

Tutorial R Módulo 10. Análise de Redundância - RDA

Disciplina de Ecologia Numérica*

Prof. Elvio S. F. Medeiros Laboratório de Ecologia
Universidade Estadual da Paraíba Campus V, João Pessoa, PB

2023-06-19

A Análise de Redundância (RDA - Redundancy Analysis) é uma técnica estatística multivariada que visa identificar a contribuição relativa de diferentes conjuntos de variáveis para a variação observada em uma variável resposta. Ela é amplamente utilizada em estudos de ecologia, biologia e outras áreas relacionadas, onde existe interesse em entender quais variáveis explicam melhor os padrões de variação em um conjunto de dados.

Índice

1	Introdução	2
1.1	RDA x PCA	3
2	Organização básica	4
2.1	Pacotes do módulo	5
2.2	Sobre os dados do PPBio	5
2.2.1	A planilha PPBio Habitat	6
2.3	Importando a planilha de trabalho	6
2.3.1	Outra forma de achar e importar uma planilha	6
2.4	Particionando as variáveis de interesse	7
2.5	Reset point	7
3	Classificação 1: Matriz comunitária	7
3.1	Dendrograma e Heatmap 1	7
3.2	Histórico das fusões 1	14

*Semestre 2023.1

4	Classificação 2: Matriz ambiental	16
4.1	Dendrograma e Heatmap 2	16
4.2	Histórico das fusões	22
5	Classificação 3: Heatmap Matriz comunitária vs. Matriz ambiental	24
5.1	Dendrograma e Heatmap 3	24
6	Explorando correlações multivariadas da matriz de habitat	29
7	Análise de Redundância	32
7.1	Transformando os dados de comunidade com o método de <code>hellinger</code>	32
7.2	Fazendo a RDA	32
7.3	Gráficos de RDA	36
8	Interpretando o biplot gerado no código anterior	40
8.1	Testando a significância dos resultados da RDA	41
9	RDA passos importantes	42
9.1	Avaliando colinearidade	42
9.2	Escolhendo as principais variáveis do modelo da RDA	43
9.2.1	Variáveis sugeridas pelo modelo	49
10	Referências	58
	Apêndices	58
	Sites consultados	58

Lista de Figuras

Lista de Tabelas

1	Nomenclatura das matrizes em AMD em relação aos atributos das colunas. . . .	8
2	Resumo sobre terminologia usada conforme definido para fins de consistência.	17

1 Introdução

A Análise de Redundância (RDA - Redundancy Analysis) é uma extensão da Análise de Componentes Principais (PCA - Principal Component Analysis), que combina técnicas de análise de regressão multivariada com métodos de análise de componentes principais.

O objetivo da RDA é encontrar uma combinação linear de variáveis explicativas (geralmente uma matriz ambiental) que maximize a correlação com a variável resposta (geralmente uma matriz comunitária). Essa combinação linear é conhecida como eixos canônicos. Além disso, a RDA também pode identificar outros eixos canônicos, cada um representando uma nova combinação linear de variáveis explicativas que é ortogonal aos eixos anteriores e também tem alta correlação com a variável resposta.

Ao realizar uma análise de redundância, é possível avaliar a importância relativa de diferentes conjuntos de variáveis explicativas e identificar quais variáveis contribuem de forma significativa para a variação na variável resposta. Essa técnica é especialmente útil quando há muitas variáveis explicativas disponíveis e é necessário identificar quais delas são as mais relevantes para o fenômeno em estudo.

A RDA oferece uma abordagem poderosa para entender os padrões de variação em dados multivariados, ajudando os pesquisadores a explorar a redundância nas informações e identificar os principais fatores que influenciam uma variável resposta específica.

1.1 RDA x PCA

A Análise de Redundância e a Análise de Componentes Principais são técnicas estatísticas multivariadas que têm abordagens diferentes e são aplicadas em contextos distintos. Vamos explorar as principais diferenças entre elas.

1. Objetivos:

- RDA: A RDA é usada para analisar a contribuição relativa de diferentes conjuntos de variáveis explicativas para a variação observada em um conjunto de variáveis resposta. Seu objetivo é encontrar combinações lineares de variáveis explicativas que tenham alta correlação com a variável resposta.
- PCA: A PCA é utilizada para explorar a estrutura de variação em um conjunto de dados multivariado sem a presença de uma variável resposta específica. Seu objetivo é reduzir a dimensionalidade dos dados, identificando as principais direções (componentes principais) ao longo das quais os dados variam.

2. Uso de variável resposta:

- RDA: A RDA requer um conjunto de variáveis resposta que sejam consideradas as variáveis dependentes do estudo (geralmente uma matriz ambiental). Ela é usada para avaliar como os conjuntos de variáveis explicativas (geralmente uma matriz comunitária) contribuem para a variação nessa variável resposta específica.
- PCA: A PCA não requer uma variável resposta. Ele analisa apenas as relações entre as variáveis explicativas e não está focada em explicar variações em uma variável resposta específica.

3. Tipo de variáveis:

- RDA: A RDA é adequada para dados em que há uma combinação de variáveis explicativas numéricas (contínuas) e categóricas (qualitativas).
- PCA: A PCA é mais adequada para dados numéricos (contínuos). Ele lida bem com variáveis contínuas, mas não é apropriado para variáveis categóricas.

4. Análise de correlação:

- RDA: A RDA busca encontrar a correlação entre os conjuntos de variáveis explicativas e a variável resposta. Ela avalia a importância relativa de cada conjunto de variáveis explicativas.
- PCA: A PCA não busca uma correlação específica com uma variável resposta. Ele busca identificar as direções ao longo das quais os dados apresentam maior variação.

5. Estrutura de dimensionamento:

- RDA: A RDA permite que seja realizada uma **análise canônica**¹ que encontre combinações lineares de variáveis explicativas para maximizar a correlação com a variável resposta. Isso significa que a RDA pode identificar múltiplos eixos canônicos, cada um representando uma combinação linear diferente.
- PCA: A PCA busca os componentes principais que explicam a maior parte da variação total nos dados, mas não está relacionado a uma variável resposta específica. Ele gera um conjunto de componentes principais ortogonais entre si.

i Em resumo, a RDA é adequada para analisar a relação entre variáveis explicativas e uma variável resposta específica, enquanto a PCA é mais útil para entender a estrutura de variação em um conjunto de dados multivariados sem considerar uma variável resposta. Ambas as técnicas têm suas aplicações distintas e podem fornecer insights valiosos em diferentes contextos de análise de dados.

2 Organização básica

```
1 rm(list=ls(all=TRUE)) ##LIMPA A MEMORIA
2 cat("\014") #limpa o console
3 dev.off() #apaga os graficos, se houver algum
```

¹São estimadas combinações lineares de variáveis explicativas que estão maximamente correlacionadas com a variável resposta. Essas combinações lineares são chamadas de eixos canônicos. Os eixos canônicos são calculados de forma a serem ortogonais uns aos outros, ou seja, eles não estão correlacionados entre si.

2.1 Pacotes do módulo

Instalando os pacotes necessários para esse módulo. Nos computadores do Laboratório de Ecologia não instale esses pacotes, eles já estão instalados.

```
1 install.packages("openxlsx")
2 install.packages("tidyverse")
3 install.packages("vegan")
4 install.packages("gplots")
5 install.packages("psych")
6 install.packages("gt")
7 install.packages("ggplot2")
```

```
1 library(tidyverse)
```

Os códigos acima, são usados para instalar os pacotes necessários para este módulo. O comando `library()` será usado para carregarmos esses pacote a medida que eles forem sendo necessários.

Para definir o diretório de trabalho usa-se os códigos abaixo. **Lembre de usar a barra “/” entre os diretórios. E não a contra-barra “\”.**

```
1 getwd()
2 setwd("C:/Seu/Diretório/De/Trabalho")
```

Alternativamente você pode ir na barra de tarefas e escolhes as opções:

SESSION -> SET WORKING DIRECTORY -> CHOOSE DIRECTORY

Usar o [RStudio Cloud](#) é uma opção para quem não quer instalar a versão para PC ².

2.2 Sobre os dados do PPBio

A planilha `ppbio` contém os dados de abundância de espécies em diferentes unidades amostrais (UA's). A base teórica dos dados do PPBio para o presente estudo pode ser vista em [Base Teórica](#). Leia antes de prosseguir.

²O RStudio Cloud é uma plataforma online que fornece um ambiente de desenvolvimento integrado para o R, permitindo que os usuários executem análises, desenvolvam código e colaborem com outras pessoas, sem a necessidade de instalar o R e o RStudio em seus próprios computadores. É uma solução conveniente e acessível, especialmente para iniciantes ou usuários que desejam compartilhar projetos e colaborar de forma eficiente

2.2.1 A planilha PPBio Habitat

Para esse módulo também usaremos a planilha `ppbioh`. Esta é uma **matriz de dados ambiental**, guardados na nO arquivo `ppbio06h.xlsx`, que traz os dados brutos de 26 localidades (UAs) em períodos diferentes (objetos) x 35 variáveis ambientais (atributos) medidas em diferentes escalas espaciais, antes de qualquer modificação. As unidades de medição incluem cm, m, °C, mg/L, %, entre outros (dados publicados por (Medeiros et al. 2008)). Esses dados tem uma alta amplitude de variação, sugerido uso de matriz transformada e/ou reescalada. As bases teóricas dos dados do PPBio para o presente estudo pode ser vista em [Base Teórica](#). Leia antes de prosseguir.

2.3 Importando a planilha de trabalho

Note que o símbolo `#` em programação R significa que o texto que vem depois dele é um comentário e não será executado pelo programa. Isso é útil para explicar o código ou deixar anotações. Ajuste a segunda linha do código abaixo para refletir “C:/Seu/Diretório/De/Trabalho/Planilha.xlsx”.

**** Em uma RDA trabalhamos com duas matrizes, uma comunitária e outra ambiental****

```
1 library(openxlsx)
2 ppbio <- read.xlsx("D:/Elvio/OneDrive/Disciplinas/_EcoNumerica/5.Matrizes/ppbio06p.xlsx",
3                   rowNames = T,
4                   colNames = T,
5                   sheet = "Sheet1")
6 ppbio_a <- read.xlsx("D:/Elvio/OneDrive/Disciplinas/_EcoNumerica/5.Matrizes/ppbio06h.xlsx",
7                    rowNames = T,
8                    colNames = T,
9                    sheet = "Sheet1")
10 #ppbio
11 #str(ppbio)
12 #ppbio_ma <- as.matrix(ppbio) #lê ppbio como uma matriz
13 #ppbio_ma
14 #str(ppbio_ma)
```

2.3.1 Outra forma de achar e importar uma planilha

```
1 getwd()
2 ppbio <- read.xlsx(file.choose(),
3                   rowNames = T, colNames = T,
4                   sheet = "Sheet1")
```

2.4 Particionando as variáveis de interesse

Para escolher quais variáveis entrar na análise:

```
1 #Lista as colunas
2 colnames(ppbio)
3 #Escolher quais colunas usar por nome
4 colnames(ppbio)[rev(order(colSums(ppbio)))] #ordena por maior soma
5 #Usar a função subset()
6 m_part <- subset(ppbio_a[, c("a.veloc", "a.temp", "a.do", "a.transp")])
7 m_part <- subset(ppbio_a[, 18:26])
```

2.5 Reset point

⚠ Aqui substitui-se uma nova matriz de dados, caso seja necessário refazer a análise com uma matriz gerada nesse código.

```
1 m_trab <- (ppbio)
2
3 #m_trab
```

① Substitua a nova matriz aqui. Caso seja necessário.

No interesse de sistematizar o uso das várias matrizes que são comumente usadas em uma AMD, a tabela a seguir (Tabela 1) resume seus tipos e abreviações.

3 Classificação 1: Matriz comunitária

Para conhecermos os dados, vamos criar uma classificação baseada na distância Bray-Curtis e UPGMA como método de fusão, a partir das matrizes de dados de interesse `ppbio**` com suas devidas relativizações e transformações (Veja Tabela 2).

3.1 Dendrograma e Heatmap 1

Tabela 1: Nomenclatura das matrizes em AMD em relação aos atributos das colunas.

Nome	Atributos (colunas)	Abreviação no R
Matriz comunitaria	Os atributos são táxons ou OTU's (Unidades Taxonômicas Operacionais) (ex. espécies, gêneros, morfotipos)	m_com
Matriz ambiental	Os atributos são dados ambientais e variáveis físicas e químicas (ex. pH, condutividade, temperatura)	m_amb
Matriz de habitat	Os atributos são elementos da estrutura do habitat (ex. macrófitas, algas, pedras, lama, etc)	m_hab
Matriz bruta	Os atributos ainda não receberam nenhum tipo de tratamento estatístico (valores brutos, como coletados)	m_brt
Matriz transposta	Os atributos foram transpostos para as linhas	m_t
Matriz relativizada	Os atributos foram relativizados por um critério de tamanho ou de variação (ex. dividir os valores de cada coluna pela soma)	m_rel
Matriz transformada	Foi aplicado um operador matemático a todos os atributos (ex. raiz quadrada, log)	m_trns
Matriz de trabalho	Qualquer matriz que seja o foco da análise atual (ex. comunitária, relativizada, etc)	m_trab

```

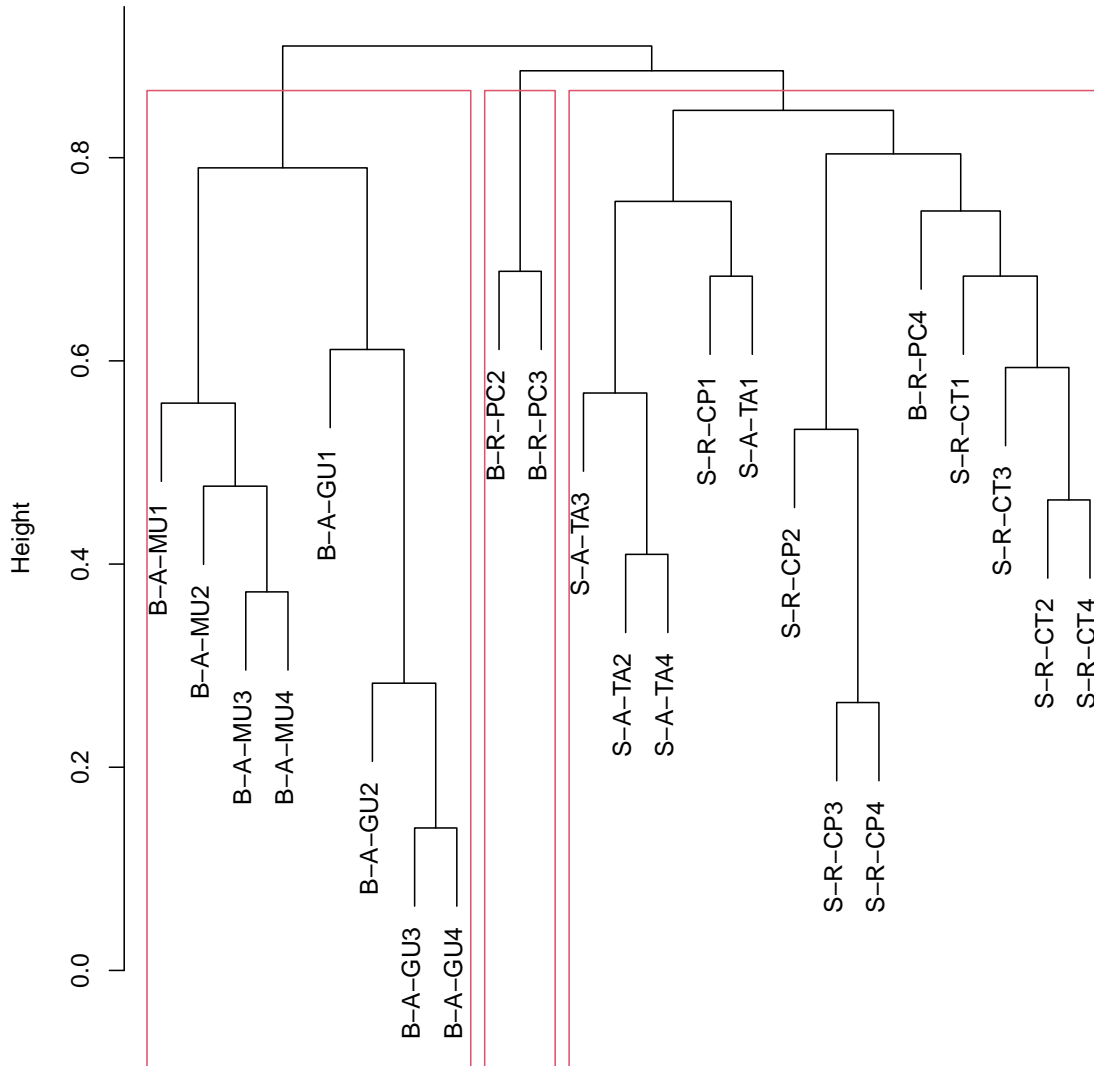
1  #Dendrograma
2  library(vegan)
3  m_trns <- asin(sqrt(decostand(m_trab,
4                                method="total", MARGIN = 2)))
5  #m_trns <- sqrt(m_trab)
6  vegdist <- vegdist(m_trns, method = "bray",
7                    diag = TRUE,
8                    upper = FALSE)
9  cluster_uas <- hclust(vegdist, method = "average")
10 plot (cluster_uas, main = "Cluster Dendrogram - Bray-Curtis da Matriz Comunitária",
11       hang = 0.1) #testar com -.01
12 rect.hclust(cluster_uas, k = 3, h = NULL)
13 #h = 0.8 fornece os grupos formados na altura h
14 as.matrix(vegdist)[1:6, 1:6]
15
16 #Heatmap
17 library("gplots")
18 heatdist <- as.matrix(vegdist)
19 col <- rev(heat.colors(999)) #rev() reverte as cores do heatmap
20 heatmap.2(x=(as.matrix(vegdist)), #objetos x objetos
21           Rowv = as.dendrogram(cluster_uas),
22           Colv = as.dendrogram(cluster_uas),
23           key = T, tracecol = NA, revC = T,
24           col = heat.colors, #dissimilaridade = 1 - similaridade
25           density.info = "none",
26           xlab = "UA´s", ylab = "UA´s",
27           mar = c(6, 6) + 0.2)
28 cluster_spp <- hclust((vegdist(t(m_trns), method = "bray",
29                             diag = TRUE,
30                             upper = FALSE)), method = "average")
31 plot (cluster_spp, main = "Dendrograma dos atributos")
32 heatmap.2(t(as.matrix(m_trns)), #objetos x atributos
33           Colv = as.dendrogram(cluster_uas),
34           Rowv = as.dendrogram(cluster_spp),
35           key = T, tracecol = NA, revC = T,
36           col = col,
37           density.info = "none",
38           xlab = "Unidades amostrais", ylab = "Espécies",
39           mar = c(6, 6) + 0.1) # adjust margin size

