Final report on the project entitled
„Exploring the role of intraspecific trait variation in assembly of terrestrial plant communities”

Table of content

1 Introduction

The main questions of the project were the following:

1. What is the amount of ITV in the studied system? How is it structured?
2. Which method is the most appropriate for integrating ITV into community assembly studies?
3. Does considering ITV change our conclusion considerably?
4. Does trait variation between sites hinder the extrapolation of our results and predictions received by using methods neglecting ITV?

The planned work was divided into five work packages:

- WP1: Components of ITV: comparisons between traits and between species
- WP2: Spatial extrapolation of results in trait-based studies
- WP3: Is methods incorporating ITV free from the Jack Horner and Narcissus effect?
- WP4: Exploring assembly rules by methods incorporating ITV: a case study
- WP5: Partitioning within population ITV into stochastic and deterministic parts

In this report first the results achieved in each WP are shown, and next a summary is given on the unplanned works related to the topic of this project.

The pandemic hinders our fieldwork, therefore data collection in WP1 and WP2 was finished later than expected. In WP4 the trait measurement needed more fieldwork than expected. The PI has worked on unplanned studies related to the project (see details below) that resulted in delays in the data analysis. These delays together resulted in the publication of some of our results are still under preparation. On the other hand, unplanned works resulted in publications in leading journals. During the work, we recognized that some of the planned analyses are meaningless. They are omitted, but overall the work done is not less than planned.

2 WP1: Components of ITV: comparisons between traits and between species

In this WP we used measurements in the Orgovány site (see WP4) and collected data from two other sites (Kunadacs, Vácrátót) with similar water-availability gradients and overlapping species pools, but different macroclimates. 34 species were selected that occur in all three sites representing the variation of occurring species in their growth form (forbs and grasses), life span (annuals and perennials), and habitat preference (species of dry grasslands and wet meadows), and specialist/generalist character (measured by a co-occurrence based specialization index; Botta-Dukát 2012). Height and leaf traits (size, SLA, LDMC) were involved in this study.

Variance partitioning analysis was carried out on a general linear model with two crossed factors: species and site (a similar analysis was done by Messier et al. 2017, but they considered more nested scales). Then variation (total sum of squares) will be partitioned within (SS_W) and between (SS_B) site components for each species separately. Trait variation of species was characterized by the relative contribution of between-site variation measured by $\ln(SS_B/SS_W)$ (analogous to aITV in Siefert et al. 2015). Finally, we explored how these two characteristics of ITV depend on characteristics of the species (growth form, life span, habitat preference and specialist character) by fitting conditional inference trees (Hothorn et al. 2006).

Comparing within- and between-site variations between traits would be meaningless because of the different measurement units. But their relative contribution ($\text{aITV} = \ln(SS_B/SS_W)$) is comparable. For SLA the between-site component is much higher than within site component ($\text{aITV} >> 0$), while for the other three traits the two components proved to be equally important (aITV values are near zero) (Figure 1).

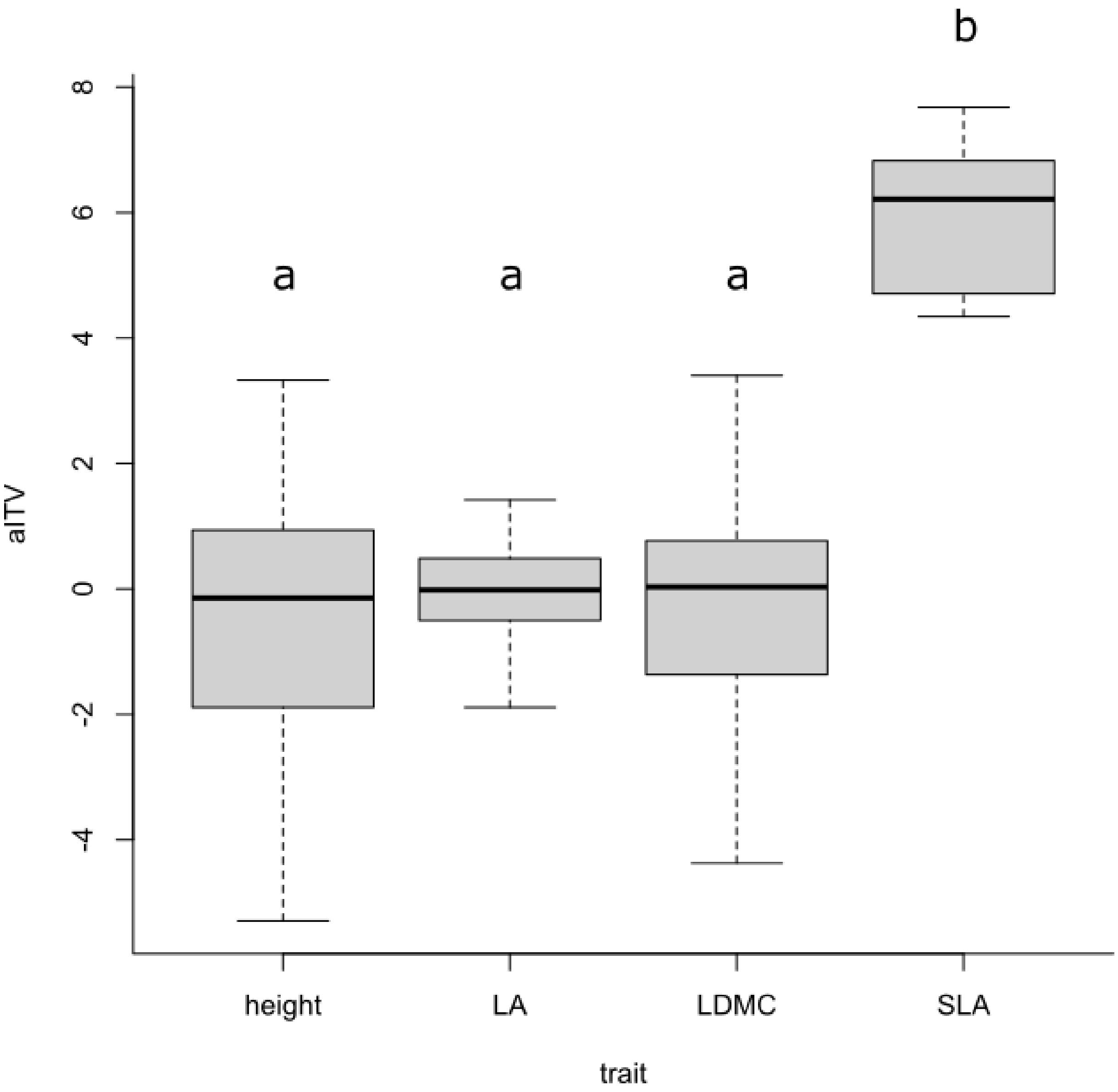


Figure 1. The relative contribution of within- and between-site variation to ITV in the four studied traits. Positive values indicate that the between-site component is larger. Different letters indicate significant differences according to the Kruskal-Wallis test and the subsequent Dunn posthoc test.

The effect of species' characteristics varies among traits and ITV components. Here only the significant differences are summarized. Within-site variation of height is higher in annual than in perennial species (Figure 2). Between-site variation of SLA is higher in generalists than in specialist species. Both within- and between-site variation of LDMC is higher in forbs than in graminoids (Figures 4-5). The difference is slightly higher in the between- than in the within-site component, therefore the relative role of the former is higher for forbs, however, the difference is only marginally significant (Figure 6).

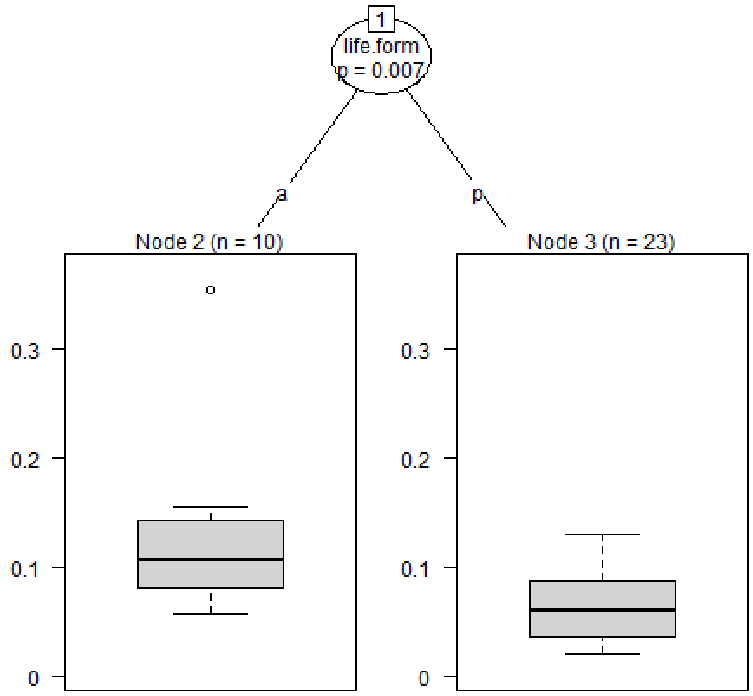


Figure 2: Effect of species characteristics on the within-site variation of height

