

УДК 574.587+578.087(282.247.431.2)+595.7(476)2

МЕТОДЫ СРАВНИТЕЛЬНОГО АНАЛИЗА ВИДОВОГО СОСТАВА НАСЕКОМЫХ РАЗЛИЧНЫХ МЕСТООБИТАНИЙ С ИСПОЛЬЗОВАНИЕМ ПРОГРАММНОЙ СРЕДЫ R

Г. Г. СУШКО¹⁾

¹⁾*Витебский государственный университет имени П. М. Машерова,
пр. Московский, 33, 210038, г. Витебск, Беларусь*

Анализ публикаций по экологии сообществ, включающих насекомых, доступных в международной базе данных PubMed Central, свидетельствует, что наиболее распространенной схемой оценки видового состава различных местообитаний является сравнение с помощью многофакторного дисперсионного анализа (PERMANOVA) и теста ANOSIM, визуализация выявленных различий с применением ординации (NMDS). Завершает схему выявления видов, в наибольшей мере обуславливающих различия в выборках, способ с использованием процедур IndVal и SIMPER. Данные приемы анализа возможно реализовать в программной среде R, которая в настоящее время является общепризнанным стандартом в научных публикациях в мире. Дан краткий обзор методов сравнительного анализа видового состава на примере насекомых за последние пять лет. Приведена последовательность команд для их реализации в программной среде R с использованием результатов собственных исследований.

Ключевые слова: видовой состав; сравнительный анализ; насекомые; экология сообществ; программная среда R.

METHODS OF THE COMPARATIVE ANALYSIS OF THE INSECT SPECIES COMPOSITION IN DIFFERENT HABITATS USING THE SOFTWARE ENVIRONMENT R

G. G. SUSHKO^a

^a*Vitebsk State University named after P. M. Masherova,
33 Moscow Avenue, Vitebsk 210038, Belarus*

Analysis of publications on the ecology of communities, in focus insects, over the past five years, available in the international database PubMed Central, showed that the most common scheme for assessing the species composition of different habitats are comparison using multivariate analysis of variance (PERMANOVA) and ANOSIM test, visualization of the differences using ordination (NMDS). The scheme is completed by identifying the species that are most responsible for the differences in the samples, using the IndVal and SIMPER procedures. All these methods can be implemented in the R software environment, which is currently a generally recognized standard in scientific publications in the world. The presented article includes a brief overview of the methods of comparative analysis of the insect species composition. The program code for this implementation in R using the results of our own research is also given.

Keywords: species composition; comparative analysis; insects; community ecology; software environment R.

Образец цитирования:

Сушко ГГ. Методы сравнительного анализа видового состава насекомых различных местообитаний с использованием программной среды R. *Журнал Белорусского государственного университета. Экология.* 2021;2:21–28.
<https://doi.org/10.46646/2521-683X/2021-2-21-28>

For citation:

Sushko GG. Methods of the comparative analysis of the insect species composition in different habitats using the software environment R. *Journal of the Belarusian State University. Ecology.* 2021;2:21–28. Russian.
<https://doi.org/10.46646/2521-683X/2021-2-21-28>

Автор:

Геннадий Геннадьевич Сушко – доктор биологических наук, доцент; заведующий кафедрой экологии и географии.

Author:

Gennadi G. Sushko, doctor of science (biology), docent; head of the department of ecology and geography.
gennadis@rambler.ru

Введение

Одной из важнейших и часто используемых процедур при оценке структурно-функциональной организации экологических систем является сравнение их видового состава. Во второй половине прошлого века в арсенале исследователей появились схемы, включающие анализ структуры доминирования или встречаемости видов с использованием специальных шкал, кластерный анализ и индексы сходства (Жаккара, Сьеренсена и др.), которые учитывают различия в выборках [1]. В связи с развитием компьютерных технологий, а также с разработкой и постоянным усовершенствованием специализированного программного обеспечения, начали появляться новые возможности, способствующие в наше время более детальному анализу данных. Однако есть и негативные моменты. В частности, серьезной проблемой можно считать недостаточное количество литературы на русском языке, посвященной методам сравнительного анализа видового состава [2; 3]. Отдельные методы можно обнаружить в монографиях В. К. Шитикова с соавторами «Макроэкология речных сообществ: концепции, методы, модели» (2012) [4], «Рандомизация и бутстреп: статистический анализ в биологии и экологии с использованием R» [5]. Детальный обзор методик сравнительного анализа видового состава ассамблей растений и животных можно найти преимущественно в англоязычной литературе [6; 7].

