

*А.В. Мощенко, Т.А. Белан, Б.М. Борисов, Т.С. Лишавская, А.В. Севастьянов
ФГБУ «ДВНИГМИ», г. Владивосток*

К МЕТОДИКЕ ВЫДЕЛЕНИЯ СООБЩЕСТВ МАКРОЗООБЕНТОСА: ПРОСТОЙ АЛГОРИТМ ДЛЯ ПРАКТИКИ ЭКОЛОГИЧЕСКОГО МОНИТОРИНГА

Введение

Классификация объектов и явлений материального мира является основой, а часто и целью множества научных исследований в самых разных областях знания. Один из авторов работы познакомился с кластерным анализом еще в начале 80-х годов прошлого века и использует его, а также дискриминантные, факторные и другие многомерные процедуры по настоящее время. В процессе работы классифицировались гранулометрические фракции, океанографические элементы, загрязняющие вещества, виды гидроидов, двустворчатых моллюсков и даже результаты иммунохимического анализа, оценивалось распределение и изменчивость миллепор, полихет, планктонных комплексов, таксоценов и сообществ рыхлых и твердых грунтов, включая обрастание корпусов судов и гидротехнических сооружений [2; 5; 7].

Естественно, ценность классификации повышается, если ее результаты удаётся выразить числом и подтвердить статистическим анализом. При этом исследователь приобретает уверенность в своих дальнейших действиях по описанию полученных групп и их свойств, а также при последующих манипуляциях с этими группами.

Оценка экологического состояния акватории на основе анализа сообществ макрозообентоса относится к многомерным (multivariate) методам мониторинга морской среды [21]. Залог успеха такого анализа – ясное и четкое представление того, что следует считать сообществом и возможность их простого и оперативного выделения.

Сообщество можно определить как «группу организмов, обитающих в определенной среде, предположительно взаимодействующих друг с другом и со средой, и которая может быть отделена от «других групп методами экологического исследования» [12]. По образному выражению Х.-Ф. Будуреску каждое сообщество представляет собой «относительный континуум между относительными перерывами континуума» или узел – скопление точек в некоем гиперпространстве, полностью или частично изолированное от других подобных скоплений [12]. Группы видов, одинаково реагирующих на факторы среды – некие узлы в многомерном континууме – это сообщества, т.е. участки континуума, ограниченные пространством перерыва континуума – экотонном (рис. 1).

Под термином сообщество (как синонимы – ассоциация, группировка) макрозообентоса авторы подразумевают совокупность донных животных второго и третьего трофического уровня, которая населяет определенный биотоп и характеризуется определенными количественными соотношениями между видами [13]. По определению, для того, чтобы выделить некую группировку и доказать что эта ассоциация является сообществом, необходимо:

1. Расклассифицировать станции (пробы) на группы и оценить статистическую значимость этой классификации;
2. Определить, являются ли выделенные ассоциации географически или экологически детерминированными со статистической оценкой значимости влияния географических и/или комплекса экологических факторов;
3. Назвать эти группировки, исходя из того, что сообщество – это континуум;
4. Выявить дифференциацию по отношению к факторам среды (населяет определенный биотоп) и оценить значимость такой дифференциации с позиций статистики.

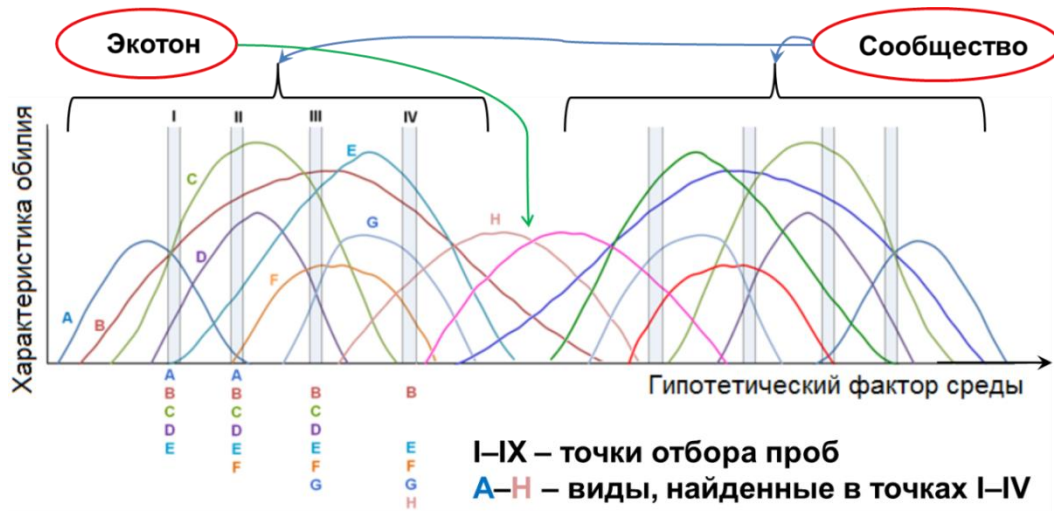


Рис. 1. Схематическое изображение изменения состава донного населения вдоль градиента гипотетического фактора среды (по [1], модифицировано)

Таким образом, целью работы является создание алгоритма, который, с одной стороны, был бы основан на принципе «необходимо и достаточно», а с другой – позволил бы относительно быстро и с минимальными «энергетическими» затратами выделять сообщества макрозообентоса. При этом каждый шаг при выполнении этой процедуры должен иметь «численное» выражение и, по необходимости, быть доказанным с позиций статистики.

Материалом для построения и отладки алгоритма послужили результаты съемок авторов, выполненные в 1992–2019 гг. в различных районах залива Петра Великого Японского моря. Это – заливы Посьета, Стрелок, Амурский и Уссурийский, акватория к северу от устья реки Туманной, бухты Рифовая, Золотой Рог, Диомид, Патрокл, Аякс и Парис, пролив Босфор Восточный (рис. 2). Методы отбора проб и их обработки были ранее неоднократно опубликованы [9; 17; 19].

Первичная подготовка данных. Для анализа потребуются так называемые видовые матрицы, в которых ряды начинаются с названия таксона, столбцы – с кода станций и/или проб, а ячейки заполнены величинами биомасс или плотностей поселения (соответственно в граммах и количестве экземпляров на квадратный метр) (рис. 3). Кроме того, необходимы координаты станций (проще – в десятичном формате) и результаты измерений факторов среды, если таковые

имеются. Удобнее всего сделать один файл в формате Excel (с расширением *.xlsx), включающий листы с биомассами, плотностями, координатами и факторами; в процессе анализа в эту книгу придется добавлять некоторые другие данные. Далее в качестве иллюстрации применения различных процедур использованы результаты съемки 2016 г.

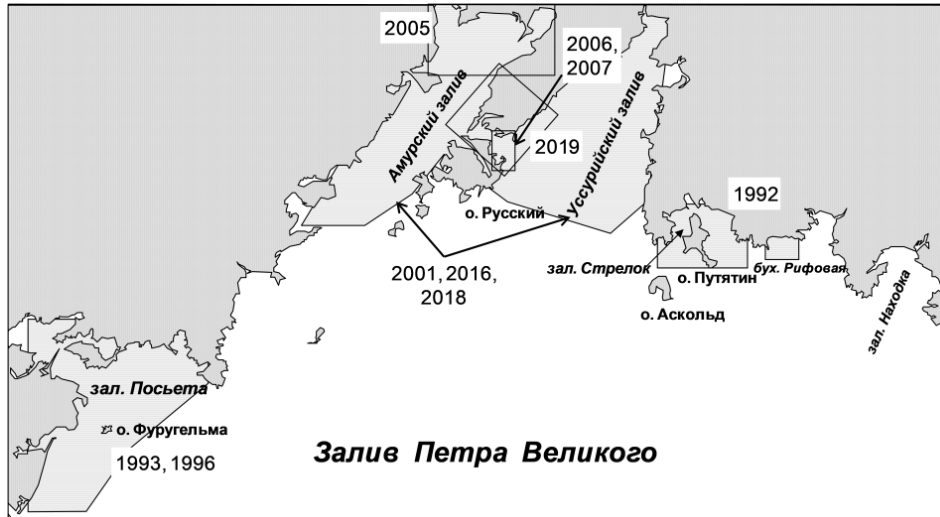


Рис. 2. Расположение районов работ в заливе Петра Великого в 1992–2019 гг.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
1	Таха	u11	u13	a16	u17	u20	u21	a26	a2	a34	a35	u40	u43	a47	a4	a50	u55	u57	u59	a62	u6	a9	Z1	Z11	Z12
2	Acanthomyx stelleri	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0
3	Actinaria fam. gen. sp.	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0	0	0
4	Ampelisca macrocephala	0	0	0	5	5	0	0	0	0	0	10	0	15	0	5	0	0	0	0	0	0	0	0	0
5	Ampharete sibirica	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0
6	Ampharete sp.	0	0	0	0	0	0	0	0	5	0	0	0	5	0	5	5	10	0	0	0	0	0	0	0
7	Ampharetidae gen. sp.	0	0	0	0	0	0	0	0	0	0	0	0	0	15	0	0	0	0	0	0	0	0	0	0
8	Amphiodia fissa	0	0	0	0	0	0	0	0	15	0	0	0	0	0	5	0	5	65	5	0	0	0	0	0
9	Anadara broughtoni	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	Anonyx sp.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0
11	Aphelochaeta pacifica	0	0	0	0	0	0	985	10	0	255	0	0	25	15	30	0	0	0	5	0	260	0	590	550
12	Arabella iricolor	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	Arca boucardi	0	0	0	0	0	0	0	0	0	0	0	0	0	70	5	0	0	0	0	0	0	0	0	0
14	Aricidea (Acmira) catherinae	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	5	0	0	0	0	0	0	0	0
15	Ascidia fam. gen. sp.	0	0	0	0	0	0	0	5	0	0	0	0	0	20	0	0	0	0	0	5	15	0	0	0

	A	B	C		A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Point	Xc	Yc	1	Point	Depth	Corg	AMBI	TPF	APC	O2	PoF1	PoF2	MeF1	MeF2	OrgF1	OrgF2	OrgF
2	11u	132.1792	43.2000	2	11u	15	0.11	2.44	2.8	13.4	4.88	-0.345	-1.485	-0.322	-1.479	-0.651	-0.096	-0.615
3	13u	132.3167	43.2000	3	13u	12	1.66	1.68	2.6	45.6	4.89	-0.413	0.533	-0.566	0.538	-0.585	0.473	-0.279
4	16a	131.8958	43.1533	4	16a	18	3.14	2.26	2.8	51.5	2.96	-0.021	0.975	-0.119	0.983	0.197	-0.146	0.100
5	17u	132.0767	43.1533	5	17u	12	0.61	3.43	3.8	18.4	4.88	0.253	-1.074	0.638	-1.075	-0.582	0.635	-0.197
6	20u	132.2400	43.1533	6	20u	30	1.10	2.25	2.8	24.0	4.58	-0.233	-0.754	-0.380	-0.834	-0.346	-0.003	-0.303
7	21u	132.3000	43.1533	7	21u	18	0.77	2.86	1.8	51.7	5.15	-0.320	0.761	-0.450	0.784	-0.203	-0.012	-0.183
8	26a	131.8483	43.1200	8	26a	21	4.03	3.19	2.8	48.4	2.58	0.140	1.031	-0.301	0.826	0.986	-0.849	0.445
9	2a	131.8217	43.2417	9	2a	10	1.58	1.81	2.8	69.6	4.23	-0.580	1.798	-0.660	1.857	-0.146	-0.274	-0.262
10	34a	131.7600	43.0850	10	34a	15	1.00	2.14	2.2	14.8	4.27	-0.333	-0.893	-0.408	-0.888	-0.344	-0.287	-0.441
11	35a	131.8183	43.0850	11	35a	25	3.23	3.05	2.4	42.2	3.70	-0.144	0.523	-0.333	0.469	0.188	-0.293	0.021
12	40u	131.9708	43.0583	12	40u	27	1.94	2.07	1.8	20.5	5.73	0.013	-0.776	-0.072	-0.852	0.139	-0.426	-0.088
13	43u	132.1583	43.0583	13	43u	27	0.20	0.92	2.6	11.0	4.86	-0.380	-1.591	-0.348	-1.570	-0.864	0.189	-0.661
14	47a	131.5890	43.0250	14	47a	16	1.26	2.84	3.2	28.0	4.45	-0.533	-0.864	-0.433	-0.780	-0.860	0.065	-0.718
15	4a	131.9333	43.2417	15	4a	5	1.78	1.62	3.2	62.2	5.23	-0.544	1.485	-0.695	1.488	-0.122	-0.211	-0.210

Рис. 3. Фрагмент видовой матрицы, и листов книги Excel с координатами станций и факторами среды: данные 2016 г. Большинство параметров среды получено на основе применения факторного анализа [8; 9; 11]

Статистические программы и модули. Для выполнения алгоритма потребуется среда **R** (модули xlsx, pvclust, scaleboot, cluster, vegan, geosphere, MASS, функция pvclust_bcdist и, по желанию, графический интерфейс пользователя – GUI – Rcmdr). Функция pvclust_bcdist, как и сам алгоритм в виде скриптовой программы для **R**, размещена на сайте ДВНИГМИ <http://ferhri.org>. «Инструкцию» по скачиванию, установке и использованию **R** можно найти в замечательной книге Р.И. Кабакова [3], которая находится в интернете в «свободном» доступе.