```

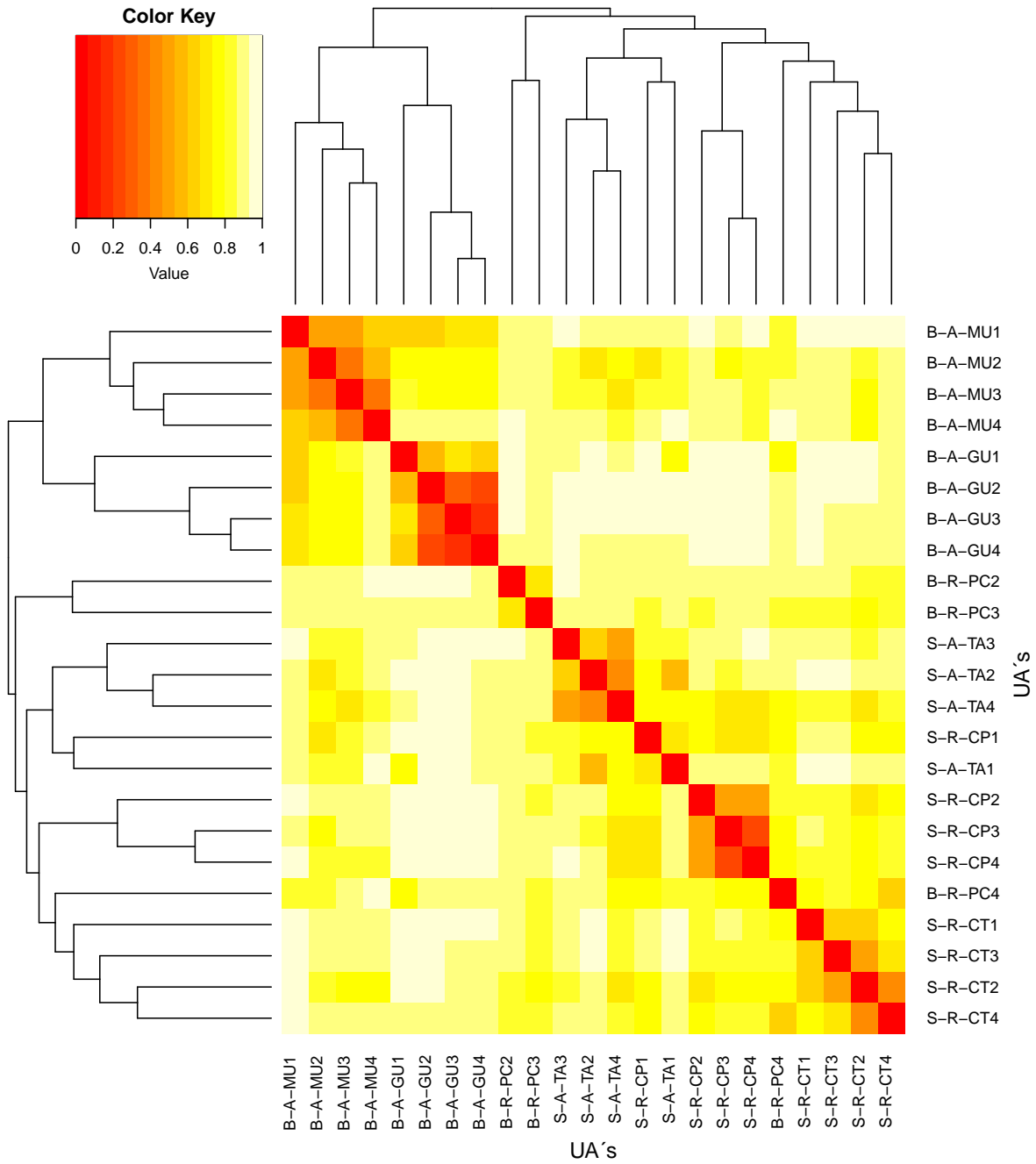
	S-R-CT1	S-R-CP1	S-A-TA1	S-R-CT2	S-R-CP2	S-A-TA2
S-R-CT1	0.000000	0.8743721	0.9338269	0.6274997	0.8106894	0.9420728

S-R-CP1 0.8743721 0.0000000 0.6833816 0.7759468 0.7726098 0.7342613
S-A-TA1 0.9338269 0.6833816 0.0000000 0.8789631 0.9178304 0.5700984
S-R-CT2 0.6274997 0.7759468 0.8789631 0.0000000 0.7280378 0.8836068
S-R-CP2 0.8106894 0.7726098 0.9178304 0.7280378 0.0000000 0.8915271
S-A-TA2 0.9420728 0.7342613 0.5700984 0.8836068 0.8915271 0.0000000

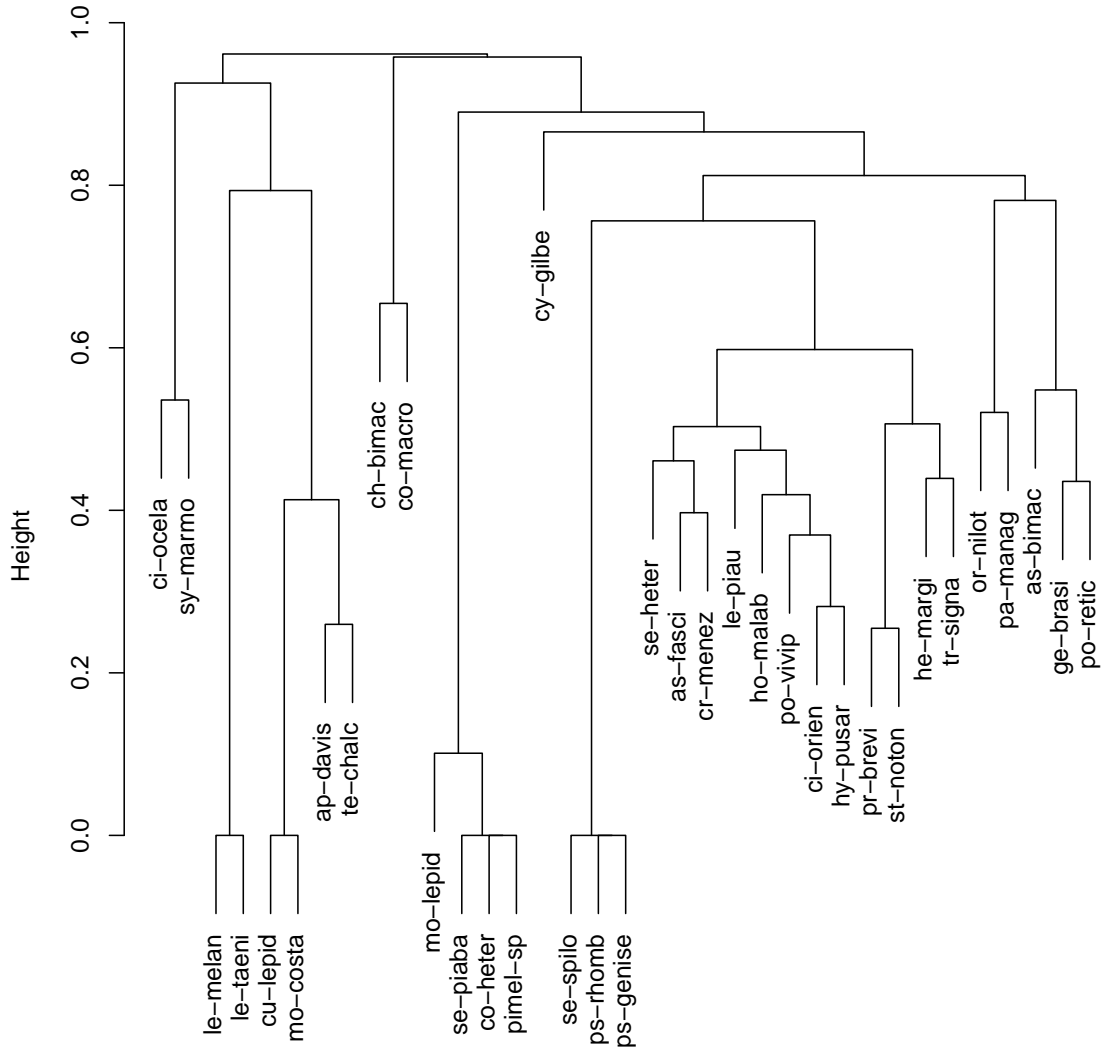
Cluster Dendrogram – Bray–Curtis da Matriz Comunitária



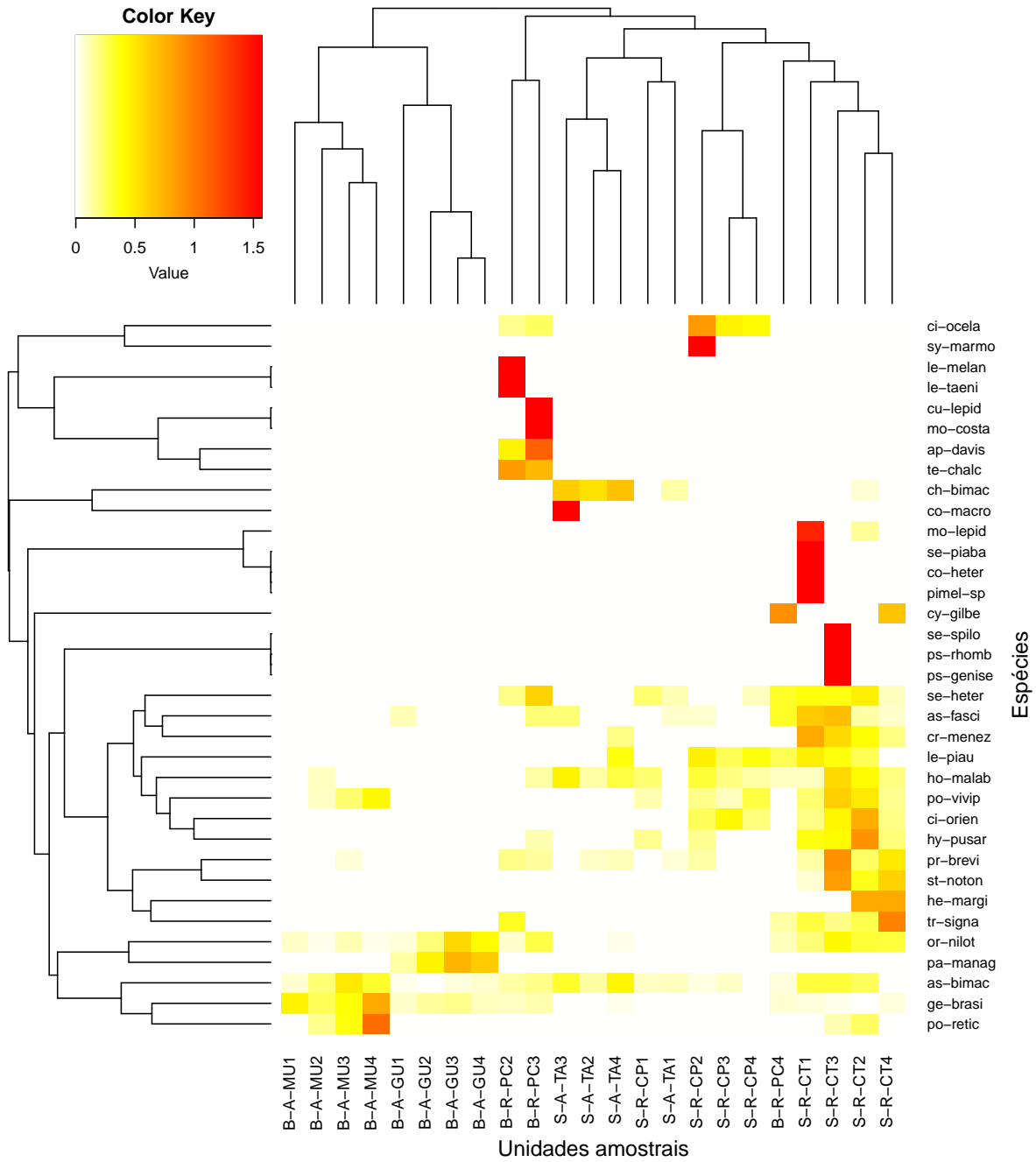
vegdist
 hclust (*, "average")



Dendrograma dos atributos



```
(vegdist(t(m_trns), method = "bray", diag = TRUE, upper = FALSE))
hclust (*, "average")
```



3.2 Histórico das fusões 1

Criamos agora o histórico das fusões dos objetos. Na tabela gerada, as duas primeiras colunas (No. e UA) representam o número (No.) atribuído a cada unidade amostral (UA). As duas

colunas subsequentes (Cluster1 e Cluster2) representam o par de objetos (indicado pelo sinal de “-”) ou grupo de objetos (indicado pela ausência do sinal de “-”) que foram agrupadas. A coluna Height, indica o valor de similaridade na qual um dado par de objetos (ou grupo de objetos) foi agrupado. O valor aproximado de Height também pode ser visualizado no eixo do dendrograma. Por último, na coluna Histórico, é mostrada a sequência das fusões da primeira até a m-1 última fusão entre os dois últimos grupos. Nesse caso, 22.

```

1 library(gt)
2 merge <- as.data.frame(cluster_uas$merge)
3 merge[nrow(merge)+1,] = c("0", "0")
4 height <- as.data.frame(round(cluster_uas$height, 2))
5 height[nrow(height)+1,] = c("1.0")
6 fusoes <- data.frame(Cluster_uas = merge, Height = height)
7 colnames(fusoes) <- c("Cluster1", "Cluster2", "Height")
8 UA <- rownames_to_column(as.data.frame(m_trns[, 0]))
9 colnames(UA) <- c("UAs")
10 No.UA <- 1:nrow(fusoes)
11 fusoes <- cbind(No.UA, UA, fusoes)
12 fusoes$Histórico <- 1:nrow(fusoes)
13 #fusoes
14 gt(fusoes)

```

No.UA	UAs	Cluster1	Cluster2	Height	Histórico
1	S-R-CT1	-20	-23	0.14	1
2	S-R-CP1	-8	-11	0.26	2
3	S-A-TA1	-17	1	0.28	3
4	S-R-CT2	-19	-22	0.37	4
5	S-R-CP2	-6	-12	0.41	5
6	S-A-TA2	-4	-10	0.46	6
7	S-R-CT3	-16	4	0.48	7
8	S-R-CP3	-5	2	0.53	8
9	S-A-TA3	-13	7	0.56	9
10	S-R-CT4	-9	5	0.57	10
11	S-R-CP4	-7	6	0.59	11
12	S-A-TA4	-14	3	0.61	12
13	B-A-MU1	-2	-3	0.68	13
14	B-A-GU1	-1	11	0.68	14
15	B-R-PC2	-15	-18	0.69	15
16	B-A-MU2	-21	14	0.75	16
17	B-A-GU2	10	13	0.76	17
18	B-R-PC3	9	12	0.79	18

19	B-A-MU3	8	16	0.8	19
20	B-A-GU3	17	19	0.85	20
21	B-R-PC4	15	20	0.89	21
22	B-A-MU4	18	21	0.91	22
23	B-A-GU4	0	0	1.0	23

No código acima, $h = 0.8$ fornece os grupos formados na altura h do eixo das distâncias do dendrograma. Ou seja, no dendrograma, o eixo y (HEIGHT, “h”) representa o valor da distancia escolhida entre os objetos ou grupos de objetos. Portanto, se dois objetos ou grupos de objetos foram agrupados num dado valor (0.8, por exemplo) no eixo `height`, isso significa que a distancia entre esses objetos é 0.8.

💡 Note que, no R, escalar os dados (argumento `scale=TRUE`) já significa centrar (argumento `center=TRUE`) e escalar, uma vez que no ambiente de programação do R base, ao escalar os dados são primeiro centrados (argumento `center=TRUE` como o padrão) (mas veja também a Tabela 2).

4 Classificação 2: Matriz ambiental

Para conhecermos os dados da matriz ambiental do PPBio, vamos criar uma classificação baseada na distância Bray-Curtis e UPGMA como método de fusão, a partir das matrizes de dados de interesse `ppbio**` com suas devidas relativizações e transformações (Veja Tabela 2).

4.1 Dendrograma e Heatmap 2

```

1 #Dendrograma
2 library(vegan)
3 #m_trns <- asin(sqrt(decostand(m_trab,
4 #                                     method="total", MARGIN = 2)))
5 m_trns <- sqrt(m_part)
6 vegdist <- vegdist(m_trns, method = "bray",
7                   diag = TRUE,
8                   upper = FALSE)
9 cluster_uas <- hclust(vegdist, method = "average")
10 plot (cluster_uas, main = "Cluster Dendrogram - Bray-Curtis para a Matriz Ambiental",
11       hang = 0.1) #testar com -.01
12 rect.hclust(cluster_uas, k = 3, h = NULL)

```

Tabela 2: Resumo sobre terminologia usada conforme definido para fins de consistência.