В наши дни для анализа данных, в том числе и в биологии, наиболее мощной платформой становится статистическая среда R. Команда разработчиков со всего мира постоянно создает и поддерживает дополнительные пакеты, значительно расширяющие ее возможности [8]. Некоторые из них разработаны и для экологических исследований. В отличие от R, большинство коммерческих программных продуктов, работающих по достаточно жестким шаблонам, далеко не всегда могут удовлетворить современного исследователя. Единственное их достоинство для биологов, не владеющих навыками программирования, – отсутствие необходимости писать программный код. Однако за последнее десятилетие вышли в свет несколько книг по R на русском языке, которые позволяют восполнить этот пробел [8–10]. На современном этапе исследований в биологии требуется комбинировать различные типы анализа и визуализировать их результаты. Поэтому, бесплатная программная среда R, которая позволяет решать широкий круг вопросов, является безусловным лидером среди свободно распространяемых систем статистического анализа. Ведущие университеты мира, аналитики крупнейших компаний и исследовательских центров постоянно используют R при проведении расчетов и создании крупных информационных проектов. Кроме того, R становится общепризнанным стандартом в научных публикациях [9].

Для экологов R также представляет широкие возможности, сочетающие как общепринятые, так и новые методики анализа данных, в том числе и касающиеся сравнительной характеристики видового состава. В частности, пакеты *vegan* (Community Ecology Package), *BiodiversityR* (Package for Community Ecology and Suitability Analysis), *labdsv* (Ordination and Multivariate Analysis for Ecology) и др. [11; 12].

Одними из наиболее популярных модельных объектов в экологических исследованиях являются насекомые вследствие их многочисленности и высокого видового богатства. В зарубежных публикациях по энтомологии и экологии насекомых применяется достаточно широкий спектр методов сравнения видового состава. Однако зачастую у исследователя, особенно начинающего, возникают проблемы выбора конкретного метода и его целесообразности в определенном случае. В связи с этим целью данной работы было обобщение основных методик сравнительного анализа видового состава насекомых, используемых в настоящее время в синэкологических исследованиях, а также методики их реализации в статистической среде R.

Материалы и методы исследования

В качестве материала для обобщения методик сравнительной оценки видового состава насекомых нескольких местообитаний послужил анализ статей, размещенных в открытом доступе в международной базе данных PubMed Central (PMC) [13]. Проанализировано 350 работ за последние 5 лет по таким показателям, как методы ординации и статистические тесты и процедуры.

Рассмотрены возможности и ограничения к применению наиболее часто используемых методов сравнительного анализа видового состава. На основании этого предпринята попытка формирования наиболее оптимальной схемы анализа и ее демонстрация на примере результатов собственных исследований с использованием общедоступного программного обеспечения R (version 2.15.2, R Development Core Team 2010) [14].

Результаты исследования и их обсуждение

Анализ научных источников выявил, что в 62,85 % публикаций, где сравнивается видовой состав насекомых нескольких (более двух) местообитаний, для визуализации используются методы ординации. В 48,57 % из их числа таким методом было неметрическое многомерное шкалирование (NMDS). Реже применялись анализ главных координат – PCoA (8,57 %) и главных компонент – PCA (5,71 %).

Для проверки достоверности различий видового состава нескольких местообитаний использовались тесты, основанные на статистических гипотезах, – это многофакторный дисперсионный анализ

PERMANOVA (25,71 %) и анализ сходства ANOSIM (22,85 %). Процедуры для выявления наиболее характерных видов, вносящих различия между ассамблеями, были выявлены в анализируемых публикациях реже. В их числе процедуры IndVal (Indicator Value) (14,28%) SIMPER (Similarity Percentages) (5,71 %).

Таким образом, наиболее распространенной схемой анализа является сравнение видового состава на наличие различий с помощью тестов, основанных на таких статистических гипотезах, как многофакторный дисперсионный анализ (PERMANOVA) и анализ сходства (ANOSIM), визуализация выявленных различий с применением ординации (NMDS). Завершает схему выявление видов в наибольшей мере обусловливающих различия в выборках, как правило, соответствующих определенным местообитаниям, с использованием процедур IndVal и SIMPER.

