

Contents

Introduction	1
Software preparation	1
Data import	1
Data preparation	2
Species Accumulation Curve	2
Species-based SAC	2
Individual-based SAC	3
SAC-based Predictions	5
Rank-Abundance Curve	7

Introduction

This script reports the code to reproduce the analyses and figures presented in the Conservation Biology class at Stanford University (Bio 144, 2020).

Software preparation

Install R software (<https://www.r-project.org>) (R Core Team 2020) if you do not have it already. Then, install and load the following packages

```
install.packages('vegan')
install.packages('bipartite')
```

```
library(vegan)
library(bipartite)
```

Data import

Download data at the following link:

<https://polybox.ethz.ch/index.php/s/ngHzOumWVsk5omz/download>

Move them to your working folder, set it up as your working directory (e.g., `mydir`)

```
setwd("/mydir")
```

and import the .csv file into R

```
maindata <- read.csv("maindata.csv", sep=",", head=T)
```

We familiarize with the dataset and see how it looks like

```
str(maindata)
```

```
## 'data.frame':    330 obs. of  3 variables:
## $ Transect: int  10 3 4 6 6 8 8 7 6 5 ...
## $ Plant   : int  16 7 2 7 8 7 10 6 26 10 ...
## $ Species : Factor w/ 80 levels "Acalypha diversifolia",...: 75 1 1 1 1 1 1 2 3 4 ...
```

```
head(maindata)
```

```
##   Transect Plant      Species
## 1      10    16      Sapranthus
## 2       3     7 Acalypha diversifolia
## 3       4     2 Acalypha diversifolia
## 4       6     7 Acalypha diversifolia
## 5       6     8 Acalypha diversifolia
## 6       8     7 Acalypha diversifolia
```

It is a `dataframe` composed of 330 observations (rows) and three variables (columns): (1) Transect is the transect number, from one to ten; (2) Plant is the count of plants occurring along each transect; (3) Species is the taxonomic name of plants.

Data preparation

To produce the Species Accumulation Curve (SAC), we will make use of the existing function `specaccum` in `vegan` package. Take first a look at what this function does and requires.

```
?specaccum
```

This function needs data organized into a matrix with species in columns, transects in rows, and species abundance (i.e., number of trees per species) as entries. We then convert our `dataframe` into such a matrix.

```
maindata$dummy <- 1
maindata$Transect <- as.character(maindata$Transect)
data_matrix <- frame2webs(maindata[,c(1,3,4)],
                          c('Transect', 'Species', 'dummy'))[[1]]
# head(data_matrix)
```

Species Accumulation Curve

We now draw our SACs taking the whole dataset. We will do so by running 10,000 permutations (all possible permutations of 10 transects are $n = 10!/(10-10)! = 3628800$) and extrapolating richness values by means of bootstrap.

