Biplots and partitioning trait-structured variation in dc-CA using Canoco 5.12

Cajo J.F. ter Braak April 2021 cajo.terbraak@wur.nl

Abstract. This note explains how the analyses of Figure 5 and Table 2 in Pinho et al. (2021, "Functional biogeography of Neotropical moist forests: trait-climate relationships and assembly patterns of tree communities", Global Ecology and Biogeography https://doi.org/10.1111/geb.13309) were obtained using Canoco 5.12. Figure 5a is a biplot of regression coefficients (CWM(traits)~climate) and Figures 5a and 5b together form a biplot of the CWM(traits) in the plots. Table 2 concerns variation partitioning in trait-environment analysis using double constrained correspondence analysis (dc-CA) with two sets of site predictor variables (climate and geography) and a single trait set. The specific aim of Figure 5 was to show more than just fourth-corner correlations, which is about the association between a single trait and a single environmental variable, by going to multi-trait, multi-environment associations (traits and composite traits along composite environmental gradients). The specific aim of Table 2 was to decompose the trait-structured variation, which is a weighted variance of community weighted means (CWMs) with respect to orthonormalized traits, on the basis of climate and geographic site predictors.

1. Biplot of coefficients of regression of CWMs on climate

Figure 5 in Pinho et al. (2021) used the plot-level data, whereas the public data are cluster-level data, in which the tree counts are summed across the plots of each cluster and the environmental variables are averaged. The cluster-level data were originally produced with the Canoco (ter Braak & Šmilauer 2018) using the *sum* option in *Project | Create derived project option | Aggregate cases*[1].

The first analysis in the Canoco project file (*dcCA climate*) is a dc-CA using the five climate variables and all six traits. Note that Canoco only allows you to perform a dc-CA after you added the transpose of the abundance table (the first table) to the project using *Data | Add new table(s) | Transpose compositional table...* There were therefore initially four data tables in the project. The project that you inspect, contains five data tables (see Fig. 3 on page 6) as we added one table (named *CWMwrtorthonormizedTraits*) that was derived from the first analysis, as explained in the section 2.2.

The *dcCA climate* analysis was obtained using *Analysis | Add new analysis | Canoco Adviser*, selecting the first four tables with focal table *AbundanceSum* and then selecting the *Double-constrained-CA (species ~ predictors x traits)* template from the *Trait Analyses | Double Constrained Analyses* section. The dc-CA analysis is visible only when the four table tables (**Y**, **E**, **T** and **transpose of Y**, say) that are required for this analysis were selected.

---

[1] Alternatively, use *average* in combination with the number of plots per cluster as *Case weights* in each analysis.

The first two graphs of this analysis (Fig. 1) are the cluster-level version of Figures 5a and 5b of Pinho et al (2021)

