A tutorial to perform fourth-corner and RLQ analyses in R

Stéphane Dray

stephane.dray@univ-lyon1.fr

July 21, 2013

Abstract

This tutorial shows how the fourth-corner and RLQ analyses can be performed using the `ade4` package (Dray and Dufour, 2007) for R. The dataset presented in the paper is used to illustrate the methods. Commands are written in red and outputs are written in blue.

Contents

1 Loading the package and the data

The different methods presented in the paper are available in the ade4 package. The first step is to load ade4 and the datasets:

```
library(ade4)
data(aravo)
```

The aravo dataset contains the abundances of 82 species in 75 sites. Species are described by 8 traits and 6 environmental variables have been measured in the sites.

```
dim(aravo$spe)
```

```
[1] 75 82
```

```
dim(aravo$traits)
```

```
[1] 82  8
```

```
dim(aravo$env)
```

```
[1] 75  6
```

2 RLQ

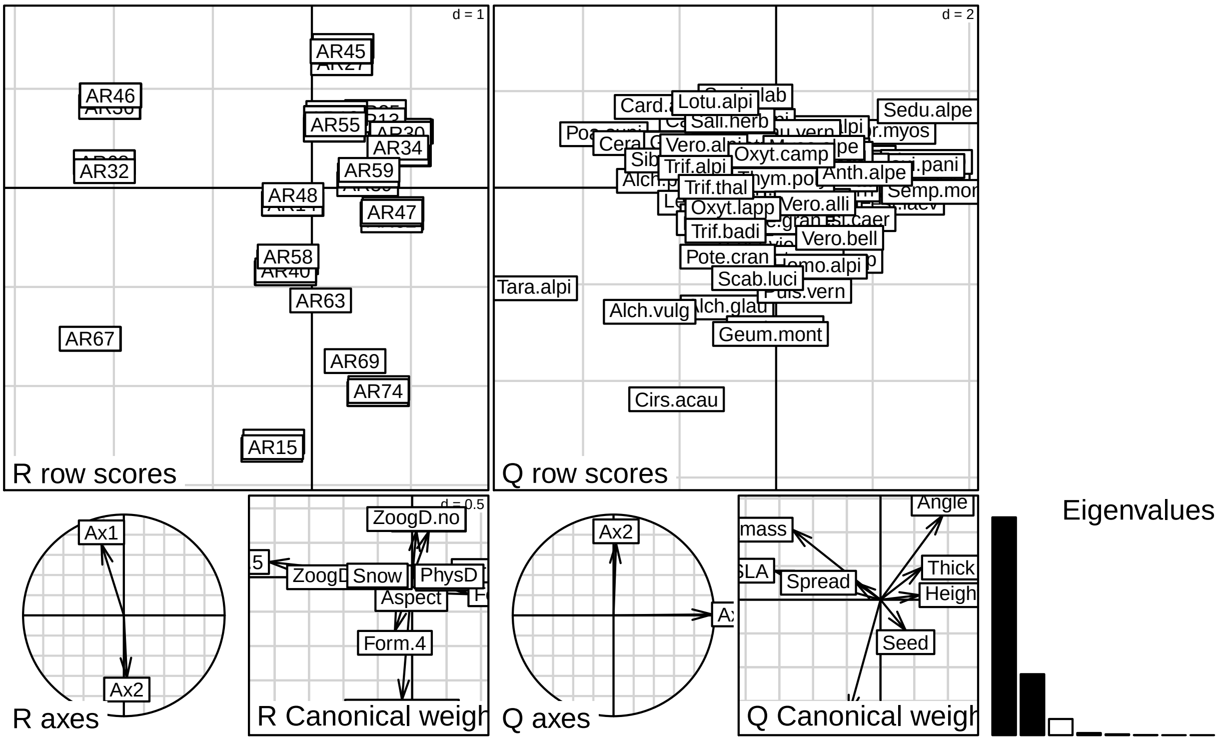
A preliminary step of RLQ analysis is to perform the separate analyses of each table. Correspondence analysis (dudi.coa) is applied to the species table. For traits data, all variables are quantitative and thus we applied a principal component analyses (dudi.pca). The environmental table contains both quantitative and categorical variables. In this case, we used the dudi.hillsmith function that allows to consider mix of different types of variables. Note that to proceed RLQ analysis, separate analyses of traits and environmental variables should be weighted by the sites and species weights derived from the previous correspondence analysis.

```
afcL.aravo <- dudi.coa(aravo$spe, scannf = FALSE)
acpR.aravo <- dudi.hillsmith(aravo$env, row.w = afcL.aravo$lw,
    scannf = FALSE)
acpQ.aravo <- dudi.pca(aravo$traits, row.w = afcL.aravo$cw,
    scannf = FALSE)
rlq.aravo <- rlq(acpR.aravo, afcL.aravo, acpQ.aravo,
    scannf = FALSE)
```

