1 Supporting Information to the paper Zelený, D. Bias in community-weighted mean analysis of
2 plant functional traits and species indicator values. *Journal of Vegetation Science.*

3 **Appendix S3. Evaluation of permutation tests using 1D simulated**

4 **community data set from Dray & Legendre (2008)**

5 **Introduction**

6 Additionally to 2D simulated community data set, I used also the simulated data set created

7 by the algorithm introduced by Dray & Legendre (2008) to evaluate the performance of

8 various permutation tests in the weighted-mean approach. This algorithm generates matrices

9 **R**, **L** and **Q** using five scenarios with various combinations of links between individual

10 matrices: scenario 1 with both **R** and **Q** linked to **L**, scenario 1N identical to scenario 1 with

11 added normal random noise, scenario 2 with **R** linked to **L**, but **Q** not linked to **L**, scenario 3

12 with **R** not linked to **L**, but **Q** linked to **L**, and scenario 4 with no links between matrices. For

13 each scenario (except 1N) I generated 1000 simulated data sets, and used each data set to

14 calculate weighted mean of species attributes and to correlate these weighted mean with

15 sample attributes (using Pearson's *r* correlation coefficient). The significance of correlation

16 was tested using *standard permutation test*, *modified permutation test* and *sequential*

17 *permutation test based on the fourth-corner statistic*. Additionally, using this simulated

18 community data sets, I conducted also a power analysis to compare the power of modified

19 permutation tests to sequential test based on the fourth-corner statistic.

20

21 **Methods of simulation study**

22 The simulated community data sets according to Dray & Legendre (2008) are constructed in

23 the following way (simplified): matrix **R** contains one randomly generated variable with

24 uniform distribution, representing sample attributes, while matrix **Q** contains one randomly

25   generated variable with uniform distribution representing species attributes. To make species

26   and sample attributes linked via species composition, matrix **L** is constructed from both

27   sample and species attributes in a way that the response of individual species abundances to

28   the environmental gradient is modelled as symmetric Gaussian curve with the optima

29   equivalent to species attributes generated in matrix **Q** and species tolerance generated as

30   random value with mean μ_{tol}. The abundance of given species in particular sample is then

31   derived as the probability of species occurrence at particular value of environmental gradient

32   (given by the sample attribute in **R**) based on constructed species response curve. For

33   technical details, see the original description in Dray & Legendre (2008). In this study, the

34   simulation algorithm was implemented into the *weimea* package (Appendix S5) as the

35   function *simul.RLQ*.

36   I created five scenarios introduced for analogical purpose in the context of the fourth-corner

37   analysis by Dray & Legendre (2008), differing from each other by the null hypothesis being

38   tested and numbered here identically to the original study. In *scenario 1*, all three matrices (**R**,

39   **L** and **Q**) are linked together by the mechanism of the simulation model described above. In

40   *scenario 1N*, normal noise was added to results of scenario 1 to increase the beta diversity of

41   compositional matrix **L** and to reduce the intensity of link between **R** or **Q** and **L**. This was

42   done by adding normal random value of mean 5 and standard deviation 1 to **R** and **Q**, and

43   normal random value of mean 0 and standard deviation 2 to values in matrix **L** (negative

44   values, which occurred in **L** after adding the random values, were replaced by zero). In

45   *scenario 2*, species composition (**L**) is linked to sample attributes (**R**), but not to species

46   attributes (**Q**). Matrices were created as in scenario 1, and afterwards the rows with species

47   attributes in matrix **Q** were permuted (cancelling the link between **L** and **Q**). In *Scenario 3*,

48   species composition (**L**) is linked to species attributes (**Q**), but not to sample attributes (**R**).

49   Matrices were created as in scenario 1, and afterwards rows with sample attributes in matrix

50   **R** were permuted (this cancels the link between **L** and **R**). Finally, in *scenario 4*, there is no

51   link between **L** and **Q**, neither between **L** and **R**. Matrices were created as in scenario 1, and

52   afterwards the rows in both matrices **R** and **Q** were permuted, cancelling all links between

53   matrices.

