Testing environmental effects on taxonomic composition with canonical correspondence analysis: alternative permutation tests are not equal

Cajo J. F. ter Braak[1] · Dennis E. te Beest[1]

Abstract

After applying canonical correspondence analysis to metagenomics data with hugely different library sizes (site totals) it became evident that Canoco and the R-packages ade4 and vegan can yield (at least up to 2022) very different *P*-values in statistical tests of the relationship between taxonomic composition (species composition) and predictors (environmental variables and/or treatments). The reason is that vegan and Canoco up to version 5.12 apply residualized response permutation (but ignore the model intercept), whereas ade4 applies predictor permutation. Predictor permutation, when extended to residualized predictor permutation, is applicable in partial constrained ordination. This paper shows by simulation that residualized response permutation can yield a very inflated Type I error rate, if the abundance data are both overdispersed and highly variable in site total. In contrast, residualized predictor permutation controlled the type I error rate and had good power, also when the predictors were skewed or binary. After square-root or log transformation of the abundance data, the differences between the permutation methods became small. Residualized predictor permutation is recommended, particularly in testing trait–environment relationships using double constrained correspondence analysis, because this method also critically depends on the species totals, which are generally highly variable. It is implemented in Canoco 5.15 and the R-code of this paper.

Keywords ade4 · Canoco · Canonical correspondence analysis · Microbiome data · Permutation testing · vegan

---

Handling Editor: Luiz Duczmal.

✉ Cajo J. F. ter Braak
    cajo.terbraak@wur.nl

1 Biometris, Wageningen University & Research, Wageningen, The Netherlands

1 Introduction

Canonical correspondence analysis (CCA) was the first method of canonical or constrained ordination to relate ecological community data to environmental and other predictor variables (ter Braak 1986, 2014; Borcard et al. 2011; Legendre and Legendre 2012). Although many alternatives have been developed since (Legendre and Legendre 2012), CCA remains attractive as it is an eigenvector method that (1) is able to analyze unimodal response of multiple species to environmental gradients—it was derived by ter Braak (1986) as an approximation to maximum likelihood fitting of a Gaussian species packing model (see Liu et al. (2020) for a recent account in the context of paleo-environment reconstruction) and (2) is suitable for analyzing environmental effects on count data, particularly for strictly compositional data in which the site total abundance is irrelevant to the research question (Greenacre 2018). The first point is also highlighted by ter Braak (1987). Using a non-parametric species packing model, ter Braak (1987) showed that CCA finds the environmental gradient that maximally separates the species niches. The environmental gradient herein is constrained to a linear combination of environmental variables. The second point follows from the derivation of (canonical) correspondence analysis as an approximation to fitting a Poissonian log-linear regression model for small deviations from row-column independence (Ihm and van Groenewoud 1984; Goodman 1986; see Appendix A1). As such, CCA takes a position in-between distance-based (McArdle and Anderson 2001) and recent model-based methods (Warton et al. 2014; Niku et al. 2019). From the transition formulas of CCA (e.g. ter Braak 1986; Appendix A1.1.2) it can be deduced that CCA is insensitive to zero-inflation in the abundance data. The abundance acts as a weight in these formulas and zero abundance thus does not count. Also, CCA is relatively insensitive to the site and species totals: where abundance appears in these formulas, the abundance is either divided by the site total or by the species total. As CCA can also be seen as a multivariate regression method (ter Braak and Verdonschot 1995), it is thus suited for detecting effects of predictors (typically treatments or environmental variables) on strictly compositional response data with many zeroes (typically species abundance data and microbiome data) for both long and short environmental gradients. A Google search (with quotes) on "canonical correspondence analysis" returned about 435.000 results, whereas the related methods of "redundancy analysis" and "canonical correlation analysis" both returned about a million.

In the context of trait–environment relationship, CCA has been extended to double constrained correspondence analysis (dc-CA) (ter Braak et al. 2018; Gobbi et al. 2022) as the method which finds linear combinations of traits and environmental variables that maximize the fourth-corner correlation, originally developed for single trait and single environmental variables (Legendre et al. 1997; Dray and Legendre 2008); dc-CA also generalizes Community Weighted Mean (CWM) trait analysis to multiple traits (Pinho et al. 2021). Moreover, the fitted inertia of CCA and dc-CA was recognized as the Rao-score test statistic of a Poissonian log-linear model in testing the species-environment and trait–environment relationships, respectively (ter Braak 2017). Score tests are asymptotically equivalent with likelihood-ratio tests and thus asymptotically most powerful. In practice, however, ecological data are often overdispersed compared to the Poisson distribution. This means that other test statistics

may be more powerful, and, more importantly, the asymptotic distribution of the inertia (times the total abundance) is not chi-square distributed, as it would be in a test of association in a contingency table. This is why permutation methods are used in practice for statistical testing.

CCA included almost from its inception Monte Carlo permutation methods for testing the statistical significance of the species-environment relation, *i.e.* to determine whether the found relation was strong in comparison to the relation found by chance. The method of permutation changed from predictor permutation (permutation of sites in the table of predictors) in the Canoco software version 2 (ter Braak 1988) to, what could be called, response permutation in later versions (ter Braak 1990; ter Braak and Šmilauer 2018), where, instead of the original response data (the abundance data), the residuals under the null hypothesis are permuted. Both permutation methods are equivalent in unweighted analyses, such as in redundancy analysis (Legendre and Legendre 2012), but are no longer equivalent in partial constrained ordination (constrained ordination conditioned on covariates) or in weighted analyses as the residuals are also standardized in this method (ter Braak 2022). It is well known that CCA is a weighted analysis as it can be carried out by a weighted redundancy analysis of transformed data, in which the weights are the margin totals of the abundance table (ter Braak and Verdonschot 1995; Legendre and Legendre 2012). The R packages ade4 (Thioulouse et al. 2018) and vegan (Oksanen et al. 2022) use predictor and residualized response permutation, respectively, in the functions randtest.pcaiv and anova.cca, respectively (Table 1). Consequently, Canoco and vegan can yield results that differ from those of ade4. This difference has gone unnoticed until we discovered it when we applied CCA using both ade4 and vegan as an alternative to log-ratio analysis to metagenomics data with huge differences in site totals (te Beest et al. 2021). We found that CCA using Canoco and vegan could yield grossly inflated type I error rates (up to 1!) in scenarios where CCA using ade4 adhered to the nominal significance level. We support the plea in Muff et al. (2021) to report actual *P*-values and levels of evidence instead of binary decision (significant, not significant). For this to be meaningful, *P*-values, also the ones obtained by permutation, should be correct.

This paper reports on the evaluation of the permutation methods in current use in CCA and of a neglected permutation method that naturally extends the predictor permutation method so that it can be used in partial constrained ordination; it is termed residualized predictor permutation (ter Braak 2022). The evaluation is mainly by simulation and application to real data. R-code of all methods and simulations is supplied in the Supporting Information and all permutation methods are implemented in Canoco version 5.15. CCA is available in a number of software packages, but we study the permutation tests implemented in ade4, Canoco and vegan only.