RLQ analysis finds coefficents (in $c1) to obtain a linear combination of traits (species scores in $lQ) and coefficients (in $l1) to obtain a linear combination of environmental variables (site scores in $lR). The covariance between these two sets scores is maximized and equal to the square root of the corresponding eigenvalue.

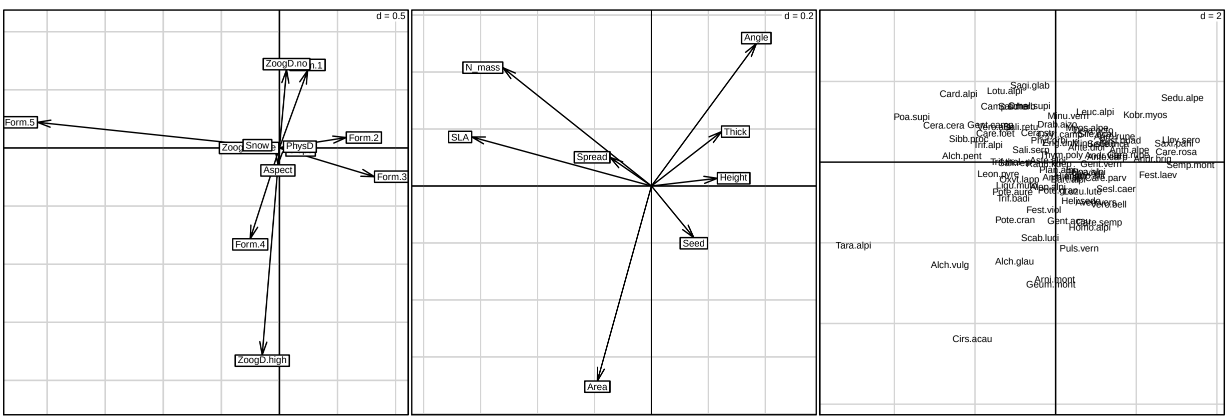
The different outputs of the analysis are obtained by the plot function:

```
plot(rlq.aravo)
```



The different figures can be obtained separately by plotting the different elements contained in the `rlq.aravo` object:

```
par(mfrow = c(1, 3))
s.arrow(rlq.aravo$l1)
s.arrow(rlq.aravo$c1)
s.label(rlq.aravo$lQ, boxes = FALSE)
```



As RLQ analysis maximizes the covariance between the traits and the environmental variables mediated by the species abundances, it is important to see how the individual parts (i.e. $cov(traits, env)^2 = var(traits) \times var(env) \times cor(traits, env)^2$) of the compromise are considered. Hence, one can compare the RLQ analysis to the separate analyses which maximize independently the structure of the trait (principal component analysis of the traits), the structure of the environment (Hill-Smith analysis of the environmental variables) and the correlation (correspondence analysis of the sites-species table). These comparisons are provided by the `summary` function.

```
summary(rlq.aravo)
```

```
Eigenvalues decomposition:
     eig  covar    sdR   sdQ   corr
1 0.4280 0.6542 0.9924 1.527 0.4318
2 0.1198 0.3462 1.0991 1.145 0.2751

Inertia & coinertia R:
   inertia   max  ratio
1    0.985 1.321 0.7457
12   2.193 2.537 0.8645

Inertia & coinertia Q:
   inertia   max  ratio
1    2.331 2.409 0.9675
12   3.641 3.907 0.9319

Correlation L:
    corr    max  ratio
1 0.4318 0.8128 0.5312
2 0.2751 0.6484 0.4243
```

The correlation is quite low for the second axis (0.2753). The variance of the environmental scores is well preseved on the first two axes (89.53 %). For the traits, the amount of variance preserved (2.341 and 3.667) in nearly equal to the amount obtained in simple principal component analysis (2.409 and 3.907).

