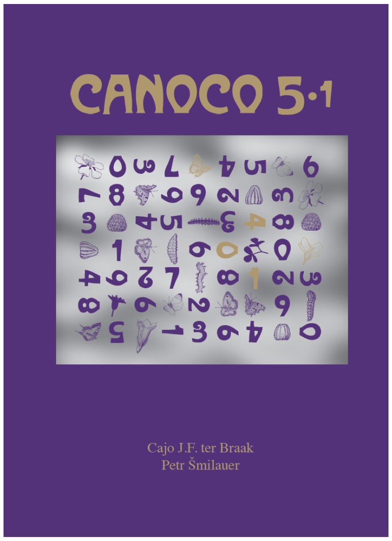
Canoco for visualization of multivariate data



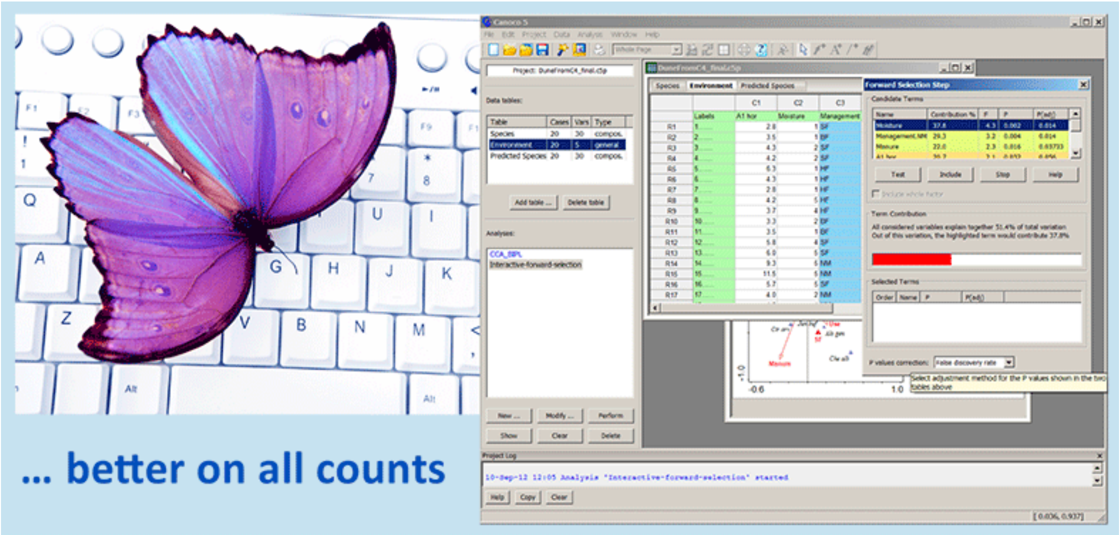
Canoco is a software package for multivariate data analysis and visualization, with an emphasis on dimension reduction (ordination), regression analysis, and the combination of the two, constrained ordination.

Canoco 5, released in 2012, is now in version 5.12. Canoco 4.0-5.0 were cited more than 10,000 times in the past 20 years (1998-2018, ISI Web of Science). Canoco 5 implements many recent advances in ordination, such as variation partitioning, co-correspondence analysis, distance-based redundancy analysis and double constrained ordination for trait-environment association, but the major advance is in user-friendliness.

In Canoco 5, data import, analyses and making graphs are integrated in a single Canoco 5 project. The Canoco Adviser helps in choosing data transformations and methods of analysis. Numerical analyses that used to take many runs, are now available through a single analysis template and the Analysis Notebook concisely summarizes the results and allows access to the full results. All analyses done on a set of data tables are now collected within Canoco 5 projects, sharing the analytical and graphing settings. Canoco 5 helps to make even better publication-quality ordination diagrams. The manual has been largely re-written and the large set of real-life examples is updated and extended to show new ways of working with multivariate data.

Introduction to Canoco

Human beings wish to detect, explain and understand pattern and structure in nature. The aim can be scientific or technical or often both. But nature can be chaotic and so pattern and structure must often be derived from noisy data-sets ranging from a few to many cases and variables. Ordination methods as provided by Canoco can help scientists to infer pattern and structure from such data sets, in particular in the form of relationships among cases and variables. The data may arise from the field or the laboratory, and can be observational or experimental. Nowadays many data-sets come from high-throughput instruments resulting in DNA profiles or expression levels or in the amounts of proteins or metabolites with associated factors or properties of the sample, such as tissue type, experimental treatments, and time of sampling.



```
10-Sep-12 12:05 Analysis 'Interactive-forward-selection' started
```

Canoco software offers methods in which data-tables and/or their constituent variables are assigned a particular role. There are **four roles**. The data-table that is to be understood in some way takes the role of **response data**. If data are subjected to analysis to find the main structure, those data are said to take the response data role. The reason is that in this method the data are explained by ordination axes (hypothetical explanatory variables or latent variables). The resulting ordination of this data-table can then be interpreted by variables in a second data-table taking the role of **supplementary data**. **Explanatory data** (in the second or a third data-table) takes a stronger role than just interpretation. As its name says, its variables are used to explain the response data. In consequence, the ordination must focus on the explained variation and on the effects that the explanatory variables have on the response data. This is achieved by constrained ordination which integrates aspects of (unconstrained) ordination and multiple regression. Constrained ordination brings in the power of regression analysis such as variation partitioning and statistical testing of effects. The final role is the **covariate data** role. Covariates (often in the same data-table as the explanatory or supplementary data) enable the detection of the main structure in the response data *after* accounting for the variation explained by these covariates. All data other than response data is often termed metadata.