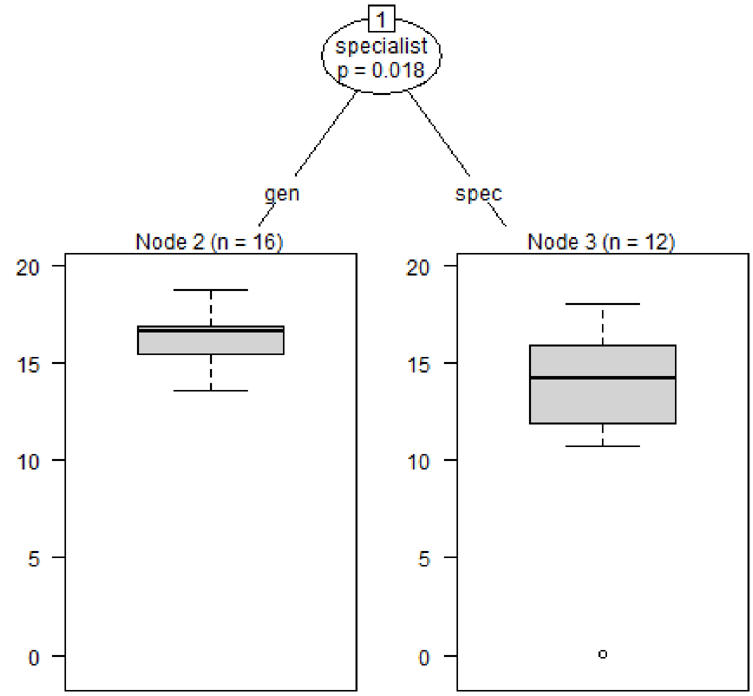


Figure 3: Effect of species characteristics on the within-site variation of SLA

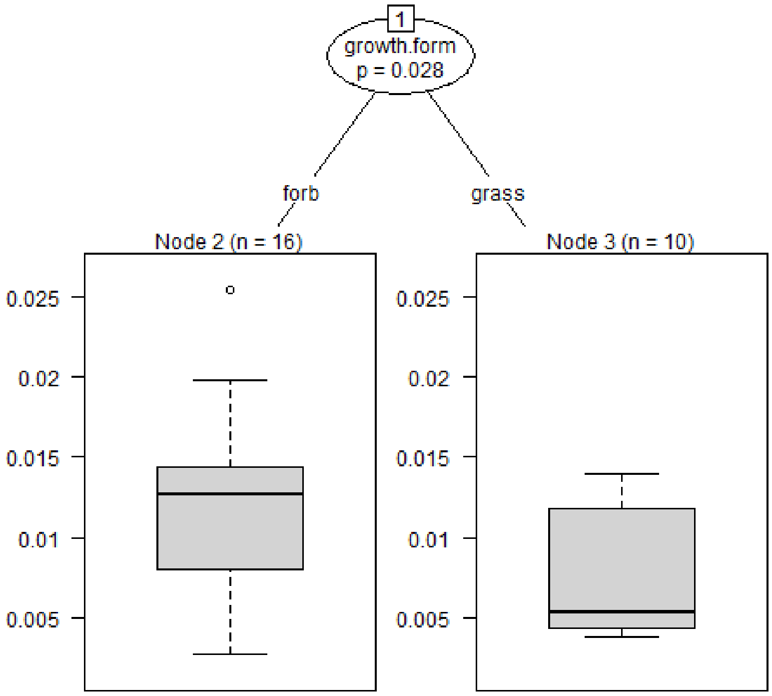


Figure 4: Effect of species characteristics on the within-site variation of LDMC

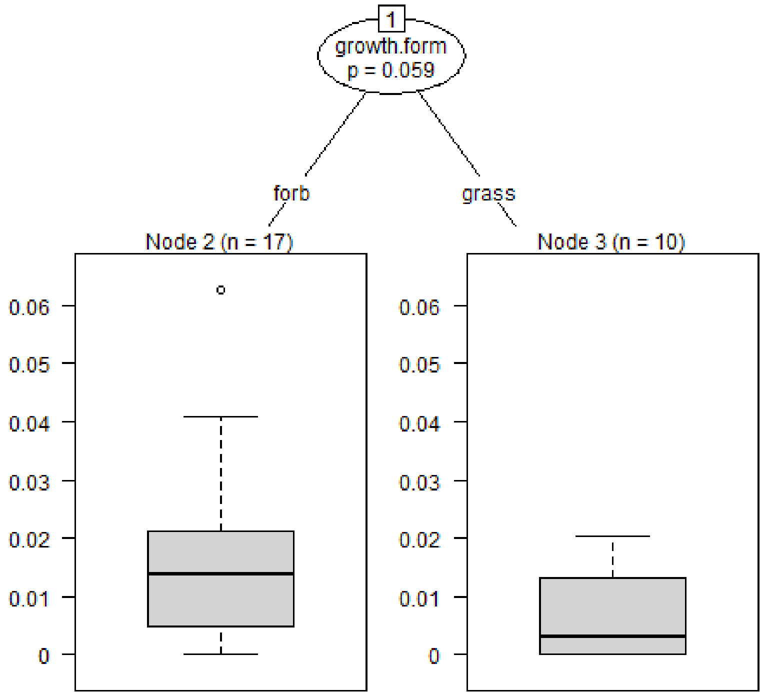


Figure 5: Effect of species characteristics on the between-site variation of LDMC

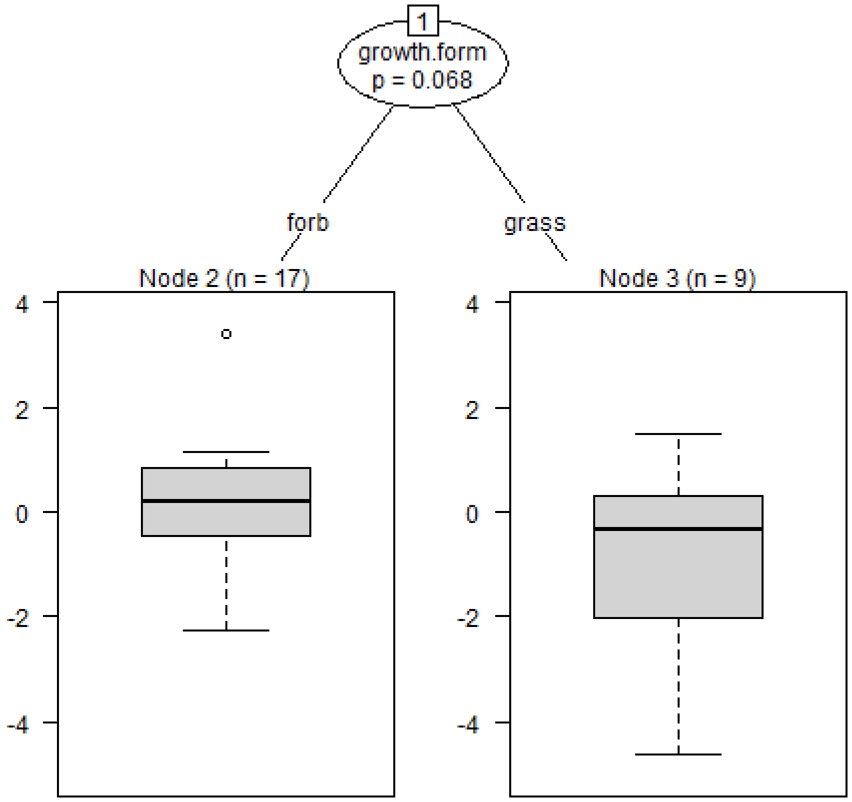


Figure 6: Effect of species characteristics on aITV of LDMC

3 WP2: Spatial extrapolation of results in trait-based studies

We have repeated the sampling had been done in Orgovány in three other sites: two sites (Kunadacs, Ásotthalom) are in the same region, while the third one (Deliblát) is situated in geographically distinct region, but represent similar abiotic conditions. In each site, plots were arranged along a productivity gradient from open sand grasslands to *Deschampsion* and *Molinion* meadows (except for Deliblát, where meadows were missing and closed tall-grass steppes were the wettest habitat). In 2x2 m plots the cover of vascular plants was visually estimated and NDVI was measured. Traits of species missing from Orgovány are measured in individuals collected from the other three sites. The same species mean trait values are used for each species.

In the analysis, we followed the methods of Lhotsky et al. with small improvements. Each trait is evaluated separately. Functional distances between species were calculated by Gower distance after log transformation. We have recognized that Rao's quadratic entropy is sensitive to outlier trait values, we replaced it with median distance (see *Unplanned results* section). Effect sizes were calculated for each plot using the method suggested by Botta-Dukát (2018). The null model was reshuffling species within the local species pool (i.e. species recorded in that site) (Götzenberger et al. 2016).

We used hierarchical general additive models (HGAM) (Pedersen et al., 2019) to check if the same effect size - NDVI relationship appears in all sites. HGAM approach means fitting six models: no trend at all (null model), site-specific trends with different smoothness (model I), site-specific trends with similar smoothness (model S), global trend + site-specific effects with different smoothness (model GI), global trend + site-specific effects with similar smoothness (model GS), and only global trend (G). The best one from the six fitted models was selected based on the AICc criteria (Zuur et al., 2009). The null model always contradicts the stress-gradient hypothesis, while the other five models may support it (depending on the shape of the fitted curve). If null-model or model G is the best, it means full transferability of local results to other sites. If models GS or GI prove to be the best, there is a common trend that allows generalization, but the local peculiarities cannot be neglected. Finally, if models S or I are the most parsimonious, we cannot make any general conclusion.

Table 1. AICc table comparing model fits for different sites.