Для ознакомления с «тонкостями» выполнения тех или иных процедур авторы использовали работы В.К. Шитикова, С.Э. Мастицкого, Г.С. Розенберга, Я. Оксанена и др. литературу [6, 15, 20]. Кроме того, желательно иметь ПП PRIMER, который значительно упрощает ряд статистических процедур (ANOSIM, SIMPER, вычисление экологических индексов); скачивание и установка его ранних версий из интернета обычно не вызывает серьезных затруднений.

Выделение сообществ. Импортируем подготовленные данные в среду R^* :

```
library(xlsx) (1),
```

```
workbook <- "d:/Название папки/.../Название файла.xlsx" (2),
```

```
A <- read.xlsx(workbook, n, row.names="xxx") (3).
```

где A – количественная характеристика вида, n – номер листа книги Excel, $row.names$ – столбец листа Excel, который содержит коды станций, названия видов или более крупных таксонов. Затем можно проверить то, что получилось:

```
fix(A) (4).
```

После просмотра данных, если, например, появились строчки или столбцы заполненные «NA», следует их удалить, можно не выходя из R , используя GUI Rcmdr. Далее необходимо преобразовать плотности и биомассы; обычно это $\log(x+1)$, корень второй или четвертой степени:

```
Alog <- log(A+1) (5),
```

```
A2 <- sqrt(A) (6),
```

```
A4 <- sqrt(sqrt(A)) (7).
```

Следующая процедура – бутстреп деревьев и расчет смещенных и несмещенных вероятностей появления узлов на дендрограмме. Обычно, в качестве метрики, мы используем коэффициент Брея-Кертиса и метод агломерации Варда, что впрочем, отнюдь не исключает других метрик и методов. Рекомендации по применению тех или иных процедур кластерного анализа можно получить, например, у Дж.-О. Кима с соавторами [4].

```
library(pvclust) (8).
```

Модуль `pvclust` не вычисляет коэффициент сходства Брея-Кертиса, поэтому включаем функцию `"pvclust_bcdist"` (открываем скрипт `pvclust_bcdist.R`, выделяем все, запускаем) и переходим к выполнению команд:

```
A.bcdist.pv_fit <- pvclust(A4, method.hclust="ward.D2",
method.dist="bray-curtis", n=1000, r=seq(.5,2,by=.1)) (9),
```

```
plot(A.bcdist.pv_fit) (10).
```

Кроме метода `ward.D2` возможно использование `average`, `single`, `complete`, `mcquitty`, `median` или `centroid`; в качестве метрик – `maximum`, `manhattan`, `euclidean`, `canberra`, `binary` или `minkowski`, `abscor`, `correlation`, `uncentered`.

* Чтобы команда (1) сработала, на компьютере должна быть установлена программа Java. Если название какой-либо папки или файла написано кириллицей, следует переключить клавиатуру на кириллицу.

В результате данной процедуры, выполненной на основе биомасс видов, на уровне 95-% несмещенных бутстреп-вероятностей станции объединяются в 5 групп I–V (рис. 4)*. Еще один кластер IV выделяется с вероятностью 92 %, а станции a2, a4, a9 и a16 остаются нерасклассифицированными. В дополнение можно отдельно получить смещенные и несмещенные вероятности для узлов деревьев в виде таблицы (результаты не приводятся):

```
library(scaleboot) (6),
```

```
fm <- sbfit(A.bcdist.pv_fit) (7),
```

```
summary(fm) (8).
```

Классификация нерасклассифицированных станций. Наличие шести плохо или совсем нерасклассифицированных точек опробования – не слишком хороший результат для выборки из 30 станций, хотя и объяснимой с позиции современных представлений о сообществах (возможно, например, они попали в области экотонов). Поэтому, определим, к каким группам эти станции сильнее всего «притягиваются» с использованием алгоритма нечеткой классификации. Смысл этой процедуры заключается в том, что объект (в нашем случае – станция) может принадлежать к нескольким группам, но тяготеть к ним с разной силой. Подключаем модули:

```
library (cluster) (9),
```

```
library (vegan) (10).
```

Транспонируем полученные ранее данные и вычисляем матрицу дистанций (метрика "bray-curtis"):

```
At4 <- t(A4) (11),
```

```
A.ch <- vegdist(At4, "bray") (12).
```

Кластеризуем данные на основе этой матрицы, строим диаграмму рассеяния и вычисляем "силу притяжения" (степень принадлежности) каждой станции к каждому кластеру, число которых определяется на основе результатов предыдущей процедуры (пять отчетливых групп и одна «сырая», всего шесть).

```
Z2.f <- fanny(A.ch, 6, memb.exp = 1.2) (13).
```

Параметр "memb.exp" (экспоненциальный вес) подбирается, его «штатный» диапазон – 1.2–2; чем этот вес больше, тем нечеткое разбиение более "размазано" [15].

```
clusplot(Z2.f, labels = 2) (14),
```

```
res.fanny <- edit(data.frame(Z2.f$membership)) (15).
```

В результате, все станции оказались расклассифицированными и группируются в 6 кластеров (рис. 5, табл. 1):

I. Бухты Золотой Рог и Диомид – Z1, Z7, Z11, Z12 и Z22;

* Полученные на разных этапах выполнения алгоритма диаграммы и др. результаты экспортируем в удобную для Вас программу и редактируем (конечно, если Вы любитель писать "заклинания", все это можно сделать и в среде R, изучив замечательную книгу Роберта И. Кабакова "R в действии...").

- II. Пролив Босфор Восточный и устье бухты Улисс – Z14, Z18, Z19 и Z23;
- III. Относительно открытые районы Амурского и Уссурийского заливов – a34, a47, a50, a62 и u11, u20, u40, u43, u55, u57, u59;
- IV. Прибрежье полуострова Эгершельда и центральная северная часть Амурского залива – a9, a26 и a35;
- V. Западное и восточное побережье Уссурийского залива – u17 и u21;
- VI. Внутренние акватории Амурского и Уссурийского заливов – a2, a4, a16, u6 и u13.

При этом станции a4, a34, Z22 и Z1 – довольно яркие примеры разной степени принадлежности, заметно тяготея, соответственно к группам IV, VI, IV и III (для большей убедительности – увеличьте "memb.exp").

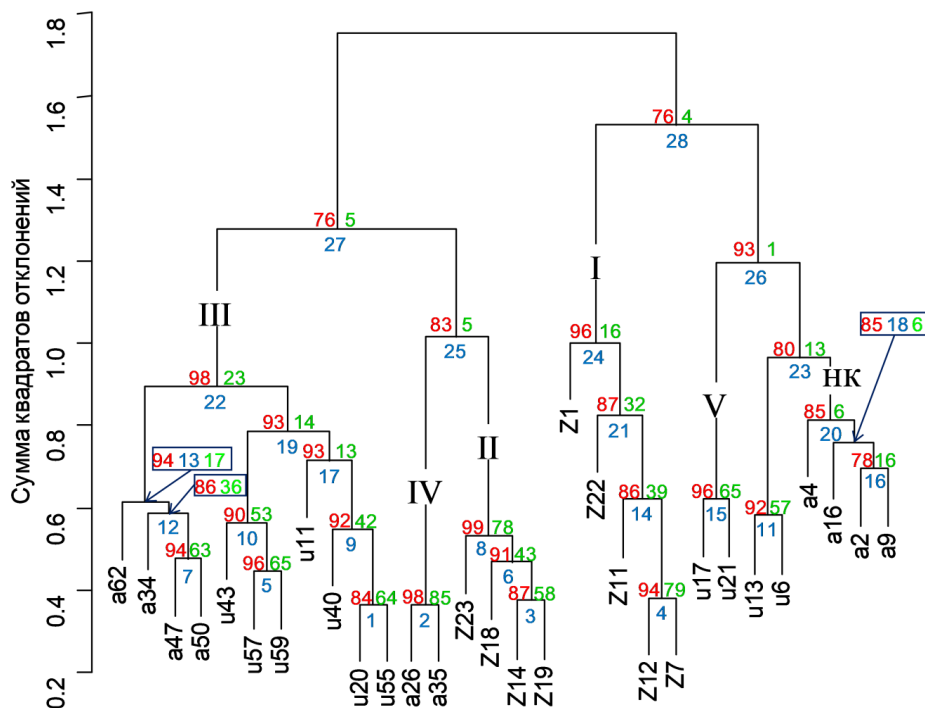


Рис. 4. Дендрограмма сходства станций по видовому составу макрозообентоса, полученная методом Варда (Ward.D2, метрика – коэффициент сходства Брея-Кертиса) на основе биомассы животных (преобразование – корень четвертой степени). В узлах дендрограммы приведен порядковый номер агломерации, приближенные и несмещенные оценки вероятности появления узла (соответственно синим, зеленым и красным цветом). Римские цифры – ветви с достоверной вероятностью >95%, нк – нерасклассифицированные станции

Использование в процедуре кластеризации плотности поселения и биомассы видов, для данных 2016 г., дает идентичные результаты, различающиеся только порядком агломерации станций в полученных кластерах и бутстреп-вероятностями. Аналогичные вариации наблюдаются и при применении разных методов, например – групповой средней (результаты не приводятся). Однако, такая «идентичность» наблюдается не всегда (в семи случаях из восьми проанализированных), поэтому желательно проводить вычисления и по биомассам и по плотностям. При таком развитии событий необходимо либо «волевое» решение, либо внимательный анализ других результатов, например процедуры SIMPER.