Canoco methods originated in the field of ecology. There, ordination analysis with Canoco can provide insights into the structure of biological communities and into the impact of natural and human-induced environmental disturbances on biological assemblages. Canoco has been used in the past to help answer research questions such as:

- How does the vegetation develop on abandoned cultivation sites within a tropical rain forest?
- How does agricultural management practice affect meadow vegetation?
- What are the effects on forest undergrowth if liming is used to mitigate the effects of acid rain?
- Do diatoms respond so strongly to lake pH that they can be used to monitor trends in acidity?
- How long does it take before an invertebrate community recovers from an application of the insecticide chlorpyrifos? How does the time to recover depend on the concentration?

From the way these questions are worded it is already fairly clear which are the response data and which data have an explanatory or supplementary role.

Canoco contains six main classes of ordination methods:

1. Methods to describe the structure in a single data-set, thus taking the response data role. For instance, the structure of a biological community or the correlation structure of a set of variables (ordination, unsupervised learning).
2. Methods to describe the structure in a single-data set (response data), and to interpret the results by external (supplementary) data (ordination, unsupervised learning, indirect gradient analysis).
3. Methods to explain one data-set (response data) by another data-set (explanatory data). For instance, to explain or to predict the abundances or expression levels from explanatory data (constrained ordination, supervised learning, direct gradient analysis).
4. Methods to describe the structure in a single data-set (response data) after accounting for the variation explained by a second data-set (covariate data). For instance, the community structure adjusted for observer and seasonal effects (partial ordination, unsupervised learning in the space orthogonal to the covariate data).
5. Methods to explain one data-set (response data) by another data-set (explanatory data), after accounting for the variation explained by a third data-set (covariate data). For instance, to explain species abundances from explanatory data, adjusted for observer and seasonal effects (partial constrained ordination, supervised learning).
6. Methods to explain one data-set (response data) by two other data-sets (predictor data for cases and for variables, referred to as explanatory data and trait data, respectively). For instance, to explain or to predict the abundances in biotic communities using environmental characteristics for cases and functional traits for species (double-constrained ordination).

Within each of these six classes, you can choose between two response models: a linear model and a unimodal model. The basic ordination methods, Principal Components Analysis and Correspondence Analysis are thus extended to constrained, partial, partial constrained, and double-constrained forms. These methods work with rectangular data-tables (they may be tall or wide) of cases by variables. Canoco 5 now integrates classical and non-metric multidimensional scaling methods that work with square symmetric matrices, such as distance matrices, similarity matrices, or dissimilarity matrices. Such matrices can either be imported or calculated from the data-tables using eleven popular measures of (dis)similarity.

The principle output of an analysis consists of:

- ordination diagrams with a numerical summary of the variation explained,
- the variation explained by the explanatory variables, if present,
- the statistical significance of the explanatory variables, if present.
- the statistical significance of each constrained ordination axis
- a ranked list of the effect sizes of the explanatory variables with their statistical significance
- a partitioning of the variation if the explanatory variables are grouped into two or three groups
- a selection (subset) out of all the explanatory variables that parsimoniously explains the response data.

Tables extracted directly from the 536 page Canoco 5.1 Manual

The tables cross reference **Example** analyses studied in the manual. The examples come from peer-reviewed articles using original data sets also included in the Canoco distribution.

Table 1.1 Standard and advanced ordination methods available via the *New* .. button

Abbreviation	Full name and aliases
PCA	Principal components analysis
RDA	Redundancy analysis, Reduced-rank regression, PCA with respect to instrumental variables (PCA-IV)
CA	Correspondence analysis, dual scaling, homogeneity analysis; includes multiple correspondence analysis
DCA	Detrended correspondence analysis
CCA	Canonical correspondence analysis, CA with respect to instrumental variables (CA-IV)
DCCA	Detrended canonical correspondence analysis
PRC	Principal response curves analysis
dc-CA	Double-constrained correspondence analysis
dc-PCA	Double-constrained principal components analysis
Var-part-*N*groups	Variation partitioning with *N* groups of explanatory variables (*N* = 2 or 3)
Var-part-PCNM	Variation partitioning where one group is represented by a distance matrix and turned into variables by PCNM (principal coordinates analysis of neighbour matrices).
FS, forwsel	Interactive / manual forward selection of explanatory variables
Calibration	Classical calibration of missing values in explanatory variables based on a constrained ordination of response against explanatory variables
Partial ...	All above methods with covariates
Hybrid ...	All above methods with a smaller than maximum number of constrained axes

Table 1.2 Specialized ordination methods available via the *New* .. button, with an example of use in chapter 6.4

Abbreviation	Full name and aliases	Example
NMDS	Nonmetric multidimensional scaling	DistBased
PCO	Principal coordinates analysis, classical scaling, metric scaling, kernel PCA	DistBased
PCNM	Principal coordinates analysis of neighbour matrices, used in variation partitioning	Oribatids
db-RDA	Distance-based redundancy analysis	DistBased
Procrustes analysis*	Compare ordinations	Spider1
CoCA	Co-correspondence analysis	CoCA
Coinertia	Co-inertia analysis	Coinertia
CVA	Canonical variate analysis	CVA
Discriminant analysis	CVA, Fisher's linear multi-group discriminant analysis	CVA
GLM	Generalized linear models, including multiple regression	CVA, Multregr
CWM-RDA	Community-weighed means RDA	Traits
Response trait-methods	Response variable-level analysis of response traits (e.g. species traits)	Traits

* available via the *Analysis* | *Add new analysis*| *Compare ordinations* menu command

Table 1.3 Available methods that are not ordination methods

Name	Calculation of ...	Example
	via the *Data*| *Add new table(s)* |.. menu command	
Diversity indices	Summary statistics of compositional data-tables	*DuneBook*
Trait averages	Community-weighted trait averages (CWM)	*Traits*
Functional diversity	Diversity of a case in terms of the traits of its response variables	*Traits*
Indicator values	Indicator values of response variables (species) for discriminating among groups (levels of a factor)	*DuneVarPart*
Phylogenetic distances	Distances among response variables (species) based on a phylogenetic tree	*RootArch*
Predict	Predicted and fitted values from a constrained ordination	*DuneBook*
	via the *Graph* | *Attribute plots* menu command	
Response curves	Nonlinear response curves via Loess, GLM, or GAM, e.g. Gaussian curves, trend curves	*Spider1*
	via the *Project* | *Create derived project* menu command	
Aggregate	Group means and relative frequencies of quantitative and qualitative variables, respectively	*SplitPlot*
Expand	Expand/inflate/vectorize a central table and associate each entry with values in the other tables	*Traits*

Canoco System Requirements and data size limits.

