

# Doing biodiversity analysis with Biodiversity.R

## Doing biodiversity analysis with Biodiversity.R

This chapter describes how the analyses presented in this manual can be performed with the Biodiversity.R software.

This special attention to Biodiversity.R does not mean that other software packages can not be used for biodiversity analysis. In 2003, a Biodiversity Analysis Package CD-ROM was produced to provide several software packages that are very good for biodiversity analysis. However, some of the software packages had a more limited scope in the analyses that they supported. Some of the software packages had not developed a graphical user interface, which caused some problems for teaching the analysis. Some types of analysis could not be performed in the software provided on the CD-ROM. For these reasons, the Biodiversity.R software was developed, including a graphical user interface. All the analyses described in this manual can be conducted with Biodiversity.R.

## What is Biodiversity.R?

Biodiversity.R is software that does all the biodiversity analyses described in this manual. It needs to be loaded into the R statistical software. R is a software that was developed to allow for many different types of statistical analysis. It is very similar to the S and S-Plus statistical software. It is free software, as is Biodiversity.R. The software is also open, so that you can check how calculations are done, and the graphics are quite advanced.

## How do I run Biodiversity.R?

The CD-ROM that is provided with this manual contains an installed version of Biodiversity.R. You can run Biodiversity.R from the CD-ROM by clicking on the file *Run-Biodiversity.bat* in the Biodiversity.R folder on the CD-ROM. Alternatively you need to install Biodiversity.R on another location on your computer first.

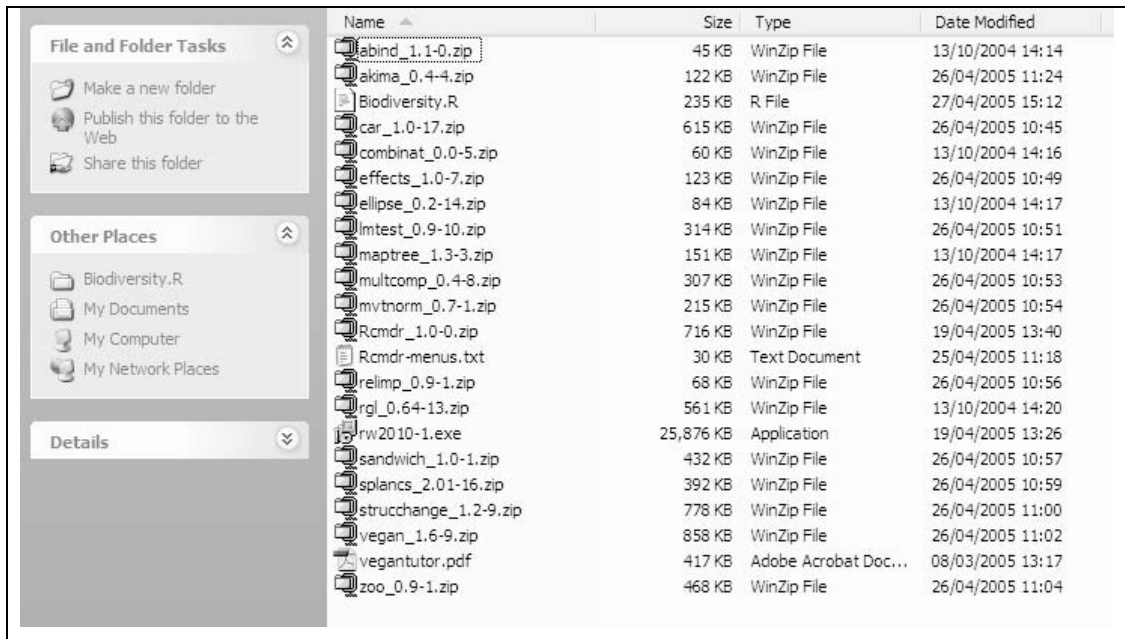
Each time that you run R, you need to load the *Rcmdr* package (see below: how do I run Biodiversity.R) to access the Biodiversity.R graphical user interface.

## How do I install Biodiversity.R?

Two options are provided for installing Biodiversity.R. When you click on *Install-Biodiversity.bat*, then all files will be installed under the *Program Files* folder of your C drive.

The alternative method is to install Biodiversity.R in steps. The files that are used during installation are all listed on the CD-ROM in the *Biodiversity.R/installation files* folder (Figure 3.1, see next page).

The first thing that you need to do is to install the R software. You can also obtain the software from the following website: <http://cran.r-project.org>. If you are using Windows, then you can download R from: <http://cran.r-project.org/bin/windows/base>. You will find a file there with a name similar to **rw2010-1.exe**. You need to run the file to install R. Note that it may be safer to close all other programs when you install R.