Краткая характеристика основных этапов сравнения видового состава. Проверку того, значимо ли отличаются группы объектов (строки, соответствующие выборкам, полученным в нескольких биотопах) на основе набора их атрибутов (столбцов, соответствующих видам) в многомерной таблице данных можно осуществить с помощью тестов, основанных на статистических гипотезах (H_0 – достоверные различия отсутствуют, H_a – есть достоверные различия). Как указано выше, чаще всего для сравнения различий видового состава используются тесты PERMANOVA и ANOSIM, которые не требуют нормального распределения. Большинство экологических данных, включающих перечень видов и их обилий, не соответствует закону нормального распределения, что обусловлено высоким количеством видов с низкой численностью, преобладанием нескольких видов и наличием большого числа нулевых значений в наборе данных [15].

Тест PERMANOVA (Permutational multivariate analysis of variance) является непараметрическим аналогом многомерного дисперсионного анализа (MANOVA). Метод получил широкое распространение в исследованиях по экологии сообществ, так как не требует нормального распределения. PERMANOVA позволяет использовать любую меру расстояния между объектами, а не только расстояние Эвклида, и терпимо относится к не независимым переменным. Используемая статистика теста представляет собой соотношение псевдо-F, аналогичное F в ANOVA [15–17]. Для выполнения PERMANOVA в R предназначена функция *adonis()* в пакете *vegan* [18].

Тест ANOSIM (Analysis Of Similarities). Эта непараметрическая процедура проверяет значимые различия между двумя или более группами на основе определенной меры расстояния. Она сравнивает ранги расстояний между группами с рангами расстояний внутри групп. На основе средних значений этих двух типов рангов формируется итоговая статистическая мера R-тест. Если $R = 1$, то сообщества очень сильно различаются, если $R = 0$ – они полностью сходны. Значения R больше 0,75 принято интерпретировать как хороший уровень различий, R больше 0,5 – наличие различий, R меньше 0,25 – как едва различимые различия. Значения R ниже 0 предполагают, что различия внутри групп больше, чем различия между группами [6]. Тест делает меньше предположений, чем MANOVA, поскольку он основан на рангах расстояний. Вследствие этого он хорошо применим для сравнения данных выборочных совокупностей по видовому составу, где сравниваются группы выборок, соответствующие различным биотопам. Для этого все группы должны иметь сопоставимые внутригрупповые дисперсии, чтобы избежать ложно значимых результатов. При сравнении выборок с высокими различиями дисперсий значения R могут быть сильно завышены, даже если нет реальных различий [19]. Для выполнения ANOSIM в R предназначена функция *anosim()* в пакете *vegan* [18].

Визуализация различий видового состава, выявленного с помощью тестов, как правило, осуществляется с помощью ординации. Ранее мы выяснили, что используются такие методы, как NMDS, PCoA и PCA. Среди них реже всего применяется анализ главных компонент, так как он требует нормального распределения, линейной зависимости и относительно коротких градиентов в структуре данных. Такие данные в экологических исследованиях редки, но возможны. К тому же улучшить ситуацию может их преобразование, например логарифмирование, что делает использование PCA аргументированным. С другой стороны, анализ главных компонент основан только на мере расстояния Эвклида, что также ограничивает его применение для сравнения видового состава [16; 17]. Анализ главных координат в этом отношении является более гибким, так как позволяет выбирать меру расстояния исследователю самостоятельно. Однако в PCoA наличие в строках таблицы данных сильно варьирующих значений (в нашем случае обилие определенного вида) может заметно различаться в ряду местообитаний и сильно влиять на ординацию, делая вклад других объектов (других видов) менее заметными [7]. Для выполнения PCoA в R предназначены функции *cmdscale()* и *ordiplot()* из пакета *vegan*, функция *pcoa()* в пакете *ape* [11; 12; 18].

Ординация NMDS (Non-metric multidimensional scaling) оказалась наиболее подходящим и часто применяемым методом для сравнения видового состава насекомых, зарегистрированных в различных биотопах. Это обусловлено тем, что неметрическое многомерное шкалирование не требует нормального распределения данных и позволяет выбрать из множества различных мер подходящее расстояние. В экологии, чаще всего, используют расстояние Брея–Кертиса, которое учитывает как наличие/отсутствие видов, так и их численность. Графики NMDS – отличный инструмент, потому что могут сжать информацию о большом количестве видов и их обилиях в различных местообитаниях в ординационном пространстве, отражая их реальные тренды биотопического распределения. На графике NMDS точки соответствуют видам. Чем ближе

точки расположены в пространстве ординации, тем более схожи их сообщества [2; 6]. Для выполнения NMDS в R предназначены функции *metaMDS()*, *ordiplot()*, *ordisurf()*, and *envfit()* пакета *vegan* [18].