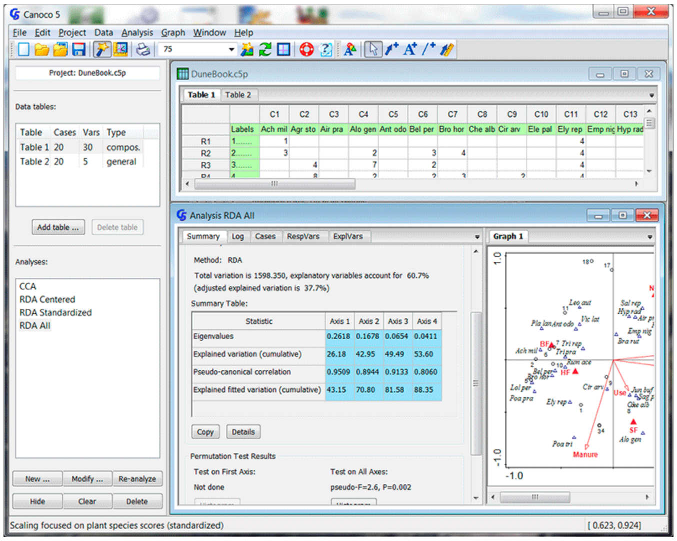
Canoco 5 requires installation on Microsoft Windows™ XP with SP2 (or SP3) or later OS versions, including Windows 10. The installer determines the type of operating system (32-bit versus 64-bit) and installs the appropriate version of the Canoco 5 executable - all other installed components are identical. The Canoco 5 project and graph files created in 32-bit and 64-bit versions are compatible. For successful work with the Canoco 5 user interface, the minimum screen resolution is 1024x768 dots, but a larger resolution is recommended.

The limits on data sizes in Canoco 5 are primarily set by the limitations of the operating system, particularly for the 32-bit version: given the maximum size of contiguous operating memory blocks that might be allocated on such systems (ca. 1 GB), none of the analyzed tables can have more than ca. 134.2 million entries. The 32-bit version of Canoco 5 is compiled with an enlarged memory space capacity, so that it can allocate in total up to 3 GB of memory. Special restrictions apply to the NMDS method, where the number of analyzed cases cannot exceed 6,800 for the 32-bit version and 50,000 for the 64-bit version. Further, the number of data-table rows and columns in both the 32-bit and 64-bit versions of Canoco 5 cannot be larger than 16.7 million, but this size is hardly achievable given the constraints discussed earlier which already imply unrealistic demands on system performance.

A Survey of Canoco

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.
  Snapshot shows the Canoco 5 workspace with an open project. At the left side, the defined data tables and defined analyses are listed, at the right side an analysis notebook is shown with its Summary page and one created ordination diagram.



Project: DuneBook.c5p

Data tables:

Table	Cases	Vars	Type
Table 1	20	30	compos.
Table 2	20	5	general

Analyses:

CCA
RDA Centered
RDA Standardized
RDA All

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
Labels	Ach mil	Agr sto	Air pra	Alo gen	Ant odo	Bel per	Bro hor	Che alb	Cir arv	Ele pal	Ely rep	Emp nig	Hyp rad
R1	1.......	1									4		
R2	2.......	3			2		3	4			4		
R3	3.......		4		7		2				4		

Method: RDA

Total variation is 1598.350, explanatory variables account for 60.7%
(adjusted explained variation is 37.7%)

Summary Table:

Statistic	Axis 1	Axis 2	Axis 3	Axis 4
Eigenvalues	0.2618	0.1678	0.0654	0.0411
Explained variation (cumulative)	26.18	42.95	49.49	53.60
Pseudo-canonical correlation	0.9509	0.8944	0.9133	0.8060
Explained fitted variation (cumulative)	43.15	70.80	81.58	88.35

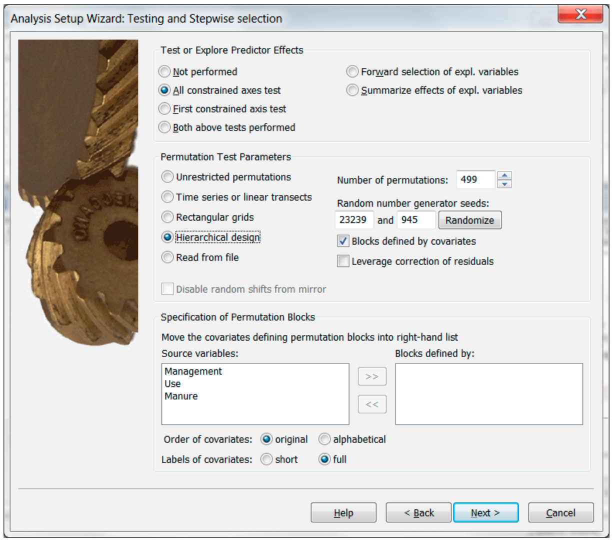
Permutation Test Results

Test on First Axis: Not done

Test on All Axes: pseudo-F=2.6, P=0.002

Scaling focused on plant species scores (standardized) [0.623, 0.924]

- 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, **Snapshot** shows the Analysis Setup Wizard at the page where the type of permutation test is chosen. Here the user is selecting which covariates define permutation blocks. Additionally, permutations for hierarchical (split-plot / repeated-measures) design were selected and the hierarchy must be characterized on the following wizard page. including the correct setup for data coming from non-trivial sampling designs.