Model	**Height**	**Leaf size**	**Seed weight**	**SLA**
null	842.660	907.476	**914.271**	914.271
G	806.478	842.532	970.602	900.406
GS	*805.943*	822.101	971.348	902.001
GI	**804.131**	**796.082**	975.666	*897.620*
S	817.553	830.121	982.056	904.107
I	807.436	*796.084*	976.916	**897.619**

For height and leaf size, the GI model proved to be the best (Table 1). For SLA model I is slightly better than model GI, but the difference is less than 0.1, therefore GI model was interpreted. (The rule of thumb is that difference in AIC less than two is negligible.) For seed weight null model proved to be the best, therefore in this trait, there is no clear trend of trait convergence/divergence along the productivity gradient.

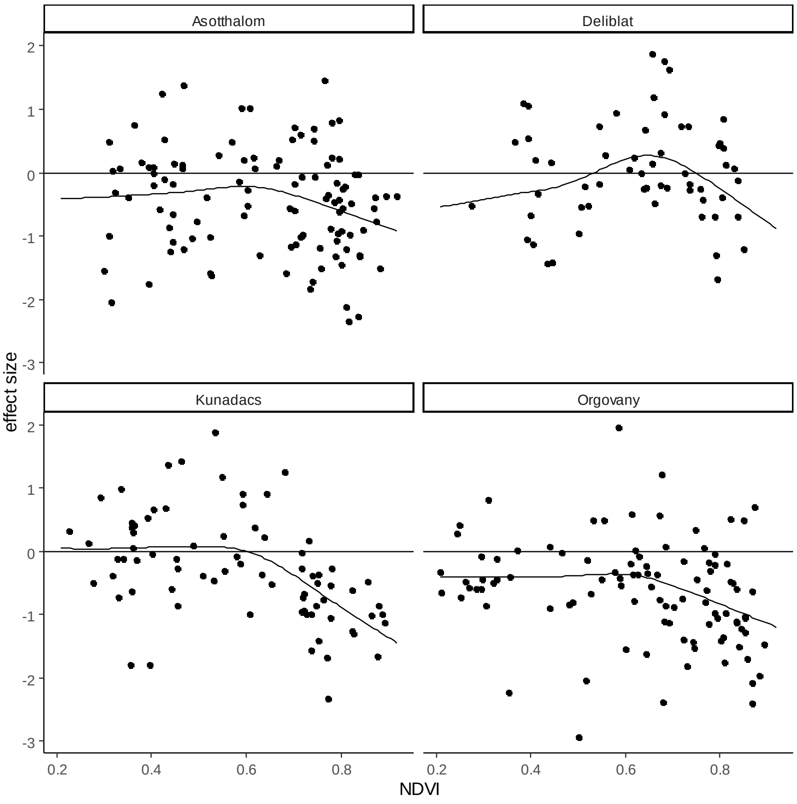


Figure 7. Effect sizes in plot-level tests of trait convergence/divergence based on model GI along productivity gradient (NDVI) for height trait in different sites

Trends of effect sizes of height (Figure 7) are similar for the four sites, but contradict the stress gradient hypothesis: effect sizes are near zero at low or medium productivity (random pattern) and decrease at high productivity (trait convergence). This stronger convergence at high productivity is probably a consequence of stronger competition for light.

In the case of leaf size, the shape of fitted curves varies considerably among sites, but there is a common trend (except in Kunadacs) of increasing effect sizes along the productivity gradient that supports the stress gradient hypothesis (Figure 8).

The trend of effect sizes calculated for SLA varies considerably among sites (Figure 9). There was no trend in Orgovány and Deliblát, the values are near zero indicating random assembly. In Ásotthalom there is a decrease at low productivity and then the curve is flat, while in Kunadacs the curve is flat first and then decreases. Both curves are decreasing and thus falsify the stress gradient hypothesis, but the first changes from divergent to random, while the second changes from random to convergent patterns.

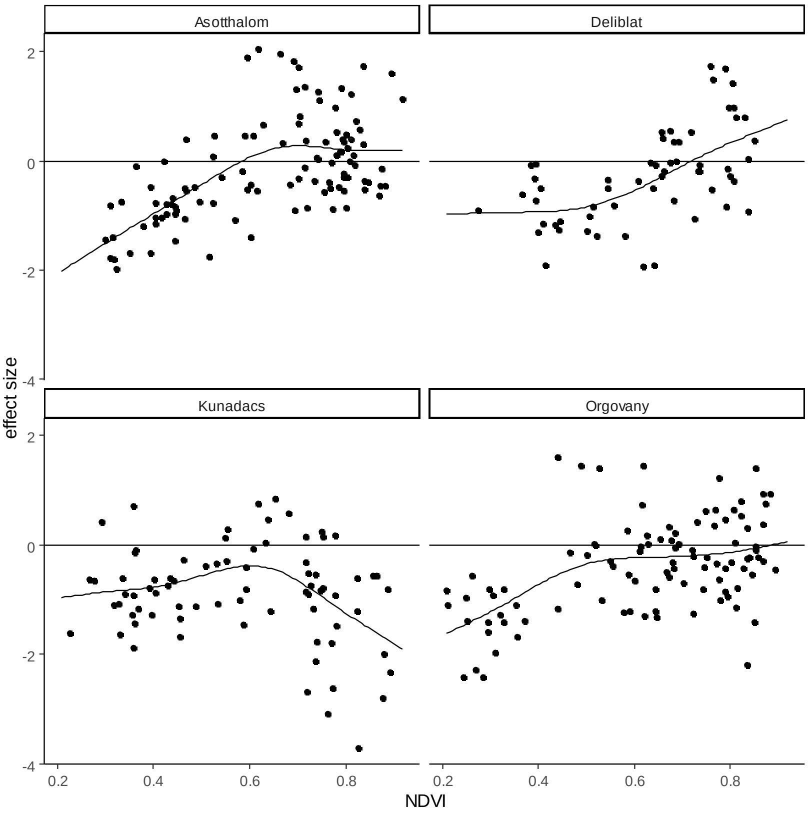


Figure 8. Effect sizes in plot-level tests of trait convergence/divergence based on model GI along productivity gradient (NDVI) for leaf size trait in different sites

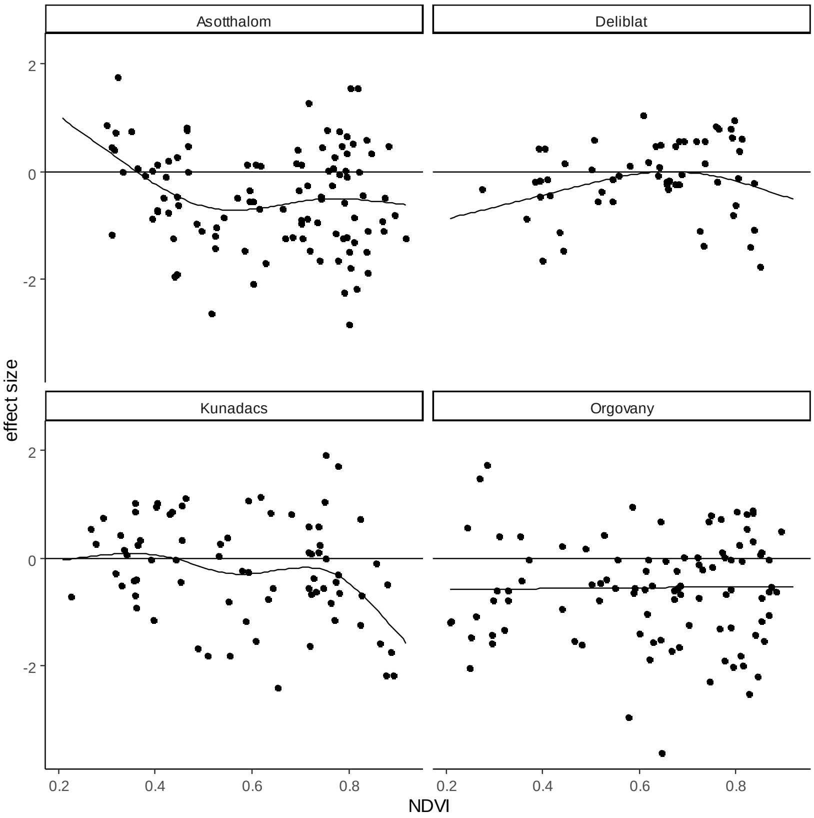


Figure 9. Effect sizes in plot-level tests of trait convergence/divergence based on model GI along productivity gradient (NDVI) for SLA trait in different sites

We had to recognize that contrary to the work plane, Traitspace is not suitable for our data, because it needs measuring traits of all occurring individuals, which restricts its use to tree communities. Joint distribution modelling also did not prove a suitable way for the analysis, because the linear trait-environment relationship is oversimplified (see details in WP3), while the more complicated models often cannot be fitted due to convergence problems.

The collateral result of the work with joint distribution models is that we proved that model-based ordination more effectively summarizes the information than its distance-based alternative. For example, in the distance-based ordination of Orgovány data 19 axes proved to be more important than the random expectation (Figure 10). Interpreting so much axis is impossible, but plotting only the first two axes may result in a considerable loss of information. In model-based ordination models with quadratic relationship and one latent variable proved to be the best (i.e. its AIC was the lowest) (Figure 11). Although we visually estimated the cover in percentage, it was converted to “number of individuals” assuming that the lowest recorded positive cover means one individual and the relationship between cover and number of individuals is linear. The transformed values can have been modelled by negative binomial distribution: the PIT residuals of the fitted models remain beyond the confidence band (Figure 12).

