Сиб. экол. журн. 1999. № 5. С. 495-500.

Красилов В.А. Этапность эволюции и её причины // Журн. общ. биол. 1973. Т. 34, № 2. С. 227-240.

Краснощеков Г.П., Розенберг Г.С. Экология в «законе» (теоретические конструкции современной экологии в цитатах и афоризмах). Тольятти: ИЭВБ РАН, 2002. 248 с.

Краснощеков Г.П., Розенберг Г.С. [Рецензия] // Самарская Лука: Бюл. 2008. Т. 17, № 1(23). С. 194-203. Рец. на кн.: Керженцев А.С. Функциональная экология / Отв. ред. Э.Г. Коломыц. М.: Наука, 2006. 259 с.

Крылов А.В. Зоопланктон равнинных малых рек. М.: Наука, 2005. 263 с.

Кудрин А.И. Очерки полевого учета // Ценологические исследования. М.: Центр системн. исслед., 1998. Вып. 6. 192 с.

Кудрин Б.И. Применение понятий биологии для описания и прогнозирования больших систем, формирующихся технологически // Электрификация металлургических предприятий Сибири. Вып. 3. Томск: Изд-во Томск. ун-та, 1973. С. 171-203.

Кудрин Б.И. Введение в технетику. Томск: Изд-во Томск. ун-та, 1993. 552 с.

Кудрин Б.И. Онтология и гносеология ценозов и их структурная устойчивость // Математическое описание ценозов и закономерности технетики. Ценологические исследования. Абакан: Центр системн. исслед., 1996. Вып. 1. С. 8-32.

Кудрин Б.И. Математика ценозов: видовое, ранговидовое, ранговое по параметру гиперболические H-распределения и законы Лотки, Ципфа, Парето, Мандельброта // Математический ап-

парат структурного описания ценозов и гиперболические Н-ограничения. Ценологические исследования. М.: Центр системн. исслед., 2002. Вып. 19. С. 357-412.

Кудрин Б.И. Мои семь отличий от Ципфа // Общ. и прикл. ценология. 2007. № 4. С. 25-33.

Кудрин Б.И., Кудряшов С.А., Фуфаев В.В. и др. Канонизация и управление видовой структурой ценоза. Принцип максимума энтропии // Докл. МОИП. Общ. биол. Морфология и генетика процессов роста и развития. М.: Наука, 1989. С. 69-75.

Левич А.П. Информация как структура систем // Семиотика и информатика. 1978. Вып. 10. С. 116-132.

Левич А.П. Структура экологических сообществ. М.: Изд-во Моск. ун-та, 1980. 181 с.

Левич А.П., Алексеев В.Л., Никулин В.А. Математические аспекты вариационного моделирования в экологии сообществ // Математическое моделирование. 1994. Т. 6, № 5. С. 55-71.

Любищев А.А. Редукционизм и развитие морфологии и систематики // Журн. общ. биол. 1977. Т. 38, № 2. С. 240-245.

Любищев А.А. Проблемы формы, систематики и эволюции организмов. Л.: Наука, 1982. 278 с.

Ляпунов А.А. О кибернетических вопросах биологии // Проблемы кибернетики. М.: Наука, 1972. Вып. 25. С. 5-40.

Максимов В.Н. О ранговых распределениях в экологии сообществ с точки зрения статистики // Изв. РАН. Сер. биол. 2004. № 3. С. 352-361.

Максимов В.Н., Булгаков Н.Г., Джабруева Л.В. Ранговые распределения размерно-морфологических групп микроводорослей в перифитоне и их связь с уровнем загрязнения водоема // Изв. РАН. Сер. биол. 1997. № 6. С. 697-704.

Малышев Л.И. Зависимость флористического богатства от внешних условий и исторических факторов // Бот. журн. 1969. Т. 54, № 8. С. 1137-1147.

Мандельброт Б. О рекуррентном кодировании, ограничивающем влияние помех // Теория передачи сообщений. М.: Изд-во иностр. лит., 1957. С. 139-157.

Мандельброт Б. Фрактальная геометрия природы. М.: Ин-т компьютерных исслед., 2002. 656 с.

Маслов А.А. Количественный анализ горизонтальной структуры лесных сообществ. М.: Наука, 1990. 160 с.

Мейен С.В. Основные аспекты типологии организмов // Журн. общ. биол. 1978. Т. 39, № 4. С. 495-508.

Меншуткин В.В. Имитационное моделирование водных экологических систем. СПб.: Наука, 1993. 158 с.

Миркин Б.М. Статистический анализ экоклинов // Бот. журн. 1971. Т. 56, № 12. С. 1772-1788.

Миркин Б.М. Теоретические основы современной фитоценологии. М.: Наука, 1985. 136 с.

Миркин Б.М. Еще раз об организмизме в фитоценологии // Бот. журн. 1989. Т. 74, № 1. С. 3-13.

Миркин Б.М. О растительных континуумах // Журн. общ. биол. 1990. Т. 51, № 2. С. 316-326.

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

Миркин Б.М., Наумова Л.Г., Соломещ А.И. Современная наука о растительности. М.: Логос, 2001. 263 с.

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

Миркин Б.М., Розенберг Г.С. Системный подход к фитоценологии // Журн. общ. биол. 1978б. Т. 39, № 2. С. 167-178.

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

Миркин Б.М., Розенберг Г.С. Толковый словарь современной фитоценологии. М.: Наука, 1983. 134 с.

Миркин Б.М., Розенберг Г.С., Кашапов Р.Ш. Статистический анализ широтных ценотопоклинов // Биол. науки. 1975. № 12. С. 19-24.

Миркин Б.М., Розенберг Г.С., Наумова Л.Г. Факторный анализ в фитоценологии. 4. Приложение к классификации растительности // Биол. науки. 1979. № 2. С. 102-106.

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

Миркин Б.М., Ямалов С.М., Наумова Л.Г. Синантропные растительные сообщества: модели организации и особенности классификации // Журн. общ. биол. 2007. Т. 68, № 6. С. 435-443.

Митчелл П. 101 ключевая идея: Экология. М.: Фаир-Пресс, 2001. 224 с.
Мэгарран Э. Экологическое разнообразие и его измерение. М.: Мир, 1992. 181 с.
Ниценко А.А. Об изучении экологической структуры растительного покрова // Бот. журн. 1969. Т. 54, № 7. С. 1002-1013.
Новаковский А.Б. Обзор современных программных средств, используемых для анализа геоботанических данных // Растительность России. 2006. № 9. С. 86-96.
Новаковский А.Б. Методы ординации в современной геоботанике // Вестн. Ин-та биол. Коми НЦ УрО РАН. 2008. № 10 (132). С. 2-8. [http://ib.komisc.ru/add/j2/files/ib_pdf/vestnik/10-132-2008.pdf].
Новаковский А.Б. Эколого-ценотические группы сосудистых растений в фитоценозах ландшафтов бассейна Верхней и Средней Печоры: Автореф. дис. … канд. биол. наук. Сыктывкар, 2009. 19 с.
Новик И.Б. Вопросы стиля мышления в естествознании. М.: Политиздат, 1975. 144 с.
Одум Ю. Основы экологии. М.: Мир, 1975. 740 с.
Одум Ю. Экология: В 2-х т. М.: Мир, 1986. Т. 1. 328 с.; Т. 2. 376 с.
Павлинов И.Я. Морфологическое разнообразие: общие представления и основные характеристики // Сб. тр. Зоол. музея МГУ. 2008. Т. 49 («Зоологические исследования»). С. 343-388.
Песенко Ю.А. Принципы и методы количественного анализа в фаунистических исследованиях. М.: Наука, 1982. 287 с.
Пианка Э. Эволюционная экология. М.: Мир, 1981. 400 с.
Поддубный А.Г. Экологическая топография популяций рыб в водохранилищах. Л.: Наука, 1971. 312 с.
Прилуцкий А.Н. Современная фитоценология: проблемы и решения // Бюл. Бот. сада-института ДВО РАН. 2007. Вып 1 (1). С. 5-23.
Проблемы экологического эксперимента (Планирование и анализ наблюдений) / Под ред. чл.-корр. РАН Г.С. Розенберг и д.б.н. Д.Б. Гелашвили; сост. и коммент. д.б.н. В.К. Шитикова. Тольятти: СамНЦ РАН; «Кассандра», 2008. 274 с.
Пузаченко Ю.Г., Пузаченко А.Ю. Семантические аспекты биоразнообразия // Журн. общ. биол. 1996. Т. 57, № 1. С. 1-43.
Пфанцагль И. Теория измерений. М.: Мир, 1976. 165 с.
Работнов Т.А. Фитоценология. М.: Изд-во Моск. ун-та, 1983. 296 с.
Раменский Л.Г. О сравнительном методе экологического изучения растительных сообществ // Дневник XII съезда естествоиспытателей и врачей. М., 1910. Т. 4. С. 389-390. [Антология экологии / Сост. и коммент. Г.С. Розенберга. Тольятти: ИЭВБ РАН, 2004. С. 33].
Раменский Л.Г. Основные закономерности растительного покрова и их изучение (На основании геоботанических исследований в Воронежской губ.) // Вестн. опытн. дела Средне-Черноозем. обл. Воронеж: Обл. ред.-изд. к-т Н.К.З., 1924. С. 37-73. [Антология экологии / Сост. и коммент. Г.С. Розенберга. Тольятти: ИЭВБ РАН, 2004. С. 34-48].
Раменский Л.Г. Введение в комплексное почвенно-геоботаническое исследование земель. М.: Сельхозгиз, 1938. 620 с.
Риклефс Р. Основы общей экологии. М.: Мир, 1979. 424 с.
Розенберг Г.С. Модели в фитоценологии. М.: Наука, 1984. 240 с.
Розенберг Г.С. Странный аттрактор в модели растительного сообщества на ранней стадии сукцессии // Математическое моделирование в биогеоценологии: Тез. докл. Всесоюз. шк. Петрозаводск: АН СССР, 1985. С. 199-201.
Розенберг Г.С. О системной экологии // Журн. общ. биол. 1988. Т. 49, № 5. С. 580-591.
Розенберг Г.С. О путях построения теоретической экологии // Успехи соврем. биол. 2005. Т. 125, № 1. С. 14-27.
Розенберг Г.С. Статистические методы в фитоценологии на рубеже тысячелетий (к 50-летию выхода монографии П. Грейг-Смита) // Актуальные проблемы геоботаники: III Всерос. шк.-конф. Лекции. Петрозаводск: КарНЦ РАН, 2007. С. 72-116.
Розенберг Г.С. Волжский бассейн: на пути к устойчивому развитию. Тольятти: Кассандра, 2009. 477 с.
Розенберг Г. С. Информационный индекс и разнообразие: Больцман, Котельников, Шеннон, Уивер… // Естественнонауч. бюл. «Самарская Лука: проблемы региональной и глобальной экологии». 2010. Т. 19, № 2. С. 4-25.

Розенберг Г.С., Краснощеков Г.П., Шитиков В.К. К созданию пространственно-распределенной базы эколого-экономических данных бассейна крупной реки (на примере Волжского бассейна) // Вопросы экологии и охраны природы в лесостепной и степной зонах. Самара: Изд-во Самар. ун-та, 1995. С. 8-15.

Розенберг Г.С., Мозговой Д.П., Гелашвили Д.Б. Экология. Элементы теоретических конструкций современной экологии. Самара: СамНЦ РАН, 1999. 396 с.

Розенберг Г.С., Смелянский И.Э. Экологический маятник (смена парадигм в современной экологии) // Журн. общ. биол. 1997. Т. 58, № 4. С. 5-19.

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

Романенко В.Д. Основы гидроэкологии. К.: Генеза, 2004. 664 с.

