Computer program review

Making more sense of the order: A review of Canoco for Windows 4.5, PC-ORD version 4 and SYN-TAX 2000

Gilliam, Frank S.* & Saunders, N. Elizabeth

Department of Biological Sciences, Marshall University, Huntington, WV 25755-2510, USA;
*Corresponding author: gilliam@marshall.edu

Abstract. Personal computers of ever-increasing speed have motivated programmers of multivariate software to adapt their programs to be run in Microsoft Windows and Macintosh platforms. Updated versions of these multivariate programs appear more and more frequently and are marketed intensively. In this review we provide a comparative analysis of the most recent versions of three analytical software packages – *Canoco for Windows 4.5*, *PC-ORD version 4* and *SYN-TAX 2000*. The three packages share two characteristics. First, the most recent versions are now compatible with the most recent Windows platforms and should therefore be accessible for use by virtually all vegetation scientists. Second, they have capabilities for numerous multivariate techniques, although each package has some unique techniques. Thus, any one of the packages will have much to offer the user.

Keywords: Analytical software; Classification; Multivariate analysis; Ordination.

Abbreviations: BC = Bray-Curtis ordination; CA (also, CoA) = Correspondence Analysis; CCA (also, CCoA) = Canonical Correspondence Analysis; CVA = Canonical Variate Analysis, also called Discriminant Analysis; DCA = Detrended Correspondence Analysis; DCCA = Detrended Canonical Correspondence Analysis; MRPP = Multi-response Permutation Procedure; NMDS (also NMS) = Non-metric Multidimensional Scaling; PCA = Principal Components Analysis; PCoA = Principal Coordinates Analysis; RDA = Redundancy Analysis; TWINSPAN = Two-way Indicator Species Analysis.

Introduction

Vegetation analysis, not unlike quantitative techniques in other disciplines, has undergone a temporal scrutiny somewhat analogous to natural selection. The early forms and techniques were novel and useful at their time of development, only to become sequentially replaced through time by later forms that contained newer and more useful characteristics. The rate of such change has been mitigated to varying degrees by the practitioners themselves (i.e. vegetation scientists), driven by a combination of true dispute with new changes along with what might be called 'research inertia' (Gilliam & Roberts 2003).

When Peet et al. (1988) attempted to "put things in order", they were responding to criticisms leveled at DCA by Wartenberg et al. (1987), who originated the phrase "putting things in order." Peet et al. (1988) did much to further the wide acceptance of DCA among vegetation scientists. Taking his cue from the now-famous catch phrase, Palmer (1993) followed by "putting things in even better order" when he demonstrated the advantages of CCA. Since that time, CCA has risen greatly in popularity.

Among the practical limitations of multivariate analysis has been the computational difficulty of manipulating data sets which are often large. The earliest techniques were developed prior to the widespread use of mainframe computers and, indeed, required laborious calculations by hand. Techniques initially designed for computers used compilers and associated languages, such as FORTRAN, on the large mainframe computers usually found on college campuses and at research centers. With the development of personal computers (PCs), multivariate analysis programs were quickly adapted for use with MS-DOS. Finally, the appearance of Microsoft® Windows-run PCs (and to a lesser extent Macintosh computers) of ever-increasing speed, along with the many updated versions of Windows, has motivated programmers of multivariate software to adapt their programs to be run in a Windows/Mac platform. Thus, a covariate with improvements in multivariate techniques in recent years has been the micro-processing capabilities of computers owned by vegetation scientists. Because of this, updated versions of multivariate software seem to appear more and more frequently and are marketed intensively.

The purpose of this review is to provide a comparative analysis of the most recent versions of three analytical software packages: *Canoco for Windows 4.5* by C.J.F. ter Braak and P. Šmilauer, *PC-ORD version 4* by B. McCune and M.J. Mefford and *SYN-TAX 2000* by J. Podani. We will discuss similarities and differences among these three packages. All three packages share two characteristics. First, the most recent versions are now compatible with the most recent Windows platforms. Thus, they should be accessible and available for use by virtually all vegetation scientists. Second, they have capabilities for numerous multivariate techniques and procedures. Accordingly, any one of the packages will have much to

offer the user. In this review, we will focus on differences between the packages themselves, rather than the choice of appropriate methods. We would like to draw readers' attention to a recently published comparative review of these same software packages, including an earlier version of SYN-TAX (SYN-TAX 5.1), that provides additional perspective to similarities and differences among them (Rejmánek & Klinger 2003).

Organization of review

This review comprises three main sections: (1) software specifications, including product histories and analytical capabilities, (2) trials of software systems on data and (3) a subjective, comparative assessment. The section on trials is presented in two parts. The first part is based on sample data supplied by the respective systems themselves. The second part of the section of trials is based on existing data taken from a long-term study at an old-growth longleaf pine stand in North Carolina. The final section, comprising a discussion of user-friendliness and cost effectiveness of the three packages, will be our subjective assessment of the results of the respective analyses and our comparisons among them. We conclude with recommendations regarding which package(s) we would deem worthy of use by vegetation scientists.

