

ОЦЕНКА БИОРАЗНООБРАЗИЯ: ПОПЫТКА ФОРМАЛЬНОГО ОБОБЩЕНИЯ

1. Общий подход к оценке биологического разнообразия

1.1. Развитие концепций и определение основных понятий

Понятие «биологическое разнообразие» за сравнительно короткий отрезок времени получило расширенное многоуровневое толкование. Собственно его биологический смысл раскрывается через представления о *внутривидовом*, *видовом* и *надвидовом (цено-тическом) разнообразии* жизни. Однако, в добавление к этому, сначала деятели охраны природы, а затем и ученые стали говорить об *экосистемном* и *ландшафтном разнообразии* как объектах сохранения, а соответственно, изучения и выделения в природе. Традиционные положения биогеографической науки также позволяют трактовать ее как науку о географических закономерностях формирования биоразнообразия. Когда на Конференции ООН по окружающей среде и развитию была принята «Конвенция о биологическом разнообразии» [3344], к которой присоединилось большинство стран на планете, сложнейшая и многообразнейшая проблема биоразнообразия приобрела еще и *политическое* звучание.

Само словосочетание «биологическое разнообразие» впервые, по-видимому, применил Г. Бэйтс в 1892 г. в работе «Натуралист на Амазонке» [1469], когда описывал свои впечатления от встречи около 700 разных видов бабочек за время часовой экскурсии. Однако основные научные концепции биоразнообразия были сформулированы лишь в середине XX в. Используя математический аппарат «статистики разнообразия», Г. Хатчинсон [3680, 3681] показал эффективность его применения в полевой экологии. Р. Маргалев [3881, 3882] на примере гидробионтов и практически одновременно Р. Мак-Артур [3863] на орнитологических объектах предложили методы количественной оценки разнообразия с теоретико-информационных позиций. Впоследствии Р. Мак-Артуром и Е. Уилсоном [3866] эти представления были развиты в виде основных положений географической экологии и иерархической биоорганизации ландшафта. Наконец, Р. Уиттекер [2718] структурировал уровни экосистемного разнообразия и исследовал зависимости биоразнообразия от факторов окружающей среды.

Хотя «разнообразие» рассматривается как основной параметр, характеризующий состояние биосферных систем, терминологические определения, представленные в литературе, имеют размытый, полуинтуитивный характер.

- «Разнообразие – это понятие, которое имеет отношение к размаху изменчивости или различий между некоторыми множествами или группами объектов» [1469];
- «Разнообразие представляет собой совокупности типов различий объектов мира (универсума) любого пространства (территории, акватории, планеты), которое выделяется на основе выбранной меры... Эта мера... в общих чертах векторизирует в бесконечность» [2987].

Биологическое разнообразие (biological diversity) – один из немногих общебиологических терминов, формулировка которого закреплена на уровне международных соглашений:

- «"Биологическое разнообразие" означает вариабельность живых организмов из всех источников, включая наземные, морские и другие водные экосистемы и экологические комплексы, частью которых они являются; это понятие включает в себя разнообразие в рамках вида, между видами и экосистемами» [3344, 3516].

Авторы серии прекрасных учебных материалов в рамках проекта Глобального экологического фонда (ГЭФ) по сохранению биоразнообразия России [672] считают, что:

- «Термин "биоразнообразие" обычно используется для описания числа, разновидностей и изменчивости живых организмов. В широком смысле этот термин охватывает множество различных параметров и является синонимом понятия "жизнь на Земле"».

И.Г. Емельянов [960] приводит более полутора десятков определений типа «разнообразие – это...», которые в целом мало отличаются от приведенных выше и основываются на следующих ключевых словах: «число биологических объектов каждого типа», «число типов объектов», «уровень отличий между объектами разного типа». Отметим, что ни в одном из определений не постулируется требование «выравниваемости» числа объектов, т.е. статистической «гомогенности», оцениваемой, в частности, информационным индексом разнообразия Шеннона.

Дать более точное определение разнообразия вряд ли возможно, поскольку попытка дальнейшей конкретизации не будет учитывать многоуровневый и целостный характер биологических явлений. По современным представлениям, можно различать (см., например, [1936, 2148, 2284, 3516]) несколько уровней организации жизни (молекулярный, генетический, клеточный, организменный, популяционный, экосистемный, биосферный), каждый из которых обладает свойственным ему специфическим биоразнообразием. Действительно, в реальной природе существует системно обусловленная иерархическая организация: в частности, геном не существует вне организма, организм – вне вида, вид – вне экосистемы, а экосистема – вне географической среды. И при этом возникает вопрос: а какие конкретно компоненты биоты следует подсчитывать, чтобы оценить ту или иную форму разнообразия? Поэтому в научном плане термин «биоразнообразие» достаточно локально относят к таким фундаментальным понятиям, как гены, особи и экосистемы, которые соответствуют трем фундаментальным, иерархически зависимым уровням организации биосферы нашей планеты.

1.2. Уровни биоразнообразия

Генетическое разнообразие определяется множественностью вариаций последовательностей комплиментарных нуклеотидов в нуклеиновых кислотах, составляющих генетический код. Оно выражается через уровни гетерозиготности, полиморфизма и другой меры генотипической изменчивости, которая вызвана адаптационной необходимостью природных популяций. Уточнению понятия «генетическое разнообразие» посвящена обширная литература [923, 2953, 3359, 3396, 4435, 4436], описывающая детально проработанные математические модели, и здесь мы не будем вдаваться в тонкости этой серьезной науки. Однако, как отмечает С. Райт (цит. по: [672]), «важнейшее заключение состоит в том, что эволюционный процесс зависит от постоянно смещающегося баланса между факторами стабильности и изменчивости и что наиболее благоприятное условие для этого – наличие тонко подразделенной структуры, в которой изоляция и перекрестная коммуникация поддерживаются в соответствующем равновесии».

Генетическое и фенетическое разнообразие является составной частью следующего иерархического слоя учета разнообразия *на уровне отдельных особей*. На этом же уровне (вернее, на уровне совокупности организмов) учитывается также разнообразие, охватывающее возрастную или групповую изменчивость внутривидовых групп, морфологические особенности и проч. В то же время в теоретических концепциях особь в качестве самостоятельной учетной оценки биоразнообразия обычно не рассматривается.

Единицы учета биоразнообразия, по Б.А. Юрцеву [3092], должны обладать автономным жизнеобеспечением, способностью к адаптивной эволюции и неограниченно длительному самоподдержанию на фоне стабильной или умеренно флуктуирующей среды, восстановлению при нарушениях и т.д. Перечисленным требованиям в качестве важных единиц учета и сохранения биоразнообразия удовлетворяют *виды*, а применительно к ограниченным территориям – представляющие вид местные *популяции*. Поэтому термин «биоразнообразие» часто рассматривают как синоним «*видового разнообразия*», в частности «богатства видов», которое есть число видов в определенном местообитании или био-

топе. Уточним, что мы понимаем *местообитание* как некоторый обособленный участок территории с совокупностью условий среды, обеспечивающей на некотором интервале времени устойчивое существование рассматриваемого организма, популяции или сообщества популяций. Линейные размеры местообитания могут варьироваться в весьма широких пределах и зависят от пространственной однородности факторов среды и степени мозаичности компонентов биосферы: для птиц, например, это может быть обширный участок леса, а для зообентоса – фрагмент донной поверхности, соизмеримый с шириной захвата скребка.

В некоторых публикациях высказывается ряд сомнений в обоснованности построения оценок разнообразия по видовому принципу. Ю. Одум [1936] считает, что виды не всегда служат лучшей единицей для оценки разнообразия, так как разные стадии жизненного цикла или разные жизненные формы одного и того же вида часто занимают разные местообитания и экологические ниши и вносят свой вклад в разнообразие (личинка и имаго чешуекрылых, стадии амфибий). Однако нельзя не признать, что именно виды имеют наиболее четкий адрес в той или иной таксономической системе, а имеющиеся базы компьютерной информации об организмах, как правило, проиндексированы по видовому принципу. Сведения по биологии и экологии видов являются основой для выработки необходимых мер их охраны в природе и сохранении в культуре (*in situ* и *ex situ*). Охраняя вид *A*, в его лице мы охраняем как отдельных особей, так и представителей рода или семейства, или редкую жизненную форму. Таким образом, вид – наиболее непротиворечивая и удобная форма группировки элементов в таксоны.

Считается, что разнообразие территориально распределенных экологических сообществ качественно отличается от точечного разнообразия в отдельных местообитаниях, поскольку представляет коллективный отклик видов на различные условия окружающей среды. В связи с этим предпринимаются попытки оценить *экосистемное разнообразие*, которое связывают с такими важнейшими теоретическими конструкциями экологии, как многомерное пространство факторов, гиперповерхность плотности распределения вида в этом пространстве («экологическая ниша») и модели конкурентной борьбы популяций за ресурсы среды [2284].

Р. Уиттекер [2718] сформулировал систему терминов и их вербальных понятий (см. табл. 1), определяющую различные формы и типы экологического разнообразия с использованием греческого алфавита. В дальнейшем эта система классификации была дополнена и уточнена Ю.И. Черновым [2943] и рядом других авторов.

Таблица 1

Формы и типы разнообразия по Р. Уиттекеру [2943]

Инвентаризационное разнообразие	Дифференцирующее разнообразие
<i>Точечное альфа-разнообразие</i> – разнообразие в пределах пробной площади или местообитания в пределах сообщества	<i>Внутреннее бета-разнообразие</i> (мозаичное разнообразие, изменение между частями мозаичного сообщества)
<i>Альфа-разнообразие</i> (внутреннее разнообразие местообитания для описания, представляющего гомогенное сообщество)	<i>Бета-разнообразие</i> (разнообразие между различными сообществами вдоль градиента среды)
<i>Гамма-разнообразие</i> (для ландшафта или серии проб, включающей более чем один тип сообщества, конкретную флору или фауну)	<i>Дельта-разнообразие</i> (географическая дифференциация, изменение сообщества вдоль климатических градиентов или между географическими регионами)
<i>Эпсилон-разнообразие</i> (для биома, географического региона, включающего различные ландшафты)	<i>Омега-разнообразие</i> (разнообразие биомов в рамках эпсилон-разнообразия)

Однако, как писал Р. Мак-Артур [3865]: «...хотя полное разнообразие – это обширное число видов в некоторой филогенетической группе в достаточно обширной географической зоне, составленной из некоторого количества местообитаний, – было предметом многих спекуляций и табулирования данных, этот аспект видового разнообразия сегодня наименее четко истолкован». Из-за отсутствия конкретных и общепринятых методик оценки различных уровней разнообразия каждый из исследователей вкладывает в их понятия свой персональный смысл. Поэтому на практике за *экосистемное* (полное) разнообразие чаще всего принимается оценка частот встречаемости разных видов на территории или в биотопе, либо подсчет частных сумм показателей относительного обилия (совокупной численности или биомассы) на разных трофических уровнях или по различным таксономическим группам. Такое «экосистемное разнообразие», в сущности, не более чем без особых затей механистическое обобщение списка видовых компонентов по сумме местообитаний.

1.3. «Кирпичики» разнообразия

Вспомним слова В.И. Вернадского – «Наука начинается тогда, когда человек подходит к явлениям природы с числом и мерой» – и рассмотрим пути количественного описания биоразнообразия в величинах, которые можно сравнивать между собой. Для этого необходимо предварительно выполнить формализацию самого понятия «разнообразие».

А.А. Протасов к своей книге [2148] привел в качестве эпиграфа афоризм Л.Н. Гумилева «Спросим себя: что доступно непосредственному наблюдателю? Оказывается, что не предмет, а границы предмета». По существу затронутой темы уместен такой парафраз: «Спросим себя: что доступно непосредственному наблюдателю экосистемы? Оказывается, что не сама экосистема, а только *таблица измерений*, привязанная к датам и точкам отбора проб». Согласно математическим традициям представим, что каждой строке такой таблицы соответствует определенный объект или явление изучаемой системы, а в столбцах таблицы размещаются значения или метки признаков. В результате получается табл. 2 типа «объект-признак»:

Таблица 2

Объекты	Даты измерений	Признак 1	Признак 2	Признак 3	...	Признак s - 1	Признак s
Объект 1	Дата 1				...		
Объект 1	Дата 2				...		
Объект 2	Дата 3				...		
...	
Объект <i>r</i>	Дата <i>k</i>				...		

Используя таблицу данных, исследователь пытается выяснить, можно ли построить на множестве наблюдений сколько-нибудь разумную (естественную, полезную) систему отношений, после чего формирует и проверяет продуктивные теоретические гипотезы о внутренних механизмах экосистемных процессов [2547]. В ходе анализа данных ищется ответ на следующие традиционные вопросы:

- Одинаковы ли два изученных объектов или они как-то отличаются друг от друга?
- За счет каких признаков можно объяснить возможные отличия объектов?
- Как можно объединить отобранные объекты в группы?
- Изменение каких признаков приводит к систематическим причинно-следственным изменениям других?
- Как можно осуществить прогноз состояния или поведения тестируемого объекта?

В экологии слишком часто ставят телегу впереди лошади: придумывают гипотезы, механизмы, гипотетические конструкции (типа экологической ниши, биоценоза, «типич-

ного» местообитания), а потом, глядя на таблицы наблюдений, ищут подходящее подтверждение этим конструкциям.

В зависимости от задач исследования и тех учетных единиц, которые используются для оценки биоразнообразия, строки и столбцы табл. 2 могут иметь различное содержание.

Объектом-носителем *фенетического* (*генетического*) разнообразия является особь (экземпляр) со свойственным ему множеством элементарных признаков. Тогда в строках табл. 2 размещается счетное множество R особей организмов определенной филогенетической группы, которые выделены и описаны в изучаемом местообитании. Круг параметров S , учитываемых при построении многомерного пространства и представленных столбцами таблицы, может быть неограниченно широк и включать произвольные фенотипические признаки, возрастные и аллометрические показатели, особенности морфологии особей и т.д. Определяющими для оценки разнообразия являются различия между особями; причем их абсолютная морфологическая сложность значения не имеет.

В пространстве признаков S фенотипическая изменчивость местообитания отображается облаком точек, каждая из которых соответствует одной особи или группе экземпляров с общей характеристикой, как это обычно бывает в систематических описаниях видов. Тогда мерой биологического разнообразия местообитания на уровне особи может служить сумма всех таксономических расстояний между точками (в математике – мера Лебега). Простейший способ измерить расстояние – евклидова дистанция, которая, если в матрице признаков используются только величины 0 и 1, трансформируется в меру Хемминга – число признаков, различающих соответствующие объекты или группы экземпляров. Для r особей простейшая мера феноразнообразия – сумма чисел различающихся признаков для $r(r-1)/2$ попарных сравнений экземпляров (групп). В литературе по таксономии [1761, 4233] обсуждаются различные коэффициенты сходства, или таксономического расстояния, а также меры, учитывающие морфологическую сложность, многие из которых эквивалентны друг другу [2435, 2011]. В общем виде идея об использовании мер в пространстве признаков для измерения фенетического разнообразия высказывается с завидным постоянством, но так и не нашла широкого применения из-за сложностей практической реализации [3460], в первую очередь из-за трудоемкости информационного описания на уровне особи.

Принципиально иной характер принимает табл. 2 при оценке *видового* разнообразия. Пусть исследователь в ходе наблюдения для некоторого местообитания в фиксированный момент времени зафиксировал наличие счетного множества P особей, каждую из которых по определенному правилу можно отнести к одному из S таксонов (групп, классов), причем все особи внутри одного таксона считаются совершенно идентичными. Данные о каждом наблюдении записываются последовательно в строки табл. 2, столбцами которой являются метки таксонов (видов), а тело таблицы составляют показатели таксономического обилия p_{ij} (например, численность особей каждого вида). В простейшем случае видовое разнообразие рассчитывается отдельно для каждой строки, причем единственная информация, используемая для оценки такого пространственно-временного среза, заключается в том, насколько велика мощность множества P и как его элементы распределяются по S таксонам. Таким образом, каков бы не был алгоритм расчета видового разнообразия, обоснованность полученной оценки напрямую зависит от обоснованности выбора системы группировки.

Экологическое сообщество феноменологически принято рассматривать как *таксоценоз*, т.е. как «систему разнообразно взаимодействующих дифференцированных по нишам популяций» [2718, с. 89]. Иными словами, предполагается естественная структурированность выделенного множества экземпляров, задающая определенные рамки его разбиения по группам. Такую группировку организмов можно выполнить как по *филетическому* (группировка по родству), так и по *типологическому* принципу. Филетическая систематика может подразделяться на иерархические уровни с серией подуровней: подвид, вид, род, семейство и т.д., вплоть до царства. Нефилетические, или типологические группировки

осуществляются по тем или иным категориям признаков, не сводимых к родству. Примером могут служить жизненные формы, стратегии жизни, ценотипы, типы метаболизма, сукцессионный статус видов (т.е. место в сукцессионных рядах или системах) и т.д.

Структуры филетического и типологического разбиений дополняют друг друга. Описание таксономического разнообразия (филума или биоты некоторой территории), обычно представляемое длинными систематическими списками, как правило, дополняется характеристикой каждого таксона по комплексу типологических признаков. Сопряженность таксономического и типологического описаний ложится в основу организации баз и банков типологической информации об организмах. Однако информация о любом типе разнообразия поступает все же через виды как фундаментальные единицы биоразнообразия.