Termo	Descrição
Transformar (ou ponderar)	Aplicação de uma única função matemática a todos os valores. Ex. Log, Raiz, etc.
Relativizar	Normalizar
Padronizar ()	Reescalar os dados para apresentarem uma média = 0 e um desvio padrão = 1, subtraindo a média de cada valor e dividindo pelo desvio padrão
Normalizar ()	Dividir pela norma. Reescalar os valores para variar entre 0 e 1
Centrar	Padronizar
Reescalar	Adicionar ou subtrair uma constante, e então multiplicar ou dividir por uma constante. Significa mudar a unidade de medida. Ex. Celsius para Fahrenheit
Escalar (X)	Relativizar. Dividir cada variável por um fator. Variáveis diferentes têm fatores de escalar diferentes
'center' (no R*)	Um valor lógico (TRUE, FALSE) indicando se as variáveis devem ser deslocadas para serem centradas em zero. O valor é passado para a função 'scale'
'scale' (no R*)	Um valor lógico (TRUE, FALSE) indicando se as variáveis devem ser re-escaladas para terem variância unitária antes da análise, é recomendável fazer a escala. Ao re-escalar, os dados também são automaticamente centrados
() termo consistente na literatura; (X) termo pode ter mais de um significado diferente; ()	(*no R), note que escalar os dados (argumento 'scale=TRUE') já significa centrar ¹⁷ (argumento 'center=TRUE') e escalar, uma vez que no ambiente de programação do R base, ao escalar os dados são primeiro centrados (argumento 'center=TRUE' como o

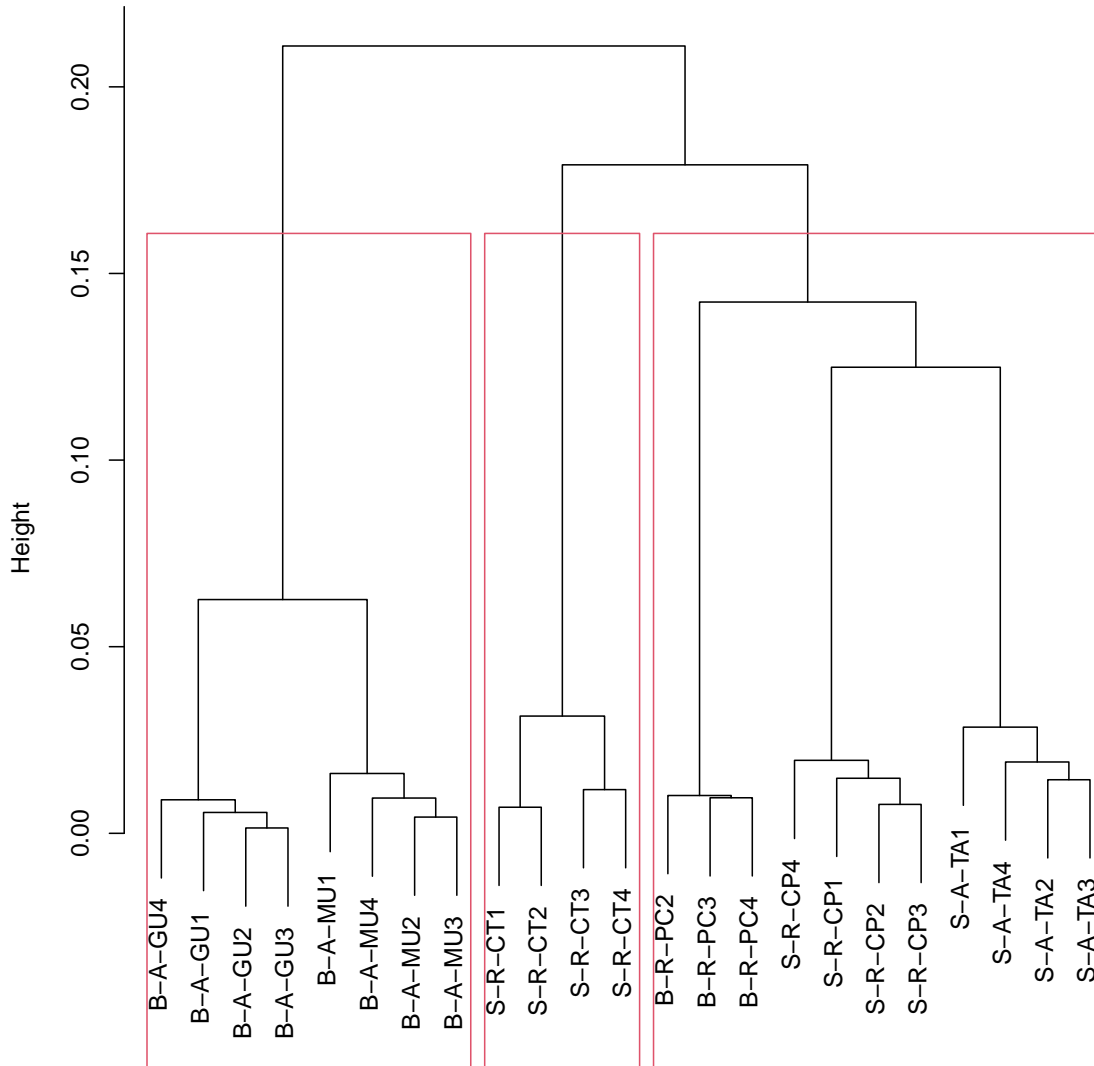
```

13 #h = 0.8 fornece os grupos formados na altura h
14 as.matrix(vegdist)[1:6, 1:6]
15
16 #Heatmap
17 library("gplots")
18 heatdist <- as.matrix(vegdist)
19 col <- rev(heat.colors(999)) #rev() reverte as cores do heatmap
20 heatmap.2(x=(as.matrix(vegdist)), #objetos x objetos
21           Rowv = as.dendrogram(cluster_uas),
22           Colv = as.dendrogram(cluster_uas),
23           key = T, tracecol = NA, revC = T,
24           col = heat.colors, #dissimilaridade = 1 - similaridade
25           density.info = "none",
26           xlab = "UA´s", ylab = "UA´s",
27           mar = c(6, 6) + 0.2)
28 cluster_spp <- hclust((vegdist(t(m_trns), method = "bray",
29                           diag = TRUE,
30                           upper = FALSE))), method = "average")
31 plot (cluster_spp, main = "Dendrograma dos atributos")
32 heatmap.2(t(as.matrix(m_trns)), #objetos x atributos
33           Colv = as.dendrogram(cluster_uas),
34           Rowv = as.dendrogram(cluster_spp),
35           key = T, tracecol = NA, revC = T,
36           col = col,
37           density.info = "none",
38           xlab = "Unidades amostrais", ylab = "Espécies",
39           mar = c(6, 6) + 0.1) # adjust margin size

```

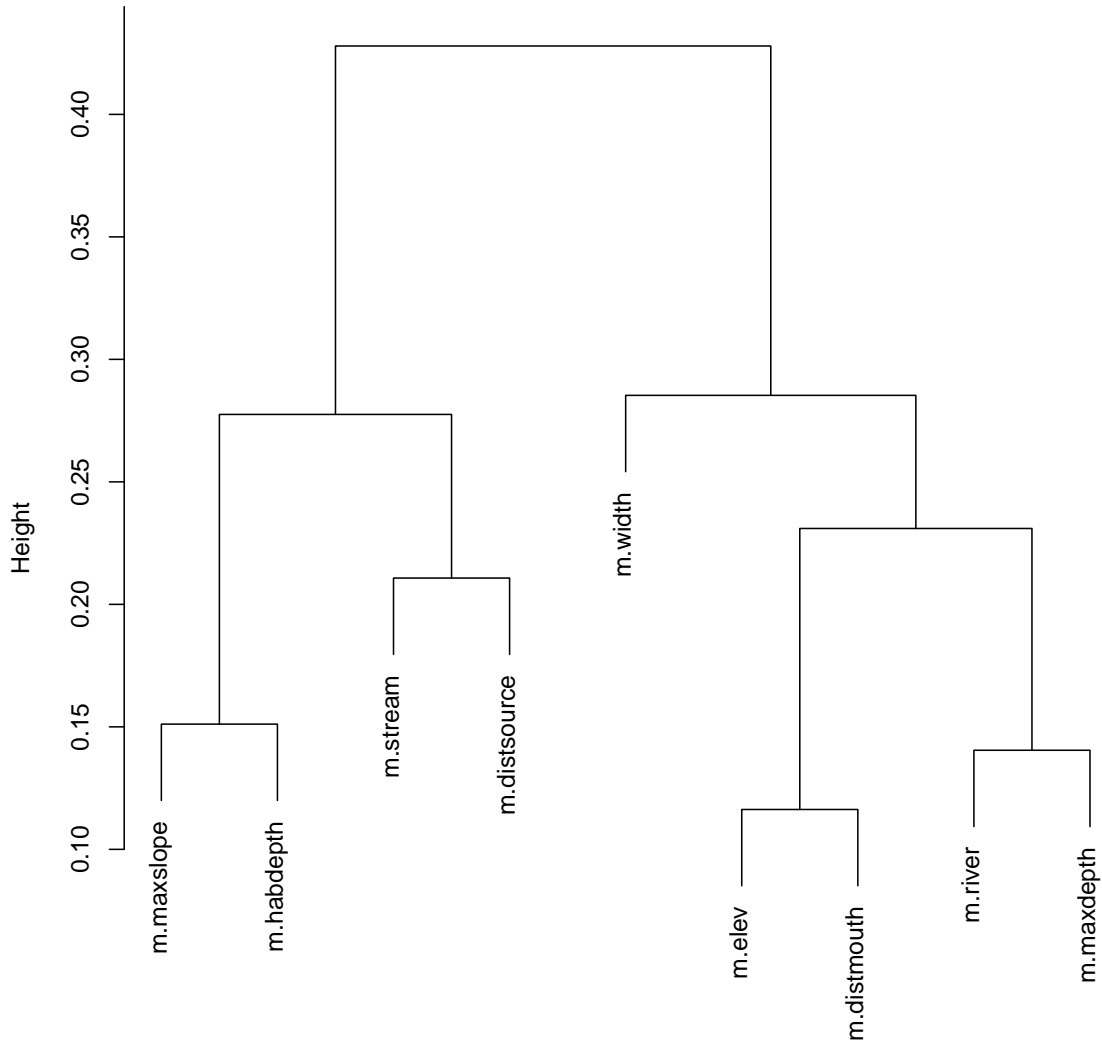
	S-R-CT1	S-R-CP1	S-A-TA1	S-R-CT2	S-R-CP2	S-A-TA2
S-R-CT1	0.00000000	0.19976326	0.16073975	0.006988768	0.18788437	0.16116410
S-R-CP1	0.199763264	0.00000000	0.14944141	0.194963386	0.01192482	0.13330395
S-A-TA1	0.160739745	0.14944141	0.00000000	0.159882118	0.13776209	0.01646547
S-R-CT2	0.006988768	0.19496339	0.15988212	0.00000000	0.18485182	0.16029566
S-R-CP2	0.187884372	0.01192482	0.13776209	0.184851815	0.00000000	0.12157239
S-A-TA2	0.161164104	0.13330395	0.01646547	0.160295661	0.12157239	0.00000000

Cluster Dendrogram – Bray–Curtis para a Matriz Ambiental

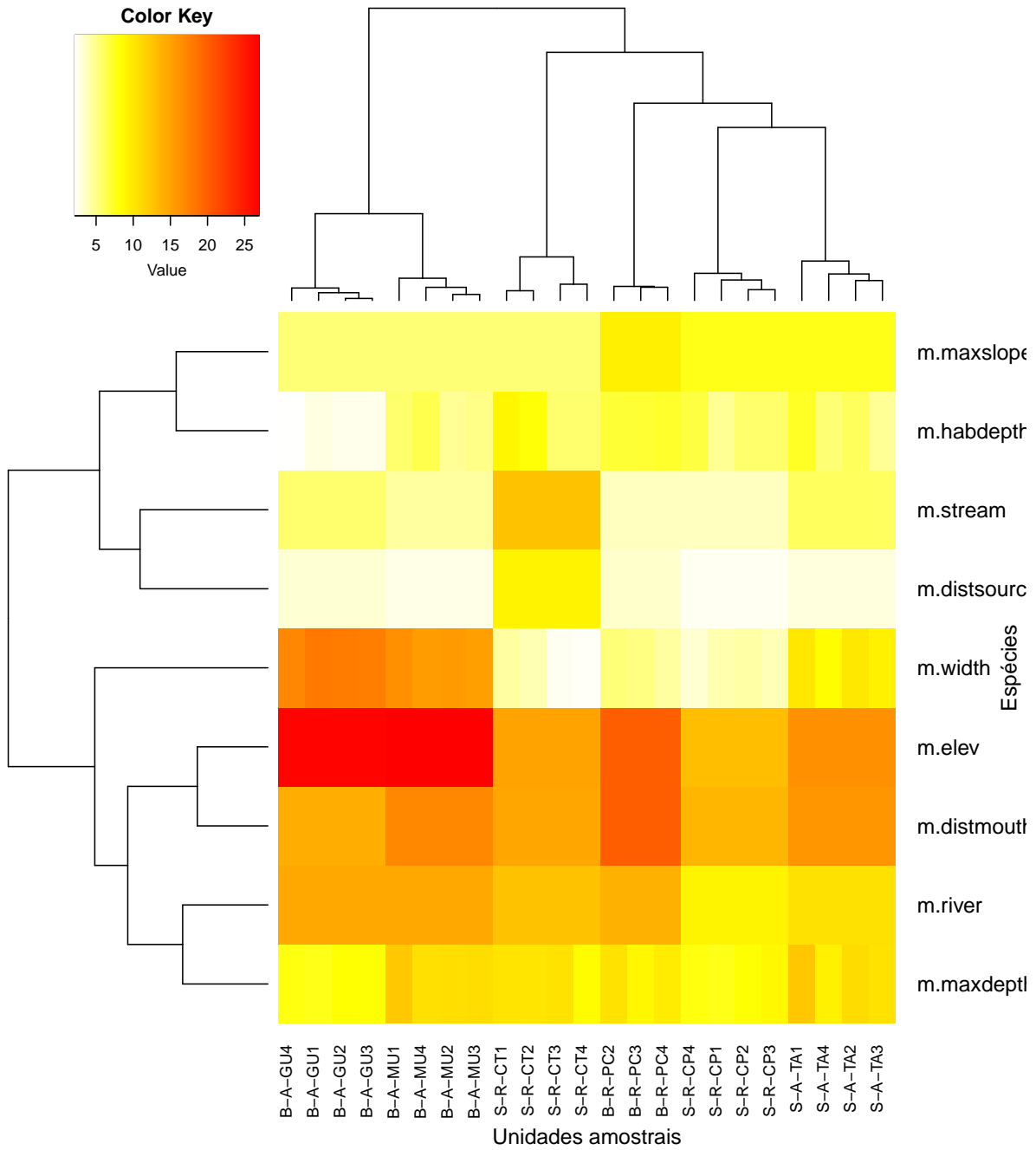


vegdist
hclust (*, "average")

Dendrograma dos atributos



```
(vegdist(t(m_trns), method = "bray", diag = TRUE, upper = FALSE))  
hclust (*, "average")
```



4.2 Histórico das fusões

Criamos agora o histórico das fusões dos objetos. Na tabela gerada, as duas primeiras colunas (No. e UA) representam o número (No.) atribuído a cada unidade amostral (UA). As duas

colunas subsequentes (Cluster1 e Cluster2) representam o par de objetos (indicado pelo sinal de “-”) ou grupo de objetos (indicado pela ausência do sinal de “-”) que foram agrupadas. A coluna Height, indica o valor de similaridade na qual um dado par de objetos (ou grupo de objetos) foi agrupado. O valor aproximado de Height também pode ser visualizado no eixo do dendrograma. Por último, na coluna Histórico, é mostrada a sequência das fusões da primeira até a m-1 última fusão entre os dois últimos grupos. Nesse caso, 22.

```

1 library(gt)
2 merge <- as.data.frame(cluster_uas$merge)
3 merge[nrow(merge)+1,] = c("0", "0")
4 height <- as.data.frame(round(cluster_uas$height, 2))
5 height[nrow(height)+1,] = c("1.0")
6 fusoes <- data.frame(Cluster_uas = merge, Height = height)
7 colnames(fusoes) <- c("Cluster1", "Cluster2", "Height")
8 UA <- rownames_to_column(as.data.frame(m_trns[, 0]))
9 colnames(UA) <- c("UAs")
10 No.UA <- 1:nrow(fusoes)
11 fusoes <- cbind(No.UA, UA, fusoes)
12 fusoes$Histórico <- 1:nrow(fusoes)
13 #fusoes
14 gt(fusoes)

```

No.UA	UAs	Cluster1	Cluster2	Height	Histórico
1	S-R-CT1	-17	-20	0	1
2	S-R-CP1	-16	-19	0	2
3	S-A-TA1	-14	1	0.01	3
4	S-R-CT2	-1	-4	0.01	4
5	S-R-CP2	-5	-8	0.01	5
6	S-A-TA2	-23	3	0.01	6
7	S-R-CT3	-22	2	0.01	7
8	S-R-CP3	-18	-21	0.01	8
9	S-A-TA3	-15	8	0.01	9
10	S-R-CT4	-7	-10	0.01	10
11	S-R-CP4	-6	-9	0.01	11
12	S-A-TA4	-2	5	0.01	12
13	B-A-MU1	-13	7	0.02	13
14	B-A-GU1	-12	11	0.02	14
15	B-R-PC2	-11	12	0.02	15
16	B-A-MU2	-3	14	0.03	16
17	B-A-GU2	4	10	0.03	17
18	B-R-PC3	6	13	0.06	18