Рябко Б.Я., Кудрин Б.И., Завалишин Н.Н. и др. Модель формирования статистической структуры биоценозов // Изв. СО АН СССР. Сер. биол. 1978. Вып. 1. С. 121-127.

Сахаров В.Б., Ильяш Л.В. Применение метода функции желательности к анализу результатов изучения действия цинка и хрома на фитопланктон Рыбинского водохранилища // Биол. науки. 1982. № 8. С. 65-68.

Сёмкин Б.И. Эквивалентность мер близости и иерархическая классификация многомерных данных // Иерархические классификационные построения в географической экологии и систематике. Владивосток: ДВНЦ АН СССР, 1979. С. 97-112.

Сёмкин Б.И., Двойченков В.И. Об эквивалентности мер сходства и различия // Исследование систем. Владивосток: ДВНЦ АН СССР, 1973. С. 18-43.

Скворцов А.В., Костюк Ю.Л. Применение триангуляции для решения задач вычислительной геометрии // Геоинформатика: Теория и практика. Вып. 1. Томск: Изд-во Томск. ун-та, 1998. С. 127-138.

Слуцкий Е. Е. Сложение случайных причин как источник циклических процессов // Вопросы конъюнктуры. 1927. Т. 3, вып. 1. С. 34-64.

Снегирёва М.С., Федюнин В.А., Гелашвили Д.Б. Фрактальные аспекты характеристики сообщества паразитических насекомых на примере наездников-ихневмоид Висимского заповедника // Биологические системы: устойчивость, принципы и механизмы функционирования: Сб. материалов III Всерос. науч.-практ. конф. с междунар. участием. В 2-х ч. Нижний Тагил: НТГСПА, 2010. Ч. 2. С 213-217.

Солнцев Л.А. Изучение видовой структуры и таксономического разнообразия рецентных и ископаемых сообществ мелких млекопитающих с позиций принципа самоподобия: Автореф. дис. … канд. биол. наук. Н. Новгород, 2009. 25 с.

Страшкраба М., Гнаук А. Пресноводные экосистемы: Математическое моделирование. М.: Мир, 1989. 373 с.

Сукачев В.Н., Дылис Н.В. (ред.). Программа и методика биогеоценологических исследований. М.: Наука, 1969. 334 с.

Суховольский В.Г., Баранчиков Ю.Н. Методы количественной оценки видового разнообразия сообществ // Энтомологические исследования в Сибири. Вып. 1. Красноярск: КФ РЭО, 1998. С. 47-54.

Тейл Г. Экономические прогнозы и принятие решений. М.: Статистика, 1971. 488 с.

Тимонин А.К. О статье А.М. Гилярова "Соотношение органицизма и редукционизма как основных методологических подходов в экологии" // Журн. общ. биол. 1989. Т. 50, № 2. С. 423-426.

Трасс Х.Х. Геоботаника. История и современные тенденции развития. Л.: Наука, 1976. 252 с.

Тутубалин В.Н., Барабашева Ю.М., Григорян А.А. и др. Математическое моделирование в экологии. Историко-методологический анализ. М.: Языки русской культуры, 1999. 208 с.

Уиттекер Р. Сообщества и экосистемы. М.: Прогресс, 1980. 327 с.

Урманцев Ю.А. Начала общей теории систем // Системный анализ и научное знание. М.: Наука, 1978. С. 7-41.

Фёдоров В.Д. Изменения в природных биологических системах. М.: Спорт и культура, 2004. 368 с.

Фёдоров В.Д., Гильманов Т.Г. Экология. М.: Изд-во Моск. ун-та, 1980. 464 с.

Фёллер В. Введение в теорию вероятностей и её приложения. 3-е изд.; в 2-х т. М.: Либроком, 2010. Т. 2. 766 с.

Фуфаев В.В. Основы теории динамики структуры техноценозов // Математическое описание ценозов и закономерности технетики. Ценологические исследования. Абакан: Центр системн. исслед., 1996. Вып. 1. С. 156-193.

Фуфаев В.В., Калашников Д.А. Ценологическая модель механизма образования H-распределений на основе рождения и гибели простых чисел // Техногенная самоорганизация и математический аппарат ценологических исследований. Ценологические исследования. М.: Центр системн. исслед., 2005. Вып. 28. С. 66-90.

Хайтун С.Д. Негауссовость социальных явлений // Социологические исследования. 1983. № 1. С. 144-152.

Хорошев А.В., Алещенко Г.М. Методы выделения геосистем с единством межкомпонентных отношений // География и природные ресурсы. 2008. № 3. С. 120-126.

Хромов-Борисов Н.Н. Синдром статистической снисходительности или значение и назначение *p*-значения // Четвертая Международная телеконф. "Проблемы и перспективы современной медицины, биологии и экологии", 2011. [http://tele-conf.ru].

Цимдинь П.А., Лиепа Р.А. Концепция речного континуума. Применение на практике // Изв. АН ЛатССР. 1989. № 5. С. 60-69.

Чайковский Ю.В. Эволюция // Ценологические исследования. М.: Центр системн. исслед., 2003. Вып. 22. 472 с.

Чефранов С.Г. Об оценке достоверности соотношений видовых фондов и локальных видовых богатств // Материалы второй междунар. науч.-практ. конф. "Актуальные проблемы экологии в условиях современного мира". Майкоп: Изд-во МГТИ, 2002. С. 145-147.

Шкорбатов Г.Л. Введение в экологию. Воронеж: Изд-во Воронеж. ун-та, 1985. 76 с.

Шитиков В.К., Зинченко Т.Д. Анализ пространственно-временной изменчивости водных экосистем при статистической обработке данных мониторинга // Проблемы экологического эксперимента: (планирование и анализ наблюдений). Тольятти: СамНЦ РАН; «Кассандра», 2008. С. 129-150. [www.ievbran.ru/kiril/Article/A30/0_v.htm].

Шитиков В.К., Зинченко Т.Д., Абросимова Э.В. Непараметрические методы сравнительной оценки видового разнообразия речных сообществ макрозообентоса // Журн. общ. биол. 2010. Т. 71, № 3. С. 263-274.

Шитиков В.К., Розенберг Г.С. Оценка биоразнообразия: попытка формального обобщения // Количественные методы экологии и гидробиологии (сб. науч. тр., посвящ. памяти А.И. Баканова). Тольятти: СамНЦ РАН, 2005. С. 91-129.

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

Шитиков В.К., Розенберг Г.С., Крамаренко С.С., Якимов В.Н. Современные подходы к статистическому анализу экспериментальных данных // Проблемы экологического эксперимента (Планирование и анализ наблюдений). Тольятти: СамНЦ РАН; Кассандра, 2008. С. 212-250.

Шорыгин А.А. Питание, избирательная способность и пищевые взаимоотношения некоторых Gobiidae Каспийского моря // Зоол. журн. 1939. Т. 18, вып. 1. С. 27-51.

Шорыгин А.А. Количественный способ изучения пищевой конкуренции у рыб // Зоол. журн. 1946. Т. 25, вып. 5. С. 45-60.

Шрейдер Ю.А. О возможности теоретического вывода статистических закономерностей текста // Проблемы передачи информации. 1967. Вып. 1. С. 57-63.

Шурганова Г.В., Иудин Д.И., Гелашвили Д.Б., Якимов В.Н. Мультифрактальный анализ видового разнообразия зоопланктоценозов Чебоксарского водохранилища // Актуальные проблемы рационального использования биологических ресурсов водохранилищ. Рыбинск: Изд-во "Рыбинский дом печати", 2005. С. 294-309.

Эйнштейн А. Принципы научного исследования // Физика и реальность. М.: Наука, 1965. С. 8-10.

Экологическое состояние реки Чапаевка в условиях антропогенного воздействия (Биологическая индикация). Экологическая безопасность и устойчивое развитие Самарской области. Вып. 3. Изд. 2-е / Отв. ред. Т.Д. Зинченко и Г.С. Розенберг. Тольятти: ИЭВБ РАН, 1997. 342 с.

Экосистема малой реки в изменяющихся условиях среды / Ред. А.В. Крылов, А.А. Бобров. М.: Т-во науч. изд. КМК, 2007. 372 с.

Эфрон Б. Нетрадиционные методы многомерного статистического анализа. М.: Финансы и статистика, 1988. 263 с.

Юрцев Б.А., Петровский В.В. Об индикационном значении флористических комплексов на Северо-Востоке СССР // Теоретические вопросы фитоиндикации. Л.: Наука, 1971. С. 15-31.

Яблонский А.И. Математические модели в исследовании науки. М.: Наука, 1986. 352 с.

Якимов В.Н., Гелашвили Д.Б., Иудин Д.И. Применение мультифрактального формализма для анализа видовой структуры // Вестн. Нижегород. ун-та. Сер. биол. 2005. Вып. 2 (10). С. 94-99.

Adler P.B., Hille Ris Lambers J., Levine J.M. A niche for neutrality // Ecol. Letters. 2007. V. 10, № 2. P. 95-104. [Синопсис статьи: http://elementy.ru/news/430460].

Aho K., Roberts D.W., Weaver T.W. Using geometric and non-geometric internal evaluators to compare eight vegetation classification methods // J. Veg. Sci. 2008. V. 19, № 4. P. 549-562.

Akatov V., Chefranov S., Akatova T. The relationship between local species richness and species pool: a case study from the high mountains of the Greater Caucasus // Plant Ecol. 2005. V. 181, № 1. P. 9-22.

Allan J.D. Faunal replacement and longitudinal zonation in an alpine stream // Limnology. 1975. V. 19. P. 1646-1652.

Allan J.D., Flecker A.S. Biodiversity conservation in running waters // BioScience. 1993. V. 43. P. 32-43.

Allen T.F., Starr T.B. Hierarchy: Perspectives for Ecological Complexity. Chicago: Univ. Chicago Press, 1982. 386 p.

Almeida-Neto M., Guimarães P., Guimarães J.P. et al. A consistent metric for nestedness analysis in ecological systems: reconciling concept and quantification // Oikos. 2008. V. 117. P. 1227-1239.

Alonso D., Etienne R.S., McKane A.J. The merits of neutral theory // Trends in Ecology and Evolution. 2006. V. 21, № 8. P. 351-356.

Alonso D., McKane A.J. Sampling Hubbell's neutral theory of biodiversity // Ecol. Letters. 2004. V. 7. P. 901-910.

Amstrup S.C., Trent L., McDonald T.L. et al. Handbook of Capture-Recapture Analysis. Princeton: Univ. Press, 2005. 313 p.

Anderson M.J. A new method for non-parametric multivariate analysis of variance // Austral Ecol. 2001. V. 26. P. 32-46.

Anderson M.J., ter Braak C.J. Permutation tests for multi-factorial analysis of variance // J. Statist. Comput. Simul. 2003. V. 73. P. 85-113.

Anderson M.J., Willis T.J. Canonical analysis of principal coordinates: a useful method of constrained ordination for ecology // Ecology. 2003. V. 84. P. 511-525.

Anscombe F.J. Sampling theory of the negative binomial and logarithmic series distributions // Biometrika. 1950. V. 37. P. 358-382.

Archibald E.E. The specific character of plant communities. II. A quantitative approach // J. Ecol. 1949. V. 37. P. 260-274.

Arita H.T., Rodriguez P. Local-regional relationships and the geographical distribution of species // Global Ecol. & Biogeogr. 2004. V. 13. P. 15-21.

Arrhenius O. Species and area // J. Ecol. 1921. V. 9. P. 95-99.

Atmar W., Patterson B.D. The measure of order and disorder in the distribution of species in fragmented habitat // Oecologia. 1993. V. 96. P. 373-382.

Austin M.P. On non-linear species response models in ordination // Vegetatio. 1976. V. 33, № 1. P. 33-41.