В рамках последующего изложения для нас важны следующие формальные отправные точки:

- биотические объекты, выделенные в исследуемом местообитании, группируются в $S(t)$ таксонов по некоторому филетическому или типологическому принципу, причем их число в каждый момент времени t определяется балансом появления новых групп особей $S^+(t)$ и исчезновения старых $S^-(t)$ за все время до момента t ;
- для каждого i -го таксона ($i = 1, 2, \dots, S$) может быть оценена величина показателя абсолютного P_i и относительного таксономического обилия

$$p_i = P_i / \sum_{i=1}^S P_i ;$$

- пространственно-временной срез биоразнообразия в каждой строке табл. 2 может характеризоваться эмпирическим распределением показателя обилия P_i (или p_i) по таксонам, которое имеет, как правило, неравномерный характер и может быть аппроксимировано подходящим аналитическим выражением.

Выделенные отправные точки можно прокомментировать следующим образом. При построении таксоценозов в подавляющем большинстве случаев используется филетическая группировка по видам, т.е. число таксонов S эквивалентно числу видов, встретившихся в местообитании. Это не является догмой, а лишь устоявшейся практикой, и вполне привлекательной альтернативой является оценка разнообразия, скажем, по типу питания. Дополнительным аргументом в пользу типологических группировок является полное игнорирование отношений сходства и (или) различия между таксонами в видовой классификации, поскольку в нашей «модели» каждый таксон занимает строго изолированную самостоятельную позицию без учета его места в реальной систематической иерархии. Предположим, например, что в одном местообитании встретилось только 10 морфологически очень близких между собой видов, принадлежащих к одной трибе и которые, при определенных условиях, легко принять эквивалентными. А в другом местообитании было найдено 9 видов, но уже каждый из них относится к различным семействам или выше. С формально видовых позиций первое местообитание считается более разнообразным, чем второе, хотя по сути явления в первом случае налицо имеет место очевидный моноценоз.

Другой вопрос состоит в том, какую из учетных характеристик сообщества в местообитании принять за оценку обилия P_i . Чаще всего в теоретических исследованиях, связанных с моделями разнообразия, оперируют численностью, т.е. $P_i \equiv N_i$, где N_i – число особей i -го вида (таксона в нашей терминологии). Однако для некоторых биологических сообществ натуральную численность определить просто невозможно, а, кроме того, численность является удобным, но не слишком адекватным показателем для оценки потребности вида в ресурсах. В частности, Р. Уиттекер [4398] отмечает: «Величина потребленного ресурса выражается продуктивностью вида... Продуктивность не относится к числу легко определяемых показателей, и поэтому мы можем сравнивать виды по некоторым другим оценкам "важности" (для популяций животных – по плотности или биомассе, для растений – по покрытию, биомассе, встречаемости, площади оснований, плотности)». Как пример, интересна работа Р. Мак-Артура [3865], который показал, что разнообразие про-

филя вертикального распределения листвы является наиболее определяющим фактором видового разнообразия птиц. При этом «распределение листьев по высоте древостоя» P_i для i -го яруса рассчитывалась как доля плотности листвы в данном диапазоне высоты над уровнем земли к общей плотности лиственной поверхности.

Если для оценки P_i вида в учетной практике изучаемой филогенетической группы применимы два одинаково важных, но различных по природе показателя, то, по нашему мнению, вполне целесообразно использовать комбинированные показатели обилия. Так в гидробиологии, где характеристика зообентоса задается комбинацией из 2-х параметров (B – биомассы и N – численности), широкое применение нашел так называемый индекс плотности Брочкой–Зенкевича $P_i = \sqrt{B_i N_i}$, в котором исходные величины берутся в абсолютных значениях [213, 3005].

Как показывает практика, весьма существенная вариация объема P_i кластеров изучаемого таксоценоза определяется, в основном, объективными экологическими причинами: эволюцией видов в условиях конкурентной борьбы за ресурсы среды и формированием наиболее устойчивой схемы структурной организации сообщества. Можно в качестве примера привести некоторые устоявшиеся соотношения численностей отдельных групп макрозообентоса в малых реках Самарской области [3005]: на 1 особь Coleoptera в среднем приходится 2 экз. Hemiptera, 3 экз. Bivalvia, 4 экз. Ephemeroptera, 8 экз. Chironomidae, 14 экз. Nematoda и 25 экз. Oligochaeta. Соотношение биомассы варьируется еще в большей степени: {1 : 3 : 10 : 50 : 400 : 2500 : 5000} для Coleoptera, Chironomidae, Oligochaeta, Bivalvia, Gastropoda, Unionidae и Dreissenidae соответственно. Разумеется, под влиянием таких факторов, как вариация условий среды конкретного биотопа, регулярные, периодические и неожиданные флуктуации и пр., эти соотношения могут существенно меняться. Однако их объективная составляющая, которая во многом определяется экологией, аллометрическими характеристиками особей отдельных видов, трофическими связями и др., остается стабильной. Немаловажным источником дисперсии объема таксонов являются субъективные ошибки определения видов (поскольку чаще всего каждый систематик уверенно работает с ограниченной группой организмов, лишь в той или иной мере согласовывая свои решения с другими специалистами).

Таким образом, удельные объемы таксонов p_1, p_2, \dots, p_s представляют собой эмпирическую совокупность независимых реализаций случайной величины ξ , характеризующуюся выборочной функцией распределения $F_s(p)$. Важно уточнить, что выборочному аналогу $F_s(p)$ можно подобрать наиболее подходящую аналитическую функцию распределения $F(\xi)$, однако любая аксиоматика, основанная на том, что это гипотетическое распределение является равномерным, не соответствует представлениям о способе существования реальных устойчивых сообществ организмов. Действительно, абсурдное предположение о равном соотношении показателей обилия вышеприведенных групп зообентоса означает для экосистемы не оптимум биоразнообразия, а признак грядущей экологической катастрофы...

1.4. Формальное определение «разнообразия»

Вернемся, однако, к уточнению понятия «разнообразия». Как пишет Ю. Одум [1936], видовое разнообразие складывается из двух компонентов:

- *видового богатства* или *плотности видов* (variety), которое характеризуется общим числом имеющихся видов;
- *выравненности* (evenness), основанной на относительном обилии или другом показателе значимости вида и положении его в структуре доминирования.

Тогда формальное определение может быть сделано следующим образом:
разнообразие есть количественный показатель экосистемы, основанный на ее предварительной декомпозиции на множество S характеристически обособленных групп и равный функционалу от двух характеристик такого разбиения – числа таксонов S и выборочной функции распределения $F_s(p)$ изучаемого показателя по этим группам.

$$R = \Psi [S, \Omega(F_s(p))], \quad (1)$$

где Ω – некоторая наиболее подходящая выборочная статистика.

В качестве конкретных дефиниций статистики Ω могут быть использованы следующие возможные классы математических функций:

- меры статистического разброса величины p_i , не основанные на предположениях о ее законе распределения (называемые непараметрическими индексами);
- критерии согласия, проверяющие гипотезу о соответствии выборочного распределения некоторой постулированной аналитической закономерности.

Статистический разброс удельных объемов таксонов p_1, p_2, \dots, p_s может быть оценен с использованием двух известных в теории вероятности формул:

$$D = \sum_{i=1}^s p_i^2 = S\sigma^2 + 1/S \quad \text{– как мера дисперсии } \sigma^2;$$

$$H = -\sum_{i=1}^s p_i \log p_i \quad \text{– как мера энтропии.}$$

Во многих работах (например, [2148]) анализируются гипотетические предположения о некотором внутрисистемном или теоретико-информационном смысле, приписываемым D и H , однако мы склонны считать, что в этом аспекте обе меры никакого специфического информационного содержания не несут. Они относятся к показателям, которые «просто измеряют гетерогенность выборки» [703], причем отличие между двумя расчетными формулами сводится лишь к тому, что вклад произведений $p_i p_i$ в результат суммирования в значительно большей мере определяется абсолютной величиной p_i , чем $p_i \log_2 p_i$.

Многочисленные формулы, основанные на различных модификациях приведенных двух мер и используемые для оценки биоразнообразия, обычно называют в экологии *индексами*. В следующем разделе будут рассмотрены только основные направления такого «индексотворчества», а пунктуальному читателю можно порекомендовать монографию А.П. Левича [1476], в которой подробно разбирается более 30 различных мер видового разнообразия.

2. Анализ индексов биологического разнообразия

2.1. Индексы, основанные на формуле дисперсии

Нетрудно заметить, что мера дисперсии D объема таксонов в точности совпадает с известным в экологии *индексом разнообразия* Е. Симпсона [1767, 4211]¹. Если использовать в качестве оценок обилия численность экземпляров ($n_i \equiv P_i$), то формулу для индекса Симпсона принято записывать как

$$D = \sum \left[\frac{n_i(n_i - 1)}{N(N - 1)} \right].$$

В этой форме записи можно принять во внимание, что знаменатель $N(N-1)/2$ означает число всех связей между парами особей независимо от их принадлежности к определенному таксону, а числитель $n_i(n_i-1)/2$ – число возможных взаимодействий между парами особей одного таксона. Тогда глубинный смысл индекса Симпсона можно объяснить [89], как отношение числа степеней свободы внутривидовых взаимодействий к общему числу степеней свободы внутренних элементов экосистемы, обеспечивающих ее единство и функционирование. Другие авторы [672] приписывают индексу Симпсона оценку вероят-

¹ Он же – индекс Херфиндаля–Хиршмана, который начиная с 1982 г. служит основным ориентиром при осуществлении антимонопольной политики в США; он же – индекс лексического разнообразия (Type Token Ratio), предложенный Дж. Юлом для оценки словарного запаса в математической лингвистике.

ности принадлежности к разным видам любых двух особей, случайно отобранных из неопределенно большого сообщества.

По мере увеличения D выравненность объема таксонов уменьшается. Поэтому индекс Симпсона часто используют в форме $(1 - D)$. Эта величина варьирует от 0 до 1 и носит название «вероятность межвидовых встреч» [4082].

Как и для любой другой статистики типа *дисперсии*, значение индекса Симпсона практически полностью зависит от удельного объема 1-2 наиболее обильных видов, так как при возведении в квадрат малых долей p_i получаются очень малые величины. Поскольку D к тому же и слабо зависит от числа таксонов S , то признается [672], что индекс Симпсона непригоден для оценки биоразнообразия или видового богатства, но является весьма чувствительным индикатором доминирования одного или нескольких видов.

Следует упомянуть еще один «клон» индекса Симпсона – «индекс полидоминантности» Вильямса (Williams W.; цит. по: [672]):

$$D^{-1} = 1/D = \sum_{i=1}^S 1/p_i^2$$

который «...многие авторы считают наилучшей мерой оценки типа распределения видовых обилий для случаев, когда число видов превышает 10» [672].

Мы уже упоминали о тесной функциональной связи между индексом Симпсона и дисперсией объема таксонов. Оценивать выравненность видового состава непосредственно с помощью суммы квадратов отклонений от среднего уровня предложил Р. Пит [4059]:

$$V_e = D - 1/S = \sum_{i=1}^S (p_i - \bar{p})^2, \text{ или с использованием численностей } V_e = \frac{\sum n_i^2 - (N^2/S)}{(1-1/S)(N-S)^2}.$$

Применение дисперсии V_e не имеет никаких преимуществ перед D , поскольку остается проблема гипертрофированного влияния на конечный результат вклада первых нескольких наиболее обильных таксонов. Однако следующим шагом Пита было использование предварительно прологарифмированных значений обилия и тогда среднеквадратичное отклонение вычисляется как $\sigma_e = \sqrt{\frac{1}{S} \sum_{i=1}^S (\log p_i - \log \bar{p})^2}$, где $\log \bar{p}$ – геометрическое среднее.

Переход к системе прологарифмированных значений обилия – здесь не просто арифметическая операция, а оператор преобразования эмпирических выборок к нормальному закону распределения, что создает предпосылки к подключению мощного математического аппарата параметрической статистики. К сожалению, эти идеи Р. Пита не нашли впоследствии должного развития.

2.2. Индексы, основанные на формуле энтропии

Оценка видовой устойчивости и разнообразия сообщества с использованием энтропии была впервые применена Р. Маргалевым [3881, 3882] и Р. Мак-Артуром [3863, 3864], в результате чего большое распространение и повсеместное признание в практической экологии получил индекс Шеннона H , иногда называемый *информационным индексом разнообразия К. Шеннона – У. Уивера* [4191]. Не располагая данными о роли У. Уивера в развитии концепций информационного анализа, мы вполне солидарны с А.А. Протасовым [2148], что по справедливости этот показатель в экологии должен называться индексом Маргалева–Мак-Артура.

В качестве некоторого обоснования связи понятия «биоразнообразие» с теоретико-информационными представлениями приведем без комментариев две цитаты:

- «Информационная интерпретация энтропийного индекса Шеннона состоит в том, что разнообразие трактуется как приходящееся на одну особь количество информации, заключенной в распределении по видам, особям или энергии по трофическим связям» [89, 3863];
- «Поддержание стабильности организма и биологических систем надорганизменных уровней интеграции определяется способностью запастись информацией из окружающей

среды в количестве не менее числа внешних нарушений, стремящихся вывести системы из области выживания... "носителем информации" могут быть показатели качественно-количественных соотношений отдельных компонентов» [960].

Остановимся на основных понятиях и расчетных формулах информационного анализа в его классическом общесистемном изложении. В теории информации [1201, 1506, 1550, 2991, 3099] в качестве меры количества переданных сигналов, оценки гомогенности распределения и возможности выбора применяется величина *энтропии*, которая представляет собой количественную меру степени неопределенности исхода случайного опыта, зависящую от распределения плотности вероятностей. Относительная или дифференциальная энтропия непрерывного распределения с функцией плотности вероятности $p(x)$ определяется как

$$H = - \int_{-\infty}^{\infty} p(x) \log_a p(x) dx.$$

Энтропия дискретного множества вероятностей (энтропия дискретной случайной величины, средняя собственная информация), соответствующая общей неопределенности всех S возможных исходов, определяется по формуле Шеннона

$$H = - \sum_{i=1}^s p_i \log_a p_i,$$

где p_i – вероятность появления дискретного события; величины $p_i, i=1, 2, \dots, s$, образуют множество вероятностей; $p_i \log_2 p_i$ – среднее количество информации, приносимое исходом i при многократном осуществлении испытания; a – основание логарифма, выбранное для оценки величины энтропии (обычно равное 2 или числу e).

Исходными данными для вычисления энтропии системы H в общем случае являются *дискретные* или *интервальные вариационные ряды*, т.е. для практических вычислений вероятности p_i заменяются частотами распределений.

При вычислении энтропии условились считать, что $0 \log_a 0 = 0$. Если в качестве основания логарифма a выбрано число 2, энтропия вычисляется в *битах*, если число e – в *нитах*. В последнем случае для вычисления энтропии в битах нужно разделить значение в нитах на $\ln 2$, что вытекает из известной формулы замены основания логарифмов.

В дискретном случае энтропия равна нулю, когда одна из вероятностей равна 1, а остальные нулю. Максимальная энтропия системы вычисляется по формуле Хартли

$$H_{max} = \log_2 S.$$

Таким образом, справедлива формула

$$0 \leq H \leq H_{max}.$$

Максимум энтропии соответствует наибольшей неопределенности или равенству вероятностей всех возможностей, т.е. при S равновероятных исходах с вероятностями $1/S$. В «практическом диапазоне» экологических разбиений H_{max} изменяется от 3,32 для 10 таксонов до 6,64 для $S = 100$.

Рассмотрим некоторые другие показатели, используемые в информационном анализе.

- Относительная энтропия. Для сравнения систем, различающихся по количеству элементов кода S , сопоставление энтропий будет не всегда корректным. В этих случаях применяется *относительная энтропия (коэффициент сжатия информации)*, вычисляемая как

$$E = H / H_{max}.$$

Относительная энтропия определяет относительную степень информационной эквивалентности системы по отношению к максимально возможной выравненности. Нетрудно заметить, что относительная энтропия в экологии соответствует предложенному Э. Пиелу [4080, 4085] *индексу выравненности экологических сообществ*: E изменяется от 0 до 1, причем $E = 1$ при равной заселенности всех таксонов.

- Избыточность. В теории информации соотношение между полным количеством информации, шумом и сохранившейся упорядоченностью системы выражается избыточностью, которая вычисляется по формуле

$$R = 1 - H / H_{\max}.$$

- Г. Тейлом [4316] в качестве показателя неравномерности распределения доходов был предложен *индекс энтропии* (энтропийный коэффициент), который более чувствительно реагирует на увеличение числа таксонов:

$$R_T = \log_a S + H, \quad \text{где } H - \text{энтропия.}$$

Дисперсию индекса Шеннона (V_H) для дискретного вариационного ряда (например, численностей видов) можно рассчитывать по формуле

$$V_H = \frac{\sum (p_i \log p_i)^2 - (\sum p_i \log p_i)^2}{N} + \frac{S-1}{2N^2}.$$

Если значения индекса Шеннона рассчитать для нескольких выборок (местообитаний), то полученное распределение величин, как правило, подчиняется нормальному закону, что дает возможность применять методы параметрической статистики, включая дисперсионный анализ. Для проверки значимости различий между выборочными совокупностями значений H Дж. Хатчинсон предложил использовать критерий Стьюдента:

$$t = \frac{H_1 - H_2}{(V_{H1} - V_{H2})^{1/2}}$$

с числом степеней свободы, определяемым по формуле

$$df = \frac{(V_{H1} + V_{H2})^2}{(V_{H1})^2 / N_1 + (V_{H2})^2 / N_2},$$

где N_1 и N_2 – общее число видов в двух выборках.

Несколько иной вычислительный аспект заложен в формулах для мер выравнимости, предложенных Р. Маргалефом и Л. Бриллюэном [3882] и основанных на подсчете разностей факториалов $\ln(n!)$. Как отмечает Р. Мак-Артур [3865], в соответствии с теоремой Стирлинга (до некоторого малого значения приближения ϵ), их запись идентична классической формуле энтропии H , и в качестве самостоятельных индексов они рассматриваться нами не будут.