Software specifications

Canoco for Windows 4.5

History of product

The first version of CANOCO was developed in 1985 on a VAX computer as an extension of the computer program DECORANA (Hill & Gauch 1980) to include CCA and its detrended form, DCA (ter Braak 1987a). The second version, released in 1987, also included PCA and its canonical form, which is also known as RDA or reduced-rank regression (ter Braak 1988). Eventually, a more user-friendly interface to CANOCO was developed to be used on PC's with MS-DOS and an additional program – *CanoDraw* – was developed for creating ordination figures. A third version was released in 1990 that added ordination diagnostics and forward selection; facilities for permutation tests also were greatly extended.

CANOCO versions beginning with 4.0 were adapted for use with a Microsoft® Windows™ interface, including Windows 98, ME, Windows NT, 2000 and XP. A console version of CANOCO 4.x is still available for other operating systems. Permutation tests were extended in versions 4.x to include tests for split-plot and related multi-level designs; scaling of species scores also was improved. *Canoco for Windows 4.5* is bundled with *CanoDraw* for Windows to replace *CanoDraw 3.1* and *CanoPost*, two programs used in version 4.0. Other programs also included in the bundle are (1) the Windows utility *WCanoImp*, which can import data from spreadsheets, (2) *CanoMerge*, which can combine numerous data tables and *PrCoord*, which allows easy access to the distance-based RDA method for users of CANOCO. The complete package is called *Canoco for Windows 4.5* (ter Braak & Šmilauer 2002).

Analytical capabilities

Canoco for Windows 4.5 performs constrained and unconstrained ordination in ecological applications, integrating ordination with regression and permutation methodology and allowing statistical modeling of ecological data. *Canoco for Windows 4.5* contains both linear and unimodal methods, including PCA, PCoA, RDA, CA, DCA, CCA and DCCA (Table 1). In addition, other ordination methods are available in *Canoco for Windows 4.5* as special cases of these six primary methods. These other methods are log-ratio analysis, CVA (discriminant analysis) and various forms of multiple regression.

Ordination diagrams can be displayed on screen immediately after an ordination has been calculated. The program *CanoDraw* takes output from any of the ordination methods of CANOCO. CanoDraw produces a wide range of graphics, including scatter plots, biplots, triplots and attribute plots (wherein the size of the symbol represents the relative measure of some attribute, e.g., importance value of a chosen species). *Canoco for Windows 4.5* can account for background variation specified by covariables (i.e., partial ordinations) and contains facilities for permutation tests, including tests of interaction effects. In addition, it has the ability to do partial analysis and hypothesis testing. *Canoco for Windows 4.5* is not designed do any form of classification or non-metric ordinations.

PC-ORD version 4

History of product

PC-ORD first appeared in October 1983 under the name *@ORD* and was designed to be run on a VAX VMS operating system. It was designed to be used for data transformation and several types of ordination, including polar (Bray-Curtis), DCA, CA, PCA and PCoA. The next version was released in July 1986 under the current name, *PC-ORD*. In addition to several changes such as adding graphics mode ordination and dropping PCoA, this version was designed to run on MS-DOS. Successive versions of PC-ORD, all for MS-DOS, appeared in 1987, 1990 and 1993. These improvements increased the MS-DOS matrix size from 100×100 (rows × columns) to mostly 250×250, added such analyses as cluster analysis, TWIN-SPAN, MRPP and NMS and increased graphics capabilities.

When version 2 was released in August 1995, PC-ORD was no longer available as shareware, as had been the case previously. Improvements in this version primarily included the addition of CCA and greatly enhanced matrix size capabilities (essentially determined by available memory). PC-ORD version 3 was released in July 1997 to be run in a 16-bit Windows environment, adding such new components as species lists, further-improved graphics, indicator species analysis, a Mantel test to test for no relationship between dissimilarity or similarity matrices, as well as more robust algorithms for DCA and CCA ordination and TWINSPAN classification.

PC-ORD version 4 was released in August 1999 (McCune & Mefford 1999). Improvements included expansion of the program to be run in a 32-bit Windows environment, precluding its use on computers with Windows 3.x and earlier. In contrast to the flexible matrix size for versions 2 and 3, matrix size for data sets in version 4 was fixed at 32000×32000

Table 1. Comparisons among three analytical software packages.