2 Methods

After introducing CCA as a form of weighted multivariate regression, in particular as redundancy analysis of transformed data (contingency ratios), this section describes residualized predictor permutation and residualized response permutation. Tables 1 and 2 summarize these permutation methods and the software and versions thereof

Table 1 Summary of methods and performance with software versions

	Permutation method		
	Residualized		
	Predictor permutation	Response permutation	
Abbreviation	RPP	RRP	LRRP*
What is permuted	Predictor residuals	Response residuals	
Assumption	Design-based randomization; no standardization of residuals	Standardization of residuals based on properties of a contingency table (independent Poisson counts)	As RRP
Software tested	ade4::randtest.pcaiv version 1.7-18** Canoco version 5.15	Canoco version 5.15	vegan::anova.cca version 2.6-2 Canoco version 3-version 5.12
Type I error inflation	None detected	Yes, if site totals are highly variable and abundance is overdispersed	Yes, slightly worse than RRP
Recommended transformation when both site variability and overdispersion are high	Square-root to increase the power in testing dimensionality (Fig. 3) and avoid sensitivity to environmental main effects (Fig. 4)	Square-root or log so as to minimize the danger of Type I error inflation (Figs. 1, 2, 3, 4)	

*Legacy Residual Response Permutation, a method which wrongly ignores the intercept during permutation. The intercept can be ignored only if the site totals are all equal

**A CCA in ade4 is available via the function pcaiv; to our knowledge, ade4 does not allow for non-trivial covariates in CCA (partial CCA)

that we compare. The section continues with the description of the simulation study and real data to evaluate these permutation methods.

We use a notation in which $\mathbf{Y} = [y_{ij}]_{n \times m}$ is the site-by-species abundance table of response data, $\mathbf{X} = [x_{ik}]_{n \times p}$ is the site-by-variable table of p environmental variables (predictors) and $\mathbf{Z} = [z_{il}]_{n \times (q+1)}$ is the site-by-covariate table of q covariate (conditioning) variables with a column of ones for the intercept. For notational convenience and without loss of generality, the abundance values are already preprocessed by division by their sum, so that the sum of all values in $\mathbf{Y}$ is 1. Also, $\mathbf{r}$ and $\mathbf{k}$ are n- and m-column vectors containing the marginal totals of $\mathbf{Y}$, and $\mathbf{R}$ and $\mathbf{K}$ are diagonal matrices with $\mathbf{r}$ and $\mathbf{k}$ on their diagonal, respectively. Finally, $\pi(.)$ denotes a random permutation of the rows of a matrix.

Table 2 Summary of methods using matrix algebra

	Permutation method*	
	Residualized	
	Predictor permutation (RPP)	Response permutation (RRP)**
Model	$\mathbf{C} = \mathbf{R}^{-1}\,\mathbf{Y}\,\mathbf{K}^{-1} \sim \mathbf{Z} + \mathbf{X}$	
What is permuted?	Predictor residuals	Response residuals
Residuals from WLS model	$\mathbf{X} \overset{\sim}{\cdot} \mathbf{Z}$	$\mathbf{C} \overset{\sim}{\cdot} \mathbf{Z}$
Notation for residuals (**Errors**)	$\mathbf{E}_{\mathbf{X}\vert\mathbf{Z}}$	$\mathbf{E}_{\mathbf{C}\vert\mathbf{Z}}$
Assumption	Design-based randomization	Model-based with $\mathrm{var}(\mathbf{E}_{\mathbf{C}\vert\mathbf{Z}}) \propto (1/\mathbf{rk}')$
Standardization of residuals	No	Yes to $\mathbf{R}^{1/2}\mathbf{E}_{\mathbf{C}\vert\mathbf{Z}}\mathbf{K}^{1/2}$ De-standardize with $\mathbf{R}^{-1/2}$ in WLS version
Transformation of **X** and **Z**	No	Yes to $\mathbf{R}^{1/2}\mathbf{X}$ and $\mathbf{R}^{1/2}\mathbf{Z}$
WLS version:		
Null model	$\mathbf{C} \overset{\sim}{\cdot} \mathbf{Z}$	$\mathbf{R}^{-1/2}\pi(\mathbf{R}^{1/2}\mathbf{E}_{\mathbf{C}\vert\mathbf{Z}}\mathbf{K}^{1/2}) \overset{\sim}{\cdot} \mathbf{Z}$
Alternative model	$\mathbf{C} \overset{\sim}{\cdot} \mathbf{Z} + \pi(\mathbf{E}_{\mathbf{X}\vert\mathbf{Z}})$	$\mathbf{R}^{-1/2}\pi(\mathbf{R}^{1/2}\mathbf{E}_{\mathbf{C}\vert\mathbf{Z}}\mathbf{K}^{1/2}) \overset{\sim}{\cdot} \mathbf{Z} + \mathbf{X}$
OLS version:		
Null model	$\mathbf{R}^{1/2}\mathbf{C}\mathbf{K}^{1/2} \sim \mathbf{R}^{1/2}\mathbf{Z}$	$\pi(\mathbf{R}^{1/2}\mathbf{E}_{\mathbf{C}\vert\mathbf{Z}}\mathbf{K}^{1/2}) \sim \mathbf{R}^{1/2}\mathbf{Z}$
Alternative model	$\mathbf{R}^{1/2}\mathbf{C}\mathbf{K}^{1/2} \sim \mathbf{R}^{1/2}\mathbf{Z} + \mathbf{R}^{1/2}\pi(\mathbf{E}_{\mathbf{X}\vert\mathbf{Z}})$	$\pi(\mathbf{R}^{1/2}\mathbf{E}_{\mathbf{C}\vert\mathbf{Z}}\mathbf{K}^{1/2}) \sim \mathbf{R}^{1/2}\mathbf{Z} + \mathbf{R}^{1/1/2}\mathbf{X}$

*Notation:
Data: $\mathbf{Y}_{n\times m}$: taxonomic composition data of m taxa in n sites; $\mathbf{X}_{n\times p}$: data on p predictors under test; $\mathbf{Z}_{n\times(q+1)}$: data on q covariates, i.e. predictors to adjust for, including a column for the intercept;
Derived data: $\mathbf{R}$: the diagonal matrix with row totals of $\mathbf{Y}$; $\mathbf{K}$: the diagonal matrix with column totals of $\mathbf{Y}$; $\mathbf{r} = \mathrm{diag}(\mathbf{R})$; $\mathbf{k} = \mathrm{diag}(\mathbf{K})$;
Model formulas:. $\sim$. denotes that the left-hand argument is regressed using ordinary least-squares (OLS);•$\overset{\sim}{\cdot}$• indicates that the left-hand argument is regressed using weighted least-squares (WLS) with row weights $\mathbf{r}$ and, for $\mathbf{C}$, also column weights $\mathbf{k}$ using the model formula on the right-hand side;
Permutation: $\pi(.)$: a permutation of the rows of a matrix
**Note that the fit of $\mathbf{C}$ due to $\mathbf{Z}$ drops out of the F-ratio for the effect of $\mathbf{X}$ as it appears in both the null model and the alternative model