2.3. Другие методы и формулы оценки выравнимости сообществ

Рассмотрим еще несколько показателей, позволяющих измерить «гетерогенность сообщества».

- Индекс Бергера–Паркера. Естественным и наиболее популярным методом оценки неравномерности распределения и выявления доминантов является деление таксонов на формальные квантильные (децильные, квинтильные, квартильные и др.) группы². Авторы индекса Бергера–Паркера не утруждали себя изощренным мудрствованием и предложили в качестве меры гетерогенности относительную значимость наиболее обильного вида: $d = p_{\max}$.

Главное достоинство индекса – простота вычисления, хотя «некоторые ученые считают этот индекс лучшей мерой разнообразия, потому что этот индекс независим от количества видов» [672] (нетрудно понять, что это далеко не так).

- Индекс Джини. Статистический показатель, разработанный итальянским экономистом К. Джини [2628], основан на анализе кумулятивной кривой накопленных долей численностей видов, которую во многих предметных областях называют *кривой Лоренца*.

² Вспомним замечательную фразу из книги «Физики шутят», отражающую основную социодемографическую закономерность: «*Двадцать процентов людей выпивают восемьдесят процентов пива*».

В прямоугольной системе координат кривая Лоренца является выпуклой вниз и проходит под диагональю прямоугольника с координатами: (0;0); (0;1); (S;1) и (S;0). Чем больше отклонение кривой Лоренца от диагонали (которая совпадает с линией абсолютно равномерного распределения долей), тем больше степень неравномерности объема таксонов.

Индекс Джини представляет собой отношение площади, заштрихованной на рис. 1 и ограниченной фактической кривой Лоренца и диагональю прямоугольника, к площади треугольника, ограниченного диагональю и осями абсцисс и ординат.

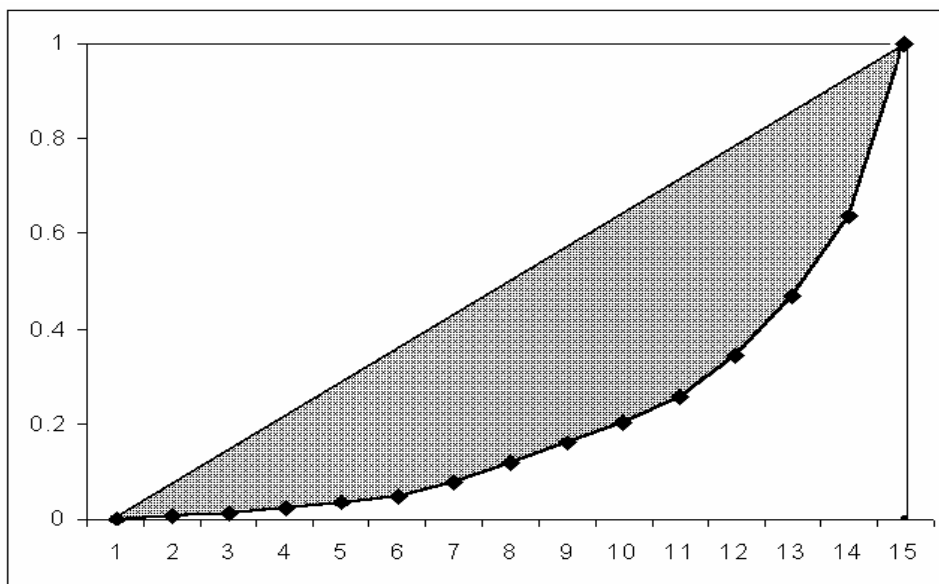


Рис. 1. Кумулятивная кривая накопленных долей численностей видов зообентоса в пробе от 10.07.1985 г. на ст. 15 реки Б. Черемшан (число видов $S = 15$, индекс Джини $G = 0,61$, энтропия $H = 3,0$)

Если представить выделенную площадь как сумму трапеций, то индекс Джини может быть также рассчитан аналитически по формуле

$$G = \frac{1}{2S-1} \sum_{i=1}^S \sum_{j=1}^S |p_i - p_j|.$$

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

- Мера разнообразия Макинтоша. В 1967 г. Р. Макинтош предложил рассматривать сообщество как точку в S -мерном гиперпространстве с координатами (n_1, n_2, \dots, n_s) . Тогда евклидово расстояние такого сообщества от начала координат можно использовать как меру его разнообразия:

$$U = \sqrt{\sum_{i=1}^S n_i^2} = N \sqrt{\sum_{i=1}^S p_i^2}.$$

Индекс Макинтоша U , как указано в [672], «сам по себе не является индексом доминирования, однако, используя его, можно рассчитать меру доминирования D , которая независима от объема выборки, и похожую на нее меру выравненности E »:

$$D_M = \frac{N-U}{N-\sqrt{N}}; \quad E_M = \frac{N-U}{N-N/\sqrt{S}} = \frac{1-\sqrt{\sum_{i=1}^S p_i^2}}{1-1/\sqrt{S}} = \frac{1-\sqrt{D}}{1-1/\sqrt{S}},$$

где D – индекс Симпсона.

2.4. Сравнительный анализ индексов разнообразия

Трудности в применении показателей разнообразия и оценке их качества заключаются в сложном и комплексном характере самой интерпретируемой величины, а также в отсутствии какой-либо объективной шкалы отсчета разнообразия, независимой от концепции, принятой исследователем на основе его субъективных (точнее, интуитивных) представлений.

В экологии не существует таких объектов и не изобретено таких «линеек», совмещение которых позволило бы путем считывания чисел со шкалы определить такие важнейшие факторы, как объем валовой продукции экосистемы, ее «биоценное качество» или темпы сукцессионных изменений. Поэтому косвенные или производные показатели, трактуемые впоследствии как неотъемлемые свойства экосистем, часто определяются путем многофакторного исчисления, алгоритм которого задается операциональной схемой (т.е. задан исчерпывающий перечень операций, которые необходимо провести, чтобы оценить данную величину). В экологии разработка таких схем – это сложный и неоднозначный аналитический процесс творческого характера. Поэтому исследователи отдают предпочтение тому или иному теоретическому предположению, следуя личной симпатии или традиции, установившейся в работе коллег. В полной мере это относится к понятию «разнообразия системы».

Ю.А. Песенко [2011] провел анализ функций разнообразия, не связанных с определенными моделями распределения видов по обилию, используя приведенные выше непараметрические критерии, и сформулировал основные требования к свойствам вычисляемых мер.

«... мера разнообразия должна быть

- количественно определена для любой выборки, содержащей хотя бы одну особь ($N > 0$; $S > 0$); разнообразие для выборки, не содержащей ни одной особи ($S = 0$), не может быть определено;
- непрерывна при $S > 0$; это подразумевает отсутствие накопленных эффектов на оценку разнообразия при измерении обилий видов, изменения долей которых в коллекции могут быть сколь угодно малы;
- неотрицательна в области своего определения при $S > 0$;
- функционально не связана с объемом выборки (не имеется в виду независимость от числа таксонов);
- возрастающей при фиксированном значении S и убывании невыравненности обилий;
- возрастающей при фиксированной невыравненности и увеличении S ».

Ю.А. Песенко проверил 6 различных показателей и пришел к выводу, что среди них только индекс полидоминантности D^{-1} удовлетворяет всем шести критериям. Мера Макинтоша E_M и индекс Симпсона D при добавлении новых видов в выборку слабо изменяются. У вероятности межвидовых встреч $(1-D)$ и индекса Шеннона H число видов является более важным фактором при небольшом числе видов (менее 10), а при увеличении числа видов возрастает роль *выравненности* (см. рис. 2).

Г. Ратледж (цит. по: [672]) на полгода позже Песенко также доказал предпочтительность меры Макинтоша E_M , но его требования к показателю разнообразия были менее строги. По мнению Ратледжа, индекс должен:

- равняться числу видов при их полной выравниваемости по обилию;
- быть меньше S , если обилия видов не выравнены;
- быть непрерывным при любых значениях p , ($0 < p_i < 1$);
- позволять разделить разнообразие сообщества на компоненты (разнообразие внутри местообитаний и разнообразие между местообитаниями).

Подробному сравнительному анализу индексов видового разнообразия посвящена также большая часть монографии Э. Мэгарран [1837] и работа В.Г. Терещенко с соавторами [2648].

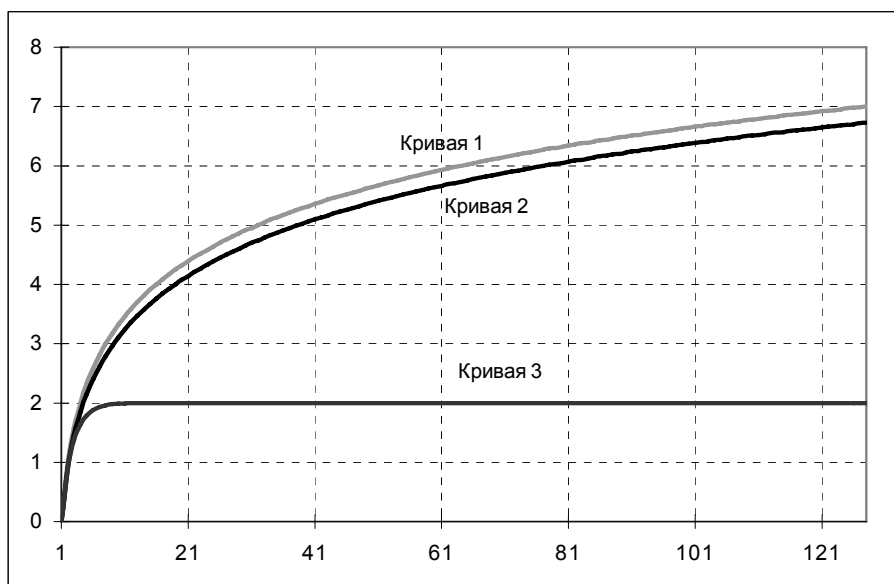


Рис. 2. Зависимость индекса Шеннона H от числа видов S :

кривая 1 – график максимально возможных значений индекса (равномерное распределение численности по таксонам $\{1, 1, 1, 1, \dots\}$); кривая 2 – умеренное доминирование (заселенность таксонов в арифметической прогрессии $\{1, 2, 3, 4, \dots\}$); кривая 3 – сильное доминирование (заселенность таксонов в геометрической прогрессии $\{1, 2, 4, 8, \dots\}$)

Выполним расчет и сопоставление основных статистических характеристик описанных индексов разнообразия для 497 проб макрозообентоса в малых реках Самарской области по данным экспедиционных наблюдений ИЭВБ РАН в 1987-2004 гг. (см. пример для 6 вполне типичных измерений в табл. 3).

Таблица 3

Значения основных индексов разнообразия сообществ макрозообентоса для 6 произвольных станций наблюдения и средние значения для массива из 497 проб

№ пп - № станции	1 - 08	2 - 03п	3 - 04	4 - 11	5 - 12	6 - 02	Среднее арифметическое	Стандартное отклонение
Водоем	Сок	Чапаевка	Сок	Маза	Сок	Ток		
Дата	30 июля 99	24 мая 90	14 июля 98	22 июня 89	26 июня 91	7 июля 94		
Число видов S	22	14	30	4	15	10	11,39	6,63
H Шеннона	4,03	3,39	2,83	2	0,752	0,409	2,256	0,79
R_T Тейла	8,49	7,19	7,74	4	4,66	3,73	5,49	1,63
E Пиелу	0,903	0,889	0,577	1	0,192	0,123	0,7	0,17
D Симпсона	0,08	0,114	0,24	0,25	0,822	0,9	0,327	0,177
d Бергера-Паркера	0,172	0,196	0,434	0,25	0,906	0,949	0,465	0,185
G Джини	0,422	0,431	0,818	0	0,912	0,92	0,579	0,182
E_M Макинтоша	0,911	0,904	0,624	1	0,126	0,074	0,679	0,191

Примечание: Жирным шрифтом отмечены значения с максимальным разнообразием; курсивом – с минимальным.

Результаты расчетов показывают, что различные индексы в целом достаточно однозначно оценивают выравненность сообществ. Основные различия сводятся в оценке разнообразия на р. Маза, где все 4 вида имеют одинаковую численность. Понятно, что индексы Пиелу, Джини и Макинтоша, оценивающие «чистую выравненность», дали здесь максимальную оценку разнообразия, несмотря на крайне бедное видовое богатство. Имеет

смысл обратить внимание на некоторую неоднозначность в трактовке разнообразия двух близких индексов Шеннона и Тейла. По нашему мнению, последний имеет серьезное преимущество в адекватности оценок, более гармонично связывая оба фактора биоразнообразия: видовое богатство и выравненность (эквитабельность).

Приведем для полноты картины матрицу коэффициентов парной корреляции перечисленных индексов (табл. 4). Рассматривая, в частности, тесноту связи индексов с числом видов (столбец 1 матрицы), можно выделить три группы оценок разнообразия:

- с сильной или отчетливой зависимостью от S (H Шеннона, R_T Тейла);
- индексы, статистически не зависящие от S (E Пиелу, E_M Макинтоша);
- остальные индексы, где эта связь прослеживается достаточно вяло.

Можно также выделить пары индексов, которые практически дублируют друг друга: D Симпсона и d Бергера-Паркера, E Пиелу и E_M Макинтоша.

Таблица 4

Матрица парных корреляций основных индексов разнообразия сообществ макрозообентоса, рассчитанная по массиву наблюдений из 497 проб

Число видов S	1							
H Шеннона	0,717	1						
R_T Тейла	0,882	0,93	1					
E Пиелу	-0,092	0,512	0,189	1				
D Симпсона	-0,487	-0,925	-0,782	-0,711	1			
d Бергера-Паркера	-0,406	-0,865	-0,686	-0,751	0,958	1		
G Джини	0,482	-0,031	0,308	-0,826	0,243	0,366	1	
E_M Макинтоша	0,038	0,606	0,313	0,972	-0,807	-0,86	-0,738	1
	S число видов	H Шенно- на	R_T Тейла	E Пиелу	D Симп- сона	d Бергера- Паркера	G Джини	E_M Макин- тоша

Примечание: Жирным шрифтом отмечены коэффициенты корреляции больше 0,5.

По результатам расчетов сформулируем два, в некотором смысле, провокационных вопроса:

- Является ли статистическая гомогенность таксоценоза (т.е. равенство видовых обилий) показателем высокого «качества» экосистемы?
- Являются ли тождественными численные выражения понятий статистической выравненности (эквитабельности) и разнообразия?

Высокую выравненность принято считать эквивалентной высокому разнообразию. «Выравненность – это единственный серьезный показатель структуры сообщества», – так считают авторы [672]. Действительно, многочисленные данные наблюдений подтверждают, что в богатых сообществах эквитабельность куда выше, чем в деградирующих. Структурный перекоп и «отрыв» лидеров-доминантов – совершенно реальный признак ухудшения качества среды, не нуждающийся в дополнительных аргументах. Однако оценка разнообразия только простым подсчетом эквитабельности мало информативна, так как ни одно сообщество не состоит из видов равной численности. Из общего числа видов какого-либо трофического уровня или сообщества в целом обычно лишь немногие бывают доминирующими, т.е. имеют значительную численность (большую биомассу, продуктивность или другие показатели), подавляющая же часть относится к редким видам (т.е. имеет низкие показатели «значимости»). Таким образом, большинство видов в сообществе малочисленны, численности других умеренны и лишь немногие обильны. А как было показано выше, и индекс D Симпсона, и индекс H Шеннона, и, тем более, индекс d Бергера-Паркера длинные «хвосты» небольших вероятностей просто-напросто игнорируют.

Меры выравненности приобрели «знаковую» популярность, на наш взгляд, по двум причинам:

- как некий «идеологический» антипод ситуациям гипертрофированного доминирования одного из видов, которые принято считать признаком «недостаточно качественной» неустойчивой экосистемы;
- эмпирическую кривую «важности видов» без каких-либо трудностей вычислительного характера можно сравнить только с равномерным распределением, характеризующимся на том же графике прямой, параллельной оси абсцисс.

Действительно, как будет показано ниже, для сопоставления реального распределения с той или иной аналитической моделью разнообразия («разломанного стержня» Мак-Артура, геометрических или логарифмических серий и пр.), более подходящей в качестве «стандарта разнообразия» экосистемы, чем гипотеза равномерного распределения, необходимо выполнить более серьезные вычисления, требующие наличия специального программного обеспечения компьютеров.

При использовании индексов *D* Симпсона, *E* Пиелу, *G* Джини и *H* Шеннона (при $S > 30$), основанных на дисперсии или энтропии обилия таксонов, следует отчетливо представлять, что эти меры оценивают только уровень равномерности распределения относительных долей, не зависят от абсолютных значений численности *N* и мало чувствительны к изменению видового богатства *S*. Следовательно, для гипотетического сообщества, где 10 000 видов делят ресурсы среды между собой на 10 000 равных долей и для двухкомпонентного сообщества, где только два вида представлены одинаковым обилием, эти показатели будут выражены примерно одним и тем же числом. Однако вопреки очевидной логике, меры эквитабельности продолжают называться «индексами разнообразия», несмотря на то, что они слабо связаны и с количеством («разнообразием?») видов *S* и с вероятностью встретить в местообитании какую-либо особь *N*.