Analysis Setup Wizard: Testing and Stepwise selection

Test or Explore Predictor Effects

○ Not performed
● All constrained axes test
○ First constrained axis test
○ Both above tests performed
○ Forward selection of expl. variables
○ Summarize effects of expl. variables

Permutation Test Parameters

○ Unrestricted permutations
○ Time series or linear transects
○ Rectangular grids
● Hierarchical design
○ Read from file

Number of permutations: 499

Random number generator seeds: 23239 and 945

☑ Blocks defined by covariates
□ Leverage correction of residuals
□ Disable random shifts from mirror

Specification of Permutation Blocks

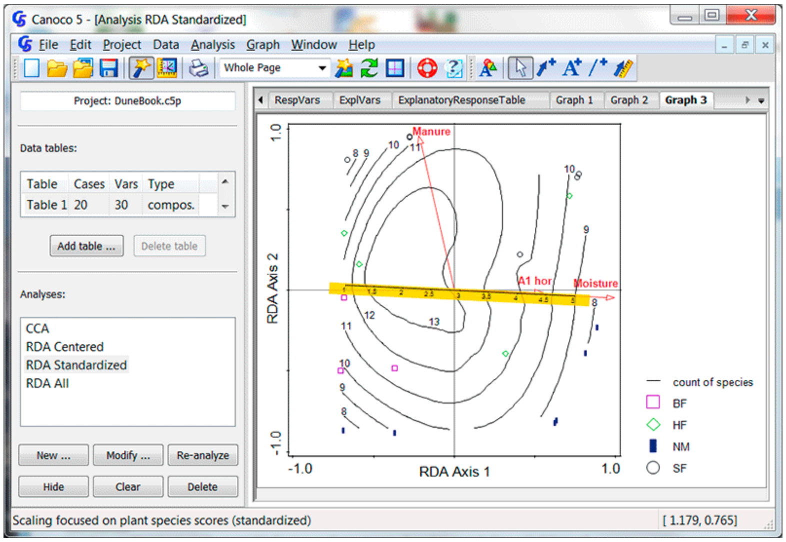
Move the covariates defining permutation blocks into right-hand list

Source variables:
Management
Use
Manure

Blocks defined by:

Order of covariates: ● original ○ alphabetical
Labels of covariates: ○ short ● full

- 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. **Snapshot** shows a diagram (attribute plot) where the variation in species richness across the ordination space is visualized using a loess smoother. Additionally, the arrows for explanatory variables are plotted, with one of the arrows (Moisture) calibrated with its original values. Individual cases are shown as symbols, classified by the management type.


Canoco 5 - [Analysis RDA Standardized]
File Edit Project Data Analysis Graph Window Help
Whole Page
RespVars ExplVars ExplanatoryResponseTable Graph 1 Graph 2 Graph 3
Add table ... Delete table
New ... Modify ... Re-analyze
Hide Clear Delete
1.0
-1.0
RDA Axis 2
-1.0
1.0
RDA Axis 1
Manure
A1 hor
Moisture
8
9
10
11
12
13
1.5
2
2.5
3
3.5
4
4.5
count of species
□ BF
◇ HF
▮ NM
○ SF

Project: DuneBook.c5p

Data tables:

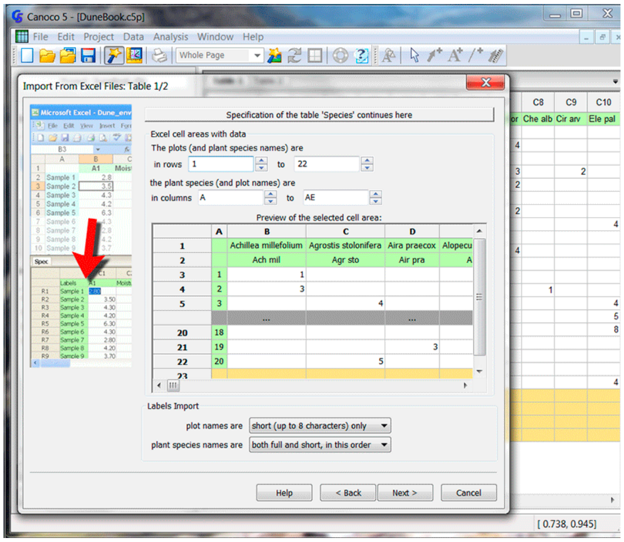
Table	Cases	Vars	Type
Table 1	20	30	compos.

Analyses:

CCA
RDA Centered
RDA Standardized
RDA All

Scaling focused on plant species scores (standardized) [1.179, 0.765]

- 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.
  Snapshot shows a page from the Excel Import wizard, where the range of input data and the multiplicity of row and column labels is specified. In the background, you can see the data notebook of the existing Canoco project. Data notebooks can be not only viewed, but also edited in a way similar to Excel spreadsheets.


Canoco 5 - [DuneBook.c5p]
File Edit Project Data Analysis Window Help
Whole Page
Microsoft Excel - Dune_env
C8 C9 C10
Che alb Cir arv Ele pal
[0.738, 0.945]

Import From Excel Files: Table 1/2

Specification of the table 'Species' continues here

Excel cell areas with data

The plots (and plant species names) are in rows 1 to 22

the plant species (and plot names) are in columns A to AE

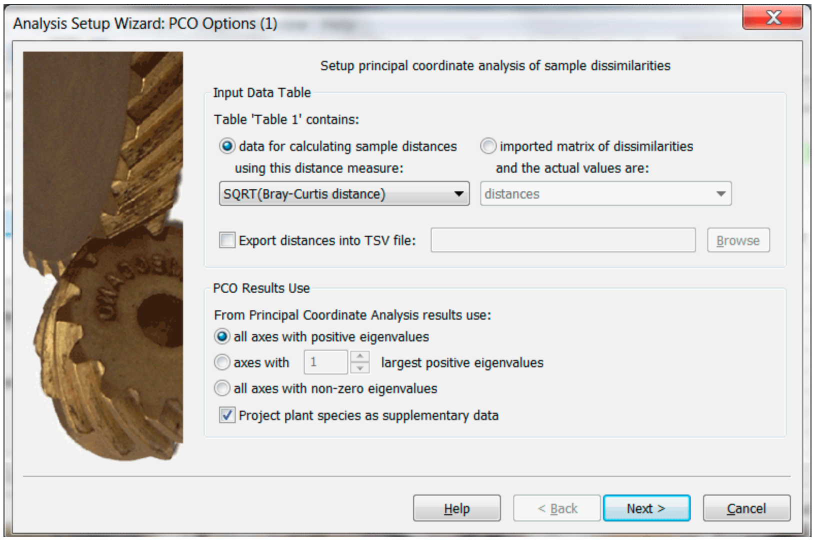
Preview of the selected cell area:

	A	B	C	D	
1		Achillea millefolium	Agrostis stolonifera	Aira praecox	Alopecu
2		Ach mil	Agr sto	Air pra	A
3	1	1			
4	2	3			
5	3		4		
		...		...	
20	18				
21	19			3	
22	20		5		
23					

Labels Import

plot names are: short (up to 8 characters) only

plant species names are: both full and short, in this order



Analysis Setup Wizard: PCO Options (1)

Setup principal coordinate analysis of sample dissimilarities

Input Data Table

Table 'Table 1' contains:

● data for calculating sample distances using this distance measure: SQRT(Bray-Curtis distance)

○ imported matrix of dissimilarities and the actual values are: distances

□ Export distances into TSV file:

PCO Results Use

From Principal Coordinate Analysis results use:

● all axes with positive eigenvalues

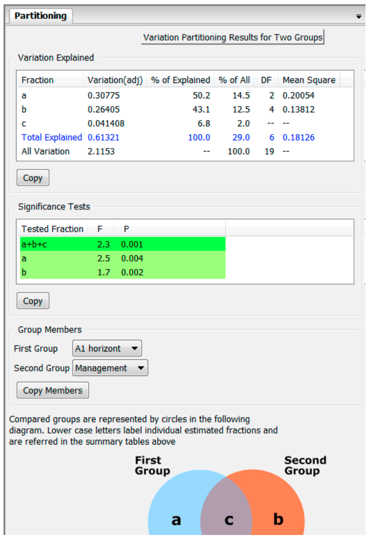
○ axes with 1 largest positive eigenvalues

○ all axes with non-zero eigenvalues

☑ Project plant species as supplementary data

•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). Similarly, non-metric multidimensional scaling is also supported.
Snapshot shows a specialized page shown in the Analysis Setup Wizard when you select an analysis template for principal coordinate analysis or distance-based RDA. On this page, you can select the distance type to calculate (or to use distances or similarities imported from elsewhere) and specify which scores to use in the following analysis steps.

- 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.
  Snapshot shows a specialized page shown in analysis notebook of a variation partitioning analysis. The variation in the response (community data) is explained by two groups of predictors and the partitioning procedure helps you to separate the unique contributions of these two groups (a, b) from their shared explanatory power. The individual fractions are described in the top table, while the central table contains tests of significance for the unique effects of the two groups and for their shared effect.



Partitioning

Variation Partitioning Results for Two Groups

Variation Explained

Fraction	Variation(adj)	% of Explained	% of All	DF	Mean Square
a	0.30775	50.2	14.5	2	0.20054
b	0.26405	43.1	12.5	4	0.13812
c	0.041408	6.8	2.0	--	--
Total Explained	0.61321	100.0	29.0	6	0.18126
All Variation	2.1153	--	100.0	19	--

Significance Tests

Tested Fraction	F	P
a+b+c	2.3	0.001
a	2.5	0.004
b	1.7	0.002

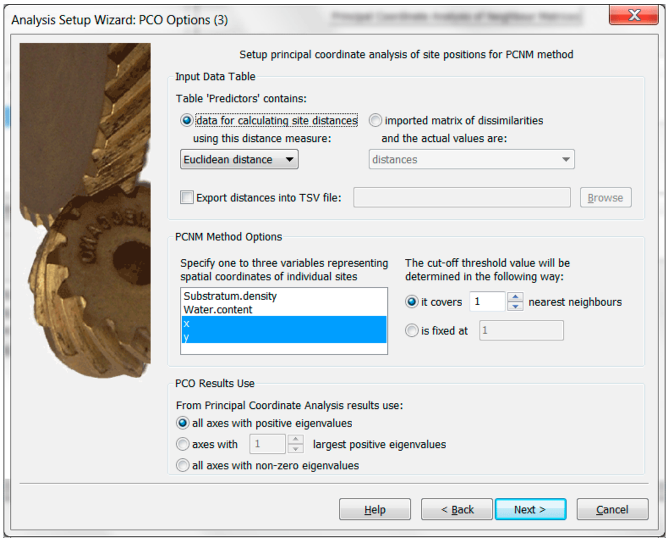
Group Members

First Group: A1 horizont

Second Group: Management

Compared groups are represented by circles in the following diagram. Lower case letters label individual estimated fractions and are referred in the summary tables above

- Principal coordinates of neighbor matrices (PCNM) methods are easily available within variation partitioning framework.
  Snapshot shows a specialized page of the Analysis Setup Wizard, where you define the coordinates of your points and choose options for PCNM.