Для констатации различий видового состава в разных биотопах достаточно результатов тестов ANOSIM или PERMANOVA. Однако если вы хотите более детально выяснить, какие виды статистически более многочисленны в одной группе по сравнению с другой и ответственны за различия видового состава, то можно выполнить такие анализы, как SIMPER или IndVal [11; 12].

SIMPER (Similarity Percentage). Это простой метод оценки того, какие таксоны (виды) вносят наибольший вклад (в %) в различия анализируемых сообществ. Для сравнения выборок чаще всего используется мера сходства Брея–Кертиса, но также можно применить и другие меры расстояния. Если анализируется более двух групп, вы можете сравнить по две группы (попарно) между собой по желанию, выбрав их из списка. Также можно объединить все выборки для выполнения общего анализа. При этом рассчитывается общий уровень различий для всех выборок [6]. Для выполнения NMDS в R предназначена функция *simper()* пакета *vegan* [18].

IndVal (Indicator Value). Выявление по встречаемости или численности небольшой группы индикаторных видов является полезным при долгосрочном мониторинге окружающей среды в целях сохранения биоразнообразия. Процедура IndVal предназначена для измерения связи между видом и биотопом или группой биотопов. Сначала вычисляется индекс IndVal между видами и каждой группой участков одного типа местообитаний, а затем выделяется группа видов, ассоциированная в наибольшей степени с определенным типом. Значимость этой связи (p) проверяется с помощью теста перестановки (permutation test) [20]. Существует специальный пакет R для выявления видов-индикаторов *indicspecies*, в котором наиболее часто используется функция *multipatt()*. Данная функция приводит к списку видов в выборках, связанных с определенным типом местообитаний [21].

Реализация основных этапов сравнения видового состава в статистической среде R. Для демонстрации охарактеризованных выше методов используем результаты наших исследований видового состава жуков жужелиц (Coleoptera, Carabidae) в трех типах сосновых лесов (PP – *Pinetum pleuroziosum*, PM – *Pinetum myrtillosum*, PL – *Pinetum ledosum*), в каждом из которых получено по 6 выборочных совокупностей. Всего выявлено 24 вида. В табл. данных число особей видов расположены в столбцах, а биотопы в строках: анализируемая матрица включает 24 переменных и 18 наблюдений (рис. 1). Более подробную информацию о создании матриц данных, графиков и загрузке пакетов в RStudio можно найти в специальной литературе [8–12].

Site.ID	Habitat	Car.vio	Car.gla	Car.cor	Car.hor	Car.gra	Car.arv	Cyc.car	Poe.lep	Poe.ver	Pte.obl
1	1 PM	19	8	0	63	1	7	9	0	0	1
2	2 PM	18	8	0	73	0	12	15	0	0	1
3	3 PM	19	7	0	52	0	13	22	0	0	1
4	4 PM	0	0	3	2	0	25	0	0	0	0
5	5 PM	0	0	0	3	0	33	1	0	3	0
6	6 PM	0	0	0	4	0	18	0	0	1	0
7	13 PP	11	0	0	6	0	17	2	0	0	0
8	14 PP	11	0	0	13	0	15	1	1	0	0
9	15 PP	18	0	0	6	0	11	2	0	1	1
10	16 PP	19	1	0	7	0	18	2	0	0	1
11	17 PP	18	0	0	6	0	8	1	1	2	0
12	18 PP	0	1	2	8	0	2	0	0	0	0
13	30 PL	0	2	0	0	3	0	0	4	0	0
14	31 PL	0	0	0	0	4	0	0	7	0	0
15	32 PL	0	0	0	1	0	0	1	1	0	0
16	33 PL	0	0	0	0	0	0	0	5	0	0
17	34 PL	0	0	0	0	2	0	0	2	0	0
18	35 PL	0	2	0	0	3	0	0	4	0	0

Рис. 1. Фрагмент анализируемой матрицы данных (скриншот RStudio)

Fig. 1. Fragment of the analyzed data matrix (screenshot RStudio)

Введя следующий код, выполним тест ANOSIM с использованием пакета *vegan* и функции *anosim()*:

```
library(vegan)
mydata<-read.csv("C:/forR/mydata.csv", sep = ",", head=T)
abund = mydata[,3:ncol(mydata)]
```



```
m_abund = as.matrix(abund)
ano = anosim(m_abund, mydata$Habitat, distance = "bray", permutations = 9999)
ano
summary(ano)
```

Результаты анализа выглядят следующим образом: $R=0.8189$, $Significance=0.0001$. Следовательно, между видовым составом жуужелиц трех биотопов выявлены достоверные и значительные различия.