В любом случае исследователь должен априори осознать, какой конкретно фактор он собирается оценивать: если просто выравнивание обилий таксонов в биоценозе, то идеальным инструментом является, скажем, индекс *E* Пиелу; если же предполагается оценка биоразнообразия, то однозначного ответа у нас нет. Согласно сформулированному выше формальному определению биоразнообразия (1), необходимо учитывать одновременно две составляющие: *видовое богатство* и *близость эмпирического распределения обилий видов к некоторому «экологическому эталону»*. Подбор функции, которая связала бы в единый работоспособный показатель оба этих фактора, представляет собой нетривиальную задачу (как и в любом другом случае системного анализа, когда конструируются комплексные критерии в условиях неопределенности целей оптимизации). Более того, вид этого соотношения может трансформироваться в зависимости от конкретных условий, в которых решается задача оценки качества³. Одним из примеров комплексной двухкритериальной оценки разнообразия является индекс энтропии (энтропийный коэффициент) *R*_T Тейла. В более общем случае можно пойти дальше и искать наилучшее решение, например, в семействе аддитивных или мультипликативных функций:

$$R = a + k_1(\log_2 S) + k_2 H, \quad R = a (\log_2 S)^{k_1} (H)^{k_2},$$

где *a*, *k*₁, *k*₂ – произвольные коэффициенты, настраиваемые в ходе массовой экспертизы.

Совершенно справедливо возражая против упрощенной идентификации разнообразия как энтропии, А. Хильберт (A. Hurlbert; цит. по: [4398]) считает, что разнообразие должно быть оценено сразу несколькими индексами и среди них нет такого, который мог бы один служить достаточно репрезентативной его мерой. Такой подход к оценке разнообразия, когда в анализе используется несколько критериев одновременно, нельзя считать научно обоснованным, поскольку неизбежно приведет к неоднозначному толкованию этого термина различными исследователями и окончательному «размытию» понятия «разно-

³ Например, качества «хорошего» стрелка определяют два фактора: «быстрый» и «меткий». В условиях ковбойской драки по сценариям вестернов «хороший» стрелок – это скорее «быстрый», чем «меткий». При спортивной стрельбе по мишеням важнее меткость, чем быстрота; в биатлоне одинаково важны оба показателя.

образия». Конечно, согласование в одном показателе двух не всегда согласующихся между собой факторов является достаточно «тонкой материей» и может вызвать неоднозначное впечатление, но тогда хотя бы становится понятным, какой конкретный смысл вкладывается в конечный результат и в какой пропорции туда включаются отдельные его составляющие. В идеале выбор мер разнообразия должен быть обоснован, исходя из продуктивности экспертных заключений по результатам их широкого практического применения для решения комплекса исследовательских задач.

3. Оценка биоразнообразия с использованием моделей распределения видового обилия

3.1. Кривые «важности» видов

Основой всех индуктивных исследований в экологии является наблюдение за поведением изучаемых биологических объектов и регистрация совокупности показателей. Реализация каждого такого наблюдения (опыта, эксперимента или количественного измерения путем взятия проб) называется *испытанием* при некоторой совокупности условий.

Поскольку каждое биологическое событие происходит с некоторой вероятностью, оно является по своей природе *случайным*. Если результат биологического испытания полностью исчерпывается каким-либо одним (и только одним) событием, то мы имеем дело с *элементарным случайным событием*. Событие, состоящее из нескольких элементарных событий, например, подсчет численностей S различных видов, определяется как *сложное случайное событие*.

Рассмотрим нестационарное поле событий, случайным образом распределенное в пространстве и времени $P_i(x, y, z, t)$, где P_i – некоторый биотический показатель обилия для подмножества особей i -го вида. Предположим, для определенности, что этот показатель связан с видовой численностью N_i , и аналогичным образом в тех же координатах пространственно-временного континуума (x, y, z, t) , представлены показатели обилия и остальных видов $i = (1, 2, \dots, S)$. Тогда биологическое разнообразие в фиксированный момент времени t для точки пространства $\{x, y, z\}$, соответствующей изучаемому местообитанию, будет соответствовать сложному случайному событию, эмпирическая реализация которого выражена ранжированной последовательностью численностей N_1, N_2, \dots, N_S .

Когда такие измерения выполнены и установлена последовательность видов по мере уменьшения⁴ их численностей, данные могут быть представлены в виде графика *ранг/обилие*, где по оси абсцисс – ранг (порядковый номер вида в ранжированном ряду), а по оси ординат – обилие вида (число особей). Линия, соединяющая точки или проходящая близко от них, называется *кривой доминирования-разнообразия*, или *важности* (importance-value) видов. Р. Уиттекер [2718, с. 106-107] считает, что форма кривых важности видов не может служить строгой основой для оценки распределения ресурсов, но она интересна как выражение отношений видов, влияющих на результат измерения разнообразия.

Кривую доминирования-разнообразия можно использовать для оценки различий в видовой структуре: чем круче падает кривая, тем меньше общее разнообразие (в смысле выравненности) и сильнее доминирование одного или нескольких видов. Так на рис. 3 легко усмотреть очевидные отличия в крутизне кривой между группами измерений 1–2 и 5–6. Значительно труднее «на глазок» оценить экологический смысл кривой значимости видов для измерения 3, характеризующего наличием одного субдоминирующего вида и максимальным видовым богатством, и 4 с очень низким числом видов, но феноменальной выравненностью.

⁴ Некоторые экологи, в частности Р. Уиттекер, при построении графиков ранг/обилие по непонятным причинам применяют обратную сортировку ряда и помещают наиболее обильный вид слева. Мы в настоящем изложении, подавляя свой внутренний протест математиков, также используем такой нетрадиционный вид графиков.

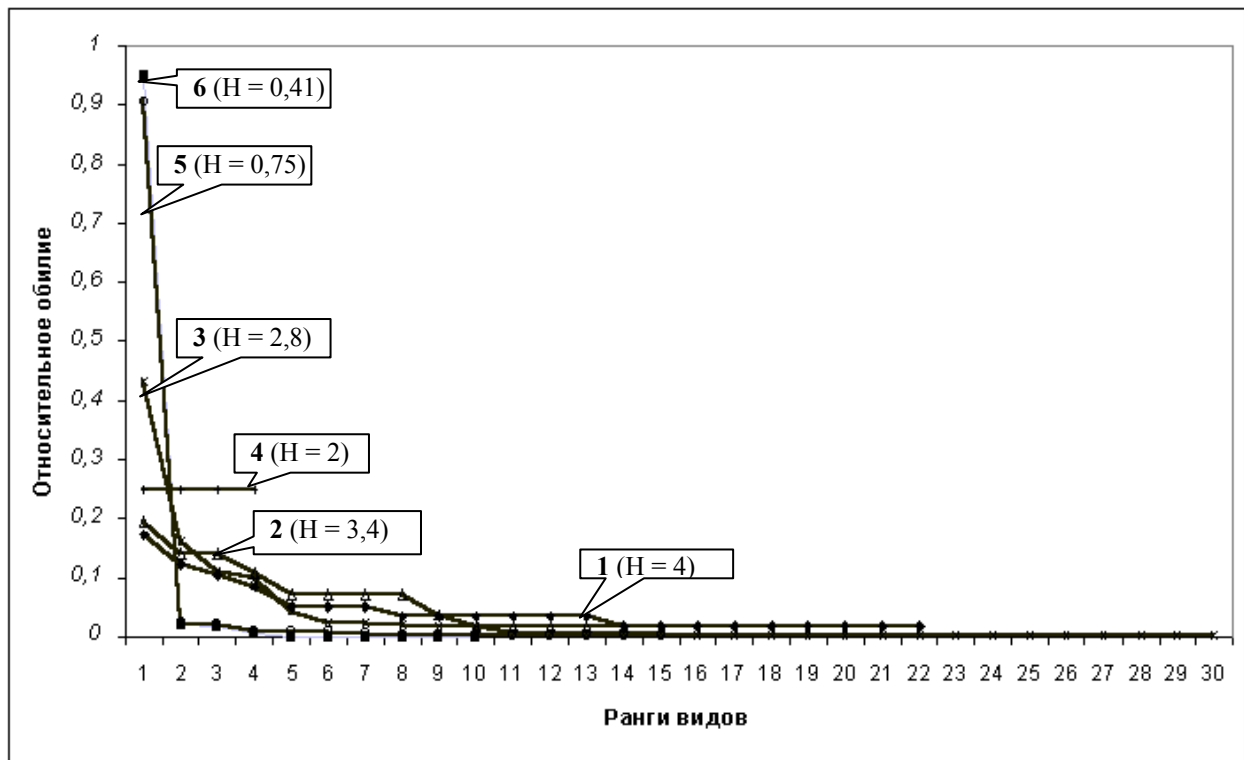


Рис. 3. Кривые доминирования-разнообразия для шести измерений численности макрозообентоса (обозначения см. в табл. 3)

Поскольку интерпретация графических кривых сильно зависит от точки зрения субъекта, ключевой момент большинство исследователей видят в подборе наиболее подходящего аналитического выражения для аппроксимации кривой важности видов (т.е. формулы, описывающей регулярную зависимость в ранжированном ряду обилий видов). Эти формулы называют «теоретическими моделями распределения видового обилия» [2718].

3.2. Основные теоретические модели распределения видового обилия

Поскольку в данном контексте термин «теоретическое» связывается скорее с экологической, нежели математической теорией, остановимся на основных идеях, которые вкладывают авторы в предлагаемые ими модели распределения.

Три нижеприведенных модели основываются на *ранговом распределении* [1837] кривой важности видов Р. Уиттекера (рис. 3):

1. Модель геометрических рядов. И. Мотомура [3968] предположил, что вид-доминант захватывает часть k некоего ограниченного ресурса, второй по обилию вид захватывает такую же долю k остатка этого ресурса, третий по обилию – k от остатка и т.д., пока ресурс не будет разделен между всеми S видами. Если это условие выполнено и если обилия видов (выраженные, например, их биомассой или числом особей) пропорциональны используемой доле ресурса, распределение этих обилий будет описываться *геометрическим рядом* (или *гипотезой преимущественного захвата экологической ниши*). Распределение Мотомуры–Мейя в координатах ранг/обилие представлено прямой, угол наклона которой зависит от константы геометрической прогрессии k . Изложенные предпосылки теории Мотомуры нельзя не считать достаточно поверхностными. Во-первых, непонятно о каком конкретно ресурсе идет речь, поскольку даже в ограниченной группе организмов для разных видов лимитирующими являются разные типы ресурсов. Во-вторых, интенсивность захвата популяцией даже глобального ресурса в большей мере зависит от аллометрии вида, а не от численности особей: например, потребление кислорода одним экзоплером Dreissenidae в тысячи раз превышает тот же показатель для Chironomidae.

2. *Модель «разломанного стержня»*. Р. Мак-Артур предложил похожую схему мотонного изменения обилий P_i в ранжированном ряду, в которой величина k не является постоянной, а изменяется на каждом шаге по некоторому нелинейному закону. На графике важности видов распределение Мак-Артура также представлено прямой, если предварительно прологарифмировать значения рангов i . Считается, что серии Мак-Артура основываются на гипотезе о случайном⁵ характере распределения ресурса: S видов особым образом разделяют среду между собой так, что они занимают не перекрывающиеся экологические ниши, причем число особей каждого вида пропорционально размеру (ширине) выделенного объема гиперпространства факторов. Мак-Артур сравнил разделение пространства ниши в пределах сообщества со случайным и одновременным разламыванием стержня на S кусков, из-за чего распределение часто называют *моделью разломанного стержня*. *Модель экспоненциально разломанного стержня*, предложенная В.Д. Федоровым [2786], основана на введении в модель Мак-Артура нового параметра – плотности вероятности обилий видов, которая в исходной модели предполагается равномерной. Согласно новой модели, на степень перекрывания ниш видов, а соответственно, и на соотношение их обилий, влияет плотность представленности организмов.

3. *Гиперболическая модель* была предложена А.П. Левичем [1473]. Теоретически выводится из модели структуры сообщества, лимитированного потребляемыми ресурсами, при логарифмическом распределении потребностей отдельных видов [1476]. Приводит к более неравномерному распределению обилия, чем предыдущие модели. Предполагается, что, по сравнению с моделью Мотомуры, гиперболическая модель лучше описывает более сложные, «целостные» сообщества, выборки большего объема или усредненные по времени или пространству данные. Интерпретация модели аналогична модели геометрических рядов, однако, согласно гиперболической модели, обилия первых видов убывают более резко, а обилия редких видов – более плавно (первая предполагает линейную, а вторая – логарифмическую зависимость потребности в лимитирующих ресурсах от ранга). А.П. Левич и В.Д. Федоров [2792] гиперболической моделью аппроксимировали ранговые распределения видов в планктонных пробах.

Две других основных модели распределения значимости видов основаны на графиках *частотного распределения*, которые отображают зависимость между числом особей (ось абсцисс) и числом видов (ось ординат), которые соответствует этой численности. Если прологарифмировать численность экземпляров ($\log_2 N$), то в первый диапазон («октаву» по терминологии Ф. Престона) попадут виды, представленные 1-2 особями, второй – 3-4 особями, третий – 5-8 особями и т.д., то есть максимальное граничное значение каждого диапазона получается удвоением предыдущего.

4. *Логнормальное распределение Престона*. Если построить частотную гистограмму числа видов, соответствующих каждой «октаве», то, как полагал Ф. Престон, полученные точки хорошо аппроксимируются гауссовой кривой нормального распределения. Однако если данные, которым соответствует гауссиана, получены из ограниченной выборки, то левая часть кривой (т.е. редкие, неучтенные виды) будет выражена нечетко. Престон назвал такую точку усечения кривой слева *линией занавеса*. «Линия занавеса» может сдвигаться влево при увеличении объема выборки, и, как считал Р. Мак-Артур, число потенциально скрытых видов, которые легко вычислить по уравнению гауссианы, является вполне приемлемой мерой видового разнообразия.

Для большинства выборок на частотных графиках выражена только часть кривой справа от моды. Это имело место и для анализируемых нами 6 измерений, малоудачные попытки аппроксимации которых гауссианой представлены на рис. 4. Только при огромном количестве данных, полученных при одновременном изучении многих классов орга-

⁵ Читатель сам впоследствии убедится, что никакого вероятностного компонента в формуле Мак-Артура нет. Происхождение мифа о «случайном» характере распределения нам неизвестно.

низмов и собранных на обширных биогеографических территориях, прослеживается полная кривая нормального распределения.

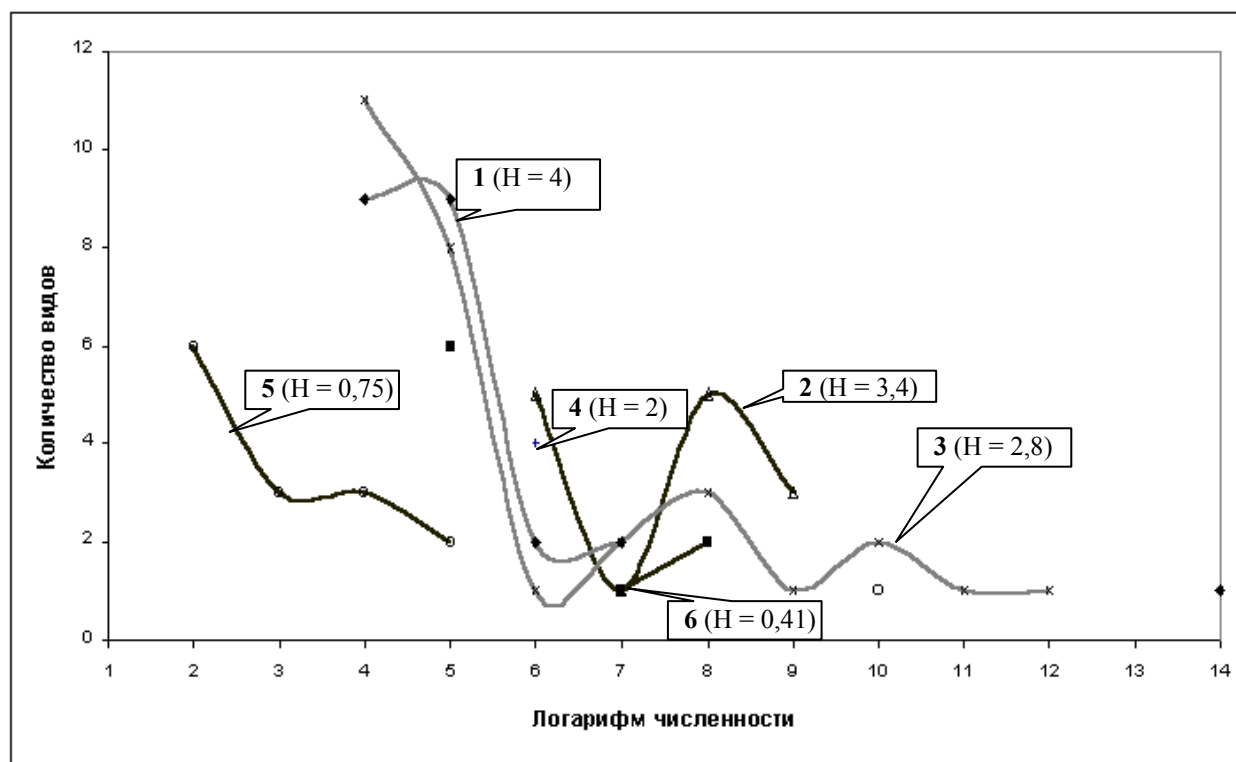


Рис. 4. Графики частотного распределения Престона для шести измерений численности макрозообентоса (обозначения см. в табл. 3)

В принципе, кроме предположений, что «так может быть», никакой экологической теории за моделью логнормального распределения Престона не скрывается. Это не мешает делать глубокомысленные выводы, что «обычно эта модель указывает на большое, зрелое и разнообразное сообщество... S-образная кривая указывает на сложный характер дифференциации и перекрывания ниш. Большинство видов в природных открытых экосистемах существует в условиях соревнования за ресурсы, а не на условиях прямой конкуренции; множество адаптации дает возможность делить ниши без конкурентного исключения из местообитания» [672]. Нам же важнее представляется следствие о том, что концептуально гипотеза Престона находится в диаметральной противоречии с постулатом выравниваемости, подробно обсуждаемом выше. Действительно, индексы Шеннона, Симпсона и другие доставляют максимум биоразнообразия в случае одинаковой численности всех видов. Эта ситуация «одинаковой численности» отображается только одной точкой на графике частотного распределения (см. измерение 4 на рис. 4), которую аппроксимировать кривой логнормального распределения совершенно бессмысленно. Если численность видов варьируется в узком интервале (скажем, от 250 до 1000 экз. или в 3 «октавах», что обычно характерно для макрозообентоса), то мы получим узковершинную гауссиану с сильно «обрезанными хвостами» и мизерным числом «потенциально скрытых видов». Впрочем, выполнять по трем точкам аппроксимацию кривой нормального распределения, вряд ли придет в голову сколько-нибудь квалифицированному статистику.