Figure 1. Ordination diagrams from double-constrained correspondence analysis (dc-CA) for 3,417 species across 59 spatial clusters of 471 Neotropical moist forest plots, showing (a) biplot of canonical weights of climate variables and scaled correlations of traits summarizing the coefficients of the multiple regressions of all CWMs of traits on the climate predictors; and (b) position (constrained scores) of samples (clusters) in the dc-CA biplot. Graphs (a) and (b) form a biplot of the CWMs of all clusters and traits. (eig = eigenvalue; rFC= fourth-corner correlation).
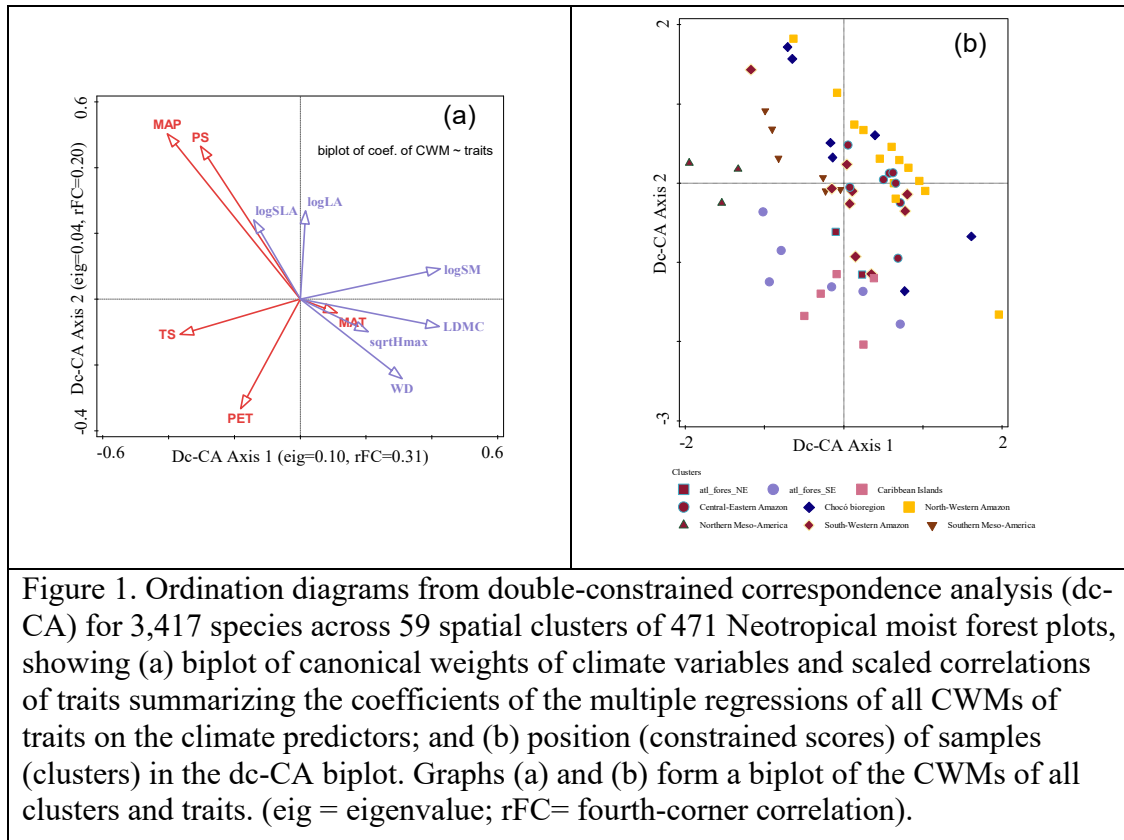
Fig 1a was obtained by *Graph | Advise on graphs...* and selecting *Regression biplot of CWM analysis with optional species scores* from the list of possible graphs. This graph option is available in Canoco 5.15 (once it is released[2]), but requires an edit of the *CanoWin5.clp* file in the Canoco 5.12 version. If you want to (re)create this type of plot, edit the file as described in the Appendix of this note (and rerun the analysis). This replacement allows the plotting of the regression weights of the predictor variables (the climate variables) together with the biplot scores of the traits[3]. This biplot was first proposed by ter Braak et al. (2018) in subsection "The biplot of **B** and **C**$_f$" of section 6.5 (page 192 or page 21 of the pdf and equations (48)-(49) with a summary below equation (51)). The regression weights are also known as canonical weights as they apply to the axes of a canonical (i.e. constrained) ordination.

---

[3] The scores can be inspected in Canoco by clicking 📚 and ⬇ and then selecting, from the drop down list, ExplVars(6) for the climate scores and SupplVars(6) for the traits. The regression weights for the climate variables are in the columns RegrE.1 and RegrE.2; the biplot scores for traits are in the columns BipS.1 and BipS.2. Note that ExplVars(6) also contains t-ratios for the regression weights (TvalE) and the biplot scores BipE.1 and BipE.2 for the climate variables. BipE and BipS constitute a biplot of the fourth-corner correlations between the climate and trait variables. Note that the biplot scores are proportional to the correlations (CorE and CorS) of the variables with the axes. The values in the score tabs differ in sign from those shown in the graph, because the second axis in the graphs was flipped by selecting *Analysis | Plot creation options* and clicking the second check box for *Flip axes*, so as to obtain a graph with maximum similarly to Fig. 5 of Pinho et al. (2021).

The biplot scores for the traits are the scaled correlations of the traits with the trait axes. Regression weights and biplot scores form a biplot of the coefficients of the weighted regression of each community-weighted mean trait on to the climate variables, in which each trait and climate variable is standardized using their weighted means (AvgE, AvgS) and standard deviation (SDE, SDS) in the scores tabs (see footnote 3 on page 2). Figure 5a in Pinho et al. (2021) is thus essentially a graphical summary of the coefficients in Table 1 (with the difference that the coefficients in Table 1 were obtained by applying mixed models and model averaging to the plot-level data, instead by weighted regression using cluster-level data). The graph was embellished by recolouring the trait labels by selecting the trait labels, clicking 🅰 and choosing the desired colour. Right clicking on a label allows one to edit the label, so as to remove the log and sqrt from the labels. Note that the canonical weights are also the coefficients that define how the climate variables are combined to define the axes of the graph.