Austin M.P. Species distribution models and ecological theory: a critical assessment and some possible new approaches // Ecol. Modelling. 2007. V. 200. P. 1-19.

Azaele S., Pigolotti S., Banavar J.R., Maritan A. Dynamical evolution of ecosystems // Nature. 2006. V. 444. P. 926-928.

Barabesi L., Fattorini L. The use of replicated plot, line and point sampling for estimating species abundances and ecological diversity // Environmental and Ecol. Statist. 1998. V. 5. P. 353-370.

Barkman J.J. Fidelity and character-species, a critical evaluation // Vegetatio. 1989. V. 85. P. 105-116.

Bascompte J., Jordano P. Plant-animal mutualistic networks: the architecture of biodiversity // Annu. Rev. Ecol. Evol. Syst. 2007. V. 38. P. 567-593.

Bastolla U., Fortuna M., Pascual-García et al. The architecture of mutualistic networks minimizes competition and increases biodiversity // Nature. 2009. V. 458. P. 1018-1020. [Синопсис статьи: http://elementy.ru/news/431066].

Bell G. The distribution of abundance in neutral communities // Amer. Nat. 2000. V. 155. P. 606-617.

Bell G. Neutral macroecology // Science. 2001. V. 293. P. 2413-2418.
Benda L., Poff L.R., Miller D. et al. Network dynamics hypothesis: spatial and temporal organization of physical heterogeneity in rivers // BioScience. 2004. V. 54. P. 413-427.
Benjamini Y., Yekutieli D. The control of false discovery rate in multiple testing under dependency // Ann. Statist. 2001. V. 29. P. 1165-1188.
Berlow E. L., Neutel A.-M., Joel E. et al. Interaction strength in food webs: issues and opportunities // J. Animal Ecol. 2004. V. 73. P. 585-598.
Bersier L.F., Sugihara G. Species abundance patterns: the problem of testing stochastic models // J. Animal Ecol. 1997. V. 66. P. 769-774.
Bertalanffy L. von. General System Theory. A critical review // General Systems. 1962. V. 7. P. 1-20 [Пер. с англ. Н.С. Юлиной: http://macroevolution.narod.ru/bertalanfi.htm].
Bezdek J.C. Numerical taxonomy with fuzzy sets // J. Math. Biol. 1974. V. 1. P. 57-71.
Bezdek J.C. Partition structures: a tutorial // Analysis of Fuzzy Information. Florida: CRC Press, Boca Raton, 1987. V. 3. P. 81-107.
Blackburn T.M., Gaston K.J. (eds.). Macroecology: Concepts and Consequences. Oxford: Blackwell, 2003. 442 p.
Blackburn T.M., Gaston K.J. There's more to macroecology than meets the eye // Global Ecol. & Biogeogr. 2006. V. 15. P. 537-540.
Blackburn T.M., Jones K.E., Cassey P., Losin N. The influence of spatial resolution on macroecological patterns of range size variation: a case study using parrots (Aves: Psittaciformes) of the world // J. Biogeogr. 2004. V. 31. P. 285-293.
Bloom S.A. Similarity indices in community studies: potential pitfalls // Mar. Ecol. 1981. V. 53. P. 7-26.
Borcard D., Legendre P. All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices // Ecol. Modelling. 2002. V. 153. P. 51-68.
Bossuyt B., Hermy M. Species turnover at small scales in dune slack plant communities // Basic Appl. Ecol. 2004. V. 5. P. 321-329.
Boyce R.L. Fuzzy set ordination along an elevation gradient on a mountain in Vermont, USA // J. Veg. Sci. 1998. V. 9. P. 191-200.
Boyce R.L., Ellison P.C. Choosing the best similarity index when performing fuzzy set ordination on binary data // J. Veg. Sci. 2001. V. 12. P. 711-720.
Bowers M.A., Brown J.H. Body size and coexistence in desert rodents: Chance or community structure? // Ecology. 1982. V. 63. P. 391-400.
Braun-Blanquet J. Eine pflanzensoziologische Exkursion durchs Unterengadin und in den schweizerischen National-park // Beitr. Geobot. Landesaufn. Schweiz. 1918. V. 4. P. 1-80.
Braun-Blanquet J. Prinzipien einer Systematik der Pflanzengesellschaften auf floristischer Grundlage // Jahrb. St. Gall. Naturwiss. Ges. 1921. V. 57. P. 305-351.
Bray J.R., Curtis J.T. An ordination of the upland forest communities of southern Wisconsin // Ecol. Monogr. 1957. V. 27. P. 325-349.
Brisse H., De Ruffray P., Grandjouan G., Hoff M. The phytosociological data base "SOPHY" // Annu. Bot. (Roma). 1995. V. 53. P. 177-223.
Brown J.H. Macroecology. Chicago: Chicago Univ. Press, 1995. 269 p.
Brown J.H., Barry J.F., Kelt D.A. Assembly rules: desert rodent communities are structured at scales from local to continental // Amer. Nat. 2000. V. 156, № 3. P. 314-321.
Brown J.H., Lomolino M.V. Biogeography. Sunderland (Massachusetts): Sinauer Ass. Inc., 1998. 691 p.
Brown J.H., Manly B.J. Restricted adaptive cluster sampling // Environmental and Ecol. Statist. 1998. V. 5. P. 49-63.
Brown J.H., Maurer B.A. Macroecology: The division of food and space among species on continents // Science. 1989. V. 243. P. 1145-1150.
Brown V.K., Southwood T.R. Secondary succession: patterns and strategies // Colonization, Succession and Stability. London: Blackwell, 1987. P. 315-337.
Brualdi R.A., Sanderson J.G. Nested species subsets, gaps, and discrepancy // Oecologia. 1999. V. 119. P. 256-264.
Bruelheide H. A new measure of fidelity and its application to defining species groups // J. Veg. Sci. 2000. V. 11. P. 167-178.
Bunge J., Fitzpatrick M. Estimating the number of species: A review // J. Amer. Statist. Ass. 1993. V. 88. P. 364-373.

Burlando B. The fractal dimension of taxonomic systems // J. Theor. Biol. 1990. V. 146. P. 99-114.
Burlando B. The fractal geometry of evolution // J. Theor. Biol. 1993. V. 163. P. 161-172.
Burnham K.P., Anderson D.R. Multimodel inference: understanding AIC and BIC in model selection // Sociol. Methods Res. 2004. V. 33, № 2. P. 261-304.
Burnham K.P., Overton W.S. Estimation of the size of a closed population when capture probabilities vary among animals // Biometrika. 1978. V. 65. P. 623-633.
Butaye J., Jacquemyn H., Honnay O., Hermy M. The species pool concept applied to forests in a fragmented landscape: dispersial limitation versus habitat limitation // J. Veg. Sci. 2002. V. 13. P. 27-34.
Butler B.J., Chazdon R.L. Species richness, spatial variation, and abundance of the soil seed bank of a secondary tropical rain forest // Biotropica. 1998. V. 30. P. 214-222.
Caley M.J., Schluter D. The relationship between local and regional diversity // Ecology. 1997. V. 71. P. 70-80.
Cao Y., Hawkins C., Larsen D.P., Van Sickle J. Effects of sample standardization on mean species detectabilities and estimates of relative differences in species richness among assemblages // Amer. Nat. 2007. V. 170, № 3. P. 381-395.
Cao Y., Larsen D.P., Hughes R.M. Evaluating sampling sufficiency in fish assemblage surveys: a similarity-based approach // Can. J. Fish. Aquatic Sci. 2001. V. 58. P. 1782-1793.
Cao Y., Larsen D.P., White D. Estimating regional species richness using a limited number of survey units // Ecoscience. 2004. V. 11. P. 23-35.
Carpenter S., Walker B., Anderies J.M., Abel N. From metaphor to measurement: resilience of what to what? // Ecosystems. 2001. V. 4. P. 765-781.
Caswell H. Community structure: a neutral model analysis // Ecol. Monogr. 1976. V. 46. P. 27-54.
Caswell H. Theory and models in ecology: a different perspective // Ecol. Modelling. 1988. V. 43. P. 33-44.
Caswell H. Matrix Population Models: Construction, Analysis, and Interpretation. Sunderland (MA): Sinauer Ass. Inc., 2001. 722 p.
Champely S., Chessel D. Measuring biological diversity using Euclidean metrics // Environmental and Ecol. Statist. 2002. V. 9. P. 167-177.
Chao A. Non-parametric estimation of the number of classes in a population // Scand. J. of Statist. 1984. V. 11. P. 265-270.
Chao A. Estimating the population size for capture-recapture data with unequal catchability // Biometrics. 1987. V. 43. P. 783-791.
Chao A. Species richness estimation // Encyclopedia of Statistical Sciences. N.Y.: Wiley, 2005. P. 7909-7916.
Chao A., Chazdon R.L., Colwell R.K., Shen T.J. A new statistical approach for assessing similarity of species composition with incidence and abundance data // Ecol. Letters. 2005. V. 8. P. 148-159.
Chao A., Chazdon R.L., Colwell R.K. Shen T.J. Abundance-based similarity indices and their estimation when there are unseen species in samples // Biometrics. 2006. V. 62. P. 361-371.
Chao A., Hwang W., Chen Y.-C., Kuo C-Y. Estimating the number of shared species in two communities // Statist. Sinica. 2000. V. 10. P. 227-246.
Chao A., Jost L., Chiang S.C. et al. A two-stage probabilistic approach to multiple-community similarity indices // Biometrics. 2008. V. 64. P. 1178-1186.
Chao A., Lee S.M., Jeng C.L. Estimating population size for capture-recapture data when capture probabilities vary by time and individual animal // Biometrics. 1992. V. 48. P. 201-216.
Chao A., Shen T.J. Nonparametric estimation of Shannon`s index of diversity when there are unseen species in sample // Environmental and Ecol. Statist. 2003. V. 10. P. 429-443.
Chave J. Neutral theory and community ecology // Ecology Letters. 2004. V. 7. P. 241-253.
Chave J., Muller-Landau H., Levin S. Comparing classical community models: theoretical consequences for patterns of diversity // Amer. Nat. 2002. V. 159. P. 1-23.
Chesson P. Mechanisms of maintenance of species diversity // Annu. Rev. Ecol. Syst. 2000. V. 31. P. 343-366.
Chytrý M., Tichý L. Diagnostic, constant and dominant species of vegetation classes and alliances of the Czech Republic: a statistical revision // Folia Fac. Sci. Nat. Univ. Masaryk. Brun. Biol. 2003. V. 108. P. 1-231.
Chytrý M., Tichý L., Holt J., Botta-Dukat Z. Determination of diagnostic species with statistical fidelity measures // J. Veg. Sci. 2002. V. 13. P. 79-90.

Clark J.S., Silman M., Kern R. et al. Seed dispersal near and far: patterns across temperate and tropical forests // Ecology. 1999. V. 80. P. 1475-1494.

Clarke K.R. Non-parametric multivariate analyses of changes in community structure // Austral. J. Ecol. 1993. V. 18. P. 117-143.

Clements F.E. Plant Succession: an Analysis of the Development of Vegetation. Washington: Carnegie Inst., 1916. 512 p.

Cody M.L. Toward a theory of continental species diversities: bird distributions over Mediterranean habitat gradients // Ecology and Evolution of Communities. Cambridge (MA): Harvard Univ. Press, 1975. P. 214-257.

Coleman B.D. On random placement and species-area relations // Math. Biosciences. 1981. V. 54. P. 191-215.

Colwell R.K., Coddington J.A. Estimating terrestrial biodiversity through extrapolation // Philosoph. Transactions of the Royal Soc. Ser. B. 1994. V. 345. P. 101-118.

Colwell R.K., Mao C.X., Chang J. Interpolating, extrapolating, and comparing incidence-based species accumulation curves // Ecology. 2004. V. 85. P. 2717-2727.