54   Three additional parameters can be modified within each scenario: average species

55   niche width (μ_{tol}), the number of samples (n) and the number of species in the data set (p). To

56   evaluate applicability of results in case of communities shaped predominantly by specialists

57   or generalists and data sets of different size, I calculated each scenario for three values of the

58   average species tolerances ($\mu_{tol} = \{10, 30, 60\}$), three sample sizes ($n = \{30, 50, 100\}$) and

59   three sizes of the species pool ($p = \{30, 50, 100\}$). For each scenario (except scenario 1N) and

60   each combination of three additional parameters I created 1000 artificial data sets, and for

61   each data set I calculated Pearson's r correlation coefficient between the community-weighted

62   mean of species attributes and sample attributes. The significance of each correlation was

63   tested using standard and modified permutation test and the proportion of results significant at

64    $\alpha = 0.05$ was calculated.

65   To evaluate the power of individual permutation tests in more detail, I performed

66   power analysis analogous to that on Fig. 4 of Dray & Legendre (2008). Using only scenario

67   1N, I created the first set of simulated data sets with different number of species (10 to 100,

68   step 10) and different number of samples (10 to 100, step 10) while keeping the average

69   species tolerances constant ($\mu_{tol} = 30$). Additionally, to see also relationship between test

70   power and beta diversity of the data set (modified by changing the niche width of simulated

71   species), I created also the second set of simulated data sets, with the number of species either

72   30, 50, or 100, and the number of samples either 30, 50 or 100, and varying the average

73   species tolerances (μ_{tol}) from 10 to 80 (with step 10). For each combination of the number of

74   species, the number of samples and species tolerance, I created 1000 simulated data sets,

75   performed modified permutation test and sequential permutation test based on the fourth-

76   corner statistic, and counted the number of significant results ($\alpha = 0.05$).

77

78   **Results of the simulation study**

79   Standard permutation test has inflated Type I error rate for scenario 2 and correct rates for

80   scenarios 3 and 4 (Appendix S3: Table S1). Modified permutation test has correct rate for

81   scenarios 2 and 4, and overly conservative rate for scenario 3 (almost no significant results).

82   Note, however, that community data simulated by 1D algorithm are more homogeneous that

83   those of 2D one. Even the most heterogeneous 1D data set has an average beta diversity 0.413

84   ($\mu_{tol} = 10$, number of species 30 and number of sites 30, see Appendix S3: Table S2), while

85   the most homogeneous 2D data set has average beta diversity 0.621 (Appendix S1: Table S1).

86   The power of the standard permutation test (measured by rejection rate for scenario 1,

87   for which the null hypothesis is true) is high for all data set sizes and both narrow and wide

88   species niches (being 1.000 in all cases, Appendix S3: Table S1). In the case of modified

89   permutation test, the power is high for simulation model with narrow species niches even for

90   smaller data sets. In the case of wider species niches, the power of the test is lower, and is

91   further decreasing with decreasing size of the data set (number of species or number of sites).

92   The lowest power recorded, 0.545, is for the scenario with the widest species niches and

93   lowest numbers of species and sites ($\mu_{tol} = 60$, number of species 30 and number of sites 30;

94   Appendix S3: Table S1).

95   Detailed power analysis based on the data set with added normal noise (scenario 1N)

96   (with $\mu_{tol} = 30$ and the numbers of species and samples ranging from 10 to 100) show that

97   both modified permutation test and sequential permutation test with the fourth-corner statistic

98   has similar power. The power decreases with decreasing number of species and samples in the

99   data set (Appendix S3: Figs. S1a, c), and also with decreasing beta diversity (Appendix S3:

100   Fig. S1b, d; note that in 1D algorithm, the beta diversity decreases with increasing species

101   tolerance).

102   **References**

103   Dray, S. & Legendre P. 2008. Testing the species traits-environment relationships: the fourth-

104   corner problem revisited. *Ecology* 89: 3400–3412.

105   ter Braak, C.J.F., Cormont, A. & Dray, S. 2012. Improved testing of species traits-