If presence-absence data were analyzed (figures not shown), the number of axes should be interpreted also much lower in model-based ordination (again one) than in the distance-based analysis (seven axes), however, the difference is not so striking here.

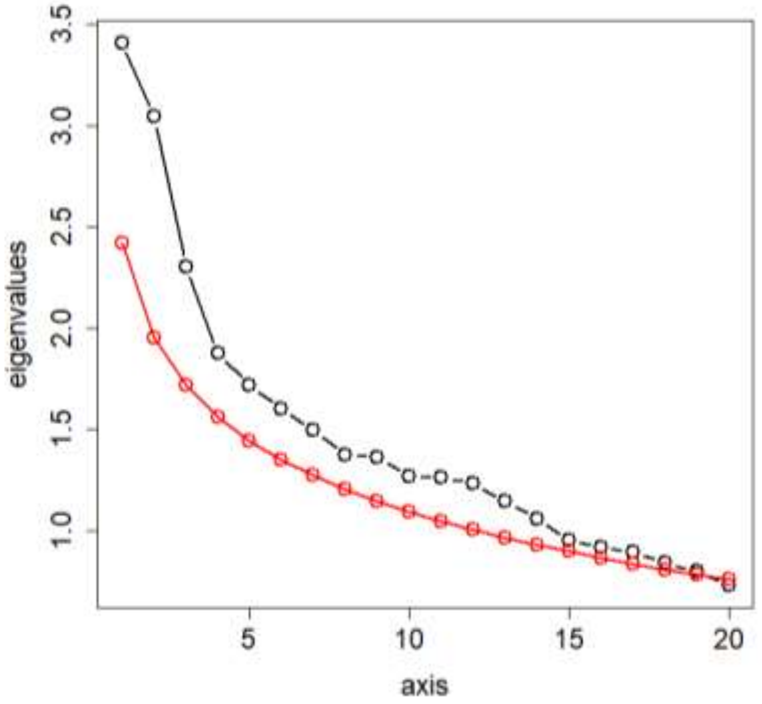


Figure 10. Eigenvalues in the ordination of cover data in the Orgovány site by metric multidimensional scaling using Bray-Curtis distance (black dots and line) compared to random expectation based on broken stick distribution. Axis with eigenvalues higher than random expectation – in this case, 19 axes - should be interpreted.

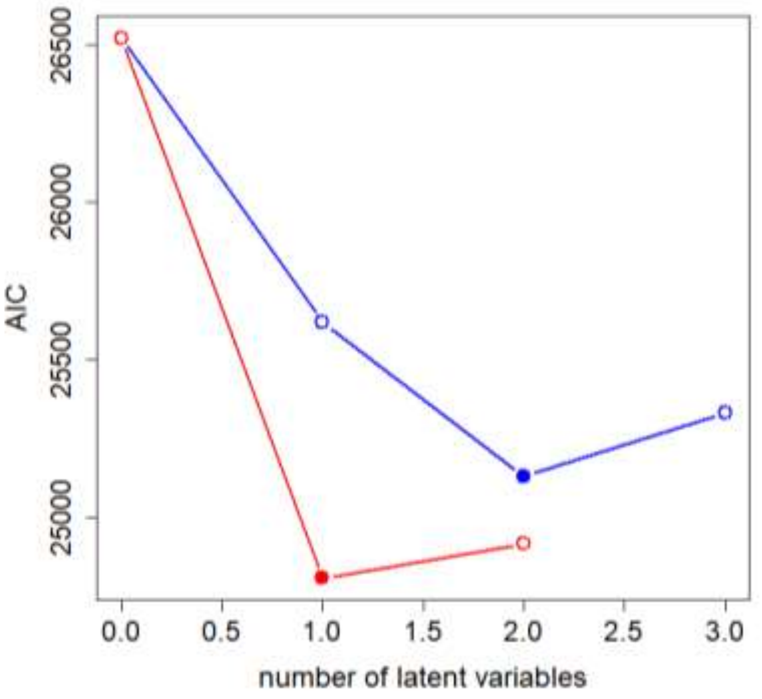


Figure 11. AIC values of model-based ordination fitted to Orgovány cover data supposing linear (blue line and dots) and quadratic response (red line and dots). Filled dots indicates the optimal number of latent variables (=ordination axes).

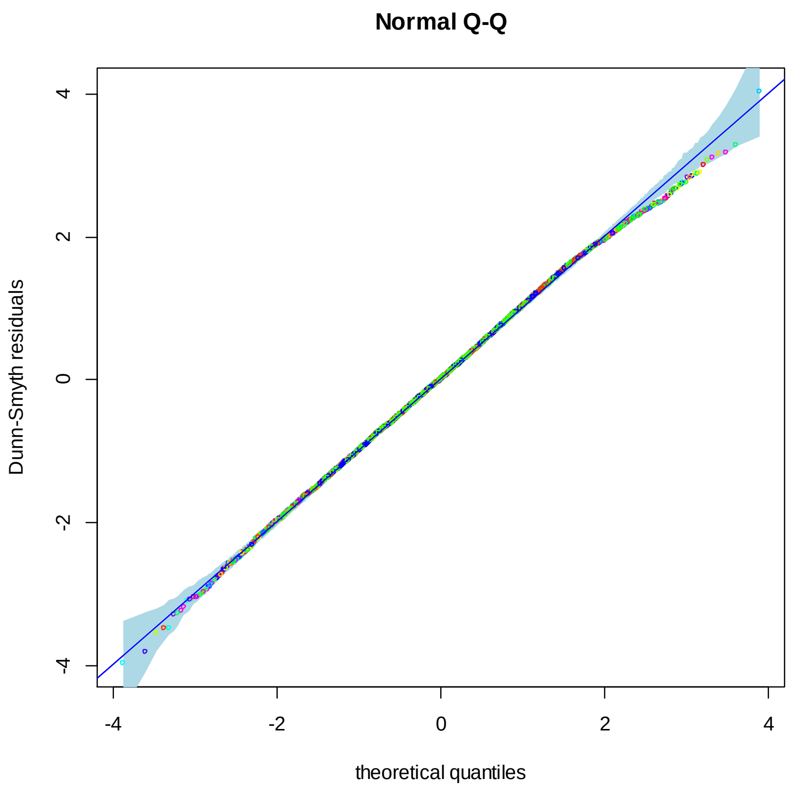


Figure 12. QQ-plot of residuals of model-based ordination fitted to cover data after transformation to “pseudo”-abundances using the negative binomial distribution.

4 WP3: Are methods incorporating ITV free from the Jack Horner and Narcissus effect?

The main result of this WP is an R package (*comsimitv*) for simulating community assembly considering intraspecific trait variation. It is an improved version of the simulation used in Botta-Dukát & Czúcz (2016). The main difference between the previous and new versions is the modular structure of the script that allows users to easily make changes in some procedures. For example, the present version contains one symmetric and two asymmetric competition kernels, but users can define further kernels. The package is available in the CRAN: https://cran.r-project.org/web/packages/comsimitv/index.html

The paper showing the new package and its possible application is under revision in *Ecology and Evolution*. The recent version of the manuscript can be find at the end of this report. Note that publication of this paper was hard because another R package, *ecolottery* (Munoz et al. 2018) has already been developed and reviewers and editors hardly understand why an alternative package is useful.

The simulation is suitable for *in silico* experiments and comparing sampling strategies, but the main aim was to check the reliability of statistical tests used for detecting assembly rules. In this topic, we focused on the T-statistics (Violle et al. 2012) because in this case all details of the analysis were already developed by Taudiere and Violle (2016). They have already checked the reliability of the statistics, but they used trait distributions created according to the null hypothesis. In this case, the type I error rate did not differ considerably from the significance level. However, when more realistic data created by our simulation is analyzed, the type I error rate proved to be higher than the significance level. The reason for the difference is that in our study type I error is defined as the probability of detecting an effect when it was switched off in the simulation (see more details in the submitted manuscript).