Category	Canoco for Windows 4.5	PC-ORD version 4	SYN-TAX 2000
Hardware and operating system specifications	Microsoft Windows 98, ME, NT, 2000, and XP	Microsoft Windows 95, 98, 2000, ME, NT, and XP	Microsoft Windows 95, 98, 2000, NT, and XP
Available media for distribution of software	CD-ROM	CD-ROM, Internet	CD-ROM
Documentation and help	500-page user's manual. More help available in the program. Further information and theory in ter Braak (1987b)	237-page user's manual. More help available in the program. Further information and theory in McCune & Grace (2002)	53-page user's manual. More help available in the program. Further information and theory in Podani (2000)
URL of supplier	www.microcomputerpower.com	www.pcord.com	www.exetersoftware.com/
License agreements and costs	New: retail—$678 additional—$428 educational—$378 additional—$238	New: retail—$299 additional—$100 educational—$199	New: retail—$450 educational—$350
	Upgrade: retail—$429 additional—$279 educational—$249 additional—$169	Upgrade:retail—$99 additional—$50 NOTE: educational price is for students only; not institutions.	Upgrade: retail—$300 educational—$200
Ordination methods	PCA, RDA, CA, DCA, CCA, DCCA, PCoA. Includes partial ordinations and statistical tests for direct ordinations	BC, CCA, DCA, NMS (= NMDS), PCA, reciprocal averaging (= CA), weighted averaging. Includes statistical tests for direct ordinations.	PCA, PCoA, CA, CCoA, RDA, NMDS
Classification methods		Cluster analysis, TWINSPAN,	Hierarchical
Other methods	Log-ratio analysis, CVA, multiple regression and species response modeling using GAM/GLM	MRPP, blocked MRPP, Mantel test, species richness, species-area curves	Canonical Correlation Analysis, CVA (these methods can be considered ordination methods)

matrices, conforming to Lotus spreadsheets (i.e., *.wk1 format). New analyses included 'autopilot' and predictive modes for NMS, species ordinations for NMS, BC, PCA, blocked MRPP and the ability to produce species lists broken down by groups.

Analytical capabilities

PC-ORD version 4 is capable of performing both ordination and classification. Ordination procedures include BC, CCA, DCA, NMS, PCA, CA and weighted averaging (Table 1). Permutation tests for CCA are available, but partial ordinations are not. Classification methods include cluster analysis, TWINSPAN and indicator species analysis. In addition, *PC-ORD version 4* can perform MRPP, blocked MRPP and Mantel tests and determine species richness and species-area curves.

Depending on the type of ordination or classification method run in *PC-ORD version 4*, a variety of graphics options is available. In the case of ordinations, options include scatter plots, biplots, triplots and overlays (which are similar to the attribute plots of CanoDraw described previously). Several types of overlays can be used, including varying symbol sizes, labels, vectors, grids and joint plots. Groups within data can be coded by colors or symbol types. Graphs produced in PC-ORD version 4 can be printed, saved to file, or pasted into other applications.

SYN-TAX 2000

History of product

SYN-TAX 2000 is the culmination of a programming project running continuously since 1980. The first version was for mainframes and only performed hierarchical clustering (Podani 1980). SYN-TAX II appeared in 1984 and included ordination. The SYN-TAX package was first available on PCs as SYN-TAX III in 1988 using MS-DOS. SYN-TAX IV was released in 1991, still using MS-DOS and was upgraded to version 5.0 in 1995. The most recent upgrade prior to SYN-TAX 2000 was version 5.1 in 1997 (Podani 1997) and was made available to both PC and Macintosh computers and added neighbor joining, canonical correspondence analysis and ordinal clustering techniques.

Analytical capabilities

SYN-TAX 2000 contains programs for data analysis with emphasis on applications in ecology and biosystematics (Table 1). Most of these programs are even applicable to fields outside ecology and systematics when reduction of dimensionality and exploration of data structure are desired. The SYN-TAX 2000 package is actually a bundle of three separate software 'modules', which appear as separate icons (shortcuts) on the user's desktop (or other selected locations on the computer).

Modules

Module HierClus: This module performs hierarchical clustering, typically represented as a dendrogram. Methods included in this module are the following (as stated on the SYN-TAX web site – see below for URL): h-SAHN and d-SAHN methods, information theory clustering based on entropy and mutual information, divisive monothetic clustering, ordinal clustering, neighbor joining, global optimization, minimum spanning trees. Module HierClus uses 33 dissimilarity coefficients, six run-time standardization procedures to produce dendrograms, minimum-spanning trees, additive trees, unrooted trees.

Module Ordin: This module performs numerous metric ordination techniques, including PCA, PCoA, CA, CCoA, RDA, Canonical Correlation Analysis and Canonical Variates (Discriminant) Analysis, as well as NMDS. Graphical output from this module includes ordination scatter plots, biplots, triplots, rotating plots, minimum spanning trees and classification polygons superimposed on ordinations, as well as Scree plots and Shepard diagrams. As with Module HierClus, the PCoA and NMDS methods Module Ordin use of 33 dissimilarity coefficients and six run-time standardization procedures.

