Appendix C:

Overview of available software

The use of any multivariate statistical method even for small datasets requires a computer program to perform the analysis. Most of the known statistical methods are implemented in several statistical packages. In this book, we demonstrate how to use the ordination methods with possibly the most widely employed package, Canoco for Windows. We also show how to use the methods not available in CANOCO (clustering, NMDS, ANOVA) with the general package Statistica for Windows. In the next paragraph, we provide information on obtaining a trial version of the CANOCO program, which you can use to work through the tutorials provided in this book, using the sample datasets (see Appendix A for information on how to obtain the datasets). We also provide an overview of other available software in tabular form and show both the freely available as well as the commercial software. The attention is focused on the specialized software packages, targeting ecologists (or biologists), so we do not cover general statistical packages such as S-Plus, SAS or GENSTAT.

The Canoco for Windows program is commercial software requiring a valid licence for its use. But we reached agreement with its distributor (Microcomputer Power, Ithaca, NY, USA), who will provide you on request with a trial version of the software, which will be functional for a **minimum** of one month. You can use it to try the sample analyses discussed in this book, using the data and CANOCO projects provided on our web site (see Appendix A). To contact Dr Richard Furnas from Microcomputer Power, write to the following E-mail address: *trial@microcomputerpower.com*.

Table 20-1 lists the main available packages, which can be used for some or many parts of the multivariate analysis of ecological data, and their functionality is compared. We do not own all the listed programs, so the information is often based on the data excerpted from their web pages. For most of them, the authors or distributors reviewed and corrected the provided information.

Table 20-1. *Overview of the functionality, relevant to the book contents, of the most often used commercial and freely available software*

	ADE4	CAP	Canoco	DECODA	ECOM	MVSP	NTSys-pc	PATN	PCORD	Primer	R	SYN-TAX	Vegan / R	ViSTA
Distribution[1]	F	C	C	C	C	C	C	C	C	C	F	C	F	F
Mac[2]	yes	no	no	no	no	no	no	no	no	no	yes	yes	yes	yes
transf[3]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	no[9]	?
stand[3]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
PCA	yes	yes	yes	no	no	yes	yes	yes	yes	yes	yes	yes	no[9]	yes
RDA	yes	no	yes	no	yes	no	no	no	no	no	yes	yes	no	yes
CA	yes	yes	yes	yes	no	yes	yes	yes	yes	no	yes	yes	yes	yes
DCA	no	yes	yes	yes	no	yes	no	yes	yes	no	no	no	yes	no
CCA	yes	no	yes	no	yes	yes	no	no	yes	no	yes	yes	yes	no
perm. tests[4]	yes	no	yes	no	yes	no	no	yes	yes	no	yes	no	yes	no
partial ord[5]	yes	no	yes	no	no	no	no	no	no	no	yes	no	yes	no
CVA	yes	no	yes	no	no	no	yes	no	no	no	yes	yes	no	no
PCoA	yes	no	yes	?	no	yes	yes	yes	no	no	yes	yes	no[9]	yes
NMDS	no	yes	no	yes	no	no	yes	yes	yes	yes	no	yes	yes	yes
Mantel test	yes	no	no	?	no	no	yes	yes	yes	yes	yes	no	yes	no
clustering[6]	yes	yes	no	no	no	yes	yes	yes	yes	yes	yes	yes	no[9]	yes
coefficients[7]	many	few	several	several	no	many	many	many	many	few	many	many	several	many
TWINSPAN	no	yes	no	yes	no	no	no	yes	yes	no	no	no	no	no
regression[8]	yes	no	yes	no?	yes	no	yes	yes	no	no	yes	no	no[9]	yes

[1] **F** – the software is provided for free, **C** – commercial software;

[2] **yes** means the program runs natively on **Mac**Intosh platforms: where **no** is shown, the program can probably be executed using one of the commercial emulator programs available from third parties; all the programs run on Microsoft Windows platform;

[3] **transf** – basic transformations (such as log or sqrt) are provided, **stand** – standardizations by data columns and rows are available;

[4] permutation tests refer here only to permutation testing of multivariate hypotheses in the framework of RDA/CCA ordination methods, as discussed in this book (but see also the Mantel test further down in the table);

[5] **yes** means that the program uses the concept of covariates (covariables);

[6] here we refer to hierarchical agglomerative **clustering**, many packages also implement the other methods;

[7] does the package provide calculation of distance (dissimilarity) or similarity coefficients in the appropriate context (such as PCoA, NMDS, clustering);

[8] multiple linear regression;

[9] this functionality is provided by the base R system, see Vegan details in the list at the end of this Appendix.