Colwell R.K., Winkler D.W. A null model for null models in biogeography // Ecological Communities: Conceptual Issues and the Evidence. Princeton: Univ. Press, 1984. P. 344-359.

Connell J.H. Diversity in tropical rain forest and coral reefs // Science. 1978. V. 199. P. 1302-1310.

Connor E.F., McCoy E.D. The statistics and biology of the species-area relationship // Amer. Nat. 1979. V. 113. P. 791-833.

Connor E.F., Simberloff D. The assembly of species communities: chance or competition? // Ecology. 1979. V. 60. P. 1132-1140.

Cormack R.M. Models for capture-recapture // Sampling Biological Populations. Fairland (MD): Intern. Cooperative Publ. House, 1979. P. 217-255.

Cormack R.M. Loglinear models for capture-recapture // Biometrics. 1989. V. 45. P. 395-413.

Cornell H.V. Species assemblages of cynipid gall wasps are not saturated // Amer. Nat. 1985. V. 126, № 4. P. 565-569.

Cornell H.V., Lawton J.H. Species interactions, local and regional processes, and limits to the richness of ecological communities: a theoretical perspective // J. Animal Ecol. 1992. V. 61. P. 1-12.

Crist T., Veech J. Additive partitioning of rarefaction curves and species–area relationships: unifying α-, β- and γ-diversity with sample size and habitat area // Ecol. Letters. 2006. V. 9. P. 923-932.

Crowley P.H. Resampling methods for computation-intensive data analysis in Ecology and Evolution // Annu. Rev. Ecol. Syst. 1992. V. 23. P. 405-447.

Curtis J.T., McIntosh R.P. An upland forest continuum in the prairie-forest border region of Wisconsin // Ecology. 1951. V. 32. P. 476-496.

Cutler A.H. Nested faunas and extinction in fragmented habitats // Conserv. Biol. 1991. V. 5. P. 496-505.

Darlington P.J. Zoogeography: Geographical Distribution of Animals. N.Y.: Wiley, 1957. 675 p.

Davenpor, J., Pugh P.J.A., McKechnie J. Mixed fractals and anisotropy in subantarctic marine macroalgae from South Georgia: implications for epifaunal biomass and abundance // Mar. Ecol. Prog. Ser. 1996. V. 136. P. 245-55.

de Caprariis P., Lindeman R.H., Collins C.M. A method for determining optimum sample size in species diversity studies // J. Intl. Assn. Math. Geol. 1976. V. 8. P. 575-581.

De Vries P.G. Sampling Theory for Forest Inventories. Berlin: Springer-Verlag, 1986. 399 p.

Dengler J. Diagnostic, constant and dominant species of the higher vegetation units of Slovakia // Bull. Eur. Dry Grassl. Group. 2008. № 4. P. 20-21.

Dengler J. Which function describes the species-area relationship best? A review and empirical evaluation // J. Biogeogr. 2009. V. 36. P. 728-744.

Dennis B., Patil G.P., Rossi O. et al. A bibliography of literature on ecological diversity and related methodology // Ecological Diversity in Theory and Practice. Fairland (MD): Intern. Cooperative Publ. House, 1979. P. 319-354.

Dennis R., Desharnais R., Cushing J., Costantino R. Nonlinear demographic dynamics: mathematical models, statistical methods, and biological experiments // Ecol. Monogr. 1995. V. 65, № 3. P. 261-281.

Dennis B., Taper M.L. Density dependence in time series observations of natural populations: estimation and testing // Ecol. Monogr. 1994. V. 64. P. 205-224.

Diamond J.M. Assembly of species communities // Ecology and Evolution of Communities. Cambridge (MA): Harvard Univ. Press, 1975. P. 342-444.
Dickey D.A., Fuller W.A. Distribution of the estimators for autoregressive time series with a unit root // J. Amer. Statist. Ass. 1979. V. 74. P. 427-431.
Diserud O.H., Ødegaard F. A multiple-site similarity measure // Biol. Letters. 2007. V. 3, № 1. P. 20-22.
Doak D.F., Bigger D., Harding E.K. et al. The statistical inevitability of stability diversity relationships in community ecology // Amer. Nat. 1998. V. 151. P. 264-276.
Dufrêne M., Legendre P. Species assemblages and indicator species: the need for a flexible asymmetrical approach // Ecol. Monogr. 1997. V. 67, № 3. P. 345-366.
Edgington E.S. Randomization Tests. N.Y.: Marcel Dekker, 1987. 341 p.
Efron B. Bootstrep methods. Another look at the Jacknife // Ann. Statist. 1979. № 7. P. 1-26.
Efron B. Bayesians, frequentists, and scientists // J. Amer. Stat. Ass. 2005. V. 100. P. 1-5.
Efron B., Thisted R. Estimating the number of unseen species: how many words did Shakespeare know? // Biometrika. 1976. V. 63. P. 435-447.
Ellenberg H., Weber H.E., Düll R. et al. Zeigerwerte von Pflanzen in Mitteleuropa // Scripta Geobotanica (Göttingen). 1991. V. 18. P. 1-248.
Elton C.S. Competition and the structure of animal communities // J. Animal Ecol. 1946. V. 15. P. 54-68.
Elton C.S. The Ecology of Invasions by Animals and Plants. London: Methuen, 1958. 181 p.
Enquist B.J., Haskell J.P., Tiffney B.H. General patterns of taxonomic diversity and biomass partitioning across tree dominated communities // Nature. 2002. V. 419. P. 610-613.
Eriksson O. The species-pool hypothesis and plant community diversity // Oikos. 1993. V. 68. P. 371-374.
Etienne R.S. A new sampling formula for neutral biodiversity // Ecol. Letters. 2005. V. 8. P. 253-260.
Etienne R.S., Alonso D. A dispersal-limited sampling theory for species and alleles // Ecol. Letters. 2005. V. 8. P. 1147-1156.
Etienne R.S., Alonso D. Neutral community theory: how stochasticity and dispersal-limitation can explain species coexistence // J. Statist. Physics. 2007. V. 128. P. 485-510.
Ewens W.J. The sampling theory of selectively neutral alleles // Theor. Popul. Biol. 1972. V. 3. P. 87-112.
Ewens W.J. Mathematical Population Genetics. Berlin; N.Y.: Springer, 1979. 325 p.
Fattorini L., Marcheselli M. Empirical investigation about statistical properties of abundance estimates based on line-intercept and network sampling of tracks // Statist. Methods and Appl. 2002. V. 11. P. 217-226.
Feinsinger P., Spears E.E., Poole R.W. A simple measure of niche breadth // Ecology. 1981. V. 62. P. 27-32.
Feoli E., Zerihun W. Fuzzy set analysis of the Ethiopian rift valley vegetation in relation to anthropogenic influences // Plant Ecol. 2000. V. 147. P. 219-225.
Fisher B.L. Improving inventory efficiency: A case study of leaf-litter ant diversity in Madagascar // Ecol. Appl. 1999. V. 9. P. 714-731.
Fisher J.A., Frank K.T. Abundance–distribution relationships and conservation of exploited marine fishes // Mar. Ecol. Prog. Ser. 2004. V. 279. P. 201-213.
Fisher R.A., Corbet, A.S., Williams C.B. The relation between the number of species and the number of individuals in a random sample of an animal population // J. Animal Ecol. 1943. V. 12. P. 42-58.
Fisher S.G., Barnes J.R., Minshall G.W. Succession in streams // Stream Ecology: Application and Testing of General Ecological Theory. N.Y.: Plenum Press, 1983. P. 7-27.
Foggo A., Frost M.T., Attrill M.J. Abundance–occupancy patterns in British estuarine macroinvertebrates // Mar. Ecol. Prog. Ser. 2003. V. 265. P. 297-302.
Freckleton R.P, Gill J.A, Noble D. et al. Large-scale population dynamics, abundance–occupancy relationships and the scaling from local to regional population size // J. Animal Ecol. 2005. V. 74. P. 353-364.
Frissell C.A., Liss W.J., Warren C., Hurley M. A hierarchical framework for stream habitat classification: viewing streams in a watershed context // J. Environmental Manag. 1986. V. 10. P. 199-214.
Gaston K.J. The Structure and Dynamics of Geographic Ranges. Oxford: Oxford Univ. Press, 2003. 304 p.
Gaston K.J., Blackburn T.M. Pattern and Process in Macroecology. Oxford: Blackwell Science, 2000. 377 p.
Gauch H.J. Multivariate Analysis in Community Ecology. Cambridge: Cambridge Univ. Press, 1982. 298 p.

Gering J.C., Crist T.O., Veech J.A. Additive partitioning of species diversity across multiple spatial scales: implications for regional conservation of biodiversity // Conserv. Biol. 2003. V. 17. P. 488-499.

Gibson D.J., Ely J.S., Collins S. L. The core-satellite species hypothesis provides a theoretical basis for Grime's classification of dominant, subordinate, and transient species // J. Ecol. 1999. V. 87. P. 1064-1067.

Gilpin M.E., Hanksi I.A. Metapopulation Dynamics: Empirical and Theoretical Investigations. London: Acad. Press, 1991. 336 p.

Gleason H.A. The individualistic concept of plant association // Bull. Torrey Bot. Club. 1926. V. 53, № 1. P. 7-26. [Пер. с англ.: Глизон Г. Индивидуалистическая концепция растительной ассоциации // Антология экологии / Сост. и коммент. Г.С. Розенберга. Тольятти: ИЭВБ РАН, 2004. С. 51-72].

Goodall D.W. Quantitative aspects of plant distribution // Biol. Rev. 1952. V. 27. P. 194-242.

Goodall D.W. Objective methods for the classification of vegetation. II. Fidelity and indicator value // Aust. J. Bot. 1953. V. 1. P. 434-456.

Goodall D.W. Numerical classification // Handbook of Vegetation Science. Pt. 5. Ordination and Classification of Communities. The Hague: Dr. W. Junk B.V. Publ., 1973. P. 575-615.

Gotelli N.J. Null model analysis of species co-occurrence patterns // Ecology. 2000. V. 81. P. 2606-2621.

Gotelli N.J., Colwell R.K. Quantifying biodiversity: Procedures and pitfalls in the measurement and comparison of species richness // Ecol. Letters. 2001. V. 4. P. 379-391.

Gotelli N.J., Ellison A.E. A primer of ecological statistics. Sunderland: Sinauer. 2004. 479 p.

Gotelli N.J., Graves G.R. Null Models in Ecology. Washington (DC): Smithonian Inst. Press, 1996. 368 p.

Gotelli N.J., McGill B.J. Null versus neutral models: what's the difference? // Ecography. 2006. V. 29. P. 793-800.

Gotelli N., Ulrich W. The empirical Bayes approach as a tool to identify non-random species associations // Oecologia. 2010. V. 162, № 2. P. 463-477.

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

Green J.L., Plotkin J.B. A statistical theory for sampling species abundances // Ecol. Letters. 2007. V. 10. P. 1037-1045.

Greve M., Chown S.L. Endemicity biases nestedness metrics: a demonstration, explanation and solution // Ecography. 2006. V. 29. P. 347-356.

Gray J.S., Ugland K.I., Lambshead J. Species accumulation and species area curves – a comment on Scheiner (2003) // Global Ecol. & Biogeogr. 2004. V. 13. P. 469-476.

Griffiths D. Local and regional species richness in North American lacustrine fish // J. Animal Ecol. 1997. V. 66. P. 49-56.

Grime J.P. Plant strategy theories: a comment on Craine (2005) // J. Ecol. 2007. V. 95, № 2. P. 227-230.

Grimm V., Wissell C. Babel, or the ecological stability discussions: an inventory and analysis of terminology and a guide for avoiding confusion // Oecologia. 1997. V. 109. P. 323-334.