Проверим различия видового состава жуужелиц другим способом, с помощью PERMANOVA (функция *adonis()* в пакете *vegan*), используя следующий код:

```
library(vegan)
mydata<-read.csv("C:/forR/mydata.csv", sep = ";", head=T)
abund = mydata[,3:ncol(mydata)]
m_abund = as.matrix(abund)
adonis(formula = m_abund ~ Habitat, data = mydata, permutations = 999)
```

На выходе мы получим традиционную таблицу результатов дисперсионного анализа. Исходя из представленных в ней результатов – PERMANOVA ($Df=2$, $F=14.66$, $P=0.01$), также можно констатировать различие видового состава.

Для визуализации различий применим NMDS с использованием функции *metaMDS* пакета *vegan*, выбрав расстояние Брея–Кертиса:

```
library(vegan)
mydata<-read.csv("C:/forR/mydata.csv", sep = ";", head=T)
dat1 <- mydata[,3:24]
dat2 <- mydata[,1:2]
nmds = metaMDS(dat1,distance="bray", k=2)
plot(nmds)
```

После использования *plot(nmds)* мы получаем диаграмму ординации (рис. 2). Таким образом, имеется определенный тренд в расположении объектов, соответствующих биотопам и выборкам, но такой график дает мало информации.

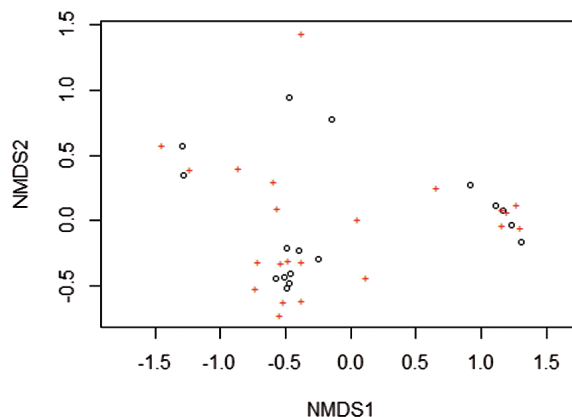


Рис. 2. Диаграмма ординации NMDS, полученная с использованием функции *plot()*

Fig. 2. NMDS ordination diagram obtained using the *plot()* function

Далее применим функцию *ordihull()*, которая дает возможность выделить области со сходными объектами. В нашем примере это будет выглядеть как *ordihull(nmds,groups=dat2\$Habitat,label=TRUE)*. Следовательно, на диаграмме можно увидеть сходство видового состава биотопов, обозначенных как PM и PP, области которых перекрываются, но это сходство не высоко, а также определить значительное отличие биотопа PL от остальных (рис. 3).

Пакет *vegan* обладает рядом других возможностей визуализации [18]. В частности, можно сделать акценты с помощью различных цветов, выделив области в виде эллипсов, центроидов и т. д.; увидеть, какие виды сконцентрированы в определенных областях диаграммы, применив *plot(nmds, type="t")* (рис. 4).

Для повышения информативности графика ординации можно комбинировать несколько функций одновременно, сначала создать пустой график (*type = «n»*), а затем, добавив виды и биотопы, выделить области и др. (рис. 5).

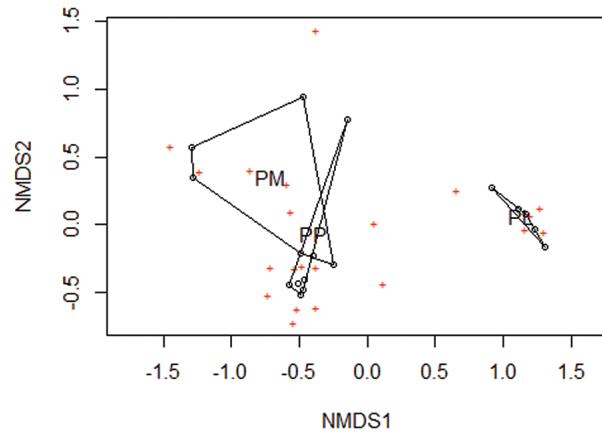


Рис. 3. Диаграмма ординации NMDS, полученная с использованием функции ordihull()