The lists of features are actual as of July 2002 and might not be fully correct. The table is followed by footnotes explaining the meaning of individual rows. Then we provide individual paragraphs for each of the listed programs, with contact information for the distributing companies or individuals and also additional comments, when needed. Five additional programs (CLUSTAN, DistPCoA, NPMANOVA, PolynomialRdaCca, and RdaCca) are listed only there, as their functionality is rather specialized. Question marks in Table 20-1 imply that we do not know about this particular aspect of program functionality.

There now follows the list of contact information for the individual producers of the software listed in Table 20-1. The information provided was accurate at the time of printing.

- **ADE-4**, is available from server of Lyon University <http://pbil.univ-lyon1. fr/ADE-4/>
- **CANOCO for Windows**, version 4.5, is distributed by Microcomputer Power, USA, <http://www.microcomputerpower.com> and by Scientia, Hungary <http://ramet.elte.hu/~scientia/>
- **CAP**, version 2.0, is distributed by Pisces Conservation Ltd., UK <http://www.pisces-conservation.com/>
- **DECODA**, version 2.05, is distributed by Anutech Pty. Ltd., Australia <http://www.anutech.com.au/TD/DECODA_WWW/welcome.html>
- **ECOM**, version 1.33, is distributed by Pisces Conservation Ltd., UK <http://www.pisces-conservation.com/>
- **MVSP**, version 3, is distributed by Kovach Computing Services, UK <http://www.kovcomp.co.uk/mvsp/index.html>
- **NTSYSpc**, version 2.1, is distributed by Exeter Software, USA <http://www.exetersoftware.com/cat/ntsyspc.html>
- **PATN** is distributed by CSIRO, Australia <http://www.cse.csiro.au/CDG/PATN/>
- **PCORD**, version 4, is distributed by MjM Software, USA <http://www.pcord.com>
- **PRIMER**, version 5, is distributed by Primer-E Ltd., UK <http://www.primer-e.com>
- The **R package**, version 4, is available from Philippe Casgrain web site <http://www.fas.umontreal.ca/BIOL/casgrain/en/labo/R/v4/>. Please, note that this software's name can be confused with the more general statistical package R, which is a non-commercial version of the S / S-Plus system (see also the Vegan package below).
- **SYN-TAX**, version 2000, is available from Scientia, Hungary <http://ramet.elte.hu/~scientia/>

- **Vegan**, version 1.5.2, is an add-on package for the R statistical system (<http://cran.r-project.org/>).
  Vegan package is available from Jari Oksanen,
  <http://cc.oulu.fi/~jarioksa/softhelp/vegan.html>
- **ViSTA**, version 6.4, is available from Prof. Forrest Young web site
  <http://forrest.psych.unc.edu/research/>
- **CLUSTAN**, version 5.0, is a package specialized in clustering and it provides almost every known algorithm of hierarchical agglomerative clustering (in addition to other methods such as K-means clustering) and calculation of many coefficients of (dis)similarity. It is available from the Clustan company, UK <http://www.clustan.com>
- **DistPCoA** is a simple program calculating principal coordinates analysis (PCoA) with optional correction for negative eigenvalues; it is also able to calculate the starting matrix of distances, using one of several available distance measures. It is available from P. Legendre's web site, at
  <http://www.fas.umontreal.ca/biol/casgrain/en/labo/distpcoa.html>
- **NPMANOVA** is a program written by M.J. Anderson, and it calculates a non-parametric multivariate ANOVA, based on a selected kind of distance (dissimilarity) measure. It supports the two-way balanced ANOVA designs only. It is available, together with other programs, at
  <http://www.stat.auckland.ac.nz/~mja/Programs.htm>
- **PolynomialRdaCca** is a simple program calculating both polynomial and linear versions of redundancy analysis or canonical correspondence analysis. It is available from P. Legendre's web site, at
  <http://www.fas.umontreal.ca/biol/casgrain/en/labo/plrdacca.html>
- **RdaCca** is a simple program calculating redundancy analysis or canonical correspondence analysis. It is available from P. Legendre's web site, at
  <http://www.fas.umontreal.ca/biol/casgrain/en/labo/rdacca.html>