

# Анализ пространственной структуры растительных сообществ с использованием пакетов R

Авторы:

*Анна Мартынова,*

*Владимир Шитиков*

<https://stok1946.blogspot.com/>

## 1. Введение

Отсутствие резких границ между фрагментами фитоценозов и наличие переходных зон между ними предопределили представления о непрерывности (континуальности) растительного покрова. Согласно индивидуалистической концепции, каждый вид растений имеет характерные диапазоны толерантности по отношению ко всей совокупности экологических условий (включая ценотические, физиологические и др.). Поскольку факторы среды обычно изменяются постепенно, как в пространстве, так и во времени, то переход от одной комбинации ценопопуляций к другой осуществляется непрерывно: одни виды постепенно уменьшают своё обилие и исчезают, а другие – появляются и начинают развиваться. Однако континуальное распределение видов по градиенту экологических факторов не исключает возможности образования более или менее выраженных дискретных группировок, как на уровне отдельного яруса, так и сообщества в целом. Все это обуславливает различные формы **неоднородности** структуры надпочвенного покрова (см. Мазинг, 1965), которая в пространственном контексте проявляется в том, что некое количественное или качественное свойство растительного сообщества статистически значимо изменяется от одного участка к другому.

Одна из важнейших задач использования математических методов для анализа геоботанических данных состоит в оценке степени выраженности, достоверности и скорости изменения плотности и видовой структуры растительного покрова относительно географических направлений и/или экологических градиентов на различных масштабных уровнях. На сегодняшний день появилось множество подходов, позволяющих получить численные оценки неоднородности пространственной структуры и использовать их для понимания экологических процессов. По одной из классификаций эти методы в зависимости от типа исходных данных могут быть разделены на точечные и площадные методы.

Точечные методы иногда оперируют неделимыми объектами - отдельными растениями, для которых известны физические координаты местоположения каждого индивида. Тогда структуру сообщества и её неоднородность можно описать с помощью статистического распределения дискретных точек. Если вычленить индивидуальные растения невозможно, то используют непрерывные переменные (проективное покрытие, обилие вида или доля эколого-ценотической группы), измеренные в случайно или регулярно выбранным точкам пробных площадей, после чего выборочные данные анализируют методами геостатистики (Савельев и др., 2012).

Площадные методы оперируют с протяженными в пространстве объектами растительного покрова: трансектами или сетками смежных ячеек, содержащими некую количественную или качественную информацию о растительном покрове (трансектный анализ, анализ сопряженных квадратов, патч-анализ и др.).

Для анализа пространственной структуры фитоценозов широко используются пакеты статистической среды R. К сожалению, нам не удалось найти полноценного обзора специализированных пакетов применительно к растительным сообществам, однако большая их часть анонсирована в обзорах «[Анализ данных по экологии и факторам окружающей среды](#)» и «[Анализ пространственных данных](#)». К ним относятся, в первую очередь, пакеты, осуществляющие анализ пространственного распределения точечных данных – spatstat, выполняющие формирование карт локальных

плотностей – graphics, rasterVis, реализующие многомерный статистический анализ и проверку статистических гипотез – vegan, ade4 и т.д. (Шитиков, Розенберг, 2014). Использование некоторых из них мы показываем ниже на конкретных примерах.

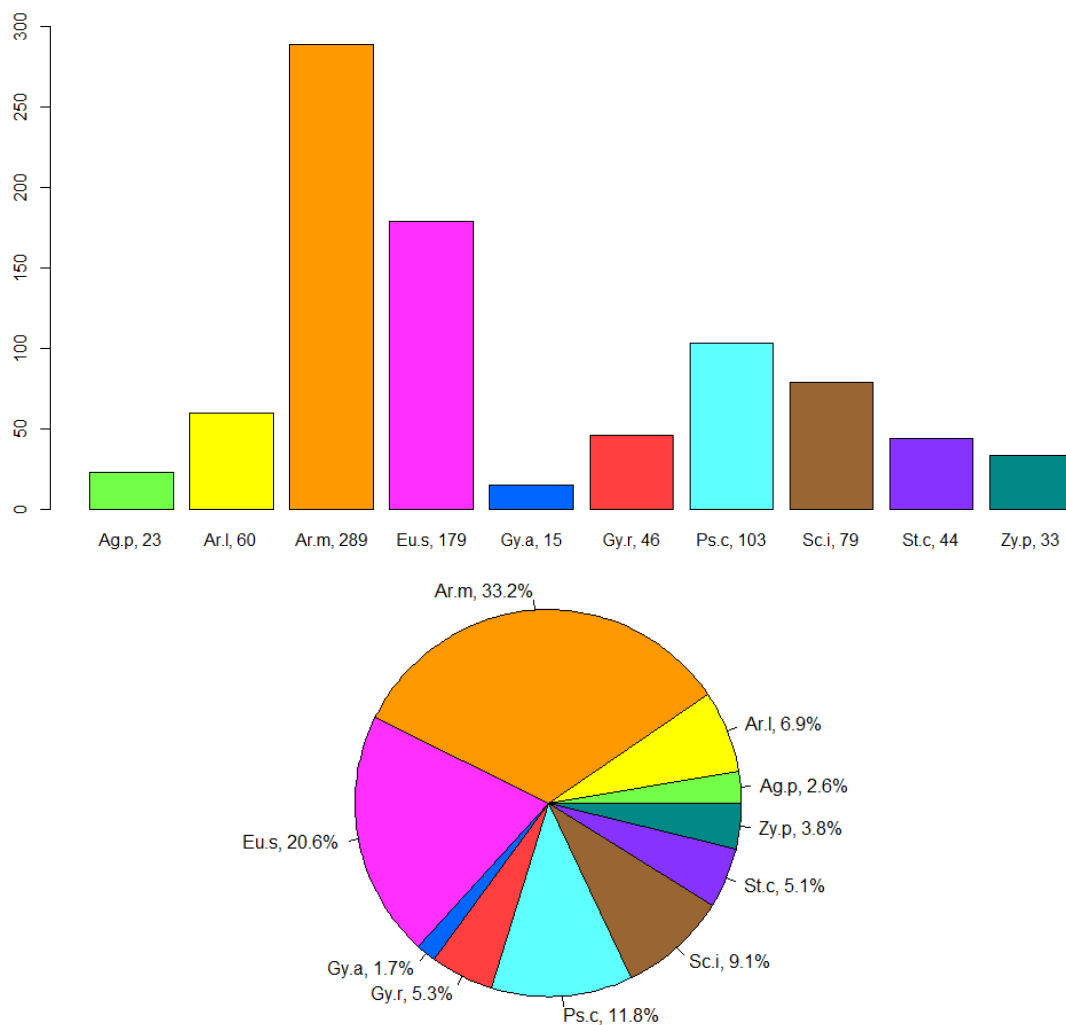
## 2. Характеристика исходных данных

Мы используем данные геоботанического описания площадки размером 8×13 м, выполненного в ходе исследований травянистого покрова в местах выхода гипсовых пород на территории Кююргазинского р-на респ. Башкортостан летом 2021 г. На учетной площадке было проведено картирование растительного сообщества и для каждого растения определены видовая принадлежность и прямоугольно-декартовы координаты. Всего было зарегистрировано 871 растений, принадлежащих к 10 видам и 6 семействам. Исходные данные представлены в виде фрейма и помещены в файл для использования в среде R, который можно скачать с общедоступного ресурса <http://www.ievbras.ru/ecostat/Kiril/R/Phyto.RData> и перенести в рабочий каталог.

В рамках нашего сообщения мы преследуем преимущественно методические цели и не касаемся задач, решаемых относительно предметной области. Поэтому общую постановку задачи исследования, полную расшифровку аббревиатуры видов растений, обсуждение полученных результатов и выводы читатель может найти в отчете А. Мартыновой «Пространственная структура растительного сообщества на обнажениях гипса», также помещенном на нашем ресурсе.

Построим некоторые диаграммы частотного распределения растений по видам:

```
# Мартынова Анна,
# Гимназия №1 им. Н.Т. Антошкина, 10 кл.
# г. Кумертау, Республика Башкортостан
# =====
# Считываем данные по координатам растений
# и их видовой принадлежности в фитоценозе
load("Phyto.RData")
str(Phyto)
'data.frame':  871 obs. of  3 variables:
 $ X : num  7.92 7.92 7.89 7.34 7.36 7.18 7.25 6.96 6.93 6.93 ...
 $ Y : num  0.04 0.1 0.18 0.07 0.25 0.18 0.31 0.36 0.29 0.18 ...
 $ Sp: Ord.factor w/ 10 levels "Ag.p"<"Ar.l"<...: 2 2 9 2 7 7 3 6 8 5 ...
# Цветовая палитра для видов
SpColors <- c(
  rgb(115,255,71,maxColorValue = 255), # Ag.p
  rgb(255,255,0,maxColorValue = 255), # Ar.l
  rgb(255,153,0,maxColorValue = 255), # Ar.m
  rgb(255,47,255,maxColorValue = 255), # Eu.s
  rgb(0,102,255,maxColorValue = 255), # Gy.a
  rgb(255,63,63,maxColorValue = 255), # Gy.r
  rgb(93,255,255,maxColorValue = 255), # Ps.c
  rgb(153,102,51,maxColorValue = 255), # Sc.i
  rgb(135,50,255,maxColorValue = 255), # St.c
  rgb(0,137,134,maxColorValue = 255)) # Zy.p
#Индекс Шеннона
SpNumber <- table(Phyto$Sp)
-sum(SpNumber/sum(SpNumber)*log2(SpNumber/sum(SpNumber)))
[1] 2.801331
# Количество растений каждого вида
barplot(SpNumber,col=SpColors,ylim=c(0,300),
  names.arg=paste0(names(SpNumber)," ",as.character(SpNumber)))
# Доля растений каждого вида в %
round(SpNumber/sum(SpNumber)*100,1)
pie(SpNumber,col = SpColors,labels = paste0(names(SpNumber)," ",
  round(SpNumber/sum(SpNumber)*100,1),"%"))
```



### 3. Анализ пространственной структуры фитоценоза

Одной из основных задач изучения пространственной структуры является визуальный анализ и графическое представление пространственной неоднородности распределения популяций с помощью карт локальных плотностей. Локальные значения плотности  $\lambda \equiv \lambda(X)$ , где  $X = (x, y)$  – произвольная точка, могут быть получены методом аппроксимации сглаживающей поверхностью (Bailey, Gatrell, 1995) на основе заданной ядерной функции (Silverman, 1986; Scott, 1992). При этом скользящее "окно" сканирует всю область исследований и суммирует вклады точек в окрестности, связанной с текущим положением курсора. В качестве ядерной функции используем функцию Епанечникова (Епанечников, 1969), а ширину "скользящего окна", равную 0,75 м по  $x$  и 1,08 м по  $y$ , определим по правилу Скотта (Scott, 1992).