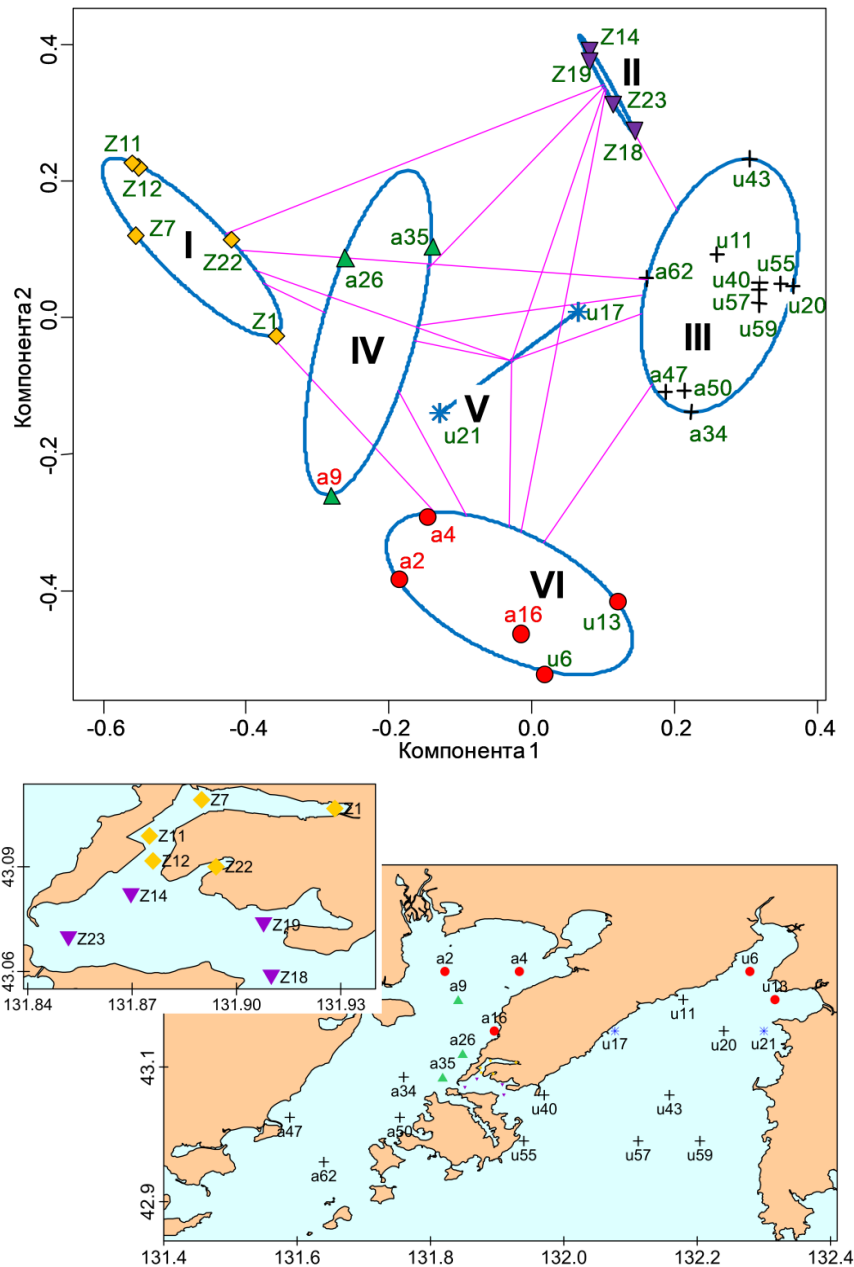


Рис. 5. Группировка станций с использованием алгоритма нечеткой классификации (компоненты 1 и 2 объясняют 31,3 % изменчивости) и распределение станций выделенных групп на исследованной акватории. Красным цветом на диаграмме рассеяния выделены номера нерасклассифицированных станций на рисунке 4; розовые линии соединяют групповые центры

ANOSIM*. Данная процедура выполняется для статистической оценки принадлежности станций к той или иной группе или, иными словами, для проверки достоверности появления фрагментов древовидной структуры. В имеющейся книге Excel создаем новый лист с результатами классификации (например: полученные группы – А, В, С, D, E, F; название колонки – Group), причем порядок положения станций в столбце должен точно совпадать с таковым в At4, чего можно добиться, сортируя данные. Импортируем его в среду **R**, как было показано выше, например, под именем Cluster, и переходим к анализу:

$$A.ch \leftarrow \text{vegdist}(At4, "bray") \quad (16),$$

* Проще всего ее выполнить в ПП PRIMER.

```

data(Cluster) (17),
attach(Cluster) (18),
A.ano <- anosim(A.ch, Group) (19),
summary(A.ano) (20).

```

Результаты процедуры ANOSIM свидетельствуют, что выделение кластеров I–VI, как по плотности, так и по биомассе, в большинстве случаев значимо с позиций статистики (табл. 2). Исключение составляют лишь те пары групп, в которых количество станций в сумме не превышает шести. О достоверности полученной классификации в целом говорят статистически значимые величины общей статистики *R*.

Таблица 1

Результаты нечеткой классификации – степень принадлежности станций к выделенным кластерам

Станция	Кластер					
	I	II	III	IV	V	VI
Z1	0,786	0,028	0,053	0,075	0,025	0,032
Z11	0,987	0,003	0,002	0,002	0,002	0,004
Z12	0,997	0,001	0,001	0,001	0,000	0,001
Z22	0,798	0,031	0,030	0,046	0,028	0,066
Z7	0,989	0,002	0,002	0,003	0,001	0,003
Z23	0,001	0,983	0,010	0,002	0,001	0,002
Z14	0,001	0,992	0,005	0,001	0,000	0,001
Z18	0,001	0,982	0,013	0,002	0,001	0,001
Z19	0,000	0,996	0,003	0,000	0,000	0,001
a34	0,010	0,054	0,752	0,133	0,015	0,037
a47	0,009	0,036	0,826	0,096	0,008	0,025
a50	0,006	0,025	0,886	0,058	0,005	0,019
a62	0,010	0,055	0,842	0,037	0,004	0,051
u11	0,017	0,055	0,841	0,044	0,026	0,017
u20	0,001	0,010	0,980	0,006	0,002	0,002
u40	0,006	0,048	0,904	0,029	0,005	0,008
u43	0,011	0,058	0,894	0,018	0,010	0,009
u55	0,003	0,017	0,959	0,013	0,004	0,004
u57	0,004	0,024	0,938	0,020	0,008	0,005
u59	0,003	0,026	0,938	0,021	0,004	0,008
a16	0,008	0,007	0,036	0,913	0,013	0,022
a2	0,010	0,005	0,015	0,942	0,005	0,023
a4	0,044	0,027	0,079	0,768	0,021	0,061
u13	0,006	0,005	0,040	0,939	0,004	0,005
u6	0,003	0,002	0,013	0,974	0,003	0,005
u17	0,000	0,000	0,001	0,001	0,997	0,000
u21	0,000	0,000	0,000	0,001	0,998	0,000
a26	0,001	0,001	0,001	0,001	0,000	0,996
a35	0,000	0,001	0,001	0,001	0,000	0,997
a9	0,009	0,003	0,008	0,029	0,002	0,949

Примечание. Жирным шрифтом выделены станции, нерасклассифицированные при помощи кластер-анализа и максимальные значения «силы притяжения».

**Результаты процедуры ANOSIM (общая статистика $R = 0,799$ и $0,818$,
 $p = 0,000$, число перестановок – 4368)**

Сравниваемая группа	Статистика R	Уровень значимости, %	Число возможных перестановок	Действительное число перестановок
1–2	0,663/0,725	1,6/0,8	126	126
1–3	0,972/0,971	0,0	4368	4368
1–4	0,431/0,503	3,6/1,8	56	56
1–5	0,782/0,745	4,8	21	21
1–6	0,836/0,884	0,8	126	126
2–3	0,581/0,636	0,1	1365	1365
2–4	0,944/0,944	2,9	35	35
2–5	1,000/1,000	6,7	15	15
2–6	0,894/0,925	0,8	126	126
3–4	0,871/0,902	0,3	364	364
3–5	0,886/0,937	1,3	78	78
3–6	0,880/0,874	0,0	4368	4368
4–5	1,000/1,000	10,0	10	10
4–6	0,508/0,59	3,6	56	56
5–6	0,691/0,782	4,8	21	21

Примечание. В числителе — для плотности, в знаменателе — для биомассы (где есть различия); p — вероятность справедливости нулевой гипотезы H_0 $R = 0$; жирным шрифтом – выделены вероятности >5 %.

Тест Мантеля. Понадобятся три массива данных:

- 1) трансформированные и транспонированные значения плотности и/или биомассы (At4);
- 2) данные по факторам среды;
- 3) координаты станций.

Первые у нас уже есть, вторые и третьи импортируем и называем, например, Env и Geo.

Формируем таблицу видов и преобразуем ее в матрицу дистанций Хеллингера:

```
At4[At4<0] <- 0 (21),
```

```
At4N <- decostand(At4, method="hellinger") (22),
```

```
D_spec <- dist(At4N) (23),
```

Создаем таблицу факторов среды, преобразуем ее в стандартизованный вид и строим матрицы евклидовых дистанций:

```
EnvS <- decostand(Env, method="standardize") (24),
```

```
D_env <- dist(EnvS) (25).
```

Формируем матрицы расстояний (км) между точками отбора проб*:

```
library(geosphere) (26),
```

```
D_geo <- as.dist(distm(Geo)/1000) (27).
```

* Если координаты представлены в стандартной форме (градусы, минуты, секунды), необходимо сначала преобразовать их в числовую форму [см. 15]

Можно выполнить графическое тестирование данных с отображением координат и кругов, пропорциональных количественной характеристике вида на каждой станции:

```
plot(Geo$Xc, Geo$Yc, asp=1, pch=21, col="white", bg="green", cex=5*
      (At4[,n])/max((At4[,n])),
      (28)
```

где n – номер вида в $At4$ (если использовать $At4$ – круги будут пропорциональны корню 4 степени из количественной характеристики, если At – самой характеристике); пример такого тестирования представлен на рисунке 6. Сам тест:

«Обилие видов» ~ «Факторы среды»:

```
mantel(D_spec, D_env) (29) .
```

«Обилие видов» ~ «Пространственное расположение»:

```
mantel(D_spec, D_geo) (30) .
```

«Факторы среды» ~ «Пространственное расположение»:

```
mantel(D_env, D_geo) (31) .
```

Связи одновременно всех трех матриц:

```
mantel.partial(D_spec, D_env, D_geo) (32) .
```

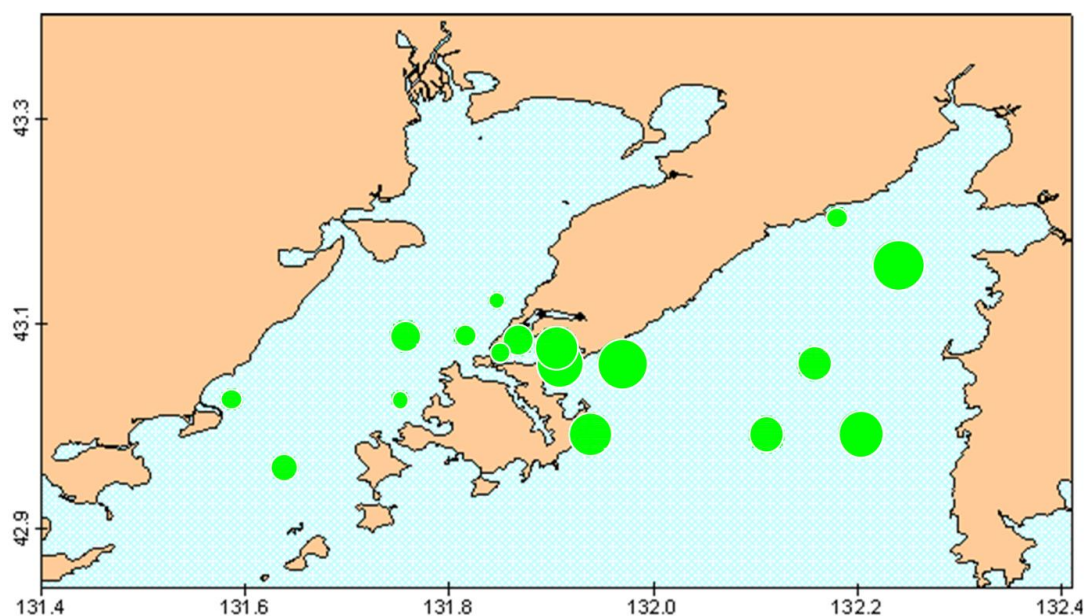


Рис. 6. Пример графического тестирования данных: размер кругов пропорционален корню четвертой степени из биомассы полихеты *Scoloplos armiger*. Пространственное распределение $At4$ наложено на карту, построенную в ППП Surfer

Результаты вычислений показывают, что определяющее и статистически значимое влияние на обилие видов макрозообентоса оказывает сумма факторов среды и, очевидно, их группирование (видов и станций, по обилию первых) обусловлено экологически, а не по каким-либо другим причинам (табл. 3). Таким образом, выделенные группировки I–VI, в соответствии с определением, являются сообществами.