Module NonHier: This module performs several non-hierarchical clustering methods, wherein samples (objects) are partitioned into mutually exclusive groups (clusters). These methods include *k*-means clustering, multiple partitioning, quick clustering, global optimization, fuzzy *c*-means clustering and ordinal clustering. Ternary plots and 'scattergrams' can be produced to illustrate fuzzy classifications graphically. The global optimization method uses 33 dissimilarity coefficients and six run-time standardization procedures.

Trials

Sample data

The three software packages were used by the second author (NES) as part of an independent study in the Department of Biological Sciences at Marshall University to learn a variety of ordination techniques. This experience is included in our review to simulate what might be encountered by first-time users of these packages. Having no other background in multivariate statistics than taking the first author's (FSG) class in plant ecology, wherein there is a brief introduction to BC, PCA, DCA, NMDS and CCA as compared in a plant ecology textbook (Barbour et al. 1999), the second author (NES) comes to this exercise not unlike a beginning graduate student.

Installation

Installation was self-explanatory for all three programs. Full applications were easily installed properly, even without reading directions. PC-ORD was installed from two floppy disks (although it is currently distributed on CD-ROM), whereas the others were installed from CD-ROM.

Canoco for Windows 4.5

The basic tutorial required no more than ca. 40 min. This included repeating sections of the tutorial until concepts behind the analyses were more completely understood. The tutorial provides accurate descriptions of screen images and processes. No difficulty was encountered while following the process of converting spreadsheet data to a CANOCO data file, accessing this data file and running DCA, CCA and RDA on the data. The project wizards were easy to use and logically organized, although the flexibility offered in the wizard is likely lost on the beginning graduate student. One example was given to access CanoDraw to produce a graph. CanoDraw is easily accessed from Canoco for Windows 4.5, but the example created a graph that was so label-intensive that the vectors were not immediately recognizable. We should add that, although it is not part of the tutorial to do so, it is possible to correct this problem with relative ease.

The help functions in the tutorial were generally complete and accurate. Running the mouse across an unknown element provided an expanded description of its function. Chapter 3 of the manual provided details of the mathematical models used in various analyses. Chapter 4 provided comprehensive information on how to construct data tables to import. Chapter 5 provided roughly the same general instructions for software use without the examples. The text seemed dense, with two to three hours spent reading through these chapters before attempting hands-on applications.

Beyond the tutorial, Chapter 8 of the manual provided examples to run the user through a broad cross-section of software functions, addressing problems that might be encountered along the way. Although this chapter appeared appropriate for a novice, the examples were inaccessible to the uninitiated user. The examples seem to have been imported directly from an earlier version of the software manual and the transition is stark. Only part of the first example (8.2.1.1 Example SPIDER1) provided the level of direction necessary for a novice to run the analysis. This provided a file path leading to the project data and step-by-step instructions on producing the graph. In no subsequent parts of example 8.2.1 were such instructions given, nor were the names of the documents on which the analyses were predicated.

In all, about 8 hr were spent learning the basics of Canoco for Windows 4.5. A drawback of the software is that, although the tutorial is deceptively easy, learning to use the software's higher functions requires a real commitment on the user's end. It bears noting that no extra assistance was sought in using the software. It is more likely that a graduate student would have access to others with a quantitative background who would be able to provide a few prompts to get them past the trouble points.

Among the assets of *Canoco for Windows 4.5* is that, if the user only needs to use the most simple functions, this software allows rapid utilization. This software is one that the graduate student can grow into as their understanding of multivariate analysis grows. The novice who needs immediate access to in-depth analysis should expect to devote considerably more time going back and forth between chapters and examples.

PC-ORD version 4

The tutorial provided by *PC-ORD version* 4 was relatively abbreviated, utilizing only one example with a single matrix. Thus, there was no opportunity to practice working with environmental data or running any sort of analysis or graphing that includes two matrices. Furthermore, the example provided (food preferences for students) had no ecological relevance. Conversely, an asset of this software program is that the layout is self-explanatory, the software is easy to navigate and the manual is organized and accessible. Thus, more examples could be utilized with ease.

The section on Modifying Data provided an excellent summary of the functions available and what these functions can offer in terms of data analysis. In this context (and others) the manual was very easy to use and follow. Beneath every instructional heading is the file path for using the function on data. Data manipulation in PC-ORD is similar to many other spreadsheet programs with which the new user is likely to be familiar. Ca. 40 min was spent learning the basics of *PC-ORD, version 4*. The new user should feel competent to run basic ordinations in this length of time and confident of their ability to master some of the software's higher functions with the manual in hand.

SYN-TAX 2000

SYN-TAX 2000 provides a very different experience for the uninitiated user, one that requires more prior knowledge than do the other two software packages. At only 53 pages, the manual provides information similar to that provided by manuals of the other packages (e.g. Quick Start, setting up data files, producing graphics, statistical theory), but its brevity limits its usefulness. The user is expected to be able to find the sample data and the 'sample analysis' is essentially a broad overview of where to find things and what impact different options will have on screen images. Sample data were found somewhat easily (the data set was appropriately labeled) and the screen images were followed according to descriptions in the manual.