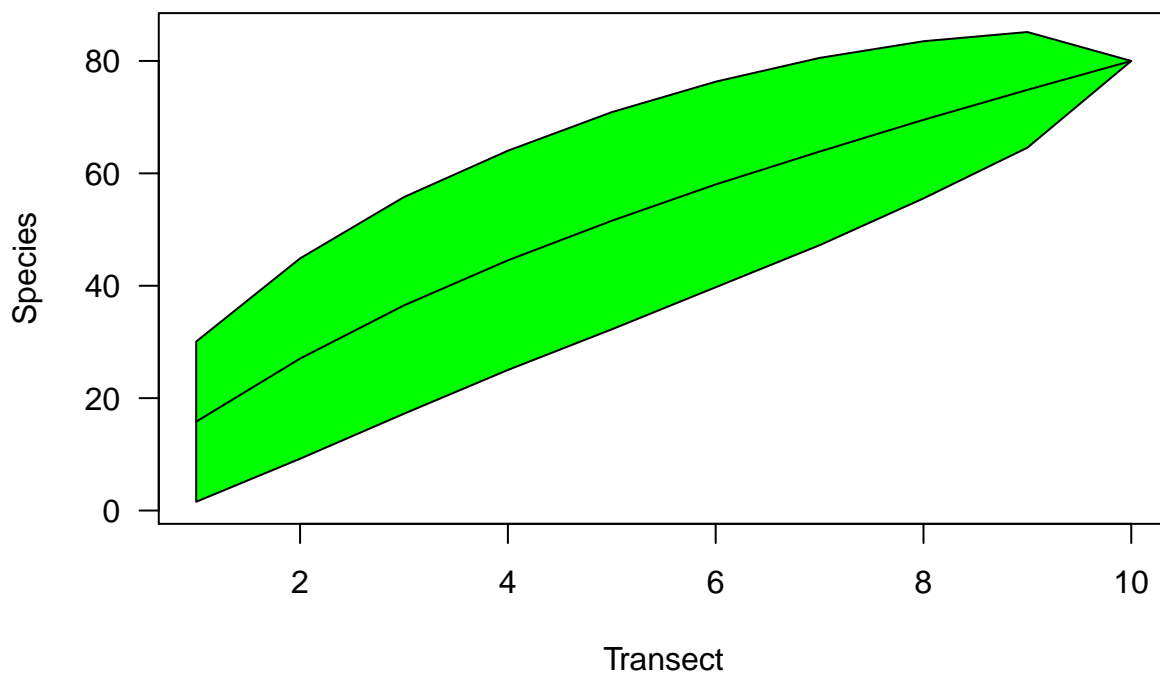
Species-based SAC

In the first case, SAC is based on adding transects (in `vegan` also called ‘sites’) in a random order and finding the mean SAC and its standard deviation from random permutations of the data (Gotelli & Colwell 2001).

```
sac <- specaccum(data_matrix,
                 method = 'random', permutations = 10000, gamma = 'boot')
sac
```

```
## Species Accumulation Curve
## Accumulation method: random, with 10000 permutations
## Call: specaccum(comm = data_matrix, method = "random", permutations = 10000,
##
##
## Sites      1.00000  2.00000  3.0000  4.00000  5.00000  6.00000  7.00000  8.00000
## Richness 15.81100 27.04020 36.5039 44.51990 51.56600 58.02260 63.87560 69.52290
## sd       7.12709  8.91256  9.6378  9.75409  9.66131  9.14421  8.33306  6.99103
##
## Sites      9.00000 10
## Richness 74.85700 80
## sd       5.14412  0
```

```
plot(sac, ci.type='polygon', ci.col='green', las = 1,
     xlab = "Transect", ylab = "Species")
```



From the shape of the curve we can deduce our goodness of sampling. Since the curve does not reach an asymptote we did more data to have a better estimation of tree species richness. Indeed, there is a certain number of unseen, missing species.

