

## Instructions for report

Choose *one* of the following data sets for your report. They are provided on the Web page of the course, <http://biol09.biol.umontreal.ca/Panama16/>, in the “Practicals in R” folder.

Data set 1 – *Gault\_forest\_reserve*. The Gault Nature Reserve vegetation data file contains all variables in a single file. The Gault Reserve is located south-east of Montréal, Canada.

1. Read the Gault data file "Gault\_Forest\_Community.txt"

```
gault = read.table( file.choose() )
```

2. Copy groups of variables into different files:

```
gault.LonLat = gault[,3:4]           # The geographic LonLat coordinate data
```

```
gault.spe = gault[,5:18]           # The species data
```

```
gault.env = gault[,19:31]         # The environmental variables
```

3. Transform the Longitude-Latitude data into flat Cartesian coordinates. Function `geoXY()` requires the coordinates to be in the order Latitude-Longitude. Required package: SoDA

```
gault.XY = geoXY(gault.LonLat[,2], gault.LonLat[,1])
```

Data set 2 – *Legendre-Gauthier practicals*. Chesapeake Bay invertebrate data through space (27 sites) and time (13 years of sampling, twice a year). They are provided on the Web page of the course, <http://biol09.biol.umontreal.ca/Panama16/>, in the “Practicals in R” folder. Instructions for analysis are found in the document “Appendix S2, R code.pdf”. Choose one research question among the many questions that can be studied with these data.

### **For the one data set you have chosen,**

1. Briefly describe the objective of the research, the sampling design, and the available data, based on the available information. Transform the data as you see fit and explain your choices.
2. Describe the multivariate structure of the community composition data using PCA or CA. Justify your methodological choices (standardization or not, scaling type 1 or 2).
3. Study the relationship between the community composition data and the environmental variables using RDA. Justify your methodological choices. Report the  $R^2$  and adjusted  $R^2$ , global test of significance, and tests of significance for the various axes. Interpret the relationships displayed by the ordination biplot or triplot.
4. Spatial structure of the community: carry out a dbMEM analysis of the community data. Use variation partitioning to describe how much of the community variation is explained by (1) a broad-scale, medium-scale, and fine-scale MEM model, or (2) by the selected dbMEM and environmental variables.

=====