Analysis Setup Wizard: PCO Options (3)

Setup principal coordinate analysis of site positions for PCNM method

Input Data Table

Table 'Predictors' contains:

● data for calculating site distances using this distance measure: Euclidean distance

○ imported matrix of dissimilarities and the actual values are: distances

□ Export distances into TSV file:

PCNM Method Options

Specify one to three variables representing spatial coordinates of individual sites

Substratum.density
Water.content
x
y

The cut-off threshold value will be determined in the following way:

● it covers 1 nearest neighbours

○ is fixed at 1

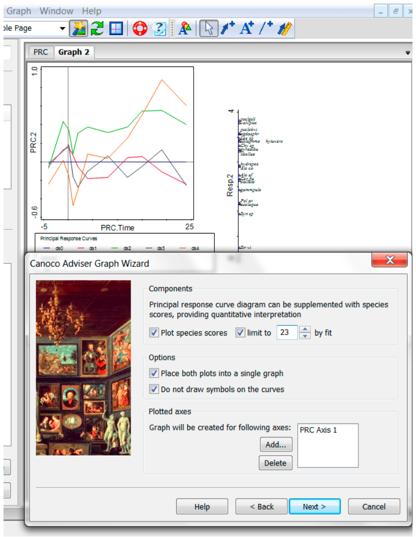
PCO Results Use

From Principal Coordinate Analysis results use:

● all axes with positive eigenvalues

○ axes with 1 largest positive eigenvalues

○ all axes with non-zero eigenvalues



Canoco Adviser Graph Wizard

Components

Principal response curve diagram can be supplemented with species scores, providing quantitative interpretation

☑ Plot species scores ☑ limit to 23 by fit

Options

☑ Place both plots into a single graph

☑ Do not draw symbols on the curves

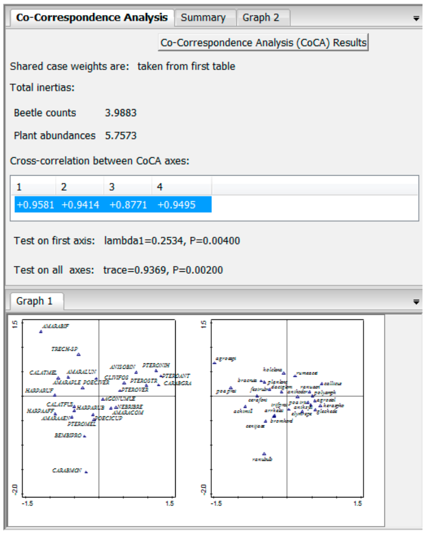
Plotted axes

Graph will be created for following axes: PRC Axis 1

- Computing, testing and graphing the Principal response curves (PRC) is now an easy task.

The PRC method is a very useful way to visualize and test the developing effects of external factors (toxic spill, ecosystem restoration actions, etc.) upon communities (response variables) across time. The above snapshot shows both the resulting diagram (in the top part) as well as the choices available in the Graph Wizard to setup a PRC diagram.

- Co-correspondence analysis (CoCA, symmetric form) is available, including Monte Carlo permutation testing.
  Co-correspondence analysis (CocA) allows you to relate the composition of two types of biotic communities, sampled at the same set of locations. The snapshot displays the results in the CoCA analysis notebook, both the dual ordination diagram (with plant and beetle communities) and the statistical summary, including Monte Carlo permutation test results.



Co-Correspondence Analysis (CoCA) Results

Shared case weights are: taken from first table

Total inertias:

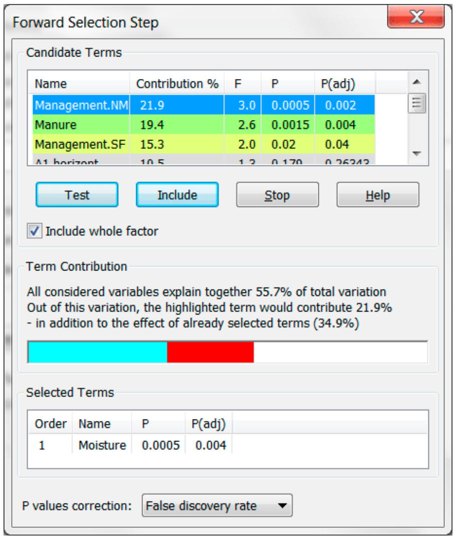
Beetle counts	3.9883
Plant abundances	5.7573

Cross-correlation between CoCA axes:

1	2	3	4
+0.9581	+0.9414	+0.8771	+0.9495

Test on first axis: lambda1=0.2534, P=0.00400

Test on all axes: trace=0.9369, P=0.00200



Forward Selection Step

Candidate Terms

Name	Contribution %	F	P	P(adj)
Management.NM	21.9	3.0	0.0005	0.002
Manure	19.4	2.6	0.0015	0.004
Management.SF	15.3	2.0	0.02	0.04
A1 horizont	10.5	1.3	0.170	0.26243

☑ Include whole factor

Term Contribution

All considered variables explain together 55.7% of total variation
Out of this variation, the highlighted term would contribute 21.9%
- in addition to the effect of already selected terms (34.9%)

Selected Terms

Order	Name	P	P(adj)
1	Moisture	0.0005	0.004

P values correction: False discovery rate

- 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).
  Stepwise selection of explanatory variables is a popular method how to sort-out results of exploratory field studies and Canoco 5 improves this procedure beyond the preceding versions by protecting users from most of the faults of stepwise selection from a large pool of predictors.

- Straightforward testing of all individual constrained axes as well as comparing results of two ordinations with Procrustes analysis is available.
  Snapshot illustrates both procedures, with the results of Procrustes rotation summarized both numerically and graphically in the lower part of the snapshot; the upper part reports on permutation tests of individual constrained (canonical) axes.

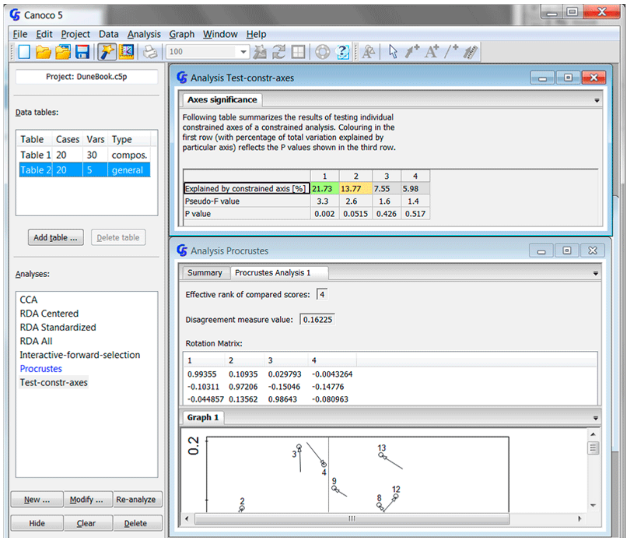


Table	Cases	Vars	Type
Table 1	20	30	compos.
Table 2	20	5	general

Following table summarizes the results of testing individual constrained axes of a constrained analysis. Colouring in the first row (with percentage of total variation explained by particular axis) reflects the P values shown in the third row.