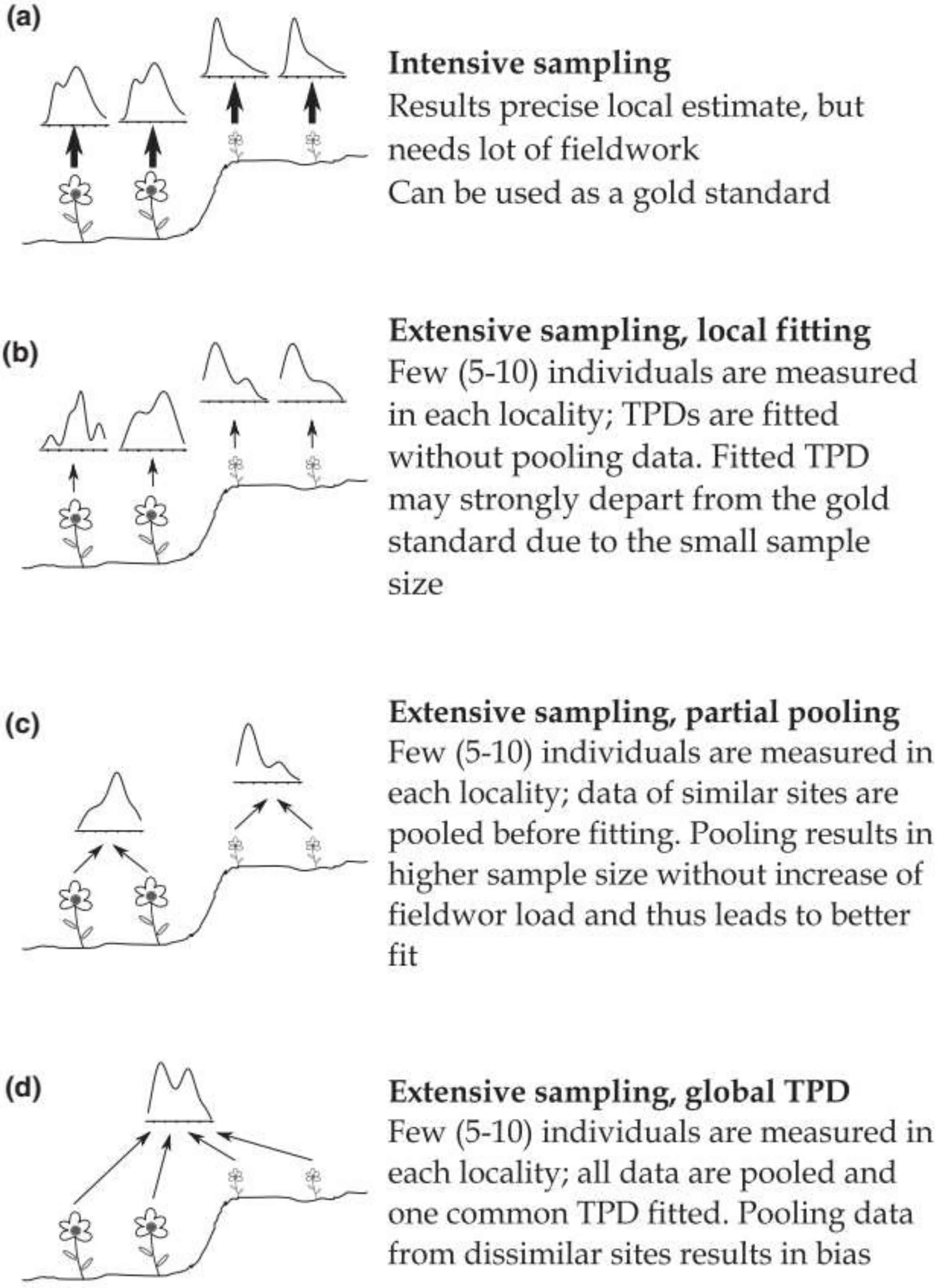
An alternative way of involving ITV in exploring assembly rules would be calculating functional diversity using this information. It needs fitting trait probability distributions (TPD) (Carmona et al. 2016, 2019b) and then calculating overlap between species' TPDs and using this overlap as the distance between species in the calculation of functional diversity, or creating community-level TPD and calculating functional diversity according to formulas proposed by Carmona et al. (2019b). In both approaches, the first step is fitting species-level TPDs. In this step, an important question emerges: what is the optimal pooling of data? Using more data results in more reliable estimates, but mixing data from sites with different TPD leads to over-estimation of within-site and under-estimation of between-site functional diversity (Figure 13). Therefore, we developed an algorithm for the automated finding of optimal pooling based on information criteria and prepared R functions implementing this method. The results of this study were published in *Global Ecology and Biogeography* (https://doi.org/10.1111/geb.13293)

Community Assembly via Trait Selection (CATS) is a promising method for detecting the effect of environmental filtering. In the original form (Shipley et al. 2006) it applies a slow entropy maximization approach, but Warton et al. (Warton et al. 2015) have shown that it is equivalent to fitting Poisson-GLM. This new CATS-regression approach not only a much quicker way of fitting the same model, but it revealed the hidden assumptions (i.e. Poisson-distribution of abundances, the log-linear relationship between trait values and abundances) and allowed relaxing them: e.g. replacing Poisson distribution with other distribution more appropriate for a given type of data, or fitting non-linear regressions. However, Warton et al. (2015) have neglected some important consequences fitting other distributions. CATS allows for estimating the relative role of environmental filtering and landscape-scale processes (Sonnier et al. 2010, Shipley 2014) by calculating adjusted Kullback-Leibler R^2 for different models. However, the formula given by Shipley (2014) is valid only for Poisson distribution. Formulas for other distributions important for modelling abundance data are summarized in our paper published in *PeerJ* (https://doi.org/10.7717/peerj.12763). In this paper, we also illustrated that landscape-level processes can be modelled by using logarithm meta-community level relative abundances as an offset (as suggested by Warton et al.), only if the fitted GLM applies the log link function. In another paper, published in the *Journal of Vegetation Science* (https://doi.org/10.1111/jvs.13154) we have shown that fitting linear models and neglecting interaction between traits leads to underestimation of the role of environmental filtering. Although in these studies mean trait values are used, they establish the use of CATS for data with ITV.

5 WP4: Exploring assembly rules by methods incorporating ITV: a case study

In this WP we have repeated the field sampling that had been done in the Orgovány site in the previous project (see details Lhotsky et al. 2016) supplemented by measuring leaf traits and height in one randomly selected individual of each species in each plot. 90 plots were sampled and 204 species are detected in plots resulting in 1933 species-plot combinations. Leaf traits (leaf size, SLA and LDMC) and vegetative height are measured for these combinations. Trait data are published in a data paper in *Acta Botanica Hungarica* (Gyalus et al. in press) and also publicly available in the PADAPT database.

To explore the role of ITV in the trends of functional diversity, we repeated the analysis of Lhotsky et al. (see details Lhotsky et al. 2016) using this new trait measurement. We applied two approaches called hereafter "mean trait" and "local measurement" approaches. In the mean trait approach, unweighted averages of traits are calculated for each species, and these traits are used in the subsequent analysis, applying exactly the same methods as Lhotsky et al. In the local measurement approach, Rao's quadratic entropy is calculated from the trait values measured in the focal plot. In this approach, we have to modify the randomization to avoid considering more than one trait value of any species in a random



Intensive sampling
Results precise local estimate, but needs lot of fieldwork
Can be used as a gold standard

Extensive sampling, local fitting
Few (5-10) individuals are measured in each locality; TPDs are fitted without pooling data. Fitted TPD may strongly depart from the gold standard due to the small sample size

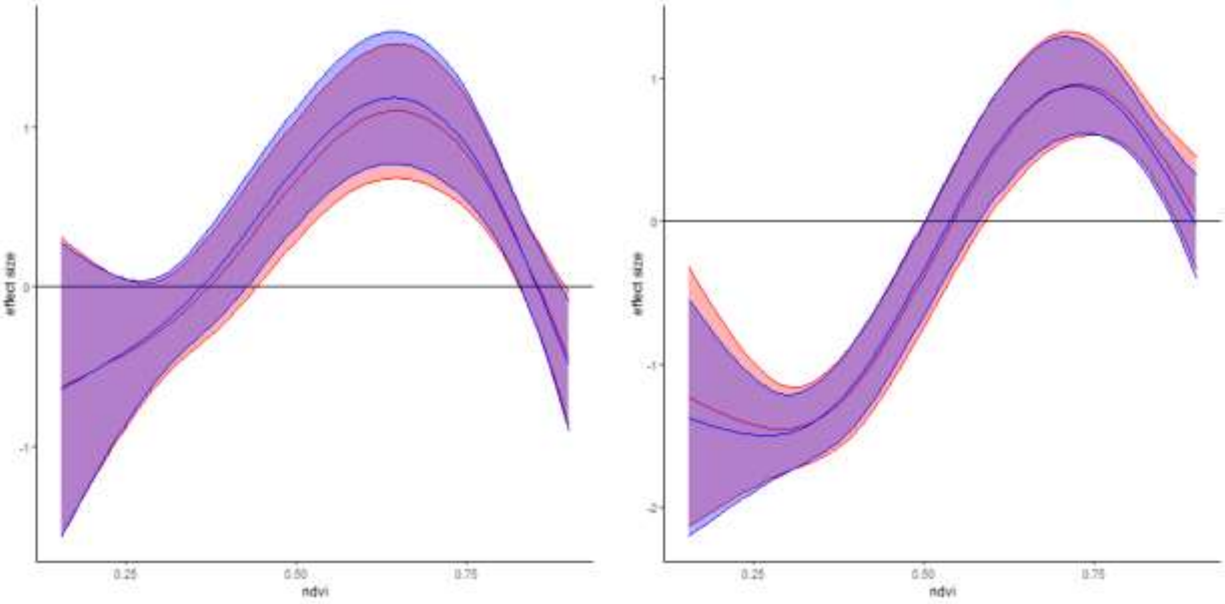
Extensive sampling, partial pooling
Few (5-10) individuals are measured in each locality; data of similar sites are pooled before fitting. Pooling results in higher sample size without increase of fieldwor load and thus leads to better fit

Extensive sampling, global TPD
Few (5-10) individuals are measured in each locality; all data are pooled and one common TPD fitted. Pooling data from dissimilar sites results in bias

Figure 13: Illustration of the different sampling strategies. In intensive sampling (a) the traits of many individuals (illustrated by thick arrows) of each species are measured in all the plots in which the species is present. This large sample size allows reliable trait probability densities (TPDs) to be fitted in each site. However, this strategy implies a great effort that is not feasible in most situations. We consider three alternative strategies where only a few individuals are measured (illustrated by thin arrows): local fitting (b), partial pooling (c) and global TPD (d). Note that sites are not evenly spaced along the environmental gradient (illustrated by the step-like slope) and only data from sites similar in the environment are pooled in the partial pooling