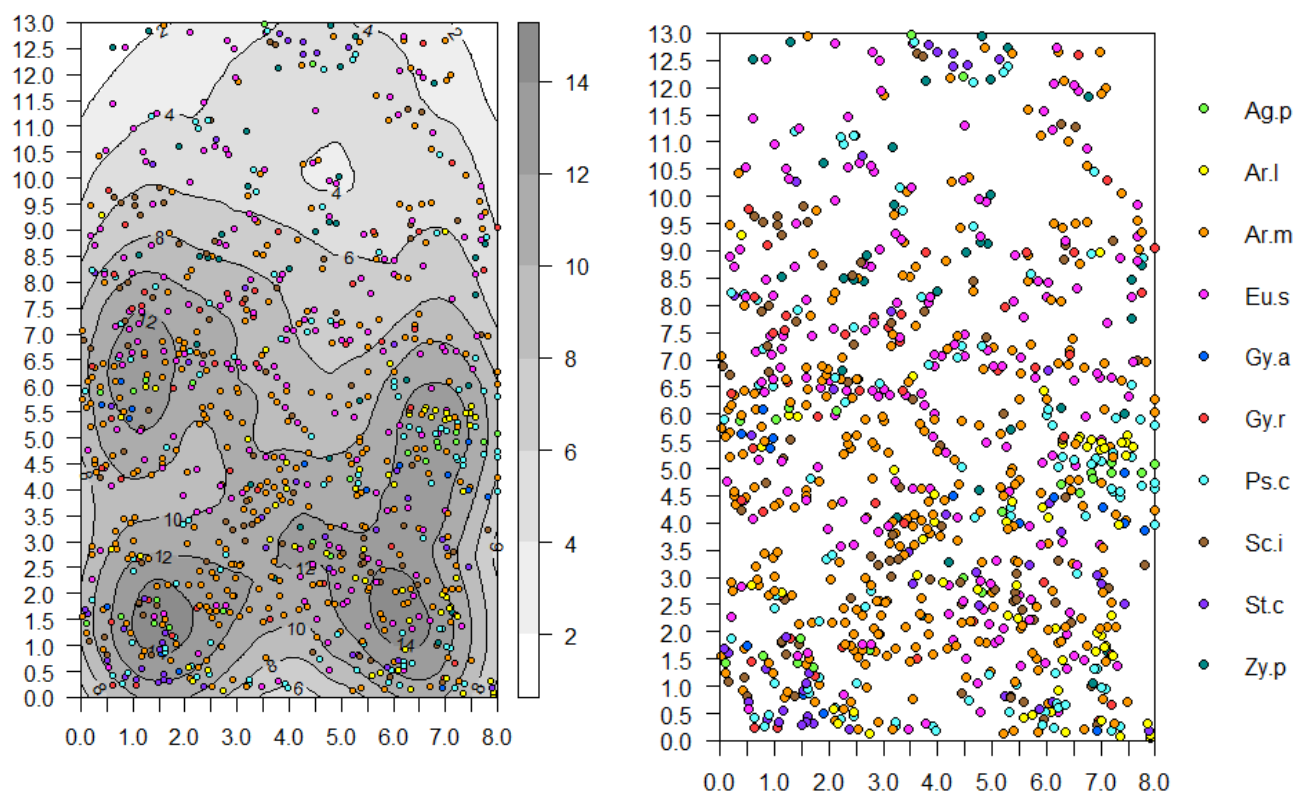
Для построения карт плотности и анализа её локальных изменений учетную площадку разобьем на квадраты размером 5×5 см (далее – элементарные квадраты), в каждом из которых значение плотности принято равным её значению в центре квадрата. Таким образом, определим локальные значения плотности, как для всей массы растений, так и для каждой популяции вида. Для анализа изменений плотности найдем ранговые (min, max) и средние (mean) её локальных значений.

```
# !!! НЕОБХОДИМО ПОДКЛЮЧИТЬ ПАКЕТ SPATSTAT !!!
library(spatstat)
# Создаем точечный паттерн (ppp-object)
PP <- ppp(x=Phyto$X, y=Phyto$Y, marks=Phyto$Sp, window=owin(c(0, 8),
  c(0, 13), unitname=c("metre", "metres")))
```

```

# Средняя плотность
intensity(unmark(PP)) # без учета видовой принадлежности
[1] 8.375
round(intensity(PP),2)# по каждому виду
Ag.p Ar.l Ar.m Eu.s Gy.a Gy.r Ps.c Sc.i St.c Zy.p
0.22 0.58 2.78 1.72 0.14 0.44 0.99 0.76 0.42 0.32
# Построение карты плотности (в формате образа)
DEN <- density(PP, kernel="epanechnikov", dimyx=c(260,160), bw.scott(PP),
  edge=TRUE, diggle=TRUE)
summary(DEN)
real-valued pixel image
260 x 160 pixel array (ny, nx)
enclosing rectangle: [3.469447e-18, 8] x [3.469447e-18, 13] metres
dimensions of each pixel: 0.05 x 0.05 metres
Image is defined on the full rectangular grid
Frame area = 104 square metres
Pixel values
  range = [0.6324174, 15.28264]
  integral = 870.7553
  mean = 8.372647
# Построение карты плотности (отрисовка изображения)
{DenColor <- colorRampPalette(c("white",
  rgb(140,140,140,maxColorValue = 255)))(8)
DenBreaks=seq(0,16,2)
plot(DEN,col=DenColor,main = NULL, breaks=DenBreaks, ribsep =0.03,
  ribwid=0.03, ribargs=list(las=1))
contour(DEN,col="black",lwd=1.5, levels=DenBreaks, labcex=0.8, add=T)
axis(side=1, pos = 0, at = seq(0, 8, 0.5), las = 1)
axis(side=2, pos = 0, at = seq(0, 13, 0.5), las = 1)
plot(PP,char=21,cols = "black",cex=0.7,lwd=1,bg=SpColors, add=T)
}
# Построение карты размещения особей
{plot(PP,main=NULL,char=21,cols =
"black",cex=1,lwd=1,leg.side="right",bg=SpColors)
axis(side=1, pos = 0, at = seq(0, 8, 0.5), las = 1)
axis(side=2, pos = 0, at = seq(0, 13, 0.5), las = 1)
}

```



Для популяций каждого вида подсчитываем количество особей  $N_i$ , среднюю плотность популяции  $\Lambda_i = N_i/S$  и долю вида в растительном сообществе  $P_i = N_i/N$  (здесь и далее  $i = 1, \dots, m$ , где  $m$  – количество видов в растительном сообществе).

