

# Improved testing of species traits–environment relationships in the fourth-corner problem

CAJO J. F. TER BRAAK,<sup>1,4</sup> ANOUK CORMONT,<sup>2</sup> AND STÉPHANE DRAY<sup>3</sup>

<sup>1</sup>*Biometris, Wageningen University and Research Centre, Wageningen, The Netherlands*

<sup>2</sup>*Alterra, Wageningen University and Research Centre, Wageningen, The Netherlands*

<sup>3</sup>*Université Lyon I, CNRS; UMR 5558, Laboratoire de Biométrie et Biologie Evolutive, 43 Boulevard du 11 Novembre 1918, Villeurbanne F-69622 France*

**Abstract.** The fourth-corner problem entails estimation and statistical testing of the relationship between species traits and environmental variables from the analysis of three data tables. In a 2008 paper, S. Dray and P. Legendre proposed and evaluated five permutation methods for statistical significance testing, including a new two-step testing procedure. However, none of these attained the correct type I error in all cases of interest. We solve this problem by showing that a small modification of their two-step procedure controls the type I error in all cases. The modification consists of adjusting the significance level from  $\sqrt{\alpha}$  to  $\alpha$  or, equivalently, of reporting the maximum of the individual  $P$  values as the final one. The test is also applicable to the three-table ordination method RLQ.

**Key words:** ecological community; fourth-corner statistic; functional ecology; permutational model; RLQ analysis; sequential test; species traits.

## INTRODUCTION

The fourth-corner problem and RLQ relate species traits (matrix  $\mathbf{Q}$ ) to environmental variables ( $\mathbf{R}$ ) via a third table ( $\mathbf{L}$ ) containing data on the abundance (or presence) of the species in a set of sites (Dray and Legendre 2008). In the fourth-corner problem, each combination of trait and environmental variable is analyzed separately (Legendre et al. 1997), whereas RLQ examines their joint multivariate relation (Dolédéc et al. 1996). Solving this problem is a first step to understanding which traits make species fit for particular environments. Traits and environmental variables cannot be correlated directly as they are measured on different observational units, namely species and sites (corresponding to rows of  $\mathbf{Q}$  and of  $\mathbf{R}$ ) respectively. The solution is to link traits and environment via the abundance table  $\mathbf{L}$ , the columns and rows of which correspond to species and sites, respectively. The problem is thus linking three data tables to produce the fourth corner, the trait-by-environment table  $\mathbf{X}$ , the elements of which measure the relationship between traits and environment (Legendre et al. 1997):

$$\begin{bmatrix} \mathbf{L} & \mathbf{R} \\ \mathbf{Q}' & \mathbf{X} \end{bmatrix}.$$

Dray and Legendre (2008) proposed a generally useful measure for  $\mathbf{X}$  and evaluated five permutation methods for statistical significance testing, including a new two-step testing procedure. However, none of these main-

tained the correct type I error in all cases of interest. We solve this problem by viewing their two-step procedure as a sequential testing procedure. We show that a small modification of their procedure actually controls the type I error in all cases. The modification consists of adjusting the significance level from  $\sqrt{\alpha}$  to  $\alpha$ . The test is also applicable to the three-table ordination method RLQ.

## THEORY, METHODS, AND RESULTS

The challenge is to develop a statistical test of  $H_0: \mathbf{X} = \mathbf{0}$  (trait and environment are unrelated) against  $H_1: \mathbf{X} \neq \mathbf{0}$  (trait and environment are somehow related). The test should control the type I error, meaning that when  $H_0$  is true, the number of false rejections is less than or equal to the nominal significance level  $\alpha$  (e.g.,  $\alpha = 0.05$ ). Dray and Legendre (2008) identified that under  $H_1$  both the links  $\mathbf{L} \leftrightarrow \mathbf{Q}$  and  $\mathbf{L} \leftrightarrow \mathbf{R}$  must exist, and that  $H_0: \mathbf{X} = \mathbf{0}$  holds if at least one of the links is absent. The null hypothesis thus consists of three null joint hypotheses: ( $\mathbf{L} \leftrightarrow \mathbf{Q}, \mathbf{L} \leftrightarrow \mathbf{R}$ ), ( $\mathbf{L} \leftrightarrow \mathbf{Q}, \mathbf{L} \leftrightarrow \mathbf{R}$ ), ( $\mathbf{L} \leftrightarrow \mathbf{Q}, \mathbf{L} \leftrightarrow \mathbf{R}$ ) (see Dray and Legendre 2008: Fig. 2). This complex null hypothesis can be tested by a sequential test procedure (Goeman and Solari 2010) on two null hypotheses,  $\mathbf{L} \leftrightarrow \mathbf{Q}$  and  $\mathbf{L} \leftrightarrow \mathbf{R}$ . Let's first test the null hypothesis  $\mathbf{L} \leftrightarrow \mathbf{Q}$  at significance level  $\alpha$ . If this null is not rejected, the overall null ( $H_0: \mathbf{X} = \mathbf{0}$ ) is not rejected. The second null hypothesis ( $\mathbf{L} \leftrightarrow \mathbf{R}$ ) is tested only if the null hypothesis  $\mathbf{L} \leftrightarrow \mathbf{Q}$  is rejected. This second null hypothesis is also tested at significance level  $\alpha$ , and, if it is (not) rejected, the overall null hypothesis is (not) rejected. The overall null hypothesis is thus rejected when both nulls ( $\mathbf{L} \leftrightarrow \mathbf{Q}$  and  $\mathbf{L} \leftrightarrow \mathbf{R}$ ) are rejected. This testing procedure fulfills the

Manuscript received 23 January 2012; accepted 28 February 2012. Corresponding Editor: B. D. Inouye.