2.1 A linear model for CCA

Whereas CCA was introduced as a method to discover unimodal response of species to environment (ter Braak 1986; ter Braak and Verdonschot 1995), it is also useful as a method to detect small log-linear environmental effects on species in count-like data. In Appendix A1 we show the link between these models. In this context, the abundance table is like a contingency table with expected values under the independence model giving by the formula $r_i k_j$ (row total r_i times column total k_j divided by the overall total, which is 1, because $\mathbf{Y}$ was divided by its total). CCA then fits a linear model to the contingency ratios $\mathbf{C} = [y_{ij}/(r_i k_j)]$ by weighted least-squares (WLS) with row

and column weights $\mathbf{r}$ and $\mathbf{k}$, respectively, and model formula $\mathbf{C} \sim \mathbf{Z} + \mathbf{X}$ (ter Braak and Verdonschot 1995). With $\mathbf{Z}$ containing a vector of ones for the intercept only, *i.e.* without covariates, the model can thus be written as

$$c_{ij} = \frac{y_{ij}}{r_i k_j} = a_{0j} + b_{j1}x_{i1} + \cdots + b_{jp}x_{ip} + e_{ij}, \tag{1}$$

where the variance of the error e_{ij} is assumed to be proportional to $1/(r_i k_j)$, which is the variance of the contingency ratio c_{ij} under row-column independence when y_{ij} is Poisson distributed (ter Braak and Šmilauer 2018). The weights used in CCA follow from this variance (for other pertinent motivations, which require a long environmental gradient, *i.e.* a strong unimodal response, instead of row-column independence, see ter Braak 1986, 1987). But note that the variance of the contingency ratio may differ from $1/(r_i k_j)$, because of overdispersion and row-column dependence. Violation of the variance assumption underlies the differences in performance between the methods of permutation testing studied in this paper. Equation (1) is slightly more general than the usual equation in which a_{0j} is replaced by 1 (ter Braak & Verdonschot 1995). The replacement is valid, as the WLS estimate of a_{0j} is 1, when the predictor variables are $\mathbf{r}$-centered (as they generally are inside the algorithms). For permutation testing, however, we need the more general formula and also the weighted residual sum of squares of the null model $\mathbf{C} \sim \mathbf{Z}$ and the alternative model $\mathbf{C} \sim \mathbf{Z} + \mathbf{X}$, RSS_0 and RSS_a, respectively, so as to be able to create a pseudo-*F* test statistic $F = ((RSS_0 - RSS_a)/p)/(RSS_a/(n - p - q - 1))$. Note that $RSS_0 - RSS_a$ is the weighted regression sum of squares due to $\mathbf{X}$ (the fitted inertia), which is, in (partial) CCA, the inertia explained by $\mathbf{X}$ after adjustment for $\mathbf{Z}$. In testing the number of ordination axes (dimensionality testing), another test statistic is often used in which RSS_a is replaced by the weighted residual sum of squares after fitting the first eigenvalue of the model for $\mathbf{X}$ with covariates $\mathbf{Z}$. In testing the second axis, the first axis is first appended to the covariates, and so on for later axes (Legendre et al. 2011). The so-calculated *P*-values for subsequent axes are adjusted for multiple testing by replacing them by their cumulative maximum (Winkler et al. 2020). This is a sequential and closed testing procedure that guarantees that an axis is judged significant only when the earlier axes are judged significant.

2.2 Residualized predictor permutation

Residualized predictor permutation was independently invented by Collins (1987) and Dekker et al. (2007) and extended by ter Braak (2022) for statistical testing in weighted regression and redundancy analysis. It has some popularity in sociology for data organized in square matrices of relatedness among *n* cases, the application for which Dekker et al. (2007) developed the method. Fieberg et al. (2020) mention and apply the method in a review of resampling methods in ecology. Randomization in experiments forms a strong basis for predictor permutation as it randomly assigns treatments ($\mathbf{X}$) to units (cases, sites). In observational studies, X-permutation can be motivated by assuming exchangeability of the rows of $\mathbf{X}$.

Residualized predictor permutation for testing the effect of $\mathbf{X}$ adjusted for $\mathbf{Z}$ has two initial steps (ter Braak 2022). With $\mathbf{C}$ as response table, it.

(1) fits the null model $\mathbf{C} \sim \mathbf{Z}$ from which RSS_0 is calculated and,
(2) calculates the residuals of $\mathbf{X}$ with respect to $\mathbf{Z}$, *i.e.* of the model $\mathbf{X} \sim \mathbf{Z}$, resulting in $\widehat{\mathbf{E}}^X$.
     It then performs the following two steps for each permutation; it
(3) permutes the rows of the $\mathbf{X}$-residuals resulting in $\pi\left(\widehat{\mathbf{E}}^X\right)$ and
(4) fits the model $\mathbf{C} \sim \mathbf{Z} + \pi\left(\widehat{\mathbf{E}}^X\right)$ from which RSS_a^{perm} is calculated.

The explained inertia due to the permuted $\mathbf{X}$-residuals is then $RSS_a^{perm} - RSS_0$, which can be used directly as a test statistic, because RSS_0 is constant across permutations so that it is monotonically related to the pseudo-F test statistic value. Both test-statistics thus yield the same P-value. Note that WLS is used in this procedure.

The advantage of permuting $\mathbf{X}$-residuals in step (3) instead of permuting $\mathbf{X}$ itself is that the correlation between $\mathbf{X}$ and $\mathbf{Z}$ remains the same (as it does in residualized response permutation) so that the explained inertia of each newly created data set examines the same contrast (specified by the projection of $\mathbf{X}$ on the orthocomplement of $\mathbf{Z}$), whereas the F-ratio in simple $\mathbf{X}$-permutation would examine different contrasts (ter Braak 1992, 2022).

We now compare this procedure with the permutation method used in the function randtest.pcaiv of the R library ade4 for $\mathbf{Z} = \mathbf{1}_n$, as randtest.pcaiv does not allow for covariates. Instead of permuting $\mathbf{X}$-residuals, ade4 permutes the raw $\mathbf{X}$, but the resulting explained inertia is the same. The reason is that in this simple situation, step (2) simply subtracts the $\mathbf{r}$-weighted column means from $\mathbf{X}$, whereas in step (4) the $\mathbf{r}$-weighted column means of the permuted $\mathbf{X}$-residuals are subtracted in the calculation of the explained inertia. The first subtraction of means (or other constants) is therefore immaterial. In conclusion, the permutation method used for CCA in ade4 is a simple form of residualized predictor permutation. This equivalence is illustrated on the real data examples.

2.3 Residualized response permutation

Residualized response permutation, also known as the Freedman and Lane (1983) method of permutation, was popularized by its implementation in Canoco version 3.1 (ter Braak 1990, 1992). It was singled out as the best out of several other response permutation versions by Anderson and Robinson (2001). Residualized response permutation is more model-based than residualized predictor permutation in that it replaces the (unknown) random error in a regression model by permutations of the regression residuals under the null model. When the error variance varies across sites (heteroscedastic errors), the regression model is transformed first so as to make the error variance constant. How the error variance varies may be partially unknown, making the transformation to constant variance debatable.

Residualized response permutation for testing the effect of $\mathbf{X}$ adjusted for $\mathbf{Z}$ also has two initial steps (ter Braak 2022). With $\mathbf{C}$ as response table, it.

(1) replaces the required WLS fitting by equivalent ordinary least-squares (OLS) fitting, i.e. unweighted least-squares, by multiplication of the rows of the tables $\mathbf{C}$, $\mathbf{X}$ and $\mathbf{Z}$ by the square-root of $\mathbf{r}$ and the columns of $\mathbf{Y}$ also by the square-root of $\mathbf{k}$, resulting in the transformed data tables

$$\mathbf{C}_w = \mathbf{R}^{\frac{1}{2}}\mathbf{C}\mathbf{K}^{\frac{1}{2}} = \mathbf{R}^{-\frac{1}{2}}\mathbf{Y}\mathbf{K}^{-\frac{1}{2}}, \ \mathbf{X}_w = \mathbf{R}^{\frac{1}{2}}\mathbf{X}, \ \text{and} \ \mathbf{Z}_w = \mathbf{R}^{\frac{1}{2}}\mathbf{Z}, \ \text{and}$$

(2) calculates the residuals of $\mathbf{C}_w$ with respect to $\mathbf{Z}_w$, i.e. of the model $\mathbf{C}_w \sim \mathbf{Z}_w$, resulting in $\widehat{\mathbf{E}}_w^C$, the residuals under the null model. In the simple case of $\mathbf{Z} = \mathbf{1}_n$, these are standardized chi-square residuals $q_{ij} = (y_{ij} - r_i k_j)/\sqrt{r_i k_j}$, which are a starting point for the computation of a CCA in Legendre and Legendre (2012). It then performs the following two steps for each permutation; it.