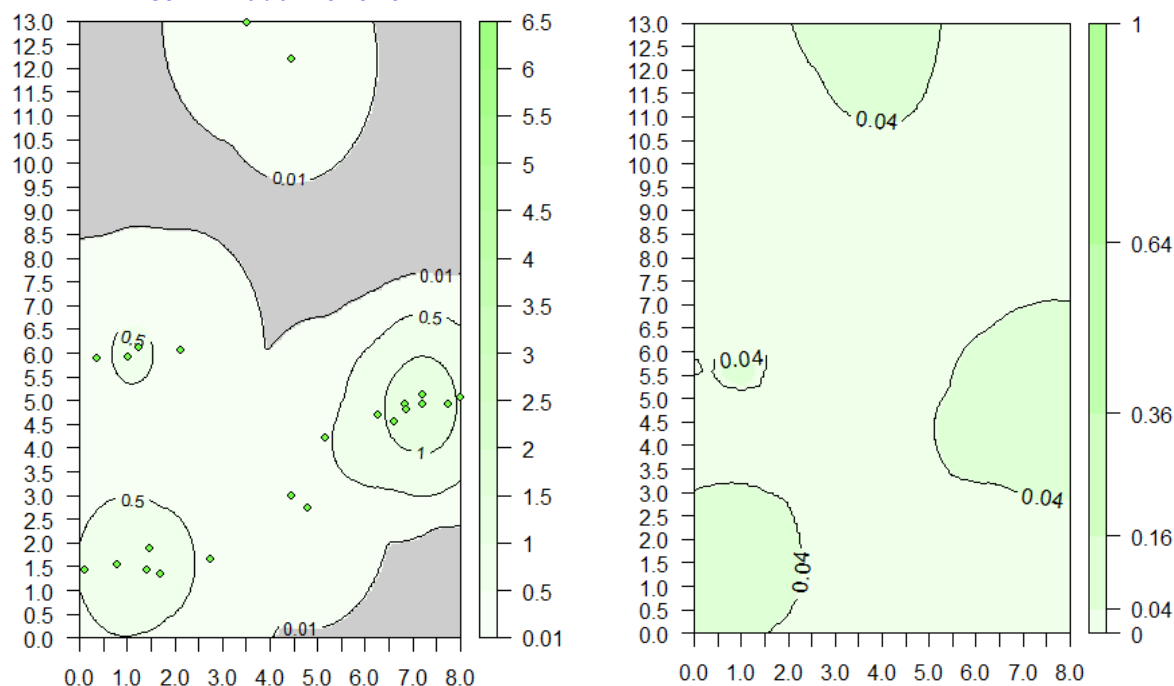
Границы популяционного поля определим путем сравнения локальных значений плотности в каждом элементарном квадрате с установленным нами пороговым значением  $0,01 \text{ экз/м}^2$  – если значение плотности меньше порогового, то данный квадрат лежит за границей популяционного поля. Площадь популяционного поля  $S_i$  оценим как сумму площадей элементарных квадратов ( $0,0025 \text{ м}^2$ ), лежащих в пределах его границы. Экологическую плотность каждой популяции вида  $\tilde{\Lambda}_i$  определим как отношение количества особей данного вида к площади популяционного поля, т.е.  $\tilde{\Lambda}_i = N_i/S_i$ . Построим популяционное поле для вида *Agropyron pectinatum*:

```
# Анализ популяционного поля вида Ag.p
Sp <- "Ag.p"; SpCol <- SpColors[1]
# Построение карты плотности отдельного вида
Sp_Pop <- Phyto[Phyto$Sp==Sp, -3]
Sp_PP <- ppp(x=Sp_Pop$X, y=Sp_Pop$Y, window=owin(c(0,8), c(0,13)),
  unitname=c("metre", "metres"))
Sp_den<-density(Sp_PP, kernel="epanechnikov", dimyx=c(260,160),
  bw.scott(PP), edge=TRUE, diggle=TRUE)
Sp_denSum<-summary(Sp_den); Sp_denSum
real-valued pixel image
260 x 160 pixel array (ny, nx)
enclosing rectangle: [3.469447e-18, 8] x [3.469447e-18, 13] metres
dimensions of each pixel: 0.05 x 0.05 metres
Image is defined on the full rectangular grid
Frame area = 104 square metres
Pixel values
  range = [-3.074464e-16, 1.197601]
  integral = 22.95038
  mean = 0.2206768
# Вычисляем площадь популяционного поля для отдельного вида Ag.p
# (считаем количество эл.квадратов плотность в которых >0.01,
# умножаем на площадь квадрата)
Sp_denV <- as.vector(Sp_den[["v"]])
round(length(Sp_denV[Sp_denV>=0.01])*0.05*0.05, 1) # в кв.м.
[1] 69.1
round(length(Sp_denV[Sp_denV>=0.01])*0.05*0.05/104*100, 1) # в % от площадки
[1] 66.4
# экологическая плотность в экземплярах на м2
round(SpNumber[Sp]/(length(Sp_denV[Sp_denV>=0.01])*0.05*0.05), 2)
Ag.p
0.33
# Прорисовка карта плотности для отдельного вида Ag.p
{Sp_denColor <- colorRampPalette(c("white", SpCol))(20)
Sp_denColor[1] <- "#CDCDCD"; Sp_denColor <- Sp_denColor[-(15:20)]
Sp_denBreaks <- sort(c(0.01, -1, seq(0.5, 6.5, 0.5)))
plot(Sp_den, col=Sp_denColor, main = NULL, breaks=Sp_denBreaks,
  zlim=c(0.01, 6.5), ribsep = 0.03, ribwid=0.03, ribargs=list(at=Sp_denBreaks[-
  1], las=1))
contour(Sp_den, col="black", levels=Sp_denBreaks, labcex=0.8, add=T)
axis(side=1, pos = 0, at = seq(0, 8, 0.5), las = 1)
axis(side=2, pos = 0, at = seq(0, 13, 0.5), las = 1)
plot(Sp_PP, chars=21, cols = "black", cex=0.8, lwd=1.5, bg=SpCol, add=T)
}
# Карта доли отдельного вида + статистика по ней
Sp_Pop <- Phyto[Phyto$Sp==Sp, -3]
Sp_PP <- ppp(x=Sp_Pop$X, y=Sp_Pop$Y, window=owin(c(0,8), c(0,13)),
  unitname=c("metre", "metres"))
```

```

Sp_den<-density(Sp_PP, kernel="epanechnikov", dimyx=c(260,160),
               bw.scott(PP), edge=TRUE, diggle=TRUE)
Sp_prob <- eval.im(round(Sp_den/DEN,3))
Sp_probColor <- colorRampPalette(c("white", SpCol)) (10)
Sp_probColor <- Sp_probColor[2:6]
plot(Sp_prob, col=Sp_probColor, main = NULL,
     zlim=c(0,1), breaks=c(0,0.04,0.16,0.36,0.64,1), ribsep =0.03,
     ribwid=0.03, ribargs=list(las=1, at=c(0,0.04,0.16,0.36,0.64,1)))
contour(Sp_prob, col="black", levels=c(0,0.04,0.16,0.36,0.64,1),
        labcex=1, add=T)
axis(side=1, pos = 0, at = seq(0, 8, 0.5), las = 1)
axis(side=2, pos = 0, at = seq(0, 13, 0.5), las = 1)
summary(Sp_prob)
real-valued pixel image
260 x 160 pixel array (ny, nx)
enclosing rectangle: [3.469447e-18, 8] x [3.469447e-18, 13] metres
dimensions of each pixel: 0.05 x 0.05 metres
Image is defined on the full rectangular grid
Frame area = 104 square metres
Pixel values
  range = [0, 0.111]
  integral = 2.342677
  mean = 0.02252575

```

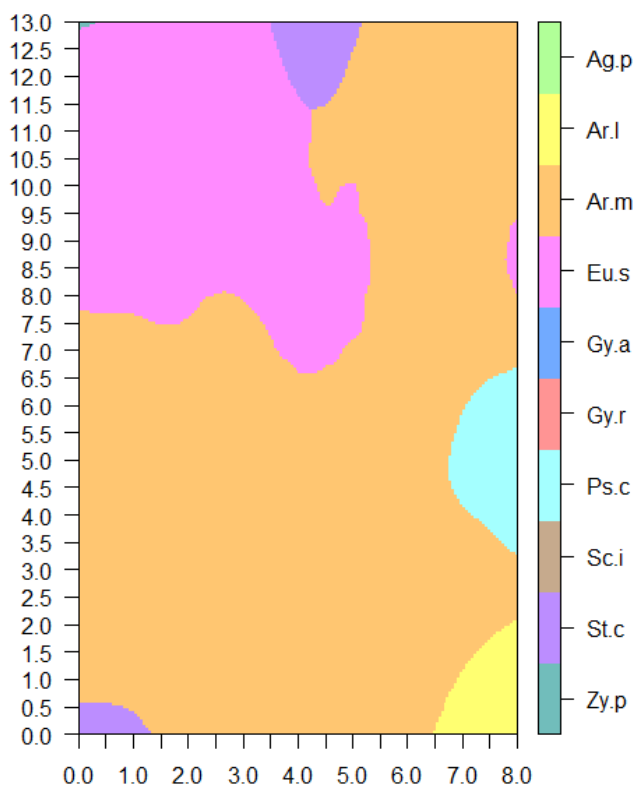


На основе данных о долях видов в растительном сообществе  $P_i$  и их локальных изменениях  $P_i$  можно выделить локально доминирующие виды, оценить площадь  $S_i$ , на которой они преобладают, и построить карту доминирующих видов.

```

# Построение карты доминирующих видов
DomColors <- c("#B1FF98", "#FFFF71", "#FFC671", "#FF8BFF", "#71AAFF", "#FF9494",
              "#A4FFF", "#C6AA8D", "#BC8DF", "#71BDBB")
Prob <- relrisk(PP, kernel="epanechnikov", dimyx=c(260,160), bw.scott(PP),
              edge=TRUE, diggle=TRUE)
Dominant <- im.apply(Prob, which.max)
Dominant <- eval.im(factor(Dominant, levels=10:1,
                          labels=sort(levels(marks(PP)), decreasing = T) ))
plot(Dominant, col=DomColors[10:1], ribsep=0.03, ribwid=0.03,
     ribargs =list(las=1), main = NULL)
axis(side=1, pos = 0, at = seq(0, 8, 0.5), las = 1)
axis(side=2, pos = 0, at = seq(0, 13, 0.5), las = 1)