Результаты теста Мантеля для оценки статистической значимости связи обилия видов макрозообентоса с факторами среды и пространственным расположением (число перестановок – 999)

Сравниваемые матрицы дистанций	Статистика Мантеля r	95-% доверительные интервалы r	p (при рандомизации)
{Обилие видов S } ~ {Факторы среды E }	0,660/0,657	$\frac{0,517 \div 0,803}{0,495 \div 0,819}$	0,001
{Обилие видов S } ~ {Расположение в пространстве G }	0,153/0,127	$\frac{0,008 \div 0,298}{-0,017 \div 0,271}$	0,042/0,070
{Факторы среды E } ~ {Расположение в пространстве G }	-0,045	-0,231 \div 0,141	0,623
Связь между тремя матрицами S , E и G одновременно	0,676/0,669	$\frac{0,523 \div 0,829}{0,524 \div 0,814}$	0,001

Примечание: в числителе – для плотности, в знаменателе – для биомассы (где есть различия), p – вероятность справедливости нулевой гипотезы $H_0: r=0$.

Наименование сообществ. В наибольшей степени сходство станций обеспечивают те виды, которые относительно обильны во всех или большинстве проб (на всех или большинстве станций). Логично, что именно по таким видам и должны именоваться сообщества, а не по тем, которые либо из-за больших размеров, либо выраженного агрегированного распределения дают максимальные средние биомассы или плотности на фоне низкой, а то и единичной встречаемости. Именно для выявления таких видов, которые вносят наибольшие вклады во внутригрупповое сходство и межгрупповое несходство, и используется процедура SIMPER (как и ANOSIM, ее проще выполнить в ПП PRIMER):

`data (At4)` (33) ,

`data (Cluster)` (34) ,

`(sim <- with(Cluster, simper(At4, Group)))` (35) ,

`summary(sim)` (36) .

Основной вклад во внутригрупповое сходство станций группы I, как по плотности, так и по биомассе, вносят *Aphelochaeta pacifica* и *Capitella capitata* (табл. 4). В группе II по плотности лидирует *Ophiura sarsi*, а по биомассе – *Macoma scarlatoi*, в группе III почти одинаковые вклады по плотности наблюдаются у *O. sarsi*, *Scoloplos armiger* и *Ennucula tenuis*, а по биомассе первый вид слегка опережает два других. В группе IV первое место занимает *Phoronopsis harmeri* и по плотности и по биомассе, в группе V 100-% вклад обеспечивают всего 4 вида макрозообентоса – полихеты *Lumbrineris longifolia*, *Sigambra bassi*, *Spiophanes bombyx* и неопределенный до вида представитель класса немуртин, причем первый из перечисленных видов несколько опережает остальные по обоим показателям. Наконец, в группе VI абсолютно лидирует *Maldane sarsi*. Очевидно, с учетом результатов процедуры ANOSIM и теста Мантеля, выделенные группы

следует рассматривать как сообщества *A. pacifica* + *C. capitata*, *O. sarsi* + *M. scarlatoii*, *O. sarsi* + *S. armiger* + *E. tenuis*, *Ph. harmeri*, *L. longifolia* и *M. sarsi*.

Таблица 4

Разложение среднего внутригруппового сходства по вкладам отдельных видов макрозообентоса

Вид	А или В	$\bar{\delta}_i$	$\delta_i/SD(\delta_i)$	CN, %	CCN, %
Плотность. Группа I. Среднее сходство: 32,11					
<i>Aphelochaeta pacifica</i>	338,0	13,98	0,94	43,53	43,53
<i>Capitella capitata</i>	65,0	11,39	1,12	35,47	79,00
<i>Schistomeringos japonica</i>	191,0	5,62	0,57	17,51	96,51
<i>Nereis</i> sp.	4,0	1,12	0,32	3,49	100,00
Биомасса. Группа I. Среднее сходство: 27,73					
<i>Aphelochaeta pacifica</i>	3,40	12,91	0,96	46,57	46,57
<i>Capitella capitata</i>	0,62	8,77	1,14	31,61	78,18
<i>Schistomeringos japonica</i>	2,13	4,40	0,60	15,86	94,05
<i>Nereis</i> sp.	9,56	1,65	0,32	5,95	100,00
Плотность. Группа II. Среднее сходство: 55,52					
<i>Ophiura sarsi</i>	618,8	8,95	4,82	16,13	16,13
<i>Aphelochaeta pacifica</i>	242,5	6,30	2,91	11,35	27,47
<i>Scoloplos armiger</i>	148,8	6,04	2,31	10,88	38,35
<i>Lumbrineris longifolia</i>	73,8	5,31	5,85	9,56	47,91
<i>Macoma scarlatoii</i>	35,0	4,29	4,60	7,73	55,64
Биомасса. Группа II. Среднее сходство: 54,01					
<i>Macoma scarlatoii</i>	274,88	13,33	4,38	24,67	24,67
<i>Ophiura sarsi</i>	35,70	8,18	4,83	15,16	39,83
<i>Scoloplos armiger</i>	5,13	4,44	2,00	8,22	48,05
<i>Aphelochaeta pacifica</i>	6,22	4,27	2,24	7,91	55,96
<i>Lumbrineris longifolia</i>	0,78	3,33	4,63	6,17	62,13
Плотность. Группа III. Среднее сходство: 42,78					
<i>Ophiura sarsi</i>	164,1	5,25	4,41	12,28	12,28
<i>Scoloplos armiger</i>	116,4	4,91	3,85	11,49	23,77
<i>Ennucula tenuis</i>	102,7	4,91	3,86	11,47	35,24
<i>Lumbrineris longifolia</i>	70,0	4,53	4,98	10,60	45,84
<i>Maldane sarsi</i>	519,1	4,10	1,34	9,58	55,42
Биомасса. Группа III. Среднее сходство: 38,67					
<i>Ophiura sarsi</i>	22,62	5,92	2,78	15,31	15,31
<i>Ennucula tenuis</i>	8,57	4,90	2,89	12,68	27,99
<i>Scoloplos armiger</i>	5,30	4,46	2,76	11,53	39,51
<i>Maldane sarsi</i>	7,63	3,78	1,33	9,77	49,28
<i>Lumbrineris longifolia</i>	1,75	3,12	3,33	8,08	57,36
Плотность. Группа IV. Среднее сходство: 52,45					
<i>Phoronopsis harmeri</i>	4000,0	13,57	1,78	25,87	25,87
<i>Aphelochaeta pacifica</i>	500,0	10,38	87,03	19,79	45,66
<i>Schistomeringos japonica</i>	36,7	5,59	28,61	10,65	56,31
<i>Maldane sarsi</i>	13,3	3,88	91,17	7,39	63,70
<i>Notomastus latericeus</i>	60,0	2,48	0,58	4,73	68,43
Биомасса. Группа IV. Среднее сходство: 44,72					
<i>Phoronopsis harmeri</i>	822,75	15,29	1,59	34,20	34,20
<i>Aphelochaeta pacifica</i>	10,42	7,81	10,33	17,46	51,66

Окончание таблицы 4.

Вид	A или B	$\bar{\delta}_i$	$\delta_i/SD(\delta_i)$	CN, %	CCN, %
<i>Schistomeringos japonica</i>	0,38	3,57	16,22	7,99	59,65
<i>Scalibregma inflatum</i>	13,15	3,22	0,58	7,20	66,85
<i>Maldane sarsi</i>	0,73	2,71	4,30	6,06	72,91
Плотность. Группа V. Среднее сходство: 39,01					
<i>Lumbrineris longifolia</i>	67,5	12,98	–	33,26	33,26
<i>Sigambra bassi</i>	15,0	8,68	–	22,25	55,51
<i>Spiophanes bombyx</i>	25,0	8,68	–	22,25	77,75
Nemertini fam.gen.sp.	10,0	8,68	–	22,25	100,00
Биомасса. Группа V. Среднее сходство: 37,94					
<i>Lumbrineris longifolia</i>	0,88	11,68	–	30,79	30,79
Nemertini fam.gen.sp.	0,43	11,05	–	29,12	59,90
<i>Sigambra bassi</i>	0,18	8,64	–	22,78	82,69
<i>Spiophanes bombyx</i>	0,28	6,57	–	17,31	100,00
Плотность. Группа VI. Среднее сходство: 31,60					
<i>Maldane sarsi</i>	442,0	16,80	1,70	53,16	53,16
<i>Sigambra bassi</i>	9,0	4,92	1,03	15,58	68,73
<i>Lumbrineris longifolia</i>	44,0	2,46	0,60	7,79	76,52
<i>Scalibregma inflatum</i>	21,0	1,82	0,60	5,77	82,29
Ascidia fam.gen.sp.	6,0	1,73	0,60	5,47	87,75
Биомасса. Группа VI. Среднее сходство: 23,98					
<i>Maldane sarsi</i>	34,97	14,69	1,60	61,26	61,26
<i>Sigambra bassi</i>	0,15	2,54	0,92	10,59	71,85
<i>Glycera</i> sp.	0,50	1,22	0,56	5,10	76,94
Ascidia fam.gen.sp.	1,67	1,19	0,58	4,95	81,89
<i>Ennucula tenuis</i>	1,17	1,13	0,32	4,70	86,59
Примечание. Приведены, по возможности, первые 5 видов. Виды расположены в порядке уменьшения процентных вкладов; δ_i – мера сходства, SD – стандартное отклонение, CN – процентный вклад, CCN – накопленные проценты.					

Описание сообществ. После выделения и наименования ассоциаций донных животных следует сделать их описания. На наш взгляд, они должны быть краткими, построенными по единой схеме и включать основные характеристики обилия и экологические индексы (R , A , B , H' , $AMBI$, $M-AMBI$)^{*} (табл. 5). Кроме того, логично оценить и статистическую значимость различий перечисленных параметров у выделенных сообществ или, хотя бы, таковую влияния фактора «кластер». Для этого, создадим в нашей книге Excel еще один лист и поместим туда значения всех этих показателей по станциям, импортируем данные в R и перейдем к выполнению непараметрических тестов (полные результаты не приводятся):

```
data(parameters) (37)
```

```
attach(Cluster) (38)
```

```
kruskal.test(A ~ Group, data = parameters) (39)
```

Воздействие фактора «кластер» оказалось значимо для большинства показателей, кроме индекса

^{*} Рассчитать биотические параметры можно, например, при помощи ПП PRIMER; для вычисления индексов $AMBI$ и $M-AMBI$, характеризующих экологическое состояние, существует программное обеспечение, свободно распространяемое в интернете [16; 18].

Пиелу и статистики Кларка (результаты теста Крускала-Уоллиса: $p=0.000-0.030$, 0.117 и 0.351). Следует отметить, что выполнить различные непараметрические тесты (и многие другие стандартные процедуры) можно в GUI Rcmdr; кроме того, этот интерфейс представляет довольно широкие графические возможности. Ниже в качестве примера приведены описания сообществ, выявленных в 2016 г.