The default graph in Canoco 5.12 is a graph of biplot scores only (footnote 3 on page 2) and provides a biplot of the fourth-corner correlations (Figure 2b). Note that MAP and PS change in position, showing that their regression coefficients differ considerably from their fourth-corner correlations. The regression coefficients are not unstable (no bouncing beta problem) and thus warrant interpretation as their variance inflation factors are not large (<5, namely 1.9 and 4.9, respectively, in the full data <3 according to Pinho et al. 2021) and their t-ratios are greater than 2 in absolute value (3.2-6.8 on the first two axes). Both Figures 2a and 2b show that a higher value of MAP thus corresponds to communities with tree species with lower values of the traits pointing approximately in the opposite direction (WD, LDMC), which are characteristic of conservative strategies. Higher PS (seasonality in precipitation) acts similar to MAP (Figure 2a), whereas its fourth-corner correlations differ from those of MAP (Figure 2b). Compared to methods that disregard correlations among the climate variables (notably RLQ), dc-CA is able to find a much stronger first gradient[4]. This gradient runs from communities with tree species with more opportunistic strategies (left) to communities with tree species with more conservative strategies (right).
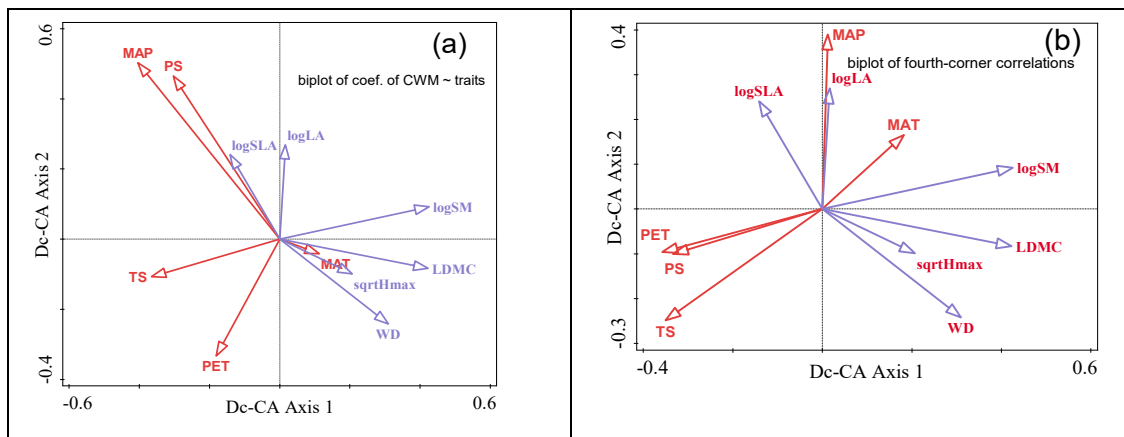

Figure 2. Biplot of traits with climate variables as canonical weights (a) and as biplot scores (b) approximating regression coefficients and fourth-corner correlations, resp.

---

[4] The first two eigenvalues of RLQ are 0.188 and 0.035, but these are not in terms of inertia of the abundance table. The squared fourth-corner correlations of trait and climate axes of first two axes are 0.06 and 0.02 in RLQ compared to 0.10 and 0.04 in dc-CA. The efficiency of RLQ compared to dc-CA (Peng et al. (2021) is thus only 0.65 for axis 1 and 0.61 for axis 2.

Fig. 1b, with clusters classified by biogeographic region, was obtained by using *Graph | Scatterplots | Clusters*. By default, Canoco plots the constrained scores as these are key to the fitted regression or ANOVA model.[5] The biogeographic region classification was created by *Project | Classifications | of clusters* and clicking *From data* with *Data Table* Environment and selecting Biogeographic region from the list of variables and, finally, clicking *Use selected classification in graphs*. To ensure that Figures 1a and 1b form a biplot, their type of scaling must agree. The scaling type of graph is shown in the bottom line of the Canoco window; it shows *Scaling of compromise type (biplot)* which is the default for trait-environment plots and, for Fig. 1b, must be set manually by clicking the scaling options 🖼 and selecting (under *Unimodal Methods*) *Symmetric scaling* and *biplot scaling* and unchecking the two check boxes for *Scale for optimal display* and *Use response variable contributions*.

By clicking ⟳ the graph is recreated with the new scaling. You can check the scaling also by right-clicking in the graph and selecting *Describe content*. With these details settled, Figures 1a and 1b form a biplot of the CWM traits in all clusters, where CWM trait is the community weighted mean of a trait (a 59 by 6 table of CWMs, similar to the plot-level CWMs in the file Data_Pinho_etal_GEB.xlsx at https://doi.org/10.5061/dryad.vq83bk3s3). This biplot of CWMs was proposed by ter Braak et al. (2018) in the ends of the first and last paragraphs of section 6.4 (pages 190/191 or pages 19/20 of the pdf and equations (18) and (42)-(44)).

---

[5] The constrained site scores are linear combinations of the climate variables (Eq (14) in ter Braak et al. (2018)). The unconstrained site scores (Eq. (12) in ter Braak et al. (2018)) are weighted averages of the (in dc-CA constrained) species scores and can be obtained by *Analysis | Plot creation options*, checking the *Use CaseR scores for constrained axes* option and recreating the graph.