```



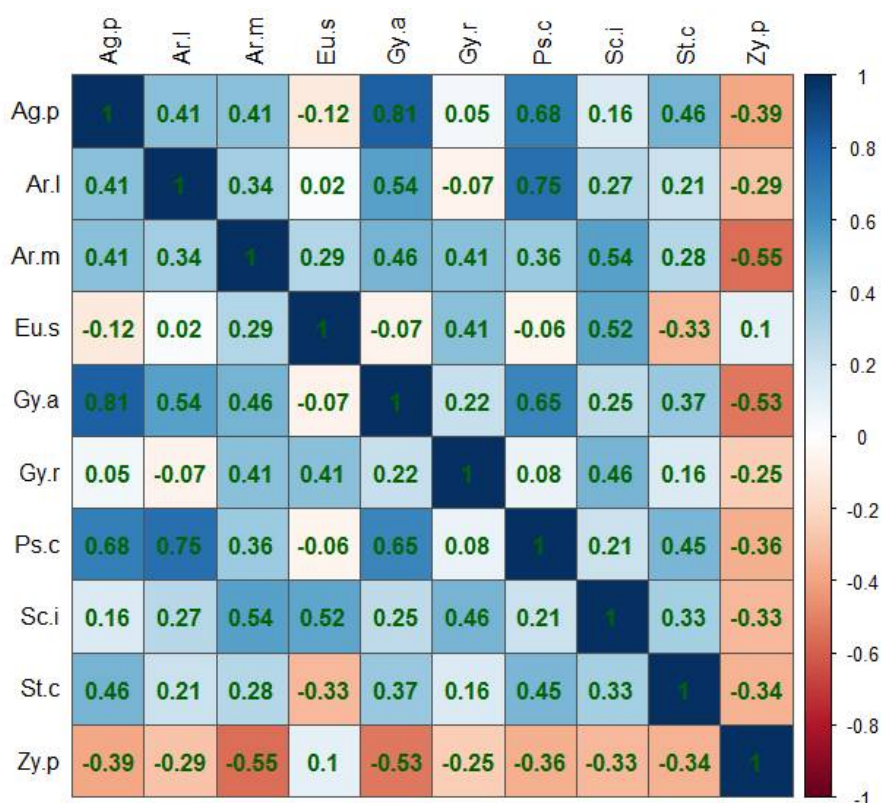
Основными видами-субдоминантами являются *Artemisia marschalliana* (Ar.m) и *Euphorbia seguieriana* (Eu.s).

```
# Площади доминирования
round(table(Dominant[["v"]])*0.05*0.05,2) # в метрах
  Zy.p  St.c  Sc.i  Ps.c  Gy.r  Gy.a  Eu.s  Ar.m  Ar.l  Ag.p
0.03  2.41  0.00  2.99  0.00  0.00  25.38  71.11  2.10  0.00
round(table(Dominant[["v"]])*0.05*0.05/104*100,2) # в % от площадки
  Zy.p  St.c  Sc.i  Ps.c  Gy.r  Gy.a  Eu.s  Ar.m  Ar.l  Ag.p
0.03  2.31  0.00  2.87  0.00  0.00  24.40  68.37  2.01  0.00
```

Для оценки взаимообусловленности встречаемости видов в пределах сообщества рассчитаем матрицу коэффициентов корреляции Пирсона локальных плотностей каждой пары видов.

```
# Построение корреляционной матрицы
S_pix <- 0.05*0.05
DEN_spl <- density(split(PP), kernel="epanechnikov", dimyx=c(260,160),
bw.scott(PP), edge=TRUE, diggle=TRUE)
df <- pairs(DEN_spl, plot=FALSE)
Sp_cor <- round(cor(df), 2)
library(corrplot)
corrplot(Sp_cor, method="color", addCoef.col="darkgreen",
addgrid.col="gray33", tl.col="black")

# Отдельные корреляционные тесты (между парой видов Gy.a- Ar.l)
cor.test(df[, "Gy.a"], df[, "Ar.l"])
  Pearson's product-moment correlation
data: df[, "Gy.a"] and df[, "Ar.l"]
t = 131.52, df = 41598, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.5351264 0.5487010
sample estimates:
  cor
0.5419491
```



#### 4. Проверка гипотезы о характере пространственного размещения точек

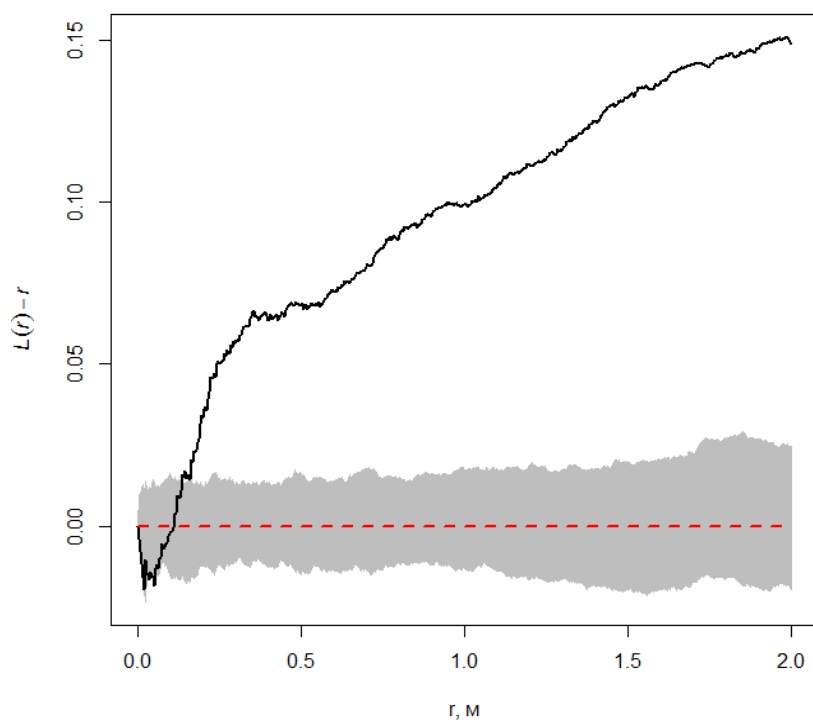
Важным аспектом анализа пространственного распределения точек является диагностика характера процесса, генерирующего события их появления. В качестве нулевой гипотезы обычно принимается предположение о полностью случайном размещении точек в изучаемой области (CSR – *Complete Spatial Randomness*), моделируемом, в частности, однородным точечным процессом Пуассона (Bivand et al., 2008). Если нулевая гипотеза отклонена, то характер размещения растений может быть агрегированный ("agg") или регулярный ("reg"). Для проверки согласия анализируемых данных и гипотетической модели выбирается подходящая описательная статистика и ее эмпирическое значение сравнивается с теоретической величиной. Типичным критерием может быть, например, распределение расстояний между ближайшими точками в круге фиксированного радиуса  $r$  с центром в произвольной точке.

Для расчетов будем использовать функцию Рипли  $K(r)$  (Ripley, 1976, 1977), которая для пуассоновского процесса равна  $K_{pois}(r) = \pi r^2$ . Однако для удобства интерпретации результатов представим их в виде трансформированной функции  $L(r) = \sqrt{K(r)/\pi} - r$  (Besag, Diggle, 1977), которая нормирована относительно 0 оси  $Y$ .

Характер размещения определим по отклонению  $L(r)$  от  $L_{CSR}(r)$ , где  $L_{CSR}(r) \equiv 0$  – функция для однородного пуассоновского процесса. Значимость отклонений  $L(r)$  от  $L_{CSR}(r)$  оценим методом огибающих оболочек (*simulation envelope test* – Besag, Diggle, 1977), для чего с использованием процедуры Монте-Карло построим 299 модельных кривых CSR. Нулевая гипотеза отвергается, если эмпирическая кривая не будет полностью лежать между нижней и верхней огибающими кривыми.

```
# Функция Рипли для сообщества (не учитывается видовая принадлежность)
{PPLest <- envelope(PP, fun=Lest, funargs=list(correction = "Ripley"),
  nsim = 299)
plot(PPLest, ~r~r, xlab="r, м", lwd=2, legend=F, main=NULL) }
```

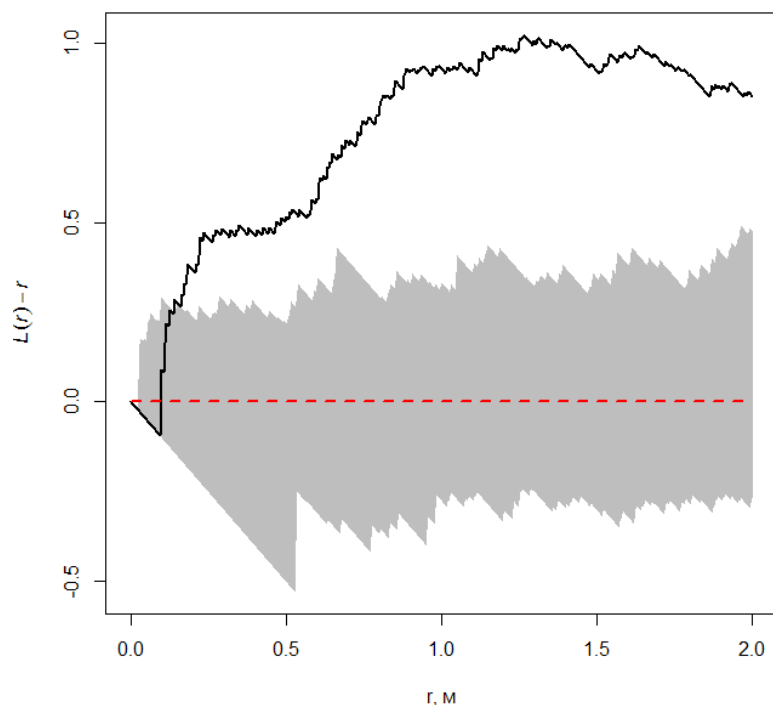




```

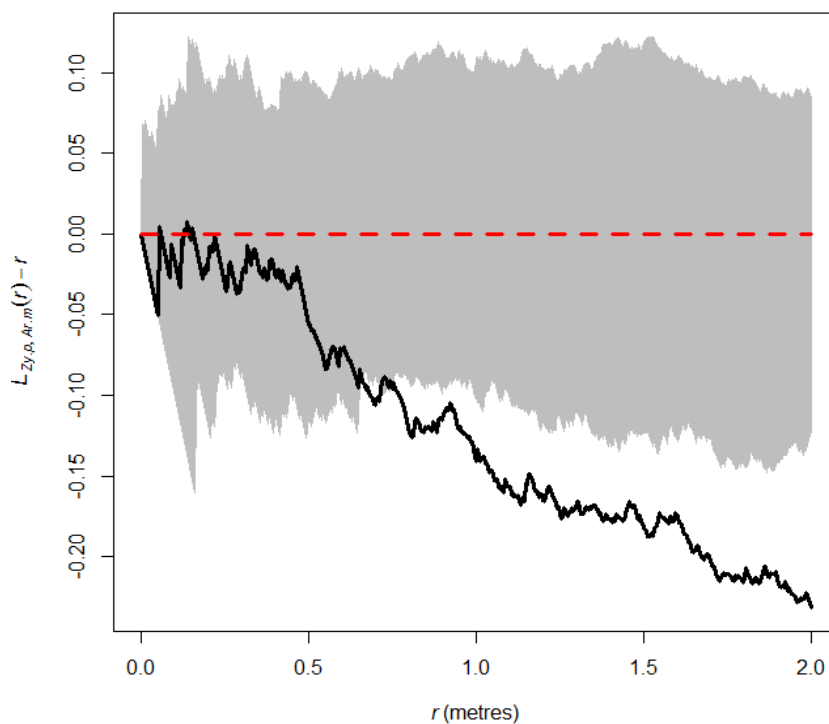
# Функция Рипли для отдельного вида (растения других видов не учитываются)
Sp <- "St.c" # Выбор вида
{Sp_Pop <- Phyto[Phyto$Sp==Sp,-3]
  Sp_PP <- ppp(x=Sp_Pop$X, y=Sp_Pop$Y, window=owin(c(0,8), c(0,13)),
unitname=c("metre", "metres"))
  Sp_Lest <- envelope(Sp_PP, fun = Lest, funargs=list(correction =
"Ripley"), nsim = 299)
  plot(Sp_Lest, .-r~r, xlab="r, м", lwd=2, legend=F, main=NULL)
}

