

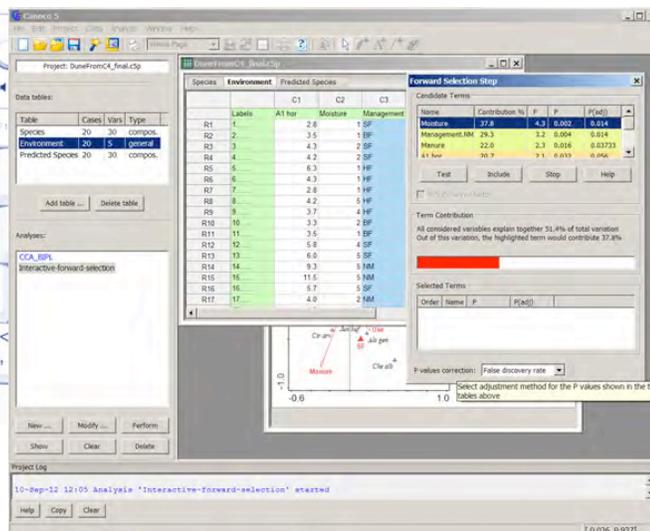
# Canoco 5 arrives!

Software for multivariate analysis and visualization

Available from:



[www.microcomputerpower.com](http://www.microcomputerpower.com)



... better on all counts:

## ✓ Powerful

multiple data tables in single Canoco 5 project, with size restricted mostly by the operating system<sup>1</sup>, supported statistical methods have no additional limits<sup>2</sup>; 32-bit or 64-bit versions installed according to Microsoft Windows™ version<sup>3</sup>; direct import from .xls and .xlsx format files

## ✓ Versatile

traditional methods (PCA, CA, DCA, RDA, CCA, DCCA) with a wide range of Monte Carlo permutation test types extended by principal coordinates analysis, db-RDA, NMDS, co-correspondence analysis; many easy-to-use analysis templates for testing significance of ordination axes, variation partitioning (including PCNM), advanced stepwise selection, principal response curves (PRC) method, analyses including functional traits, discriminant analysis, and others

## ✓ User-friendly

intuitive user interface complemented by Canoco Adviser™ expert system suggesting analyses, values for their options, ordination diagrams to create, and offering their interpretation

## ✓ Integrated

just a single application for data import, analysis and graph creation; multiple analyses on the same dataset share settings, classifications, groups, and series definitions in a project

## ✓ Affordable

pricing of new licenses and of upgrades is the same as for the previous version 4.5

To learn more

[www.canoco.com](http://www.canoco.com)

To get a trial version or to order

[www.microcomputerpower.com](http://www.microcomputerpower.com)

<sup>1</sup> number of rows or columns in each table is limited to 16.4 million entries

<sup>2</sup> except NMDS method, with a maximum of 6,800 cases on 32-bit version and 50,000 cases on 64-bit version

<sup>3</sup> supported versions: Windows XP SP2 or SP3, Vista (32- or 64-bit), Windows 7 (32- or 64 bit), Windows 8 (32- or 64-bit)