106   environment relationships in the fourth-corner problem. *Ecology* 93: 1525–1526.
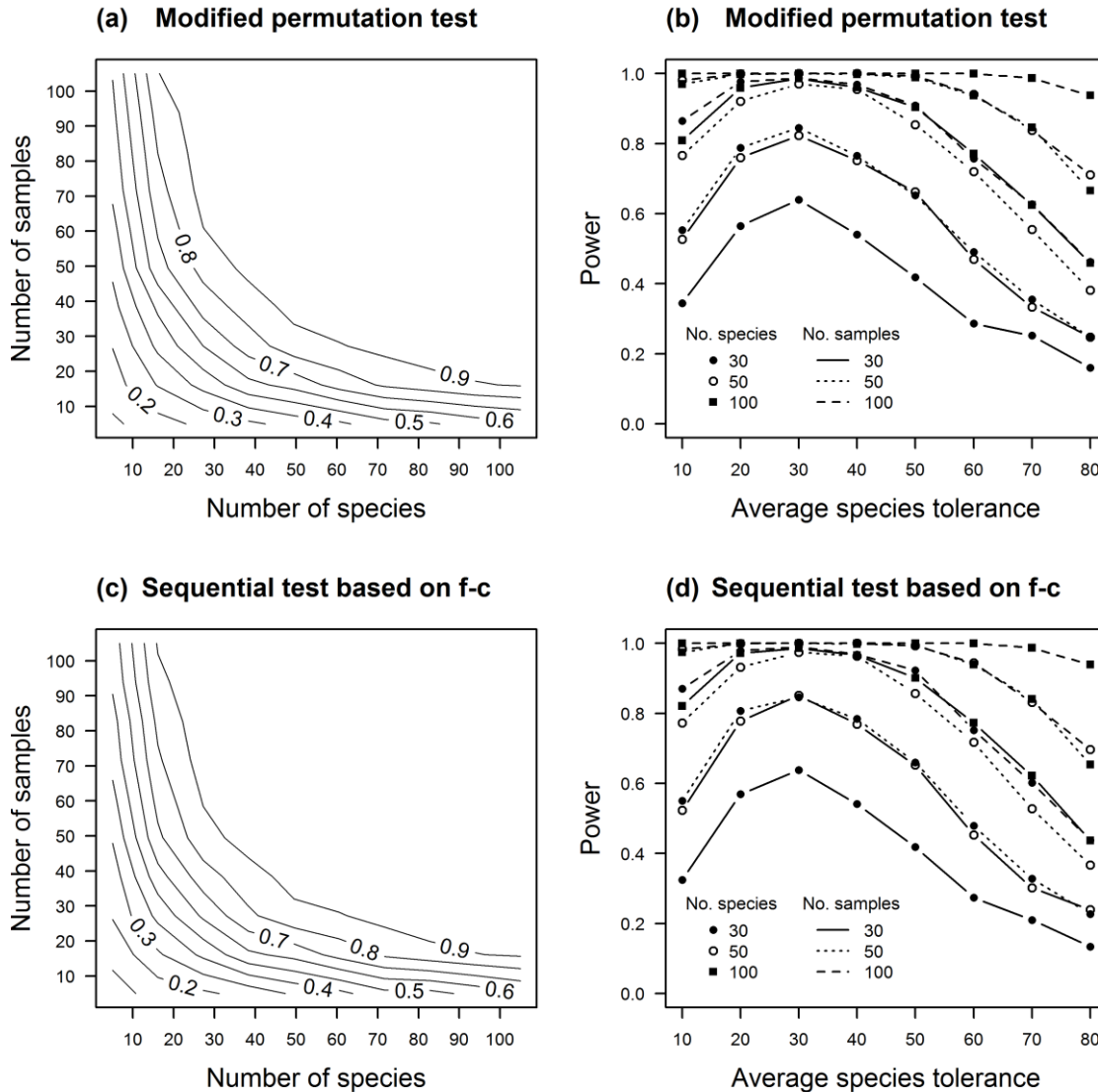
107

Appendix S3: Table S1

109 Evaluation of standard and modified permutation test examining the relationship of weighted
110 mean of species attributes with sample attributes, based on the simulation study designed
111 according to 1D simulated model of Dray & Legendre (2008). The simulation varies three
112 parameters: the average simulated species niche width (μ_{tol}), the number of sites in the data set
113 (# of sites) and the number of species in the data set (# of species). Rejection rates represent
114 the proportion of significant results at $\alpha = 0.05$ (using 1000 simulated data sets). In scenario 1,
115 all three matrices (**R**, **L** and **Q**) are linked together (this represents the power test for both
116 standard and modified permutation test). In scenario 2, **R** is linked to **L**, but **Q** is not linked to
117 **L**, while in scenario 3 **Q** is linked to **L** but **R** is not linked to L. In scenario 4, both **R** and **Q**
118 are not linked to **L**. The bold numbers are those used on Appendix S3: Fig. S2.