5. Логарифмические серии Фишера. Наконец, первой по хронологии попыткой описать отношение между числом видов и числом особей этих видов была гипотеза о том, что выборка видов некоторой территории описывается отрицательным биномиальным распределением, высказанная группой английских математиков и энтомологов с участием Р. Фишера [3455]. Они предложили аппроксимировать это распределение логарифмическими сериями (лог-серии, log-series). Модель логарифмического распределения была

впервые применена Фишером как теоретическая модель для описания распределения видов в коллекциях моли. Этой модели и статистике разнообразия было посвящено подробное исследование Л. Тейлора с соавторами [4307]. Логарифмическими сериями, ряд которых характеризуется малым числом обильных видов и большой долей «редких», с наибольшей вероятностью можно описать такие сообщества, структура которых определяется одним или немногими экологическими факторами. Этой модели, например, может соответствовать распределение обилий видов растений наземного яруса в хвойных культурах в условиях низкой освещенности [1837].

Расчетные формулы всех представленных основных теоретических моделей разнообразия приведем в табл. 5.

Таблица 5

Расчетные формулы различных моделей распределения численностей видов N_i ($i = 1, 2, \dots, S$, где S – число видов)

Модель	Расчетная формула	Обозначения
1. Модель геометрических рядов Мотомуры	$N_i = \frac{Nk(1-k)^{i-1}}{1-(1-k)^S};$	$N = \sum_{i=1}^S N_i$, k – параметр модели (отношение геометрического ряда)
2. Модель разломанного стержня Мак-Артура	$N_i = \sqrt{S} \cdot \sum_{l=i}^S (1/l) / S$	
3. Экспоненциальная модель Левича	$N_i = N_1 i^b$	b – параметр модели (мера выравниваемости видов по обилию)
4. Логарифмически нормальное (лог-нормальное) распределение Престона	$S_r = S_m e^{-r-\sigma^2}$	S_r – теоретическое число видов в октаве, расположенной в r -октавах от модальной октавы; S_m – число видов в модальной октаве; σ – стандартное отклонение теоретической лог-нормальной кривой, выраженное в числе октав.
5. Логарифмические серии Фишера	$S = a(x + x^2/2 + x^3/3 + \dots) = -a \ln(1-x)$	a и x – параметры распределения (индексы разнообразия); $a = N(1-x)/x$

Если изобразить каждую из моделей в виде графиков ранг/обилие, можно увидеть переход от геометрического ряда к модели «разломанного стержня» (используем для иллюстрации этого на рис. 5 гипотетическую картинку из работы Р. Уиттекера [4398]). При геометрическом распределении доминируют немногие виды при очень низкой численности большинства, при логарифмическом и логнормальном распределении виды со средним обилием становятся все более и более обычными; в распределении, описываемом моделью «разломанного стержня», обилия видов распределены с максимально возможной в природе равномерностью.

Оценка параметров представленных моделей в общем случае осуществляется с использованием компьютера стандартными методами нелинейного оценивания, оптимизируя при этом, например, сумму квадратов отклонений между теоретическими и экспериментальными значениями обилий видов. Для оценки параметров двух моделей при расчетах без применения персонального компьютера может быть использована процедура приведения их к линейному виду [672]:

$$\log N_i = \log K + (i-1) \log a \quad \text{– для модели геометрических рядов;}$$

$$\log N_i = \log K - b \log i \quad \text{– для гиперболической модели;}$$

где N_i – относительное обилие вида с рангом i ; K (обилие для 1-го таксона), a , b – параметры моделей. После линеаризации оценка параметров и их стандартных ошибок проводится обычным методом наименьших квадратов для линейной регрессии.

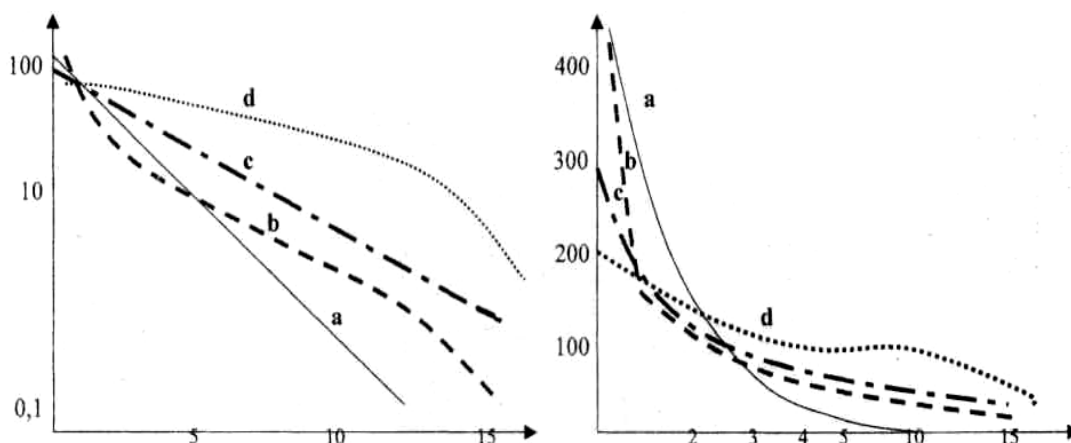


Рис. 5. Различные формы кривых доминантного разнообразия или важностей

Число индивидуумов на ординате размещено против числа видов в последовательности от более к менее обильным на абсциссе. Четыре типа кривых иллюстрируются на двух картинках: **a)** геометрический ряд; **b)** логнормальное распределение; **c)** логарифмическая серия; **d)** распределение Мак-Артура. Кривые рассчитаны для гипотетического примера – 1000 индивидуумов 20 видов. Для кривой «а» $c = 0,5$; для кривой «в» $r = 0,2$; $\sigma = 2,26$; для кривой «с» $a = 3,542$.

Слева при логарифмическом выражении числа индивидуумов последовательность видов линейна. Геометрический ряд отражается прямой линией, а логнормальное и мак-артурово распределение дают сигмоидальные кривые с различной степенью наклонности.

Справа число индивидуумов выражено в линейной шкале, а последовательность видов в логарифмической. Мак-артурово распределение выглядит как прямая линия, другие же формы – как S-образные кривые.

Следует заметить, что, независимо от типа распределения видов по обилию в любой конкретной пробе, такое распределение обязательно имеет «хвост», образованный редкими видами, представленными в пробе единичными особями. Их попадание в пробу следует рассматривать, как совершенно случайное событие, а поэтому оно, скорее всего, должно подчиняться распределению Пуассона. С точки зрения статистики вопрос, как отличить редкие виды (по отношению к которым ранжирование по обилию вряд ли имеет экологический смысл) от «нередких», является весьма не простым и во всяком случае для его решения нельзя предложить какой-либо рецепт, пригодный на все случаи жизни.

Описанные выше теоретические модели не охватывают всего разнообразия функций аппроксимации эмпирических распределений видового обилия: согласно *принципа множественности моделей* В.В. Налимова, для каждой конкретной экологической ситуации возможно построение неограниченного множества математических моделей примерно одинаковой достоверности. Тем не менее многими исследователями предпринимаются попытки подобрать к реальным сообществам другие теоретические модели (таких работ насчитывается более 30). Однако тут возникают естественные вопросы: *ну и что?* Какие конкретные результаты по оценке точечного биоразнообразия можно извлечь из результатов аппроксимаций распределения (кроме естественного чувства удовлетворения их авторов от «модельных игр»)? Имеющиеся публикации не дают на это четкого ответа, а лишь смутные предположения, что тот или иной параметр модели может быть как-то связан с разнообразием.

3.3. Использование моделей распределения для оценки биоразнообразия

В общем случае использование моделей подобного типа происходит по двум возможным механизмам:

- Из всего множества моделей выбирается одна, наилучшим образом аппроксимирующая имеющиеся экспериментальные данные. Сама эта модель или один из ее параметров используется в качестве оценки биоразнообразия. Например, если в ходе подбора теоретических распределений, наилучшее согласие с опытом имела модель Мак-Артура, то можно принять гипотезу о высоком разнообразии местообитаний.

- Выбирается некоторый глобальный «стандарт» аналитического распределения (например, та же модель Мак-Артура, а остальные модели вообще не рассматриваются) и в качестве критерия разнообразия принимается степень расхождения двух кривых: эмпирической и теоретической.

Рассмотрим сначала метод оценки видового богатства Р. Кемптона и Л. Тейлора [4306], учитывающий распределение видовых обилий, но не требующий в явном виде соответствия какой-либо модели. Предложенный ими индекс Q представляет собой меру наклона кривой накопленного видового обилия в ее 2-3 квартилях и обеспечивает измерение разнообразия сообщества, не отдавая предпочтения ни очень обильным, ни очень редким видам. Для расчета индекса по оси абсцисс откладывается обилие видов в логарифмическом масштабе (\log_{10}), а по оси ординат – накопленное число видов. Предполагается, что между 1 и 4 квартилями кумулятивная кривая хорошо аппроксимируется прямой, имеющей угол Q наклона к оси абсцисс, который и принимается за оценку разнообразия. Аналитическая формула для расчета индекса Q по эмпирическим данным имеет вид:

$$Q = \frac{0,5n_{R_1} + \sum_{R_1+1}^{R_2-1} n_r + 0,5n_{R_2}}{\log_{10}(R_2 / R_1)},$$

где n_r – общее число видов с обилием R ; R_1, R_2 – нижний и верхний квартили; n_{R_1} – число особей в классе, соответствующем R_1 ; n_{R_2} – число особей в классе, соответствующем R_2 . Некоторые исследователи находят, что Q связан с параметром a логарифмического распределения Фишера: $Q = 0,371 S/a$. Процедура расчета индекса Q страдает теми же проблемами, что и логнормальная модель Престона: при относительно равномерном распределении численности видов она просто вырождается, утрачивая всякую корректность.

Для используемых нами 6 примеров ни в одном из случаев не «прорисовывалась» S -образная кумулятивная кривая, предсказываемая Р. Кемптоном и Л. Тейлором (см. рис. 6.)

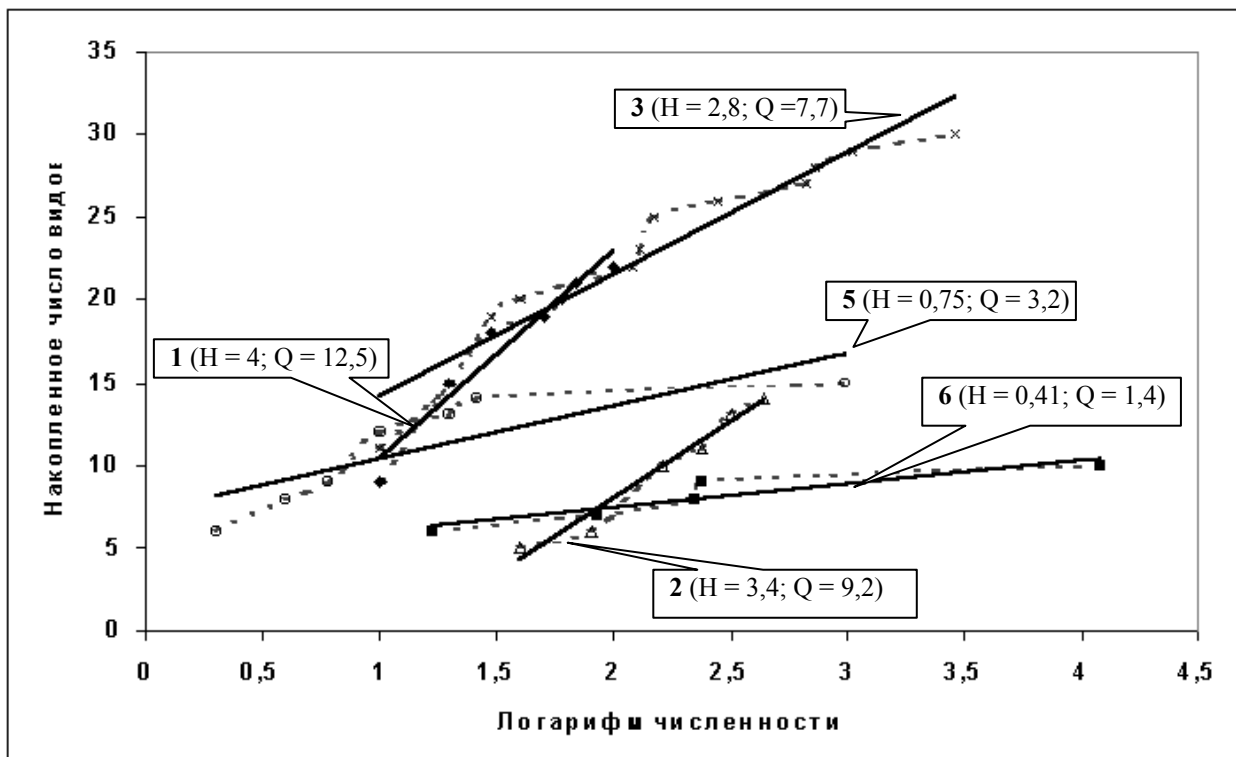


Рис. 6. Графики линейной аппроксимации кумулятивной кривой для пяти измерений численности макрозообентоса (обозначения см. в табл. 3)

Несколько ранее индекс, основанный на сходной идее, был предложен Р. Уиттекером, при этом он учитывал всю кривую видовых обилий, что, по мнению специалистов

[672], «дает ошибки на обоих концах распределения». В нашем случае использование версии Уиттекера оказалось значительно надежнее и технологичнее: ввиду того что доминанты и субдоминанты – явление для макрозообентоса относительно редкое, то, после отбрасывания длинного «хвоста» 4-й квантили, для нахождения угла наклона Q у нас оставалось обычно 1-2 точки. Однако сам метод анализа кумулятивной кривой накопленных частот не продемонстрировал ни ощутимых преимуществ, ни серьезных расхождений в результатах оценки разнообразия (см. рис. 6) по сравнению с индексом Шеннона.

При сопоставлении эмпирического распределения и «стандартизованной» теоретической кривой степень их соответствия можно оценивать по различным критериям статистического анализа:

- квадрату коэффициента линейной корреляции между реальными и теоретическими значениями обилия (% общей дисперсии, объясненной моделью) или сумме квадратов отклонений реальных значений от модельных;
- суммарной площади фигуры, образованной разностью двух распределений, как это применяется при расчете индекса Джини (см. рис. 1) ;
- максимальному расхождению между эмпирической и теоретической кривыми (по аналогии с критерием согласия Колмогорова).

Здесь, а также в выборе того «стандарта» аналитического распределения, с которым будут сравниваться все эмпирические данные, видятся как большое поле для «индексотворчества», так и трудности последующей унификации методик.

Предложим один из возможных подходов к оценке биоразнообразия графоаналитическим методом. Проведем на графике важности видов (см. рис. 7) с прологарифмированной осью ординат прямую через точки с максимальной $\ln(N_1)$ и минимальной $\ln(N_S)$ численностью видов.