(3) permutes the rows of the $\mathbf{C}_w$-residuals resulting in $\pi\left(\widehat{\mathbf{E}}_w^C\right)$,

(4) fits the model $\pi\left(\widehat{\mathbf{E}}_w^C\right) \sim \mathbf{Z}_w$ from which RSS_0^{perm} is calculated, and

(5) fits the model $\pi\left(\widehat{\mathbf{E}}_w^C\right) \sim \mathbf{Z}_w + \mathbf{X}_w$ from which RSS_a^{perm} is calculated.

The explained inertia of the permuted $\mathbf{C}_w$-residuals is then $RSS_a^{perm} - RSS_0^{perm}$ and is used to calculate the pseudo-F test-statistic. The explained inertia and pseudo-F test statistic are not monotonically related in residualized response permutation, as RSS_0^{perm} varies across permutations in this method. The pseudo-F test statistic is a better test statistic than the explained inertia in that it better adheres to the nominal significance level (Anderson and Robinson 2001). The transformation in step (1) is motivated by the variance of the contingency ratios of a contingency table under independence (ter Braak and Šmilauer 2018), but their variance may differ from $1/(r_i k_j)$ for other models and overdispersed data, which may lead to invalid P-values (ter Braak 2022).

We now compare this procedure with the permutation method used in the function anova.cca of the R library vegan. Starting from the standardized chi-square residuals $[q_{ij}]$, $\mathbf{r}$-centered $\mathbf{X}$ and $\mathbf{Z}$, the function anova.cca uses an efficient algorithm to calculate the explained inertia due to $\mathbf{X}$ in a permutation. However, the function does not include an intercept in $\mathbf{Z}$, so that RSS_a^{perm} is the same or slightly higher and the pseudo F test statistic the same or slightly smaller than in residualized response permutation, yielding a P-value that is systematically somewhat smaller, unless the site totals are all equal. This introduces a small amount of Type I error rate inflation in CCA with unequal $\mathbf{r}$ in scenarios where residualized response permutation adheres to the nominal significance level. However, the error is minor compared to the differences between residualized response and predictor permutation (e.g. Fig. 1).

Canoco versions 3.1 up to 5.12 used an algorithm different from that in vegan, but yield precisely the same P-value when these programs were given the same set of permutations. We show the equivalence by including the version without intercept in our R-code.

2.4 Simulation study

The simulation study is designed to discover when residualized response permutation (the current (2022) method of permutation in vegan and Canoco, see Table 1) does and
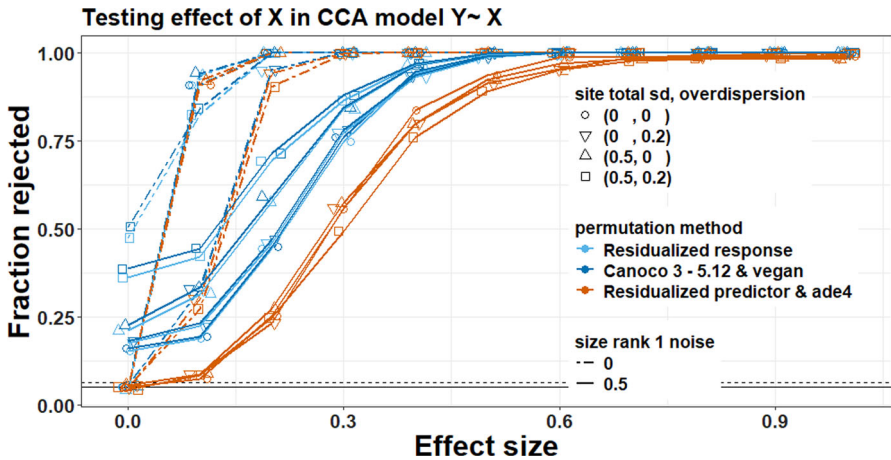
Fig. 1 Influence of noise types on the rejection rates (Type I error rate if effect size = 0, power otherwise) of three permutation methods for testing the effect of twelve predictors (**X**) on abundance data (**Y**) using CCA with the model **Y** ~ **X** ($n = 30$, $m = 50$, $p = 12$) (data generated using the loglinear simulation model of Appendix A2; Effect size = overall effect size; noise types: (1) site total sd = standard deviation of the site main effect, (2) overdispersion = overdispersion parameter of the negative binomial (0 = Poissonian), (3) size rank 1 noise = size of the effect of an unobserved predictor that is independent of **X**). The horizontal solid line is at the nominal significance threshold; rates (from 1000 simulations) above the dashed line (at 0.064) are significantly greater than 0.05. A small amount of jitter was applied to the effect size points so as to avoid a total overlap of points

does not maintain the nominal significance level and to demonstrate that residualized predictor permutation provides a solution to the issues found and that it is about equally powerful in realistic (*i.e.* noisy) data as residualized response permutation when the latter maintains the nominal level. The study is carried out in three series of simulation (Tables 3 and 4) by varying the data properties of the abundance data and by varying the applied data transformation (none, square-root, log).

The first series of simulations is about testing the overall null hypothesis (**X** had no effect on **Y**), the second series is about testing the dimensionality of the effects, and the third series is on the (lack of) sensitivity of CCA to log-linear environmental main effects (environmental effects that are common to all species). CCA is suited for strictly compositional data, if it is insensitive to such effects.

Data with $n = 30$ sites, $m = 50$ species and $p = 12$ predictors was simulated using a model with two true constrained ordination axes. A detailed description and R code can be found in Appendix A2. In short, all predictors are multivariate normal. The first four predictors define the first axis but, in the data, have added noise so that 10% of their variance was noise. The second set of four predictors define the second axis, but, in the data, have some contribution from the first axis so that their correlation with the first axis is 0.7. The third set of four predictors contains noise variables without any effect on the species, but have a correlation of 0.7 with the second axis. These correlations with the previous axis make it more difficult for methods to maintain the nominal significance level when testing dimensionality. The subsequent predictors

Table 3 Aims of the simulation study with null hypotheses and test statistic

Series	1		2	3
Figure	1	2	3	4

Aim, null hypothesis and test statistic

Aim	How the fraction rejected* by each permutation method depends on:			
	effect size, site total variability and noise level (overdispersion and rank 1 noise)	effect size and data transformation for high site variability and noise	effect size of an axis when there is a stronger earlier axis, for high site variability and noise	effect size of environmental main effect, so as to determine whether CCA is sensitive to main effects**
Null Hypothesis tested	Overall null		Dimensionality	Overall null
	The environmental variables have no effect on the taxonomic composition*** of the ecological community	As in Fig. 1	The environmental variables have no effect on the taxonomic composition*** beyond the effects collected in axis 1	
Test statistic(s)	Total inertia			
			Inertia of axis that is tested	

i.e. Type I error rate if effect size $= 0$ and power if effect size > 0

**An environmental main effect implies equal environmental effects on the species (on the log-linear scale); all species-specific environmental effects are then identical. The theory of CCA says that it should be insensitive to site main effects (so that it is suited for strictly compositional data) and we check here whether this property holds or does not hold true with and without data transformation

***Composition in the sense that only fractions are of interest and their row total (*i.e.* the site total) is not

within each set have a correlation of 0.7. Additional structured noise was generated from an independent third axis (rank 1 noise).