2. Decomposition of trait-structured variation using dc-CA

2.1 Introduction

Variation partitioning aims to decompose the variation in a matrix, e.g. a site × species abundance table, explained by a set of predictors into parts. With two sets of predictor variables, sets A and B, there are three parts (Borcard, Legendre & Drapeau 1992; Okland & Eilertsen 1994; Peres-Neto *et al.* 2006):

a. variation is that is unique to A

b. variation is that is unique to B

c. shared variation

Part a is the part of the variation that can be explained by the predictors in set A but not by those in set B, part b is the part of the variation that can be explained by the predictors in set B but not by those in set A, and the shared variation is the variation that can be explained by both sets. The residual variation, variation that cannot be explained by the combined set of predictors can be considered as fourth part.

Here we consider partitioning of the trait-structured variation in trait-environment analysis using dc-CA and weighted redundancy analysis (RDA) so as to reproduce Table 1 which is a copy of Table 2 of Pinho et al. (2021).

In this table, we consider one set of traits and two sets of environmental variables and decompose the trait-structured variation in the parts due to two sets of environmental variables. The traits are the six functional traits logLA, logSLA, logSM, sqrtHmax, LDMC and WD. For the variation partitioning (without external computations by Excel or R) an initial dc-CA was performed using all trait variables. In general, it would be wise to use all environmental variables (thus including both the climate and geography variables), but this is not needed in this case as there are no missing values in the data and we have already the *dcCA climate* analysis.

Fig. 3 is a screenshot of the Canoco project *NeotropicTreeTraitEnv_dcCA.c5p* where you can find all entries of Table 1 (clicking the two Copy buttons copies the Variation Explained and the Significance Tests to the clipboard and can so be pasted in Excel, Word or similar). The *P*-values differ numerically, but not in meaning, as they are obtained by Monte Carlo permutation with a different number of permutations (this project used 999 random permutations).
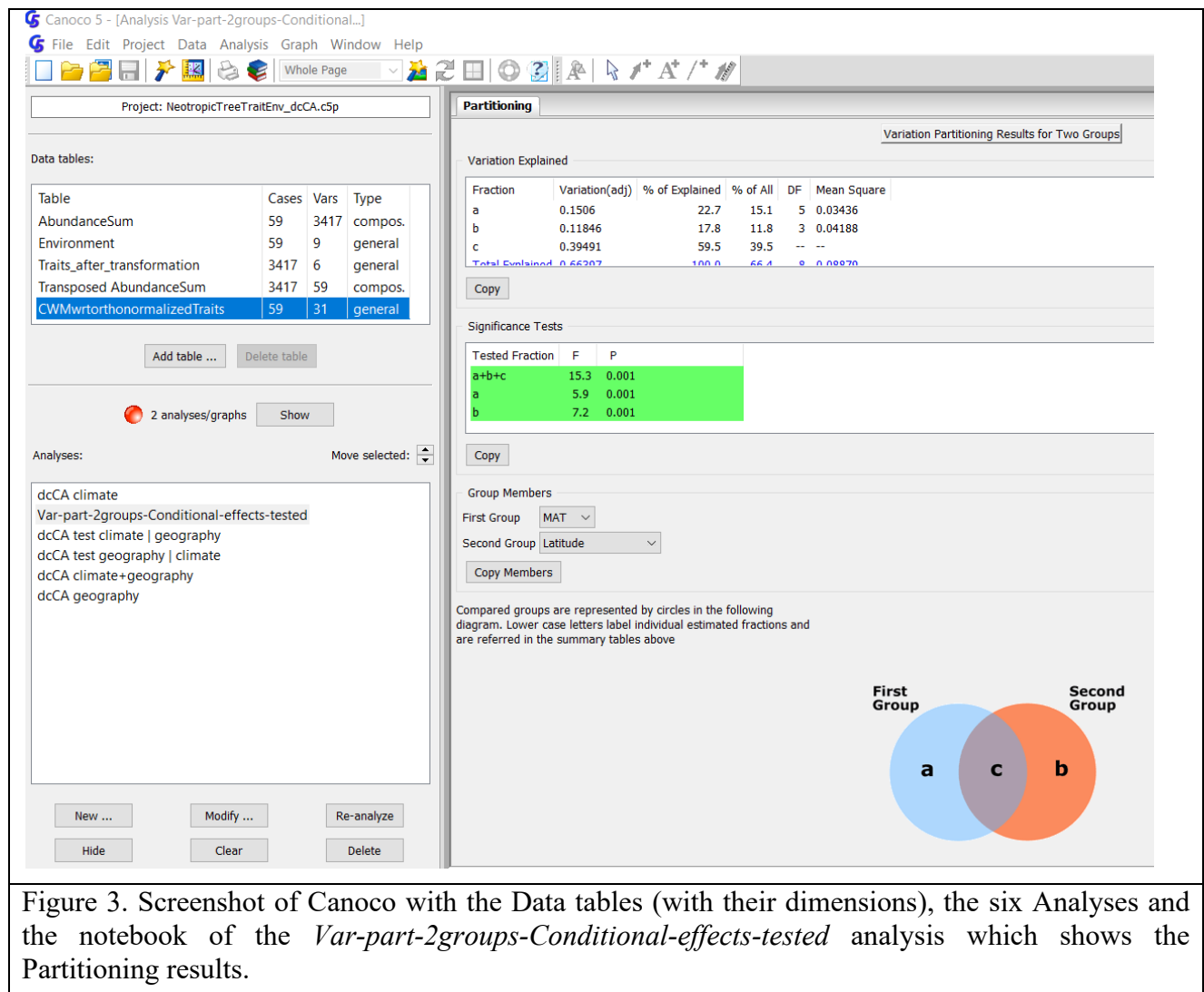
Section 2.2 gives a cookbook recipe of how perform the decomposition of the trait-structured variation. Section 2.3 provided the rationale for the recipe by outlining that a dc-CA can be seen as a combination of two analyses:

1. An analysis of the trait-structured variation in terms of the environment (schematically: CWM~E)

2. An analysis of the environmentally structured variation in terms of the traits (schematically: SNC~T).

Section 2.3 details all steps that are carried out in a dc-CA in Canoco (ter Braak & Šmilauer 2018).

Table 1. [Table 2 of Pinho et al 2021]. Variation partitioning of the trait-structured variation in the dc-CA with all traits, showing the unique and shared effects of geography (latitude, longitude, and degrees from equator) and climate (MAP, MAT, TS, PS, PET) in aggregated samples of tree communities across Neotropical moist forests (N = 59). The trait-structured variation is a weighted variance of the CWMs with respect to orthonormalized traits with the sample total as weight.

Component	Variation (Adj R2)	% of Explained	DF	Mean Square	F	P
Climate (unique)	0.15	22.7	5	0.03	5.9	0.0005
Geography (unique)	0.12	17.8	3	0.04	7.2	0.0045
Shared	0.39	59.5	--	--		--
Total Explained	0.66	100	8	0.09	15.3	0.0005

Figure 3. Screenshot of Canoco with the Data tables (with their dimensions), the six Analyses and the notebook of the *Var-part-2groups-Conditional-effects-tested* analysis which shows the Partitioning results.

2.2 How to perform the variation decomposition using Canoco 5.1x

The initial analysis *dcCA climate* in the Canoco project is a dc-CA with all traits and the five climate variables.

Step 2.1A: Place the RespVars(3) scores of the initial dc-CA to a data table

Display the scores of the *dcCA climate* analysis by clicking on the icon 📚.[6] The trait-structured variation is available as the first six columns of the *RespVars(3)* tab (six is the dimensionality of the traits; see *Summary* of step 6: *Supplementary vars: 6 traits [DF = 6])*. The columns *Resp.1 – Resp.6* are the community weighted means of the orthonormalized traits for each of the clusters (ClA1 – ClA59). For the decomposition of the trait-structured variation we must place these scores in a new data table, so that they can be analysed in the role of response data using any of the variation partitioning templates in Canoco. As we also need the column *RespW*, which contains cluster totals, it is most convenient to copy the whole sheet.

[motivation] In terms of the next section (section 2.3), $p^* = 6$ and the columns *Resp.1 – Resp.6* form **R*** and the column *RespW* is $\mathbf{w}_E$.

There are several ways how to place the scores into a new table. One convenient way is via Excel (or comparable spreadsheet program) and other is via Canoco only. The small advantage of the way via Excel is that one does not need to specify the number of the number of variables in the new Canoco data table.

For both ways:

After clicking the 📚 icon in the Canoco toolbar to show all score tabs, select the *RespVars (3)* tab. Copy all scores in the *RespVars (3)* table by right-clicking its empty top left corner and using the *Copy expanded* command. The advantage of *Copy expanded* over just *Copy scores* is that the copy includes any cases that were deleted from the analysis or cases that were not available for analysis due to missing values; such cases have NA values in the copied data.

Route via Excel (recommended)

Open Excel and open a blank workbook (*File|New|Blank* workbook) or blank sheet (click + in the one before bottom row of an existing workbook). Put the cursor in the A1 cell and paste (Ctrl-p) and save the workbook as Excel Workbook, say with the name `Rstar.xlsx`.

In Canoco, select *Data | Add new table(s) | Import from Excel* (the short-cut is with Alt pressed down: Alt-d,a,i, Enter). Now import the `Rstar.xlsx` (all columns or just the first $p^*+1=7$ columns, i.e. the first columns up to and including the column labelled *RespW*) as you import normally a data table or a project from Excel data.

In the Canoco project `NeotropicTreeTraitEnv_dcCA.c5p` the new data table was given the name `CWMwrtorthnormalTraits`.