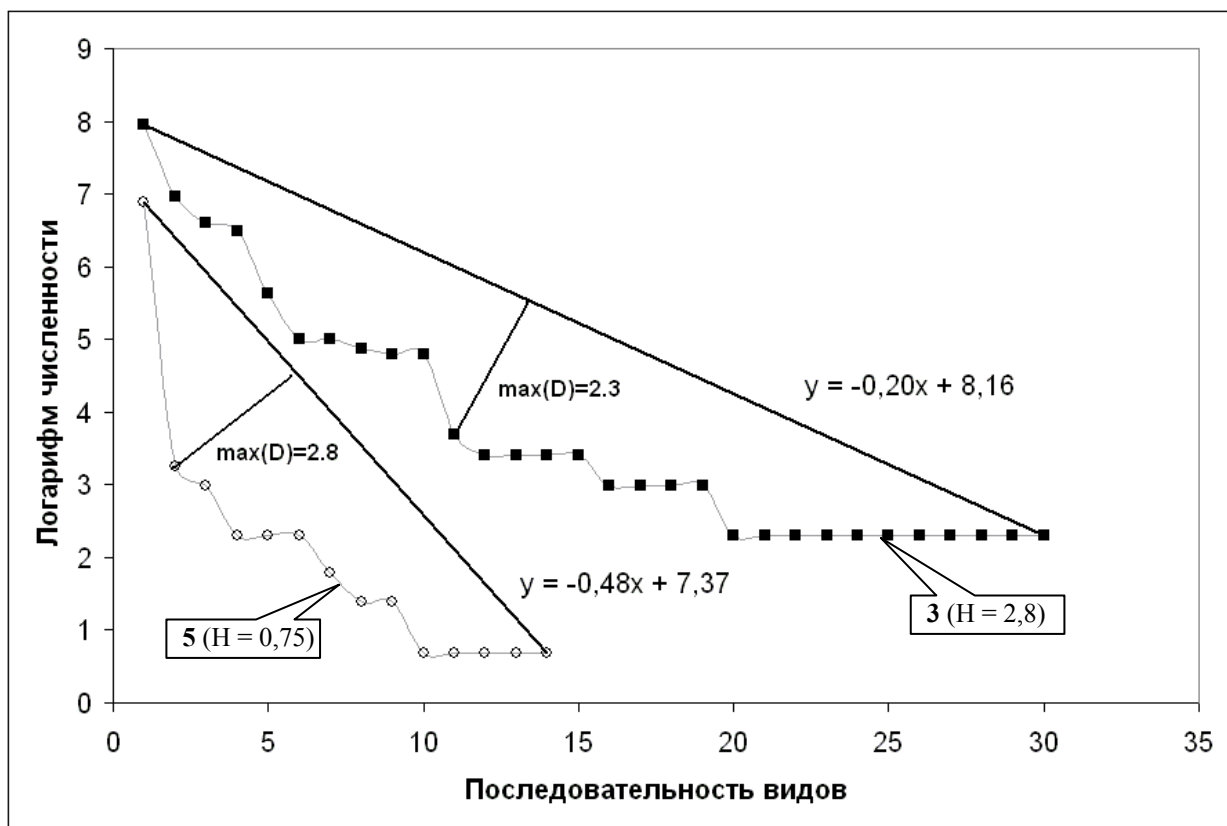


Рис. 7. Графоаналитический метод оценки биоразнообразия для двух измерений численности макрозообентоса (обозначения см. в табл. 3)

Длина L_R двух отрезков ломаной от $\ln(N_1)$ до $\ln(N_S)$ и далее до пересечения с осью абсцисс будет соответствовать общему биотическому богатству, учитывающему как численность, так и количество видов:

$$L_R = \ln N_s + \sqrt{(\ln N_1 - \ln N_s)^2 + S^2} .$$

Восстановим из каждой промежуточной точки $\ln(N_2), \dots, \ln(N_{S-1})$ перпендикуляр и найдем значение максимального расстояния $\max(D)$ от эмпирической кривой до линии $[\ln(N_1) - \ln(N_S)]$, оценивающего степень доминирования в сообществе. Тогда показатель разнообразия может иметь вид:

$$R = L_R - \lambda \max(D) ,$$

где λ – коэффициент, нормирующий относительную важность каждого из факторов. Если принять значение параметра $\lambda = 3$, то получим численные значения показателя для двух измерений, представленных на рис. 7:

$$\text{для станции 3} - R = 32,8 - 3 \cdot 2,3 = 25,9;$$

$$\text{для станции 5} - R = 16,1 - 3 \cdot 2,84 = 7,5 .$$

Несмотря на некоторые элементы логической обоснованности показателя R (которой, к сожалению, не отличаются многие описанные выше способы оценки разнообразия), мы привели его конструкцию в значительной степени только для того, чтобы проиллюстрировать, сколь легко рождаются новые и новые индексы, при этом мало задумываясь, в каких условиях применим тот или иной метод оценки, каковы требования к составу исходных данных и как следует интерпретировать полученные результаты.

Шкала вновь разработанного комплексного показателя тогда становится осознаваемой, когда на ней выделяется некоторое количество «характерных» точек или интервалов (точка наибольшей благожелательности, диапазоны толерантности, предельно допустимые значения, критические или пороговые точки, диапазон необратимых кризисных явлений, катастрофический экстремум и пр.). Только после этого числа, представленные в шкале предлагаемого индекса, будучи сопоставлены с ее характерными точками, приобретают определенную смысловую идентификацию, а показатель становится «работоспособным». А это – длительный и трудный процесс всестороннего анализа спроектированного критерия на массовых экспериментальных данных.

За последние 30 лет у экологов в отношении к моделям ранговых распределений произошла определенная «переоценка ценностей». Вообще говоря, серьезная критика в адрес этого подхода высказывалась с самого начала их внедрения в практику экологических исследований. В частности, в ответ на такую критику модели разломанного стержня со стороны Е. Пиелу еще в 1966 г., Р. Мак-Артур признал, что «модель служит грубым приближением экологии сообществ, которой нужно позволить умереть естественным путем» (цит. по: [1476]). Теперь, почти 40 лет спустя, мы, по-видимому, являемся свидетелями этого «умирания», причем сам подход к моделированию в этом несколько не виноват. Можно выделить следующие причины такой тенденции:

- инфляция облика модного ранее «модельерства» в среде научного «общественного мнения»;
- желание быстро получить сомнительный результат (рассчитав, например, индекс Шеннона);
- отсутствие тщательно проработанных методик по процедурам построения и интерпретации моделей и адаптированных для экологов соответствующих пакетов программ;
- низкая математическая и компьютерная квалификация практикующих экологов.

Основываясь на общем принципе «маятника», авторы надеются на ренессанс идеи моделирования распределений в ближайшие десятилетия.

4. Методы оценки обобщенного биоразнообразия подмножества местообитаний

4.1. Предварительные замечания

Не будем лишать себя удовольствия привести сначала две достаточно обширных цитаты общесистемного характера из работ почитаемых нами авторов:

«Исходные посылки:

1. Будем считать, что весь эволюционирующий мир можно рассматривать как множество текстов. Когда мы говорим о биосфере, то текстами оказываются отдельные особи, виды и другие составляющие биосферы. Когда рассматриваем ноосферу, то текстами становятся сознания людей как в их личностных, так и коллективных проявлениях.

...4. Изначально все возможные смыслы мира как-то соотношены с линейным континуумом Кантора – числовой осью μ , на которой в порядке возрастания их величин расположены все вещественные числа.

...6. Распаковывание (появление текстов) осуществляется вероятностной взвешиваемостью оси μ , т.е. разным ее участкам приписывается разная мера.

7. Соответственно, семантика каждого конкретного текста задается своей функцией распределения (плотностью вероятности) – $p(\mu)$. В общем случае можно говорить о текстах, определяемых функцией распределения вероятности, задаваемой на многомерном пространстве.

Функция $p(\mu)$ оказывается тем окном, через которое нам дана возможность всматриваться в мир смыслов».

В.В. Налимов [1851] с некоторыми сокращениями.

«Точечные разнообразия, т.е. разнообразия проб, взятых в точке, почти всегда бессмысленны... Так как организация, выраженная через разнообразие, развертывается в пространстве, измерения не имеют большого смысла, если они не представлены в виде спектра: разнообразие как функция пространства»⁶.

Р. Маргалев [1642].

В предыдущих разделах мы определили экологические наблюдения как эмпирическую выборку из нестационарного поля событий, случайным образом распределенных в пространстве и времени. Не вторгаясь в более «тонкие материи», обсудим отличия подробно рассмотренных выше оценок точечного биоразнообразия в пробе, от расчета совокупного биоразнообразия в некотором подмножестве местообитаний (α -разнообразие гомогенного сообщества, γ -разнообразие ландшафта, ϵ -разнообразие биома и т.д.). Во всех своих греко-буквенных проявлениях второй вариант связан с определенными процедурами математического усреднения результатов наблюдений в пятимерном параллелепипеде с осями (P, x, y, z, t) . В этом гиперпространстве распределяются данные в виде облака из M обобщаемых точек ($j = 1, 2, \dots, M$), координаты которых соответствуют показателям обилия P_{ij} для i -го вида (или экологической группы особей), полученным в момент времени t для области пространства $\{x, y, z\}$.

С использованием методов OLAP-технологии имеется возможность выполнить анализ отдельных срезов многомерного куба эмпирических данных секущими плоскостями и отображать полученные выборки в двух- или трехмерном евклидовом пространстве. Например, задавшись временным интервалом Δt и диапазоном глубины (высоты) Δz , мы можем получить трехмерную диаграмму плотности распределения i -го вида в географических координатах (x, y) , либо плоскостную карту с изономами его обилия. Однако подоб-

⁶ Можно добавить: «и времени».

ный подход не обеспечивает расчет необходимых исходных компонент для оценки биоразнообразия, т.е. величин S и p_i в выражении (1) на с.98.

При переходе к многомерному варианту оценки по сумме местообитаний смысл основных параметров биоразнообразия существенно трансформируется:

- количество видов S , однозначно определенное для пробы, рассчитывается в этом случае статистическими методами с учетом многомерной плотности распределения вероятности нахождения каждого вида в каждый момент времени в каждой точке пространства;
- показатель обилия P_i для i -го вида в самом общем виде трактуется как интеграл от дифференциальной *популяционной плотности* по всей анализируемой области пространственно-временного континуума.

Не будем здесь приводить запись достаточно очевидных теоретических уравнений интегральных моделей для S и P_i , что было бы неуместным математическим кокетством, принимая во внимание весьма малую вероятность исчисления дифференциальных плотностей в практических условиях. Поэтому последующее изложение посвящено обсуждению эвристических методов обобщения реальных экспериментальных данных для расчета списка видов и популяционного обилия произвольного экологического сообщества.

4.2. Как рассчитать плотность популяции?

Рассматривая некоторое подмножество местообитаний, мы неизбежно приходим к уточнению понятия *средневзвешенной плотности популяции* – вероятного числа особей или моментной биомассы популяции в расчете на единицу площади или объема. В принципе, с плотностью может быть связана и любая другая однозначно и четко регистрируемая величина, *линейно* отражающая численность особей: содержание ДНК (РНК); число фекалий животных на единицу маршрута; количество силуэтов птиц, появляющихся на фоне диска Луны за единицу времени и др.

Анализ пространственного распределения популяционной плотности связан с тремя наиболее фундаментальными задачами экологии:

- установление особенностей структурного развития биоценоза и определение точных границ для каждой составляющей его популяции;
- построение моделей причинно-следственной связи между факторами внешней среды и изменением популяционной плотности (то, что уже 50 лет назад называлось «градиентным анализом» [4399, 4400]);
- оценка показателей средневзвешенной мультипопуляционной плотности организмов и подход к измерению γ -разнообразия территории.

Наиболее сложной проблемой при определении показателей численности вида является определение точных границ популяции. Часто при проведении экологических исследований за границы популяций принимают резкие географические барьеры, непреодолимые для организмов, или административно-территориальные границы. Если эти границы все же очерчены, то появляется возможность различать *абсолютную плотность* (параметр популяции на единицу всего пространства) и *экологическую плотность* (параметр популяции на единицу заселенного популяцией пространства).

Одно из основных затруднений при измерении и выражении плотности популяции возникает в связи с тем, что особи распределяются в пространстве неравномерно. При этом, как правило, интерес представляют следующие вопросы:

- как классифицировать наблюдаемое размещение?
- как оценить плотность объектов по небольшому числу измерений?

Под классификацией понимается отнесение конкретной совокупности объектов к одному из трех типов размещения [1761, 2284]:

- регулярное (детерминированное равномерное или стохастическое равномерное);

- групповое (детерминированное неравномерное);
- случайное (стохастическое неравномерное).

При ответе на *первый вопрос* для целей классификации обычно предполагается, что любая модель пространственного распределения ансамбля точек есть реализация некоторого случайного процесса (например, однородного точечного процесса Пуассона, как модели совершенно случайного размещения). Далее проверяется согласие имеющихся данных с выдвигаемой гипотезой, после чего гипотеза о случайности либо принимается, либо отвергается. Для проверки гипотезы могут быть использованы различные статистические критерии (например, критерии Холгейта, Хопкинса, Кокса, Эберхарда, однородности Диггла и другие традиционные или более экзотические критерии). Другим наиболее общим тестом для такой оценки является также определение *меры диссонанса* (различия между реальным и теоретическим регулярным размещением) или расчет индексов, предложенных разными авторами: Ю. Одумом [1936], Т. Соутвудом [4246], М. Мориситой [3956], С. Ивао [3689] и другими (подробнее см., например, [1767]).

Оценка типа распределения ансамбля точек нашла свое должное место в фитоценологических исследованиях, однако трудно предположить необходимость подобных упражнений в отношении, например, сообществ макрозообентоса. Во-первых, сама конфигурация облака точек в водоеме не является стохастической, а определяется планом гидробиологических экспедиций и местоположением станций отбора проб, и, во-вторых, появление или отсутствие *i*-го вида в точке $\{x, y\}$ определяется в значительной мере биотопической обстановкой, а не стремлением вида к некоей пространственной регулярности.

Вернемся ко *второму вопросу*: как все же рассчитать характер распределения плотности популяции по небольшому числу измерений? Вследствие континуального характера градиентов изменения как субстрата (окружающей среды), так и элементов биосферы, наиболее подходящим аппаратом для этого являются вероятностные методы. Способы получения вероятностного описания размещения популяции в экосистеме теоретически хорошо известны. Например, для анализа и классификации пространственного размещения можно применять разнообразный математический аппарат двухмерной аппроксимации, круги Вороного, сглаживание по Биалсу [3203] и др. Для статистического моделирования можно применять многомерный анализ, логистическую регрессию и другие аналогичные модели [3005], использующие данные о значениях факторов среды в точках измерения.

К сожалению, весь этот мощный математический аппарат остается некоторой отстраненной экзотикой, когда речь заходит о практической работе с гидробиологическими данными. Зато широко распространены многочисленные отчетные реликвии об обилии того или иного вида в водоеме, претендующие на глубокое обобщение экспедиционных данных, а на деле оказывающиеся «лукавой цифрой». Рассмотрим в этой связи те стандартные рекомендации, которые даются прикладной статистикой, и область применимости этих рекомендаций к типичным гидробиологическим измерениям.

Абсолютная плотность (*A*) организмов на исследуемом *однородном* участке определяется как среднее арифметическое количества зарегистрированных в границах каждого местообитания (x_i) организмов (данного вида):
$$A = \frac{\sum x_i}{n},$$

где *n* – количество исследованных пробных площадок ($i \in [1, n]$).

Показателем, который характеризует точность получаемых результатов, является *ошибка оценки плотности* $SE_D = S/n^{1/2}$, которая определяется через понятие ее *вариансы*

(или оценки дисперсии):
$$S^2 = \frac{\sum x_i^2 - (\sum x_i)^2}{n-1},$$
 где $\sum x_i^2$ – сумма квадратов количеств организмов в любой из *n* пробных площадок, а $(\sum x_i)^2$ – квадрат суммы этих величин.

Достоверными считаются результаты, для которых отношение A/SE_D оказывается более или равно 5, т.е. величина ошибки должна составлять не более 20% величины самого показателя.

Рассмотрим небольшой, но очень типичный пример⁷. По данным экспедиционных исследований в 27 пробах, сделанных в разное время на 8 станциях р. Байтуган, вид *Telopelopia* (сем. Chironomidae) встретился только в 6 пробах со значениями численности 1350, 120, 60, 50, 40 и 10 экз./м². Поскольку р. Байтуган – небольшой водоток, не имеющий существенных перепадов условий среды и находящийся в зоне умеренного антропогенного воздействия, то все 27 измерений можно считать однородными. Элементарная статистическая обработка дает следующие формальные результаты: мода – 0, медиана – 0, среднее арифметическое – 60, стандартная ошибка среднего – 50. Очевидно, что ошибка среднего практически совпадает с самим значением среднего арифметического, которое в данном случае трудно интерпретировать как «центр распределения» или «оценку математического ожидания», поскольку величина A находится где-то в районе последнего десятичного интервала.

А как же вычислять средние? Общее понятие среднего по Колмогорову [1957] таково: средней величиной для чисел X_1, X_2, \dots, X_n является любая функция $G\{(F(X_1)+F(X_2)+\dots+F(X_n))/n\}$, где F – строго монотонная функция, G – функция, обратная к F . Если $F(x) = x$, то среднее по Колмогорову – это среднее арифметическое, если $F(x) = \ln x$, то среднее геометрическое, если $F(x) = 1/x$, то среднее гармоническое и т.д. При этом показано, что для большинства шкал измерения обилия видов из всех средних по Колмогорову наиболее устойчивыми относительно сравнения являются только степенные средние и среднее геометрическое. В частности, среднее геометрическое численности вида *Telopelopia* в р. Байтуган равно 2,7 экз./м², что, на наш взгляд, соответствует более обоснованному значению, не в столь очевидной степени зависящему от единственного сильно завышенного значения ($x = 1350$ экз./м²).

К сожалению, за 100 лет развития количественной экологии [378] (в том числе и гидроэкологии) так и не появились внятных методических разработок, как нужно рассчитывать средние показатели обилия для последовательности однородных наблюдений и как следует оценивать степень этой однородности. И когда наши уважаемые коллеги [3056] нашли, например, что в р. Прони численность вида *Cladotanytarsus mancus* около 3 экз./м², а по нашим данным для малых рек бассейна Нижней Волги (средним из 560 проб), численность этого вида оказывается в 55 раз больше, то трудно сказать, чем объясняются различия в показателях: существенными экологическими отличиями рек Рязанской и Самарской областей или некоторыми отдельными аномальными пробами, сильно влияющими на подсчет арифметического среднего.

⁷ К этому конкретному примеру (как и ко всем реальным примерам) специалист-эколог всегда найдет множество «объяснений». Так, вид *Telopelopia* – трудно определяемый таксон, часть проб – определяли мало квалифицированные лаборанты или студенты, в некоторый период (1350 экз./м²) возникли специфические условия местообитания и пр. Но среднюю подсчитать хочется!..

Приведем другой, сходный, но еще более абстрактный пример [430, 724, 2292, с. 24]. На скамейке сидело 5 человек. Двое были бродягами, имущество которых оценивалось по 25 центов, третий был рабочим с доходом в 2000\$, четвертый владел 15 000\$, а пятый был миллионером с чистым доходом в 5 млн.\$. Таким образом, мода равна 25 центам (характеризует 40% выборки), медиана – 2000\$, средняя геометрическая – 180\$, а средняя арифметическая – 1 003 400,1\$. Если взять за основу моду, то средний актив этой странной группы на скамейке должен быть оценен в 1-1,5\$ (что обидно для всех, кроме бродяг). Остальные оценки выглядят еще хуже. Очевидно, нет меры, которая была бы адекватна этим «коллегам по скамейке» (отсутствует центральная тенденция).