Сообщество I *A. pacifica* + *C. capitata* характеризуется наименьшими среди остальных группировок средними величинами числа видов, биомассы, индексов Шеннона-Винера и статистики Кларка и, в тоже время, довольно высокой плотностью поселения (табл. 5). Это очень бедная группировка ($H' = 0,80 \pm 0,31$) с плохо выровненным ранговым распределением видов ($e = 0,331 \pm 0,120$), в которой преобладают относительно мелкие животные ($W = -0,063 \pm 0,060$). Всего в ассоциации найдено 13 видов, среди которых чаще других встречались *A. pacifica* и *C. capitata* (80,0 %), причем первый доминировал по плотности (338 ± 215 экз./м²). По биомассе лидировал относительно редкий *Nereis* sp. (частота встречаемости – 40 %, биомасса – $9,6 \pm 10,3$ г/м²). На долю видов – позитивных индикаторов загрязнения и эвтрофикации (5 из 13) приходится около 96 и 98 % от общей плотности и биомассы сообщества. В соответствии с величинами *AMBI* и *M-AMBI* на разных станциях сообщество находится в состоянии умеренного и сильного повреждения, его экологический статус варьирует от обедненного до плохого.

Для группировки II *O. sarsi* + *M. scarlatoi* характерны более высокие значения всех биотических параметров (табл. 5). Всего в ней отмечено 46 видов беспозвоночных, а 9 видов – *A. pacifica*, *Chaetozone setosa*, *Eteone longa*, *Glycera capitata*, *L. longifolia*, *S. armiger*, *Schistomeringos japonica*, *M. scarlatoi* и *O. sarsi* имели 100-% встречаемость, по плотности – лидировал *O. sarsi* (619 ± 271 экз./м²), по биомассе – *M. scarlatoi* ($274,9 \pm 97,6$ г/м²). Это разнообразное сообщество ($H' = 2,66 \pm 0,23$) с умеренно выровненным распределением видов по рангам ($e = 0,603 \pm 0,052$) и доминированием относительно крупных животных ($W = 0,127 \pm 0,031$). На долю положительных индикаторов загрязнения приходится почти 31 % от общей плотности, но только 2,6 % от биомассы. Степень повреждения варьирует от легкой до умеренной, а экологический статус – от хорошего до высокого.

Ассоциация III *O. sarsi* + *S. armiger* + *E. tenuis* – третья по обилию макрофауны и самая разнообразная по ее составу ($H' = 2,86 \pm 0,22$). Ранговое распределение видов – умеренно выровненное ($e = 0,645 \pm 0,053$), а статистика Кларка изменяется в очень широких пределах (табл. 5). Общее число видов – 85, причем четыре из них найдены во всех пробах. Это сами *O. sarsi*, *S. armiger* и *E. tenuis*, а также *L. longifolia*. Самый многочисленный вид – *Phoronida* fam. gen. sp., найденный только на 3 из 11 станций (895 ± 627 экз./м²), доминант по биомассе – *B. rostratus* (1 станция, $100,0 \pm 104,8$ г/м²). Вклад положительных индикаторов загрязнения в общие показатели не превышает соответственно 1 и 0.1 %. Степень повреждения изменяется от ненарушенной до слегка нарушенной, а экологический статус – от хорошего до высокого.

Количественные и структурные характеристики сообществ макрозообентоса на исследованной акватории

Сообщество	r	A , экз./м ²	B , г/м ²	H'	e	W	$AMBI$	$M-AMBI$
I. <i>Aphelochaeta pacifica</i> + <i>Capitella capitata</i>	$\frac{4 \pm 1}{1-6}$	$\frac{619 \pm 225}{10-1275}$	$\frac{16,3 \pm 11,2}{0,1-55,7}$	$\frac{0,80 \pm 0,31}{0,00-1,58}$	$\frac{0,331 \pm 0,120}{0,000-0,613}$	$\frac{-0,063 \pm 0,060}{-0,174-0,099}$	$\frac{5,22 \pm 0,38}{4,29-6,00}$	$\frac{0,16 \pm 0,05}{0,01-0,25}$
II. <i>Ophiura sarsi</i> + <i>Macoma scarlatoi</i>	$\frac{22 \pm 4}{16-33}$	$\frac{1589 \pm 501}{915-2860}$	$\frac{382,5 \pm 123,4}{121,4-581,2}$	$\frac{2,66 \pm 0,23}{2,30-3,16}$	$\frac{0,603 \pm 0,052}{0,533-0,732}$	$\frac{0,127 \pm 0,031}{0,049-0,173}$	$\frac{2,70 \pm 0,25}{2,15-3,18}$	$\frac{0,69 \pm 0,06}{0,58-0,83}$
III. <i>Ophiura sarsi</i> + <i>Scoloplos armiger</i> + <i>Ennucula tenuis</i>	$\frac{22 \pm 1}{16-30}$	$\frac{2710 \pm 942}{235-8975}$	$\frac{235,3 \pm 107,2}{32,8-1235}$	$\frac{2,86 \pm 0,22}{1,56-3,58}$	$\frac{0,645 \pm 0,053}{0,340-0,829}$	$\frac{0,099 \pm 0,052}{-0,186-0,355}$	$\frac{1,97 \pm 0,20}{0,77-2,84}$	$\frac{0,75 \pm 0,02}{0,60-0,88}$
IV. <i>Phoronopsis harmeri</i>	$\frac{15 \pm 1}{13-17}$	$\frac{4957 \pm 2473}{1045-7780}$	$\frac{894,0 \pm 650,7}{97,5-1902}$	$\frac{1,52 \pm 0,80}{0,84-2,83}$	$\frac{0,384 \pm 0,189}{0,220-0,692}$	$\frac{0,049 \pm 0,016}{0,029-0,074}$	$\frac{3,36 \pm 0,30}{3,05-3,83}$	$\frac{0,45 \pm 0,07}{0,39-0,57}$
V. <i>Lumbrineris longifolia</i>	$\frac{10 \pm 5}{6-13}$	$\frac{403 \pm 357}{150-655}$	$\frac{5,5 \pm 0,2}{5,4-5,7}$	$\frac{2,19 \pm 0,21}{2,04-2,33}$	$\frac{0,727 \pm 0,249}{0,551-0,902}$	$\frac{0,025 \pm 0,288}{-0,179-0,228}$	$\frac{3,15 \pm 0,40}{2,86-3,43}$	$\frac{0,48 \pm 0,05}{0,44-0,52}$
VI. <i>Maldane sarsi</i>	$\frac{11 \pm 3}{6-22}$	$\frac{639 \pm 139}{235-935}$	$\frac{885,0 \pm 585,1}{32,7-2695}$	$\frac{1,56 \pm 0,61}{0,48-3,52}$	$\frac{0,461 \pm 0,142}{0,185-0,789}$	$\frac{0,146 \pm 0,099}{-0,055-0,374}$	$\frac{1,80 \pm 0,14}{1,61-2,26}$	$\frac{0,51 \pm 0,09}{0,38-0,83}$
Все данные	$\frac{16 \pm 2}{1-33}$	$\frac{1938 \pm 453}{10-8975}$	$\frac{377,3 \pm 117,9}{0,1-2695}$	$\frac{2,09 \pm 0,20}{0,00-3,58}$	$\frac{0,536 \pm 0,043}{0,000-0,902}$	$\frac{0,073 \pm 0,029}{-0,186-0,374}$	$\frac{2,80 \pm 0,25}{0,77-6,00}$	$\frac{0,56 \pm 0,04}{0,01-0,88}$

Примечание. В числителе — среднее значение \pm ошибка репрезентативности, в знаменателе — диапазон изменчивости, R — среднее число видов в пробе (индекс Маргалефа), A — плотность поселения, B — биомасса, H' и e — индексы Шеннона-Винера и Пиелу, W — статистика Кларка, $AMBI$ и $M-AMBI$ — индексы, характеризующие степень нарушения сообщества и его экологический статус.

Сообщество IV *Ph. harmeri* показывает самые высокие значения параметров обилия, но существенно уступает ассоциации *O. sarsi* + *S. armiger* + *E. tenuis* в разнообразии донных животных (табл. 5). Всего в группировке обнаружено 25 видов, причем 5 из них – *Ph. harmeri* (абсолютный доминант как по плотности, так и по биомассе – 4000 ± 2392 экз./м² и 822.8 ± 640.0 г/м²), *A. pacifica*, *Sch. japonica*, *S. inflatum*, *M. sarsi* – были найдены во всех пробах. Выраженное доминирование мелких форонид обуславливает не только низкие значения статистики Кларка (0.049 ± 0.016), но и таковые индексов H' и e (соответственно 1.52 ± 0.80 и 0.384 ± 0.189). Вклад положительных индикаторов загрязнения в общую плотность и биомассу несколько больше, чем в предыдущем сообществе и составляет почти 11 и чуть более 1 %. Степень повреждения изменяется от слегка до умеренно нарушенной, а экологический статус – от умеренного до хорошего.

Группировка V *L. longifolia* имеет самую низкую биомассу, небольшое среднее число видов в пробе и плотность (табл. 5). Вместе с тем, с позиций теории информации, это умеренно разнообразное сообщество с выровненным распределением видов по рангам ($H' = 2.19 \pm 0.21$ и $e = 0.727 \pm 0.249$), причем индекс Пиелу, в среднем, достигает здесь наибольшего значения среди найденных ассоциаций. Статистика Кларка изменяется, как и в группировке *O. sarsi* + *S. armiger* + *E. tenuis*, в очень широких пределах ($W = 0.025 \pm 0.288$). Всего в сообществе найдено 15 видов, четыре из них – *L. longifolia*, *S. bombyx*, *S. bassi* и неопределенный до вида представитель класса немертин – обнаружены на обеих станциях. По плотности доминировала полихета *Chaetozone setosa*, по биомассе – гастропода *Patelloida heroldi* (соответственно 208 ± 293 экз./м² и 1.2 ± 1.6 г/м², встречаемость – 50 %); положительные индикаторы загрязнения не найдены. Как и у сообщества форонид, степень повреждения варьирует от слегка до умеренно нарушенной, экологический статус – умеренный.

Ассоциация VI *M. sarsi* имеет весьма высокую биомассу при относительно небольшой плотности и числе видов в пробе (табл. 5). В среднем, это бедное сообщество с плохо выровненным ранговым распределением видов и довольно высокой статистикой Кларка ($H' = 1.56 \pm 0.61$, $e = 0.461 \pm 0.142$, $W = 0.146 \pm 0.099$), при этом все эти параметры показывают очень высокую изменчивость. В группировке найдено 36 видов, но только *M. sarsi* была встречена на всех станциях, доминируя по плотности (442 ± 183 экз./м²), а наибольшие средние биомассы были отмечены у редких двустворчатых моллюсков *A. boucardi* и *A. broughtoni* (280.9 ± 314 и 258.5 ± 289 ; $F_q = 20$ %). Положительные индикаторы загрязнения встречались единично. Сообщество находится в слегка нарушенном состоянии, экологический статус варьирует от обедненного до высокого.

Таким образом, выполненные выше процедуры позволили расклассифицировать станции,

показать, что выделенные ассоциации экологически детерминированы и, следовательно, являются сообществами макрозообентоса, назвать и описать эти сообщества. Результаты данных операций, в «нужных местах», оказались значимыми с позиций статистики. Однако важной задачей мониторинга является и выявление «главных» факторов, определяющих развитие той или иной группировки.