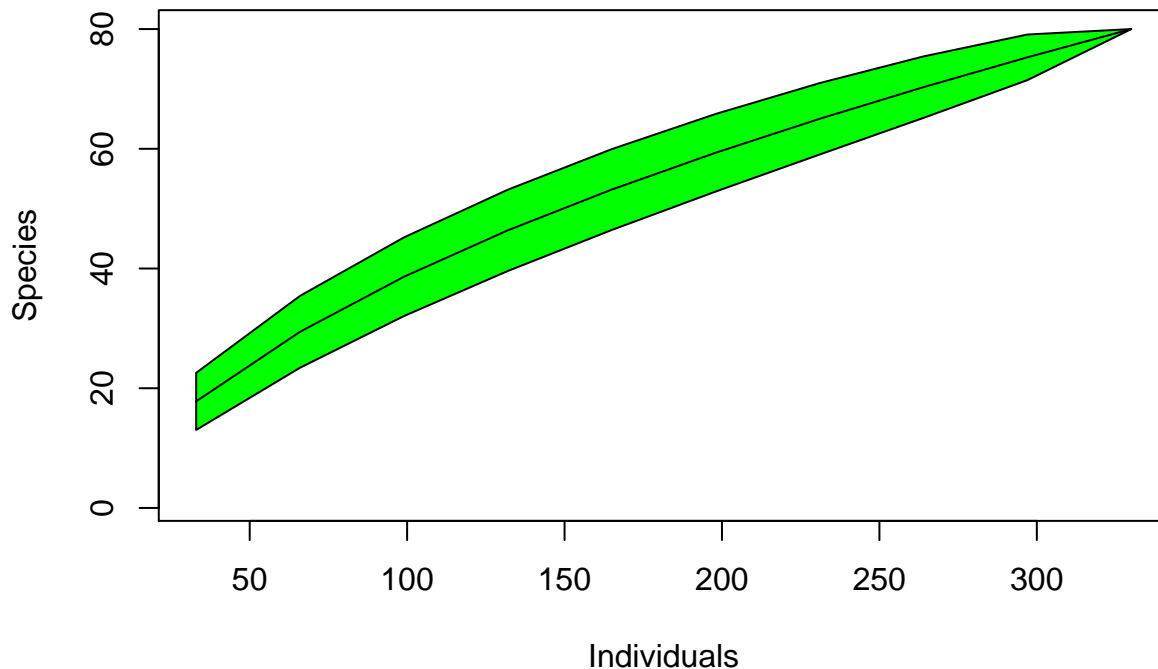
Individual-based SAC

In the second case, we will draw a SAC according to a rarefaction method that is based on sampling individuals (i.e., trees) rather than transects (Hurlbert 1971).

```
sac <- specaccum(data_matrix,
                 method = 'rarefaction', permutations = 10000, gamma = 'boot')
sac
```

```
## Species Accumulation Curve
## Accumulation method: rarefaction
## Call: specaccum(comm = data_matrix, method = "rarefaction", permutations = 10000,
##
##
## Sites      1.0000  2.0000  3.0000  4.0000  5.0000  6.0000  7.000  8.0000
## Individuals 33.0000 66.0000 99.0000 132.0000 165.0000 198.0000 231.000 264.0000
## Richness    17.8053 29.4306 38.6249 46.3796 53.1888 59.3384 65.004 70.2952
## sd          2.3845  3.0002  3.2845  3.3941  3.3753  3.2400  2.981  2.5673
##
## Sites      9.0000 10
## Individuals 297.0000 330
## Richness    75.2801 80
## sd          1.9051  0
```

```
plot(sac, ci.type='polygon', ci.col='green', xvar = "individuals",
     xlab = "Individuals", ylab = "Species")
```



We obtain a similar curve to the previous one. We can see how the number of observed species increases by increasing the number of transects but it does not ‘saturate’, indicating that species richness should continue to increase with sampling more than 10 transects.

We estimate (`?specpool`) the number of missing species being around 99 ± 10 , which corresponds to c 25% of observed species pool (i.e., 80 species).

```
specpool(data_matrix)[,7:8]
```

```
##          boot  boot.se  
## All 99.17644 10.48784
```

Furthermore, if we sample less than 10 transects, the estimated species number is reduced too.

```
poolaccum(data_matrix)$means[,c(2,6)]
```

```
##          S Bootstrap  
## [1,] 36.89 45.55926  
## [2,] 44.59 55.30988  
## [3,] 51.18 63.48284  
## [4,] 57.78 71.56546  
## [5,] 64.67 80.11972  
## [6,] 69.07 85.42296  
## [7,] 74.33 91.96027  
## [8,] 80.00 99.17644
```

With 3 transects we would have estimated 46 species (out of 37 observed), with 4 57, and so on, until 99 estimated species out of 80 observed over 10 transects.

SAC-based Predictions

Therefore, how many species can we observe by increasing the number of transects? How many new transects are needed to sample the estimated 99 species?

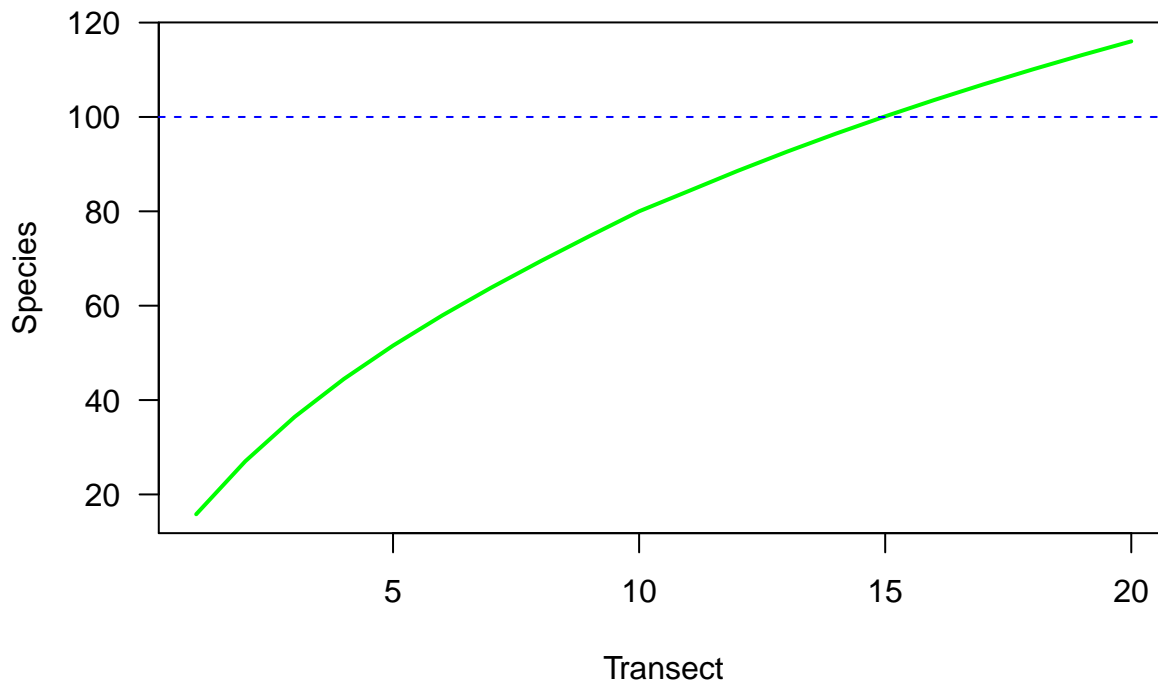
```
sac <- specaccum(data_matrix, method = 'exact')
```

```
## Warning in cor(x > 0): the standard deviation is zero
```

```
sac.fit <- fitspecaccum(sac, 'lomolino')  
sac.pred <- predict(sac.fit, 11:20); sac.pred
```

```
## [1] 84.26879 88.56517 92.62555 96.47302 100.12729 103.60535 106.92201  
## [8] 110.09020 113.12137 116.02565
```

```
plot(c(sac.fit$richness, sac.pred), type = 'l', lwd = 2, las = 1,  
     col = 'green', xlab = 'Transect', ylab = 'Species')  
abline(h = 100, lty = 2, col = 'blue')
```



By sampling one more transect we observe 84 species, 89 with two more transects, 93 with three, 96 with four, and 100 with five. Thus, with fifteen total transects we may have a quite good sample of our forest.