The abundance data were drawn from a negative binomial distribution with mean (μ) log-linearly defined by the three axes, random site and species main effects and a intercept of log(10). The variance was $\mu + \phi\mu^2$ where ϕ is the overdispersion compared to the Poisson distribution. In the second series (on dimensionality testing), the first dimension had Gaussian responses with exponentially distributed tolerances with mean 1. In the third series, with $n = 60$ sites and $m = 100$ species, environmental main effects were introduced in the model of the first series by making the log-linear site main effects correlated with the first predictor (x_1) for a fixed value (0.5) of the standard deviation of these site effects. In the first series, the site main effect are pure size effects (Greenacre 2017), independent of the predictors.

In the first series, the rejection rate of the permutation methods was studied as a function of (1) the effect size, with the first axis having twice the effect of the second or no effect for effect size zero, (2) the amount of overdispersion ϕ, (3) the variability

Table 4 Set-up of the simulation study: data generation and parameters that are varied. For details see Appendix A2 and Table A1 in the Supplementary material

Series	1		2	3
Figure	1	2	3	4

Data generation and transformation

Abundance data $\mathbf{Y}$	Negative binomial with mean μ and variance $\mu + \phi\,\mu^2$			
Sites	30			60
Species	50			100
Model	If effect size > 0, μ depended on log-linear effects of rank 2 by 8 environmental variables[*]. In addition, there were 4 environmental variables without any effect on the species. Environmental variables were correlated up to a correlation of 0.7	As in Series 1 but first axis had Gaussian responses for the species with exponentially distributed tolerance with parameter 1		As in Series 1, but with environmental main effects
Site main effects	random: $N(0,\sigma_1^2)$ σ_1 is the parameter for site total variability**			Correlated with first environmental variable
Species main effects	random: $N(\log(10), 0.25)$			
Transformation	none	none, square-root, $\log(y+1)$		

Parameter settings

Effect size	0–1		0–1.5	Fixed to 0
σ_1**	0 and 0.5	Fixed to 0.5		
Overdispersion ϕ	0 and 0.2	Fixed to 0.2		0,0.2
Rank 1 noise***	0,0.5		Fixed to 0.5	0,0.5

*The effects matrix of regression coefficients $[b_{jk}]_{j=1...m;\ k=1...p}$ in Eq. 1 had thus rank 0 (no effect) or rank 2 (effect > 0)

**σ_1 is the standard deviation (sd) of the log-linear site main effects; this is approximately the coefficient of variation of the site totals

***Effect of a third axis that was independent of the environmental variables in the data (an unconstrained axis), yielding structured noise in $\mathbf{Y}$ as compared to the unstructured noise by the negative binomial distribution

of the site main effects, expressed by their standard deviation σ_1, (4) the size of the rank 1 noise, and (5) the type of transformation of $\mathbf{Y}$ (none, square-root or $\log(y + 1)$). In the second series, their rejection rate for the second and third axis was studied in dependence of (1) the effect size of the second dimension, (2) type of test statistic, and (3) transformation of the abundance data. In the third series, the rejection rate was studied similarly in dependence of the correlation of x_1 with the site main effects.

For each scenario, 1000 data sets were generated, with 199 random permutations applied to each data set. Each permutation used in residualized predictor permutation

was the inverse of that in residualized response permutation. (The inverse, π^{-1}, of a permutation π applied to a permutation π yields the identity, i.e. no permutation). Applying the inverse permutation to the predictors yields the same data as applying the permutation to the response except for a rearrangement of rows, so that the resulting P-values are identical when the site totals would be all equal, despite the small number of permutations. All series were again run with the predictor data binarized, or exponentiated, so that the predictor data were skewed like concentration data.

2.5 Case studies

For the first case study, we searched for an ecological data set in which residual response and predictor permutation lead to a different conclusion on whether the predictor tested has any effect on the taxonomic composition when CCA is used. A data set we found is the public Dryad data of Pinho et al. (2021). As the full data on 417 forest plots and 3417 species across South America were proprietary and could not be disclosed, the plots were aggregated in to 59 clusters of plots that are within 50 km of another. In order to make sure that the analysis—a double constrained correspondence analysis—of the aggregated data would result in a similar ordination diagram as the same analysis of the original data, the abundance data were summed over plots in a cluster, whereas the environmental data were averaged. In the aggregated data on Dryad, the site totals (i.e. the cluster total) ranged from 35 to 36,915 trees, particularly as some biodiversity hotspots were highly sampled (the number of plots per cluster is not in the data but we know that it ranged from 1 to about 100). This make these data suited as an example in our paper because the site total herein is of no interest to any research question. The environmental data consisted of five climatic variables and three geographic variables. As an illustration, we wish to investigate whether forest composition shows an east–west gradient after taking account of the climatic variables which measure the current climate. Such a difference could be caused by the historical biogeography of the region (Pinho et al. 2021). In particular, we test the null hypothesis that longitude has no effect on the tree composition with and, for completeness, without adjustment for the climatic variables.

In addition, 24 public data sets available in the R packages ade4, vegan, mvabund (Wang et al. 2012) and TraitEnvMLMWA (ter Braak 2019) are analyzed without transformation and after square-root and log-transformation by the R code developed for this paper and by the current permutation methods for CCA in ade4 and vegan. These analyses allow to compare residualized response and predictor permutation in real data sets.