4.3. Оценка видового богатства совокупности местообитаний

Если в тех же пространственно-временных координатах (P, x, y, z, t) рассматривать в отдельности все множество биологических объектов, то можно сформировать список видов $i = 1, 2, \dots, S$, встретившихся в параллелепипеде описаний. Хотя факт встречаемости произвольного вида – более надежная категория, чем оценка среднего обилия, для каждого описания, сделанного в разное время или в разных точках территории (или водоема), списки видов обычно также сильно различаются. С одной стороны, эти описания можно рассматривать как выборку из генеральной совокупности всех возможных сочетаний видов *локального актуального пула*, распределение вероятностей которой определено сочетанием экологических факторов в данной точке. С другой стороны, отмечаются существенные стохастические отклонения распределения видов от их равновесного состояния, определяемого локальными градиентами среды, которые имеют как пространственную, так и временную случайные составляющие. В связи с этим общая проблема моделирования распределения встречаемости видов по эмпирическим пробам состоит в том, что, не имея дополнительной информации, мы не можем точно установить, чем вызвано отсутствие данного вида в данной точке – неблагоприятными условиями среды, при которых развитие вида принципиально невозможно, либо какими-то случайными факторами при наличии общих благоприятных условий.

Для описания видового состава совокупности местообитаний обычно используются *списки видов*, т.е. полный перечень видов, который встретился в гиперкубе данных. Будучи дискретными, эти списки не всегда удобны для анализа и моделирования в силу следующих причин. *Во-первых*, инвентаризация видового богатства должна применяться только в отношении резистентных видов, которые имеют очевидные свидетельства их размножения, за вычетом временных и случайных видов [3865]. К сожалению, формально-математическими методами без экспертного участия опытного биолога невозможно отличить редкие и малочисленные виды от «временных и случайных». *Во-вторых*, размер списка видов сильно зависит от общего количества сделанных проб M . Так, в работах по общей теории систем и математической лингвистике [2041] показано, что для вычисления дискретных накопленных вероятностей в частотном словаре наиболее эффективно использование непрерывной функции распределения Вейбулла, а общий объем тезауруса системы в простейшем случае аппроксимируется экспоненциальной зависимостью:

$$S_M = S_0 e^{-kM},$$

где S_M – количество фраз «словаря» для системы из M элементов, k и S_0 – коэффициенты модели регрессии.

Описано несколько способов оценки разнообразия для совокупности местообитаний на основе данных по присутствию или отсутствию видов [672].

- *Мера Уиттекера* рассчитывается по формуле

$$\beta_W = \frac{S_M}{a_M} - 1,$$

где S_M – общее число видов, зарегистрированных в системе после выполнения M измерений, a_M – среднее число видов в одном измерении.

- *Меры Ратледжа*. Мера β_R учитывает общее видовое богатство и степень совпадения видов:

$$\beta_R = \frac{S_M^2}{2r + S_M},$$

где S – общее число видов во всех выборках, а r – число пар видов с перекрывающимся распределением.

Другие меры Ратледжа, β_I и β_E , основаны на теории информации и были упрощены для качественных данных и равного размера выборок:

$$\beta_I = \log(T) - (1/T) \sum b_i \log(b_i) - (1/T) \sum a_j \log(a_j); \quad \beta_E = e^{\beta_I} - 1,$$

где b_i – число выборок вдоль трансекты (некоторая произвольная линия, проведенная в многомерном параллелепипеде данных), в котором представлен i -й вид, a_j – видовое богатство j -й выборки, а $T = \sum b_i = \sum a_j$.

- Еще одна группа мер, так или иначе, связывает видовое богатство с линейными размерами пространства местообитаний. Так, обоснована [2263] модель «число видов – площадь изучаемой территории (V)»:

$$V = \frac{2p_v}{(1-p_v)} - A_0 \ln(1-p_v) \cdot e^{-ka_v},$$

где A_0 и k – коэффициенты модели, $p = a_v / S_v$, а обозначения S_v и a_v имеют тот же смысл, что и для меры Уиттекера.

- Мера Коуди β_C разработана для исследования изменений в сообществе птиц вдоль «средового градиента»:

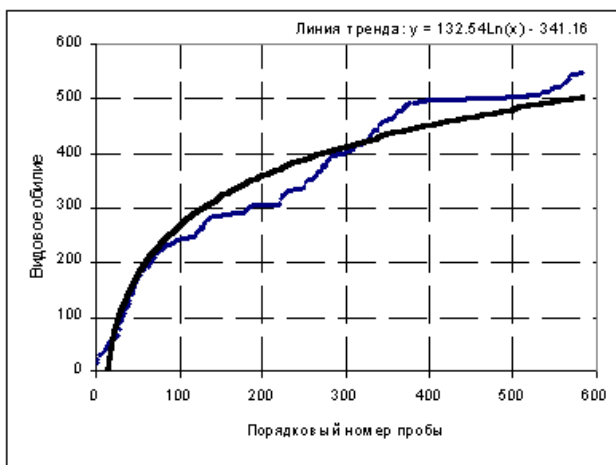
$$\beta_C = 0,5[g(S) - l(S)],$$

где $g(S)$ – число видов, прибавившихся вдоль «градиента местообитаний», а $l(S)$ – число видов, утраченное на той же трансекте.

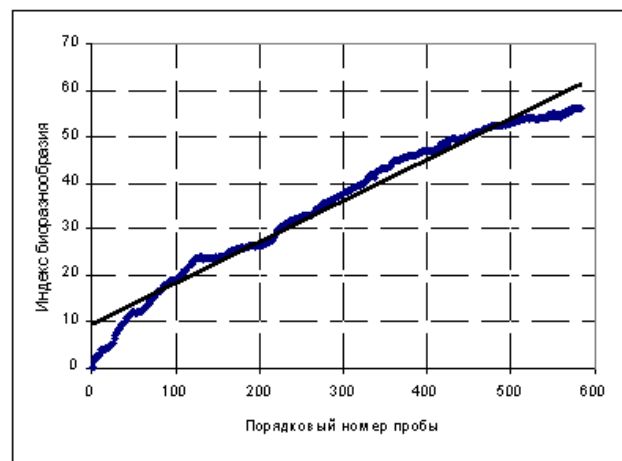
- Мера Уилсона и Шмиды β_T нормирует меру Коуди на среднее видовое богатство выборок a , входящее в меру Уиттекера:

$$\beta_T = \beta_C / a.$$

В работе Э. Мэгарран [1837] приводится оценка совокупности различных вышеупомянутых формул с целью определить лучший показатель β -разнообразия и делается вывод, что большинству критериев удовлетворяет именно мера Уиттекера. На рис. 8 представлена динамика изменения числа обнаруженных видов макрозообентоса S_t по мере взятия проб за период с 10.07.85 г. по 29.09.04 г. ($t = 1, 2, \dots, 570$) в малых реках Самарской области (фиг. А). В этих же условиях среднее число видов в пробе изменялось незначительно (от 8,9 до 15). Поскольку трудно себе представить, что реальное биоразнообразие экосистемы за 20 лет возросло в десятки раз (см. фиг. Б), то нельзя рассматривать меру Уиттекера в качестве сколько-нибудь серьезной оценки относительного видового богатства вследствие ее сильной зависимости от размеров экспериментальной выборки M .



А. Динамика изменения числа видов S



Б. Динамика изменения индекса биоразнообразия по Уиттекеру β_w

Рис. 8. Изменение видового богатства и «биоразнообразия» макрозообентоса в малых реках Самарской области

4.4. Формирование списка видов с учетом их обилия

При оценке встречаемости особей какого-либо вида учитывается только их наличие или отсутствие в отобранных пробах. При этом как обнаружение единственного экземпляра, так и большого их количества, трактуется одинаково: «особи данного вида встречаются в биотопе». Таким образом, виды, характерные для сообщества и случайные для него, получают одинаковый статус, что значительно обесценивает результаты формирования видового списка и последующий его анализ.

Исходя из этого, необходим строгий количественный критерий, оценивающий экологическую роль вида и позволяющий определить: является ли конкретный вид характерным для сообщества или он находится вне «локального актуального пула», что позволяет исключить его при сравнительной оценке видового состава в различных условиях. Обычно роль такого критерия отводится встречаемости b_i , усредненным популяционным показателям (например, доле численности вида в общей численности всех найденных особей $p_i/\Sigma p_i$) или индексам доминантности, таким как индекс Паляя-Ковнацкого ($b_i p_i / \Sigma p_i$).

Как считает В.Ф. Шуйский с соавторами [3056], степень характерности вида для сообщества нельзя оценивать на основе их ранжирования по значениям популяционных показателей, усредненным за период наблюдения. Вид может встречаться регулярно и являться вполне характерным для сообщества, но иметь сравнительно низкую популяционную плотность. Например, по нашим данным, *Sphaeriastrum rivicola* занимает видное место по встречаемости в малых реках Самарской области, но серьезно уступает многим другим видам по удельной численности. Кроме того, по абсолютным значениям долей численности или индекса доминирования не всегда удается выделить и отбросить подмножество случайных или второстепенных видов. Например, согласно установленных градаций доминирования диапазон индекса Ковнацкого менее 0,01 должны иметь совсем уж «плохонькие» адоминанты группы В [213]. В случае небольших эмпирических выборок это возможно и так. Но при наличии весьма репрезентативного материала резко уменьшается масштаб показателей, основанных на долях: только 44 вида макрозообентоса из 546, найденных в малых реках Самарской области, имели значение индекса Паляя-Ковнацкого более 0,01 (и лишь 6 видов – более 1).

С учетом сказанного, предлагается [3056] оценивать степень характерности вида для сообщества по той достоверности, с которой среднее значение популяционной плотности отличается от нуля. Проверку нулевой гипотезы авторы предлагают осуществлять с использованием критерия Стьюдента $t = A/m_A$, где A – средняя арифметическая популяционной плотности за период наблюдений, m_A – ошибка средней арифметической. Количество степеней свободы $\nu = n - 2$, где n – общее количество проб, отобранных в тестируемой части водоема. Далее, в зависимости от уровня значимости ξ , задаваемого исследователем, все виды делятся на две категории:

- виды, *характерные для сообщества*, средняя популяционная плотность которых отличается от нуля с достоверностью, превышающей доверительный уровень $(1 - \xi)$;
- виды, *не характерные для сообщества*, для которых нулевая гипотеза о равенстве нулю средней популяционной плотности статистически не отвергается⁸.

⁸ Заметим, что дискуссия о роли *характерных видов* в фитоценологическом классификационном подходе, связанном с именем Ж. Браун-Бланке, была весьма острой в третьей четверти XX века. В настоящее время считается, что концепция «характерных видов» (предпочитающих тот или иной синтаксон, встречающихся исключительно в его сообществах или имеющих там более высокое постоянство) имеет весьма относительную ценность [1767]. Характерные виды в фитоценологических исследованиях все чаще играют диагностическую роль только в отдельных районах, т.е. являются *локальными*. По-видимому, нет никаких оснований в гидроэкологических исследованиях придавать им бóльшую роль, чем в фитоценологических...

Без сомнения, использование формальных статистических критериев может существенно повысить обоснованность формирования локального актуального пула видов. Однако требуются дополнительные исследования устойчивости критерия Стьюдента в условиях различного характера распределения наблюдаемых значений популяционной плотности у разных членов сообщества. Принимая во внимание традиционный для биологических компонентов логнормальный характер распределения показателей обилия, в качестве эффективной профилактической меры можно предложить использование во всех случаях статистической обработки данных *предварительное логарифмирование* показателей обилия [170], что обычно приводит к порождению нормально распределенной величины. Другим вариантом «нормализации» выборок P_i является использование известного углового преобразования Фишера [2744], при котором частотные оценки вероятностей имеют ошибку, почти не зависящую от самих вероятностей. Используемая при этом функция $\arcsin(2P - 1)$ ведет себя почти так же, как используемая в байесовских подходах функция $\ln(P/(1 - P))$, но в то же время при P , близких к 0 или 1, не вырождается, устремляясь в бесконечность, а ограничена $\pm\pi/2$.

4.5. Оценка β -разнообразия на основе мер сходства подмножеств измерений

Основная концепция, заложенная в понятие биоразнообразия (см. раздел 1.1), заключается в раздельной оценке «разнообразия *внутри и между местообитаниями*» [3344, 3516]. «Внутреннее биоразнообразие» для каждого измерения рассчитывается одним из способов, описанным выше (например, с использованием шенноновской энтропии). «Биоразнообразие между местообитаниями» (т.е. пробами) оценивается в терминах видового сходства рассматриваемых объектов. Для этого рассчитывается матрица R размерностью $m \times m$ метрического расстояния в пространстве видов S между всеми парами m точек параллелепипеда данных, после чего используется широко описанная в литературе технология кластерного анализа [170, 1761, 1955, 2011, 3005].

В разделе «Анализ бета-разнообразия: сравнение, сходство, соответствие сообществ» учебного пособия [672] подробно представлены все этапы кластеризации:

- выбор меры для оценки «соответствия» пар объектов в пространстве видов и подробный разбор многочисленных индексов сходства (по Жаккару, Сьёренсену, Чекановскому и др.);
- построение матрицы сходства выборочных совокупностей R ;
- выделение кластеров с использованием корреляционных плеяд Терентьева и традиционных иерархических дендрограмм.

Не хватает только одного важнейшего звена – с помощью каких, собственно, показателей можно оценить уровень β -разнообразия. Впрочем, таких предположений нет и в других, известных нам литературных источниках, кроме исходного указания Р. Уиттекера на необходимость измерять β -разнообразие в полусменах (half change) – смены половины видового состава вдоль некоторого экологического градиента [2718].

В экологической литературе некоторым «брендом» методического качества работы является неизменное приведение и попытка истолкования дендрограмм *кластерного анализа*, полученных с использованием меры сходства по Сьёренсену, либо иных подобных формул. Вряд ли можно придумать более примитивную схему математической обработки, механистически заимствованную из флористики начала прошлого века.

Во-первых, при расчете меры сходства между пробами нет никакой нужды сводить количественную шкалу, в которой измерено подавляющее большинство гидробиологических показателей к информативно более ослабленной номинальной шкале. Слишком много труда экологов и гидробиологов вкладывается в подсчет значений видовых численностей гидробионтов, чтобы потом огрублять исходные данные в мере Сьёренсена до статистически сомнительного факта простой встречаемости видов.

Во-вторых, специфика детерминированного кластерного анализа связана с невозможностью использования полного объема повторяющихся данных для одного и того же

местообитания [105], так как метод работает только с непересекающимися и необъединяемыми строками исходной матрицы наблюдений (см. табл. 2). Если для каждого классифицируемого объекта (станции, створа, участка, реки и т.д.) сделано *более чем одно* измерение, то перед формированием матрицы сходства R все пробы одного местообитания надо предварительно как-то адекватно «усреднить». Общепринятой экологической практикой является, как было описано в разделах 4.4 и 4.5, простое «сложение» списка видов во всех пробах по местообитаниям. Поскольку видовой состав местообитания стремительно расширяется при увеличении количества взятых проб, то полученная матрица сходства R отражает скорее степень нерегулярности проведения экспедиционных исследований, чем реальную оценку экологической изменчивости анализируемых местообитаний. Все это значительно обесценивает достоверность результатов последующего кластерного анализа.

Вероятностные методы оценки расстояний между объектами, в отличие от детерминированных формул мер сходства, позволяют обрабатывать весь массив исходных данных, допускают наличие ошибок и неполноту знаний о сравниваемых объектах и оперируют с плотностью распределения вероятностей переменных [105]. Было показано [3005], что наилучшей оценкой расстояния между центрами тяжести двух областей с векторами средних значений $m(P_1)$ и $m(P_2)$ является обобщенное расстояние Махалонобиса, а в качестве меры сходства двух подмножеств наблюдений целесообразно использовать T -критерий Хотеллинга.

Далее легко сделать вывод, что искать вид математического выражения для параметра β -разнообразия в дебрях дендрограмм иерархического кластерного анализа – мало перспективное занятие. Предположим, что нам удалось сформировать адекватную матрицу расстояний R для $m = 150$ местообитаний и среди многочисленных алгоритмов классификации нам попался такой, которому мы склонны полностью доверять. Но, получив развесистое дерево агрегирования объектов из десятка ярусов и нескольких сотен ветвей, мы неизбежно остановимся в нерешительности: а какое конкретное свойство дендрограммы принять за оценку β -разнообразия. Некоторые исследователи рекомендуют «проползти по дереву с линейкой» и измерить относительную длину ветвей на каком-то заранее оговоренном уровне (срезе), но аргументы против этого подхода настолько очевидны, что нет особой нужды их формулировать.

В то же время основные идеи факторного анализа и многомерного шкалирования дают широкий простор для выбора вполне обоснованного критерия разрешающей способности матрицы R . Например, оценка обобщенной дисперсии «соответствия» между сравниваемыми объектами, а следовательно, мера β -разнообразия экологического сообщества может быть рассчитана как отношение максимального и некоторого базового собственного значения симметричной матрицы сходства R [1687]:

$$V(R) = \lambda_{\max} / \lambda_k,$$

где λ_k – k -е собственное значение в ранжированном ряду, $k \cong 5$. Еще более обоснованным является использование одного из вариантов *показателя стресса*, отражающего уровень расхождений между матрицами исходных и результирующих расстояний при многомерном шкалировании [939].

Каждый из двух этапов оценки биоразнообразия (т.е. внутри и между местообитаниями) по отдельности имеет значительную методологическую и практическую ценность, но остается открытым вопрос: каким образом можно на основе этих различных по смыслу показателей оценить общее биоразнообразие совокупности местообитаний?