Fig. 3. NMDS ordination diagram obtained using the ordihull() function

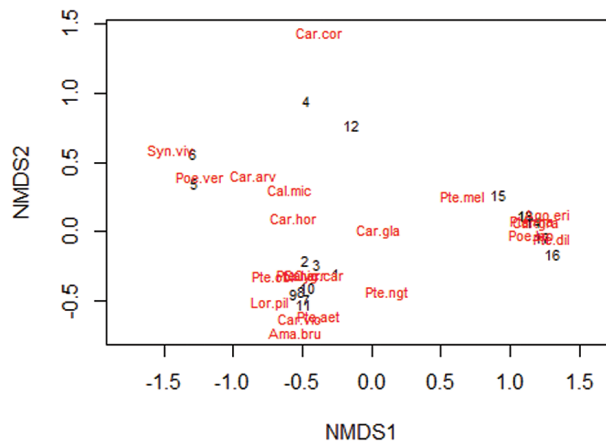


Рис. 4. Диаграмма ординации NMDS, полученная с использованием функции plot(type='t')

Fig. 4. NMDS ordination diagram obtained using the plot(type="t") function

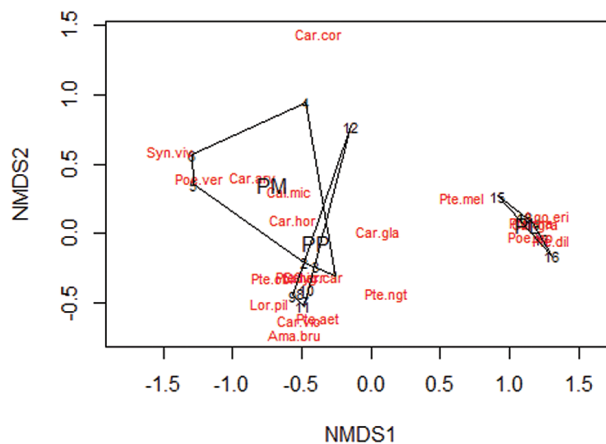


Рис. 5. Диаграмма ординации NMDS, полученная с использованием нескольких функций

Fig. 5. NMDS ordination diagram obtained using the combined function

Следующим этапом анализа различий видового состава может быть выявление видов, наиболее тесно ассоциированных с определенным биотопом, или их группой. Сначала применим процедуру IndVal пакета indicpecies. Для этого используем следующий код:

```
library(indicpecies)
mydata<-read.csv("C:/forR/mydata.csv", sep = ",", head=T)
```

```
abund = mydata[,3:ncol(mydata)]
habitat = mydata$Habitat
inv = multipatt(abund, habitat, func = "r.g", control = how(nperm=999))
summary(inv)
```

Результаты анализа демонстрируют, что 2–3 вида (программа сокращает названия видов до первых трех букв названия рода и названия вида) являются наиболее характерными для каждого из трех местообитаний (табл.).

Значения индекса IndVal могут варьировать от 0 (не является индикаторным видом) до 1 (идеальный индикатор) при статистической значимости $p < 0,01$ [20]. В нашем примере большинство выявленных видов обладают высокой достоверной индикаторной значимостью ($p < 0,01$) и высокими значениями индекса IndVal.

Таблица

Результаты анализа индикаторных видов (IndVal)

Table

Indicator Species Analysis Results (IndVal)

Вид	Биотоп	Значения индекса IndVal	Уровень значимости (p)
<i>Agonum ericeti</i> (Ago. eri)	PL	0,954	0,001
<i>Pterostichus rhaeticus</i> (Pte. rha)	PL	0,843	0,001
<i>Pterostichus diligens</i> (Pte. dil)	PL	0,832	0,002
<i>Calathus micropterus</i> (Cal. mic)	PM	0,779	0,004
<i>Carabus hortensis</i> (Car. hor)	PM	0,606	0,044
<i>Calathus erratus</i> (Cal. err)	PP	0,901	0,002
<i>Amara brunnea</i> (Ama. bru)	PP	0,824	0,006

Далее применим другой метод, который также используют для выявления видов, которые наиболее характерны для конкретного местообитания – SIMPER. Метод можно реализовать с помощью функции *simper()* пакета *vegan*:

```
library(vegan)
mydata<-read.csv("C:/forR/mydata.csv", sep = ";", head=T)
abund = mydata[,3:ncol(mydata)]
m_abund = as.matrix(abund)
sim = simper(m_abund, mydata$Habitat)
sim
summary(sim)
```