3 Fourth-corner

Fourth-corner analysis can be used to test the associations between individual traits and environmental variables. To obtain a test with a correct type I error, results of model 2 (permutation of sites, i.e. rows) and 4 (permutation of species, i.e. columns) should be combined. While the former version of `ade4` required three steps (tests with the two models and combination of the results), the combined approach can now be runned by setting the `modeltype` argument to 6. Note that we used a very high number of repetitions (`nrepet <- 49999`) to have enough power in corrected tests. This is time-consuming and could be modified to speed up the different analyses (e.g., `nrepet <- 999`). Note that by default, the `nrepet` argument of the `fourthcorner` function is set to 999.

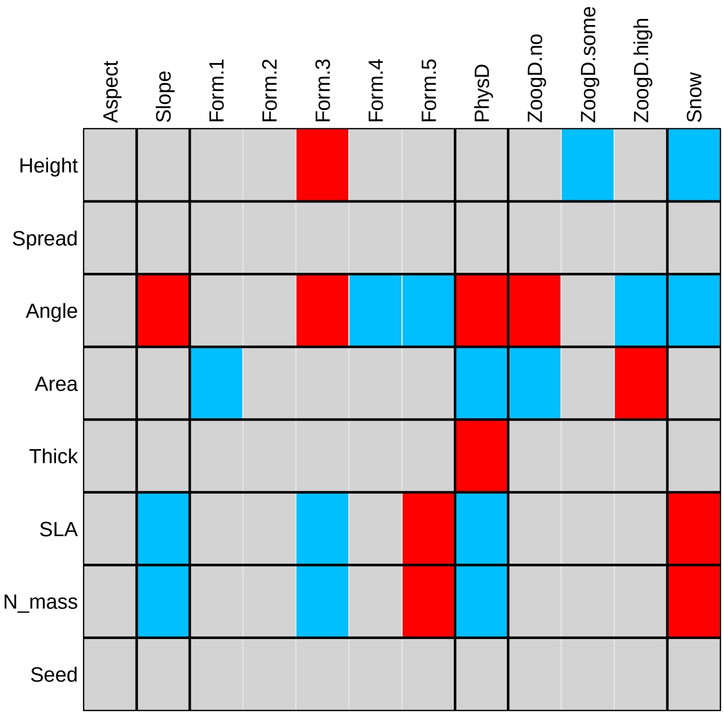
```
nrepet <- 49999
four.comb.aravo <- fourthcorner(aravo$env, aravo$spe,
    aravo$traits, modeltype = 6, p.adjust.method.G = "none",
    p.adjust.method.D = "none", nrepet = nrepet)
```

Then, results can be plotted. Blue cells correspond to negative significant relationships while red cells correspond to positive significant relationships (this can be modified using the argument `col`). In this example, there are some associations between categorical traits and quantitative environmental variable which can be measured in three different ways (Legendre et al., 1997). These three methods correspond to three possible values of the `stat` argument in the `plot` and `print` functions:

- `stat="D2"`: the association is measured between the quantitative variable and each category separately. A correlation coefficient is used to indicate the strength of the association between the given category and the small or large values of the quantitative variable.
- `stat="G"`: the association between the quantitative variable and the whole categorical variable is measured by a global statistic (F).
- `stat="D"`: the association is estimated between the quantitative variable and each category separately by a measure of the within-group homogeneity. The strength of the association is indicated by the dispersion of the values of the quantitative variable for a given category.

In the rest of the tutorial, we focus on the `D2` statistic. The correction of p-values by a sequential procedure leads to significant associations if the maximal pvalue is lower than $\alpha = 0.05$. When $\alpha = 0.05$, there are only 26 significant associations (while there are 51 when $\alpha = \sqrt{0.05}$, i.e. the biased version proposed by Dray and Legendre (2008)).

```r
plot(four.comb.aravo, alpha = 0.05, stat = "D2")
```



Now, we adjusted p-values for multiple comparisons (here we used the `fdr` method using the `p.adjust.4thcorner` function).