19	B-A-MU3	15	16	0.12	19
20	B-A-GU3	9	19	0.14	20
21	B-R-PC4	17	20	0.18	21
22	B-A-MU4	18	21	0.21	22
23	B-A-GU4	0	0	1.0	23

5 Classificação 3: Heatmap Matriz comunitária vs. Matriz ambiental

5.1 Dendrograma e Heatmap 3

```

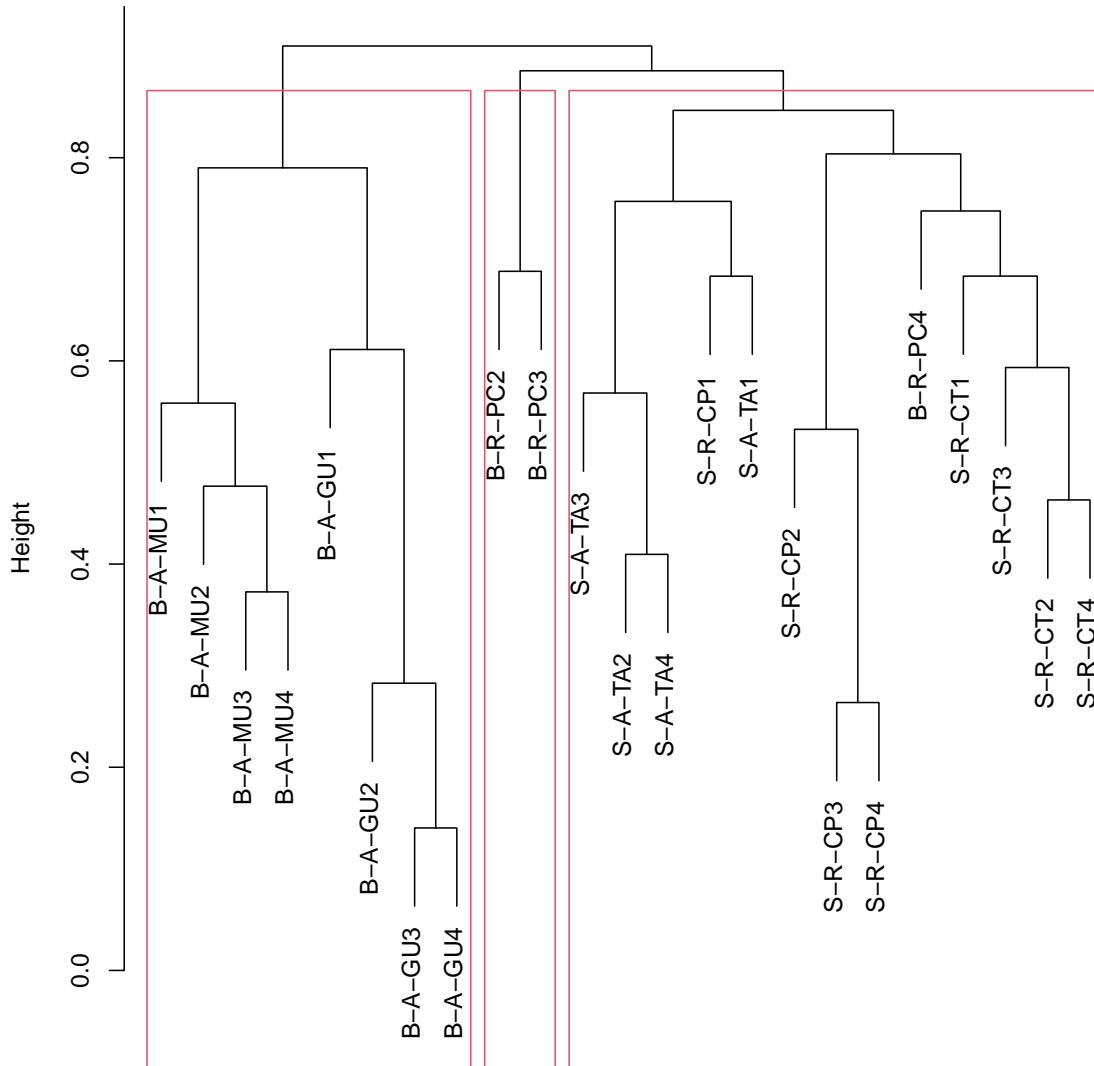
1 #Dendrograma
2 library(vegan)
3 library(gplots)
4 # Dendrograma para as UA's da matriz comunitária
5 m_trns_c <- asin(sqrt(decostand(ppbio,
6                       method="total", MARGIN = 2)))
7 vegdist_c <- vegdist(m_trns_c, method = "bray",
8                       diag = TRUE,
9                       upper = FALSE)
10 cluster_uas_c <- hclust(vegdist_c, method = "average")
11 plot (cluster_uas_c, main = "Cluster Dendrogram - Bray-Curtis da Matriz Comunitária",
12       hang = 0.1) #testar com -.01
13 rect.hclust(cluster_uas_c, k = 3, h = NULL)
14 #h = 0.8 fornece os grupos formados na altura h
15 as.matrix(vegdist_c)[1:6, 1:6]
16
17 # Dendrograma para as variáveis morfológicas da matriz ambiental
18 m_trns_a <- sqrt(m_part)
19 cluster_mfg_a <- hclust((vegdist(t(m_trns_a),
20                               method = "bray",
21                               diag = TRUE,
22                               upper = FALSE)), method = "average")
23 plot (cluster_mfg_a, main = "Dendrograma dos atributos")
24
25 # Heatmap Comunidade vs. Morfologia
26 heatmap.2(t(as.matrix(m_trns_a)), #objetos x atributos
27           Colv = as.dendrogram(cluster_uas_c),
28           Rowv = as.dendrogram(cluster_mfg_a),

```

```
29     key = T, tracecol = NA, revC = T,  
30     col = col,  
31     density.info = "none",  
32     xlab = "Unidades amostrais", ylab = "Morfologia",  
33     mar = c(6, 9) + 0.1) # adjust margin size
```

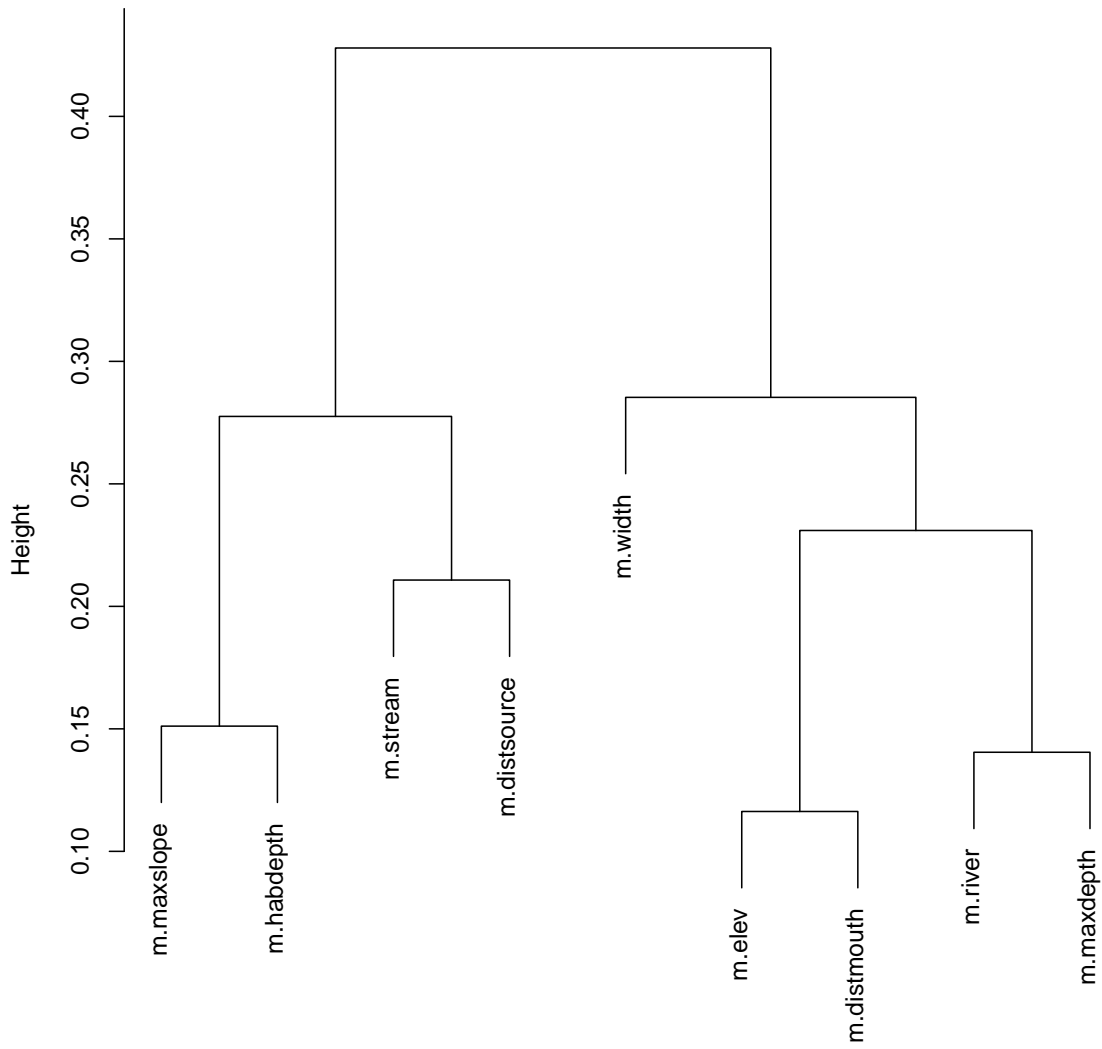
	S-R-CT1	S-R-CP1	S-A-TA1	S-R-CT2	S-R-CP2	S-A-TA2
S-R-CT1	0.0000000	0.8743721	0.9338269	0.6274997	0.8106894	0.9420728
S-R-CP1	0.8743721	0.0000000	0.6833816	0.7759468	0.7726098	0.7342613
S-A-TA1	0.9338269	0.6833816	0.0000000	0.8789631	0.9178304	0.5700984
S-R-CT2	0.6274997	0.7759468	0.8789631	0.0000000	0.7280378	0.8836068
S-R-CP2	0.8106894	0.7726098	0.9178304	0.7280378	0.0000000	0.8915271
S-A-TA2	0.9420728	0.7342613	0.5700984	0.8836068	0.8915271	0.0000000

Cluster Dendrogram – Bray–Curtis da Matriz Comunitária

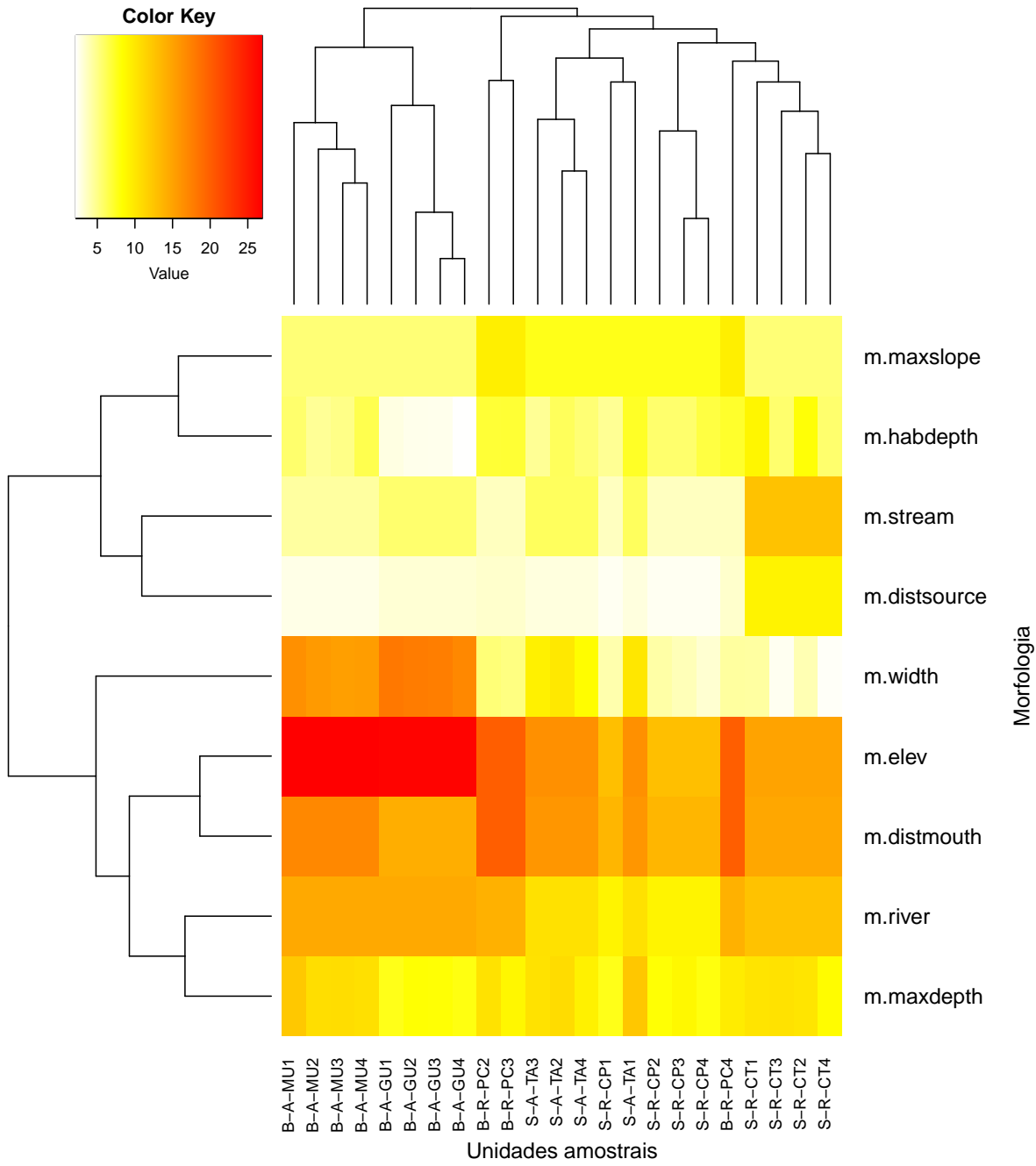


vegdist_c
 hclust (*, "average")

Dendrograma dos atributos



```
(vegdist(t(m_trns_a), method = "bray", diag = TRUE, upper = FALSE))  
hclust (*, "average")
```



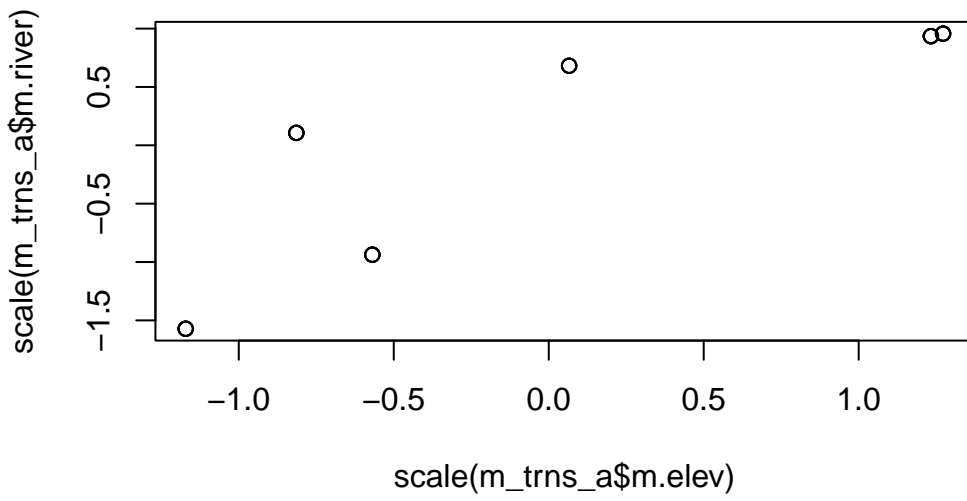
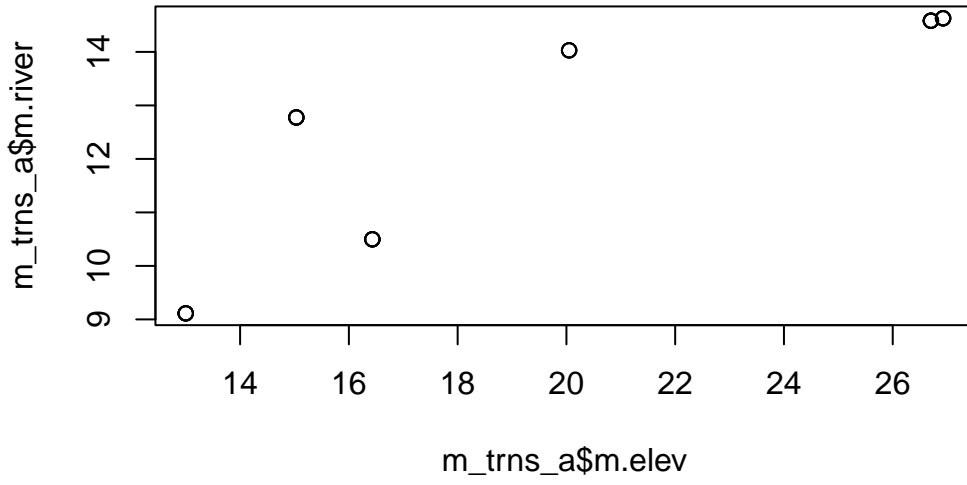
6 Explorando correlações multivariadas da matriz de habitat

Colinearidade entre variáveis é um fenômeno em que duas ou mais variáveis independentes em uma análise de redundância (RDA) estão altamente correlacionadas entre si. Isso significa que existe uma relação linear forte entre essas variáveis, tornando difícil distinguir o efeito individual de cada uma delas na variável dependente.

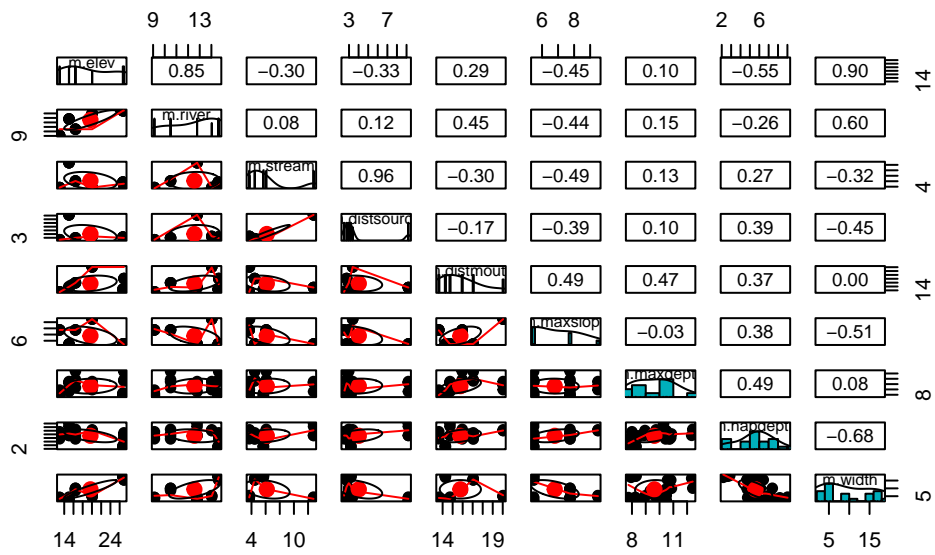
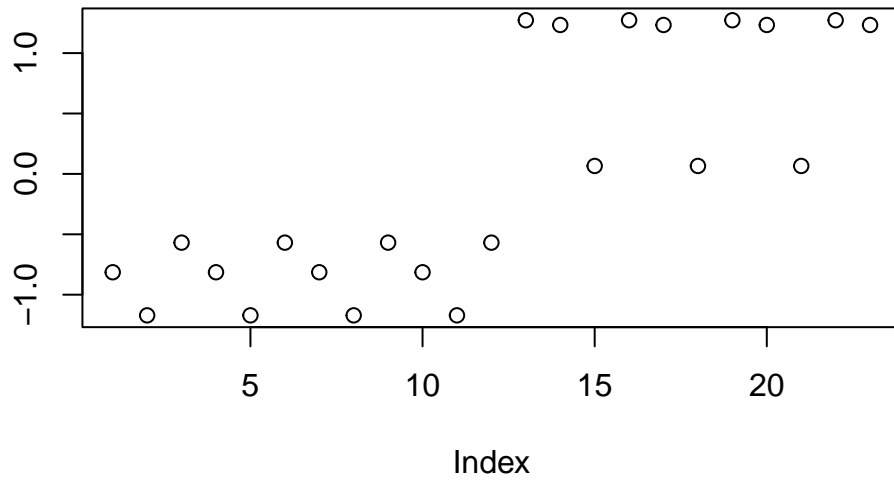
A colinearidade pode afetar negativamente a análise de redundância, pois torna os resultados menos confiáveis e interpretações mais difíceis. Quando as variáveis independentes estão altamente correlacionadas, os coeficientes de regressão (R^2) associados a cada variável podem se tornar instáveis ou mesmo incoerentes.