---

[6] With Canoco versions before 5.15, you must make sure that the *Use response variable score* option in the *Edit scaling options* icon 🔲 in the Canoco toolbar is switched off (it is, if you never used it).

Route without Excel

Write down/recall the number of columns up to and including *RespW* in the *RespVars(3)* tab.

Select the *Data | Add new table(s) | Empty tables* menu command (Alt-d,a,e, Enter). The dialog box that appears asks to specify what the rows and columns represent: *Each table row represents a sample* (Case); *Each table column represents a*: type hereafter, for example, 'score', and in the next line 'scores', and give the table a name, e.g. CWMwrtorthnormalTraits. Click *Done* button and move[7] to the new table (it has a single numerical column C1 filled with 0 values), right click the C1 heading and choose *Insert scores from Clipboard*. This inserts all the columns you have copied to Clipboard before the zeroes and that last column can be then deleted[8].

Step 2.1B: Analyse the new table using a variation decomposition template

Perform a variation decomposition on the first $p*=6$ columns of the new table CWMwrtorthnormalTraits using weighted RDA with case weights from the $p*+1^{th}$ column of the table. To set the weights, the quick wizard mode must to off (use the 🪄 icon for this). Then start the Canoco Advisor in the usual way, for example via *Analysis | Add new analysis | Canoco Advisor* (Alt a,a, Enter) and deselect all data tables except the new table and the environmental data table. Go to the next advisor page and set the focal table to the new table. Go to the next advisor page and expand the section *Variation Partitioning Analyses*, where you select one template, depending on the analysis you want. In the example, the first template *Var-part-2groups-Conditional-effects-tested (scores~environmental variables)* is chosen and confirmed with *Finish*.

To replicate the partitioning analysis, put the five climate variables in the first group and the three geography variables in the second group. In the next page, select Set weights for cases ("clusters") and Select response variables ("scores"). In the next page, keep only the first $p*=6$ variables (*Resp.1*, ..., *Resp.6*). In the next page, click *Import weights* and select the *RespW* column in the CWMwrtorthnormalTraits table. Keep RDA in step 1, but **change** the ***Center and standardize*** to **Center.** No further changes to the defaults are needed.

Beware: do not automatically include all *Resp.** columns before the *RespW* column, because $p*$ might be a bit smaller. See Summary of step 6 - it says: *Supplementary vars: 6 traits [DF = 6]* but if there are factor variables it may say, for example, 10 traits [DF = 6].

---

[7] Alternatively, resize the table via *Data | Resize table*. It shows already the number of samples (Cases) as the rows were samples (Cases). But increase the number of columns from 1 to $p*+1$, in the example 6. Now copy just the first $p*+1$ columns from the RespVars(3) tab and paste these into the new, empty table. Ignore the warning that no row labels are copied.

[8] If you copied the whole RespVars(3) table with row names, you get a warning "One or more of the values cannot be stored in numeric variables [...]. Do you want to continue?". In this situation, click *Yes* and delete both the first column C1 which is completely yellow and empty and the last column that is filled with 0 values.

This results of the variation decomposition are in Table 1. Case-based tests of significance of the conditional effects A|B (read as: A conditional on B, *i.e.* giving the effects of A on the response, adjusted for the effects of B) and B|A are also present. Yet, to obtain the more reliable max test, a species-based test of significance is also needed. For this two separate dc-CA with environmental covariate data must be used, as demonstrated by the *dcCA test climate | geography* and *dcCA test geography | climate* analyses. The result of the max is listed on the *Double constrained CA* tab under the name *Combined test*. Both fractions a and b are significant as judged by the max test. Apparently the climate variables have an additional effect over what is coded as geography and reversely. Note also that the shared part is large.

Numerical check on the variation partitioning

A numerical check on the above procedure is as follows. The eigenvalues of the three RDAs in the analysis *Var-part-2groups-Conditional-effect-tested* should be equal to those of step 4 (*Explain constrained sample scores by environmental variables*) of the three dc-CAs. To check this, click on the analysis *Var-part-2groups-Conditional-effect-tested* so that its notebook is visible, click on the 🌍 icon and next on the *Summary* tab. This allows you to scroll through the three RDAs performed for this variation decomposition (A+B, A|B and B|A, called *SharedEffect* [a misnomer of their joint effect], *Group1Effect* and *Group2Effect*).

In the example, the eigenvalues of the

1. *SharedEffect* are 0.4205, 0.1883, 0.0729 and 0.017 which are identical to the first four eigenvalues of step 4 in the *dcCA climate+geography* analysis (the fifth analysis).
2. *Group1Effect* are 0.1045, 0.0499, 0.0115 and 0.0041 which are indeed equal to the first four eigenvalues of step 4 of the *dcCA test climate | geography* analysis.
3. *Group2Effect* are 0.0687, 0.0562, 0.0008 and 0.1049 which are indeed equal to the first fourth eigenvalues of step 4 of the *dcCA test geography | climate* analysis. Note that the fourth eigenvalue is that of an unconstrained axis.