```

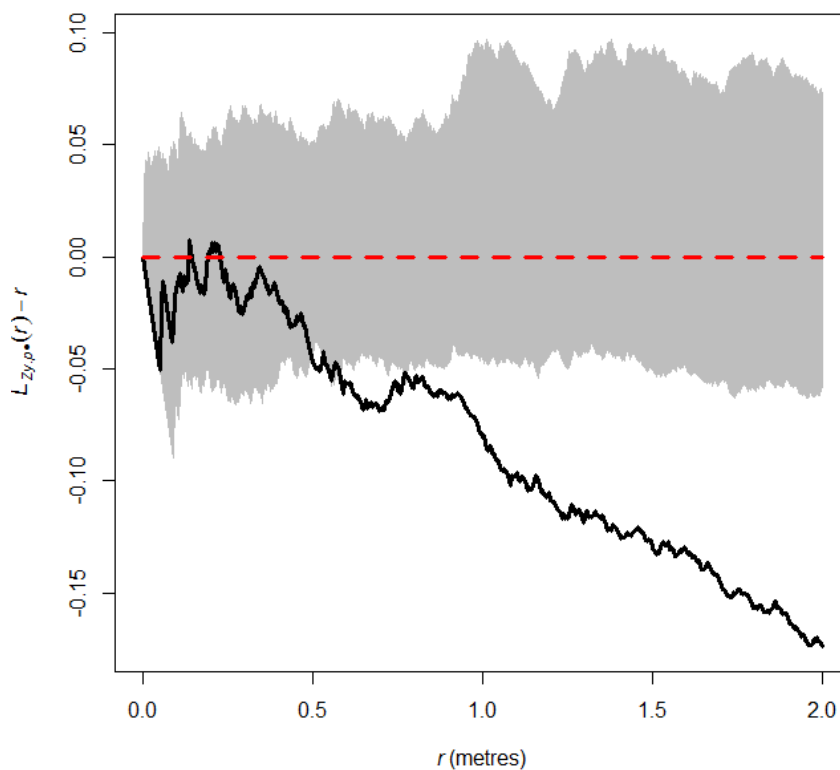


Перекрестная (cross-type) версия  $K$ -функции Рипли равна ожидаемому числу точек вида  $v$  (сверх случайного их числа), расположенных на расстоянии  $r$  от каждой точки вида  $u$ . Отклонение эмпирической кривой  $K_{uv}$  от теоретической кривой  $r^2$  предполагает эффект взаимного влияния экземпляров сравниваемых видов при их пространственном размещении.

```
# кросс-функции Рипли
# !!! (сравниваемые виды указываются в аргументе) !!!
{Pop_LD <- envelope(PP, fun=Lcross, funargs = list("Zy.p", "Ar.m",
  correction = "Ripley"), nsim=299)
plot(Pop_LD, .~r~r, lwd=3, main=NULL, legend=F)}
```



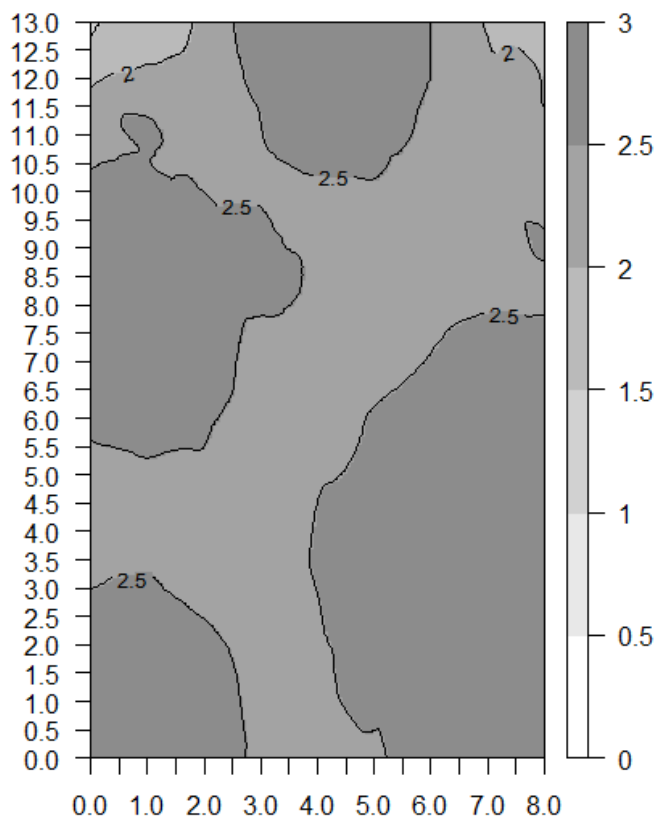
```
{Pop_LD <- envelope(PP, fun=Ldot, funargs = list("Zy.p",
  correction = "Ripley"), nsim=299)
plot(Pop_LD, .~r~r, lwd=3, main=NULL, legend=F)}
```



## 5. Группировка по видовому составу и разнообразию

Представляет интерес рассмотреть, как изменяется по территории пробной площадки видовое разнообразие растений, оцениваемое по индексу Шеннона:

```
# Создаем точечный паттерн (ppp-object)
PP <- ppp(x=Phyto$X, y=Phyto$Y, marks=Phyto$Sp, window=owin(c(0,8),
  c(0,13), unitname=c("metre", "metres")))
Sp_Names <- names(table(Phyto$Sp))
# Создаем набор из карт долей всех видов (используем split-ppp)
DEN <- density(PP, kernel="epanechnikov", dimyx=c(260,160), bw.scott(PP),
  edge=TRUE, diggle=TRUE)
DEN_spl <- density(split(PP), kernel="epanechnikov", dimyx=c(260,160),
  bw.scott(PP), edge=TRUE, diggle=TRUE)
for (Sp in Sp_Names) DEN_spl[[Sp]]<-DEN_spl[[Sp]]/DEN
# Построение карты локальных изменений индекса Шеннона
for (Sp in Sp_Names){ # логарифмируем доли с учетом log2(0)=0
  VSP <- DEN_spl[[Sp]][["v"]]
  VSP[VSP<0] <- 0
  VSP[VSP>0] <- VSP[VSP>0]*log2(VSP[VSP>0])
  DEN_spl[[Sp]][["v"]] <- VSP
}
for (Sp in Sp_Names){ # считаем индекс Шеннона, суммирование
  if (Sp=="Ag.p") H <- -DEN_spl[[Sp]] else H <- H-DEN_spl[[Sp]]
}
# отрисовка карты
{HColor <- colorRampPalette(c("white",
  rgb(140,140,140,maxColorValue = 255)))(6)
HBreaks=seq(0,3,0.5)
plot(H,col=HColor,main = NULL, breaks=HBreaks, zlim=c(0,3), ribsep =0.03,
  ribwid=0.03, ribargs=list(las=1))
contour(H,col="black",lwd=1.5, levels=HBreaks, labcex=0.8, add=T)
axis(side=1, pos = 0, at = seq(0, 8, 0.5), las = 1)
axis(side=2, pos = 0, at = seq(0, 13, 0.5), las = 1)
}
```



```
# статистика по локальным значениям индекса Шеннона
summary(H)
real-valued pixel image
260 x 160 pixel array (ny, nx)
enclosing rectangle: [3.469447e-18, 8] x [3.469447e-18, 13] metres
dimensions of each pixel: 0.05 x 0.05 metres
Image is defined on the full rectangular grid
Frame area = 104 square metres
Pixel values
  range = [1.356471, 2.923984]
  integral = 261.7434
  mean = 2.516764
```

Таким образом, среднее значение индекса Шеннона  $H = 2,80$ , а его локальные значения изменяются от 1,36 до 2,92.