Определение ведущих факторов среды. Наиболее «прогрессивной» процедурой выявления таких факторов является канонический анализ соответствий в самом его нелинейном выражении (Constrained CA, CCA; на самом деле, базируясь на расстояниях χ^2 и представляя взвешенное линейное отображение, он каноничен лишь в силу терминологических традиций). При выполнении этой процедуры генерируется множество промежуточных результатов, которые мы пропустим. Полная модель, включающая все факторы среды, ее дисперсионный анализ и различные «полезные» диаграммы получаются при выполнении следующих команд:

```
vareA.cca <- cca(At4 ~ ., Env) (40),
```

```
vareA.cca (41),
```

```
plot(vareA.cca, display = "sites", type = "t") (42),
```

```
plot(vareA.cca, type = "t") (43),
```

```
plot(vareA.cca) (44),
```

```
anova(vareA.cca) (45).
```

Диаграмма рассеяния, получаемая по выражению (42), отображает только коды станций, команда (43) показывает коды станций, векторы факторов и среды названия всех видов (как оставить названия лишь самых «важных» из них – [см. 20]). При включении скрипта (44) вместо названия видов выводятся их «центры тяжести» (центроиды), формула (45) воспроизводит дисперсионный анализ для модели, которая включает все факторы среды. Структуру связей последних с осями CCA, получаемых по характеристикам обилия видов, можно получить при помощи выражений:

```
ef <- envfit(vareA.cca, Env, permu = 999); ef (46),
```

а на диаграмму рассеяния вывести только «значимые» параметры, введя ограничения по уровню значимости (p.max):

```
plot(ef, p.max = 0.05) (47).
```

При этом можно заметить, например, что многие параметры явно коррелируют друг с другом и могут быть «опущены». Для их сокращения используем пошаговую процедуру, которая выбирает «наилучшую» модель на основе информационного критерия Акаике (по умолчанию):

```
mod <- step(vareA.cca, scope = formula(vareA.cca),  
test = "perm"); mod (48).
```

В результате получаем некую новую модель:

$$At4 \sim P1 + P2 + \dots + Pk \quad (49),$$

где P_i – оставшиеся после сокращения факторы. Необходимо подчеркнуть, что применять эту процедуру следует с большой осторожностью, проверяя получаемые результаты на каждом шаге, чтобы не удалить параметры, которые могут быть весьма информативны для объяснения результатов. Можно, например, вначале оценить скоррелированность переменных и отбросить те из них, которые сильно связаны с другими и, в то же время, плохо интерпретируются (например, после предварительного снижения размерности методами факторного анализа). Кроме того, на наш взгляд, в модели следует оставлять и параметры с маргинальными вероятностями ($0,05 < p < 0,1$).

Для данных 2016 г. (биомасса) итоговая модель включает общий уровень химического загрязнения донных отложений, параметр, характеризующий терригенный (речной) сток, летнее (минимальное) содержание кислорода у дна и глубину (соответственно TPF , $MeF2$, $O2$, Dth). При анализе плотности вместо TPF в модель «встраивается» параметр RoF_1 , идентичный ему по смыслу. Дисперсионный анализ:

$$\text{mod} <- \text{cca}(At4 \sim TPF + MeF2 + O2 + Dth, Env) \quad (50),$$

$$\text{anova}(\text{mod}) \quad (51),$$

$$\text{anova}(\text{mod}, \text{by} = "term", \text{step}=200) \quad (52),$$

$$\text{anova}(\text{mod}, \text{by} = "margin", \text{perm}=500) \quad (53),$$

$$\text{anova}(\text{mod}, \text{by}="axis", \text{perm}=1000) \quad (54).$$

При использовании выражения (51) получаем результат, аналогичный таковому из (45), но оценивается модель с сокращенным числом переменных. Скрипт (52) воспроизводит дисперсионный анализ I типа, когда все характеристики анализируются последовательно, причем в том порядке, в котором они стоят в формуле, и этот порядок влияет на конечный результат: первые параметры оказываются более «значимыми», чем последующие. Команда (53) приводит к выполнению теста ANOVA III типа, оценивающего так называемые краевые эффекты, которые независимы от порядка термов в формуле, но коррелированные переменные получают завышенные значения вероятности. Применение (54) оценивает значимость осей ССА с позиций статистики. Следующие команды выполняют подгонку векторов факторов среды к осям ССА и показывают их структуру:

$$ef <- \text{envfit}(\text{mod}, Env, \text{permu} = 999); ef \quad (55).$$

Результаты дисперсионного анализа суммированы в таблице 6: статистически значимыми оказались сама модель в целом, включение в нее глубины, уровня химического загрязнения, содержания кислорода, терригенного стока и две первые оси ССА из четырех. Кроме того, мы получили весовые коэффициенты перечисленных переменных для осей ССА (табл. 7). Следует подчеркнуть, что отчетливая и статистически значимая дифференциация сообществ по

отношению к факторам среды является подтверждением результатов теста Мантеля и, соответственно, вывода о том, что выделенные группы являются сообществами.

Таблица 6

Результаты ССА: дисперсионный анализ (ANOVA) для полных моделей, моделей с сокращенным числом параметров и статистическая оценка их термов по отдельности, влияние краевых эффектов и значимость осей (число перестановок – 999)

Параметр	<i>Df</i>	χ^2	<i>F</i>	Pr(> <i>F</i>)
Общая модель: $Bt4/At4 \sim \text{Depth} + TPF/PoF_1 + MeF_2 + O_2$				
Модель	4	1,495/1,285	1,764/1,917	0,001
Остатки	25	5,296/4,191		
Последовательное включение (тип I)				
Глубина	1	0,463/0,342	2,187/2,041	0,001
TPF/PoF_1	1	0,356/0,398	1,679/2,374	0,006/0,001
O_2	1	0,302/0,211	1,423/1,260	0,020/0,073
MeF_2	1	0,375/0,334	1,769/1,992	0,002/0,001
Остатки	25	5,296/4,191		
Краевые эффекты (тип III)				
Глубина	1	0,340/0,267	1,605/1,593	0,002/0,004
TPF/PoF_1	1	0,315/0,307	1,485/1,830	0,016/0,001
O_2	1	0,302/0,211	1,423/1,260	0,021/0,065
MeF_2	1	0,341/0,285	1,608/1,702	0,004/0,002
Остатки	25	5,296/4,191		
Оси				
ССА1	1	0,549/0,464	2,590/2,767	0,001
ССА2	1	0,416/0,392	1,963/2,339	0,001
ССА3	1	0,272/0,229	1,284/1,369	0,153/0,063
ССА4	1	0,259/0,200	1,221/1,192	0,071/0,139
Остатки	25	5,296/4,191		

Примечание. В числителе – результаты для биомассы, в знаменателе – для плотности.

Таблица 7

Результаты анализа соответствий и процедуры многомерного шкалирования: корреляция факторов среды и осей, полученных по биомассам/плотностям найденных видов (число перестановок — 999)

Параметр	ССА1	ССА2	r^2	Pr(> <i>r</i>)
Depth	-0,981/-0,490	-0,194/-0,872	0,668/0,553	0,001/0,002
TPF/PoF_1	-0,029/0,635	-0,999/-0,772	0,466/0,780	0,002/0,001
O_2	0,020/-0,881	0,999/0,473	0,635/0,503	0,001
MeF_2	0,860/0,928	-0,511/0,371	0,776/0,739	0,001
Параметр	NMDS1	NMDS2	r^2	Pr(> <i>r</i>)
Depth	-0,252/-0,045	0,968/0,999	0,269/0,374	0,016/0,002
TPF/PoF_1	0,999/0,982	0,035/0,188	0,478/0,432	0,001/0,002
O_2	-0,875/-0,941	0,485/0,338	0,528/0,545	0,001
MeF_2	0,482/0,305	-0,876/-0,953	0,541/0,670	0,001

Примечание. В числителе – результаты для биомассы, в знаменателе – для плотности.

Однако четыре независимых переменных – все же слишком много для содержательной интерпретации, их число можно снизить, если внимательно посмотреть на структуру факторных осей (табл. 7). Каждую ось формируют по две переменные с близкими по величине, стремящимися к единице, но противоположными по знаку весовыми коэффициентами. Следовательно, можно оставить всего две переменные, например *TPF* и *MeF2*. Или *TPF* и глубину (как интегральный фактор, характеризующий гидрологический режим, что не всегда верно) и т.д.

Визуализацию дифференциации сообществ под воздействием факторов среды и результаты подгонки поверхностей последних нагляднее выполнять при помощи нелинейного многомерного шкалирования (NMDS). Этот метод, судя по опыту применения, «менее линейен», чем ССА и диаграммы рассеяния получаются более сжатыми. Так как скрипты (57) – (61) повторяют таковые для ССА, не будем заострять на них внимание.

```
library(MASS) (56),
```

```
vareA.mds <- metaMDS(At4, trace = FALSE) (57),
```

```
vareA.mds (58),
```

```
ef <- envfit(vareA.mds, Env, permu = 999); ef (59),
```

```
plot(vareA.mds, display = "sites", type = "t") (60),
```

```
plot(ef, p.max = 0.05) (61),
```

Структура корреляций осей NMDS с факторами среды представлена в таблице 7. Следующие команды позволят отобразить на диаграммах рассеяния поверхности факторов:

```
tmp <- with(Env, ordisurf(vareA.mds, TPF, add = TRUE, (62),
col = "red"))
```

```
tmp <- with(Env, ordisurf(vareA.mds, Dth, add = TRUE, (63),
col = "orange"))
```

```
tmp <- with(Env, ordisurf(vareA.mds, O2, add = TRUE, (64),
col = "green4"))
```

```
tmp <- with(Env, ordisurf(vareA.mds, MeF2, add = TRUE, (65).
col = "blue"))
```

Результаты визуализации представлены на рисунках 7 и 8, а приуроченность сообществ к определенным диапазонам различных факторов – в таблице 8. Далее, очевидно, необходимо сделать некое описание этой привязанности ассоциаций донных животных, например, следующим образом.

Сообщество I *A. pacifica* + *C. capitata* развивается в бухтах Золотой Рог и Улисс, исключительно при экстремальном и высоком уровне химического загрязнения и низком содержании кислорода у дна, причем не только в летний период. Именно эти факторы, судя по

результатам ССА и NLMS, имеют ведущее значение для его дифференциации (табл. 6, 8, рис. 7, 8). Соседствующая с ней группировка II *O. sarsi* + *M. scarlatoi* (пролив Босфор Восточный) привязана к большим глубинам, более аэробным условиям и умеренному загрязнению.

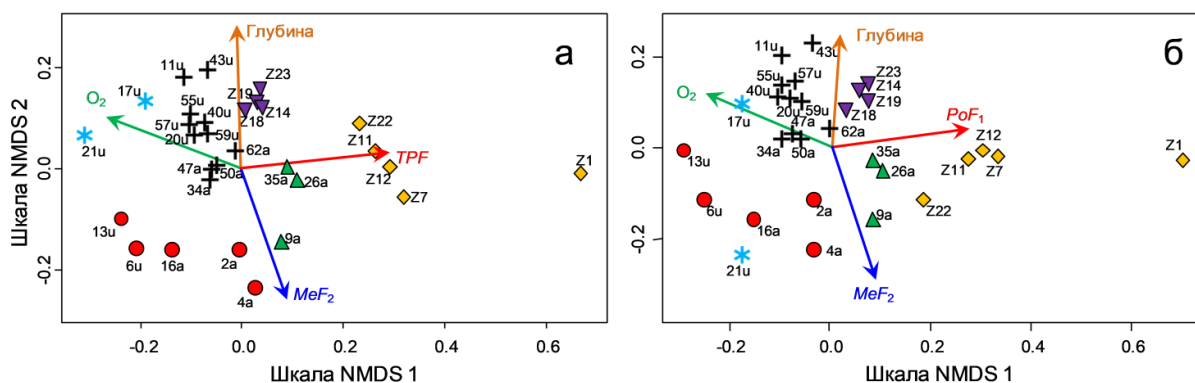


Рис. 7. Результаты процедуры неметрического многомерного шкалирования (NMDS) (преобразование – корень 4 степени, стресс – 0,138 и 0,136) и градиенты основных факторов среды: а – биомасса, б – плотность