A colinearidade pode ser detectada por meio de várias medidas, como o coeficiente de correlação (R^2) entre as variáveis independentes ou o fator de inflação da variância (VIF). Se for identificada colinearidade, é necessário tomar medidas para lidar com ela. Isso pode incluir a remoção de uma ou mais variáveis independentes altamente correlacionadas ou a combinação de variáveis em um único fator.

```
1 plot(m_trns_a$"m.elev", m_trns_a$"m.river")
2 plot(scale(m_trns_a$"m.elev"), scale(m_trns_a$"m.river"))
3 plot((m_trns_a$"m.elev" - mean(m_trns_a$"m.elev")) / sd(m_trns_a$"m.elev"))
4
5 library(psych)
6 pairs.panels(m_trns_a[,1:9],
7             method = "pearson", # correlation method
8             scale = FALSE, lm = FALSE,
9             hist.col = "#00AFBB", pch = 19,
10            density = TRUE, # show density plots
11            ellipses = TRUE, # show correlation ellipses
12            alpha = 0.5
13            )
```



$a\$m.elev - \text{mean}(m_trns_a\$m.elev) / \text{sd}(m_trns$



7 Análise de Redundância

7.1 Transformando os dados de comunidade com o método de hellinger

```
1 hell <- decostand(ppbio, method = "hellinger")
2 m_trns_a <- sqrt(m_part)
```

A transformação usando o **método the Hellinger** calcula uma padronização total da matriz de dados comunitária, onde, se as unidades amostrais são linhas e as espécies são colunas, todos os valores em uma linha são divididos pela soma da linha e, em seguida, é tirada a raiz quadrada dos valores. A padronização de Hellinger é uma padronização convexa que simultaneamente ajuda a minimizar os efeitos das unidades amostrais com valores totais muito diferentes entre espécies.

7.2 Fazendo a RDA

```
1 rda <- rda(hell ~ ., m_trns_a) # variavel de resposta ~ variavel explanatória
2 summary(rda)
3 #str(rda)
4 modelo_rda <- rda$call
5 modelo_rda
```

Call:

```
rda(formula = hell ~ m.elev + m.river + m.stream + m.distsource + m.distmouth + m.maxsl
```

Partitioning of variance:

	Inertia	Proportion
Total	0.6376	1.0000
Constrained	0.4862	0.7626
Unconstrained	0.1513	0.2374

Eigenvalues, and their contribution to the variance

Importance of components:

	RDA1	RDA2	RDA3	RDA4	RDA5	RDA6	RDA7
Eigenvalue	0.1649	0.1263	0.07956	0.06226	0.03295	0.01049	0.007194
Proportion Explained	0.2586	0.1980	0.12479	0.09765	0.05168	0.01645	0.011284
Cumulative Proportion	0.2586	0.4567	0.58144	0.67909	0.73078	0.74723	0.758511
	RDA8	PC1	PC2	PC3	PC4	PC5	PC6

Eigenvalue	0.002619	0.04492	0.03190	0.02025	0.01436	0.01106	0.01016
Proportion Explained	0.004109	0.07045	0.05004	0.03177	0.02252	0.01735	0.01594
Cumulative Proportion	0.762619	0.83307	0.88311	0.91488	0.93740	0.95475	0.97068
	PC7	PC8	PC9	PC10	PC11	PC12	
Eigenvalue	0.006627	0.006149	0.001784	0.001618	0.001420	0.0008005	
Proportion Explained	0.010394	0.009644	0.002798	0.002538	0.002228	0.0012556	
Cumulative Proportion	0.981079	0.990723	0.993521	0.996059	0.998287	0.9995425	
	PC13	PC14					
Eigenvalue	0.0002513	4.040e-05					
Proportion Explained	0.0003942	6.337e-05					
Cumulative Proportion	0.9999366	1.000e+00					

Accumulated constrained eigenvalues

Importance of components:

	RDA1	RDA2	RDA3	RDA4	RDA5	RDA6	RDA7
Eigenvalue	0.1649	0.1263	0.07956	0.06226	0.03295	0.01049	0.007194
Proportion Explained	0.3391	0.2597	0.16363	0.12805	0.06777	0.02157	0.014796
Cumulative Proportion	0.3391	0.5988	0.76243	0.89048	0.95825	0.97982	0.994613
	RDA8						
Eigenvalue	0.002619						
Proportion Explained	0.005387						
Cumulative Proportion	1.000000						

Scaling 2 for species and site scores

- * Species are scaled proportional to eigenvalues
- * Sites are unscaled: weighted dispersion equal on all dimensions
- * General scaling constant of scores: 1.935249

Species scores

	RDA1	RDA2	RDA3	RDA4	RDA5	RDA6
ap-davis	-0.0031620	-0.0048824	-0.0110049	0.093593	-0.068101	0.0059371
as-bimac	0.2532136	-0.3124256	0.1416298	0.059634	0.052751	0.0505050
as-fasci	0.0023602	0.0301488	-0.0863601	0.104203	0.022376	0.0654845
ch-bimac	0.2591420	-0.5253309	-0.2327575	-0.128133	0.056150	-0.0025083
ci-ocela	0.1452054	0.1788332	0.0081576	-0.125397	-0.182390	-0.0054260
ci-orien	0.1485460	0.2172991	-0.0246427	-0.093503	-0.007111	0.0306808
co-macro	0.0056272	-0.0133312	-0.0059647	-0.001866	0.001920	-0.0027535
co-heter	0.0021713	0.0036533	-0.0008304	0.002924	0.005287	0.0122382
cr-menez	0.0201166	0.0288392	-0.0201653	0.024919	0.061365	0.0379788
cu-lepid	-0.0013364	-0.0032992	-0.0056660	0.054729	-0.038285	0.0049398
cy-gilbe	0.0075912	0.0607007	-0.0558598	0.176400	-0.023957	-0.1368110

ge-brasi	-0.3412941	-0.1146370	0.4755923	0.062138	0.013636	0.0202855
he-margi	0.0037201	0.0083677	-0.0048951	0.006810	0.016194	0.0023251
ho-malab	0.1583100	0.0801276	-0.0438524	-0.061179	-0.028562	-0.0123754
hy-pusar	0.0591939	0.0843114	-0.0230111	0.019845	0.045930	0.0787908
le-melan	-0.0011347	-0.0009522	-0.0032922	0.023765	-0.018288	0.0005572
le-piau	0.0572735	0.0669133	-0.0106613	-0.030844	-0.029327	-0.0039150
le-taeni	-0.0008024	-0.0006733	-0.0023280	0.016805	-0.012931	0.0003940
mo-costa	-0.0002916	-0.0007200	-0.0012364	0.011943	-0.008354	0.0010779
mo-lepid	0.0153181	0.0261378	-0.0064633	0.020994	0.038374	0.0837101
or-nilot	-0.4004543	0.1306247	-0.1709455	0.125155	0.113756	0.0083875
pa-manag	-0.6061411	0.0222687	-0.2037387	-0.220476	-0.029783	0.0306444
pimel-sp	0.0053185	0.0089487	-0.0020342	0.007162	0.012950	0.0299773
po-retic	-0.0136644	-0.0116037	0.2017852	0.016084	0.086108	-0.0251859
po-vivip	0.2992231	0.4439822	0.1166819	-0.169611	0.091254	0.0023115
pr-brevi	0.0882396	0.0648018	-0.0904496	0.115900	0.113825	-0.0130077
ps-rhomb	0.0010222	0.0034565	-0.0025553	0.002365	0.006240	-0.0041103
ps-genise	0.0010222	0.0034565	-0.0025553	0.002365	0.006240	-0.0041103
se-heter	0.1284778	0.1073805	-0.0917259	0.224843	-0.125813	0.0717309
se-piaba	0.0179046	0.0301257	-0.0068480	0.024110	0.043596	0.1009185
se-spilo	0.0010222	0.0034565	-0.0025553	0.002365	0.006240	-0.0041103
st-noton	0.0376188	0.0976932	-0.0635617	0.074562	0.185679	-0.0350874
sy-marmo	0.0103219	0.0103782	0.0015469	-0.010811	-0.011405	0.0060782
te-chalc	-0.0092160	-0.0113525	-0.0297112	0.237453	-0.176359	0.0116444
tr-signa	0.0367759	0.0997503	-0.0760832	0.191037	0.103992	-0.0103384

Site scores (weighted sums of species scores)

	RDA1	RDA2	RDA3	RDA4	RDA5	RDA6
S-R-CT1	0.17970	0.156638	-0.08629	0.342788	0.4035001	1.75031
S-R-CP1	0.40555	0.214638	0.04913	-0.007242	-0.2312790	0.58568
S-A-TA1	0.36039	-0.619092	-0.29443	0.077366	-0.0394443	0.38825
S-R-CT2	0.28668	0.421059	-0.12267	0.049677	0.5720424	0.44985
S-R-CP2	0.41139	0.576296	-0.02432	-0.595618	-0.4851551	-0.23091
S-A-TA2	0.34037	-0.774492	-0.45350	-0.353282	0.1094824	-0.39290
S-R-CT3	0.21802	0.419735	-0.17093	0.165103	0.7866870	0.11028
S-R-CP3	0.41146	0.364174	0.07085	-0.582330	-0.4854795	-0.07357
S-A-TA3	0.35909	-0.728565	-0.30021	-0.261080	0.0605021	-0.07166
S-R-CT4	-0.03531	0.487586	-0.42742	0.590600	0.8751521	-1.41844
S-R-CP4	0.37264	0.748237	0.03980	-0.641120	-0.3641386	-0.40160
S-A-TA4	0.32964	-0.719561	-0.20884	-0.209133	0.1307171	-0.08917
B-A-MU1	-0.38147	-0.196706	0.86058	0.128704	-0.0778885	-0.15182
B-A-GU1	-0.70741	-0.086218	0.03338	-0.195696	-0.0619169	0.39735

B-R-PC2	0.02206	-0.071998	-0.01301	0.869858	-0.6528132	0.09833
B-A-MU2	-0.01557	-0.236449	0.82735	0.016805	0.1638010	0.01741
B-A-GU2	-0.86405	0.071061	-0.25799	-0.384771	-0.1284332	-0.10179
B-R-PC3	-0.06180	-0.006191	-0.11797	0.819380	-0.7536655	0.67026
B-A-MU3	0.05079	-0.219372	0.68718	0.023678	0.2629406	0.04611
B-A-GU3	-0.80446	0.080084	-0.38822	-0.251725	0.0297358	-0.08667
B-R-PC4	-0.05481	0.122046	-0.23416	0.795358	-0.3776719	-1.19962
B-A-MU4	-0.05600	-0.047197	0.97083	-0.088030	0.2623861	-0.27961
B-A-GU4	-0.76691	0.044287	-0.43914	-0.309291	0.0009388	-0.01604

Site constraints (linear combinations of constraining variables)

	RDA1	RDA2	RDA3	RDA4	RDA5	RDA6
S-R-CT1	0.18984	0.319415	-0.07261	0.25563	0.462238	1.07001
S-R-CP1	0.39387	0.341782	0.03231	-0.36167	-0.393723	0.13484
S-A-TA1	0.34967	-0.650776	-0.26972	-0.25836	-0.078132	0.25992
S-R-CT2	0.17636	0.333246	-0.12808	0.27425	0.537348	0.73035
S-R-CP2	0.40174	0.403932	0.06021	-0.42078	-0.443879	0.23657
S-A-TA2	0.35145	-0.783892	-0.27536	-0.15636	-0.007751	0.30537
S-R-CT3	0.12949	0.437854	-0.32369	0.29955	0.790412	-0.52067
S-R-CP3	0.39118	0.503594	0.01370	-0.46897	-0.401486	-0.12948
S-A-TA3	0.33356	-0.790217	-0.35356	-0.11060	0.113804	-0.16322
S-R-CT4	0.15339	0.394505	-0.28292	0.31873	0.847383	-0.38770
S-R-CP4	0.41426	0.654037	0.02925	-0.57489	-0.326964	-0.36234
S-A-TA4	0.35481	-0.616825	-0.35833	-0.22081	0.233336	-0.56756
B-A-MU1	-0.10948	-0.236390	0.85920	0.04447	0.032926	0.24271
B-A-GU1	-0.77508	-0.047464	-0.21268	-0.25784	-0.101352	0.42476
B-R-PC2	-0.04087	-0.034297	-0.11859	0.85603	-0.658731	0.02007
B-A-MU2	-0.10670	-0.247016	0.82133	0.07921	0.177749	-0.09601
B-A-GU2	-0.78677	0.004261	-0.25265	-0.27612	-0.074058	0.16556
B-R-PC3	-0.02104	-0.051936	-0.08919	0.86152	-0.602665	0.07776
B-A-MU3	-0.10471	-0.139735	0.80181	0.01519	0.223152	-0.37052
B-A-GU3	-0.78742	0.034200	-0.26532	-0.28991	-0.047048	0.04056
B-R-PC4	-0.03263	0.130089	-0.15737	0.76705	-0.522754	-0.52887
B-A-MU4	-0.08135	-0.076584	0.86359	-0.05771	0.177412	-0.14409
B-A-GU4	-0.79355	0.118218	-0.32132	-0.31762	0.062782	-0.43803

Biplot scores for constraining variables

	RDA1	RDA2	RDA3	RDA4	RDA5	RDA6
m.elev	-0.835824	-0.28357	0.45714	0.05373	0.0000987	-0.082609

```

m.river      -0.786512 -0.07793  0.31367  0.48435  0.1891617 -0.002859
m.stream     0.089774  0.21578 -0.33735  0.22083  0.8417884  0.265768
m.distsource 0.118244  0.38272 -0.27956  0.39620  0.7343292  0.250473
m.distmouth  -0.004434 -0.37609  0.24749  0.80424 -0.3165319 -0.176485
m.maxdepth   0.297290 -0.53131  0.39190  0.31985  0.2495003  0.149012
m.habdepth   0.696301  0.10174  0.08606  0.48504  0.0186011  0.305801
m.width      -0.759696 -0.45051  0.34823 -0.30998  0.0306750  0.025372

```

```

rda(formula = hell ~ m.elev + m.river + m.stream + m.distsource +
     m.distmouth + m.maxslope + m.maxdepth + m.habdepth + m.width,
     data = m_trns_a)

```

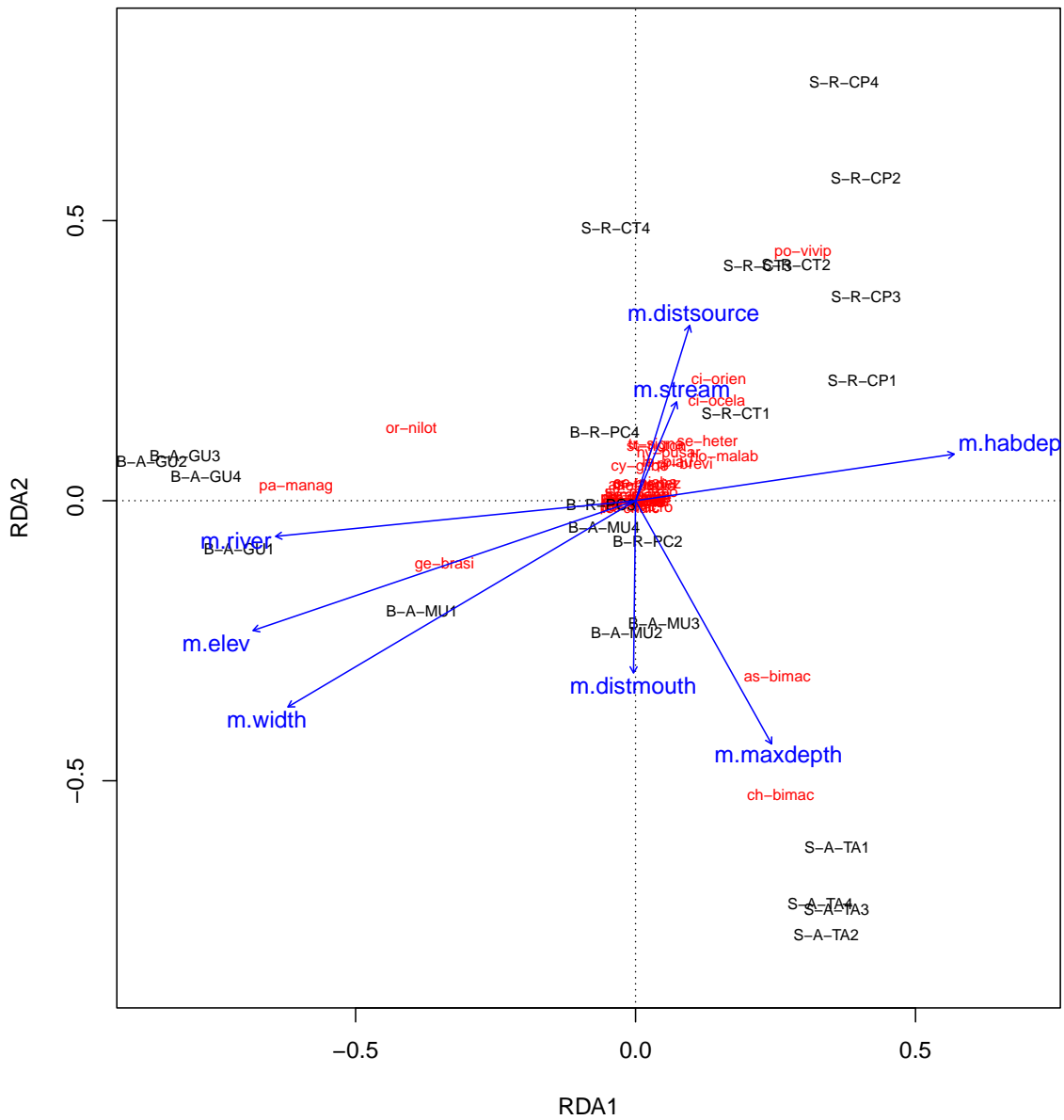
Note que o modelo proposto para a RDA para as principais variáveis explanatórias da matriz de comunidades é dado por modelo_rda.