<sup>4</sup> E-mail: cajo.terbraak@wur.nl

conditions of Theorem 1 in Goeman and Solari (2010). By corollary, the test controls the type I error of the test on the overall null hypothesis ( $H_0: \mathbf{X} = \mathbf{0}$ ). Without the theorem, we may argue as follows. Because we do not execute the second test when the first is nonsignificant, the overall type I error is at most  $\alpha$  (the level of the first test), when  $\mathbf{L} \leftrightarrow \mathbf{Q}$  is true. When the first test is judged significant, we either commit a type I error of  $\alpha$  (when  $\mathbf{L} \leftrightarrow \mathbf{Q}$  is true) or we take the right decision (when  $\mathbf{L} \leftrightarrow \mathbf{Q}$  is true) and proceed with the second test. When this test is judged significant, we either commit a type I error of  $\alpha$  (when  $\mathbf{L} \leftrightarrow \mathbf{R}$  is true) or we take the right decision (when  $\mathbf{L} \leftrightarrow \mathbf{R}$  is true). The overall type I error can thus not be larger than  $\alpha$ . This result can be phrased as “once the first hypothesis is rejected, the second hypothesis can be treated as a completely new test” (Romano and Wolf quoted in Goeman and Solari 2010).

So far, the hypotheses were tested in the order  $\mathbf{L} \leftrightarrow \mathbf{Q}$  and  $\mathbf{L} \leftrightarrow \mathbf{R}$ . Would we get the same outcome by changing the order? Not necessarily, if different test statistics are used in the two steps. This potential asymmetry in the procedure is unwanted and can be avoided by using the same test statistic in both steps. If both test statistics are set to the fourth-corner test statistic of Dray and Legendre (2008), both testing orders yield the same result. Note that the test on  $\mathbf{L} \leftrightarrow \mathbf{Q}$  then already involves data table  $\mathbf{R}$  and the test on  $\mathbf{L} \leftrightarrow \mathbf{R}$  involves data table  $\mathbf{Q}$ . The overall null is then rejected when the  $P$  values of both tests are less than  $\alpha$ . Phrased differently, the significance level ( $P$  value) of the overall test is the maximum of the two individual  $P$  values.

The test proposed here is remarkably similar to the two-step approach of Dray and Legendre (2008). Instead of multiplying the individual  $P$  values, the maximum must be taken. In terms of pre-set significance levels, the significance level must be changed from  $\sqrt{\alpha}$  to  $\alpha$ . Taking the maximum of the individual  $P$  values is what happens to be currently implemented in the `R` procedure `combine.4thcorner` of the `ade4` package (Dray and Dufour 2007). The “combine” test proposed by Dray and Legendre (2008) would require setting the alpha argument of the `plot.4thcorner` function to `sqrt(0.05)`, but we showed in this report that  $\alpha$  should be 0.05 instead. This result was first used in Cormont et al. (2011) for the linear trait–environment method.

We evaluated the procedure using the simulation study performed by Dray and Legendre (2008). The Appendix shows that the type I error is controlled in the new, sequential approach (it is lower than or around 0.05 in scenarios 2–4), but is up to five times too large (0.25) in the two-step approach of Dray and Legendre (2008). The sequential approach has good power ( $>0.88$ ) with large numbers of species (100), reasonable power ( $>0.60$ ) for 50 species and some power ( $>0.40$ ) for 30 species (scenario 1N); the corresponding powers for the invalid two-step approach are 0.99, 0.90 and 0.78, respectively.

#### CONCLUSION

We showed that a simple modification to the two-step procedure in Dray and Legendre (2008) yields a valid test. With a prescribed significance level  $\alpha$ , both tests ( $\mathbf{L} \leftrightarrow \mathbf{Q}$  and  $\mathbf{L} \leftrightarrow \mathbf{R}$ ) should be executed at level  $\alpha$  instead of  $\sqrt{\alpha}$ . In terms of  $P$  values, the maximum of the individual  $P$  values gives the final  $P$  value (significance level).

Our result is quite general; no mention is made on how the individual tests are carried out, except that they use the same test statistic. Thus our result applies not only to permutation tests in the fourth-corner problem and RLQ, but also to bootstrap and parametric tests in these and related methods.

#### LITERATURE CITED

- Cormont, A., C. C. Vos, C. A. M. Turnhout, R. P. B. Foppen, and C. J. F. ter Braak. 2011. Using life-history traits to explain bird population responses to changing weather variability. *Climate Research* 49:59–71.
- Dolédéc, S., D. Chessel, C. J. F. ter Braak, and S. Champely. 1996. Matching species traits to environmental variables: a new three-table ordination method. *Environmental and Ecological Statistics* 3:143–166.
- Dray, S., and A. B. Dufour. 2007. The `ade4` package: implementing the duality diagram for ecologists. *Journal of Statistical Software* 22:1–20.
- Dray, S., and P. Legendre. 2008. Testing the species traits environment relationships: the fourth-corner problem revisited. *Ecology* 89:3400–3412.
- Goeman, J. J., and A. Solari. 2010. The sequential rejection principle of familywise error control. *Annals of Statistics* 38:3782–3810.
- Legendre, P., R. Galzin, and M. L. Harmelin-Vivien. 1997. Relating behavior to habitat: solutions to the fourth-corner problem. *Ecology* 78:547–562.

#### SUPPLEMENTAL MATERIAL

##### Appendix

Results of the simulation study performed in Dray and Legendre (2008) for the two-step procedure of Dray and Legendre (2008) and the new sequential testing procedure (*Ecological Archives* E093-136-A1).