**Figure 3.1.** All the files used during installation are provided in the *Installation files* folder

During the installation, make sure that you opt to install the support files for `library(tcltk)`.

After you have installed R, then you can run it. R is a language that is run by typing in commands. It does not have an extensive user interface.

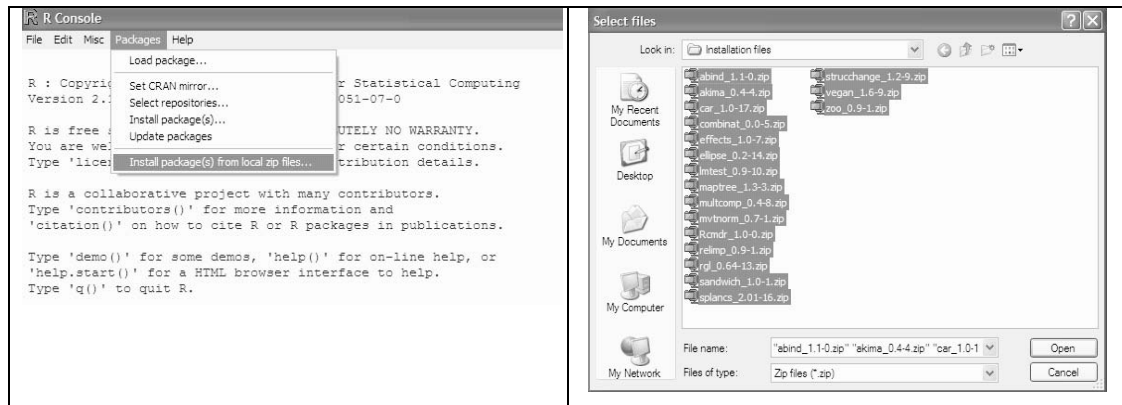
Some additional software can be loaded onto R. These addins are called packages or libraries. Some packages already come with the installation version of R. Some other packages need to be downloaded.

You have various options to install these additional packages. The first option is install the additional packages from the CD-ROM. Within

the R program, you need to go to the top menu, choose Packages, and then choose menu option: Install packages from local zip files.... Figure 3.2 shows how this can be done.

Note that we will not put many images of menus in this manual, so that some space will be saved. We will describe the selection shown in Figure 3.2 as: “You select the R menu option: Packages > Install package(s) from local zip files...”.

You will obtain a list of available packages under the *Installation files* folder of the CD-ROM. You can install all the packages by selecting all of them at the same time (click on the CTRL and A keys simultaneously).



**Figure 3.2.** Installing other packages onto R. This menu option is described in the text as: “Packages > Install package(s) from local zip files...”.

The alternatives are to download the packages from R and then install them manually afterwards, or to install the packages directly from the CRAN website. You can download the files for the different packages from the same website you downloaded the R package itself from: <http://cran.r-project.org>. You will see a link to “contributed packages”. If you are using Windows, then you can download the packages from: <http://cran.r-project.org/bin/windows/contrib>.

You only need to install a package once after installing R. When you run R, you will not be able to access the functions of contributing packages, unless you follow the menu option of: Packages > Load package..., or if you give the appropriate command in R of loading the package. For the `vegan` library, you need to select this package after menu option: Packages > Load package... or you need to type `library(vegan)` each time that you started using R and wanted to access the functions of this library.

Note that we will use a different font for results and commands in R and Biodiversity.R. For example, `library(vegan)` points to a command in R. Any information in this font will appear in the R console.

The packages that you need to install into R are:

- abind
- akima
- car
- combinat
- effects
- ellipse
- lme4
- lme4tree
- multcomp
- mvtnorm
- rcmdr
- relimp
- rgl
- sandwich
- splancs
- strucchange
- vegan
- zoo

These are quite a few packages, but this also means that various types of analysis can be done in R. Especially important is `vegan` as it allows

for many of the analyses of Biodiversity.R. `Rcmdr` is also an important package as it allows for the R-commander graphical interface. (For the advanced user: The R-commander runs best under Single-Document Interface. You can set this option by setting “MDI = no” in file *C:\Program Files\R\rw2010\etc\Rconsole.*)

The final (and probably most complicated) step in installing Biodiversity.R is to copy two files into the *library\Rcmdr\etc* folder of R. In case that you installed R under the program files of your C drive, then you can find the *Rcmdr* directory under *C:\Program files\R\rw2010\library\Rcmdr\etc*. You need to put the following files into this directory: *Biodiversity.R* and *Rcmdr-menus.txt*. The *Rcmdr-menus.txt* will already exist in the library, so you need to **replace** the *Rcmdr-menus.txt* with the file that is provided on the CD-ROM. You could use a program such as the Windows Explorer to copy the files.