Guisan A., Thuillier W. Predicting species distribution: offering more than simple habitats models // Ecol. letters. 2005. V. 8. P. 993-1009.

Guisan A., Zimmermann N.E. Predictive habitat distribution models in ecology // Ecol. Modelling. 2000. V. 135. P. 147-186.

Gurevitch J., Morrow L.L., Wallace A., Walsh J.S. A meta-analysis of field experiments on competition // Amer. Nat. 1992. V. 140. P. 539-572.

Hanski I. Dynamics of regional distribution: the core and satellite species hypothesis // Oikos. 1982. V. 38. P. 210-221.

Harris B. Determining the bounds on integrals with applications to cataloging problems // Ann. Math. Statist. 1959. V. 30. P. 521-548.

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

Harvey P.H., Colwell R.K., Silvertown J.W., May R.M. Null models in ecology // Annu. Rev. Ecol. Syst. 1983. V. 14. P. 189-211.

Hausdorf B., Hennig C. Null model tests of clustering of species, negative co-occurrence patterns and nestedness in meta-communities // Oikos. 2007. V. 116. P. 818-828.

Hawkes H.A. River zonation and classification // River Ecology. Oxford: Blackwell Sci. Publ., 1975. P. 312-374.

Hayes A.F. Permutation test is not distribution-free: testing H0: ρ = 0 // Psychol. Methods. 1996. V. 1. P. 184-198.

He F., Gaston K.J., Connor E.F., Srivastava D.S. The local-regional relationship: immigration, extinction, and scale // Ecology. 2005. V. 86. P. 360-365.

He F., Legendre P. On species-area relation // Amer. Nat. 1996. V. 148. P. 719-737.

Hellmann J.J., Fowler G.W. Bias, precision, and accuracy of four measures of species richness // Ecol. Appl. 1999. V. 9. P. 824-834.

Heltshe J., Forrester N.E. Estimating species richness using the jackknife procedure // Biometrics. 1983. V. 39. P. 1-11.

Herben T. Correlation between richness per unit area and the species pool cannot be used to demonstrate the species pool effect // J. Veg. Sci. 2000. V. 11. P. 123-126.

Hill M.O. Diversity and evenness: a unifying notation and its consequences // Ecology. 1973a. V. 54. P. 427-432.

Hill M.O. Reciprocal averaging: an eigenvector method of ordination // J. Ecol. 1973b. V. 61. P. 237-249.

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.

Hill M.O. Computerised matching of relives and association tables, with an application to the British national vegetation classification // Vegetatio. 1989. V. 83. P. 187-194.

Hill M.O., Gauch H.G. Detrended correspondence analysis: an improved ordination technique // Vegetatio. 1980. V. 42. P. 47-58.

Holdridge L.R., Grenke W.C., Hatheway W.H. et al. Forest Environments in Tropical Life Zones: a Pilot Study. Oxford: Pergamon Press, 1971. 747 p.

Horn H.S. Measurement of «overlap» in comparative ecological studies // Amer. Nat. 1966. V. 100. P. 419-424.

Horner-Devine M.C., Lage M., Hughes J. B., Bohannan B.J.M. A taxa-area relationship for bacteria // Nature. 2004. V. 432. P. 750-753.

Hubbell S.P. The Unified Neutral Theory of Biodiversity and Biogeography. Princeton: Univ. Press, 2001. 448 p.

Hubbell S.P. Modes of speciation and the lifespans of species under neutrality: a response to the comment of Robert E. Ricklefs // Oikos. 2003. V. 100. P. 193-199.

Hugueny B., Paugy D. Unsaturated fish communities in African rivers // Amer. Nat. 1995. V. 146. P. 162-169.

Huisman J., Olff H., Fresco L.F. A hierarchical set of models for species response analysis // J. Veg. Sci. 1993. V. 4. P. 37-46.

Hulten E. Outline of the History of Arctic and Boreal Biota During the Quaternary Period. Stockholm: Bokforlags Aktiebolaget Thule, 1937. 345 p.

Hurlbert S.H. The non-concept of species diversity: a critique and alternative parameters // Ecology. 1971. V. 52. P. 577-586.

Hutchinson G.E. Homage to Santa Rosalia or Why are there so many kinds of animals? // Amer. Nat. 1959. V. 93. P. 145-59. [Пер. с англ.: Хатчинсон Дж. Дань Санта Розалии, или Почему так много видов животных? // Антология экологии / Сост. и коммент. Г.С. Розенберга. Тольятти: ИЭВБ РАН, 2004. С. 263-281].

Hutchinson G.E., MacArthur R.H. A theoretical ecological model of size distributions among species // Amer. Nat. 1959. V. 93. P. 117-125.

Inagaki H., Lenoir A. Une etude d'ecologie evolutive: application de la loi de Motomura aux fourmis // Bull. Ecol. 1974. V. 5, № 3. P. 207-219.

Izsak J., Szeidl L. Quadratic diversity: its maximization can reduce the richness of species // Environmental and Ecol. Statist. 2002. V. 9. P. 423-430.

Jancey R.C. Species ordering on a variance criterion // Vegetatio. 1979. V. 39. P. 59-63.

Jongman R.H.G., ter Braak C.J.F., van Tongeren O.F.R. Data Analysis in Community and Landscape Ecology. Wageningen (The Netherlands): Pudoc, 1987. 299 p. [Пер. с англ.: Джонгман Р.Г.Г., тер Браак С.Дж.Ф., ван Тонгерен О.Ф.Р. Анализ данных в экологии сообществ и ландшафтов. М.: РАСХН. 1999. 306 с.].

Jorgensen S.E., Mejer H., Nielsen S.N. Ecosystem as self-organizing critical systems // Ecol. Modelling. 1998. V. 111. P. 261-268.
Jost L. Entropy and diversity // Oikos. 2006. V. 113. P. 363-375.
Jost L. Partitioning diversity into independent alpha and beta components // Ecology. 2007. V. 88, № 10. P. 2427-2439.
Jost L. GST and its relatives do not measure differentiation // Molecular Ecol. 2008. V. 17. P. 4015-4026.
Karr J.R. Biological integrity: a long-neglected aspect of water resource management // Ecol. Appl. 1991. V. 1. P. 66-84.
Kassen R., Buckling A., Bell G., Rainey P.B. Diversity peaks at intermediate productivity in a laboratory microcosm // Nature. 2000. V. 406. P. 508-512.
Kenkel N. Trends and interrelationship in boreal wetland vegetation // Can. J. Bot. 1987. V. 65. P. 12-22.
Keylock C. Simpson diversity and the Shannon-Wiener index as special cases of a generalized entropy // Oikos. 2005. V. 109. P. 203-207.
Koleff P, Gaston K.J., Lennon J.J. Measuring beta diversity for presence–absence data // J. Animal Ecol. 2003. V. 72. P. 367-382.
Kolkwitz R., Marsson M. Grundsätze für die biologishe Beurtheilung des Wassers hach seiner Flora und Fauna // Mitt. aus der konigl. Prüfungsanst. für Wasserversorg. und Abwasserreinig. 1902. H. 1. P. 33-72.
Krebs C.J. Ecological Methodology. N.Y.: Harper Collins Publ., 1989. 654 p.
Krishna A., Guimarães P.R., Jordano P., Bascompte J. A neutral-niche theory of nestedness in mutualistic networks // Oikos. 2008. V. 117. P. 1609-1618.
Lambshead P.J., Platt H.M. Structural patterns of marine benthic assemblages and their relationship with empirical statistical models // Proc. of the nineteenth European marine biology symp. (ed. P.E. Gibbs). Cambridge: Cambridge Univ. Press, 1985. P. 371-380.
Lambshead P.J., Platt H.M., Shaw K.M. The detection of differences among assemblages of marine benthic species based on an assessment of dominance and diversity // J. Nat. Hist. 1983. V. 17. P. 859-874.
Lande R. Statistics and partitioning of species diversity, and similarity among multiple communities // Oikos. 1996. V. 76. P. 5-13.
Lande R., Engen S., Sæther B.E. Stochastic Population Dynamics in Ecology and Conservation. Oxford (NY): Oxford Univ. Press, 2003. 212 p.
Lau K.-S. Characterization of Rao's quadratic entropies // Sankhya. 1985. V. 47. P. 295-309.
Lawlor L.R. Structure and stability in natural and randomly constructed competitive communities // Amer. Nat. 1980. V. 116. P. 394-408.
Lecordier C., Lavelle P. Application du modele de Motomura aux peuplements de vers de terre: signification et limites // Rev. Ecol. et Biol. Sol. 1982. V. 19, № 2. P. 177-191.
Legendre P., Anderson M.J. Distance-based redundancy analysis: testing multispecies responses in multifactorial ecological experiments // Ecol. Monogr. 1999. V. 69. P. 1-24.
Legendre P., Gallagher E. Ecologically meaningful transformations for ordination of species data // Oecologia. 2001. V. 129. P. 271-280.
Legendre P., Legendre L. Numerical Ecology. Amsterdam: Elsevier Sci. BV, 1998. 853 p.
Lehman A., Overton J.M., Leathwick J.R. Grasp: generalized regression analysis and spatial prediction // Ecol. Modelling. 2002. V. 157. P. 189-207.
Leibold M.A., Mikkelson G.M. Coherence, species turnover, and boundary clumping: elements of metacommunity structure // Oikos. 2002. V. 97. P. 237-250.
Leibold M.A., Holyoak M., Moquet N. et al. The metacommunity concept: a framework for multi-scale community ecology // Ecol. Letters. 2004. V. 7. P. 601-613.
Leigh E.G. Neutral theory: a historical perspective // J. Evolut. Biol. 2007. V. 20. P. 2075-2091.
Lennon J.J., Kole P., Greenwood J.J., Gaston K.J. The geographical structure of British bird distributions: diversity, spatial turnover and scale // J. Animal Ecol. 2001. V. 70. P. 966-979.
Levins R. Some demographic and genetic consequences of environmental heterogeneity for biological control // Bull. Entomol. Soc. Amer. 1969. V. 15. P. 237-240.
Lhommel J.-P., Winkel T. Diversity–stability relationships in community ecology: Re-examination of the Portfolio effect // Theor. Popul. Biol. 2002. V. 62. P. 271-279.
Ljung G.M., Box G.E.P. On a measure of a lack of fit in time series models // Biometrika. 1978. V. 65. P. 297-303.