The tutorial for *SYN-TAX 2000* is quite helpful in its explanation of graphics manipulation. Like *PC-ORD version 4* and unlike *Canoco for Windows 4.5,* SYN-*TAX 2000* provides comprehensive, path-specific directions for manipulating the graphics. Ca. 1 hr was spent learning the basics of *SYN-TAX 2000*. In this length of time new users should feel somewhat confident of their ability to master this software quickly.

Real data

We used CCA on a single data set to make direct comparisons between software packages. CCA was selected because (1) it has been shown to provide reliable assessments of spatial and temporal patterns of vegetation (Palmer 1993), (2) it is used with increasing frequency by vegetation scientists and (3) by using both species and environmental data, it allows for more direct comparisons of outcomes among the software packages (i.e. we were able to compare axis scores for sample plots, individual species and environmental variables).

The data set is from a study carried out on the Boyd Tract, a 66-ha old-growth stand of longleaf pine forest in the Sandhills of North Carolina (Gilliam et al. 1993; Gilliam & Platt 1999).

Woody stem sampling ($\geq$ 2.5 cm DBH) was carried out in plots established randomly on two distinct site types within the Boyd Tract: mesic uplands and xeric lowlands. This apparent contradiction of site conditions (i.e., uplands are generally more xeric than lowlands) is the result of unconsolidated Cretaceous clay deposits overlying unconsolidated Cretaceous sand deposits. A total of 60 circular 0.04-ha sample plots (11.3 m radius) were established on the Boyd Tract, 32 in the lowland and 28 in the upland. Environmental data collected for each plot were soil analyses for pH (pH_w and pH_s for H_2O- and KCl-extractable pH, respectively), particle size distribution (% sand, silt, clay) and organic matter (loss-on-ignition). Published results have shown the following. Among 19 tree species tallied for the Boyd Tract, dominant woody species in mesic plots were *Quercus velutina*, *Carya* spp. and large, sparse *Pinus palustris* trees. Xeric plots showed high densities of *Quercus laevis* with the large *P. palustris*, as well as higher frequencies of smaller *P. palustris* stems. *T*-tests indicated significant differences between site types in soil texture (higher clay in the upland and higher sand in the lowland), pH (higher in the upland) and organic matter (also higher in the upland).

Prior to this review, these data had not been subjected to CCA. Accordingly, the first question to be addressed is: *does CCA provide the same conclusions regarding site-related species and soil differences as those of previous studies?* In the context of this comparative review, the second question to be addressed is: *how do CCA results compare among the software packages?* Default settings were used for each.

Results of CCA on Boyd Tract data confirm earlier conclusions, with lowland and upland plots generally well-separated, but with some degree of overlap (Figs. 1-3). Further, *Quercus laevis* (QULA) showed a high affinity for lowland plots, whereas *Quercus velutina* (QUVE), *Carya* species (CARY) and *Cornus florida* (COFL) showed a high affinity for upland plots; *Pinus palustris* (PIPA), which was ubiquitous among areas of the Boyd Tract, exhibited little site affinity. The environmental gradients produced by CCA, however, contradict the importance of soil texture, especially clay content, reported in earlier studies. The most important environmental factors, measured as vector length, were pH_s and organic matter (both generally higher in upland plots), followed in order by sand, pH_w, silt and clay. Thus, CCA was able to uncover vegetation-environment relationships that were not apparent from previous analyses.

The range of values for axes scores of plots, species and environmental factors varied greatly among the software packages. Compare, for example, CCA Axis 1 from *Canoco for Windows 4.5* (–1.5 to 3) versus that from *SYN-TAX 2000* (–0.8 to 0.8) and CCA Axis 2 from *Canoco for Windows 4.5* (–2.5 to 2.5) versus that from *PC-ORD version 4* (–0.6 to 1). The range of axis scores for species was quite different from that for plots and environmental factors in *PC-ORD version 4*. As a result, graphing plot scores alone with environmental vectors (Fig. 2a) yielded interpretable patterns, whereas graphing these along with species scores greatly compressed the environmental vectors, as well as the range of variation along the *y*-axis (Fig. 2b).