3 Results

3.1 Simulation study

Residualized predictor permutation (red lines in Figs. 1, 2, 3, 4) controls the Type I error rate of the test of the effect of the predictors on the abundance data in all scenarios
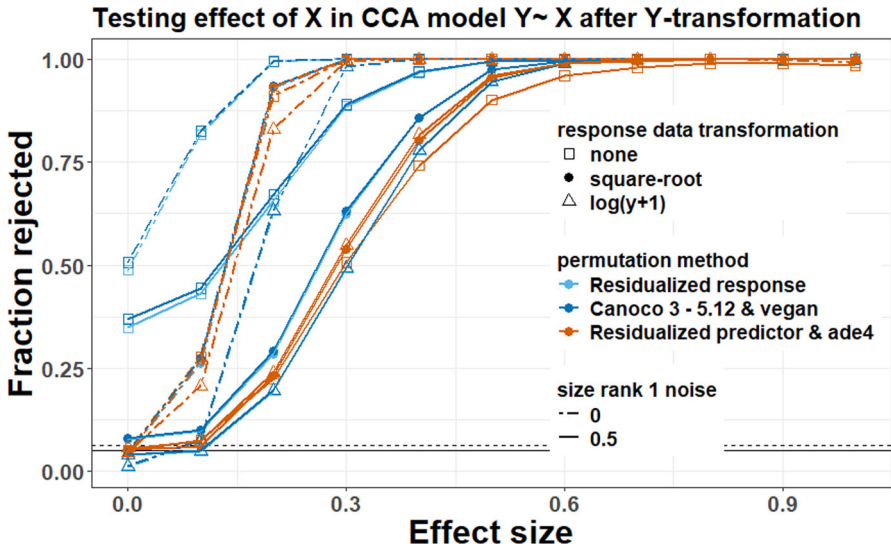
Fig. 2 The influence of data transformation and noise on the rejection rates (Type I error rate if effect size = 0, power otherwise) of three permutation methods for testing the effect of twelve predictors (**X**) on transformed abundance data with CCA using model $\mathbf{Y} \sim \mathbf{X}$ ($n = 30, m = 50, p = 12$). Data generated using the loglinear simulation model with overdispersion 0.2 and a standard deviation of 0.5 of the site main effect. For effect size and size rank 1 noise, see legend Fig. 1. The horizontal solid line is at the nominal significance threshold; rates (from 1000 simulations) above the dashed line (at 0.064) are significantly greater than 0.05
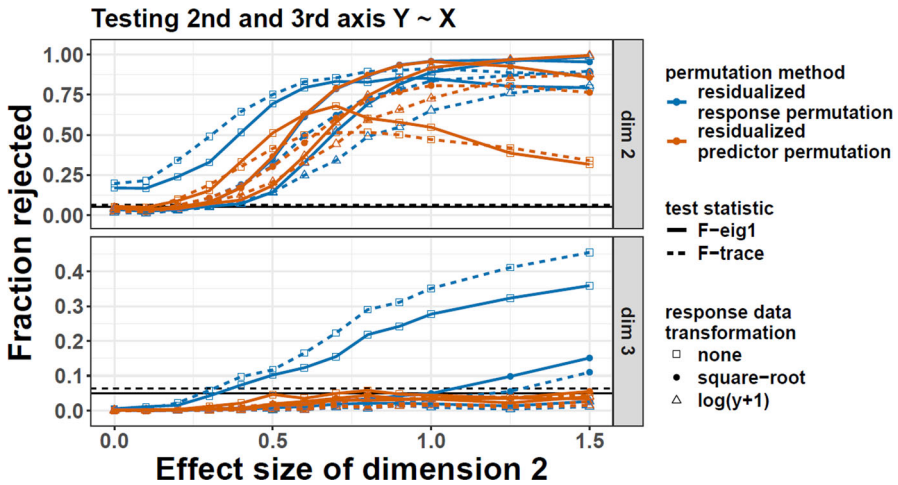

Fig. 3 Rejection rates of testing the second and third axes against the effect size of the second axis by CCA using two alternative test statistics (F_{eig} and F_{trace}) with $n = 30, m = 50, p = 12$. The rejection rate for the first axis, which had Gaussian response in this simulations series, was close to 1 everywhere. Data generated using overdispersion 0.2, a standard deviation of 0.5 of the site main effect and rank 1 noise of 0.5. The horizontal solid line is at the nominal significance threshold; rates (from 1000 simulations) above the dashed line (at 0.064) are significantly greater than 0.05
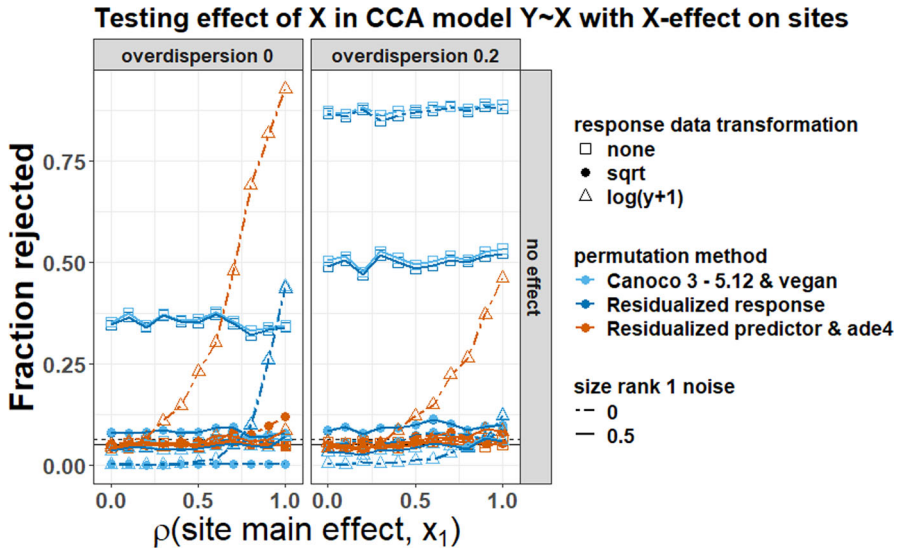
Fig. 4 Type I error rate of testing the effect of twelve predictors (**X**) on transformed abundance data with CCA using model $\mathbf{Y} \sim \mathbf{X}$ ($n = 60$, $m = 100$, $p = 12$) in relation to the correlation of one of the predictors (x_1) with the log-linear site main effect, with the influence of data transformation, overdispersion and rank 1 noise. Data generated as in Fig. 1 with effect size $= 0$, except that the site main effects were made correlated with x_1. The standard deviation of the site main effects was 0.5. For size rank 1 noise, see legend Fig. 1. The horizontal solid line is at the nominal significance threshold; rates (from 1000 simulations) above the dashed line (at 0.064) are significantly greater than 0.05

of Fig. 1 (effect size $= 0$) and its power increases naturally with the effect size (Fig. 1). Residualized response permutation (dark blue lines in Figs. 1, 2, 3, 4) gives grossly inflated Type I error rates in the presence of variability in the site totals, overdispersion and/or additional structured (rank 1) noise (effect size $= 0$ in Fig. 1) except when there is little more noise than Poissonian noise (Fig. 1: all dashed blue lines start at around (0,0.05) in Fig. 1 and almost coincide with the dashed red lines with same symbols, except for the dashed dark blue and light blue lines with blue squares which start around (0,0.50), whereas the corresponding dashed red line with red squares starts at (0,0.05), as required). Additional unstructured noise behaved similar to an increase in the overdispersion and thus gave similar Type I error inflation (results not shown). The implementations without correction for the intercept (light blue lines in Figs. 1, 2, 3, 4) in Canoco versions 3-5.12 and in anova.cca in the R-library vegan (up to the current version which is version 2.6-2) led to slightly more Type I error inflation than its corrected form (*i.e.* residualized response permutation).

Whereas residualized response permutation gave a highly inflated type I error rate on untransformed data (Fig. 1), it showed no and only a minor type I error inflation after square-root or $\log(y + 1)$ transformation of the abundance data (Fig. 2). Its small type I error inflation for the square-root transformed data with rank 1 noise was compensated by a higher power than for the other transformations (Fig. 2: the solid blue line with circles is above the corresponding solid red line for all effect sizes in

Fig. 2). As highly variable count data are often square-rooted or, as recommended in the Canoco adviser, log-transformed, the type I error rate inflation with residualized response permutation is not as dramatic in practice as it is in Fig. 1.

With residualized predictor permutation, the fraction of rejections was almost independent of the data transformation, including no transformation, particularly in the case with rank 1 noise (Fig. 2). The square-root transformation had equal or slightly higher power than the log transformation in this series. After transformation, the difference between residualized response permutation and its current implementation in Canoco and vegan is so small that their lines and points are hardly distinguishable in Fig. 2.

In testing the second and third dimension, residualized response permutation showed type I error rate inflation without response data transformation (Fig. 3) and neither method could reach a power of 1 for a large effect size of the second dimension (Fig. 3). After a square-root or log transformation, large effect sizes resulted in a power close to 1 and no type I error rate inflation was observed (except for testing the third dimension on square-rooted data using residualized response permutation when the effect size of dimension 2 was large). The test statistic based on the first eigenvalue performed better for dimensionality testing in this series than the one based on the trace (all inertia explained by $\mathbf{X}$ given $\mathbf{Z}$).