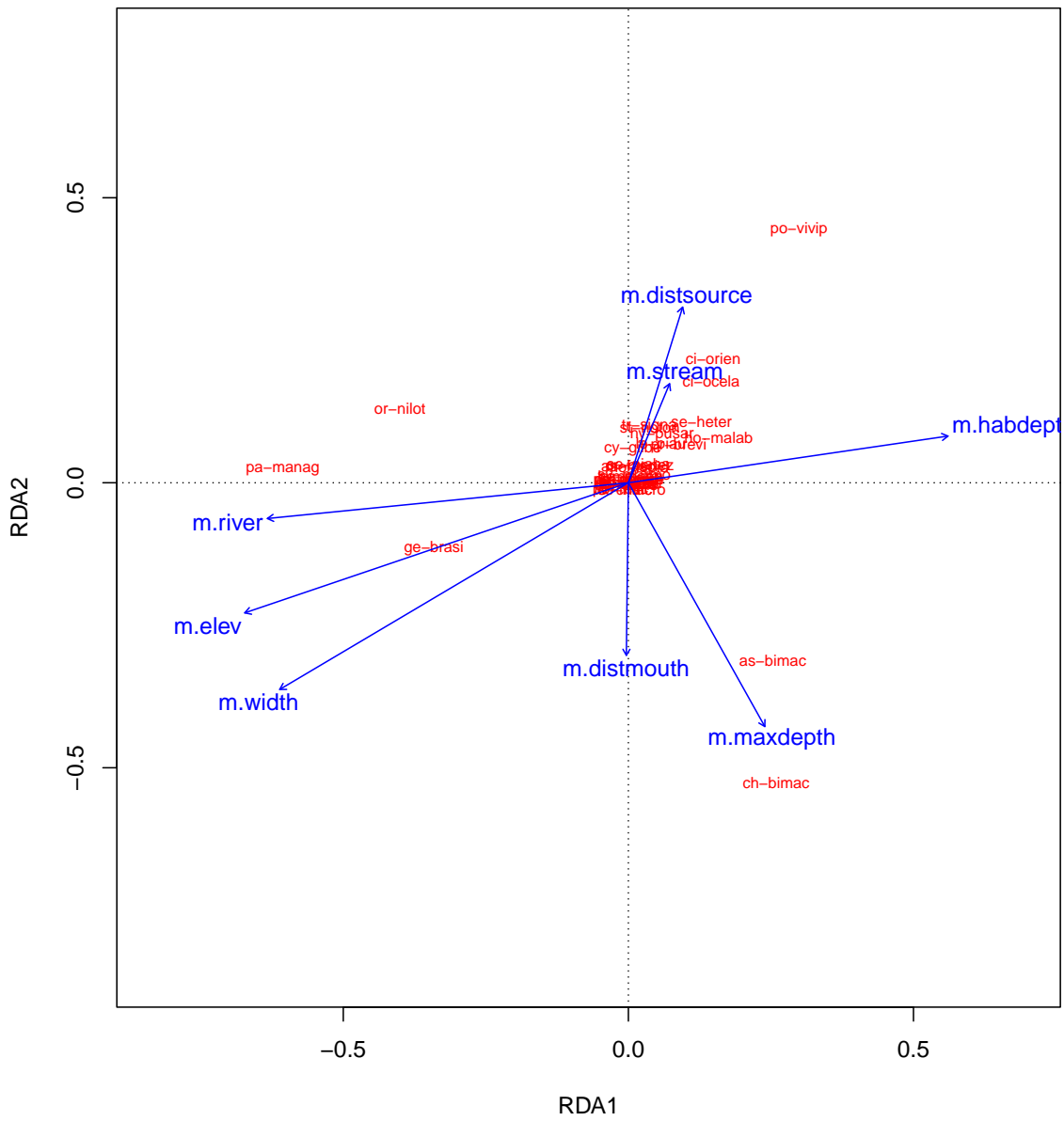
7.3 Gráficos de RDA

```

1  #Gráfico triplot basico
2  plot(rda, choices = c(1,2))
3  plot (rda, display = c("sp","bp"))
4
5  #Ajustes no gráfico triplot
6  model <- ordiplot(rda, type = "none", scaling = 2, xlab = "RDA1", ylab = "RDA2") # type="point"
7  points(rda, col = "black", pch = 21, bg = "gray", cex = 1)
8  text(rda, col = "black", cex = 0.7, pos = 1)
9  points(rda, dis = "bp", col = "red", pch = 21, cex = 1)
10 text(rda, dis = "bp", col = "red", cex = 0.9, pos = 2)
11 points(rda, dis = "sp", col = "blue", pch = 3, cex = 1)
12 text(rda, dis = "sp", col = "blue", cex = 0.8, pos = 1)
13
14 # Show only the first 5 species with higher values
15 #high_species <- order(abs(scores(rda, display = "species")[, 1]), decreasing = TRUE)[1:5]
16 # Get the indices of the first 5 species with higher values on the first axis
17 #text(rda, display = "species", select = high_species, col = "darkblue", pos = 4) # Display

```





8 Interpretando o biplot gerado no código anterior

Um gráfico de Análise de Redundância (RDA) é uma representação visual das relações entre as variáveis explicativas e as variáveis resposta em um modelo de RDA. Ele permite entender a estrutura dos dados e identificar padrões e relações importantes.

Existem diferentes tipos de gráficos que podem ser gerados a partir de uma análise de RDA, mas aqui vamos nos concentrar em um dos mais comuns: o gráfico de dispersão dos pontos (biplot) das amostras no espaço dos primeiros dois eixos canônicos (RDA1 e RDA2).

No gráfico de RDA, cada amostra é representada como um ponto no espaço bidimensional, onde os eixos representam combinações lineares das variáveis explicativas que maximizam a correlação com as variáveis resposta. Os pontos podem ser coloridos ou rotulados de acordo com outras variáveis de interesse.

Ao interpretar um gráfico de RDA, você deve observar os seguintes aspectos:

1. Posicionamento das amostras: As amostras que estão próximas umas das outras no gráfico tendem a ser semelhantes em termos de suas respostas às variáveis explicativas. Amostras que estão distantes umas das outras indicam diferenças significativas nas respostas.
2. Direção e ângulo dos vetores: Os vetores representam as variáveis explicativas no espaço do gráfico. Eles indicam a direção e a força da relação entre cada variável explicativa e as variáveis resposta. Um vetor mais longo e mais alinhado com uma direção específica indica uma relação mais forte. Aproximadamente, o ângulo entre dois vetores reflete a correlação entre as duas variáveis explicativas.
3. Contribuição das variáveis explicativas: As variáveis explicativas que estão localizadas mais próximas das amostras no gráfico têm uma contribuição mais importante na explicação das variações nas variáveis resposta. Portanto, você pode identificar quais variáveis têm um papel mais relevante no modelo.
4. Grupos ou padrões: Padrões, agrupamentos ou estruturas identificáveis no gráfico podem revelar relações importantes entre as variáveis explicativas e a variável resposta. Por exemplo, grupos de amostras próximas podem indicar uma resposta semelhante a determinadas variáveis explicativas.

É importante lembrar que a interpretação de um gráfico de RDA deve ser feita considerando o contexto específico do estudo e o conhecimento sobre as variáveis envolvidas. Além disso, é recomendado realizar análises estatísticas complementares para confirmar e validar as interpretações obtidas a partir do gráfico de RDA.

8.1 Testando a significância dos resultados da RDA

```
1 RsquareAdj(rda)
2 anova.cca(rda) #significância global
3 anova.cca(rda, by="axis") #sign. dos eixos da RDA
4 anova.cca(rda, by="terms") #sign. das var. explanatórias
```

```
$r.squared
[1] 0.7626193
```

```
$adj.r.squared
[1] 0.6269732
```

```
Permutation test for rda under reduced model
Permutation: free
Number of permutations: 999
```

```
Model: rda(formula = hell ~ m.elev + m.river + m.stream + m.distsource + m.distmouth + m.maxx)
```

	Df	Variance	F	Pr(>F)
Model	8	0.48622	5.6221	0.001 ***
Residual	14	0.15135		

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Permutation test for rda under reduced model
```

```
Forward tests for axes
```

```
Permutation: free
Number of permutations: 999
```

```
Model: rda(formula = hell ~ m.elev + m.river + m.stream + m.distsource + m.distmouth + m.maxx)
```

	Df	Variance	F	Pr(>F)
RDA1	1	0.164893	15.2531	0.001 ***
RDA2	1	0.126253	11.6788	0.001 ***
RDA3	1	0.079562	7.3598	0.001 ***
RDA4	1	0.062259	5.7592	0.001 ***
RDA5	1	0.032952	3.0482	0.045 *
RDA6	1	0.010487	0.9701	0.931
RDA7	1	0.007194	0.6655	0.951
RDA8	1	0.002619	0.2423	0.994
Residual	14	0.151346		

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Permutation test for rda under reduced model
Terms added sequentially (first to last)
Permutation: free
Number of permutations: 999
```

```
Model: rda(formula = hell ~ m.elev + m.river + m.stream + m.distsource + m.distmouth + m.max
```

	Df	Variance	F	Pr(>F)	
m.elev	1	0.142240	13.1577	0.001	***
m.river	1	0.065920	6.0978	0.001	***
m.stream	1	0.039957	3.6962	0.002	**
m.distsource	1	0.105074	9.7197	0.001	***
m.distmouth	1	0.104716	9.6865	0.001	***
m.maxdepth	1	0.004786	0.4428	0.918	
m.habdepth	1	0.009555	0.8839	0.512	
m.width	1	0.013971	1.2923	0.241	
Residual	14	0.151346			

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

9 RDA passos importantes

9.1 Avaliando colinearidade

Colinearidade entre variáveis inflaciona sua importância conjunta. Função variance inflation factors. Como regra geral, **valores de colinearidade acima de 20 são considerados altos** ((James et al. 2013), pag.101).

```
1 #Função variance inflation factors
2 vif.cca(rda)
```

	m.elev	m.river	m.stream	m.distsource	m.distmouth	m.maxslope
	219.692198	298.967146	387.520711	797.919482	41.486429	NA
	m.maxdepth	m.habdepth	m.width			
	5.809314	6.512516	186.121218			

```
1 # remover valores com colinearidade maior que 20
```

9.2 Escolhendo as principais variáveis do modelo da RDA

Extract AIC from a Fitted Model Description Computes the (generalized) Akaike An Information Criterion for a fitted parametric model.

Entramos a nossa RDA como um objeto representando um modelo GLM[Um Modelo Linear Generalizado (GLM), é um framework estatístico que permite modelar relações entre variáveis dependentes e independentes, mesmo quando os dados não seguem uma distribuição normal. O GLM estende o modelo de regressão linear tradicional para lidar com diferentes tipos de respostas, incluindo variáveis binárias, contagens, proporções e outras distribuições não gaussianas. No GLM, a resposta é modelada em relação a uma combinação linear das variáveis independentes, usando uma função de ligação que relaciona a média da resposta à combinação linear das variáveis predictoras. Além disso, o GLM permite que a variância da resposta seja modelada em termos de funções de variância, que são específicas para cada tipo de distribuição.] que será usado como modelo inicial na Regressão Múltipla Stepwise (SMR).

```
1 # GLM
2 rda1 <- step(rda, scope = formula(rda), scale = 0, direction = "both", steps = 1000, test =
```

```
Start: AIC=-26.45
```

```
hell ~ m.elev + m.river + m.stream + m.distsource + m.distmouth +
      m.maxslope + m.maxdepth + m.habdepth + m.width
```

```
Step: AIC=-26.45
```

```
hell ~ m.elev + m.river + m.stream + m.distsource + m.distmouth +
      m.maxdepth + m.habdepth + m.width
```

	Df	AIC	F	Pr(>F)
- m.maxdepth	1	-27.913	0.3311	0.965
- m.habdepth	1	-26.965	0.9340	0.525
<none>		-26.451		
- m.width	1	-26.420	1.2923	0.280
- m.stream	1	-25.858	1.6709	0.110
- m.distsource	1	-25.525	1.8994	0.065 .
- m.distmouth	1	-24.765	2.4334	0.025 *
- m.river	1	-23.390	3.4460	0.010 **
- m.elev	1	-22.822	3.8815	0.005 **

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Step: AIC=-27.91
```

```
hell ~ m.elev + m.river + m.stream + m.distsource + m.distmouth +
```

m.habdepth + m.width

	Df	AIC	F	Pr(>F)
- m.habdepth	1	-28.430	0.9990	0.515
<none>		-27.913		
- m.width	1	-27.874	1.3909	0.225
- m.stream	1	-27.197	1.8801	0.065 .
- m.distsource	1	-26.597	2.3264	0.020 *
+ m.maxdepth	1	-26.451	0.3311	0.985
- m.distmouth	1	-24.504	3.9771	0.005 **
- m.elev	1	-23.626	4.7157	0.005 **
- m.river	1	-22.579	5.6335	0.005 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Step: AIC=-28.43

hell ~ m.elev + m.river + m.stream + m.distsource + m.distmouth +
m.width

	Df	AIC	F	Pr(>F)
- m.width	1	-28.506	1.3958	0.190
<none>		-28.430		
+ m.habdepth	1	-27.913	0.9990	0.415
+ m.maxdepth	1	-26.965	0.3530	0.980
- m.stream	1	-25.740	3.6190	0.005 **
- m.elev	1	-24.504	4.7024	0.005 **
- m.distsource	1	-24.342	4.8486	0.005 **
- m.river	1	-20.380	8.7678	0.005 **
- m.distmouth	1	-20.145	9.0224	0.005 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Step: AIC=-28.51

hell ~ m.elev + m.river + m.stream + m.distsource + m.distmouth

	Df	AIC	F	Pr(>F)
<none>		-28.506		
+ m.width	1	-28.430	1.3958	0.160
+ m.habdepth	1	-27.874	0.9799	0.455
+ m.maxdepth	1	-27.128	0.4379	0.925
- m.elev	1	-21.731	7.8974	0.005 **
- m.river	1	-21.088	8.6031	0.005 **
- m.distsource	1	-21.003	8.6976	0.005 **

```
- m.stream      1 -20.599 9.1536 0.005 **
- m.distmouth   1 -19.944 9.9086 0.005 **
```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```
1 summary(rda1)
```

Call:

```
rda(formula = hell ~ m.elev + m.river + m.stream + m.distsource + m.distmouth, data = m)
```

Partitioning of variance:

	Inertia	Proportion
Total	0.6376	1.0000
Constrained	0.4579	0.7182
Unconstrained	0.1797	0.2818

Eigenvalues, and their contribution to the variance

Importance of components:

	RDA1	RDA2	RDA3	RDA4	RDA5	PC1	PC2
Eigenvalue	0.1647	0.1218	0.0784	0.06136	0.03158	0.05057	0.03279
Proportion Explained	0.2584	0.1911	0.1230	0.09624	0.04953	0.07931	0.05143
Cumulative Proportion	0.2584	0.4495	0.5724	0.66868	0.71821	0.79753	0.84896
	PC3	PC4	PC5	PC6	PC7	PC8	
Eigenvalue	0.02103	0.01958	0.01563	0.01090	0.008561	0.006846	
Proportion Explained	0.03298	0.03071	0.02452	0.01709	0.013428	0.010738	
Cumulative Proportion	0.88193	0.91264	0.93716	0.95425	0.967677	0.978415	
	PC9	PC10	PC11	PC12	PC13	PC14	
Eigenvalue	0.004576	0.003095	0.001922	0.001733	0.001085	0.0006922	
Proportion Explained	0.007177	0.004854	0.003015	0.002717	0.001701	0.0010858	
Cumulative Proportion	0.985592	0.990446	0.993461	0.996179	0.997880	0.9989660	
	PC15	PC16	PC17				
Eigenvalue	0.0005869	6.746e-05	4.945e-06				
Proportion Explained	0.0009205	1.058e-04	7.756e-06				
Cumulative Proportion	0.9998864	1.000e+00	1.000e+00				

Accumulated constrained eigenvalues

Importance of components:

	RDA1	RDA2	RDA3	RDA4	RDA5
Eigenvalue	0.1647	0.1218	0.0784	0.06136	0.03158
Proportion Explained	0.3598	0.2660	0.1712	0.13400	0.06896
Cumulative Proportion	0.3598	0.6258	0.7970	0.93104	1.00000