The direction of the axes also varied among software packages, wherein upland plots occurred on the positive side of CCA Axis 1 for *Canoco for Windows 4.5* and *PC-ORD version 4*, but on the negative side for *SYN-TAX 2000*. In
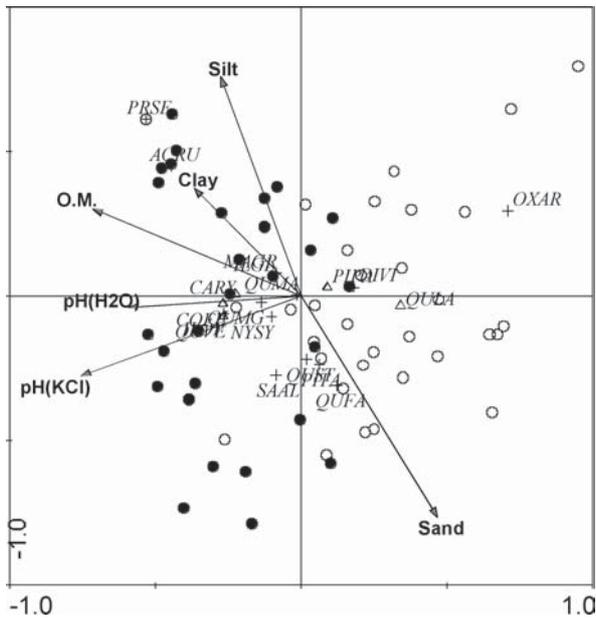
Fig. 1. CCA triplot of Boyd Tract data, as produced by *Canoco for Windows 4.5 and CanoDraw*.

addition, sand content was oriented toward the negative end of CCA Axis 2 for *Canoco for Windows 4.5*, but was oriented toward the positive end for *PC-ORD version 4* and *SYN-TAX 2000*.

Despite these differences, the results of all three packages led to the same conclusions regarding elevation-related differences in species composition and environmental gradients. In other words, as one would expect, the solutions for all three packages were identical, except for different scaling and reversal of axes.

Discussion

User-friendliness

We consider user-friendliness from the two experiences reported in the Trials; (1) use of tutorials for learning by the beginning user and (2) analysis of actual data by the more experienced user. *Canoco for Windows 4.5* provided an excellent tutorial that familiarizes the user with the basics of the software operation. Beyond that, the manual's size and dense text was at times a bit overwhelming for the first-time user, who was required to save a lot of documents whose functions were not always immediately clear. On the other hand, the software appeared to allow far more user-defined options than *PC-ORD version 4* or *SYN-TAX 2000*.

PC-ORD version 4 provided a tutorial that can be described as minimalist at best. On the other hand, the package had a very accessible software format, accompanied by an excellent, well-written and easily understood manual. For the novice, this program felt the most user-friendly. If it had provided the same level of analysis and flexibility as *Canoco for Windows 4.5*, *PC-ORD version 4* would certainly be pre-
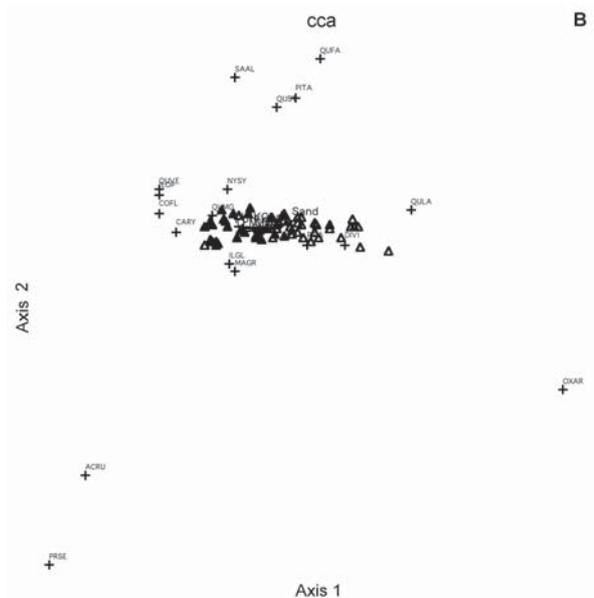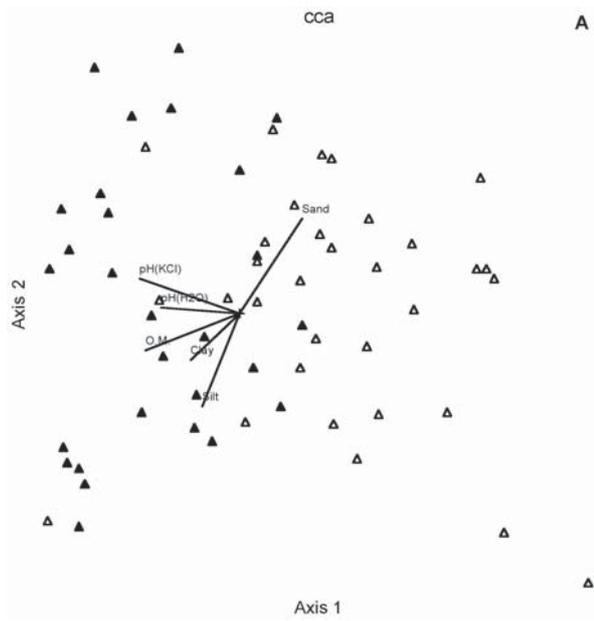

Fig. 2. CCA triplot of Boyd Tract data, as produced by *PC-ORD version 4*. **A**. Plotted without species scores; **B**. Plotted with species scores.

ferred.