Другой интересной задачей является выделение участков, занятых различными ассоциациями видов. Выясним, сколько вообще в ячейках сетки встречается микрогруппировок, отличающихся присутствием/отсутствием отдельных видов и обозначаемых 10-разрядным двоичным кодом.

```
# Формирование микрогруппировок видов
# Генерация бинарных кодов видовых комбинаций
DEN_spl <- density(split(PP), kernel="epanechnikov", dimyx=c(260,160),
bw.scott(PP), edge=TRUE, diggle=TRUE)
H[["v"]][] <- ""
for (Sp in Sp_Names){
  VSP <- DEN_spl[[Sp]][["v"]]
  H[["v"]][VSP<0.01] <- paste0(H[["v"]][VSP<0.01], "0")
  H[["v"]][VSP>=0.01] <- paste0(H[["v"]][VSP>=0.01], "1")
}
# Число ячеек, занятые микрогруппировками видов
(NgBin <- rev(sort(table(H[["v"]]),)))
111111111 011101111 111111110 111101111 101101111 011101110 101100111
 15679      6041      4072      2789      2563      2188      2044
011111111 001101111 001101101 001100111 111111101 001100100 001101010
 2013      1993      850      369      237      216      80
101101101 001100101 101111110 011101110 111101110 111101111 001101101
 73      71      47      47      46      41      36
001100101 111101101 001101001 001101000 011111110 001100001 101100101
 33      29      20      10      6      6      1
```

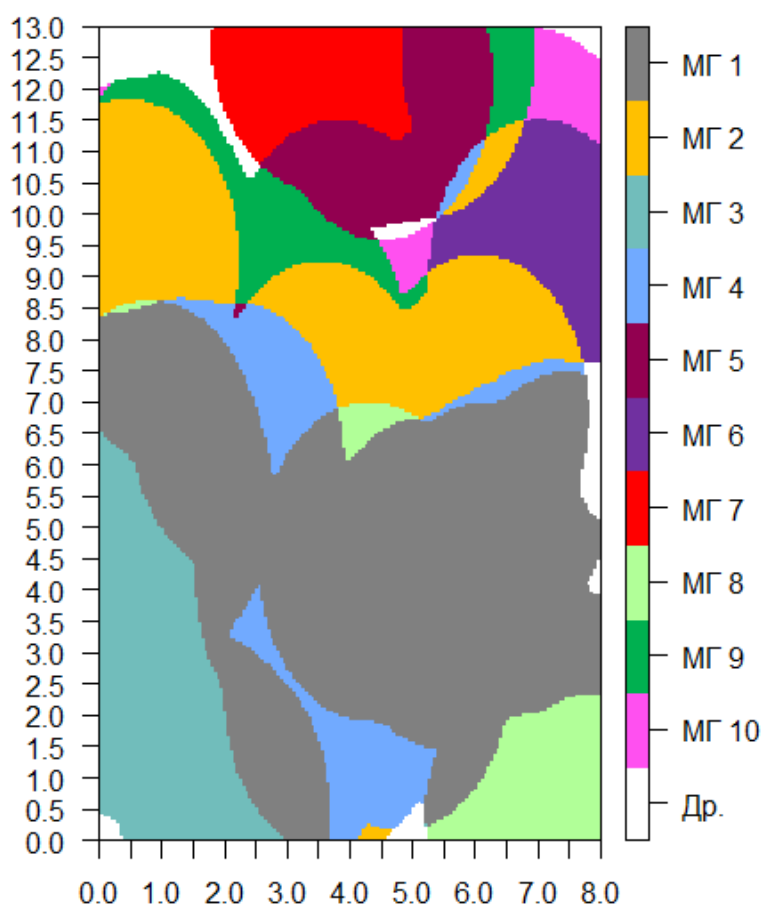
Всего насчитывается 28 микрогруппировок из общего возможного их числа 1024. Наибольшей по встречаемости является микрогруппировка, в которой представлены все 10 видов изучаемого растительного сообщества (111111111). В остальных комбинациях символами 0 обозначены один или несколько выпадающих видов. Некоторые микрогруппировки встречаются редко и могут создавать «шум» при построении мозаичной карты. Поэтому лучше ограничить их список разумным количеством, ориентируясь, например, на долю занимаемой площади. В нашем случае будем рассматривать первые 10 микрогруппировок, каждая из которых занимает площадь более  $1 \text{ м}^2$ , а остальные 18 микрогруппировок с меньшей площадью объединим под меткой "Другие".

```
Fs_OSN <- c("МГ 1", "МГ 2", "МГ 3", "МГ 4", "МГ 5", "МГ 6", "МГ 7", "МГ 8",
"МГ 9", "МГ 10" )
VSP <- H[["v"]]
for (i in 1:10) VSP[VSP==names(NgBin[i])] <- Fs_OSN[i]
VSP[!(VSP %in% Fs_OSN)] <- "Др."
H[["v"]] <- VSP
Fs_OSN <- c(Fs_OSN, "Др.")
H <- eval.im(factor(H, levels=rev(Fs_OSN)))
```

```

# Площади, занятые микрогруппировками видов
round(rev(sort(table(VSP)))*0.05*0.05,2)
VSP
  МГ 1  МГ 2  МГ 3  МГ 4  МГ 5  МГ 6  МГ 7  МГ 8  МГ 9  Др. МГ 10
39.20 15.10 10.18  6.97  6.41  5.47  5.11  5.03  4.98  3.42  2.12
# Доли %, занятые микрогруппировками видов
round(rev(sort(table(VSP)))*0.05*0.05/104*100,1)
VSP
  МГ 1  МГ 2  МГ 3  МГ 4  МГ 5  МГ 6  МГ 7  МГ 8  МГ 9  Др. МГ 10
 37.7  14.5   9.8   6.7   6.2   5.3   4.9   4.8   4.8   3.3   2.0
# Построение мозаичной карты
Fs_Colors <- c("#FFFFFF",
               "#FF50F0", "#00B050", "#B1FF98", "#FF0000", "#7030A0",
               "#920050", "#71AAFF", "#71BDBB", "#FFC000", "#808080")
plot(H, main=NULL, col=Fs_Colors, ribsep=0.03, ribwid=0.03,
     ribargs=list(las=1))
axis(side=1, pos = 0, at = seq(0, 8, 0.5), las = 1)
axis(side=2, pos = 0, at = seq(0, 13, 0.5), las = 1)

```



На карте видно, что большинство элементарных участков средней и нижней части содержит полный состав видов (МГ 1). *Zygophyllum pinnatum* выпадает из растительного сообщества в нижнем левом углу площадки (МГ 3), а в центре и справа нижней части также выпадают из сообщества *Gypsophila altissima* (МГ 4), *Agropyron rectinatum* (МГ 8) или оба этих вида (МГ 2). В верхней части площадки наблюдается более сложная мозаичная картина.

## 6. Анализ пространственной изменчивости видового состава с использованием индексов zeta ( $\zeta$ ) -разнообразия

Видовое разнообразие отдельного однородного участка ( $\alpha$ ) или общее разнообразие ландшафта ( $\gamma$ ) являются статичными показателями, из которых наиболее интерпретируемыми оценками является число видов  $S_\alpha$  и  $S_\gamma$ . Модели разнообразия, оценивающие изменчивость видового состава между элементами каждого масштабного уровня (т.е. *мозаичное разнообразие* между выборками из одного местообитания или сравнительное  $\beta$ -разнообразие между отдельными сообществами), имеют более сложный характер и позволяют детально проанализировать механизмы совместного распределения видов в пространстве и времени. Композиционная, т.е. основанная на частотах, степень видовой дифференциации между участками традиционно измеряется с использованием показателей, основанных на парных сравнениях ( $n = 2$ ) встречаемости видов (например, Diserud и Odegaard 2007, Jost et al., 2010, Baselga 2013), эффективность которых горячо оспаривается (Magurran, Henderson, 2010<sup>1</sup>). Эти возражения сводятся к трем аргументам: а) при достаточно большом числе участков эти меры становятся нечувствительными, устремляясь к 0 или 1; б) не всегда корректно оценивается сравнительная роль общих и редких видов; в) среднее из парных мер сходства не всегда является экологически достаточной и статистически интерпретируемой.

Дзета-разнообразие (*zeta- или  $\zeta$ -diversity*, см. Hui, McGeoch, 2014) было разработано как концепция для количественной оценки числа видов, общих для нескольких участков, которая предоставляет информацию о полном спектре редких, промежуточных и распространенных видов, поскольку все они вносят определенный вклад при анализе композиционной дифференциации (*turnover*). Дзета-разнообразие  $\zeta_n$  определяется как количество видов, общих для  $n$  участков, причем  $n$  называется порядком  $\zeta$ . Например,  $\zeta_1$  - среднее число видов на одном участке ( $\alpha$ -компонент),  $\zeta_2$  – среднее число видов, общих для любых двух участков ( $\zeta_1 - \zeta_2$  отражает  $\beta$ -компонент),  $\zeta_3$  – для любых трех участков и т.д.

Поскольку мы имеем только одну пробную площадку, то для подсчета индексов  $\zeta$ -разнообразия разобьем ее на 15 равных участков ( $3 \times 5$ ) площадью около  $7 \text{ м}^2$ .

```
# Считываем данные по координатам растений и видовой принадлежности
load("Phyto.RData")
# Прописываем в исходный фрейм номера участков по вертикали и горизонтали
library(classInt)
Phyto$ClX <- findCols(classIntervals(Phyto$X, n = 3, style = "equal"))
Phyto$ClY <- findCols(classIntervals(Phyto$Y, n = 5, style = "equal"))
Phyto$ClXY <- paste(Phyto$ClX, Phyto$ClY, sep="_")
library(reshape2)
Phyto.spec <- dcast(Phyto, ClXY ~ Sp, fun.aggregate = length)
  ClXY Ag.p Ar.l Ar.m Eu.s Gy.a Gy.r Ps.c Sc.i St.c Zy.p
1  1_1    5    3   31    6    3    6   15    9   16    0
2  1_2    0    2   27   14    0    3    4    3    0    0
3  1_3    4    3   36   17    4    8    7    8    2    1
4  1_4    0    1    8   12    0    4    4   12    2    4
5  1_5    0    0    2   10    0    0    4    1    1    3
6  2_1    1    2   32   15    0    2    8    2    2    1
7  2_2    3    8   29   11    1    2    5   11    4    3
8  2_3    0    1   20   23    0    5    5    5    0    0
9  2_4    0    0   11   12    0    3    4    2    1    7
10 2_5    2    0    3    7    0    0    4    2    6    5
11 3_1    0   20   25   11    1    3   14   11    4    2
12 3_2    8    5   23    9    6    1   17    6    5    0
13 3_3    0   14   19   13    0    3    8    2    1    4
14 3_4    0    1   13   12    0    5    3    3    0    2
15 3_5    0    0   10    7    0    1    1    2    0    1
```