Следовательно, результаты анализа, помещенные в табл., демонстрирует совокупные вклады (в %) видов в структуру сообществ (в целях экономии места табл. не приводится; далее указаны только виды, имеющие значение для итогового анализа). В отличие от IndVal, SIMPER сравнивает относительные обилия видов в парах биотопов и выявляет те, которые вносят наибольшие различия, а не указывает на их связь с биотопом. В нашем примере такими парами были PM и PP, PM и PL, PP и PL. Наибольший вклад в различия (average dissimilarity) видового состава в паре PM и PP вносят виды *Pterostichus niger* (12 %), *Carabus hortensis* (10 %), *Calathus micropterus* (7 %), в паре PM и PL – виды *Agonum ericeti* (39 %), *Pterostichus niger* (10 %), *Carabus hortensis* (10 %), в паре PP и PL – *Agonum ericeti* (42 %) и *Pterostichus oblongopunctatus* (8 %). Вклад остальных видов составил менее 1%.

Заключение

В заключение следует отметить, что методология сравнительного анализа видового состава различных местообитаний достаточно ограничена. Это обусловлено не соответствием в большинстве случаев количественных показателей обилия видов закону нормального распределения. Наиболее распространенными в наши дни схемами оценки видового состава являются сравнение с помощью многофакторного дисперсионного анализа (PERMANOVA) и теста ANOSIM, ординации (NMDS) и статистических процедур IndVal и xSIMPER. Все эти методы возможно реализовать в программной среде R, которая в настоящее время является общепризнанным стандартом в научных публикациях в мире и мощным инструментом, сочетающим все рассмотренные виды анализа данных.

Библиографические ссылки

1. Песенко ЮА. *Принципы и методы количественного анализа в фаунистических исследованиях*. Минск: Наука; 1982.
2. Сушко ГГ. Методы многомерного анализа данных в синэкологии насекомых. *Журнал Белорусского государственного университета. Экология*. 2020; 1:38–45.
3. Сушко ГГ. Использование методов анализа данных в энтомологических исследованиях. В: *Сборник статей III Международной научно-практической конференции, памяти В. А. Цинкевича*. Минск: ГНПО «НПЦ НАН Беларуси по биоресурсам»; 2019. с. 381–383.
4. Шитиков ВК, Зинченко ТД, Розенберг ГС. *Макроэкология речных сообществ: концепции, методы, модели*. Тольятти: Кассандра; 2012.
5. Шитиков ВК, Розенберг ГС. *Рандомизация и бутстреп: статистический анализ в биологии и экологии с использованием R*. Тольятти: Кассандра; 2013.
6. Clarke KR. Non-parametric multivariate analysis of changes in community structure. *Australian Journal of Ecology*. 1993;18:117–143.
7. Legendre P, Legendre L. *Numerical Ecology*. 2nd edition. Amsterdam: Elsevier; 1998.
8. Кабаков РИ. *R в действии. Анализ и визуализация данных в программе R*. Москва: ДМК Пресс; 2014.
9. Мاستицкий СЭ, Шитиков ВК. *Статистический анализ и визуализация данных с помощью R*. Москва: ДМК Пресс; 2015.
10. Джеймс Г, Уиттон Д, Хасти Т, Тибширани Р. *Введение в статистическое обучение с примерами на языке R*. Москва: ДМК Пресс; 2016.
11. Borcard D, Gillet F, Legendre P. *Numerical Ecology with R*. Wien: Springer Nature; 2018.
12. Herve M. Aide-mémoire de statistique appliquée à la biologie. Construire son étude et analyser les résultats à l'aide du logiciel R. 2016.
13. PubMed Central (PMC) [Internet]. [Cited 2020 August 15]. Available from: <https://www.ncbi.nlm.nih.gov/pmc/>
14. R Development Core Team (2010) R: A language and environment for statistical computing. R Foundation for Statistical Computing. [Internet]. Vienna: Austria. Available from: <https://www.R-project.org/>
15. Anderson MJ, Walsh DC. PERMANOVA, ANOSIM, and the Mantel test in the face of heterogeneous dispersions: what null hypothesis are you testing? *Ecological monographs*. 2013;83(4):557–574.
16. McCune B, Grace JB. *Analysis of ecological communities*. Glendeden Beach: MjMSoftware Design; 2002.
17. Ramette A. Multivariate analyses in microbial ecology. *FEMS. Microbiological Ecology*. 2007; 62(2):142–160.
18. Oksanen J. *Multivariate Analysis in Ecology*. Lecture Notes. Finland: University of Oulu; 2004.
19. Warton DI, Wright TW, Wang Y. Distance-based multivariate analyses confound location and dispersion effects. *Methods in Ecology and Evolution*. 2012;3:89–101.
20. Dufrene M, Legendre P. Species assemblages and indicator species: the need for a flexible asymmetrical approach. *Ecological Monographs*. 1997;67:345–366.
21. De Caceres M, Legendre P. Associations between species and groups of sites: indices and statistical inference. *Ecology*. 2009;90(12):66–74.