Still, we notice that the curve does not saturate yet. Indeed, the derivate of the SAC at 20 transects is still much higher than zero.

```
specslope(sac.fit, 20)
```

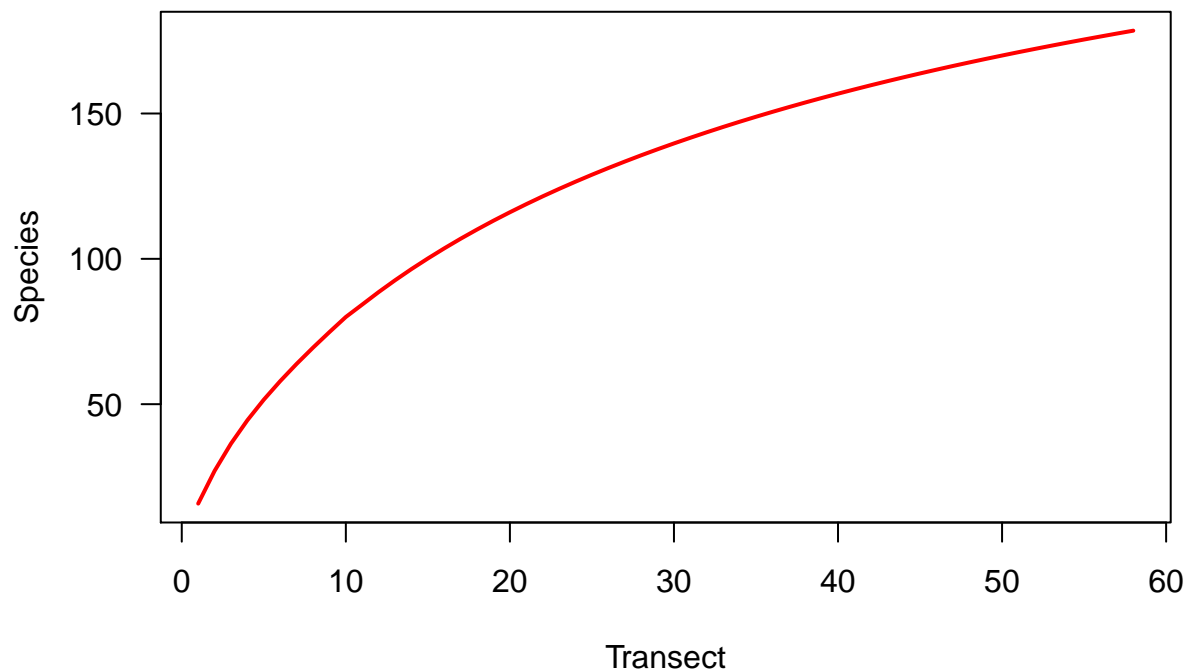
```
## [1] 2.843932
```

We estimate that the curve will start saturating (i.e., increasing rate smaller than one) with 48 additional transects.

```
specslope(sac.fit, 21:60)
```

```
## [1] 2.7302940 2.6243435 2.5253289 2.4325955 2.3455701 2.2637486 2.1866853
## [8] 2.1139849 2.0452949 1.9803004 1.9187187 1.8602954 1.8048010 1.7520277
## [15] 1.7017868 1.6539068 1.6082312 1.5646170 1.5229334 1.4830602 1.4448870
## [22] 1.4083120 1.3732415 1.3395887 1.3072735 1.2762214 1.2463634 1.2176355
## [29] 1.1899778 1.1633348 1.1376546 1.1128886 1.0889916 1.0659211 1.0436373
## [36] 1.0221030 1.0012831 0.9811446 0.9616566 0.9427899
```

```
sac.pred <- predict(sac.fit, 11:58)
plot(c(sac.fit$richness,sac.pred), type = 'l', lwd = 2, las = 1,
     col = 'red', xlab = 'Transect', ylab = 'Species')
```



```
sac.pred[48]
```

```
## [1] 178.4879
```

We notice how the number of newly discovered species decreases with further increasing the number of transects. With a total of 58 transects we estimate a forest richness of 178 species.

Rank-Abundance Curve