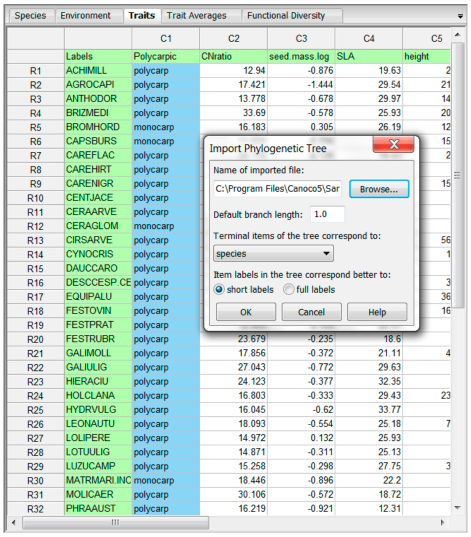
	1	2	3	4
Explained by constrained axis [%]	21.73	13.77	7.55	5.98
Pseudo-F value	3.3	2.6	1.6	1.4
P value	0.002	0.0515	0.426	0.517

Effective rank of compared scores: 4

Disagreement measure value: 0.16225

Rotation Matrix:

1	2	3	4
0.99355	0.10935	0.029793	-0.0043264
-0.10311	0.97206	-0.15046	-0.14776
-0.044857	0.13562	0.98643	-0.080963



		C1	C2	C3	C4	C5
	Labels	Polycarpic	CNratio	seed.mass.log	SLA	height
R1	ACHIMILL	polycarp	12.94	-0.876	19.63	2
R2	AGROCAPI	polycarp	17.421	-1.444	29.54	21
R3	ANTHODOR	polycarp	13.778	-0.678	29.97	14
R4	BRIZMEDI	polycarp	33.69	-0.578	25.93	20
R5	BROMHORD	monocarp	16.183	0.305	26.19	12
R6	CAPSBURS	monocarp				15
R7	CAREFLAC	polycarp				2
R8	CAREHIRT	polycarp				
R9	CARENIGR	polycarp				15
R10	CENTJACE	polycarp				
R11	CERAARVE	polycarp				
R12	CERAGLOM	monocarp				
R13	CIRSARVE	polycarp				56
R14	CYNOCRIS	polycarp				1
R15	DAUCCARO	polycarp				
R16	DESCCESP.CE	polycarp				3
R17	EQUIPALU	polycarp				36
R18	FESTOVIN	polycarp				16
R19	FESTPRAT	polycarp				
R20	FESTRUBR	polycarp	23.679	-0.235	18.6	
R21	GALIMOLL	polycarp	17.856	-0.372	21.11	4
R22	GALIULIG	polycarp	27.043	-0.772	29.63	
R23	HIERACIU	polycarp	24.123	-0.377	32.35	
R24	HOLCLANA	polycarp	16.803	-0.333	29.43	23
R25	HYDRVULG	polycarp	16.045	-0.62	33.77	
R26	LEONAUTU	polycarp	18.093	-0.554	25.18	7
R27	LOLIPERE	polycarp	14.972	0.132	25.93	
R28	LOTUULIG	polycarp	14.871	-0.311	25.13	
R29	LUZUCAMP	polycarp	15.258	-0.298	27.75	3
R30	MATRMARI.INC	monocarp	18.446	-0.896	22.2	
R31	MOLICAER	polycarp	30.106	-0.572	18.72	
R32	PHRAAUST	polycarp	16.219	-0.921	12.31	

Name of imported file: C:\Program Files\Canoco5\Sar

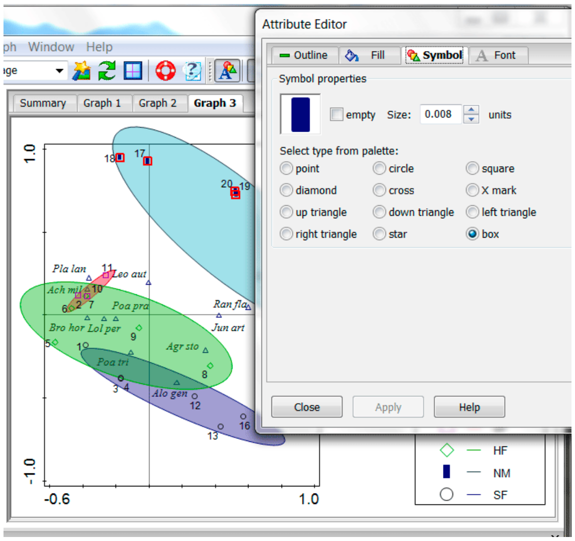
Default branch length: 1.0

Terminal items of the tree correspond to: species

Item labels in the tree correspond better to: ● short labels ○ full labels

•One can easily work with species functional traits or import the data on phylogenetic relatedness of species.
Snapshot of the Canoco 5 workspace displays (in the background) a data notebook with multiple data tables. The Species data table contains information about community composition, while the currently visible Traits table shows the known traits for individual species. You can also calculate, for individual cases, the averages of individual traits (implemented also for factorial traits) and analyse such table directly. Further, when doing analyses for individual community species acting as cases, it might be appropriate to perform phylogenetic correction during the tests. Canoco 5 supports this in the form of the Desdevises method and for this it allows importing a phylogenetic tree in Newick format and transforming it into a matrix of patristic distances, as illustrated by the dialog box in the front of the snapshot. In Canoco 5 you can also easily make nice biplots of analyses with traits and environmental variables.