Loreau M., Naeem S., Inchausti P. et al. Biodiversity and ecosystem functioning: current knowledge and future challenges // Science. 2001. V. 294. P. 804-808.
Lorenz M.O. Methods of measuring the concentration of wealth // J. Amer. Statist. Ass. 1905. V. 9. P. 209-219.
Lotka A. The frequency distribution of scientific productivity // J. Wash. Acad. Sci. 1926. V. 16. P. 317-323.
Lyons J. The length of stream to sample with a towed electrofishing unit when fish species richness is estimated // N. Amer. J. Fish. Manag. 1992. V. 12. P. 198-203.
MacArthur R.H. On the relative abundance of bird species // Proc. Natl. Acad. Sci. USA. 1957. V. 45. P. 293-295.
MacArthur R.H. On the relative abundance of species // Amer. Nat. 1960. V. 94. P. 25-36.
MacArthur R.H. Patterns of species diversity // Biol. Rev. 1965. V. 40. P. 510-533. [Пер. с англ.: Мак-Артур Р. Модели видового разнообразия // Антология экологии / Сост. и коммент. Г.С. Розенберга. Тольятти: ИЭВБ РАН, 2004. С. 297-330].
MacArthur R.H. Geographical Ecology. Pattern in the Distribution of Species. N.Y.: Harper Row Publ., 1972. 269 p.
MacArthur R.H., Wilson E.O. An equilibrium theory of insular zoogeography // Evolution. 1963. V. 17, № 4. P. 373-387.
Magurran A.E. Measuring Biological Diversity. 2nd ed. Oxford: Blackwell Publ., 2004. 260 p. [Пер. с англ.: Мэгарран Э. Экологическое разнообразие и его измерение. М.: Мир, 1992. 181 с.].
Magurran A.E., Henderson P.A. Explaining the excess of rare species in natural species abundance distributions // Nature. 2003. V. 422. P. 714-716.
Malécot G. Les mathématiques de l'hérédité. Paris: Masson & Cie, 1948. 63 p.
Manly B.F. Randomization, Bootstrap and Monte Carlo Methods in Biology. London: Chapman & Hall, 2007. 455 p.
Manni F., Guerard E., Heyer E. Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by using Monmonier's algorithm // Human Biol. 2004. V. 76, № 2. P. 173-190.
Mantel N. The detection of disease clustering and a generalized regression approach // Cancer Res. 1967. V. 27. P. 209-220.
Mao C.X., Colwell R.K. Estimation of species richness: mixture models, the role of rare species, and inferential challenges // Ecology. 2005. V. 86, № 5. P. 1143-1153.
Mao C.X., Colwell R.K., Chang J. Estimating species accumulation curves using mixtures // Biometrics. 2005. V. 61. P. 433-441.
Mao C.X., Lindsay B.G. Estimating the number of classes in multiple populations: a geometric analysis // Canad. J. Statist. 2004. V. 32, № 3. P. 1-12.
Marsili-Libelli S. Fuzzy clustering of ecological data // Coenoses. 1989. V. 4. P. 95-106.
Maunder M.N. Population viability analysis, based on combining integrated, Bayesian, and hierarchical analyses // Acta Oecologica. 2004. V. 26. P. 85-94.
Maurer B.A., McGill B.J. Neutral and non-neutral macroecology // Basic and Appl. Ecol. 2004. V. 5. P. 413-422.
Maurer B.A., Taper M.L. Connecting geographical distributions with population processes // Ecol. Letters. 2002. V. 5. P. 223-231.
May R.M. Stability and Complexity in Model Environments. Princeton: Princeton Univ. Press, 1973. 292 p.
May R.M. Patterns of species abundance and diversity // Ecology and Evolution of Communities. Cambridge (MA): Harvard Univ. Press, 1975. P. 81-120.
May R.M. The search for patterns in the balance of nature – advances and retreats // Ecology. 1986. V. 67. P. 1115-1126.
McCann K.S. The diversity–stability debate // Nature. 2000. V. 405. P. 228-233.
McCullagh P., Nelder J.A. Generalized Linear Models. London: Chapman & Hall, 1989. 511 p.
McCune B. Non-parametric habitat models with automatic interactions // J. Veg. Sci. 2006. V. 17. P. 819-830.
McCune B., Grace J.B., Urban D.L. Analysis of Ecological Communities. Gleneden Beach (Oregon): MjM Software, 2002. 285 p.
McGill B.J. Strong and weak tests of macro-ecological theory // Oikos. 2003a. V. 102. P. 679-685.
McGill B.J. A test of the unified theory of biogeography // Nature. 2003b. V. 422. P. 881-885.

McGill B.J., Etienne R.S., Gray J.S. et al. Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework // Ecol. Letters. 2007. V. 10. P. 995-1015.

McGill B.J., Maurer, B.A., Weiser M.D. Empirical evaluation of the neutral theory // Ecology. 2006. V. 87. P. 1411-1423.

McIntosh R.P. An index of diversity and the relation of certain concepts to diversity // Ecology. 1967a. V. 48. P. 392-404.

McIntosh R.P. The continuum concept of vegetation // Bot. Rev. 1967b. V. 130. P. 130-187.

Mielke P.W. Meteorological applications of permutation techniques based on distance functions // Handbook of Statistics. V. 4. Nonparametric Methods. Amsterdam: North-Holland, 1984. P. 813-830.

Mikloš I., Podani J. Randomization of presence absence matrices: comments and new algorithms // Ecology. 2004. V. 85. P. 86-92.

Miller R. I., White P.S. Considerations for preserve design based on the distribution of rare plants in Great Smoky Mountains National Park, U.S.A. // J. Environmental Manag. 1986. V. 10. P. 119-124.

Milne B.T. Motivation and benefits of complex systems approaches in ecology // Ecosystems. 1998. V. 1. P. 449-456.

Minchin P.R. An evaluation of the relative robustness of techniques for ecological ordination // Vegetatio. 1987. V. 67. P. 1167-1179.

Minshall G.W. Stream ecosystem theory: a global perspective // J. North. Amer. Benthol. Soc. 1988. V. 7, № 4. P. 263-288.

Monmonier M. Maximum-difference barriers: An alternative numerical regionalization method // Geogr. Anal. 1973. V. 3. P. 245-261.

Monod J. La technique de culture continue theorie et applications // Annls. Inst. Pasteur. Paris, 1950. V.79. P. 390-401.

Montgomery D.R. Process domains and the river continuum concept // J. Amer. Water Res. Ass. 1999. V. 35. P. 397-410.

Moran P.A.P. Notes on continuous stochastic phenomena // Biometrika. 1950. V. 37. P. 17-23.

Morisita M. Measuring of interspecific association and similarity between communities // Memoirs of the Faculty of Sci. of Kyushu Univ. Ser. E. Biology. 1959. V. 3. P. 65-80.

Morisita M. Composition of the I-index // Res. Pop. Ecol. 1971. V. 13. P. 1-27.

Motomura I. On the statistical treatment of communities [in Japanese] // Zool. Mag. 1932. V. 44. P. 379-383.

Mouquet N., Munguia P., Kneitel J.M., Miller T.E. Community assembly time and the relationship between local and regional species richness // Oikos. 2003. V. 103. P. 618-626.

Mueller-Dombois D., Ellenberg H. Aims and Methods of Vegetation Ecology. N.Y.: John Willey & Sons, 1974. 547 p.

Muneepeerakul R., Levin S.A., Rinaldo A., Rodriguez-Iturbe I. On biodiversity in river networks: a trage-off metapopulation model and comparative analysis // Water Resources Res. 2007. V. 43, W07426. 11 p.

Nielsen A., Bascompte J. Ecological networks, nestedness and sampling effort // J. Ecol. 2007. V. 95. P. 1134-1141.

O'Hara R.B. Species richness estimates: How many species can dance on the head of a pin? // J. Animal Ecol. 2005. V. 74. P. 375-386.

Oksanen J., Minchin P.R. Instability of ordination results under changes in input data order: explanations and remedies // J. of Vegetation Sci. 1997. V. 8. P. 447-454.

Oksanen J., Minchin P.R. Continuum theory revisited: what shape are species responses along ecological gradients? // Ecol. Modelling. 2002. V. 157. P. 119-129.

Olszewski T.D. A unified mathematical framework for the measurement of richness and evenness within and among multiple communities // Oikos. 2004. V. 104. P. 377-387.

Otis D.L., Burnham K.P., White G.C., Anderson D.R. Statistical inference from capture data on closed animal populations // Wildlife Monogr. 1978. V. 62. P. 1-135.

Palmer M.W. The estimation of species richness by extrapolation // Ecology. 1990. V. 71. P. 1195-1198.

Pareto V. Cours d'économic politique. Lausanne; Paris: Rouge and Cie, 1897. V. 1. 299 p.

Partel M., Zobel M., Zobel K., van der Maarel E. The species pool and its relation to species richness: evidence from Estonian plant communities // Oikos. 1996. V. 75. P. 111-117.

Patil G.P., Taille C. An overview of diversity // Ecological Diversity in Theory and Practice. Fairland (MD): Intern. Cooperative Publ. House, 1979. P. 3-27.
Patil G.P., Taille C. Diversity as a concept and its measurement // J. Amer. Statist. Ass. 1982. V. 77. P. 548-567.
Patterson B.D. The principle of nested subsets and its implications for biological conservation // Conserv. Biol. 1987. V. 1. P. 323-334.
Patterson B.D., Atmar W. Nested subsets and the structure of insular mammalian faunas and archipelagos // Biol. J. Linn. Soc. 1986. V. 28. P. 65-82.
Patterson B.D., Atmar W. Analyzing species composition in fragments // Bonner Zool. Monogr. 2000. H. 46. P. 9-24.
Pedersen T.C., Baattrup-Pedersen A., Madsen T.V. Effects of stream restoration and management on plant communities in lowland streams // Freshwater Biol. 2006. V. 51. P. 161-179.
Peterson D.L., Parker V.T. (eds.). Ecological Scale – Theory and Applications. N.Y.: Columbia Univ. Press, 1998. 608 p.
Perry J.A., Schaeffer D.J. The longitudinal distributions of or riverine benthos: a river discontinuum? // Hydrobiologia. 1987. V. 148. P. 257-268.
Pianka E.R. On lizard species diversity: North American flatland deserts // Ecology. 1967. V. 48, № 3. P. 333-351.
Pielou D.P., Pielou E.C. Association among species of infrequent occurrence: the insect and spider fauna of *Polyporus betulinus* (Bulliard) Fries // J. Theor. Biol. 1968. V. 21. P. 202-216.
Pielou E.C. Niche width and niche overlap: a method for measuring them // Ecology. 1972. V. 53, № 4. P. 687-692.
Pielou E.C. Ecological Diversity. N.Y.: Wiley, 1975. 165 p.
Pielou E.C. The Interpretation of Ecological Data: A Primer on Classification and Ordination. N.Y.: Wiley, 1984a. 263 p.
Pielou E.C. Probing multivariate data with random skewers: a preliminary to direct gradient analysts // Oikos. 1984b. V. 42. P. 161-165.
Pignatti S., Mengarda F. Un nuovo procedimento per l'elaborazione delle tabelle fitosociologiche // Acad. Nat. dei Lincei. Ser. 8. 1962. V. 32, № 2. P. 215-221.
Pillar V.D. A randomization-based solution for vegetation classification and homogeneity testing // Coenoses. 1996. V. 11. P. 29-36.
Pillar V.D. The bootstrapped ordination reexamined. // J. Vegetation Sci. 1999a. V. 10. P. 895-902.
Pillar V.D. On the identification of optimal plant functional types // J. Vegetation Sci. 1999б. V. 10. P. 631-640.
Pillar V.D., Orlóci L. Fuzzy components in community level comparisons. // Computer Assisted Vegetation Analysis. Dordrecht: Kluwer, 1991. P. 87-93.
Platt J.R. Strong inference // Science. 1964. V. 146. P. 347-353.
Plotkin J.B., Muller-Landau H.C. Sampling the species composition of a landscape // Ecology. 2002. V. 83. P. 3344-3356.
Plotkin J.B., Potts M.D., Leslie N. et al. Species-area curves, spatial aggregation, and habitat specialization in tropical forests // J. of Theor. Biol. 2000. V. 207. P. 81-99.
Prentice I.C. Non-metric ordination methods in ecology // J. Ecol. 1977. V. 65. P. 85-94.
Preston F.W. The commonness and rarity of species // Ecology. 1948. V. 29. P. 254-283.
Preston F.W. The canonical distribution of commonness and rarity of species // Ecology. 1962. V. 43. P. 185-215, 410-432.
Preston F.W. Pseudo-lognormal distributions // Ecology. 1981. V. 62. P. 355-364.
Quinn G.P., Keough M.J. Experimental Design and Data Analysis for Biologists. Cambridge: Cambridge Univ. Press, 2002. 520 p.
Raaijmakers J.G. Statistical analysis of the Michaelis–Menten equation // Biometrics. 1987. V. 43. P. 793-803.
Rao C.R. The use and interpretation of principal component analysis in applied research // Sankhya. 1964. Ser. A. V. 26. P. 329-358.
Rao C.R. Diversity and dissimilarity coefficients: a unified approach // Theor. Popul. Biol. 1982. V. 21. P. 24-43.
Rao C.R. A review of canonical coordinates and an alternative to correspondence analysis using Hellinger distance // Qüestiió. 1995. V. 19. P. 23-63.