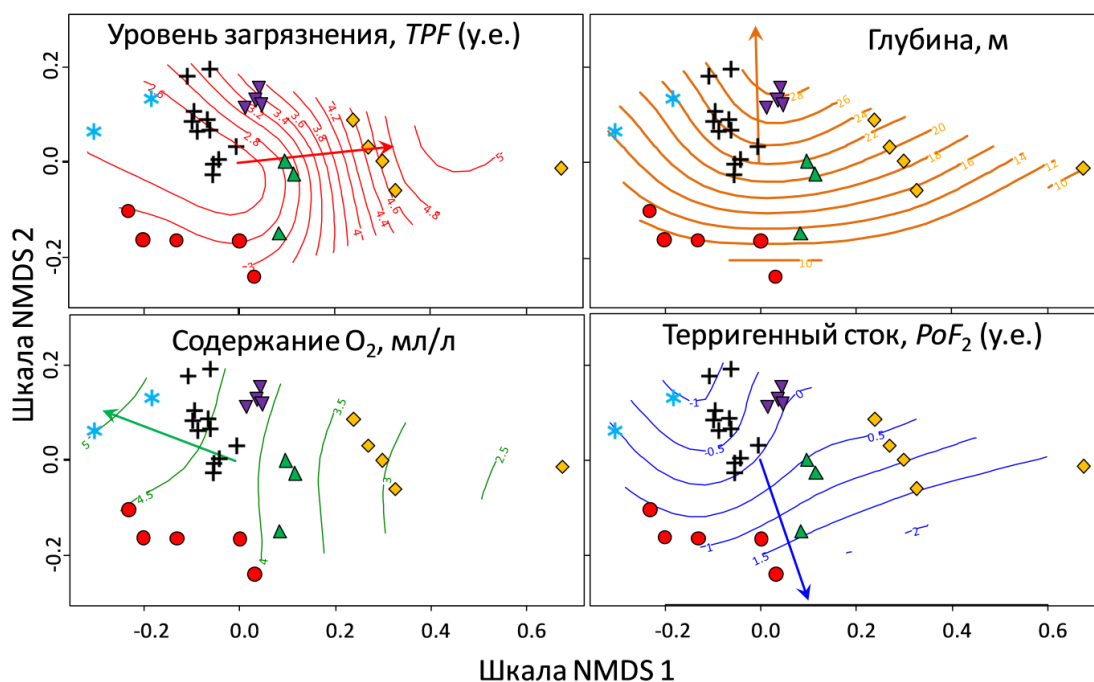


Рис. 8. Ординация станций методом NMDS (биомасса) и результаты подгонки поверхностей для факторов среды

Ассоциация III *O. sarsi* + *S. armiger* + *E. tenuis* формируется в открытых районах Амурского и Уссурийского заливов, где для большинства станций характерен низкий уровень загрязнения, а концентрации кислорода в придонном слое воды относительно высоки даже летом. Другая черта акватории, занимаемой этой группировкой – довольно низкая доля частиц тонких фракций (< 0,01 мм). Сообщество IV *Ph. harmeri* приурочено к области выраженной летней гипоксии (рис. 9). Этот дефицит кислорода возникает из-за микробиологического окисления избыточной биомассы диатомовых водорослей в условиях слабой динамики вод и

при низкой интенсивности фотосинтетически активной радиации [14]. Соответственно, в конце периода цветения фитопланктона и сразу после него, на этом участке может наблюдаться кратковременное, но сильное (в разы) повышение содержания органического углерода в донных осадках [10].

Таблица 8

Параметры среды и сообщества макрозообентоса

Сообщество	Глубина, м	Содержание частиц <0,1 мм, %	Содержание С _{орг.} , мг/г	TPF, усл. ед.	Содержание O ₂ , мл/л
I. <i>Aphelocheata pacifica</i> + <i>Capitella capitata</i>	17 (9–24)	53,8 (43,2–62,5)	5,68 (3,66–10,20)	4,6 (4,0–5,0)	3,05 (2,09–4,15)
II. <i>Ophiura sarsi</i> + <i>Macoma scarlatoi</i>	31 (27–37)	47,4 (40,6–52,3)	3,84 (2,78–4,90)	3,7 (3,0–4,0)	4,06 (3,52–4,54)
III. <i>Ophiura sarsi</i> + <i>Scoloplos armiger</i> + <i>Ennucula tenuis</i>	24 (15–38)	19,7 (11,0–39,8)	1,18 (0,11–3,47)	2,5 (1,8–3,2)	4,75 (3,96–5,73)
IV. <i>Phoronopsis harmeri</i>	21 (17–25)	49,9 (42,2–59,1)	3,45 (3,08–4,03)	2,5 (2,4–2,8)	3,36 (2,58–3,79)
V. <i>Lumbrineris longifolia</i>	15 (12–18)	35,0 (18,4–51,7)	0,69 (0,61–0,77)	2,8 (1,8–3,8)	5,02 (4,88–5,15)
VI. <i>Maldane sarsi</i>	11 (5–18)	56,3 (45,6–69,6)	1,99 (1,58–3,14)	2,9 (2,6–3,2)	4,34 (2,96–5,23)

Примечание. Для каждого параметра среды приведено среднее значение и диапазон изменений (в скобках).

Сообщество V *L. longifolia* найдено всего на двух станциях, которые сильно отличаются по уровню загрязнения, содержанию алевропелитов (что делает дифференциацию группировки сомнительной), а их сходство проявляется в низких концентрациях органического углерода и высоких – кислорода. Группировка VI *M. sarsi* явно приурочена к областям повышенного влияния терригенного стока, что отражается, в частности в очень высоком содержании частиц тонких фракций, хотя средняя глубина здесь минимальна (табл. 8, рис. 9).

Таким образом, суммируя результаты ССА, NMDS и анализа дифференциации выявленных сообществ относительно факторов среды, можно сделать вывод о том, что состав, структура и состояние макрозообентоса на исследованной акватории определяется, прежде всего, уровнем химического загрязнения и особенностями придонного гидрологического режима, который во многом обусловлен интенсивностью терригенного стока.

В заключение приведем скрипт еще для одной весьма полезной процедуры, позволяющей сопоставить результаты ординаций, получаемых на основе разных параметров (например, по биомассам и плотностям поселения) или методов (NMDS, ССА) – теста Прокруста или "прокрустовом" преобразовании геометрии точек (рис. 10):

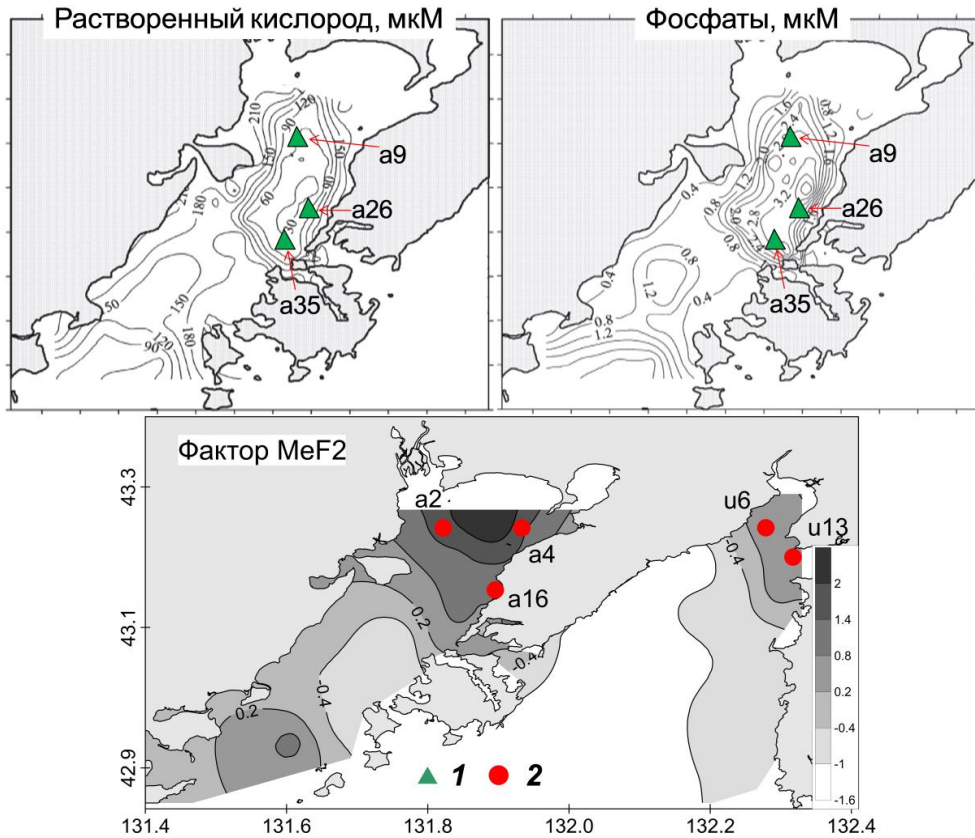


Рис. 9. Распределение некоторых океанографических элементов у дна в августе 2008 г. и значений фактора MeF_2 «терригенный сток» в 2016 г. (вверху и внизу соответственно по: Тищенко и др., [14] и Мощенко и др., [9]). Условные обозначения: 1 и 2 – сообщества IV *Ph. harmeri* и VI *M. sarsi*

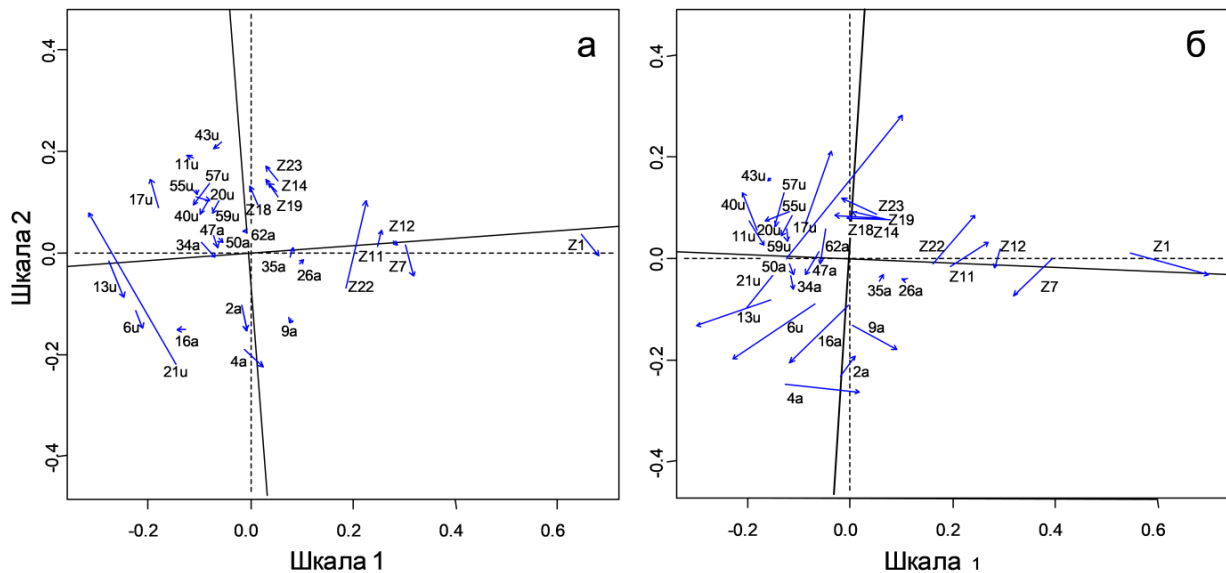


Рис. 10. Суперпозиция ординаций станций по биомассе и плотности, полученных методом NMDS, и по биомассе, полученных методами NMDS и CCA (соответственно а и б). Сумма квадратов прокруста m^2 соответственно 0,210 и 0,493; стрелки: а – начало – плотность, конец – биомасса; б – начало – CCA, конец – NMDS

```
pro <- procrustes(vareA.mds, vareB.mds) (66),
```

```
pro <- procrustes(vareA.cca, vareA.mds) (67),
```

```
pro (68) ,
```

```
plot(pro, type = "t") (69) ,
```

```
plot(pro, kind = 2) (70) .
```

Команды (66) и (67) сравнивают ординации, (68) – выводит значение прокрустовой суммы квадратов, представляющее собой минимум суммы квадратов расстояний между двумя ординациями (рис. 10). Выражение (69) строит эти диаграммы с осями разных ординаций и стрелками, показывающими «перемещение» точек (станций) при использовании разных параметров или методов (рис. 10). Скрипт (70) позволяет получить по-станционный график разностей остатков (не приводится).