When you will have completed all these steps, Biodiversity.R will have been installed.

In short, you need to follow these steps to install and run Biodiversity.R:

1. Install R
2. Install the required contributing packages
3. Copy *Biodiversity.R* and *Rcmdr-menus.txt* into *library\Rcmdr\etc*

## How do I run Biodiversity.R?

Obviously you need to run R first.

To run Biodiversity.R with the menu options, you will need to load the R-commander, either by the command `library(Rcmdr)` or by menu option: Packages > Load package...

**Each time that you want to use Biodiversity.R, you need to load the `Rcmdr` package again after launching R.**

## The species and environmental datasets

After you have loaded Biodiversity.R, you will not be able to conduct any type of analysis before you have chosen the species and environmental datasets. As we saw in chapter 2 on data preparation, the analyses in this manual will be conducted with these two datasets.

The R-commander was designed to only use one dataset. This dataset is called the active dataset. Biodiversity.R uses two datasets, the species and environmental datasets. When developing it the decision was taken to make the environmental dataset of Biodiversity.R always to be the active dataset of the R-commander. When a new dataset is chosen to be the new active dataset for the R-commander, then it also becomes the new environmental dataset of Biodiversity.R. With the menu options of the R-commander, various manipulations can be done on the active dataset, including importing an active dataset and saving an active dataset. You can do the same manipulations to the species dataset, but first you need to set the species dataset to also be the active dataset (and thus also the environmental dataset).

## An example of an analysis in Biodiversity.R

Descriptions of the way in which each type of analysis can be done in R are provided at the end of each chapter. As an example and a test whether you installed Biodiversity.R correctly, you could run following analysis. This example calculates a species accumulation curve for the dune dataset. This dataset is provided with `vegan`. The result that you will obtain is shown in the chapter on species richness.

You need to follow these menu options:

Biodiversity > Environmental Matrix > Select environmental matrix

→Select the dune.env dataset

Biodiversity > Community Matrix > Select community matrix

→Select the dune dataset

Biodiversity > Analysis of biodiversity > Species accumulation curves

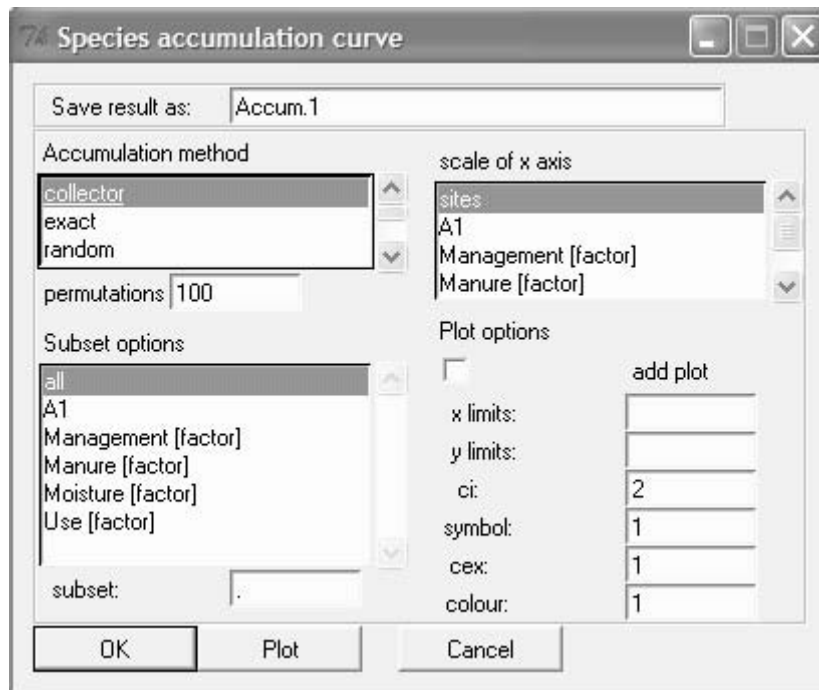
→Accumulation method: exact

→Press on the “OK” button

→Press on the “Plot” button

The last window that allows you to calculate a species accumulation curve is shown in Figure 3.3.