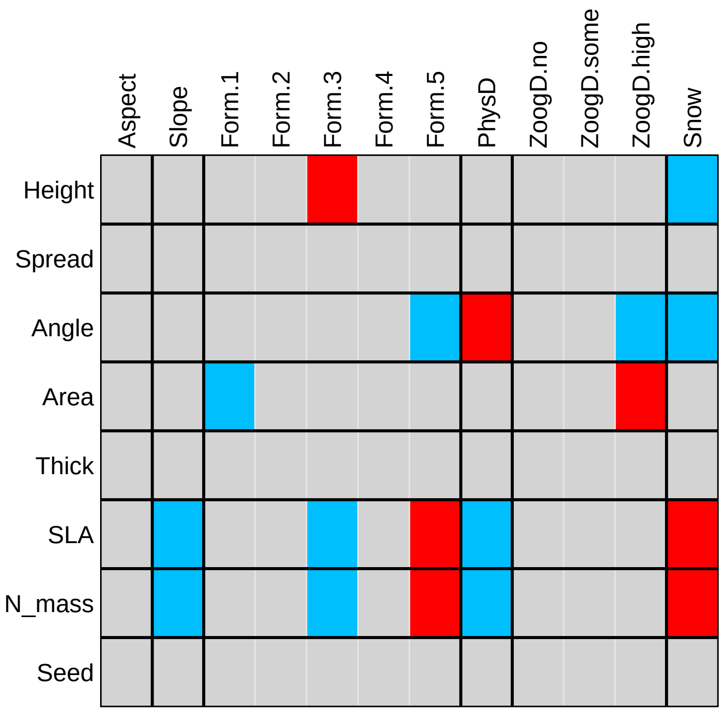
```r
four.comb.aravo.adj <- p.adjust.4thcorner(four.comb.aravo,
    p.adjust.method.G = "fdr", p.adjust.method.D = "fdr")
```

Note that, adjusted p-values can be obtained directly using the `fourthcorner` function:

```r
fourthcorner(aravo$env, aravo$spe, aravo$traits, modeltype = 6,
    p.adjust.method.G = "fdr", p.adjust.method.D = "fdr",
    nrepet = nrepet)
```

When adjusted p-values are used, there are 18 significant associations:

```r
plot(four.comb.aravo.adj, alpha = 0.05, stat = "D2")
```



4 Combining both approaches

First, a multivariate test can be applied to evaluate the global significance of the traits-environment relationships. This test is based on the total inertia of the RLQ analysis :

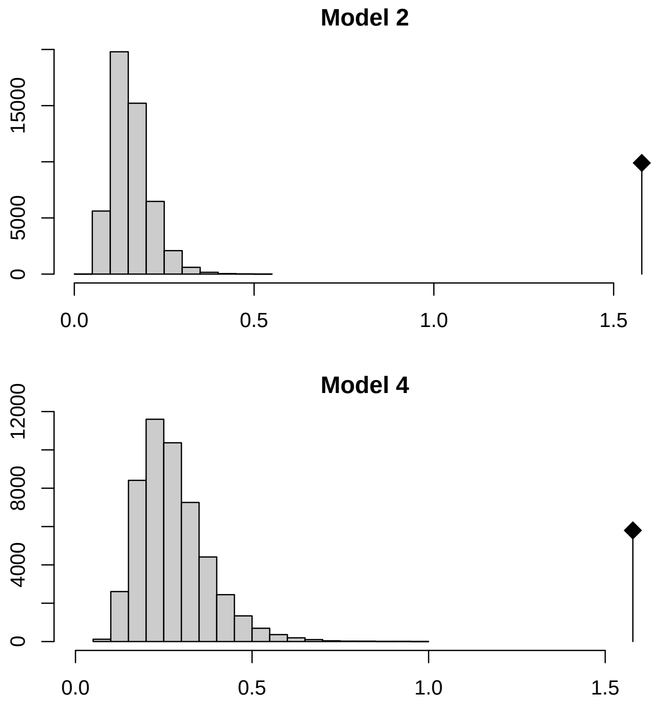
```
testrlq.aravo <- randtest(rlq.aravo, modeltype = 6, nrepet = nrepet)
```