SYN-TAX 2000 might be described as the working man's software package: not a lot to say, but it gets the job done. The scant 53 pages of the manual, along with its lack of an index, limit its usefulness, especially compared to the somewhat encyclopedic manuals of the other two packages. To rank these for ease-of-use by a novice, we place *SYN-TAX 2000* a close second behind *PC-ORD version 4*, primarily for the paucity of its user's manual. *Canoco for Windows 4.5* was a bit
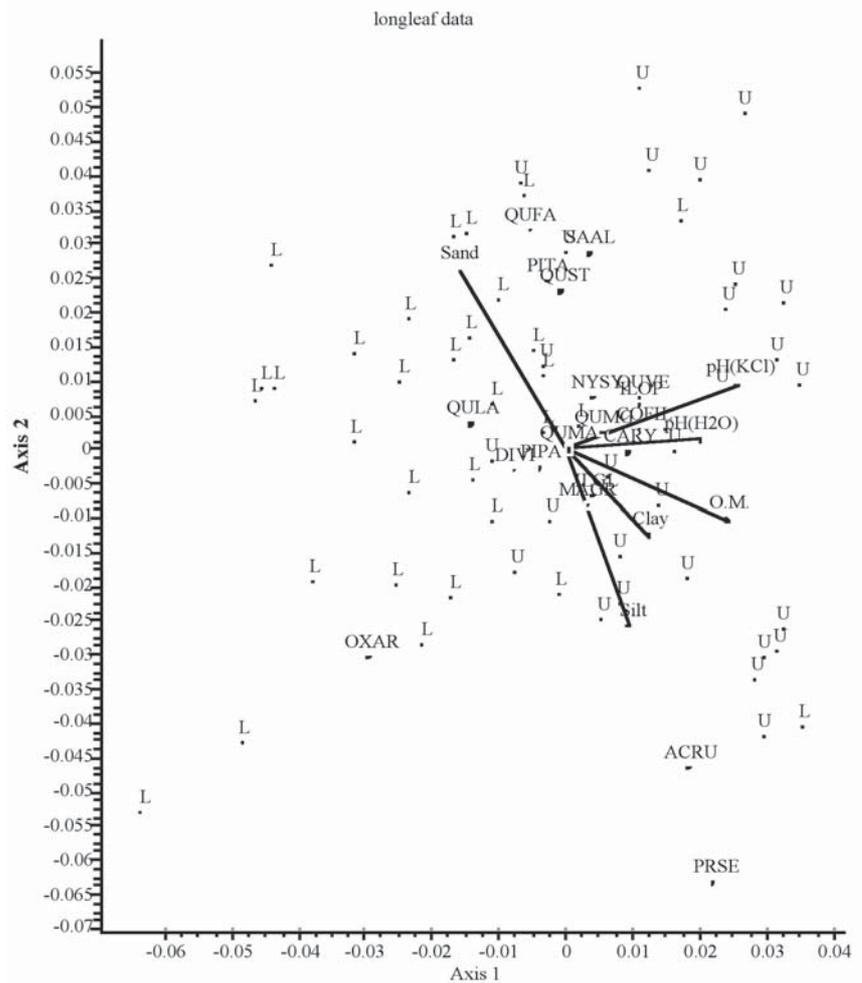
Fig. 3. CCA triplot of Boyd Tract data, as produced by *SYN-TAX 2000*.

overwhelming for a beginner.

For analysis of actual data, our consideration of user-friendliness must begin with data input. The addition of *WCanoImp* into the *Canoco for Windows 4.5* bundle makes data input from an existing (in this case, Microsoft Excel) spreadsheet very easy, whereas additional steps were required in preparing data for analysis with the other two packages. Because all packages are Windows-based, getting through the options for CCA was done with relatively equal ease.

Once the analysis was completed, the graphics and further analytical options were more extensive for *Canoco for Windows 4.5* (Fig. 1) and *PC-ORD version 4* (Fig. 2a, b) than for *SYN-TAX 2000* (Fig. 3). Furthermore, the quality of graphics provided by the packages varied greatly. Again, additions to *Canoco for Windows 4.5* (namely *CanoDraw*) represent substantial improvements in graphics capabilities and ease of use. For both graphics user-friendliness and quality, we would rank them in the order: *PC-ORD version 4 > Canoco for Windows 4.5 > SYN-TAX 2000*.