CCA is insensitive to environmental main effects, at least on untransformed or square-rooted data (Fig. 4: all lines are about horizontal, except those with triangles indicating the $\log(y + 1)$ transformation). After log-transformation, there is some sensitivity when there is no additional rank 1 noise (the red and blue dashed lines with triangles in Fig. 4), with residual predictor permutation being more sensitive than residual response permutation. The sensitivity decreases with increasing noise due to overdispersion (compare left and right panels in Fig. 4) and rank 1 noise (compare solid and dashed red lines in both panels of Fig. 4). As in Figs. 2, 3, the type I error inflation of residual response permutation using untransformed data (the six out of eight blue lines with blue squares in Fig. 4; the two exceptions are the two dashed ones in the left panel which are without any overdispersion and which are hardly visible as they show no type I error inflation and overlap with the red lines) largely disappears after square-root and log-transformation (Fig. 4: all blue lines without squares).

With predictor data made binary or skewed, the results on type I error rate and power were qualitatively similar to those of the normal predictor data, except that the power was less and that, for the skewed predictor data, there was sensitivity to site main effects without response transformation when there was overdispersion (Appendix A3).

3.2 Case studies

With residualized response permutation, the effect of longitude on Neotropical forest composition in the first case study is significant with and without adjustment for climate (Table 5). However, with residualized predictor permutation the conditional effect of latitude is no longer significant ($P = 0.40$). By repeated simulation of a random predictor (so that we know for sure that it has no effect) we verified that residualized response permutation shows Type I error inflation whereas residualized predictor permutation does not (Table 5).

Table 5 P-values for statistically testing the east–west gradient (Longitude) in neotropical forest composition ($n = 59$ clusters; $m = 3417$ tree species) with and without adjustment for climate by residualized response (RRP) and predictor (RPP) permutation in comparison with the rejection rates of such tests in 1000 simulations of a completely random predictor (199 permutations)

Adjustment by climate	Longitude		Random predictor	
	Yes	No	Yes	No
RRP	0.005	0.005	0.830	0.888
RPP	0.400	0.010	0.058	0.059

These results suggest that there is strong statistical evidence that tree composition varies with longitude, but little evidence that it varies with longitude after present-day climate is taken into account. However, after square-root or log transformation there is some evidence for the conditional effect of longitude with residualized predictor permutation ($P = 0.065$ and 0.045, respectively, Appendix A4). A somewhat speculative ecological conclusion is that historical biogeography plays a bigger role for the subdominant tree species than for the dominant ones. The firm statistical conclusion is that residualized response permutation cannot be trusted in data with such a wide range of site totals (in this extreme case, not even after square-root or log-transformation, see Appendix A4), whereas residualized predictor permutation maintains the nominal significance level.

In the additional 24 analyses of the real data from R packages, the differences in P-values between residualized response and predictor permutation or between types of transformation were unremarkable as most would lead to the same conclusion about the evidence for the effects investigated (Appendix A5). Exceptions are as follows. In the Revisit data of the TraitEnvMLMWA package with highly variable species totals, the relation between abundance and the species trait of this data is highly significant using residual response permutation, but, using residual predictor permutation, only clearly seen after square-root or log-transformation. In the solberg data of the mvabund package the site totals are almost constant, but the species totals are so variable that it is only after square-root or log-transformation that the differences between treatments are evident. Apparently, the treatments have effects only on the sub-dominant species. The differences between permutation methods was minor in this data and all the other data sets (Appendix A5).

Supplied with the same set of Monte Carlo permutations, the P-values issued by vegan (version 2.6-2) are exactly equal to those of residualized response permutation without accounting for the intercept, and those issued by ade4 (version 1.7-18) are exactly equal to those of residualized predictor permutation (Appendix A5).

4 Discussion

The permutation test in CCA used in vegan and Canoco 5.12 is based on the assumption that the greater the site total, the more accurately the fraction of each species in a site is

estimated. However, the test is not robust against violation of this assumption, as this paper shows by simulating overdispersed count data with highly variable site totals. In contrast, the permutation test used in ade4 is not dependent on this assumption, does not suffer from Type I error rate inflation where the other test does, and is approximately equally powerful in realistic (*i.e.* noisy) data when both tests maintain the nominal significance level. With a slight correction of the first and an extension to covariate data for the second test, the tests are known as residualized response and residualized predictor permutation tests.

The difference between Canoco/vegan and ade4 remained unnoticed so far, presumably because their *P*-values can be expected to differ as both are Monte Carlo tests. Moreover, they do not differ that much in practice, for the reason that highly variable data are often square-rooted or, as recommended in the Canoco adviser, log-transformed.

Without data transformation, residualized response permutation yielded grossly inflated type I error rates in testing the number of constrained axes. The power of both methods did not reach 1 for a large effect size of the tested axis, a feature that we do not fully understand. After square-root or log-transformation, both methods performed much better in terms of both power and Type I error rate. The test statistic based on the first eigenvalue performed better than that based on the trace (*i.e.* based on all constrained eigenvalues). The choice of test statistic may need further research, as ter Braak (2022) concluded the opposite in his simulations of weighted redundancy analysis.

After log-transformation, CCA becomes somewhat sensitive to environmental (log-linear) main effects, as noted by te Beest et al. (2021). Log-transformation is therefore less suited for strictly compositional data, because the predictor variables may accidentally be correlated to the site total which is a technical artefact in such data and would then be detected as true effects, even if it is a pure size effect. This issue does not appear of much importance in ecological applications outside (current) metagenomics, also because the data are typically rather noisy, in which case there is hardly any sensitivity. If not a sampling or other technical artefact, a regression analysis of site totals against the environmental variables may complement a CCA.

Significance tests of trait–environment relationships require both a site-level and a species-level test to control the type I error. The tests are combined by taking the maximum of their *P*-values (max test) as first proposed using permutation tests with the fourth-corner correlation as the test statistic (ter Braak 2019), which is a special case of double constrained correspondence analysis relating a single trait with a single environmental variable. However, the *P*-values are similar to those in the papers on the fourth-corner correlation (Peres-Neto et al. 2017; ter Braak 2019) only when (residualized) predictor permutation is used, in which the environmental and trait data are permuted in the site- and species-level test, respectively. Alternatively, one can permute the row of the abundance table in the site-level test and columns of this table in the species-level test, but this does not make it (residualized) response permutation, as the weights (the margin totals of the abundance table) come with the permutation and there is no link with standardized chi-square residuals. Species totals generally vary much more than site totals. As in the related analysis of CWMs, the analysis proceeds often without data transformation, so that residualized response permutation

will likely turn very liberal. The message of this paper is thus even more relevant to double constrained correspondence analysis as it is to canonical correspondence analysis.

Declarations

Conflict of interest CtB is the first author of Canoco. He has academic, but no financial, interest in Canoco.