Scaling 2 for species and site scores

* Species are scaled proportional to eigenvalues

* Sites are unscaled: weighted dispersion equal on all dimensions

* General scaling constant of scores: 1.935249

Species scores

	RDA1	RDA2	RDA3	RDA4	RDA5	PC1
ap-davis	-0.0033512	0.0035144	-0.0139013	-0.090906	-0.0663101	0.0368148
as-bimac	0.2532217	-0.2882728	0.1441978	-0.056717	0.0583303	0.4243202
as-fasci	0.0015022	0.0314901	-0.0935819	-0.107800	0.0405842	0.0263011
ch-bimac	0.2600316	-0.5332807	-0.2215477	0.111375	0.0541427	-0.0226344
ci-ocela	0.1444617	0.1771794	-0.0010665	0.129885	-0.1767754	-0.0563130
ci-orien	0.1484989	0.2101962	-0.0307628	0.095778	0.0061738	-0.0198549
co-macro	0.0058862	-0.0123529	-0.0049901	0.002657	0.0009629	0.0004574
co-heter	0.0018652	0.0043128	-0.0023131	-0.003146	0.0079909	0.0067873
cr-menez	0.0195201	0.0294230	-0.0228101	-0.025306	0.0689498	0.0223463
cu-lepid	-0.0019546	0.0020497	-0.0081080	-0.053021	-0.0386754	0.0257982
cy-gilbe	0.0088258	0.0403037	-0.0428435	-0.185455	-0.0539628	-0.2171832
ge-brasi	-0.3432171	-0.0944993	0.4716819	-0.060093	0.0125413	-0.0765878
he-margi	0.0037115	0.0085819	-0.0046028	-0.006260	0.0159009	-0.0103363
ho-malab	0.1598940	0.0905801	-0.0411976	0.075784	-0.0340882	0.0114178
hy-pusar	0.0585524	0.0980691	-0.0302738	-0.010290	0.0609033	0.0476471
le-melan	-0.0008542	0.0008958	-0.0035434	-0.023172	-0.0169021	0.0065835
le-piau	0.0565389	0.0601312	-0.0134085	0.027995	-0.0280772	-0.0220594
le-taeni	-0.0006040	0.0006334	-0.0025055	-0.016385	-0.0119516	0.0046553
mo-costa	-0.0004265	0.0004473	-0.0017693	-0.011570	-0.0084397	0.0056296
mo-lepid	0.0132743	0.0306935	-0.0164622	-0.022388	0.0568701	0.0452624
or-nilot	-0.3988053	0.1270259	-0.1679580	-0.121967	0.1148366	-0.0319211
pa-manag	-0.6059159	0.0132580	-0.2097879	0.219741	-0.0196834	-0.0060990
pimel-sp	0.0045688	0.0105641	-0.0056660	-0.007706	0.0195737	0.0166255
po-retic	-0.0159458	-0.0186255	0.2036677	-0.022912	0.0732801	0.0476898
po-vivip	0.2965955	0.4257332	0.1103178	0.173865	0.0884645	0.0277666
pr-brevi	0.0913350	0.0668588	-0.0807986	-0.109326	0.1110585	-0.0424792
ps-rhomb	0.0012874	0.0029768	-0.0015966	-0.002171	0.0055155	0.0011951
ps-genise	0.0012874	0.0029768	-0.0015966	-0.002171	0.0055155	0.0011951
se-heter	0.1269011	0.1191922	-0.1061640	-0.223708	-0.1033018	0.1259165
se-piaba	0.0153808	0.0355642	-0.0190746	-0.025941	0.0658947	0.0559696
se-spilo	0.0012874	0.0029768	-0.0015966	-0.002171	0.0055155	0.0011951
st-noton	0.0404846	0.0936102	-0.0502070	-0.068280	0.1734446	-0.0717140
sy-marmo	0.0102587	0.0123095	0.0004222	0.012366	-0.0100359	-0.0074642

te-chalc	-0.0085141	0.0089285	-0.0353174	-0.230955	-0.1684662	0.0834576
tr-signa	0.0381676	0.1024098	-0.0700306	-0.185740	0.0982160	-0.1189638

Site scores (weighted sums of species scores)

	RDA1	RDA2	RDA3	RDA4	RDA5	PC1
S-R-CT1	0.17794	0.17407	-0.11165	-0.34215	0.55066	0.59343
S-R-CP1	0.40442	0.23165	0.03283	0.02697	-0.19952	0.63387
S-A-TA1	0.36145	-0.63072	-0.28597	-0.10629	-0.01026	0.22225
S-R-CT2	0.28664	0.42814	-0.12956	-0.03117	0.62283	0.28838
S-R-CP2	0.41105	0.57920	-0.04165	0.62974	-0.49453	-0.29052
S-A-TA2	0.34232	-0.80942	-0.42968	0.31635	0.07837	-0.35035
S-R-CT3	0.21941	0.42589	-0.16656	-0.14433	0.82057	0.15138
S-R-CP3	0.41110	0.36579	0.05733	0.60720	-0.48404	0.09590
S-A-TA3	0.36066	-0.75017	-0.28063	0.23528	0.04592	0.02711
S-R-CT4	-0.03168	0.48020	-0.40119	-0.58252	0.80059	-1.03319
S-R-CP4	0.37055	0.73976	0.01721	0.66131	-0.38435	-0.43926
S-A-TA4	0.33122	-0.73861	-0.18688	0.18475	0.11428	0.10099
B-A-MU1	-0.38305	-0.18344	0.87161	-0.13187	-0.11434	-0.59580
B-A-GU1	-0.70824	-0.08820	0.02406	0.19258	-0.02897	0.18393
B-R-PC2	0.02355	-0.04081	-0.01622	-0.86189	-0.65533	0.23714
B-A-MU2	-0.01727	-0.22412	0.84152	-0.01623	0.13278	0.26472
B-A-GU2	-0.86422	0.05816	-0.26839	0.38495	-0.12580	-0.22247
B-R-PC3	-0.06162	0.02542	-0.13758	-0.81471	-0.71121	0.40611
B-A-MU3	0.04940	-0.21047	0.70126	-0.02505	0.24019	0.42788
B-A-GU3	-0.80399	0.06771	-0.39601	0.25211	0.03831	-0.05670
B-R-PC4	-0.05423	0.11219	-0.22910	-0.82732	-0.45991	-0.64325
B-A-MU4	-0.05907	-0.04333	0.98305	0.08215	0.20938	-0.09679
B-A-GU4	-0.76635	0.03110	-0.44779	0.31012	0.01436	0.09524

Site constraints (linear combinations of constraining variables)

	RDA1	RDA2	RDA3	RDA4	RDA5	PC1
S-R-CT1	0.16308	0.37708	-0.20224	-0.27504	0.69866	0.59343
S-R-CP1	0.39928	0.47910	0.01643	0.48131	-0.39061	0.63387
S-A-TA1	0.34891	-0.73223	-0.29579	0.15752	0.05708	0.22225
S-R-CT2	0.16308	0.37708	-0.20224	-0.27504	0.69866	0.28838
S-R-CP2	0.39928	0.47910	0.01643	0.48131	-0.39061	-0.29052
S-A-TA2	0.34891	-0.73223	-0.29579	0.15752	0.05708	-0.35035
S-R-CT3	0.16308	0.37708	-0.20224	-0.27504	0.69866	0.15138
S-R-CP3	0.39928	0.47910	0.01643	0.48131	-0.39061	0.09590

S-A-TA3	0.34891	-0.73223	-0.29579	0.15752	0.05708	0.02711
S-R-CT4	0.16308	0.37708	-0.20224	-0.27504	0.69866	-1.03319
S-R-CP4	0.39928	0.47910	0.01643	0.48131	-0.39061	-0.43926
S-A-TA4	0.34891	-0.73223	-0.29579	0.15752	0.05708	0.10099
B-A-MU1	-0.10250	-0.16534	0.84936	-0.02275	0.11700	-0.59580
B-A-GU1	-0.78570	0.01719	-0.27203	0.28494	-0.02552	0.18393
B-R-PC2	-0.03077	0.03227	-0.12763	-0.83464	-0.60882	0.23714
B-A-MU2	-0.10250	-0.16534	0.84936	-0.02275	0.11700	0.26472
B-A-GU2	-0.78570	0.01719	-0.27203	0.28494	-0.02552	-0.22247
B-R-PC3	-0.03077	0.03227	-0.12763	-0.83464	-0.60882	0.40611
B-A-MU3	-0.10250	-0.16534	0.84936	-0.02275	0.11700	0.42788
B-A-GU3	-0.78570	0.01719	-0.27203	0.28494	-0.02552	-0.05670
B-R-PC4	-0.03077	0.03227	-0.12763	-0.83464	-0.60882	-0.64325
B-A-MU4	-0.10250	-0.16534	0.84936	-0.02275	0.11700	-0.09679
B-A-GU4	-0.78570	0.01719	-0.27203	0.28494	-0.02552	0.09524

Biplot scores for constraining variables

	RDA1	RDA2	RDA3	RDA4	RDA5	PC1
m.elev	-0.837057	-0.28161	0.4638	-0.06528	-0.02497	0
m.river	-0.787075	-0.06706	0.3185	-0.49160	0.18145	0
m.stream	0.091081	0.21413	-0.3329	-0.21588	0.88792	0
m.distsource	0.119356	0.38915	-0.2797	-0.38522	0.77956	0
m.distmouth	-0.004124	-0.36176	0.2587	-0.82438	-0.35011	0

```
1 modelo_rda1 <- rda1$call
2 modelo_rda1
```

```
rda(formula = hell ~ m.elev + m.river + m.stream + m.distsource +
    m.distmouth, data = m_trns_a)
```

Aqui fizemos uma regressão múltipla com o comando `step()` e escolhemos ambas `forward selection` e `backward selection` (argumento `direction = "both"`) para a criação de um modelo de entrada das variáveis que melhor explicam a matriz comunitária. Outras opções são, `direction = "backward"` ou `direction = "forward"`.

Note que o modelo proposto pela RDA para as principais variáveis explanatórias da matriz de comunidades é dado por `modelo_rda1`

9.2.1 Variáveis sugeridas pelo modelo

Criamos uma nova RDA com apenas as variáveis explanatórias sugeridas pelo modelo.

```
1 #/results: hold
2
3 #(matriz de dados comunitária ~ variáveis da matriz ambiental)
4 rda2 <- rda(formula = hell ~ m.elev + m.river + m.stream + m.distsource + m.distmouth, data = m)
5 #(variáveis de resposta ~ variáveis explanatórias sugeridas pelo modelo stepwise regression)
6 summary(rda2)
```

Call:

```
rda(formula = hell ~ m.elev + m.river + m.stream + m.distsource + m.distmouth, data = m)
```

Partitioning of variance:

	Inertia	Proportion
Total	0.6376	1.0000
Constrained	0.4579	0.7182
Unconstrained	0.1797	0.2818

Eigenvalues, and their contribution to the variance

Importance of components:

	RDA1	RDA2	RDA3	RDA4	RDA5	PC1	PC2
Eigenvalue	0.1647	0.1218	0.0784	0.06136	0.03158	0.05057	0.03279
Proportion Explained	0.2584	0.1911	0.1230	0.09624	0.04953	0.07931	0.05143
Cumulative Proportion	0.2584	0.4495	0.5724	0.66868	0.71821	0.79753	0.84896
	PC3	PC4	PC5	PC6	PC7	PC8	
Eigenvalue	0.02103	0.01958	0.01563	0.01090	0.008561	0.006846	
Proportion Explained	0.03298	0.03071	0.02452	0.01709	0.013428	0.010738	
Cumulative Proportion	0.88193	0.91264	0.93716	0.95425	0.967677	0.978415	
	PC9	PC10	PC11	PC12	PC13	PC14	
Eigenvalue	0.004576	0.003095	0.001922	0.001733	0.001085	0.0006922	
Proportion Explained	0.007177	0.004854	0.003015	0.002717	0.001701	0.0010858	
Cumulative Proportion	0.985592	0.990446	0.993461	0.996179	0.997880	0.9989660	
	PC15	PC16	PC17				
Eigenvalue	0.0005869	6.746e-05	4.945e-06				
Proportion Explained	0.0009205	1.058e-04	7.756e-06				
Cumulative Proportion	0.9998864	1.000e+00	1.000e+00				

Accumulated constrained eigenvalues

Importance of components:

	RDA1	RDA2	RDA3	RDA4	RDA5
Eigenvalue	0.1647	0.1218	0.0784	0.06136	0.03158
Proportion Explained	0.3598	0.2660	0.1712	0.13400	0.06896
Cumulative Proportion	0.3598	0.6258	0.7970	0.93104	1.00000

Scaling 2 for species and site scores

* Species are scaled proportional to eigenvalues

* Sites are unscaled: weighted dispersion equal on all dimensions

* General scaling constant of scores: 1.935249

Species scores

	RDA1	RDA2	RDA3	RDA4	RDA5	PC1
ap-davis	-0.0033512	0.0035144	-0.0139013	-0.090906	-0.0663101	0.0368148
as-bimac	0.2532217	-0.2882728	0.1441978	-0.056717	0.0583303	0.4243202
as-fasci	0.0015022	0.0314901	-0.0935819	-0.107800	0.0405842	0.0263011
ch-bimac	0.2600316	-0.5332807	-0.2215477	0.111375	0.0541427	-0.0226344
ci-ocela	0.1444617	0.1771794	-0.0010665	0.129885	-0.1767754	-0.0563130
ci-orien	0.1484989	0.2101962	-0.0307628	0.095778	0.0061738	-0.0198549
co-macro	0.0058862	-0.0123529	-0.0049901	0.002657	0.0009629	0.0004574
co-heter	0.0018652	0.0043128	-0.0023131	-0.003146	0.0079909	0.0067873
cr-menez	0.0195201	0.0294230	-0.0228101	-0.025306	0.0689498	0.0223463
cu-lepid	-0.0019546	0.0020497	-0.0081080	-0.053021	-0.0386754	0.0257982
cy-gilbe	0.0088258	0.0403037	-0.0428435	-0.185455	-0.0539628	-0.2171832
ge-brasi	-0.3432171	-0.0944993	0.4716819	-0.060093	0.0125413	-0.0765878
he-margi	0.0037115	0.0085819	-0.0046028	-0.006260	0.0159009	-0.0103363
ho-malab	0.1598940	0.0905801	-0.0411976	0.075784	-0.0340882	0.0114178
hy-pusar	0.0585524	0.0980691	-0.0302738	-0.010290	0.0609033	0.0476471
le-melan	-0.0008542	0.0008958	-0.0035434	-0.023172	-0.0169021	0.0065835
le-piau	0.0565389	0.0601312	-0.0134085	0.027995	-0.0280772	-0.0220594
le-taeni	-0.0006040	0.0006334	-0.0025055	-0.016385	-0.0119516	0.0046553
mo-costa	-0.0004265	0.0004473	-0.0017693	-0.011570	-0.0084397	0.0056296
mo-lepid	0.0132743	0.0306935	-0.0164622	-0.022388	0.0568701	0.0452624
or-nilot	-0.3988053	0.1270259	-0.1679580	-0.121967	0.1148366	-0.0319211
pa-manag	-0.6059159	0.0132580	-0.2097879	0.219741	-0.0196834	-0.0060990
pimel-sp	0.0045688	0.0105641	-0.0056660	-0.007706	0.0195737	0.0166255
po-retic	-0.0159458	-0.0186255	0.2036677	-0.022912	0.0732801	0.0476898
po-vivip	0.2965955	0.4257332	0.1103178	0.173865	0.0884645	0.0277666
pr-brevi	0.0913350	0.0668588	-0.0807986	-0.109326	0.1110585	-0.0424792
ps-rhomb	0.0012874	0.0029768	-0.0015966	-0.002171	0.0055155	0.0011951
ps-genise	0.0012874	0.0029768	-0.0015966	-0.002171	0.0055155	0.0011951
se-heter	0.1269011	0.1191922	-0.1061640	-0.223708	-0.1033018	0.1259165

se-piaba	0.0153808	0.0355642	-0.0190746	-0.025941	0.0658947	0.0559696
se-spilo	0.0012874	0.0029768	-0.0015966	-0.002171	0.0055155	0.0011951
st-noton	0.0404846	0.0936102	-0.0502070	-0.068280	0.1734446	-0.0717140
sy-marmo	0.0102587	0.0123095	0.0004222	0.012366	-0.0100359	-0.0074642
te-chalc	-0.0085141	0.0089285	-0.0353174	-0.230955	-0.1684662	0.0834576
tr-signa	0.0381676	0.1024098	-0.0700306	-0.185740	0.0982160	-0.1189638

Site scores (weighted sums of species scores)

	RDA1	RDA2	RDA3	RDA4	RDA5	PC1
S-R-CT1	0.17794	0.17407	-0.11165	-0.34215	0.55066	0.59343
S-R-CP1	0.40442	0.23165	0.03283	0.02697	-0.19952	0.63387
S-A-TA1	0.36145	-0.63072	-0.28597	-0.10629	-0.01026	0.22225
S-R-CT2	0.28664	0.42814	-0.12956	-0.03117	0.62283	0.28838
S-R-CP2	0.41105	0.57920	-0.04165	0.62974	-0.49453	-0.29052
S-A-TA2	0.34232	-0.80942	-0.42968	0.31635	0.07837	-0.35035
S-R-CT3	0.21941	0.42589	-0.16656	-0.14433	0.82057	0.15138
S-R-CP3	0.41110	0.36579	0.05733	0.60720	-0.48404	0.09590
S-A-TA3	0.36066	-0.75017	-0.28063	0.23528	0.04592	0.02711
S-R-CT4	-0.03168	0.48020	-0.40119	-0.58252	0.80059	-1.03319
S-R-CP4	0.37055	0.73976	0.01721	0.66131	-0.38435	-0.43926
S-A-TA4	0.33122	-0.73861	-0.18688	0.18475	0.11428	0.10099
B-A-MU1	-0.38305	-0.18344	0.87161	-0.13187	-0.11434	-0.59580
B-A-GU1	-0.70824	-0.08820	0.02406	0.19258	-0.02897	0.18393
B-R-PC2	0.02355	-0.04081	-0.01622	-0.86189	-0.65533	0.23714
B-A-MU2	-0.01727	-0.22412	0.84152	-0.01623	0.13278	0.26472
B-A-GU2	-0.86422	0.05816	-0.26839	0.38495	-0.12580	-0.22247
B-R-PC3	-0.06162	0.02542	-0.13758	-0.81471	-0.71121	0.40611
B-A-MU3	0.04940	-0.21047	0.70126	-0.02505	0.24019	0.42788
B-A-GU3	-0.80399	0.06771	-0.39601	0.25211	0.03831	-0.05670
B-R-PC4	-0.05423	0.11219	-0.22910	-0.82732	-0.45991	-0.64325
B-A-MU4	-0.05907	-0.04333	0.98305	0.08215	0.20938	-0.09679
B-A-GU4	-0.76635	0.03110	-0.44779	0.31012	0.01436	0.09524

Site constraints (linear combinations of constraining variables)

	RDA1	RDA2	RDA3	RDA4	RDA5	PC1
S-R-CT1	0.16308	0.37708	-0.20224	-0.27504	0.69866	0.59343
S-R-CP1	0.39928	0.47910	0.01643	0.48131	-0.39061	0.63387
S-A-TA1	0.34891	-0.73223	-0.29579	0.15752	0.05708	0.22225
S-R-CT2	0.16308	0.37708	-0.20224	-0.27504	0.69866	0.28838