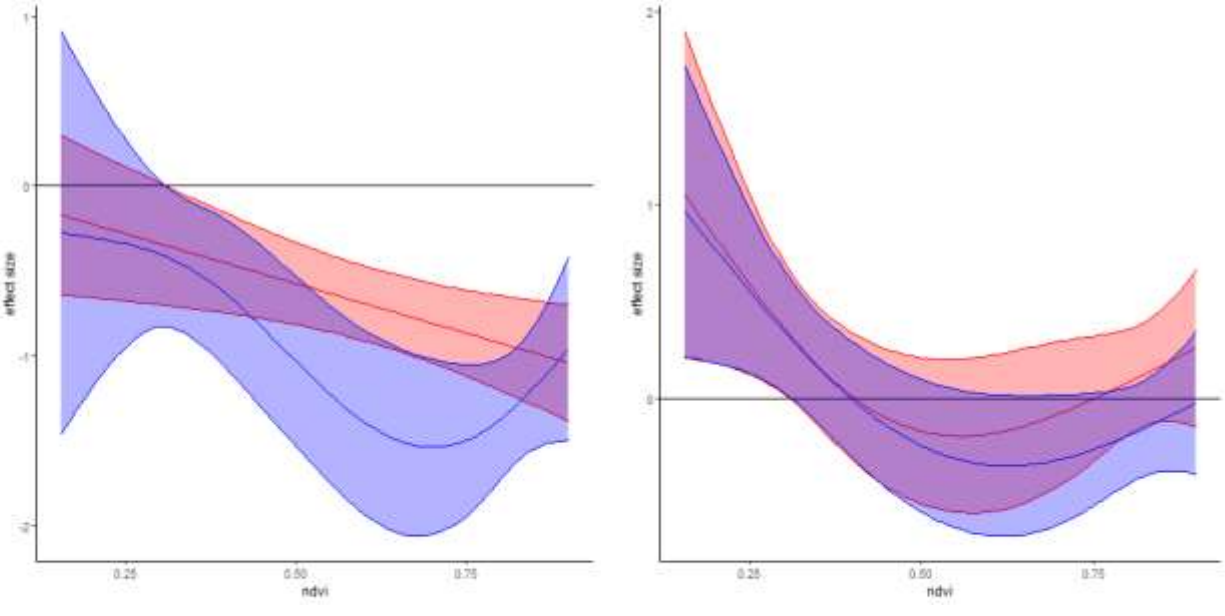
community. Therefore, we applied a two-step procedure. First, for each non-zero abundance, one species is chosen from the list of detected species (an estimate of the regional pool). Each species can be chosen only once. In the second step, one of the measured trait values within each chosen species was randomly chosen. Otherwise, the method was the same as in the other approach.

The differences between approaches in fitted trends of effect sizes were marginal (Figure 14), suggesting that in this case incorporating ITV does not give additional insight into the assembly processes.



a) Height

b) leaf size



c) SLA

d) LDMC

Figure 14. Fitted trends along productivity gradients and their 95% confidence intervals of effect sizes calculated from species means (red) and local trait measurements (blue).

Figure 15: Effect size of SLA along the productivity (NDVI) gradient calculated from SLA values used by Lhotsky et al. (old data set), and used in WP2 (new dataset). The red line is the trend fitted by GAM.

Figure 16: Effect size of SLA along the productivity (NDVI) gradient calculated from SLA values used in WP2 (new dataset) or using the same traits except for *Poa bulbosa* whose trait set to the value used in Lhotsky et al. Black line is the trend of actual data, while the red line is the trend fitted to effect sizes calculated from “old trait dataset”. Filled dots show plots where *Poa bulbosa* occurs.

Variation of trait values along environmental gradients is only one possible source of ITV. Due to ITV, species means obtained from different databases or calculated from different measurements may differ. In the case of SLA, the trend calculated for Orgovány data in WP2 considerably differ from the trend shown by Lhotsky et al. In both cases the same abundances and similar methods, but different trait datasets were used. To exclude the possible effect of small differences in the methods, we repeated the analysis with the new trait dataset, using the same methods as Lhotsky et al. The difference between trends did not disappear (Figure 15).

For exploring which species are responsible for the differences, we created “hybrid” datasets changing the SLA of each species in the new trait dataset with the value used by Lhotsky et al. Then we calculated the departure of the trend of effect size when the new and the hybrid datasets were used from the trend of effect size when the old dataset was used.

Based on this analysis, changes in the SLA of *Poa bulbosa* and *Secale sylvestre* proved to be the two most important reasons of changing trends (Figures 16-17). If trait values of these species change to the value used in Lhotsky et al, the difference in trends becomes small (Figure 18). Interestingly, both values of these species were measured by our research group in the same region, but in different years. SLA of these two annual grasses considerably varies among the years, and in some plots, they can reach high cover. These two facts together explain the strong effect of their changing SLA on the trends of effect size.

6 WP5: Partitioning within population ITV into stochastic and deterministic parts

The analysis originally planned in this section assumes that CWM is a good estimate of optimal trait value in the given environment. During the work with CATS models, we have recognized that this assumption is rarely satisfied. To demonstrate this fact, let us imagine that the trait-abundance relationship can be described by a Gaussian curve with an optimum at μ, and tolerance width of σ. If the range of trait values was not finite, CWM would be modelled by a mean of random values from a normal distribution with parameters μ and σ. In this case, the expected value of CWM would be μ. But if trait values ranges from *a* to *b* where at least one border is finite, the normal distribution has to be replaced by a truncated normal distribution and thus expected value of CWM is

$$E(CWM) = \mu + \frac{\varphi\left(\frac{a-\mu}{\sigma}\right) - \varphi\left(\frac{b-\mu}{\sigma}\right)}{\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)}\sigma$$

where φ and Φ are the density and distribution function of standard normal distribution.

Because of the possible bias due to the unsatisfied assumption, we cancelled the planned work. Instead, we fitted TPDs with our new method that automatically chooses the optimal level of pooling. For most species, the optimal pooling resulted in several TPDs indicating the presence of local adaptation. We plan to further studies using these fitted TPDs to separate adaptation to the local environment (groups within species are separated along the productivity gradient) from the other sources.

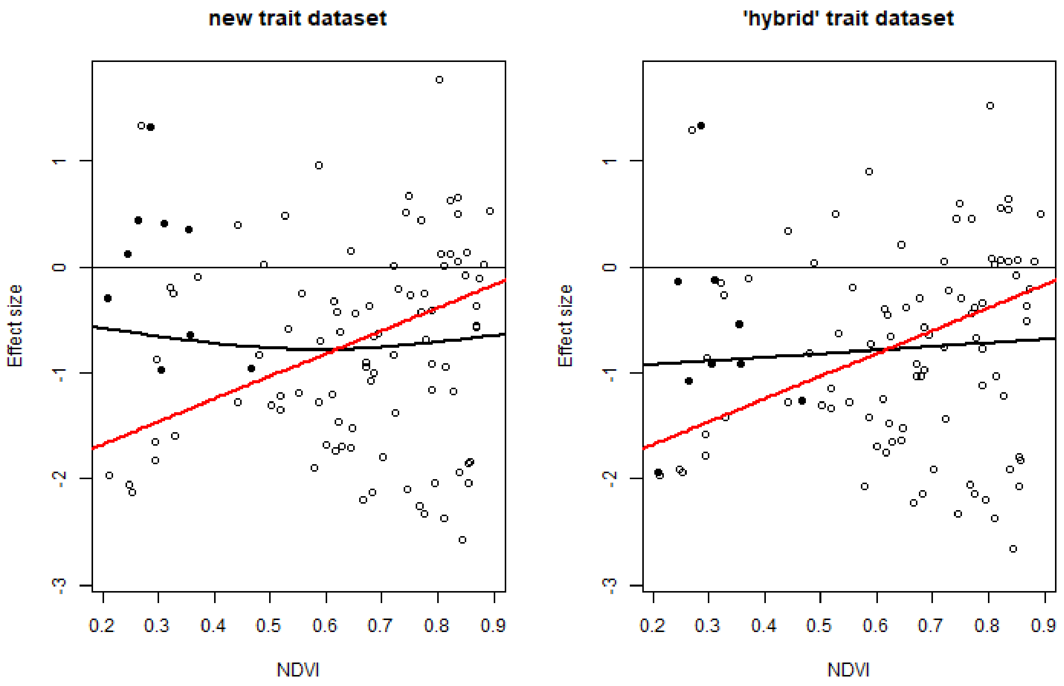


Figure 17: Effect size of SLA along the productivity (NDVI) gradient calculated from SLA values used in WP2 (new dataset) or using the same traits except for *Secale sylvestre* whose trait set to the value used in Lhotsky et al. Black line is the trend of actual data, while the red line is the trend fitted to effect sizes calculated from “old trait dataset”. Filled dots show plots where *Secale* occurs.