			Rejection rate for individual scenarios							
			Standard permutation test				Modified permutation test			
μ_{tol}	# of sites	# of species	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 1	Scenario 2	Scenario 3	Scenario 4
10	30	30	1.000	0.750	0.047	0.059	0.999	0.039	0.008	0.038
10	30	50	1.000	0.762	0.051	0.049	1.000	0.049	0.007	0.044
10	30	100	1.000	0.711	0.049	0.061	1.000	0.042	0.003	0.066
10	50	30	1.000	0.803	0.047	0.046	1.000	0.051	0.007	0.043
10	50	50	1.000	0.799	0.047	0.051	1.000	0.043	0.004	0.049
10	50	100	1.000	0.808	0.051	0.040	1.000	0.057	0.004	0.044
10	100	30	1.000	0.864	0.045	0.046	1.000	0.053	0.003	0.031
10	100	50	1.000	0.845	0.053	0.043	1.000	0.046	0.004	0.039
10	100	100	1.000	0.861	0.043	0.044	1.000	0.050	0.001	0.049
30	30	30	1.000	0.862	0.052	0.050	0.894	0.054	0.001	0.054
30	30	50	1.000	0.872	0.058	0.049	0.969	0.066	0.001	0.049
30	30	100	1.000	0.888	0.055	0.045	0.999	0.048	0.002	0.043
30	50	30	1.000	0.913	0.043	0.035	0.918	0.055	0.002	0.047
30	50	50	1.000	0.904	0.056	0.054	0.989	0.051	0.000	0.049
30	50	100	1.000	0.904	0.053	0.047	0.999	0.049	0.000	0.042
30	100	30	1.000	0.931	0.040	0.049	0.908	0.051	0.000	0.046
30	100	50	1.000	0.922	0.035	0.052	0.978	0.056	0.000	0.044
30	**100**	**100**	**1.000**	**0.935**	**0.062**	**0.046**	**0.999**	**0.049**	**0.001**	**0.051**
60	30	30	1.000	0.961	0.045	0.047	0.554	0.055	0.000	0.051
60	30	50	1.000	0.968	0.055	0.047	0.717	0.037	0.000	0.053
60	30	100	1.000	0.957	0.057	0.057	0.852	0.057	0.000	0.059
60	50	30	1.000	0.970	0.050	0.048	0.547	0.055	0.000	0.057
60	50	50	1.000	0.968	0.041	0.045	0.710	0.045	0.000	0.052
60	50	100	1.000	0.976	0.051	0.057	0.867	0.061	0.000	0.046
60	100	30	1.000	0.979	0.042	0.048	0.530	0.046	0.001	0.040
60	100	50	1.000	0.973	0.051	0.037	0.704	0.044	0.003	0.059
60	100	100	1.000	0.980	0.041	0.045	0.863	0.039	0.000	0.054

119

Appendix S3: Table S2

Beta diversity of the matrix of species composition for each combination of species tolerance × number of sites × number of species in the 1D simulation model of Dray & Legendre (2008), calculated as a variation in the dissimilarity matrix with squared values of Whittaker's index of association among samples. Beta diversity is reported separately for scenarios without added noise (scenarios 1, 2, 3 and 4) and with added normal random noise (scenario 1N). The simulation varies three parameters: the average simulated species niche width (μ_{tol}), the number of sites in the data set (# of sites) and the number of species in the data set (# of species). The bold numbers are beta diversities of the data set for the combination of parameters used on Appendix S3: Fig. S2.

| μ_{tol}	# of sites	# of species	Beta diversity of matrix **L** for scenarios 1, 2, 3, 4	Beta diversity of matrix **L** for scenario 1N
10	30	30	0.413±0.065	0.475±0.018
10	30	50	0.410±0.053	0.471±0.014
10	30	100	0.408±0.041	0.468±0.009
10	50	30	0.409±0.061	0.474±0.014
10	50	50	0.409±0.048	0.471±0.011
10	50	100	0.407±0.038	0.467±0.007
10	100	30	0.409±0.054	0.474±0.010
10	100	50	0.412±0.045	0.470±0.007
10	100	100	0.408±0.033	0.468±0.006
30	30	30	0.141±0.032	0.433±0.017
30	30	50	0.140±0.028	0.429±0.013
30	30	100	0.141±0.024	0.427±0.010
30	50	30	0.139±0.030	0.433±0.014
30	50	50	0.140±0.025	0.430±0.011
30	50	100	0.140±0.020	0.428±0.008
30	100	30	0.140±0.029	0.434±0.010
30	100	50	0.141±0.024	0.430±0.008
30	**100**	**100**	**0.142±0.017**	**0.427±0.005**
60	30	30	0.019±0.005	0.400±0.017
60	30	50	0.020±0.005	0.396±0.013
60	30	100	0.019±0.004	0.394±0.009
60	50	30	0.019±0.005	0.400±0.013
60	50	50	0.019±0.004	0.397±0.010
60	50	100	0.019±0.003	0.394±0.007
60	100	30	0.019±0.005	0.399±0.009
60	100	50	0.019±0.004	0.396±0.007
60	100	100	0.019±0.003	0.394±0.005