4.6. Оценка биоразнообразия с учетом «емкости» биоценоза

Можно задаться таким вопросом: почему, говоря о разнообразии, не учитывается такое важнейшее понятие, как «емкость» биоценоза, т.е. максимально возможное число видов (или экологических групп), способных существовать в этих условиях среды? Ведь в таком случае более разнообразной будет не тот биоценоз, в котором встречено больше ви-

дов (или экологических группировок), а тот где число этих видов (группировок) будет приближаться к максимально возможной «емкости».

К настоящему времени убедительно доказана несостоятельность концепции «избыточных», второстепенных видов [2581], исчезновение которых считалось вполне допустимым и безопасным для экосистемы. Выяснилось, что нельзя произвольно исключать из сообщества и считать «ненужными» какие-либо виды только на том основании, что они имеют сравнительно низкие популяционные показатели (численность, встречаемость) или не проявляют очевидного взаимодействия с другими видами [989]. Любое, даже внешне незначительное изменение видового состава биоты может спровоцировать сукцессионные процессы или нарушить стабильность природных экосистем. Поэтому важнейшим биоиндикационным показателем для выяснения последствий антропогенного воздействия является количественная мера изменения исходного видового состава сообщества, т.е. доля уровня текущего биоразнообразия по отношению к некоторому эталонному уровню, характеризующему «емкость» биоценоза.

Оценить степень «наполнения биоценозной емкости» можно, видимо, единственным способом: путем сравнения видового состава в двух состояниях, т.е. при наличии изучаемого воздействия и в его отсутствии, причем обязательно при прочих равных условиях. Только в этом случае все различия наблюдаемых значений этих характеристик можно считать обусловленными именно изучаемым воздействием, а не какими-либо иными неучтенными факторами. В отечественной экологической литературе [3056] принята следующая терминология для идентификации этих двух состояний:

- «импактные условия» (англ. *impact conditions*), при наличии воздействия, где имеют место импактные значения абиотических факторов и импактное состояние биоценоза;
- «фоновые условия», которые полностью исключают само изучаемое воздействие, а во всем остальном идентичны первым.

На первый взгляд, меру изменения видового состава сообщества легко установить, просто сравнив список всех видов, особи которых были обнаружены в сообществе фонового биотопа и в идентичных биотопах в импактных условиях антропогенного воздействия. Следует предварительно отметить, что ни один из представленных выше индексов биоразнообразия не только не ориентирован на подсчет редких видов, но и просто игнорирует их существование. Меры сходства, по Съеренсену, Жаккару и другими, использующие только величины 0 и 1, страдают крайностью противоположного свойства: факт наличия или отсутствия случайных (лабильных) видов имеет такую же значимость, что и для ключевых доминантов, исчезновение которых коренным образом меняет экологическую ситуацию в биотопе.

Существенным ограничением является также симметричный характер традиционных мер подобия, в то время как при сравнении списков видов значимость логического события «вид есть в фоновых условиях, но отсутствует в импактных» неравнозначна событию «вид есть в импактных условиях, но отсутствует в фоновых».

Пусть U – перечень из n видов, зарегистрированных в фоновых условиях, который мы будем полагать за эталон «биоценозной емкости». Аналогично V – перечни видового богатства для каждого из местообитаний, предположительно находящихся в импактных условиях, по которым проверяется гипотеза о снижении биоразнообразия. Для каждого j -го вида, выделенного в фоновых условиях, определим индекс его «доминирования в биоценозе» IP_j , являющийся наиболее обоснованной, на наш взгляд, модификацией из семейств индексов, предложенных В.А. Брочкой и Л.А. Зенкевичем [213]:

$$IP_j = b_j \ln(n_j + 1)_{cp},$$

где b_j – встречаемость; $b_j = m_j / m$, m_j – число проб, в которых был найден вид j , m – общее число проб, из которых формируется перечень видов U ; $\ln(n_j + 1)_{cp}$ – средняя из логарифмов численностей особей j -го вида.

В системном анализе и математической лингвистике наиболее известной метрикой

сходства является *функция Левенштайна* L , которую также называют расстоянием редактирования [3624]. Значение L равно минимальному количеству операций, которые преобразуют один список в другой, если считать все типы операций, такие, как удаление, добавление и замены строки списков равноправными. Особый интерес представляет метрика Левенштайна, если стоимости различных типов ошибок не совпадают.

Рассмотрим в общем виде формальный подход к определению расстояния между списками видов с использованием метрики Левенштайна. Под списком будем понимать последовательность строк некоторого словаря A . Расширим словарь A пустой строкой ϵ .

Определение 1. Выравниванием двух списков U из m строк и V из n строк будем называть два списка \bar{U} и \bar{V} длиной l , $l = \max(m, n)$, которые получаются из исходных списков U и V в результате добавления пустых строк ϵ таким образом, чтобы в обоих списках \bar{U} и \bar{V} не было пустых строк в одинаковых позициях.

Пусть для любой пары строк α и β из множества $A \cup \epsilon$ определена стоимость $\delta(\alpha \rightarrow \beta)$ замены строки α на β . Введение пустой строки ϵ и процедуры выравнивания списков позволяет свести любые операции удаления и добавления к единой операции замены. В частности, замена $\alpha \rightarrow \epsilon$ соответствует удалению α , а $\epsilon \rightarrow \beta$ – вставке β .

Определение 2. Мерой Левенштайна различия двух списков U и V будем называть минимальную общую стоимость операций по их выравниванию с учетом частной стоимости операции над каждой парой строк:

$$L = - \min \{ \delta(\bar{U}, \bar{V}) \} = - \sum_{i=1}^l \delta(\bar{u}_i \rightarrow \bar{v}_i).$$

С использованием индексов доминирования видов IP_j определим стоимости операций выравнивания для подсчета функции Левенштайна следующим образом:

- для $\alpha \rightarrow \beta$ и $\epsilon \rightarrow \epsilon$ примем $\delta = 0$, т.е, если вид одновременно присутствует (или отсутствует) в фоновых и импактных условиях, то стоимость такого события равна нулю;
- для операции выравнивания $\alpha \rightarrow \epsilon$ стоимость определяется как $\delta = k \cdot IP_j$ для фона, т.е. если в импактных условиях имеет место исчезновение вида, зарегистрированного в фоновом биотопе, то на это событие налагается штраф, пропорциональный индексу плотности вида в фоновых условиях $k \cdot IP_j$, где k – повышающий коэффициент, который мы приняли равным 3;
- стоимость операции $\epsilon \rightarrow \beta$ определяется как $\delta = (0 - IP_j)$ для импактного биотопа, т.е. если в импактных условиях имеется вид, отсутствующий в фоновом биотопе, то вычисляемое расстояние уменьшается на величину IP_j индекса плотности вида в импактных условиях.

График изменения функции Левенштайна, вычисляемой по описанному алгоритму и показывающей отличие видового состава макрозообентоса для каждой из 14 станций р. Сок от фоновых условий в р. Байтуган, представлен на рис. 9. Чем больше кривая L приближается к оси абсцисс, тем ближе видовое разнообразие тестируемого местообитания к «биоценозной емкости». На построенном графике в значительно более отчетливой мере, чем с использованием традиционных биотических показателей, показано относительное снижение биоразнообразия в районе станций 7-9 и 11, обусловленное воздействием антропогенных факторов.

4.7. Расчет показателей γ -биоразнообразия по сумме описаний

Предположим, что нами найден вектор осредненных абсолютных численностей по списку видов, характерному для некоторого подмножества измерений. Но что является в этих условиях показателем γ -разнообразия: совокупное биоразнообразие биотопа или среднее от мгновенных оценок разнообразия?

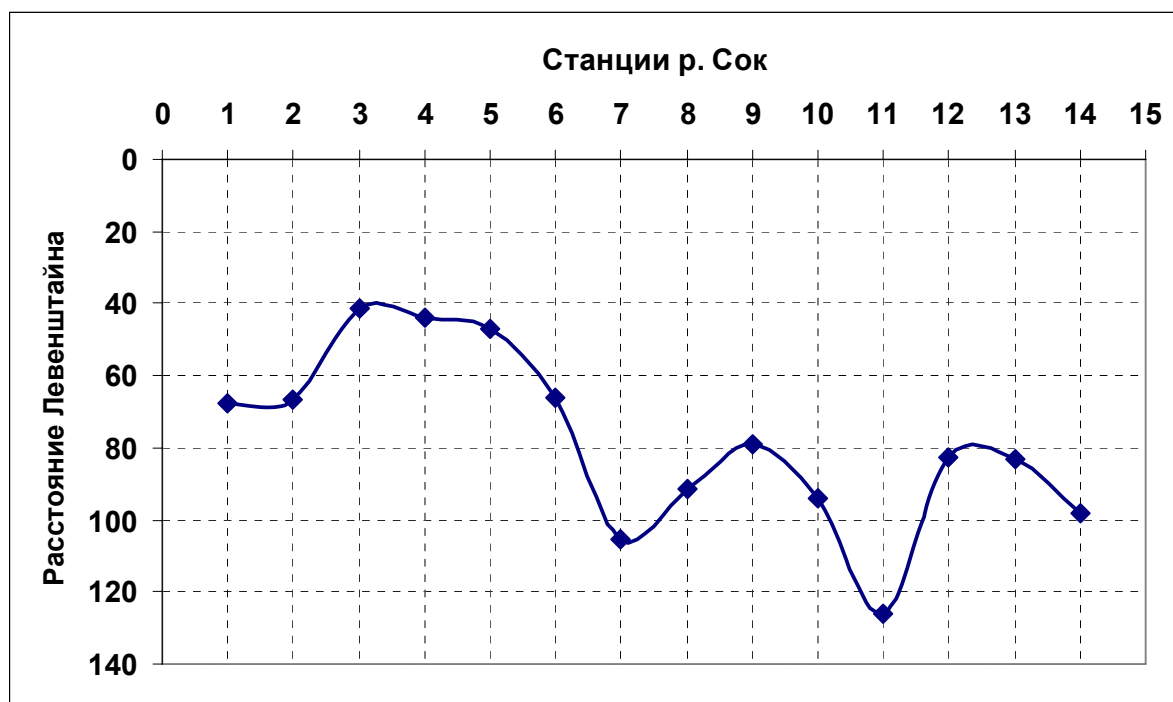


Рис. 9. График отличий видового состава по станциям р. Сок от фоновых условий

Существующие в литературе характеристики γ -разнообразия мало способствуют его пониманию. В основополагающем документе [3517] отмечено только, что «экосистемное разнообразие (т.е. гамма-разнообразие) труднее измерить, чем видовое или генетическое, поскольку границы сообществ и экосистем менее дискретны». Однако проблема не только в трудностях поиска границ, но и в трактовке концепции аддитивности.

Согласно Д. Шультру и Р. Риклефсу [4249], γ -разнообразие может быть рассчитано следующим образом:

$$\gamma = \alpha \times \beta \times n,$$

где n – общее число местообитаний или ключевых участков, α и β – среднее соответствующего компонента разнообразия на ландшафтную единицу. Очевидно, в этом уравнении γ -разнообразие может увеличиваться, когда α - и β -разнообразия могут остаться постоянными. В любом случае совокупное разнообразие больше разнообразия в любом из составляющих местообитаний.

Диаметрально другой подход для оценки γ -разнообразия заключается в вычислении сложной энтропии:

$$H'' = -\sum_i^M P_i \sum_j^N p_{ij} \log_2 p_{ij},$$

где P_i – априорная вероятность появления i -го объекта (местообитания); p_{ij} – вероятность отнесения образца i -го объекта к j -му таксона. При $i = 1$ $P_j = 1$ и формула сводится к простой энтропии Шеннона. Поскольку при $i > 1$ $P_j < 1$, то γ -разнообразие не может превышать разнообразие в любом из составляющих местообитаний.

Продемонстрируем все изложенное выше на конкретном примере (см. табл. 6). Вниманию предлагается два варианта расчета биоразнообразия по Шеннону:

- первое значение $H = 4,02$ рассчитано из предположения, что биоразнообразие местообитания составляют 63 вида с численностями, равными средним значениям для 7 проб;
- второе значение $H = 2,01$ рассчитано по формуле сложной энтропии, которая из предположения о равновероятности всех проб сводится к среднему арифметическому из индексов Шеннона для каждого наблюдения.

Фрагмент таблицы данных по семи пробам макрозообентоса, взятых в разное время на ст. 3 р. Байтуган

№ п/п	Наименование вида	Средняя числен- ность	Дата взятия проб и численность видов						
			17.07. 1991	18.07. 1991	19.07. 1991	24.09. 1992	30.07. 1992	28.06. 1993	29.07. 1999
1	Arachnoidae	6,43		30	15				
2	<i>E. ponderosa</i>	5,71	40						
3	<i>Euglesa</i> sp.	11,43	80						
4	<i>Pisidium inflatum</i>	11,43	80						
5	<i>Pisidium</i> sp.	0,71			5				
6	Ceratopogonidae	13,57	80	10	5				
7	<i>Culicoides</i> sp.	11,43			80				
8	<i>Brillia modesta</i>	1,14							8
9	<i>Chironomus</i> sp.	5,71	40						
10	<i>Corynoneura</i> sp.	4,29		30					
11	<i>Cricotopus bicinctus</i>	5,43		20			10		8
12	<i>Cricotopus (Isocladius)</i> sp.	1,43					10		
13	<i>Cricotopus</i> gr. <i>tremulus</i>	1,43					10		
14	<i>C. (Isocladius) trifasciatus</i>	54,86	160						224
15	<i>Eukiefferiella claripennis</i>	1,14							8
16	<i>E. gr. gracei</i>	72,00							504
17	<i>Eukiefferiella</i> sp.	1,14							8
18	<i>Mesocricotopus</i> sp.	23,57	160		5				
19	<i>Micropsectra</i> gr. <i>praecox</i>	24,29	160		10				
20	<i>Micropsectra</i> sp.	6,86							48
21	<i>Monodiamesa bathyphila</i>	11,43	80						
22	<i>Nanocladius rectinervis</i>	2,29							16
23	<i>Odontomesa fulva</i>	31,14	200				10		8
24	<i>Orthocladius</i> sp.	3,57			25				
25	<i>Orthocladius (Orth.) clarki</i>	1,14							8
26	<i>Paracladopelma</i> gr. <i>camptolabis</i>	5,00			35				
27	<i>Paracladius alpicola</i>	2,29							16
28	<i>Procladius choreus</i>	2,86				20			
52	<i>Dicranota bimaculata</i>	62,14	80		15	280	60		
53	Nematoda	2,86		20					
54	Oligochaeta	245,86	560	40	35	220	850		16
55	<i>Leuctra fusca</i>	5,71		40					
56	<i>Nemoura cinerea</i>	0,71			5				
57	<i>Nemoura</i> sp.	5,71	40						
58	<i>Pericoma</i> sp.	5,71	40						
59	Simuliidae	268,29		150	480				1248
60	<i>Simulium</i> sp.	8,57				10	30	20	
61	<i>Nemotelus</i> sp.	1,43				10			
62	<i>Odontomyia</i> sp.	7,14		50					
63	<i>Limnophilus flavicornis</i>	2,86		20					
Индекс Шеннона		$\frac{4,02}{2,01}$	3,5	1,99	1,45	2,97	1,89	0,59	2,24
Всего видов S		14,4	20	19	19	14	10	2	17

Каждый из этих двух подходов имеет своих авторитетных сторонников, хотя лично нам вторая версия расчета представляется гораздо более симпатичной. Однако на уровне простых логических рассуждений ни обосновать, ни опровергнуть любую из представленных гипотез оценки γ -разнообразия невозможно.

При оценке γ -разнообразия, как и в других описанных выше случаях, остается актуальной задача гармоничного учета относительной доли как ключевых для функционирования экосистемы обильных доминирующих видов, так и редких и малочисленных видов, являющихся компонентом сообщества, наиболее чувствительным к изменениям факторов среды. Для решения этой проблемы используются различные уравнения для оценки совокупного разнообразия, но наиболее адекватными считаются нелинейные (логарифмические) функции, повышающие значение мало распространенных элементов и, наоборот, снижающие (подавляющее) значение широко распространенных фоновых компонент, а также позволяющие вводить условные определители [576].

Некоторые малоутешительные выводы

1. Оценка биоразнообразия в пространстве видов в значительной мере некорректна, в первую очередь потому, что никак не учитывается морфологическое, функциональное, экологическое сходство/различие между самими видами. Каждый вид представляется как изолированный таксон, информационно равноудаленный от всех остальных.
2. Индексы разнообразия, базирующиеся на лучшей (по некоторым данным) мере энтропии по Шеннону, основываются на экологически сомнительной концепции: в качестве эталона принимается экосистема с равными обилиями всех видов, что не вполне соответствует структуре реальных природных сообществ.
3. Модели разнообразия, основанные на тех или иных аналитических формулах распределения популяционной плотности, остаются лишь интерпретацией весьма частных гипотез их авторов и не могут служить фундаментом для формальной строгой методики оценки биоразнообразия.
4. Приходится констатировать, что со времен Р. Уиттекера, заложившего вербальную основу понятий разнообразия для совокупности местообитаний, сколько-нибудь строгих методов количественной оценки β - или γ -разнообразия до настоящего времени не выработано.
5. Отсутствуют также общепринятые методические разработки для расчета средней популяционной плотности и списка характерных видов произвольного сообщества организмов, которые должны являться неотъемлемыми «кирпичиками» построения любой количественной концепции биоразнообразия.

Чтобы как-то компенсировать несколько нарочитый негатив наших выводов, приведем фразу Мишеля Монтеня: *«Для тех, кто никуда не плывет, не бывает попутного ветра»*.

В заключение авторы считают своим приятным долгом выразить искреннюю благодарность д.б.н., проф. Протасову А.А., д.б.н., проф. Шуйскому В.Ф., д.б.н. Зинченко Т.Д. и к.б.н., доц. Крамаренко С.С., которые внимательно прочитали рукопись и дали ценные замечания и своевременные советы.