The results are highly significant:

```
testrlq.aravo
```

```
class: krandtest
Monte-Carlo tests
Call: randtest.rlq(xtest = rlq.aravo, nrepet = nrepet, modeltype = 6)

Number of tests:   2

Adjustment method for multiple comparisons:   none
Permutation number:   49999
     Test   Obs Std.Obs   Alter Pvalue
1 Model 2 1.578   26.79 greater  2e-05
2 Model 4 1.578   13.36 greater  2e-05

other elements: adj.method call comb.pvalue
```

```
plot(testrlq.aravo)
```



The total inertia of RLQ analysis is equal to the S_{RLQ} multivariate statistic defined in Dray and Legendre (2008). This statistic is returned by the `fourthcorner2` function:

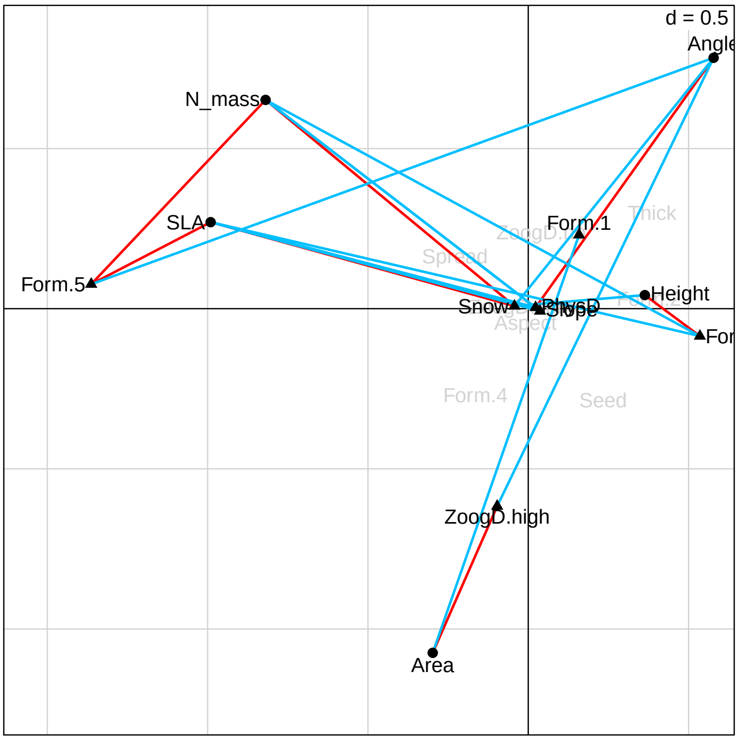
```
Srlq <- fourthcorner2(aravo$env, aravo$spe, aravo$traits,
    modeltype = 6, p.adjust.method.G = "fdr", nrepet = nrepet)
```

```
Srlq$trRLQ
```

```
Monte-Carlo test
Call: fourthcorner2(tabR = aravo$env, tabL = aravo$spe, tabQ = aravo$traits,
    modeltype = 6, nrepet = nrepet, p.adjust.method.G = "fdr")

Observation: 1.578

Based on 49999 replicates
Simulated p-value: 2e-05
Alternative hypothesis: greater

    Std.Obs Expectation    Variance
  26.739535    0.157577    0.002823
```

Both approaches can be combined if RLQ scores are used to represent traits and environmental variables on a biplot. Then, significant associations revealed by the fourthcorner approach can be represented using segments (blue lines for negative associations, red lines for positive associations, see the argument `col`). Only traits and environmental variables that have at least one significant association are represented. Here, we apply this method using adjusted pvalues for multiple comparisons and a significant level $\alpha = 0.05$.

```
plot(four.comb.aravo.adj, x.rlq = rlq.aravo, alpha = 0.05,
    stat = "D2", type = "biplot")
```