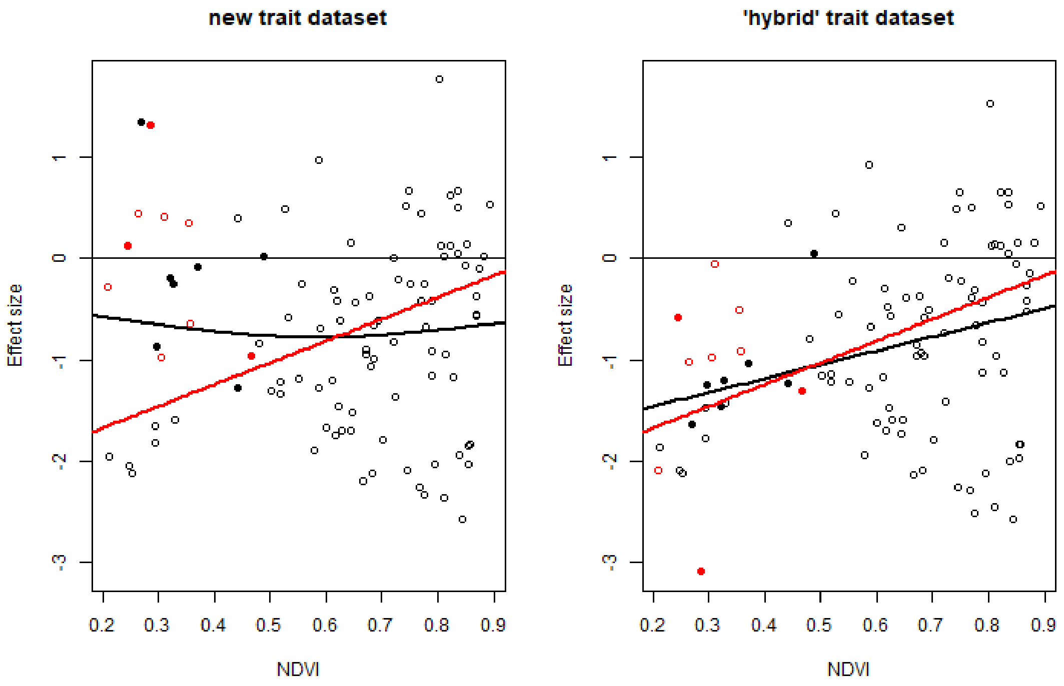


Figure 17: Effect size of SLA along the productivity (NDVI) gradient calculated from SLA values used in WP2 (new dataset) or using the same traits except for *Poa bulbosa* and *Secale sylvestre* whose trait set to the value used in Lhotsky et al. Black line is the trend of actual data, while the red line is the trend fitted to effect sizes calculated from “old trait dataset”. Filled dots show plots where *Poa bulbosa*, while red dots where *Secale* occurs.

7 Unplanned works

7.1 ITV in competition experiments

Carmona and colleagues (Carmona et al. 2019a) had published the results of a well-designed experiment examining the effect of functional traits on the outcome of the competition. Their main conclusions were that (1) trait hierarchies better predict the strength of competition than trait differences, and (2) the consideration of intraspecific trait variation improves our ability to predict the strength of competition. They applied multi-model inference by averaging regression coefficients. This approach is strongly criticized because it averages regression coefficients for which the meaning is conditional on the other predictors included in the model. We re-analysed their data using alternative methods and our results only partly support the original conclusions that trait hierarchy proved to be important in height, but not in the other three traits. The paper was published *Journal of Ecology* (https://doi.org/10.1111/1365-2745.13666). Beyond the re-analysis, the main message of our paper is that traits measured on the level of individuals involved in the experiments are not only drivers but also consequences of competitive interactions (Figure 18).

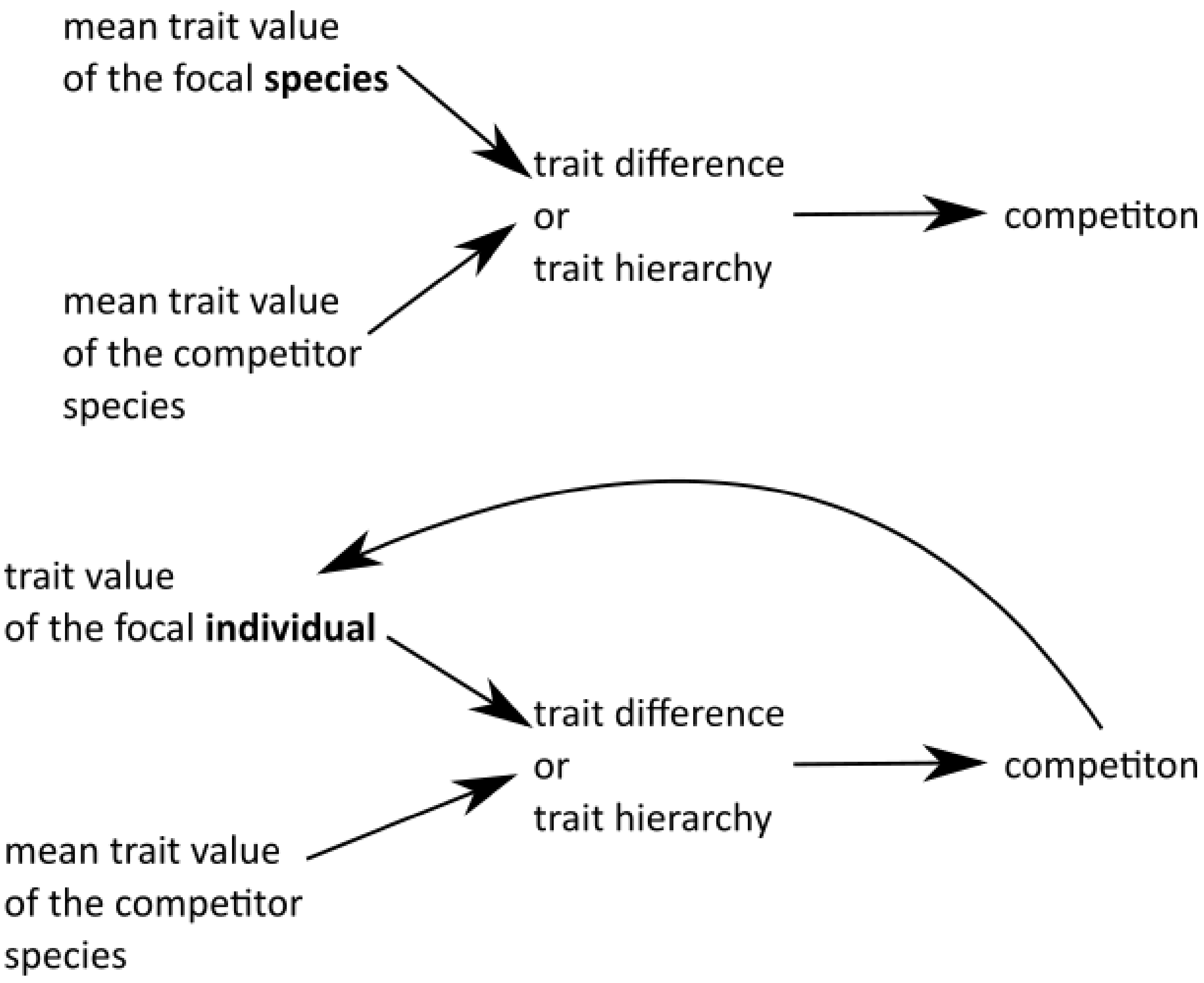


Figure 18: Traits measured in focal individuals may be both drivers and consequences of competition, which makes the interpretation of the relationship more difficult.

7.2 CV is not suitable for ratio-type traits

Comparing within-species variations of traits can be used in testing ecological theories. In these comparisons, it is useful to remove the effect of the difference in mean trait values, therefore measures of relative variation, most often the coefficient of variation (CV), are used. The studied traits are often calculated as the ratio of the size or mass of two organs: e.g. SLA is the ratio of leaf size and leaf mass. Often the inverse of these ratios is also meaningful; for example, the inverse of SLA is often referred to as LMA (leaf mass per area). Relative variation of a trait and its inverse should not considerably differ. However, it is illustrated that using the coefficient of variation may result in differences that could influence the interpretation, especially if there are outlier trait values. Using data collected in WP4, we compared the alternative ways of measuring the relative variation of traits. The alternative way for estimating CV from the standard deviation of log-transformed data assuming log-normal distribution and Kirkwood's geometric coefficient of variation free from this problem, but they proved to be sensitive to outlier values. The quartile coefficient of variation performed best in the tests: it gives the same value for a trait and its inverse and it is not sensitive to outliers. The manuscript has been submitted to Scientific Reports and its pre-print is available in http://dx.doi.org/10.1101/2022.10.13.512014.

7.3 Among-species variation of correlation between leaf traits

The leaf economy spectrum is a well-known pattern of leaf traits among species on the global scale: leaf traits are changes in a correlated way. However, at final spatial scales (e.g. within communities) or within species these correlations often become weaker or even non-significant. However, if there is a physical relationship in the background of correlation, we expect that correlation remains relatively stable across scales.

Specific leaf area (SLA) and leaf dry matter content (LDMC) are two often measured leaf traits. Their negative correlation is a part of correlations forming the leaf economy spectrum. The background of this correlation is that both traits are related to the cell wall : protoplasm ratio. Due to this physical connection, we expected that this correlation is stable across scales.

Trait measurements in WP4 resulted in 2050 SLA and LDMC measurement pairs (species-plot combinations). Correlation and type II regression between log-transformed traits are calculated over several scales: for raw measurement in all species-plot combinations, within plots and species, and for aggregated values.

The slope of the relationship only slightly changed when the relationship was calculated within plots instead of the whole gradient (Figure 19) or when aggregated values are analyzed instead of raw measurements (Figure 20, 21). The only exception is the within-species relationship (Figure 22). For almost one-third of species, the correlation becomes non-significant when calculated within species (Figure 23). Grasses were over-represented in this group. The slope significantly varies among species even after excluding species with non-significant correlation. A main, but not the only source of this variation is the difference between herbs and grasses (Figure 24). This exception calls attention to the limits of extrapolating trait correlations across scales.