<sup>1</sup> Подробности и литературные ссылки см. (Шитиков и др., 2012)

```
Phyto.specB <- as.data.frame(ifelse(Phyto.spec[, -1]==0, 0, 1))
# За координаты XY примем средние координаты всех растений участка
library(tidyverse)
Phyto.XY <- group_by(Phyto, ClXY) %>%
  summarise(x = mean(X), y = mean(Y)) %>% as.data.frame()
```

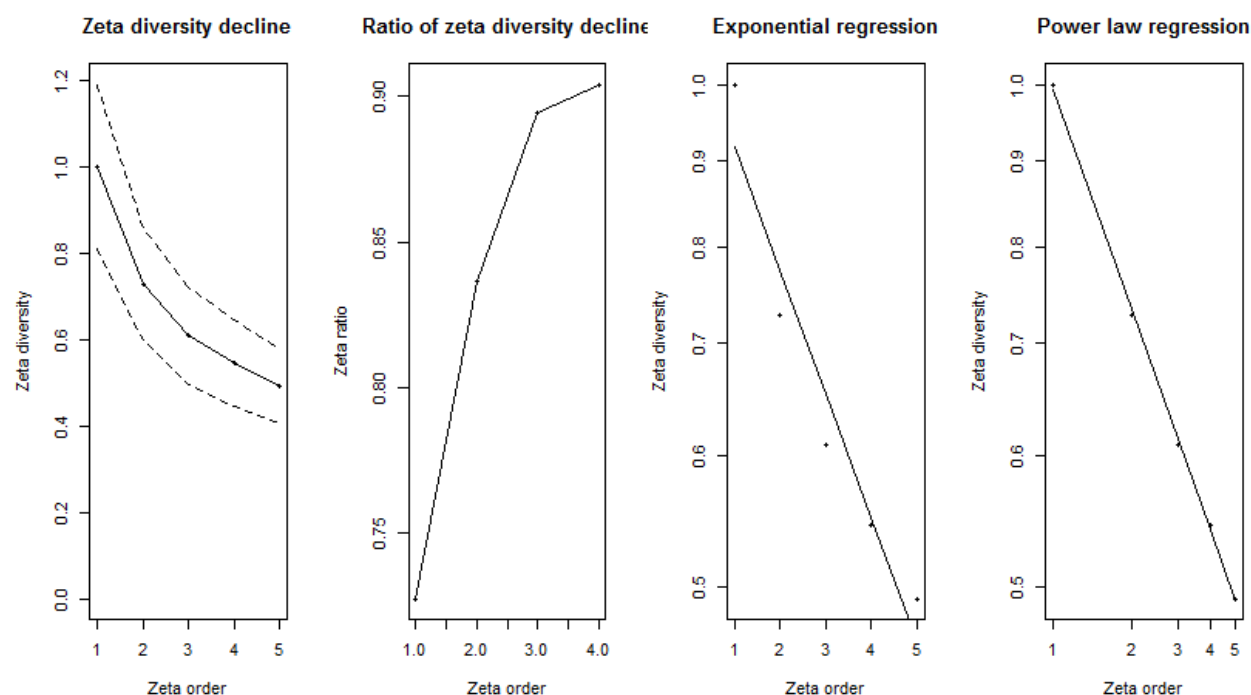
Рассчитаем последовательность значений  $\zeta_n$  порядка  $n$  от 1 до 5. Вычисления можно легко осуществить с помощью пакета R `zetadiv` (Latombe et al., 2017).

```
library(zetadiv)
Value.list <- list()
for (n in 1:5) Value.list[[n]] <- Zeta.order.ex(Phyto.specB, order = n)
matrix(unlist(Value.list), nrow=5, byrow = TRUE,
       dimnames = list(1:5, c("order", "combinations", "zeta.val", "sd")))
  order combinations zeta.val sd
1      1             15 7.800000 1.4735768
2      2            105 6.552381 1.2708337
3      3            455 5.782418 1.0900626
4      4           1365 5.263736 0.9592332
5      5           3003 4.892441 0.8451489
```

Здесь  $\text{combinations} = C_m^n = m!/[n!(m-n)!]$ , т.е. число сочетаний из  $m = 15$  участков по  $n$ ,  $\text{zeta.val}$  - среднее число видов, встретившихся совместно в этих комбинациях.

По определению,  $\zeta$ -разнообразие уменьшается с ростом порядка  $n$ . Поскольку виды, общие для  $n + 1$  участков, являются подмножеством из общих видов для  $n$  участков, то отношение  $\zeta$ -разнообразия соседних порядков ( $\zeta_{n+1}/\zeta_n$ ) представляет собой вероятность того, что вид, обнаруженный на  $n$  участках, встретится на еще одном дополнительном участке. Т.е., если обобщить, то  $\zeta_{n+1}/\zeta_n$  - доля видов, общих для  $n$  участков, которые также встретятся совместно на  $n + 1$  участках, что является мерой активности растительного сообщества к потенциальному расселению за пределы обследованной области.

```
# Анализ скорости уменьшения  $\zeta$ -разнообразия
Zeta.decline.ex(Phyto.specB, orders = 1:5)
Zeta.decline.mc(Phyto.specB, orders = 1:5, sam = 500, normalize = "Jaccard")
```

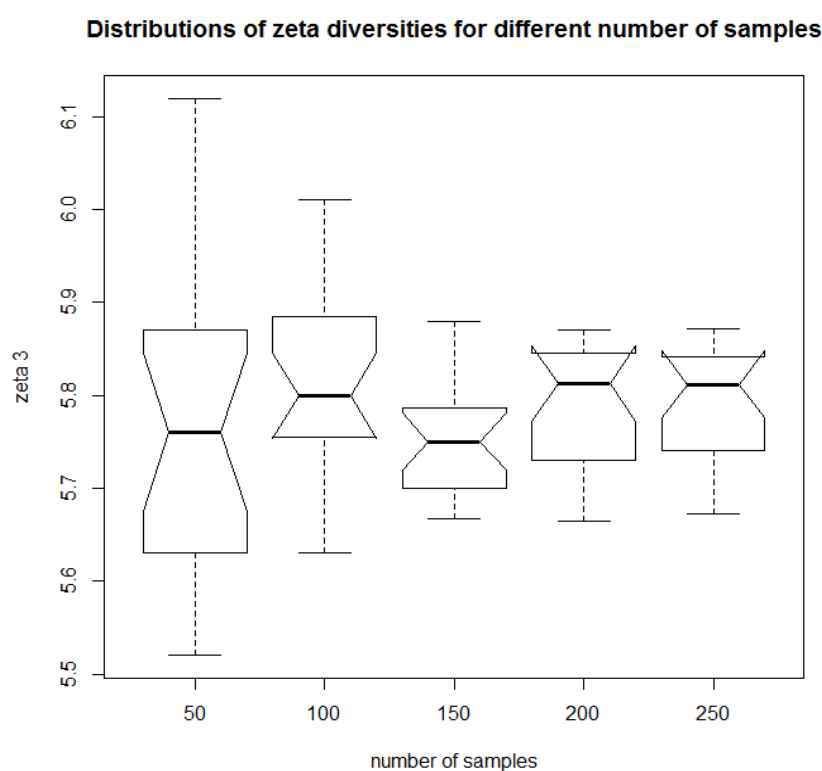


Использованные выше функции вычисляют последовательность относительных значений  $\zeta$ -разнообразия для заданного диапазона порядков  $n$ , темпы уменьшения которых

(*decline*) аппроксимируются экспоненциальной и степенной моделями. Выводятся графики различных зависимостей, включающих также отношение  $\zeta_n/\zeta_{n-1}$  (*ratio*). Функция `Zeta.decline.ex` основана на полной матрице наблюдений и использует формулу для числа сочетаний, а `Zeta.decline.mc` реализует процесс Монте-Карло перебора различных случайных комбинаций участков. В пакете `zetadiv` используются две нулевые модели для реализации процессов имитации: элементы матрицы встречаемости видов по участкам (1: присутствие; 0: отсутствие) переставляются при ограничениях на сохранение сумм строк или сумм столбцов такими же, что и в матрицах наблюдений.

Как большинство других показателей биоразнообразия,  $\zeta$ -разнообразие чувствительно к характеру выборки, т.е. числу рассматриваемых участков, их масштабу и др. Функция `Zeta.sam.sensitivity` выполняет имитации и вычисляет оценку чувствительности  $\zeta$ -разнообразия заданного порядка  $n$  в зависимости от размеров выборки (количества скоплений или участков).