An expert can also derive the components of the variance decomposition directly from the three dc-CA analyses without covariates (analysis 1, 5 and 6), using the Canonical eigenvalues (double-constrained) available by clicking *Details* in the *Summary* page of Step 6 of the dc-CAs or from those of the analyses 3, 4, 5.

Notes

1. Peres-Neto et al. (2006) proposed a permutational form of adjustment of R^2 for weighted analyses, for the reason that the standard adjustment is based on the fact that a random predictor has an average R^2 of $1/(n-1)$ in an unweighted analysis but that this quantity may be different in a weighted analysis (Peres-Neto et al. 2006). In the

permutational form, this quantity is estimated by a simulation in which the rows of the predictors are permuted. The permutational analysis is available in Canoco 5.1x by the checking the box Permutation-based adjustment of explained variation in *Edit | Settings | Canoco5options / Actions*. The permutation-based adjustment with the *Var-part-2groups-Simple-effects-tested* template follows the Peres-Neto et al. (2006) proposal exactly in the sense that it performs the permutations in analyses without covariates. Permutation-based adjustment with the *Var-part-2groups-Conditional-effects-tested* template performs the permutations in conditional analyses and may yield slightly different results. Both permutation-adjusted versions yield adjusted R^2 values that are very close to those of Table 1, despite the fact that the analysis uses highly unequal weights.

2. Decomposition of the environmentally structured variation by two sets of traits works analogously. It requires a weighted RDA of **S*** on the sets of traits, with **S*** defined in section 2.3. **S*** is available in the *RespVars(1)* tab of an initial dc-CA that then must contain all environmental variables you want to include in the decomposition.

2.3 Variation decomposition in dc-CA using weighted RDA

Recall from Peng et al. (2021) the 6 step algorithm of the dc-CA analysis and its motivation:

Without any statistical testing and variation decomposition, the analysis [i.e. dc-CA] can be performed in a single step (a singular value decomposition), but often it is also of interest how well the traits and the environmental variables each explain the community data using single constrained ordinations (steps 1 and 3 below) and how well the environmental variables explain the trait-structured variation and how well the traits explain the environmentally structured variation (Peres-Neto et al., 2017). The full analysis thus needs four steps. Two more steps are required a) to determine what the maximum strength of the trait-environment relation could have been for a given abundance matrix **Y** (step 5 below) and b) to glue the results of all steps together (step 6). The six steps for the dc-CA are given below:

1. The first step is a canonical correspondence analysis (CCA; **Y~E**) constraining the species composition by environmental variables (**E**). This step allows to determine how much variation in the abundance values can be explained by the environmental variables. It also gives species scores (**S***) to be used in the next step; **S*** is an $m \times q^*$ table of scores with q^* the rank of the environmental variables and its entries are species niche centroids (SNC, Peres-Neto et al., 2017) of the orthonormalized environmental variables. **S*** represents the environmentally structured variation in the abundance table and is available as the first q^* columns in the *RespVars(1)* scores tab in a dc-CA analysis.

2. The second step is a weighed RDA (**S*~T**), with the traits in **T**, which turns the single constrained ordination of step 1 into a double constrained ordination (Ter Braak et al. 2018a). This step is a species-level analysis, which gives a significance test of whether the traits modulate the species-environment relationship and a measure of how well the traits explain the environmentally

structured variation, namely the percentage variance in the species niche centroids (in S*) that is explained by the traits.

3. The third step is a CCA ($\mathbf{Y}^\mathrm{T}$~$\mathbf{T}$) constraining the transposed species composition by traits. This step allows to determine how much variation in the abundance values can be explained by the traits. It also gives case scores ($\mathbf{R}^*$) to be used in the next step; $\mathbf{R}^*$ is an $n \times p^*$ table of scores with p^* the rank of the trait data and its entries are community weighted means (Peres-Neto et al. 2017) of the orthonormalized traits. $\mathbf{R}^*$ represents the trait-structured variation in the abundance table and is available as the first p^* columns in the *RespVars(3)* scores tab in a dc-CA analysis.

4. The fourth step is a weighted RDA ($\mathbf{R}^*$~$\mathbf{E}$), which turns the single constrained ordination of step 3 into a double constrained ordination that is identical to the one obtained in step 2 (Ter Braak et al. 2018a). This step is a case-level (community-level) analysis, which gives a significance test of whether the environmental variables modulate the species-trait relationship and a measure of how well the environmental variables explain the trait-structured variation, namely the percentage variance in the community weighted means (in $\mathbf{R}^*$) that is explained by the environmental variables.