Another approach is provided by the `fourthcorner.rlq` function and consists in testing directly the links between RLQ axes and traits (`typetest="Q.axes"`) or environmental variables (`typetest="R.axes"`).

```
testQaxes.comb.aravo <- fourthcorner.rlq(rlq.aravo, modeltype = 6,
    typetest = "Q.axes", nrepet = nrepet, p.adjust.method.G = "fdr",
    p.adjust.method.D = "fdr")
testRaxes.comb.aravo <- fourthcorner.rlq(rlq.aravo, modeltype = 6,
    typetest = "R.axes", nrepet = nrepet, p.adjust.method.G = "fdr",
    p.adjust.method.D = "fdr")
```

```
print(testQaxes.comb.aravo, stat = "D")
```

```
Fourth-corner Statistics
------------------------
Permutation method  Comb. 2 and 4  ( 49999  permutations)

Adjustment method for multiple comparisons:   fdr
call:  fourthcorner.rlq(xtest = rlq.aravo, nrepet = nrepet, modeltype = 6,      typetest = "Q.axes", p.adjust.method.G = "fdr

---

             Test Stat       Obs Std.Obs     Alter  Pvalue
1  AxcR1 / Height    r  0.166172  1.6580 two-sided  0.0969
2  AxcR2 / Height    r  0.005482  0.0974 two-sided  0.9247
3  AxcR1 / Spread    r -0.095280 -0.9973 two-sided 0.32498
4  AxcR2 / Spread    r  0.034635  0.5099 two-sided 0.61842
5   AxcR1 / Angle    r  0.253857  2.5567 two-sided 0.00826
6   AxcR2 / Angle    r  0.152269  2.1152 two-sided  0.0327
7    AxcR1 / Area    r -0.119618 -1.2661 two-sided 0.20862
8    AxcR2 / Area    r -0.210731 -2.9267 two-sided 0.00278
9   AxcR1 / Thick    r  0.168446  1.6856 two-sided   0.092
10  AxcR2 / Thick    r  0.058806  0.8700 two-sided 0.39106
11    AxcR1 / SLA    r -0.434786 -4.4161 two-sided   2e-05
12    AxcR2 / SLA    r  0.063480  0.8974 two-sided 0.37562
13 AxcR1 / N_mass    r -0.360835 -3.6381 two-sided   6e-05
14 AxcR2 / N_mass    r  0.139740  1.9484 two-sided 0.04946
15   AxcR1 / Seed    r  0.107564  1.0502 two-sided 0.30206
16   AxcR2 / Seed    r -0.059631 -0.8104 two-sided 0.42866
           Pvalue.adj
1              0.1938
2              0.9247
3   0.472698181818182
4            0.659648
5             0.03304    *
6             0.10464
7             0.37088
8  0.0148266666666667    *
9              0.1938
10  0.481304615384615
11            0.00032 ***
12  0.481304615384615
13            0.00048 ***
```