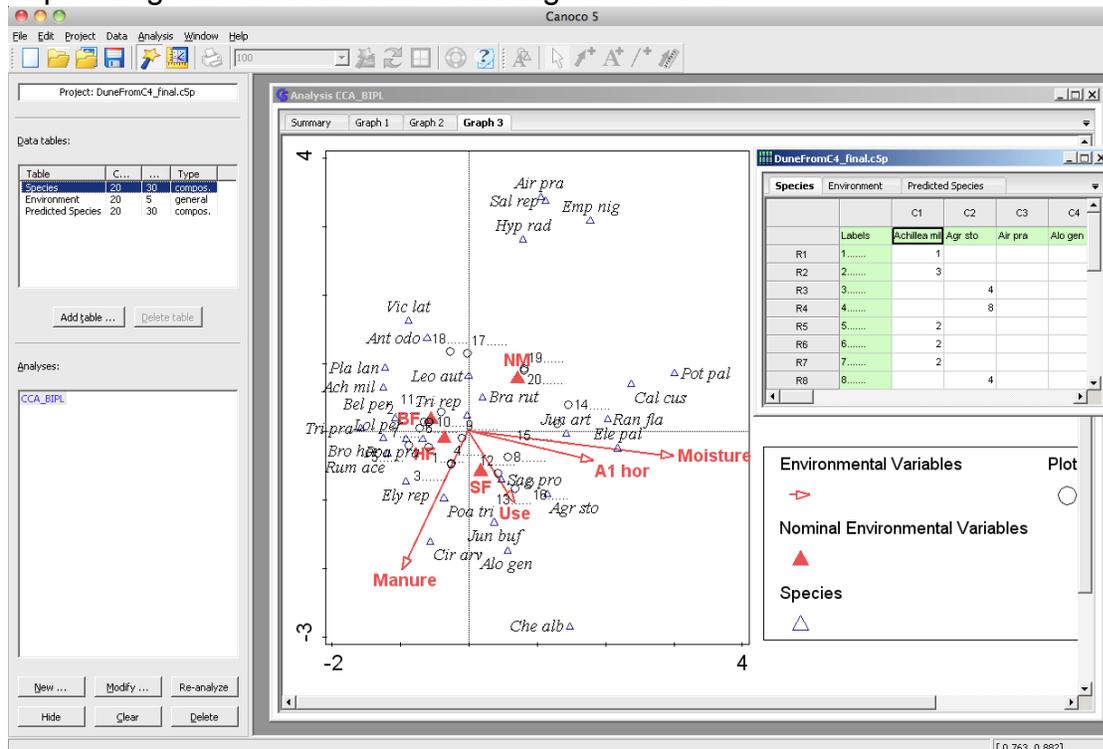
# Overview of Canoco 5

The main features of Canoco 5 are summarized in the following points:

- Analytical and graphing capabilities are integrated with an easy-to-use spreadsheet data editor in a single program. All analyses done on a set of data tables are now collected within Canoco 5 projects, sharing the analytical and graphing settings.
- All statistical methods offered by Canoco for Windows 4.5 are available, including DCA, CA, CCA, DCCA, PCA, and RDA methods including their partial variants, with Monte Carlo permutation tests for constrained ordination methods, including the correct setup for data coming from non-trivial sampling designs.
- All visualization tools of CanoDraw 4.x are available (including the loess, GLM and GAM models for the visualization of data attributes in ordination space) and improved.
- Data can be entered within the program itself or easily imported from Excel (.XLS or .XLSX formats) or from Canoco 4.x data files. Labels no longer need to be shortened to 8 characters, but these brief forms are still available (and can be even automatically generated from the long ones) for display in diagrams. Standard coding of factors (categorical predictors) is now used, dummy (0/1) variables are generated internally.
- Principal coordinate analysis (PCoA) and distance based RDA (db-RDA) are now easily accessible, with new distance measures added (11 distance types in total, including Bray-Curtis, Gower distance, and Jaccard coefficients).
- Variation partitioning is easily accessible for two or three groups of predictors including calculations of individual fractions of explained variation, based either on partial or non-partial analyses.
- Principal coordinates of neighbor matrices (PCNM) methods are easily available within variation partitioning framework.
- Computing, testing and graphing the Principal response curves (PRC) is now an easy task.
- Co-correspondence analysis (CoCA, symmetric form) is available, including Monte Carlo permutation testing.
- Stepwise selection of predictors was visually enhanced and provides support for protection against Type I error inflation (preliminary test with all predictors, adjustment of p values by one of three methods).
- Straightforward testing of all individual constrained axes as well as comparing results of two ordinations with Procrustes analysis is available.
- One can easily work with species functional traits or import the data on phylogenetic relatedness of species.
- Visualization capabilities were enhanced by the semi-transparent fill colour attributes, calibration axes for arrows associated with a variable in ordination diagrams or enclosing ellipses as an alternative to enclosing polygons. Additional types of export in JPEG, TIFF, and PDF formats added.
- Every step of your work is supported by the context-sensitive help system and by the Canoco Adviser – an expert system that helps you to select a proper analytical method for your research question, a correct type of ordination model (linear vs. unimodal) and an appropriate visualization of the results.
- Advanced users can combine multiple methods in a single analysis, including generalized linear models to correlate scores etc.



- Triplot diagram on an Intel Mac running CrossOver:



#### 4. Is Canoco 5 a 64-bit program?

There is now just one application within the Canoco 5 package and the installer installs the 64-bit version on 64-bit Windows and the 32-bit version on 32-bit Windows. The files read and stored are compatible between 32-bit and 64-bit versions of Canoco 5.

#### 5. How does Canoco 5 deal with the stability issues reported by Oksanen & Minchin (1997)?

The bug in the original DECORANA (Hill, 1979) code for Detrended Correspondence Analysis has been studied and was already repaired by Canoco 4, 15 years ago. It turned out that the change suggested by Oksanen and Minchin (1997) did not make the smoothing routine order-invariant. Another change was required as well, as agreed upon by Oksanen & Minchin in subsequent correspondence (see <http://cc.oulu.fi/~jarioksa/pages/isbug.htm>).

#### 6. Does Canoco 5 provide NMDS?

Yes, the non-metric multidimensional scaling (NMDS or NMS) method is now fully integrated into Canoco 5, replacing the add-on program WinKyst from the previous version. The WinKyst program remains available as a free alternative solution for performing the NMDS analyses. In addition, Canoco 5 provides a wide array of ordination methods that operate directly on the data, instead of on derived similarity or dissimilarity (distance) measures, as NMDS does. The advantage of these methods over NMDS include:

- The resulting ordination readily provides information on the level of individual variables in the data table.

- The ordination can be focused on the effect of particular explanatory variable(s) – in the form of constrained ordination – so that it shows e.g. the community response to explanatory variables.
- Background variation can be removed from the ordination by using covariate data, so that the ordination can display new information rather than trivial or uninteresting variation, or variation that is already well understood.