Renkonen O. Statisch-oëkologische Untersuchungen uëber die terrestiche Kaferwelt der finnischen Bruchmoore // Ann. Zool. Soc. Zool.-Bot. Fennicae "Vanamo". 1938. V. 6. P. 1-123.
Renshaw E. Modelling Biological Population in Space and Time. Cambridge: Cambridge Univ. Press, 1991. 422 p.
Rényi A. On measures of entropy and information // Proc. of the 4th Berkeley Symp. on Mathematical Statistics and Probability. Berkley: California Univ. Press, 1961. V. 1. P. 547-561.
Resh V.H., Brown A.V., Covich A.L. et al. The role of disturbance in stream ecology // J. Noth. Amer. Benthol. Sot. 1988. V. 7. P. 433-455.
Ricklefs R.E. Community diversity: relative roles of local and regional processes // Science. 1987. V. 235. P. 167-171.
Ricklefs R.E. A comment on Hubbell's zero-sum ecological drift model // Oikos. 2003. V. 100. P. 185-192.
Ricklefs R.E., Lau M. Bias and dispersion of overlap indices: results of some Monte Carlo simulations // Ecology. 1980. V. 61. P. 1019-1024.
Ricklefs R.E., Schluter D. Species diversity: regional and historical influences // Species Diversity in Ecological Communities: Historical and Geographical Perspectives. Chicago: Chicago Univ. Press, 1993. P. 350-362.
Ricotta C. Additive partition of parametric information and its associated β-diversity measure // Acta Biotheor. 2003. V. 51. P. 91-100.
Ricotta C., Avena G.C. On the information-theoretical meaning of Hill's parametric evenness // Acta Biotheor. 2002. V. 50. P. 63-71.
Roberts D.W. Ordination on the basis of fuzzy set theory // Vegetatio. 1986. V. 66. P. 123-131.
Roberts D.W. A dynamical systems perspective on vegetation theory // Vegetatio. 1987. V. 69. P. 27-33.
Roberts D. Vegetation classification in R, for labdsv ver. 1.1-1, vegetation ecology package. 2005. [http://www.cran.r-project.org. unpubl].
Roberts D.W. Statistical analysis of multidimensional fuzzy set ordinations // Ecology. 2008. V. 89, № 5. P. 1246-1260.
Roberts D.W. Comparison of multidimensional fuzzy set ordination with CCA and DB-RDA // Ecology. 2009. V. 90, № 9. P. 2622-2634.
Rodriguéz-Gironés M.A., Santamarıa L. A new algorithm to calculate the nestedness temperature of presence-absence matrices // J. Biogeogr. 2006. V. 33. P. 924-935.
Roleček J., Tichý L., Zelený D., Chytrý M. Modified TWINSPAN classification in which the hierarchy respects cluster heterogeneity // J. Vegetation Sci. 2009. V. 20. P. 596-602.
Rosenberg D.M., Resh V.H. (eds.). Freshwater Biomonitoring and Benthic Macroinvertebrates. London: Chapman and Hall, 1993. 488 p.
Rousseau R., van Hecke P., Nijssen D., Bogaert J. The relationship between diversity profiles, evenness and species richness based on partial ordering // Environmental and Ecol. Statist. 1999. V. 6. P. 211-223.
Ruokolainen K., Tuomisto H. Beta-diversity in tropical forests // Science. 2002. V. 297. P. 1439.
Sale P.F. Stock-recruit relationships and regional coexistence in a lottery competitive system: A simulation study // Amer. Nat. 1982. V. 120. P. 139-159.
Sanders H. Marine benthic diversity: a comparative study // Amer. Nat. 1968. V. 102. P. 243-282.
Sanderson J.G. Null model analysis of communities on gradients // J. Biogeogr. 2004. V. 31. P. 879-883.
Sankaran M., McNaughton S.J. Determinants of biodiversity regulates compositional stability of communities // Nature. 1999. V. 401. P. 691-693.
Schaffer W.M., Ellner S., Kot M. Effects of noise on some dynamical models in ecology // J. Math. Biol. 1986. V. 24. P. 479-523.
Scheiner S.M. Six types of species–area curves // Global Ecol. & Biogeogr. 2003. V. 12. P. 441-447.
Scheuring I. The fractal nature of vegetation and the species-area relation // Theor. Popul. Biol. 1991. V. 39. P. 170-177.
Scheuring I., Riedi R.H. Application of multifractals to the analysis of vegetation patterns // J. Vegetation Sci. 1994. V. 5. P. 489-496.
Schlosser I.J. Stream fish ecology: a landscape perspective // BioScience. 1991. V. 41. P. 704-712.
Schluter D. A variance test for detecting species associations, with some example applications // Ecology. 1984. V. 65. P. 998-1005.

Schoener T.W. The species-area relation within archipelagos: models and evidence from island land birds // Proc. of the 16th Intern. Ornithological Conf. Canberra: Austral. Acad. Sci., 1976. P. 629-642.
Seber G.A. The Estimation of Animal Abundance and Related Parameters. Caldwell; New Jersey: Blackburn Press, 1982. 654 p.
Shockley W. On the statistics of individual variations of productivity in research laboratories // Proc. Inst. Radio Engineers. 1957. V. 45. P. 279-290.
Simberloff D. Use of rarefaction and related methods in ecology // Biological Data in Water Pollution Assessment: Quantitative and Statistical Analyses. Philadelphia: Amer. Soc. for Testing and Materials, 1978. P. 150-165.
Simberloff D., Martin J.L. Nestedness of insular avifaunas – simple summary statistics masking complex species patterns // Ornis Fenn. 1991. V. 68. P. 178-192.
Smith E.P., van Belle G. Nonparametric estimation of species richness // Biometrics. 1984. V. 40. P. 119-129.
Snegireva M.S. Fractal species' structure characteristic of parasitic communities // Types of Strategy and Not Only… (Materials of the Fourth Russian-Polish School of Young Ecologists; Togliatti, September). Togliatti: Kassandra, 2010. P. 57-58.
Sohn M.W. Distance and cosine measures of niche overlap // Social Networks. 2001. V. 23. P. 141-165.
Soininen J. The ecological characteristics of idiosyncratic and nested diatoms // Protist. 2008. V. 159. P. 65-72.
Sokal R., Sneath P. Principles of Numerical Taxonomy. San Francisco: W.H. Freeman, 1963. 573 p.
Sokal R.R., Rohlf F.J. Biometry: The Principles and Practice of Statistics in Biological Research. N.Y.: W.H. Freeman, 1981. 887 p.
Southwood T.R. Ecological Methods with Particular Reference to the Study of Insect Populations. London: Methuen, 1966. 391 p.
Southwood T.R. The concept and nature of the community // Organisation of Communities. 27th Symp. of the British Ecological Society. Oxford: Blackwell Sci. Publ., 1986. P. 3-27.
Srivastava D. Using local-regional richness plots to test for species saturations: pitfalls and potentials // J. Animal Ecol. 1999. V. 68. P. 1-16.
Statzner B., Higler B. Stream hydraulics as a major determinant of benthic invertebrate zonation patterns // Freshwater Biol. 1986. V. 16. P. 127-139.
Stohlgren T.J., Quinn J.F., Ruggiero M., Waggoner G. Status of biotic inventories in U.S. National Parks // Biol. Conserv. 1995. V. 71. P. 97-106.
Stone L., Roberts A. The checkerboard score and species distributions // Oecologia. 1990. V. 85. P. 74-79.
Stone L., Roberts A. Competitive exclusion or species aggregation? An aid in deciding // Oecologia. 1992. V. 91. P. 419-424.
Strong D.R., Simberloff D. Straining at gnats and swallowing ratios: character displacement // Evolution. 1981. V. 35. P. 810-812.
Sugihara G. Minimal community structure: an explanation of species abundance patterns // Amer. Nat. 1980. V. 116. P. 770-787.
Sugihara G., Ye H. Complex systems: Cooperative network dynamics // Nature. 2009. V. 458. P. 979-980.
Szafer W., Pawlowski B. Die Pflanzenassoziationen des Tatra-Gebirges. A. Bemerkungen über die angewandte Arbeitstechnik // Die Pflanzen assoziationen des Tatra-Gebirges. 1927. Teil III. P. 1-12.
Tanaka R. New regression formula to estimate the whole population for recapture-addicted small mammals // Res. Popul. Ecol. 1967. V. 9. P. 83-94.
Taneja I. On generalized information measures and their applications // Advances in Electronics and Electron Physics. London: Acad. Press, 1989. V. 76. P. 327-413.
Taylor D.R., Aarssen Z.W., Zochle C. On the relationship between r/K selection and environmental carrying capacity: a new habitat template for plant life history strategies // Oikos. 1990. V. 58. P. 239-250.
Taylor D.R., Kempton R.A., Woiwod J.P. Diversity statistics and log-series model // J. Animal Ecol. 1976. V. 45. P. 255-271.
ter Braak C.J. Principal components biplots and alpha and beta diversity // Ecology. 1983. V. 64. P. 454-462.
ter Braak C.J. Correspondence analysis of incidence and abundance data: properties in terms of a unimodal response model // Biometrics. 1985. V. 41. P. 859-873.

ter Braak C.J. Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis // Ecology. 1986. V. 67. P. 1167-1179.

ter Braak C.J. Partial canonical correspondence analysis // Bock H.H. (Ed.): Classification and related methods of data analysis. Amsterdam: North Holland, 1988. P. 551-558.

ter Braak C.J. Canonical community ordination. Part I: Basic theory and linear methods // Ecoscience. 1994. V. 1. P. 127-140.

Terborgh J.W., Faaborg J. Saturation of bird communities in the West Indies // Amer. Nat. 1980. V. 116. P. 178-195.

Thompson S.K. Sampling. N.Y.: John Wiley and Sons, 1992. 360 p.

Thompson S.K., Seber G. A. Adaptive Sampling. N.Y.: John Wiley and Sons, 1996. 288 p.

Thoms M.C., Parsons M. Identifying spatial and temporal patterns in the hydrological character of the Condamine-Balonne River, Australia, using multivariate statistics // River Res. & Appl. 2003. V. 19. P. 443-457.

Thorp J.H., Thoms M.C., DeLong M.D. The riverine ecosystem synthesis: biocomplexity in river networks across space and time // River Res. & Appl. 2006. V. 22. P. 123-147.

Tichy L., Chytry M. Statistical determination of diagnostic species for site groups of unequal size // J. Vegetation Sci. 2006. V. 17. P. 809-818.

Tilman D. Biodiversity: Population versus ecosystem stability // Ecology. 1996. V. 77. P. 350-363.

Timi J.T., Poulin R. Different methods, different results: temporal trends in the study of nested subset patterns in parasite communities // Parasitology. 2008. V. 135. P. 131-138.

Tjørve E. Shapes and functions of species-area curves: a review of possible models // J. Biogeogr. 2003. V. 30. P. 827-835.

Tokeshi M. Niche apportionment or random assortment: species abundance patterns revisited // J. Animal Ecol. 1990. V. 59. P. 1129-1146.

Tokeshi M. Species abundance patterns and community structure // Adv. Ecol. Res. 1993. V. 24. P. 111-186.

Tokeshi M. Species coexistence: ecological and evolutionary perspectives. N.Y.: Wiley-Blackwell, 1999. 454 p.

Townsend C.R. The patch dynamics concept of stream community ecology // J. North Amer. Benthological Soc. 1989. V. 8. P. 36-50.