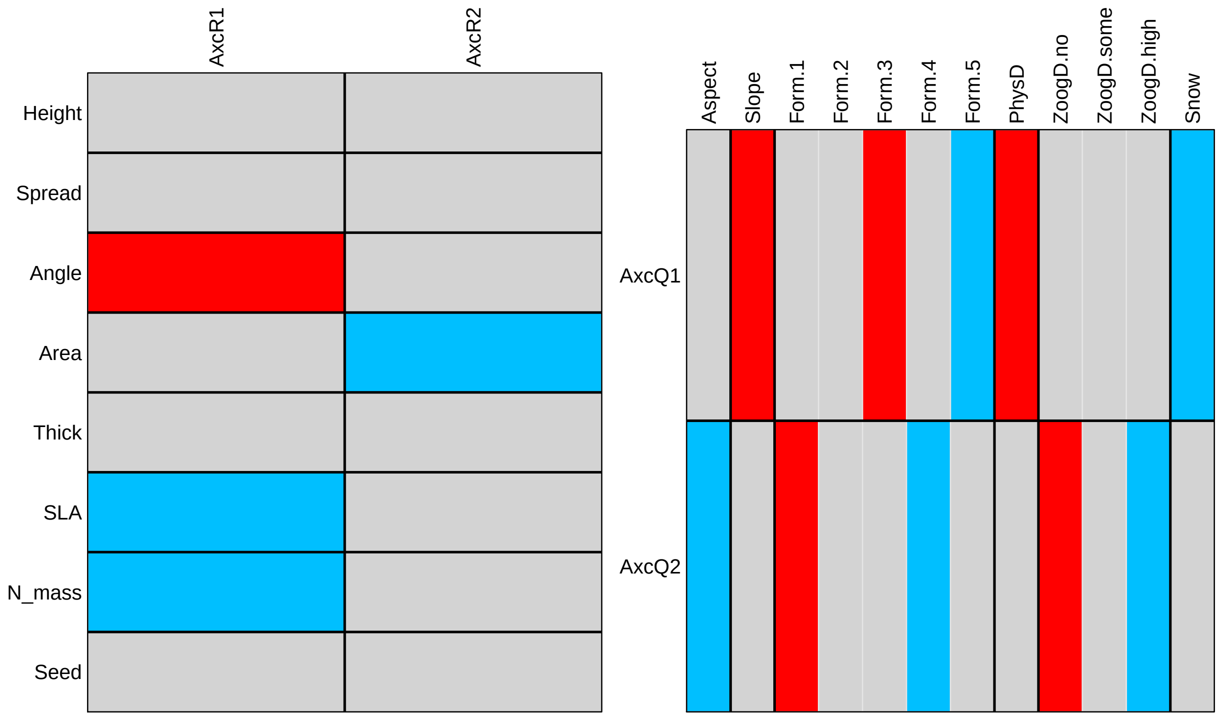
```
14  0.131893333333333
15  0.472698181818182
16  0.489897142857143
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
print(testRaxes.comb.aravo, stat = "D")
```

```
Fourth-corner Statistics
------------------------
Permutation method  Comb. 2 and 4  ( 49999  permutations)

Adjustment method for multiple comparisons:   fdr
call:  fourthcorner.rlq(xtest = rlq.aravo, nrepet = nrepet, modeltype = 6,      typetest = "R.axes", p.adjust.method.G = "fdr
---

                 Test   Stat      Obs Std.Obs     Alter  Pvalue
1      Aspect / AxcQ1      r -0.01259 -0.2301 two-sided  0.8181
2       Slope / AxcQ1      r  0.22652  4.1595 two-sided   2e-05
3      Form.1 / AxcQ1 Homog.  0.23597 -0.1475      less  0.4639
4      Form.2 / AxcQ1 Homog.  0.08358 -0.3476      less 0.42034
5      Form.3 / AxcQ1 Homog.  0.26129 -0.9018      less 0.19488
6      Form.4 / AxcQ1 Homog.  0.10294 -1.6170      less 0.01574
7      Form.5 / AxcQ1 Homog.  0.13521 -0.5296      less 0.32994
8       PhysD / AxcQ1      r  0.30944  5.6603 two-sided   2e-05
9     ZoogD.no / AxcQ1 Homog.  0.56986  2.8684      less 0.99978
10 ZoogD.some / AxcQ1 Homog.  0.30680 -1.5000      less 0.05672
11 ZoogD.high / AxcQ1 Homog.  0.12234 -1.3253      less 0.05572
12       Snow / AxcQ1      r -0.49740 -9.1299 two-sided   2e-05
13     Aspect / AxcQ2      r -0.09833 -2.2865 two-sided 0.02102
14      Slope / AxcQ2      r -0.06159 -1.2263 two-sided 0.22278
15     Form.1 / AxcQ2 Homog.  0.17597 -1.5871      less  0.0399
16     Form.2 / AxcQ2 Homog.  0.09112 -0.0249      less 0.55324
17     Form.3 / AxcQ2 Homog.  0.28641 -0.1094      less 0.48694
18     Form.4 / AxcQ2 Homog.  0.18260  0.8655      less 0.80482
19     Form.5 / AxcQ2 Homog.  0.24224  1.9512      less  0.9591
20      PhysD / AxcQ2      r  0.06799  1.1106 two-sided 0.27244
21   ZoogD.no / AxcQ2 Homog.  0.31611 -2.6342      less  0.0018
22 ZoogD.some / AxcQ2 Homog.  0.37582  0.1249      less 0.56468
23 ZoogD.high / AxcQ2 Homog.  0.24031  2.0419      less  0.9611
24       Snow / AxcQ2      r  0.12693  1.2298 two-sided  0.2239
           Pvalue.adj
1   0.908356363636364
2             0.00016 ***
3   0.618533333333333
4             0.63051
5             0.33408
6  0.0539657142857143   .
7   0.465797647058824
8             0.00016 ***
9             0.99978
10  0.123752727272727
11  0.121570909090909
12            0.00016 ***
13 0.0560533333333333   .
14            0.35824
15            0.09576   .
16   0.69882947368421
17  0.687444705882353
18  0.919794285714286
19            0.99378
20            0.40866
21            0.00864  **
22            0.71328
23            0.99978
24            0.35824
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

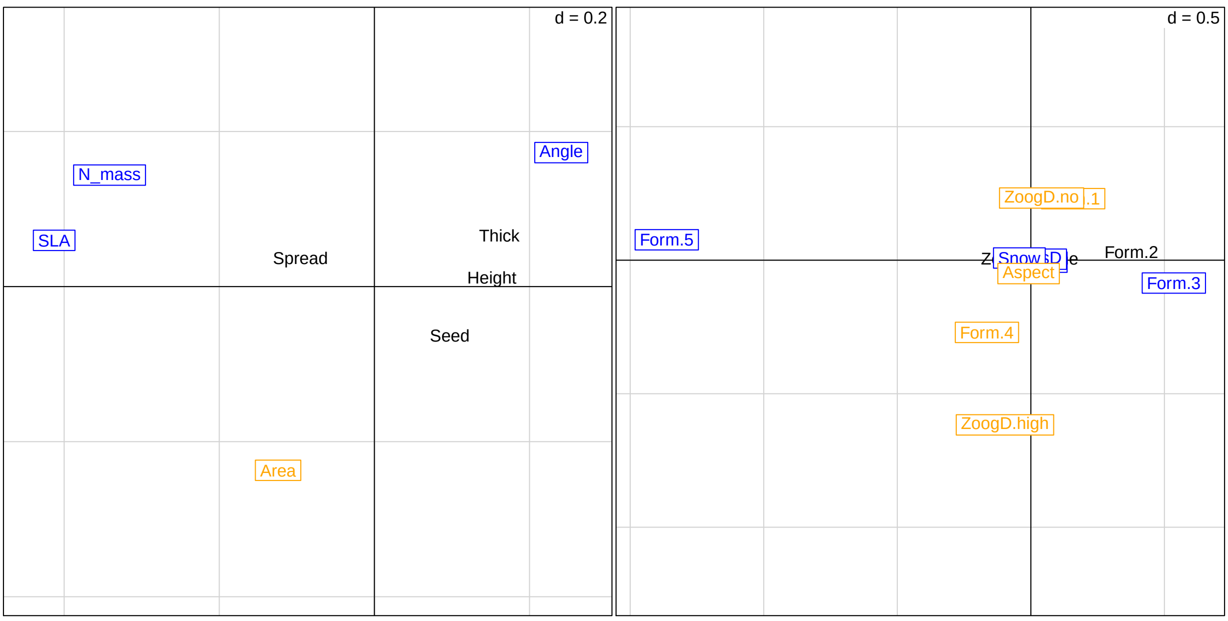
Results can be represented using a table with colors indicating significance :