5. The fifth step is a CA of $\mathbf{Y}$ to find scores for cases and species that maximize the fourth-corner correlation without any constraints on rows and columns of $\mathbf{Y}$.

6. The last step is used for calculating the dc-CA scores from steps 1-4, in which all resulting summary statistics and ordination scores are collected. The number of double constrained axes is equal to the minimum of p^* and q^*.

For variation decomposition using two sets of environmental variables $\mathbf{A}$ and $\mathbf{B}$, we need to decompose the trait-structured variation in the abundance table, i.e. we need to decompose the variation in $\mathbf{R}^*$ due to $\mathbf{A}$ and $\mathbf{B}$. This will require three RDA analyses, either three RDA with formulas $\mathbf{R}^*$~$\mathbf{A}$+$\mathbf{B}$, $\mathbf{R}^*$~$\mathbf{A}|\mathbf{B}$ and $\mathbf{R}^*$~$\mathbf{B}|\mathbf{A}$ or three RDA with formulas $\mathbf{R}^*$~$\mathbf{A}$+$\mathbf{B}$ with $\mathbf{R}^*$~$\mathbf{A}$ and $\mathbf{R}^*$~$\mathbf{B}$. Both versions thus allow variation decomposition, but the former set allows testing of the conditional effects ($\mathbf{A}|\mathbf{B}$ and $\mathbf{B}|\mathbf{A}$), whereas the latter allows testing of simple or marginal effects ($\mathbf{A}$ and $\mathbf{B}$). To get the same decomposition as the decomposition based on (three analyses using) single step dc-CA, each RDA must be weighted by the sample totals $\mathbf{w}_\mathrm{E}$ and must only be centred (even centering is not strictly needed as $\mathbf{R}^*$ is already centred using these weights). The sample totals $\mathbf{w}_\mathrm{E}$ are available in the *RespW* column of the *RespVars(3)* tab. Variation decomposition of dc-CA can thus be obtained by variance decomposition using weighted RDAs as detailed as recipe in section 2.2. Decomposition of the environmentally structured variation by two sets of traits works analogously. It requires a weighted RDA of $\mathbf{S}^*$ on the sets of traits.

3. Useful links

www.canoco.com

www.canoco5.com

canoco5.com/index.php/resources

www.microcomputerpower.com

4. Appendix

This appendix applies to Canoco 5.12 only. Backup the existing *CanoWin5.clp* file in the Canoco 5 installation folder, e.g. by copying the file and renaming the copy to *CanoWin5_old.clp* and then edit the file *CanoWin5.clp* by inserting the text below. The insert must be after line 10063 reading "(create$". Do not forget to include the closing right parenthesis ")", which matches the first left parenthesis "(" of the insertion. In summary, insert the following 11 lines after line 10063 in *CanoWin5.clp*:

```
(make-instance of ADVISED-GRAPH
 (plot-type   ordin)
 (graph-name  (format nil "REGRESSION BIPLOT of CWM analysis with optional %s scores"
           ?singCol))
 (graph-desc  (format nil
           "summarize the regression of CWM of %s onto %s"
           ?plurColTraits ?plurColExpl))
 (required    7 3)
 (optional    1)
 (subsetable  1)
 )
```

Note that this insertion adds the regression biplot of the CWM analysis only to the first of the set of double-constrained analysis templates offered by the *Canoco Adviser,* i.e. excluding those including forward selection of environmental predictors and traits or testing individual axes or the double-constrained analyses based on a linear model (dc-PCA). The graph will be available for all those analysis templates in Canoco version 5.15, once it is released.

5. References

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

Okland, R.H. & Eilertsen, O. (1994) Canonical correspondence analysis with variation partitioning: some comments and an application. *J. Veg. Sci.,* **5,** 117-126.https://doi.org/10.2307/3235645

Peng, F.-J., ter Braak, C.J.F., Rico, A. & Van den Brink, P.J. (2021) Double constrained ordination for assessing biological trait responses to multiple stressors: A case study with benthic macroinvertebrate communities. *Science of the Total Environment,* **754,** 142171.https://doi.org/10.1016/j.scitotenv.2020.142171

Peres-Neto, P.R., Legendre, P., Dray, S. & Borcard, D. (2006) Variation partitioning of species data matrices: Estimation and comparison of fractions. *Ecology,* **87,** 2614-2625.https://doi.org/10.1890/0012-9658(2006)87[2614:VPOSDM]2.0.CO;2

ter Braak, C.J.F. & Šmilauer, P. (2018) *Canoco reference manual and user's guide: software for ordination (version 5.10).* Microcomputer Power, Ithaca, USA

ter Braak, C.J.F., Šmilauer, P. & Dray, S. (2018) Algorithms and biplots for double constrained correspondence analysis. *Environmental and Ecological Statistics,* **25,** 171-197.https://doi.org/10.1007/s10651-017-0395-x