Note that there are separate “OK” and “Plot” options for different menus. When you press “OK”, then the analysis is calculated and the results are saved under name the specified for the result. The “Plot” options apply to the name of the result that appears on top. It is not necessary to redo the analysis to plot results, and you can also plot information from earlier analyses by changing the name of the result on top.



**Figure 3.3.** The window interface for the calculation of species accumulation curves

The alternative option to using the menu interface is to type in the following commands:

```
Accum.1 <- accumresult(dune,
  method='exact')
Accum.1
plot(Accum.1)
```

**Note that R is case sensitive:** `Plot(accum.1)` will not work!

If you chose the first option, then you will see that the same commands were generated by the R-commander. The R-commander is actually an interface that allows you to find appropriate commands that can be used in R. If you are very familiar with the commands, then you could perform some analyses more quickly by directly typing them in R.

You can see that R commands consist of functions (such as `accumresult` or `specaccum`), followed by arguments that are put between brackets. The command `specaccum(dune)` will calculate a species accumulation for the dune dataset, for instance.

## Importing files into Biodiversity.R

Files of various formats can be imported into R. We usually prefer to use text files (files with an extension of `.txt`), where the columns are separated by tabs. You can create such files from Excel by choosing the menu option in Excel: File > Save as... and then putting “text (tab delimited)”.

Be aware that text fields (also names of variables and sites) can only contain one continuous text without spaces. If you have words separated by spaces (such as the genus and species of a plant name), then you need to change these into a continuous text such as `changed_into_continuous_text`, `changedintocontinuous_text` or `changed.into.continuous.text`.

Datasets in R have one particular feature in not having a name for the column that contains the row names of the data. So if your first column of data determines the name of each site, then you should not have any contents in the first row of data with the column names. As a result, the number of fields in the first row of the file will be one less than the other rows of the file. You can check for this feature by looking at the files `dune.txt` and `dune.env.txt`, provided under the *Data* folder on the CD-ROM, or look at figures 3.4 and 3.5. If the first row has the same number of columns, then R will conclude that you have not chosen to give row names to your data.

	Achmil	Agrsto	Airpra	Alogen	Antodo	Belper	Brarut	Brohor	Calcus	Chealb
X1	1	0	0	0	0	0	0	0	0	0
X2	3	0	0	2	0	3	0	4	0	0
X3	0	4	0	7	0	2	2	0	0	0
X4	0	8	0	2	0	2	2	3	0	0
X5	2	0	0	0	4	2	2	2	0	0

**Figure 3.4.** Part of the `dune.txt` file with the species data for the dune meadow dataset. Note that the column with the names of the sites has no name.

	A1	Moisture	Management	Use	Manure
X1	2.8	1	SF	Haypastu	4
X2	3.5	1	BF	Haypastu	2
X3	4.3	2	SF	Haypastu	4
X4	4.2	2	SF	Haypastu	4
X5	6.3	1	HF	Hayfield	2

**Figure 3.5.** Part of the `dune.env.txt` file with the environmental data for the dune meadow dataset. Note that the column with the names of the sites has no name.

Once you have created the file, then you can import it by menu option: Data > Import data > From text file... and then browse to the right place of your file. You can look at your data by clicking on the “View Dataset” button below the top menu of the R-commander.

### Learning more about the commands of R and the contributed packages

As we saw above, R is statistical software that is driven by functions. You may therefore increase the utility of this software by learning more about its functions, especially finding out options, calculation methods and related functions.

There are various options of learning about the functions.

One option is to type in the name of the function by R menu option: Help > R functions (text)... and then typing in “`specaccum`”, for example.

The second option is to go to R menu option: Help > Html help. It may especially be useful to check then under the Packages and the Search Engine.

The third option is to type a ? before the function in R. For example `?specaccum`

The fourth option is to just type the function in R, for example `specaccum`. In that case, you will obtain the actual programming code that was used for the function.

Note that some new functions were created for Biodiversity.R. Documentation for these functions is not provided through the R console, but you can access these files from the CD-ROM that comes with this manual under the *Biodiversity.R* folder (click on Biodiversity.R help.html)

## Citing R

When you use R for data analysis, you need to cite using the base package and the additional packages in the publications where you report the results of the data analysis. Information on citation is provided by the `citation` function:

```
> citation()
To cite R in publications use:
  R Development Core Team (2005). R: A language and environment for
  statistical computing.
  R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-
  900051-07-0,
  URL http://www.R-project.org.

> citation("vegan")
To cite package 'vegan' in publications use:
  Oksanen, J., Kindt, R. & O'Hara, R.B. (2005). vegan: Community
  Ecology Package
  version 1.6-9. http://cc.oulu.fi/~jarioksa/.
```