Cost effectiveness

For new licenses and upgrades (in brackets), costs are the following (Table 1):

Canoco for Windows 4.5	$378.00	($249.00)
SYN-TAX 2000	$350.00	($200.00)
PC-ORD version 4	$199.00	($99.00)

Assuming that all packages offered the same analyses, *PC-ORD version 4* would clearly be the most cost-effective, with new licenses and upgrades being almost one-half the cost of the other two. However, it is even more cost-effective considering the greater breadth of analyses it provides.

Recommendations

We stated in the Introduction that we would conclude our review with recommendations regarding which package(s) we would deem worthy of use by our colleagues in the field of vegetation science. At the risk of sounding equivocal, we suggest that all three software packages are worthy. Indeed, we are heartened by the breadth of choice of excellent analytical tools available to vegetation scientists.

We would like to emphasize the difficulty in providing a ranking of these software packages, because of the vast differences in analytical capabilities. Indeed, direct comparisons cannot be made for several methods, the importance of which vary greatly amongst plant ecologists. For example, *Canoco for Windows 4.5* was designed solely for ordination and, thus, would not be applicable for classification, which is well-developed in *PC-ORD version* 4 and *SYN-TAX 2000*. Another example of such differences would be NMDS, a multivariate method of recently-increasing use among vegetation scientists. This is not available in *Canoco for Windows 4.5*, but is available in the other two. Of these, NMDS is particularly well-developed in *PC-ORD version 4*. Since only *Canoco for Windows 4.5* is capable of partial analysis, it is not possible to perform variance partitioning (e.g., Borcard et al. 1992).

To return to our analogy with natural selection, these three software packages are examples of the positive results of the selective pressures of ongoing scrutiny. We look forward to the new 'species' of the future.

References

Barbour, M.G., Burk, J.H., Pitts, W.D., Gilliam, F.S. & Schwartz, M.W. 1999. *Terrestrial plant ecology*. 3rd. ed. Benjamin/Cummings, Menlo Park, CA, US.

Borcard, D., Legendre, P. & Drapeau, P. 1992. Partialling out the spatial component of ecological variation. *Ecology* 73: 1045-1055.

Gilliam, F.S. & Platt, W.J. 1999. Effects of long-term fire exclusion on tree species composition and stand structure in an old-growth longleaf pine forest. *Plant Ecol*. 140: 15-26.

Gilliam, F.S. & Roberts, M.R. (eds.) 2003. *The herbaceous layer in forests of Eastern North America*. Oxford University Press, New York, NY, US.

Gilliam, F.S., Yurish, B.M. & Goodwin, L.M. 1993. Community composition of an old-growth longleaf pine forest: relationship to soil texture on tree species. *Bull. Torrey Bot. Club* 120: 287-294.

Hill, M.O. & Gauch Jr., H.G. 1980. Detrended correspondence analysis: an improved ordination technique. *Vegetatio* 42: 47-58.

McCune, B. & Grace, J.B. 2002. *Analysis of ecological communities*. MjM Software Design, Gleneden Beach, OR, US.

McCune, B. & Mefford, M.J. 1999. *PC-ORD. Multivariate analysis of ecological data, version 4*. MjM Software Design, Gleneden Beach, OR, US.

Palmer, M.W. 1993. Putting things in even better order: the advantages of canonical correspondence analysis. *Ecology* 74: 2215-2230.

Peet, R.K., Knox, R.G., Case, J.S. & Allen, R.B. 1988. Putting things in order: the advantages of detrended correspondence analysis. *Am. Nat*. 131: 924-934.

Podani, J. 1980. Computer programs for ecological and taxonomical classifications (in Hungarian). *Abstr. Bot*. 6: 1-158.

Podani, J. 1997. SYN-TAX 5.1: A new version for PC and Macintosh computers. *Coenoses* 12: 149-152.

Podani, J. 2000. *Introduction to the exploration of multivariate biological data*. Backhuys, Leiden, NL.

Rejmánek, M. & Klinger, R. 2003. CANOCO 4.5 and some comparisons with PC-ORD and SYN-TAX. *Bull. Ecol. Soc. Am*. 84: 79-74.

ter Braak, C.J.F. 1987a. *CANOCO - a FORTRAN program for community ordination by [partial] [detrended] [canonical] correspondence analysis, principal components analysis and redundancy analysis*. Version 2.1. ITI-TNO, Wageningen, NL.

ter Braak, C.J.F. 1987b. *Unimodal models to relate species to environment*. Agricultural Mathematics Group, Wageningen, NL.

ter Braak, C.J.F. 1988. CANOCO – an extension of DECORANA to analyze species-environment relationships. *Vegetatio* 75: 159-60.

ter Braak, C.J.F. & Šmilauer, P. 2002. *CANOCO reference manual and user's guide to Canoco for Windows: Software for Canonical Community Ordination (version 4.5)*. Microcomputer Power, Ithaca, NY, US.

Wartenberg, D., Ferson, S. & Rohlf, F.J. 1987. Putting things in order: a critique of detrended correspondence analysis. *Am. Nat*. 129: 434-448.