References

Anderson MJ, Robinson J (2001) Permutation tests for linear models. Aust N Z J Stat 43:75–88. https://doi.org/10.1111/1467-842X.00156

Borcard D, Gillet F, Legendre P (2011) Numerical ecology with R. Springer, New York

Collins MF (1987) A permutation test for planar regression. Aust J Stat 29:303–308. https://doi.org/10.1111/j.1467-842X.1987.tb00747.x

Dekker D, Krackhardt D, Snijders TAB (2007) Sensitivity of MRQAP tests to collinearity and autocorrelation conditions. Psychometrika 72:563–581. https://doi.org/10.1007/s11336-007-9016-1

Dray S, Legendre P (2008) Testing the species traits environment relationships: the fourth-corner problem revisited. Ecology 89:3400–3412. https://doi.org/10.1890/08-0349.1

Fieberg JR, Vitense K, Johnson DH (2020) Resampling-Based Methods for Biologists. PeerJ. https://doi.org/10.7717/peerj.9089

Freedman DA, Lane D (1983) A nonstochastic interpretation of reported significance levels. J Bus Econ Stat 1:292–298. https://www.jstor.org/stable/1391660

Gobbi M, Corlatti L, Caccianiga M, ter Braak CJF, Pedrott L (2022) Hay meadows' overriding effect shapes ground beetle functional diversity in mountainous landscapes. Ecosphere. https://doi.org/10.1002/ecs2.4193

Goodman LA (1986) Some useful extensions of the usual correspondence analysis approach and the usual log-linear models approach in the analysis of contingency tables. Int Stat Rev 54:243–270. https://doi.org/10.2307/1403053

Greenacre M (2017) 'Size' and 'shape' in the measurement of multivariate proximity. Methods Ecol Evol 8:1415–1424. https://doi.org/10.1111/2041-210X.12776

Greenacre M (2018) Compositional data analysis in practice. CRC Press, Boca Raton

Ihm P, van Groenewoud H (1984) Correspondence analysis and Gaussian ordination. Compstat Lectures 3:5–60

Legendre L, Legendre P (2012) Numerical ecology. Elsevier, Amsterdam

Legendre P, Galzin RG, Harmelin-Vivien ML (1997) Relating behavior to habitat: solutions to the fourth-corner problem. Ecology 78:547–562. https://doi.org/10.2307/2266029

Legendre P, Oksanen J, ter Braak CJF (2011) Testing the significance of canonical axes in redundancy analysis. Methods Ecol Evol 2:269–277. https://doi.org/10.1111/j.2041-210X.2010.00078.x

Liu M, Prentice IC, ter Braak CJF, Harrison SP (2020) An improved statistical approach for reconstructing past climates from biotic assemblages. Proc R Soc A. https://doi.org/10.1098/rspa.2020.0346

McArdle BH, Anderson MJ (2001) Fitting multivariate models to community data: a comment on distance-based redundancy analysis. Ecology 82:290–297. https://doi.org/10.1890/0012-9658(2001)082[0290:FMMTCD]2.0.CO;2

Muff S, Nilsen EB, O'Hara RB, Nater CR (2021) Rewriting results sections in the language of evidence. Trends Ecol Evol. https://doi.org/10.1016/j.tree.2021.10.009

Niku J, Hui FKC, Taskinen S, Warton DI (2019) gllvm: fast analysis of multivariate abundance data with generalized linear latent variable models in R. Methods Ecol Evol 10:2173–2182. https://doi.org/10.1111/2041-210X.13303

Oksanen J et al (2022) vegan: community ecology package. R package version 2.6-2. http://CRAN.R-project.org/package=vegan

Peres-Neto PR, Dray S, ter Braak CJF (2017) Linking trait variation to the environment: critical issues with community-weighted mean correlation resolved by the fourth-corner approach. Ecography 40:806–816. https://doi.org/10.1111/ecog.02302

Pinho BX, Tabarelli M, ter Braak CJF, Wright SJ, Arroyo-Rodríguez V, Benchimol M, Engelbrecht BMJ, Pierce S, Hietz P, Santos BA, Peres CA, Müller SC, Wright IJ, Bongers F, Lohbeck M, Niinemets Ü, Slot M, Jansen S, Jamelli D, de Lima RAF, Swenson N, Condit R, Barlow J, Slik F, Hernández-Ruedas MA, Mendes G, Martínez-Ramos M, Pitman N, Kraft N, Garwood N, Guevara Andino JE, Faria D, Chacón-Madrigal E, Mariano-Neto E, Júnior V, Kattge J, Melo FPL (2021) Functional biogeography of Neotropical moist forests: Trait–climate relationships and assembly patterns of tree communities. Global Ecol Biogeog 30:1430–1446. https://doi.org/10.1111/geb.13309

te Beest DE, Nijhuis EH, Möhlmann TWR, ter Braak CJF (2021) Log-ratio analysis of microbiome data with many zeroes is library size dependent. Mol Ecol Resour 21:1866–1874. https://doi.org/10.1111/1755-0998.1339

ter Braak CJF (1986) Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. Ecology 67:1167–1179. https://doi.org/10.2307/1938672

ter Braak CJF (1987) The analysis of vegetation-environment relationships by canonical correspondence analysis. Vegetatio 69:69–77. https://doi.org/10.1007/BF00038688

ter Braak CJF (1988) CANOCO - a FORTRAN program for canonical community ordination by [partial] [detrended] [canonical] correspondence analysis, principal components analysis and redundancy analysis (version 2.1). Report LWA-88-02. Agricultural Mathematics Group, Wageningen. http://edepot.wur.nl/248698

ter Braak CJF (1990) Update notes: CANOCO version 3.1. Agricultural Mathematics Group, Wageningen. http://edepot.wur.nl/250652

ter Braak CJF (1992) Permutation versus bootstrap significance tests in multiple regression and ANOVA. In: Jöckel K-H, Rothe G, Sendler W (eds) Bootstrapping and related techniques. Springer, Berlin. http://edepot.wur.nl/249346

ter Braak CJF (2014) History of canonical correspondence analysis. In: Blasius J, Greenacre M (eds) Visualization and verbalization of data. Chapman and Hall, London

ter Braak CJF (2017) Fourth-corner correlation is a score test statistic in a log-linear trait–environment model that is useful in permutation testing. Environ Ecol Stat 24:219–242. https://doi.org/10.1007/s10651-017-0368-0

ter Braak CJF (2019) New robust weighted averaging- and model-based methods for assessing trait–environment relationships. Methods Ecol Evol 10:1962–1971. https://doi.org/10.1111/2041-210X.13278

ter Braak CJF (2022) Predictor versus response permutation for significance testing in weighted regression and redundancy analysis. J Stat Comput Simul 92:2041–2059. https://doi.org/10.1080/00949655.2021.2019256

ter Braak CJF, Šmilauer P (2018) Canoco reference manual and user's guide: software for ordination (version 5.10). Microcomputer Power

ter Braak CJF, Verdonschot PFM (1995) Canonical correspondence analysis and related multivariate methods in aquatic ecology. Aquat Sci 57:255–289. https://doi.org/10.1007/BF00877430

ter Braak CJF, Šmilauer P, Dray S (2018) Algorithms and biplots for double constrained correspondence analysis. Environ Ecol Stat 25:171–197. https://doi.org/10.1007/s10651-017-0395-x

Thioulouse J, Dray S, Dufour A-B, Siberchicot A, Jombart T, Pavoine S (2018) Multivariate analysis of ecological data with ade4. Springer, New York

Wang Y, Naumann U, Wright ST, Warton DI (2012) mvabund: an R package for model-based analysis of multivariate abundance data. Methods Ecol Evol 3:471–474. https://doi.org/10.1111/j.2041-210X.2012.00190.x

Warton D, Foster S, Death G, Stoklosa J, Dunstan P (2014) Model-based thinking for community ecology. Plant Ecol. https://doi.org/10.1007/s11258-014-0366-3

Winkler AM, Renaud O, Smith SM, Nichols TE (2020) Permutation inference for canonical correlation analysis. Neuroimage 220:117065. https://doi.org/10.1016/j.neuroimage.2020.117065