- 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.
  Snapshot illustrates the enclosing ellipses created for four types of management, displaying their separation in ordination space. The ellipses are filled with semi-transparent colours, allowing one to read the other information within the diagram. The snapshot also displays a part of the Attribute Editor, which can be used to modify attributes of any component of Canoco 5 graphs.

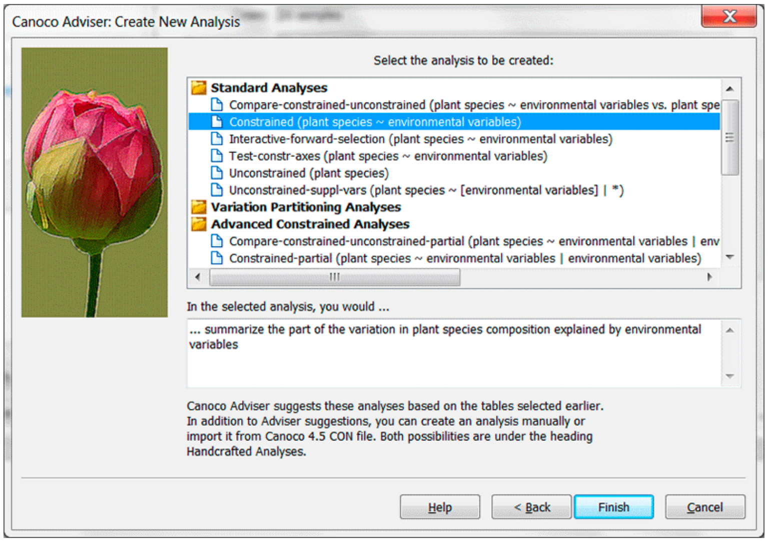


Symbol properties

☐ empty Size: 0.008 units

Select type from palette:

○ point ○ circle ○ square
○ diamond ○ cross ○ X mark
○ up triangle ○ down triangle ○ left triangle
○ right triangle ○ star ● box


Canoco Adviser: Create New Analysis

Select the analysis to be created:

Standard Analyses
- Compare-constrained-unconstrained (plant species ~ environmental variables vs. plant spe
- Constrained (plant species ~ environmental variables)
- Interactive-forward-selection (plant species ~ environmental variables)
- Test-constr-axes (plant species ~ environmental variables)
- Unconstrained (plant species)
- Unconstrained-suppl-vars (plant species ~ [environmental variables] | *)

Variation Partitioning Analyses

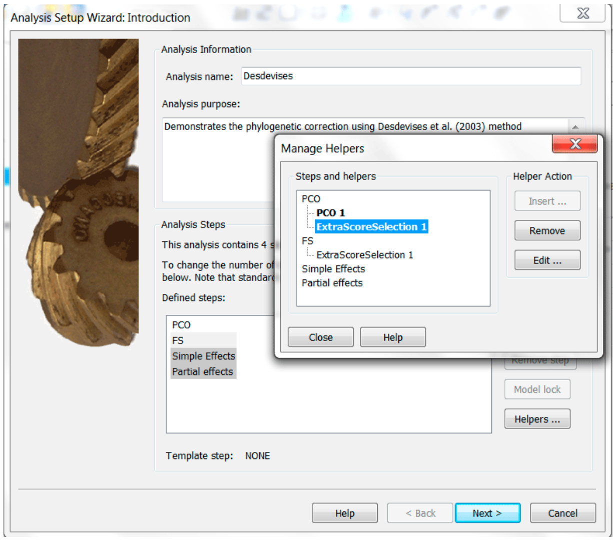
Advanced Constrained Analyses
- Compare-constrained-unconstrained-partial (plant species ~ environmental variables | env
- Constrained-partial (plant species ~ environmental variables | environmental variables)

In the selected analysis, you would ...

... summarize the part of the variation in plant species composition explained by environmental variables

Canoco Adviser suggests these analyses based on the tables selected earlier.
In addition to Adviser suggestions, you can create an analysis manually or
import it from Canoco 4.5 CON file. Both possibilities are under the heading
Handcrafted Analyses.

- Every step of your work is supported by an 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, correct type of ordination model (linear vs. unimodal), data transformation, or appropriate visualization of the results. It even advises you how to interpret ordination diagrams you create with the help of Graph Wizard.
  Snapshot shows the last page of the New Analysis Wizard, where - based on the inspection of the chosen data tables - the expert system of Canoco 5 (Canoco Adviser) suggests appropriate analyses. Canoco Adviser provides a helping hand also at other stages of your work with Canoco 5 (choosing linear vs. unimodal ordination method, choosing transformation for your response data as well as for the predictors, choosing appropriate diagram types for your analysis, interpreting graph contents, etc.)


Analysis Setup Wizard: Introduction

Analysis Information

Analysis name: Desdevises

Analysis purpose:

Demonstrates the phylogenetic correction using Desdevises et al. (2003) method

Analysis Steps

This analysis contains 4 s

To change the number of
below. Note that standar

Defined steps:

PCO
FS
Simple Effects
Partial effects

Template step: NONE

Manage Helpers

Steps and helpers

PCO
- **PCO 1**
- ExtraScoreSelection 1

FS
- ExtraScoreSelection 1

Simple Effects
Partial effects

•Advanced users can combine multiple methods in a single analysis, including generalized linear models to correlate scores etc.
Snapshot shows the first page of the Analysis Setup Wizard for a customized analysis, where you define individual analysis steps and optionally add specialized data handling for individual steps, with the help of so-called helpers. For standard analyses, such advanced work is not needed, but it gives you power and flexibility if you are an expert user of ordination methods

For details about ordering Canoco and its availability, please see our website:

[www.microcomputerpower.com]

The Canoco support website is at:

[www.canoco5.com]

General Canoco pages are at:

[www.canoco.com]