## 7. Is Canoco 5 able to analyse my data set with 20,000 samples and 600 species?

Definitely, and much larger ones as well! There are always limits to the size of data one can analyze on a personal computer, but in version 5, Canoco itself enforces only the limit of 16.7 million table rows and 16.7 million table columns, which is not very limiting. Operating system limitations, however, impose much harsher limits in the case of 32-bit Windows. Standard applications in 32-bit Windows cannot allocate more than 2 GB in total, with Canoco 5 using special provisions to extend its data space to 3 GB (after proper Windows re-configuration). A single, contiguous memory block cannot be allocated with a size greater than *ca* 1 GB, so Canoco 5 data tables are limited to roughly 134 million entries. Yet another limitation usually comes from the processor speed, making analyses of really big data unbearably slow (but you can always leave the computer working overnight).

## 8. How does Canoco 5 compare to PC-ORD version 6?

Canoco 5 is software primarily focused on ordination methods, while PC-ORD offers a wider range of methods, including cluster analysis, TWINSpan, species-area curves, Mantel test and similar distance-based permutation tests.

We believe, however, that due to its focus on ordination, Canoco does a better job in this field in terms of offered statistical methods, user friendliness of the software and the quality of the graphical output. PC-ORD version 6 (as of October 2012) does not support partial constrained and unconstrained analyses, needed to address many of the non-trivial research questions of today: PC-ORD has neither support for selecting a subset of explanatory variables reasonably describing the community variation, nor the ability to easily evaluate simple as well as conditional effects of individual predictors collected in the project. Additionally, PC-ORD does not support correct permutation tests for datasets coming from several frequently encountered design types, including repeatedly measured permanent plots or data coming from linear transects or rectangular sampling grids.

The Canoco package also includes *Canoco 5*, the 496 page manual. The Canoco manual includes comprehensive discussion of many published papers. In addition, the Canoco installation includes many of the data sets used in the papers. Users are invited to reproduce analyses found in the papers, explore the data further and have the tools and discussion at hand to help them get the most from Canoco. Canoco is also supported by a pair of excellent textbooks, available separately, describing Canoco and its application:

- Jongmann, ter Braak, van Tongeren (1995) *Data Analysis in Community and Landscape Ecology*
- Šmilauer & Lepš (2014) *Multivariate Analysis of Ecological Data using CANOCO 5*

Naturally, both programs continue to develop. Version 6 of PC-ORD added linear constrained ordination (redundancy analysis - RDA) that was missing in previous versions. Canoco 5 adds the non-metric multidimensional scaling (NMDS) that was previously available only as a (free) add-on package. Canoco 5 features the symmetric form of co-correspondence analysis to relate two biotic communities together.

Both Canoco 5 and PC-ORD are valuable analytical tools which can be used to corroborate one another where they offer overlapping functionality. Each complements the other with analyses not available in the other package.

We would like to mention here some of the gross mistakes in the PC-ORD – Canoco comparison table published on PC-ORD web site: <http://home.centurytel.net/~mjm/canoco.htm> (as of October 2012). This comparison concerns the previous version of Canoco (4.5), so newly added features of Canoco 5 are naturally not covered (no strict size limits either in analyses or in graphing, import from CSV files, presence of NMDS, storage of distance matrix), but some of the claims were not correct even for version 4.5 and may be perpetuated into any new comparisons.

Also important is the comparison of prices, where MjM boasts of a substantially cheaper product, but the comparison made for a single user license compares a personal license of PC-ORD with a Canoco price for private companies. A fairer comparison would be among educational licenses purchased by an academic department for its employee. As the PC-ORD pages tell you, if the license should stay at the institution which bought it for you, the price for a single user is 299 + 100 USD, which does not compare so well with a single educational license of 368 USD for Canoco 5 (same as for Canoco 4.5). Similarly, the student licenses are offered for Canoco as additional educational licenses for 238 USD, which does not compare so badly with the 199 USD for PC-ORD. It is only fair to add that shipping costs are smaller for PC-ORD.

As a rather minor point, we shall note that Canoco 4.5 had (and continues to have in version 5) the “successional vectors” called series collections in PC-ORD, that monotonic transformations (namely log and sqrt) are available, as are the diversity indices computed for individual cases, and the detection of outliers with leverage statistics for individual predictors as well as for their whole set.

## 9. How does Canoco 5 compare to the Vegan package in R?

The [Vegan package in R](#) implements most of the methods available in Canoco 5 as well as a few more. As [R](#) is an open system, those Canoco methods missing in [Vegan](#) can be found in other packages, related or unrelated to [Vegan](#). However, the R environment is primarily command-oriented and to import data, perform analysis or create simple ordination diagrams, commands calling functions are needed, often with a complex set of function parameters to correctly choose from. We see the R environment as excellent for a statistically-oriented ecologist, but for researchers focused on their research subjects, using [Vegan](#) in the R environment has a steep learning curve, which must be climbed repeatedly for someone evaluating their own data after long periods of work in the field or in lab. There is a relatively user-friendly interface in the [Rcmdr](#) package, offering in combination with the DiversityR package some of the multivariate methods of [Vegan](#), but only for a small subset, excluding all the advanced methods and advanced options.