131

(a) **Modified permutation test**

(b) **Modified permutation test**

(c) **Sequential test based on f-c**

(d) **Sequential test based on f-c**

132

Appendix S3: Figure S1

The detailed power analysis of modified permutation test (a, b) and sequential permutation test based on the fourth-corner statistic (c, d). Panels (a) and (c) show the dependence of the power (contour) on given combination of the number of species (*x*-axis) and the number of samples (*y*-axis), all with fixed $\mu_{tol} = 30$. Panels (b) and (d) show dependence of power (*y*-axis) on species tolerance (μ_{tol}, ranging from 10 to 80) for a given combination of the number of species and the number of samples. Power is expressed as the proportion of results significant at $\alpha = 0.05$ from 1000 simulated data sets based on scenario 1N from Dray & Legendre (2008).

(a) Weighted-mean approach   **(b) Fourth-corner approach**

143
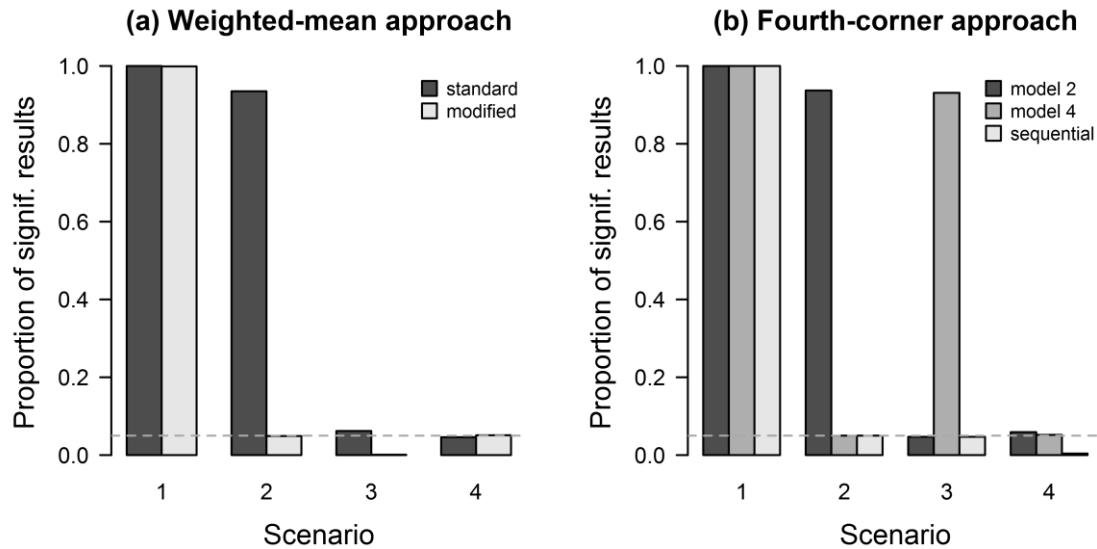
Appendix S3: Figure S2

Comparison of permutation tests used in (a) weighted-mean approach with Pearson's
correlation, and (b) the fourth-corner approach. Based on simulated community data sets
according to 1D simulated model of Dray & Legendre (2008), using 1000 simulated data sets
with $\mu_{tol} = 30$, the number of species $p = 100$ and the number of samples $n = 100$. The
proportion of significant results represents the rejection rate at $\alpha = 0.05$ for standard and
modified permutation test in weighted-mean approach and model 2, model 4 and sequential
permutation test in the fourth-corner approach. The values in the barplot for the fourth-corner
approach (panel b) are from Table 1 in Dray & Legendre (2008) for model 2 and model 4
tests, and from Table A1 in ter Braak et al. (2012) for the sequential permutation test.