Townsend C.R., Scarsbrook M.R., Dolédec S. Quantifying disturbance in streams: alternative measures of disturbance in relation to macroinvertebrate species traits and species richness // J. North Amer. Benthological Soc. 1997. V. 16. P. 531-544.

Tsallis C. Possible generalization of Boltzmann-Gibbs statistics // J. Stat. Phis. 1988. V. 52. P. 479-487.

Tsallis C., Mendes R., Plastino A. The role of constraints within generalized nonextensive statistics // Physica. 1998. V. 261. P. 534-554.

Turner M.G. Landscape ecology: the effect of pattern on process // Annu. Rev. Ecol. Syst. 1989. V. 20. P. 171-197.

Ugland K.I., Gray J.S., Ellingsen K.E. The species-accumulation curve and estimation of species richness // J. Animal Ecol. 2003. V. 72. P. 888-897.

Ulrich W. Models of relative abundance distributions I: model fitting by stochastic models // Pol. J. Ecol. 2001. V. 49. P. 145-157.

Ulrich W., Almeida-Neto M., Gotelli N.C. A consumer's guide to nestedness analysis // Oikos. 2009. V. 118. P. 3-17.

Ulrich W., Gotelli N.J. Null model analysis of species nestedness patterns // Ecology. 2007. V. 88. P. 1824-1831.

van der Maarel E. Transformation of cover-abundance values in phytosociology and its effects on community similarity // Vegetatio. 1979. V. 39. P. 97-114.

Vannote R.L., Minshall G.W., Cummins K.W. et al. The river continuum concept // Can. J. Fish. Aquatic Sci. 1980. V. 37. P. 130-137.

Veech J.A., Summerville K.S., Crist T.O., Gering J.C. The additive partitioning of species diversity: recent revival of an old idea // Oikos. 2002. V. 99. P. 3-9.

Volkov I., Banavar J.R., Hubbell S.P., Maritan A. Neutral theory and relative species abundance in ecology // Nature. 2003. V. 424. P. 1035-1037.

Volterra V. Lecions sur la Theorie Mathematique de la Lutte pour via. Paris: Gauthier Villars, 1931. [Пер. с фран.: Вольтерра В. Математическая теория борьбы за существование. М.: Наука, 1976. 286 с.].

Wagner H.H., Wildi O., Ewald K.C. Additive partitioning of plant species diversity in an agricultural mosaic landscape // Lands. Ecol. 2000. V. 15. P. 219-227.

Webb T.J., Aleffi I.F., Amouroux J.M. et al. Macroecology of the European soft sediment benthos: insights from the MacroBen database // Mar. Ecol. Prog. Ser. 2009. V. 382. P. 287-296.

Weber H.E., Moravec J., Theurillat J.-P. International code of phytosociological nomenclature. 3rd ed. // J. Vegetation Sci. 2000. V. 11. P. 739-768.

Weiher E., Keddy P.A. (eds.). Ecological Assembly Rules: Perspectives, Advances, Retreats. Cambridge: Cambridge Univ. Press, 1999. 393 p.

Weiner J., Solbrig O.T. The meaning and measurement of size hierarchies in plant populations // Oecologia. 1984. V. 61. P. 334-336.

Westhoff V., van der Maarel E. The Braun-Blanquet approach // Handbook of Vegetation Science. Pt. 5. Ordination and Classification of Communities. The Hague: Dr. W. Junk B.V. Publ., 1973. P. 617-737.

Whittaker R.H. Dominance and diversity in land plant communities // Science. 1965. V. 147. P. 250-260.

Whittaker R.H. Gradient analysis of vegetation // Biol. Rev. 1967. V. 42. P. 207-264.

Whittaker R.H. Evolution and measurement of species diversity // Taxon. 1972. V. 21. V. 213-251. [Пер. с англ.: Уиттекер Р. Эволюция и измерение видового разнообразия // Антология экологии / Сост. и коммент. Г.С. Розенберга. Тольятти: ИЭВБ РАН, 2004. С. 331-378].

Williams B.K., Nicholls J.D., Conroy M.J. Analysis and Management of Animal Populations: Modeling, Estimation, and Decision Making. N.Y.: Acad. Press, 2001. 817 p.

Williams E.J. Use of scores for the analysis of association in contingency tables // Biometrika. 1952. V. 39. P. 274-289.

Wilson J.B. Methods for detecting nonrandomness in species co-occurrences: a contribution // Oecologia. 1987. V. 73. P. 579-582.

Wilson J.B., Gitay H., Stee J.B., King W.M. Relative abundance distributions in plant communities: effects of species richness and of spatial scale // J. Vegetation Sci. 1998. V. 9. P. 213-220.

Wilson M.V. A statistical test of the accuracy and consistency of ordinations // Ecology. 1981. V. 62. P. 8-12.

Wilson M.V., Mohler C. L. Measuring compositional change along gradients // Vegetatio. 1983. V. 54. P. 129-141.

Wilson M.V., Shmida A. Measuring beta diversity with presence-absence data // J. Ecol. 1984. V. 72. P. 1055-1064.

Winemiller K.O., Pianka E.R. Organization in natural assemblages of desert lizards and tropical fishes // Ecol. Monogr. 1990. V. 60. P. 27-55.

Wolda H. Similarity indices, sample size and diversity // Oecologia. 1981. V. 50. P. 296-302.

Wright D.H., Patterson B., Mikkelson G. et al. A comparative analysis of nested subset patterns of species composition // Oecologia. 1998. V. 113. P. 1-20·

Wright D.H., Reeves J.H. On the meaning and measurement of nestedness of species assemblages // Oecologia. 1992. V. 92. P. 416-428.

Yakimov V.N., Bossuyt B., Iudin D.I., Gelashviliy D.B. Multifractal diversity-area relationship at small scales in dune slack plant communities // Oikos. 2008. V. 117. P. 33-39.

Yachi S., Loreau M. Biodiversity and ecosystem productivity in a fluctuating environment: The insurance hypothesis // Proc. Natl. Acad. Sci. 1999. V. 96. P. 1463-1468.

Yodzis P. The stability of real ecosystems // Nature. 1981. V. 289. P. 674-676.

Yodzis P. Competition, mortality, and community structure // Community Ecology. N.Y.: Harper and Row, 1986. P. 480-491.

Zadeh L.A. Fuzzy sets // Information and Control. 1965. V. 8. P. 338-353.

Zahl S. Jackknifing an index of diversity // Ecology. 1977. V. 58. P. 907-913.

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

Zhang J.T. A combination of fuzzy set ordination with detrended correspondence analysis – one-way to combine multi-environmental variables with vegetation data // Vegetatio. 1994. V. 115. P. 115-121.

Zipf J.K. Human Behaviour and the Principle of Least Effort. Cambridge: Addison-Wesley Press, 1949. 574 p.

Zobel M. The relative role of species pools in determining plant species richness: an alternative explanation of species coexistence // Trends Ecol. Evol. 1997. V. 12. P. 266-269.

Zobel M., van der Maarel E., Dupre C. Species pool the concept, its determination and significance for community restoration // Appl. Veg. Sci. 1998. V. 1. P. 55-66.

Zurlini G., Riitters K.H., Zaccarelli N., Petrosillo I. Patterns of disturbance at multiple scales in real and simulated landscapes // Landscape Ecol. 2007. V. 22. P. 705-721.

Шитиков Владимир Кириллович
Зинченко Татьяна Дмитриевна
Розенберг Геннадий Самуилович

МАКРОЭКОЛОГИЯ РЕЧНЫХ СООБЩЕСТВ: КОНЦЕПЦИИ, МЕТОДЫ, МОДЕЛИ

Технический редактор *О.Л. Носкова*

Издательство «Кассандра»
445061, г. Тольятти, ул. Индустриальная, д. 7
Тел/факс: (8482) 570-004

Подписано в печать с оригинал макета 28.06.2011 г.
Формат 60x84 1/16 Бумага офсетная. Печать офсетная.
Усл. печ. л. 21,5
Тираж 300 экз. Заказ № 111

Отпечатано в типографии ООО «Кассандра»

СПИСОК ИСПОЛЬЗОВАННЫХ ПРОГРАММНЫХ СРЕДСТВ

EstimateS – статистическая оценка богатства видов и распределения видов [R. Colwell, 1994-2005, http://viceroy.eeb.uconn.edu/estimates].

SPADE – прогнозирование числа видов и видового разнообразия [A. Chao, T. Shen, 2003-2005, http://chao.stat.nthu.edu.tw].

SDR-IV – видовое разнообразие и богатство [R. Seaby M., P. Henderson, www.pisces–conservation.com**].**

EcoSim – нуль-модели для анализа экологических сообществ [N. Gotelli, G. Entsminger, http://homepages.together.net/~gentsmin/ecosim.htm].

PAST – набор разнообразных статистических функций для анализа палеонтологических данных [Ø. Hammer, D. Harper, P. Ryan, http://folk.uio.no/ohammer/past].

PC-ORD – многомерный анализ экологических данных [B. McCune, M. Mefford, http://www.pcord.com].

PARTITION – аддитивное разложение на компоненты видового разнообразия [J. Veech, T. Crist, http://www.users.muohio.edu/cristto/partition.htm].

R - объектно-ориентированная статистическая среда [русскоязычная документация и методические материалы А.Б. Шипунова, http://herba.msu.ru/shipunov/software/r/r-ru.htm], в том числе:

vegan – пакет анализа экологии сообществ [J. Oksanen et al., http://vegan.r-forge.r-project.org];

labdsv – пакет ординации и многомерного анализа в экологии [D. Roberts, 2005, http://ecology.msu.montana.edu/labdsv/R];

fso – пакет ординации нечетких множеств [D. Roberts, http://www.nku.edu/%7Eboycer/fso/].

NTC (Nestedness Temperature Calculator) – расчет матричной температуры вложенности [W. Atmar, B. Patterson, 1995, http://aics-research.com/nestedness/tempcalc.html].

Комплекс FORTRAN-программ для макроэкологических исследований [W. Ulrich, www.uni.torun.pl/~ulrichw], в том числе:

Nestedness – расчет показателей матричной вложенности;

PatchOccupancy – оценка критериев пространственного распределения видов;

CoOccurrence – анализ совместного распространения видов;

RAD – подбор моделей рангового распределения обилия видов.

WinTWINS – двухсторонний анализ видов-индикаторов TWINSPAN для Windows [M. Hill, P. Šmilauer, http://www.canodraw.com/wintwins.htm].

CANOCO – канонический анализ и ординация сообществ [C. ter Braak, P. Šmilauer, 1998, http://www.microcomputerpower.com, http://www.pri.wur.nl/uk/products/canoco].

JUICE – управление, анализ и классификация фитоценологических данных [L. Tichy, J. Holt, www.sci.muni.cz/botany/juice.htm].

MULTIV – многомерный анализ, бутстреп и рандомизационные тесты [V. Pillar, http://ecoqua.ecologia.ufrgs.br/ecoqua/software.html].

Multi_mantel – многомерная матричная регрессия на основе процедуры Мантеля и проверка гипотез [L. Revell, http://anolis.oeb.harvard.edu/~liam/programs].

HOF – анализ экологических градиентов по модели Хьюсмана-Олфа-Фреско [J. Oksanen, http://cc.oulu.fi/~jarioksa/softhelp/hofprog.html].

SAM – пространственный анализ в макроэкологии [T. Rangel et al., http://purl.oclc.org/sam].

NesAnova – иерархический многомерный дисперсионный анализ при каноническом анализе [P. Legendre, http://www.bio.umontreal.ca/casgrain/en/labo/nesanova.html].

Barrier – выделение локальных географических областей по методу Монмоньера [F. Manni, E. Guerard, http://www.mnhn.fr/mnhn/ecoanthropologie/software/barrier.html].