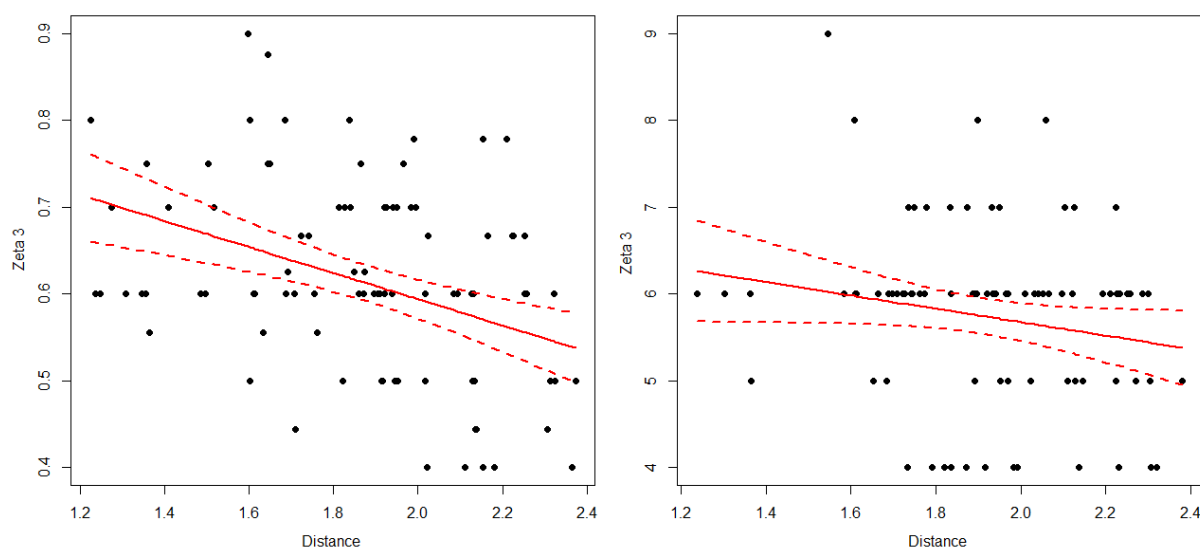
```
Zeta.sam.sensitivity(Phyto.specB, order = 3,
  sam.seq = seq(50, 250, 50), reps = 20, plot = TRUE, notch = TRUE)
```



Увеличение расстояния между участками обычно приводит к снижению сходства видового состава, что в экологии трактуется как "затухание подобия на расстоянии" (*distance decay of similarity*), а в биогеографии – как пространственная автокорреляция. Чтобы оценить, как  $\zeta$ -разнообразие уменьшается с увеличением среднего расстояния между случайными выборками, в пакете используются модели парной аппроксимации по мотивам уравнений статистической физики. Функция `Zeta.ddecay` вычисляет затухание  $\zeta$ -разнообразия заданного порядка в зависимости от расстояния, используя обобщенную линейную модель, либо обобщенную аддитивную модель с ограничениями.

```
Zeta.ddecay(Phyto.XY, Phyto.specB, sam = 100, order = 3,
  method.glm = "glm.fit2", confint.level = 0.95,
  trsf = "log", normalize = "Jaccard")
Zeta.ddecay(Phyto.XY, Phyto.specB, sam = 100, order = 3,
  method.glm = "glm.fit2", confint.level = 0.95,
  trsf = "log", normalize = "Euclidean")
```



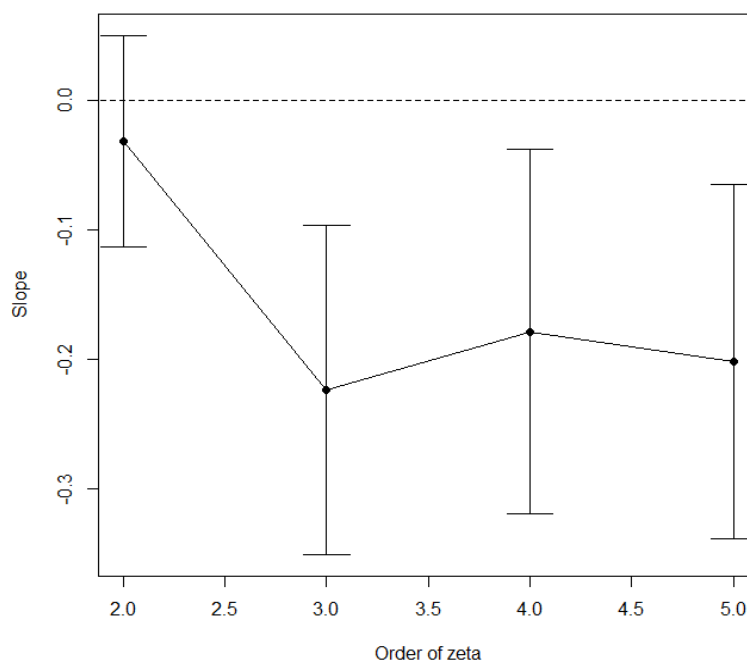


Если использовать параметр `normalize = "Jaccard"`, то в качестве расстояния между случайными участками используется индекс Жаккара и оценивается скорость смены видов непосредственно по экологическому градиенту (график слева); при `normalize = "Euclidean"` вычисляется, как сходство видового состава уменьшается с географическим расстоянием по Евклиду между каждой парой участков (справа).

Можно также оценить, как меняется угол наклона этих прямых для  $\zeta$ -разнообразия различных порядков с использованием обобщенной линейной модели.

```
Zeta.ddecays(Phyto.XY, Phyto.specB, sam = 100,
             orders = 2:5, plot = TRUE, confint.level = 0.95)
```

Distance decay of zeta diversity



При увеличении размеров пробных площадок видовое богатство на каждом участке, а также количество общих видов для  $n$  из них, будут увеличиваться. Серьезные исследования предусматривают "выборочный иерархический дизайн", при котором участки самого мелкого масштаба вложены в участки большего размера. В пакете `zetadiv` разработаны функции объединения выборок для формирования более крупных кластеров и анализа изменений  $\zeta$ -разнообразия при различных формах дизайна.

```

# Функция объединения m ближайших участков
Rescale = function (m) {
  data.rescale <- rescale.min.dist(Phyto.XY, Phyto.specB, m = m)
  Value.list <- list()
  for (n in 1:5)
    Value.list[[n]] <- Zeta.order.ex(data.rescale[,-(1:3)], order = n)
  matrix (unlist(Value.list), nrow=5, byrow = TRUE,
    dimnames = list(1:5, c("order", "combinations", "zeta.val", "sd")))
}
Rescale (2)
  order combinations zeta.val      sd
1      1             7 9.000000 1.000000
2      2            21 8.285714 0.9023778
3      3            35 7.800000 0.7194769
4      4            35 7.485714 0.5621089
5      5            21 7.285714 0.4629100
Rescale (3)
  order combinations zeta.val      sd
1      1             5  9.2 1.0954451
2      2            10  8.5 1.0801234
3      3            10  7.9 0.8755950
4      4             5  7.4 0.5477226
5      5             1  7.0      NaN

```

Функция `rescale.min.dist` осуществляет агрегацию  $m$  ближайших участков. Очевидно, что  $\zeta$ -разнообразие (до некоторых пределов) увеличивается при агрегировании участков.

В представленных публикациях (Hui, McGeoch, 2014; McGeoch et al., 2019) подробно представлена концепция  $\zeta$ -разнообразия и показано, как ее интерпретация позволяет внести "свежую струю" в различные традиционные методы анализа, основанные на матрице встречаемости видов, включая кривые накопления, моделирование распределения частот заселения и оценку доли эндемичных комбинаций. Если параллельно с биоценотическими наблюдениями проводился мониторинг различных переменных окружающей среды, то показано (Latombe et al., 2017), как оценка их вклада в  $\zeta$ -разнообразии позволяет более обоснованно реализовать концепцию моделирования обобщенных различий (GDM – *Generalised Dissimilarity Modelling*) применительно к нескольким участкам.

Дополнительную информацию можно получить на [сайте](#) австралийской исследовательницы McGeoch со звучным именем Melodie.

### Список рекомендуемой литературы

*Мазинг В. В.* Об изучении мозаичности и комплексности растительного покрова // Изв. АН ЭстССР, ср. биол. 1965. Т. 14. №1. С. 345-368.

*Савельев А.А., Мухарамова С.С., Пулюгин А.Г. и др.* Геоestatистический анализ данных в экологии и природопользовании (с применением пакета R). Казань: Казанский университет, 2012. 120 с. URL: <http://gis-lab.info/docs/saveliev2012-geostat.pdf>

*Шутиков В.К., Розенберг Г.С.* Рандомизация и бутстреп: статистический анализ в биологии и экологии с использованием R. Тольятти: Кассандра, 2014. 314 с. URL: <https://stok1946.blogspot.com/p/blog-page.html>

*Bailey T.C., Gatrell A.* Interactive spatial data analysis. – Harlow, England: Longman Scientific & Technical, 1995. 413 p.

*Silverman B.W.* Density estimation for statistics and data analysis. London: Chapman and Hall, 1986. 175 p.

*Scott D.W.* Multivariate density estimation. Theory, Practice and Visualization. New York: John Wiley & Sons Ltd, 1992. 384 p.

*Епанечников В.А.* Непараметрическая оценка многомерной плотности вероятности // Теория вероятности и ее применение, 1969. – Т. 14 (1) – С. 156-151

*Bivand R., Pebesma E.J., Gomez-Rubio V.* Applied spatial data analysis with R. N.Y.:Springer, 2008. 374 p.

*Ripley B.D.* Modelling spatial patterns // Journal of the Royal Statistical Society. Series B. 1977. № 39. P. 172-212.

*Ripley B.D.* The second-order analysis of stationary point processes // Journal of Applied Probability. 1976. № 13. P. 255-266.

*Besag J.E., Diggle P.J.* Simple Monte Carlo tests for spatial pattern // Applied Statistics, 1977. V. 26. P. 327-333.

*Шутиков В.К., Зинченко Т.Д., Розенберг Г.С.* Макроэкология речных сообществ: концепции, методы, модели. Тольятти:Кассандра, 2012. 257 с. URL: <http://www.ievbras.ru/ecostat/Kiril/Download/Maec.pdf>

*Hui C., McGeoch M. A.* Zeta diversity as a concept and metric that unifies incidence-based biodiversity patterns. // The American Naturalist. 2014. V. 184(5). P. 684-694. doi:10.1086/678125.

*Latombe G., Hui C., McGeoch M. A.* Multi - site generalised dissimilarity modelling: using zeta diversity to differentiate drivers of turnover in rare and widespread species // Methods in Ecology and Evolution. 2017. V. 8(4). P. 431-442. doi:10.1111/2041-210X.12756.

*McGeoch M. A., Latombe G., Andrew N.R., Nakagawa S., Nipperess D.A., Roigé M., Marzinelli E.M., Campbell A.H., Vergés A., Thomas, T.* Measuring continuous compositional change using decline and decay in zeta diversity // Ecology. 2019. V. 100:e02832. doi:10.1002/ecy.2832.