S-R-CP2	0.39928	0.47910	0.01643	0.48131	-0.39061	-0.29052
S-A-TA2	0.34891	-0.73223	-0.29579	0.15752	0.05708	-0.35035
S-R-CT3	0.16308	0.37708	-0.20224	-0.27504	0.69866	0.15138
S-R-CP3	0.39928	0.47910	0.01643	0.48131	-0.39061	0.09590
S-A-TA3	0.34891	-0.73223	-0.29579	0.15752	0.05708	0.02711
S-R-CT4	0.16308	0.37708	-0.20224	-0.27504	0.69866	-1.03319
S-R-CP4	0.39928	0.47910	0.01643	0.48131	-0.39061	-0.43926
S-A-TA4	0.34891	-0.73223	-0.29579	0.15752	0.05708	0.10099
B-A-MU1	-0.10250	-0.16534	0.84936	-0.02275	0.11700	-0.59580
B-A-GU1	-0.78570	0.01719	-0.27203	0.28494	-0.02552	0.18393
B-R-PC2	-0.03077	0.03227	-0.12763	-0.83464	-0.60882	0.23714
B-A-MU2	-0.10250	-0.16534	0.84936	-0.02275	0.11700	0.26472
B-A-GU2	-0.78570	0.01719	-0.27203	0.28494	-0.02552	-0.22247
B-R-PC3	-0.03077	0.03227	-0.12763	-0.83464	-0.60882	0.40611
B-A-MU3	-0.10250	-0.16534	0.84936	-0.02275	0.11700	0.42788
B-A-GU3	-0.78570	0.01719	-0.27203	0.28494	-0.02552	-0.05670
B-R-PC4	-0.03077	0.03227	-0.12763	-0.83464	-0.60882	-0.64325
B-A-MU4	-0.10250	-0.16534	0.84936	-0.02275	0.11700	-0.09679
B-A-GU4	-0.78570	0.01719	-0.27203	0.28494	-0.02552	0.09524

Biplot scores for constraining variables

	RDA1	RDA2	RDA3	RDA4	RDA5	PC1
m.elev	-0.837057	-0.28161	0.4638	-0.06528	-0.02497	0
m.river	-0.787075	-0.06706	0.3185	-0.49160	0.18145	0
m.stream	0.091081	0.21413	-0.3329	-0.21588	0.88792	0
m.distsource	0.119356	0.38915	-0.2797	-0.38522	0.77956	0
m.distmouth	-0.004124	-0.36176	0.2587	-0.82438	-0.35011	0

```
1 #str(rda2)
```

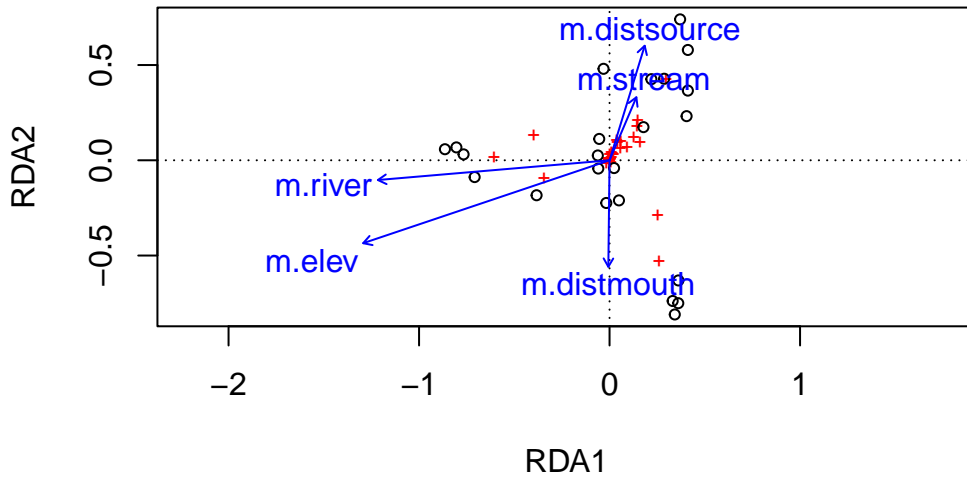
Avaliamos a colinearidade do novo modelo de RDA e testamos a significância dos novos resultados

```
1 #Colinearidade
2 vif.cca(rda2)
3 #ANOVAS
4 anova(rda2, perm.max = 999)
5 anova(rda2, by="axis", perm.max = 999)
6 anova(rda2, by="terms", perm.max = 999)
7 anova(rda2, by="margin", perm.max = 999)
```

```

8 summary(rda2)
9 #Triplot
10 simplified_model <- ordiplot(rda2, type = "points", scaling = 2)

```



```

      m.elev      m.river      m.stream m.distsource  m.distmouth
104.052842  108.865795   50.598866  125.910397    5.112757

```

Permutation test for rda under reduced model

Permutation: free

Number of permutations: 999

Model: rda(formula = hell ~ m.elev + m.river + m.stream + m.distsource + m.distmouth, data =

	Df	Variance	F	Pr(>F)	
Model	5	0.45791	8.6658	0.001	***
Residual	17	0.17966			

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Permutation test for rda under reduced model

Forward tests for axes

Permutation: free

Number of permutations: 999

Model: rda(formula = hell ~ m.elev + m.river + m.stream + m.distsource + m.distmouth, data =

	Df	Variance	F	Pr(>F)	
RDA1	1	0.164744	15.5888	0.001	***
RDA2	1	0.121819	11.5270	0.001	***
RDA3	1	0.078405	7.4190	0.001	***
RDA4	1	0.061361	5.8063	0.001	***
RDA5	1	0.031579	2.9881	0.002	**
Residual	17	0.179658			

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Permutation test for rda under reduced model

Terms added sequentially (first to last)

Permutation: free

Number of permutations: 999

Model: rda(formula = hell ~ m.elev + m.river + m.stream + m.distsource + m.distmouth, data =

	Df	Variance	F	Pr(>F)	
m.elev	1	0.142240	13.4594	0.001	***
m.river	1	0.065920	6.2376	0.001	***
m.stream	1	0.039957	3.7809	0.002	**
m.distsource	1	0.105074	9.9426	0.001	***
m.distmouth	1	0.104716	9.9086	0.001	***
Residual	17	0.179658			

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Permutation test for rda under reduced model

Marginal effects of terms

Permutation: free

Number of permutations: 999

Model: rda(formula = hell ~ m.elev + m.river + m.stream + m.distsource + m.distmouth, data =

	Df	Variance	F	Pr(>F)	
m.elev	1	0.083461	7.8974	0.001	***
m.river	1	0.090918	8.6031	0.001	***
m.stream	1	0.096736	9.1536	0.001	***
m.distsource	1	0.091917	8.6976	0.001	***
m.distmouth	1	0.104716	9.9086	0.001	***
Residual	17	0.179658			

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Call:

rda(formula = hell ~ m.elev + m.river + m.stream + m.distsource + m.distmouth, data = m

Partitioning of variance:

	Inertia	Proportion
Total	0.6376	1.0000
Constrained	0.4579	0.7182
Unconstrained	0.1797	0.2818

Eigenvalues, and their contribution to the variance

Importance of components:

	RDA1	RDA2	RDA3	RDA4	RDA5	PC1	PC2
Eigenvalue	0.1647	0.1218	0.0784	0.06136	0.03158	0.05057	0.03279
Proportion Explained	0.2584	0.1911	0.1230	0.09624	0.04953	0.07931	0.05143
Cumulative Proportion	0.2584	0.4495	0.5724	0.66868	0.71821	0.79753	0.84896
	PC3	PC4	PC5	PC6	PC7	PC8	
Eigenvalue	0.02103	0.01958	0.01563	0.01090	0.008561	0.006846	
Proportion Explained	0.03298	0.03071	0.02452	0.01709	0.013428	0.010738	
Cumulative Proportion	0.88193	0.91264	0.93716	0.95425	0.967677	0.978415	
	PC9	PC10	PC11	PC12	PC13	PC14	
Eigenvalue	0.004576	0.003095	0.001922	0.001733	0.001085	0.0006922	
Proportion Explained	0.007177	0.004854	0.003015	0.002717	0.001701	0.0010858	
Cumulative Proportion	0.985592	0.990446	0.993461	0.996179	0.997880	0.9989660	
	PC15	PC16	PC17				
Eigenvalue	0.0005869	6.746e-05	4.945e-06				
Proportion Explained	0.0009205	1.058e-04	7.756e-06				
Cumulative Proportion	0.9998864	1.000e+00	1.000e+00				

Accumulated constrained eigenvalues

Importance of components:

	RDA1	RDA2	RDA3	RDA4	RDA5
Eigenvalue	0.1647	0.1218	0.0784	0.06136	0.03158
Proportion Explained	0.3598	0.2660	0.1712	0.13400	0.06896
Cumulative Proportion	0.3598	0.6258	0.7970	0.93104	1.00000

Scaling 2 for species and site scores

- * Species are scaled proportional to eigenvalues
- * Sites are unscaled: weighted dispersion equal on all dimensions
- * General scaling constant of scores: 1.935249

Species scores

	RDA1	RDA2	RDA3	RDA4	RDA5	PC1
ap-davis	-0.0033512	0.0035144	-0.0139013	-0.090906	-0.0663101	0.0368148

as-bimac	0.2532217	-0.2882728	0.1441978	-0.056717	0.0583303	0.4243202
as-fasci	0.0015022	0.0314901	-0.0935819	-0.107800	0.0405842	0.0263011
ch-bimac	0.2600316	-0.5332807	-0.2215477	0.111375	0.0541427	-0.0226344
ci-ocela	0.1444617	0.1771794	-0.0010665	0.129885	-0.1767754	-0.0563130
ci-orien	0.1484989	0.2101962	-0.0307628	0.095778	0.0061738	-0.0198549
co-macro	0.0058862	-0.0123529	-0.0049901	0.002657	0.0009629	0.0004574
co-heter	0.0018652	0.0043128	-0.0023131	-0.003146	0.0079909	0.0067873
cr-menez	0.0195201	0.0294230	-0.0228101	-0.025306	0.0689498	0.0223463
cu-lepid	-0.0019546	0.0020497	-0.0081080	-0.053021	-0.0386754	0.0257982
cy-gilbe	0.0088258	0.0403037	-0.0428435	-0.185455	-0.0539628	-0.2171832
ge-brasi	-0.3432171	-0.0944993	0.4716819	-0.060093	0.0125413	-0.0765878
he-margi	0.0037115	0.0085819	-0.0046028	-0.006260	0.0159009	-0.0103363
ho-malab	0.1598940	0.0905801	-0.0411976	0.075784	-0.0340882	0.0114178
hy-pusar	0.0585524	0.0980691	-0.0302738	-0.010290	0.0609033	0.0476471
le-melan	-0.0008542	0.0008958	-0.0035434	-0.023172	-0.0169021	0.0065835
le-piau	0.0565389	0.0601312	-0.0134085	0.027995	-0.0280772	-0.0220594
le-taeni	-0.0006040	0.0006334	-0.0025055	-0.016385	-0.0119516	0.0046553
mo-costa	-0.0004265	0.0004473	-0.0017693	-0.011570	-0.0084397	0.0056296
mo-lepid	0.0132743	0.0306935	-0.0164622	-0.022388	0.0568701	0.0452624
or-nilot	-0.3988053	0.1270259	-0.1679580	-0.121967	0.1148366	-0.0319211
pa-manag	-0.6059159	0.0132580	-0.2097879	0.219741	-0.0196834	-0.0060990
pimel-sp	0.0045688	0.0105641	-0.0056660	-0.007706	0.0195737	0.0166255
po-retic	-0.0159458	-0.0186255	0.2036677	-0.022912	0.0732801	0.0476898
po-vivip	0.2965955	0.4257332	0.1103178	0.173865	0.0884645	0.0277666
pr-brevi	0.0913350	0.0668588	-0.0807986	-0.109326	0.1110585	-0.0424792
ps-rhomb	0.0012874	0.0029768	-0.0015966	-0.002171	0.0055155	0.0011951
ps-genise	0.0012874	0.0029768	-0.0015966	-0.002171	0.0055155	0.0011951
se-heter	0.1269011	0.1191922	-0.1061640	-0.223708	-0.1033018	0.1259165
se-piaba	0.0153808	0.0355642	-0.0190746	-0.025941	0.0658947	0.0559696
se-spilo	0.0012874	0.0029768	-0.0015966	-0.002171	0.0055155	0.0011951
st-noton	0.0404846	0.0936102	-0.0502070	-0.068280	0.1734446	-0.0717140
sy-marmo	0.0102587	0.0123095	0.0004222	0.012366	-0.0100359	-0.0074642
te-chalc	-0.0085141	0.0089285	-0.0353174	-0.230955	-0.1684662	0.0834576
tr-signa	0.0381676	0.1024098	-0.0700306	-0.185740	0.0982160	-0.1189638

Site scores (weighted sums of species scores)

	RDA1	RDA2	RDA3	RDA4	RDA5	PC1
S-R-CT1	0.17794	0.17407	-0.11165	-0.34215	0.55066	0.59343
S-R-CP1	0.40442	0.23165	0.03283	0.02697	-0.19952	0.63387
S-A-TA1	0.36145	-0.63072	-0.28597	-0.10629	-0.01026	0.22225
S-R-CT2	0.28664	0.42814	-0.12956	-0.03117	0.62283	0.28838

S-R-CP2	0.41105	0.57920	-0.04165	0.62974	-0.49453	-0.29052
S-A-TA2	0.34232	-0.80942	-0.42968	0.31635	0.07837	-0.35035
S-R-CT3	0.21941	0.42589	-0.16656	-0.14433	0.82057	0.15138
S-R-CP3	0.41110	0.36579	0.05733	0.60720	-0.48404	0.09590
S-A-TA3	0.36066	-0.75017	-0.28063	0.23528	0.04592	0.02711
S-R-CT4	-0.03168	0.48020	-0.40119	-0.58252	0.80059	-1.03319
S-R-CP4	0.37055	0.73976	0.01721	0.66131	-0.38435	-0.43926
S-A-TA4	0.33122	-0.73861	-0.18688	0.18475	0.11428	0.10099
B-A-MU1	-0.38305	-0.18344	0.87161	-0.13187	-0.11434	-0.59580
B-A-GU1	-0.70824	-0.08820	0.02406	0.19258	-0.02897	0.18393
B-R-PC2	0.02355	-0.04081	-0.01622	-0.86189	-0.65533	0.23714
B-A-MU2	-0.01727	-0.22412	0.84152	-0.01623	0.13278	0.26472
B-A-GU2	-0.86422	0.05816	-0.26839	0.38495	-0.12580	-0.22247
B-R-PC3	-0.06162	0.02542	-0.13758	-0.81471	-0.71121	0.40611
B-A-MU3	0.04940	-0.21047	0.70126	-0.02505	0.24019	0.42788
B-A-GU3	-0.80399	0.06771	-0.39601	0.25211	0.03831	-0.05670
B-R-PC4	-0.05423	0.11219	-0.22910	-0.82732	-0.45991	-0.64325
B-A-MU4	-0.05907	-0.04333	0.98305	0.08215	0.20938	-0.09679
B-A-GU4	-0.76635	0.03110	-0.44779	0.31012	0.01436	0.09524

Site constraints (linear combinations of constraining variables)

	RDA1	RDA2	RDA3	RDA4	RDA5	PC1
S-R-CT1	0.16308	0.37708	-0.20224	-0.27504	0.69866	0.59343
S-R-CP1	0.39928	0.47910	0.01643	0.48131	-0.39061	0.63387
S-A-TA1	0.34891	-0.73223	-0.29579	0.15752	0.05708	0.22225
S-R-CT2	0.16308	0.37708	-0.20224	-0.27504	0.69866	0.28838
S-R-CP2	0.39928	0.47910	0.01643	0.48131	-0.39061	-0.29052
S-A-TA2	0.34891	-0.73223	-0.29579	0.15752	0.05708	-0.35035
S-R-CT3	0.16308	0.37708	-0.20224	-0.27504	0.69866	0.15138
S-R-CP3	0.39928	0.47910	0.01643	0.48131	-0.39061	0.09590
S-A-TA3	0.34891	-0.73223	-0.29579	0.15752	0.05708	0.02711
S-R-CT4	0.16308	0.37708	-0.20224	-0.27504	0.69866	-1.03319
S-R-CP4	0.39928	0.47910	0.01643	0.48131	-0.39061	-0.43926
S-A-TA4	0.34891	-0.73223	-0.29579	0.15752	0.05708	0.10099
B-A-MU1	-0.10250	-0.16534	0.84936	-0.02275	0.11700	-0.59580
B-A-GU1	-0.78570	0.01719	-0.27203	0.28494	-0.02552	0.18393
B-R-PC2	-0.03077	0.03227	-0.12763	-0.83464	-0.60882	0.23714
B-A-MU2	-0.10250	-0.16534	0.84936	-0.02275	0.11700	0.26472
B-A-GU2	-0.78570	0.01719	-0.27203	0.28494	-0.02552	-0.22247
B-R-PC3	-0.03077	0.03227	-0.12763	-0.83464	-0.60882	0.40611
B-A-MU3	-0.10250	-0.16534	0.84936	-0.02275	0.11700	0.42788

B-A-GU3	-0.78570	0.01719	-0.27203	0.28494	-0.02552	-0.05670
B-R-PC4	-0.03077	0.03227	-0.12763	-0.83464	-0.60882	-0.64325
B-A-MU4	-0.10250	-0.16534	0.84936	-0.02275	0.11700	-0.09679
B-A-GU4	-0.78570	0.01719	-0.27203	0.28494	-0.02552	0.09524

Biplot scores for constraining variables

	RDA1	RDA2	RDA3	RDA4	RDA5	PC1
m.elev	-0.837057	-0.28161	0.4638	-0.06528	-0.02497	0
m.river	-0.787075	-0.06706	0.3185	-0.49160	0.18145	0
m.stream	0.091081	0.21413	-0.3329	-0.21588	0.88792	0
m.distsource	0.119356	0.38915	-0.2797	-0.38522	0.77956	0
m.distmouth	-0.004124	-0.36176	0.2587	-0.82438	-0.35011	0

Fizemos uma RDA. Pergunta... Existe uma alta colinearidade, acima de 20, entre algumas das variáveis explicativas do modelo sugerido pela SMR. É recomendável que essas variáveis sejam removidas.

10 Referências

- James, G., D. Witten, T. Hastie, e R. Tibshirani. 2013. An Introduction to Statistical Learning with Applications in R. Página 426. Book, Springer.
- Medeiros, E. S. F., M. J. Silva, e R. T. C. Ramos. 2008. Application of catchment- and local-scale variables for aquatic habitat characterization and assessment in the Brazilian semi-arid region. *Neotropical Biology and Conservation* 3:13–20.

Apêndices

Sites consultados

https://youtu.be/_bpQ16uUqs4
<https://youtu.be/O5HFxgl2WA4>
<https://youtu.be/wXFDIgaTdLw>
<https://youtu.be/YkpO2ESf0qo>