В нашем случае, как при изменении характеристики обилия, так и метода, существенное «перемещение», которое может изменить окончательные выводы, показывает только станция u21 (рис. 10). При использовании плотности поселения вместо биомассы, равно как ССА вместо NMDS, группировка V *L. longifolia* (всего две станции, u17 и u21) плохо дифференцирована от остальных сообществ. Здесь можно вернуться к алгоритму нечетких множеств, сделать ординацию более «размазанной», увеличив экспоненциальный вес, и оценить силу притяжения станций u17 и u21 к другим группам.

Такая операция приводит к расщеплению кластера V: станция u17 отходит к III, а u21 – к VI группировке. Это решение весьма логично сразу по нескольким соображениям. Во-первых, кластер III объединит все станции с низким содержанием тонких фракций, а *L. longifolia* входит в пятерку видов, вносящих наибольший вклад во внутригрупповое сходство данной ассоциации (результаты SIMPER). Во-вторых, станция u21 с высоким содержанием алевропелитов найдет свое место в группировке VI приуроченной к грунтам со сходным количеством таких частиц, при этом *S. bassi*, весьма заметная в сообществе V, занимает второе место по значимости вклада во внутригрупповое сходство и в ассоциации VI. Наконец, такое расщепление просто следует из локализации станций в Уссурийском заливе и их расположения по отношению к распределению значений фактора терригенного стока (см. рис. 5, 9).

В задачи работы не входило выделение и описание сообществ макрозообентоса рассмотренной акватории, поэтому мы изначально усложнили работу – выделили шесть ассоциаций донных животных, вместо пяти. Явные «звонки», предупреждающие о неправильности принимаемых решений, звучали и из результатов процедуры SIMPER и, особенно, при оценке приуроченности сообществ к различным условиям среды: по всем гидробиологическим канонам включение в одну группу станций с совершенно разным грунтом «не лезет ни в какие ворота». Цель такого усложнения была двояка – с одной стороны, подчеркнуть необходимость осторожности и внимательности при выполнении различных процедур и тестов с последующей формулировкой выводов, с другой – показать, что очень важно

последовательно, шаг за шагом, выполнить их все: результаты последнего – прокрустова – теста тому подтверждение. Следует также отметить, что скомпилированный нами алгоритм – весьма быстрая процедура; при некотором опыте за рабочий день можно «обработать» результаты нескольких съемок. Раньше на это уходило до нескольких дней, а то и недель, причем всегда оставалось сомнение в правильности выбранного решения. В то же время это далеко не полный набор методов, которые можно, а иногда и нужно, применять: с их многообразием можно познакомиться в уже упомянутых работах Я. Оксанена, Р.И. Кабакова, В.К. Шитикова, С.Э. Мاستицкого, Г.С. Розенберга [3, 6, 15, 20] и, естественно, на просторах интернета.

Выводы

Разработанный алгоритм позволяет довольно легко и оперативно производить выделение сообществ донных животных и получать статистические подтверждения справедливости полученной классификации, выявлять основные факторы среды, ведущие к дифференциации этих группировок, что также сопровождается оценками с позиций статистики.

Классификация, выраженная в численной форме и подтвержденная результатами статистического анализа, придает исследователю уверенность в своих дальнейших действиях по описанию полученных групп и их свойств, а также при последующих манипуляциях с этими группами, например, при разработке методов оценки качества морской среды на основе структуры самих сообществ.

Важным условием использования данного алгоритма является осторожность и внимательность при выполнении различных процедур и тестов, а также соблюдение последовательности и полноты их применения. Естественно, область приложения скомпилированных операций значительно шире, чем только макрозообентос: алгоритм может быть использован при анализе структуры самых разнообразных природных комплексов, как растительных, так и животных.

ЛИТЕРАТУРА

1. Бигон М., Харпер Дж., Таунсенд К. Экология. Особи, популяции и сообщества. Том 2. – М.: Мир, 1989. – 477 с.
2. Звягинцев А.Ю., Мощенко А.В. Морские техноэкосистемы энергетических станций. – Владивосток: Дальнаука, 2010. – 343 с.
3. Кабаков Р.И. **R** в действии. Анализ и визуализация данных в программе **R**. – М.: ДМК Пресс, 2014. – 588 с.
4. Ким Дж., Клекка У.Р., Мьюлер К.У. Факторный, дискриминантный и кластерный анализ. – М.: Финансы и статистика, 1989. – 312 с.
5. Латыпов Ю.Я., Даутова Т.Н., Мощенко А.В. Принципы и методы классификации квидарий. – Владивосток: Дальнаука, 1998. – 244 с.
6. Мاستицкий С.Э., Шитиков В.К. Статистический анализ и визуализация данных с помощью **R**. –

Электронная книга, адрес доступа: <http://r-analytics.blogspot.com>. – 2014. – 401 с.

7. Мощенко А.В. Роль микромасштабной турбулентности в распределении и изменчивости бентосных животных. – Владивосток: Дальнаука, 2006. – 321 с.

8. Мощенко А.В., Белан Т.А., Ивин В.В. Сообщества макрозообентоса восточной части пролива Босфор Восточный (залив Петра Великого Японского моря) // Изв. ТИНРО. – 2018. – Т. 193. – С. 112–142.

9. Мощенко А.В., Белан Т.А., Борисов Б.М., Лишавская Т.С., Севастьянов А.В. Современное загрязнение донных отложений и экологическое состояние макрозообентоса в прибрежной зоне Владивостока (залив Петра Великого Японского моря) // Изв. ТИНРО. – 2019. – Т. 196. – С. 155–181.

10. Мощенко А.В., Белан Т.А., Лишавская Т.С., Севастьянов А.В., Борисов Б.М. Многолетняя динамика концентраций приоритетных поллютантов и общего уровня химического загрязнения прибрежных акваторий Владивостока (залив Петра Великого Японского моря) // Изв. ТИНРО. – 2020. – Т. 200. В печати.

11. Мощенко А.В., Белан Т.А., Олейник Е.В. Влияние загрязнения на сообщества рыхлых грунтов в заливе Петра Великого Японского моря: выбор переменных // Ecological studies and the state of the ecosystem of Amursky Bay and the estuarine zone of the Razdolnaya River (Sea of Japan). Volume 2. – Vladivostok: Dalnauka, 2009. – P. 173–204.

12. Несис К.Н. Общие экологические понятия в приложении к морским сообществам. Сообщество как континуум // Биология океана. Т. 2. Биологическая продуктивность океана. – М.: Наука, 1977. – С. 5–13.

13. Пропп М.В. Экология прибрежных донных сообществ Мурманского побережья Баренцева моря. – Л.: Наука, 1971. – 128 с.

14. Тищенко П.Я., Лобанов В.Б., Звалинский В.И., Сергеев А.Ф., Волкова Т.И., Колтунов А.М., Михайлик Т.А., Сагалаев С.Г., Тищенко П.П., Швецова М.Г. Сезонная гипоксия Амурского залива (Японское море) // Изв. ТИНРО. – 2011. – Т. 165. – С. 136–157.

15. Шитиков В.К., Розенберг Г.С. Рандомизация и бутстреп: статистический анализ в биологии и экологии с использованием R. – Тольятти: Кассандра, 2013. – 314 с.

16. AZTI Marine Biotic Index. – Электронный источник: <http://ambi.azti.es>.

17. Belan T.A. Benthos abundance pattern and species composition in conditions of pollution in Amursky Bay (the Peter the Great Bay, the Sea of Japan // Mar. Pollut. Bull. – 2003. – Vol. 49, № 9. – P. 1111–1119.

18. Borja Á., Mader J., Muxika I. Instructions for the use of the AMBI index software (Version 5.0) // Revista de Investigacion Marina. – 2012. – Vol. 19, № 3. – P. 71–82.

19. Moshchenko A.V., Belan T.A. Near-bottom environmental conditions and macrobenthos of the inner part of Amursky Bay (Peter the Great Bay, Japan Sea) // Pacific Oceanography. – 2005. – Vol. 3, № 2. – P. 121–136.

20. Oksanen J. Multivariate Analysis of Ecological Communities in R: vegan tutorial. – URL: cc.oulu.fi/~jarioksa/opetus/metodi/vegantutor.pdf. – 2015. – 43 p.

21. UNEP: Statistical analysis and interpretation of marine community data: Reference methods for Marine Pollution Studies. – 1995. – № 64. – 75 p.

УДК 504.42.

Экологическое состояние морской среды и макрозообентоса у южной оконечности полуострова Муравьева-Амурского / Мощенко А.В., Белан Т.А., Борисов Б.М., Лишавская Т.С., Севастьянов А.В. // *Труды ДВНИГМИ. – 2020.*

Абстракт. Представлена последовательность процедур и тестов, необходимых и достаточных для выделения группировок донных животных, которая основана на концепции сообщества как «относительного континуума между относительными перерывами континуума». Алгоритм включает следующие этапы. (1) Выделение сообществ – кластерный анализ различными методами агломерации на основе метрики Брея-Кертиса и процедуры бутстрепинга для оценки вероятности появления узла на дендрограмме. (2) Классификация станций, нерасклассифицированных на первом этапе, с использованием алгоритма, основанного на теории нечетких множеств. (3) Статистическая оценка принадлежности проб (станций) к той или иной группе – процедура ANOSIM. (4) Оценка статистической значимости связи обилия видов с факторами среды и географическим положением – тест Мантеля. (5) Наименование сообществ – выделение видов, вносящих наибольший вклад во внутрigrупповое сходство – процедура SIMPER; (6) Определение ведущих факторов среды, ответственных за формирование сообществ методом канонического анализа соответствий (Constrained CA) с последовательным удалением статистически незначимых параметров на основе применения теста перестановок и пошагового дисперсионного анализа. (7) Визуализация дифференциации сообществ и воздействия факторов среды при помощи процедуры нелинейного многомерного шкалирования – процедура NMDS и ряд других процедур и тестов. Использование данного алгоритма иллюстрируется рядом примеров, полученных при анализе реальных данных.

Ключевые слова: сообщество, макрозообентос, кластерный анализ, бутстреп, ANOSIM, SIMPER, тест Мантеля, CCA, NMDS

Табл. 8. Ил. 10. Библ. 21

To the methodology of the differentiation of macrozoobenthic soft-bottom communities: simple algorithm for ecological monitoring practice / Moshchenko A.V., Belan T.A., Borisov B.M., Lishavskaya T.S., Sevastyanov A.V. // *FERHRI Issues. – 2020.*

Abstract. A sequence of procedures and tests that are indispensable and sufficient for extraction of bottom-animals groupings is presented. This sequence is based on conception of community as «relative continuum between relative continuum discontinuities». Algorithm includes the following steps. (1) Community extraction using cluster analysis with different methods of agglomeration, Bray-Curtis distance as a metrics and bootstrapping for evaluation of probability of the nod appearance at the dendrogram. (2) Classification of stations that were not classified at the first stage using algorithm based on the theory of fuzzy sets. (3) Statistical estimation of belonging of samples (stations) to one or another group (ANOSIM). (4) Evaluation of statistical significance of species abundance connection with environmental factors and geographical location (Mantel's test). (5) Designation of communities – selection of the species that input maximally into withingroup similarity (SIMPER); (6) Determination of the leading environmental factors responsible for communities formation using canonical correspondence analysis (Constrained CA) with sequential removal statistically insignificant parameters on the base of permutation test and stepwise ANOVA. (7) Visualization of differentiation of communities and effects of environmental factors by means of non-linear multidimensional scaling (NLMS) and some other procedures and tests. Employment of this algorithm is illustrated by a number of examples obtained under analysis of real data.

Key words: community, macrozoobenthos, cluster analysis, bootstrapping, ANOSIM, SIMPER, Mantel's test, CCA, NMDS

Tabl. 8. Fig. 10. Bibl. 21