References

1. Pesenko Yu.A. *Principy i metody kolichestvennogo analiza v faunisticheskikh issledovaniyah* [Principles and methods of quantitative analysis in faunistic research]. Minsk: Nauka; 1982. Russian.
2. Sushko GG. Methods of multivariate data analysis in insect synecology. *Journal of the Belarusian State University. Ecology*. 2020;1:38–45. Russian.
3. Sushko GG. *Ispol'zovanie metodov analiza dannyh v jentomologicheskikh issledovaniyah* [Use of data analysis methods in entomological research]. In: *Sbornik statej III Mezhdunarodnoj nauchno-prakticheskoy konferencii, pamjati V. A. Cinkevicha*. Minsk: GNPO «NPTS NAN Belarusi po bioresursam»; 2019. p. 381–383. Russian.
4. Shitikov VK, Zinchenko TD, Rozenberg GS. *Makroekologija rechnyh soobshhestv: koncepcii, metody, modeli* [Macroecology of river communities: concepts, methods, models]. Togliatti: Kassandra; 2012. Russian.
5. Shitikov VK, Rozenberg GS. *Randomizacija i butstrep: statisticheskij analiz v biologii i jekologii s ispol'zovaniem R* [Randomization and bootstrap: a statistical analysis in biology and ecology with R use]. Togliatti: Kassandra; 2014. Russian.
6. Clarke KR. Non-parametric multivariate analysis of changes in community structure *Australian Journal of Ecology*. 1993;18:117–143.
7. Legendre P, Legendre L. *Numerical Ecology*. 2nd edition. Amsterdam: Elsevier; 1998.
8. Kabacoff RI. *R in Action*. Data analysis and graphics with R. Shelter Island, New York: Manning Publications Co.; 2012.
9. Mastitsky SE, Shitikov VK. *Statisticheskij analiz i vizualizacija dannyh s pomoshh'ju R* [Statistical analysis and data visualization using R]. Moscow: DMK Press; 2015. Russian.
10. James G, Witten D, Hastie T, Tibshirani R. *An Introduction to Statistical Learning*. New York: Springer, 2013.
11. Borcard D, Gillet F, Legendre P. *Numerical Ecology with R*. Wien: Springer Nature; 2018.
12. Herve M. Aide-mémoire de statistique appliquée à la biologie. Construire son étude et analyser les résultats à l'aide du logiciel R. 2016.
13. PubMed Central (PMC) [Internet]. [Cited 2020 August 15]. Available from: <https://www.ncbi.nlm.nih.gov/pmc/>
14. R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. [Internet]. Vienna: Austria; 2010. Available from: <https://www.R-project.org/>
15. Anderson MJ, Walsh DC. PERMANOVA, ANOSIM, and the Mantel test in the face of heterogeneous dispersions: what null hypothesis are you testing? *Ecological monographs*. 2013;83(4):557–574.
16. McCune B, Grace JB. *Analysis of ecological communities*. Glendeden Beach: MjMSoftware Design; 2002.
17. Ramette A. Multivariate analyses in microbial ecology. *FEMS. Microbiological Ecology*. 2007; 62(2):142–160.
18. Oksanen J. *Multivariate Analysis in Ecology*. Lecture Notes. Finland: University of Oulu; 2004.
19. Warton DI, Wright TW, Wang Y. Distance-based multivariate analyses confound location and dispersion effects. *Methods in Ecology and Evolution*. 2012;3:89–101.
20. Dufrene M, Legendre P. Species assemblages and indicator species: the need for a flexible asymmetrical approach. *Ecological Monographs*. 1997;67:345–366.
21. De Caceres M, Legendre P. Associations between species and groups of sites: indices and statistical inference. *Ecology*. 2009;90(12):66–74.

Статья поступила в редколлегию 08.02.2021.
Received by editorial board 08.02.2021.