```
par(mfrow = c(1, 2))
plot(testQaxes.comb.aravo, alpha = 0.05, type = "table",
    stat = "D2")
plot(testRaxes.comb.aravo, alpha = 0.05, type = "table",
    stat = "D2")
```



Significance with axes can also be reported on the factorial map of RLQ analysis. Here, significant associations with the first axis are represented in blue, with the second axis in orange, with both axes in green (variables with no significant association are in black):

```
par(mfrow = c(1, 2))
plot(testQaxes.comb.aravo, alpha = 0.05, type = "biplot",
    stat = "D2", col = c("black", "blue", "orange", "green"))
plot(testRaxes.comb.aravo, alpha = 0.05, type = "biplot",
    stat = "D2", col = c("black", "blue", "orange", "green"))
```



References

S. Dray and A. B. Dufour. The ade4 package: implementing the duality diagram for ecologists. *Journal of Statistical Software*, 22(4):1–20, 2007.

S. Dray and P. Legendre. Testing the species traits-environment relationships: the fourth-corner problem revisited. *Ecology*, 89:3400–3412, 2008.

P. Legendre, R. Galzin, and M. L. Harmelin-Vivien. Relating behavior to habitat: solutions to the fourth-corner problem. *Ecology*, 78(2):547–562, 1997.