7.4 Weighted median of distances is more robust against outliers than Rao's quadratic entropy

During data exploration, we noticed that some species has extreme (outlier) trait value, and in plots where these species occur the functional diversity is high, and effect size indicates strong divergence. It may be an artefact: since the outlier trait value is far from the rest of the traits in the species pool when this species presents, the mean distance among traits of co-occurring species will be high, even if the community is randomly assembled. This artefact can be excluded, when repeating the analysis

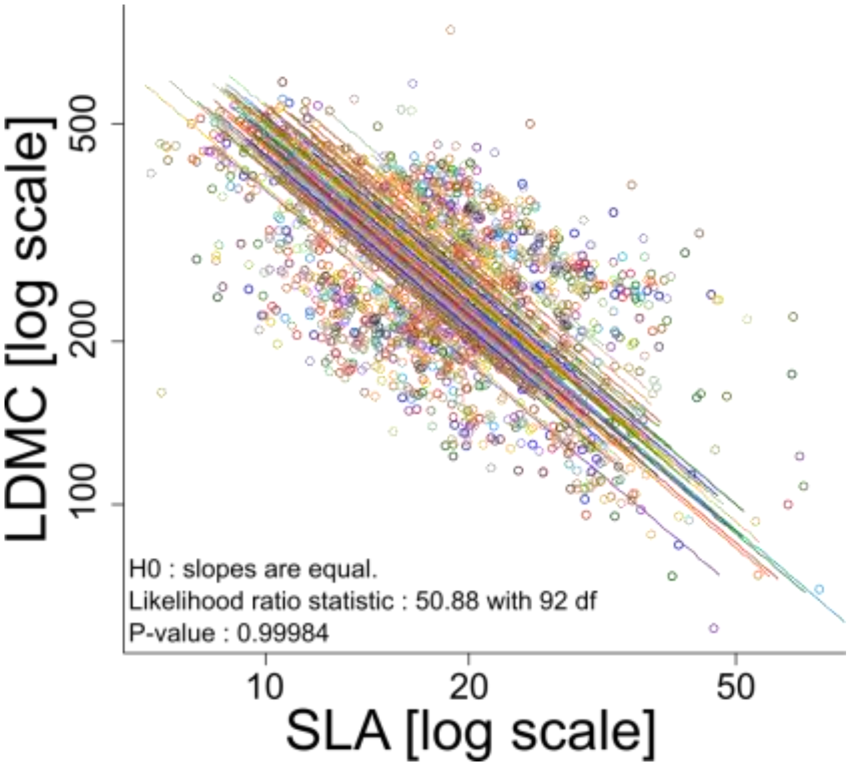


Figure 19: Within-plot Type II regression between SLA and LDMC.

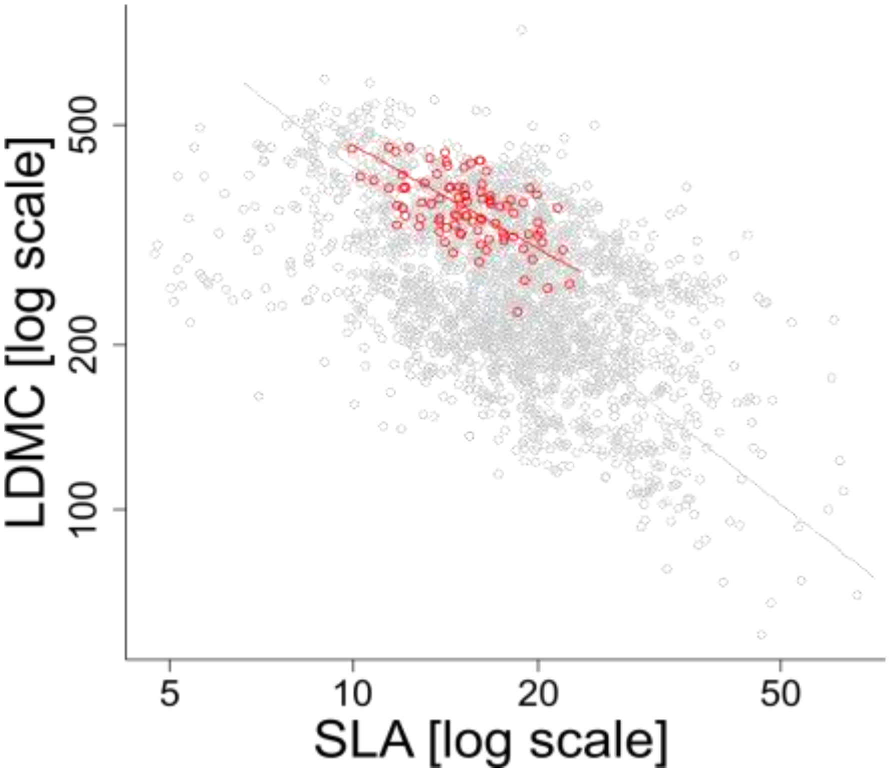


Figure 20: Type II regression between SLA and LDMC. Gray dots = original measurements; red dots = community weighted means

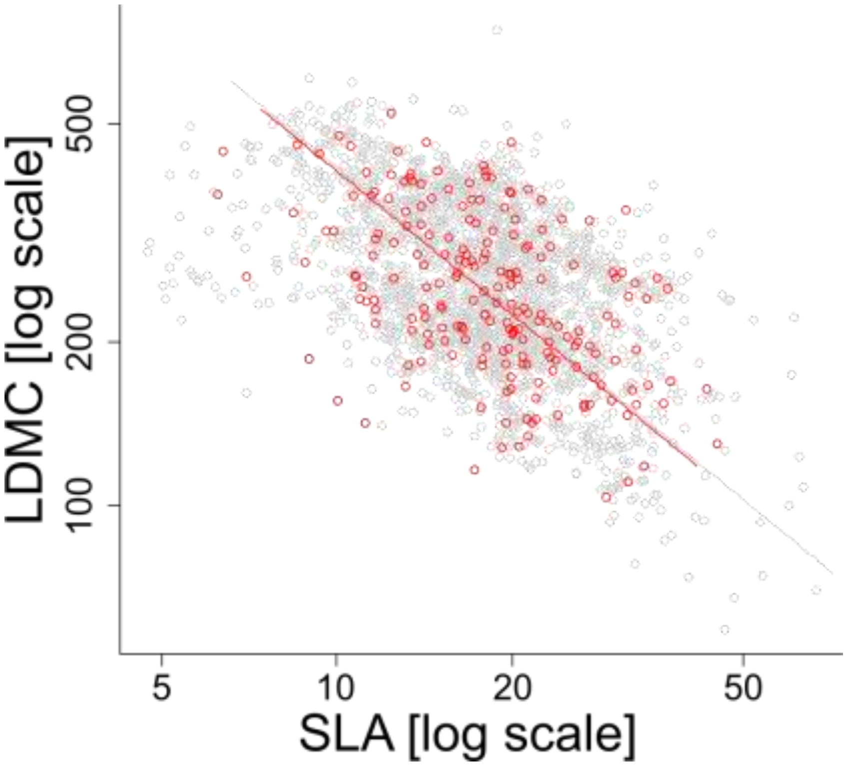


Figure 21: Type II regression between SLA and LDMC. Gray dots = original measurements; red dots = species' trait centroids

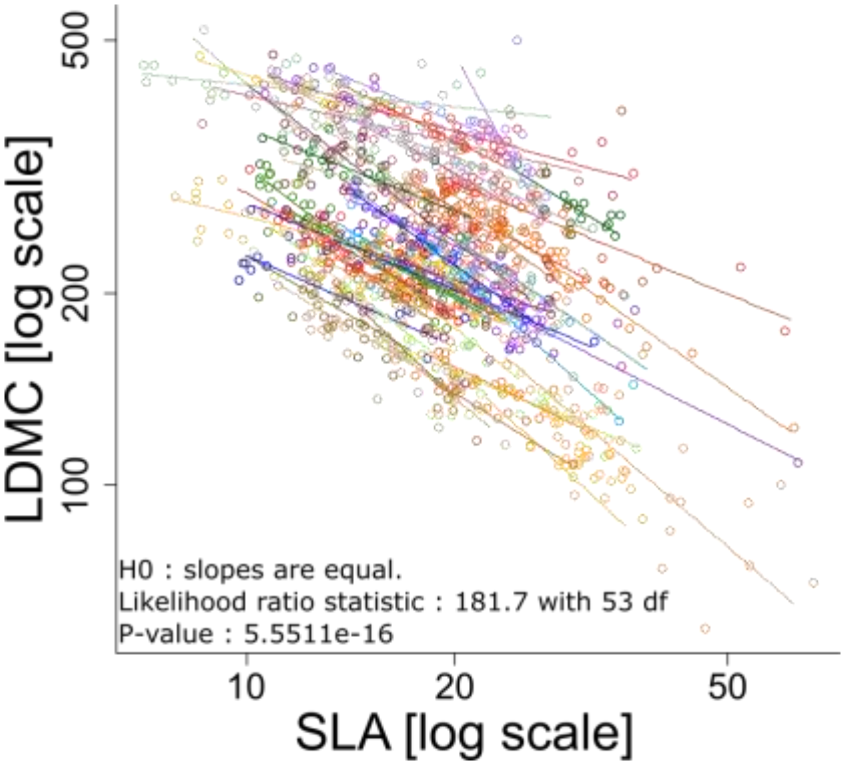


Figure 22: Within-species Type II regression between SLA and LDMC. Species with non-significant correlation are excluded

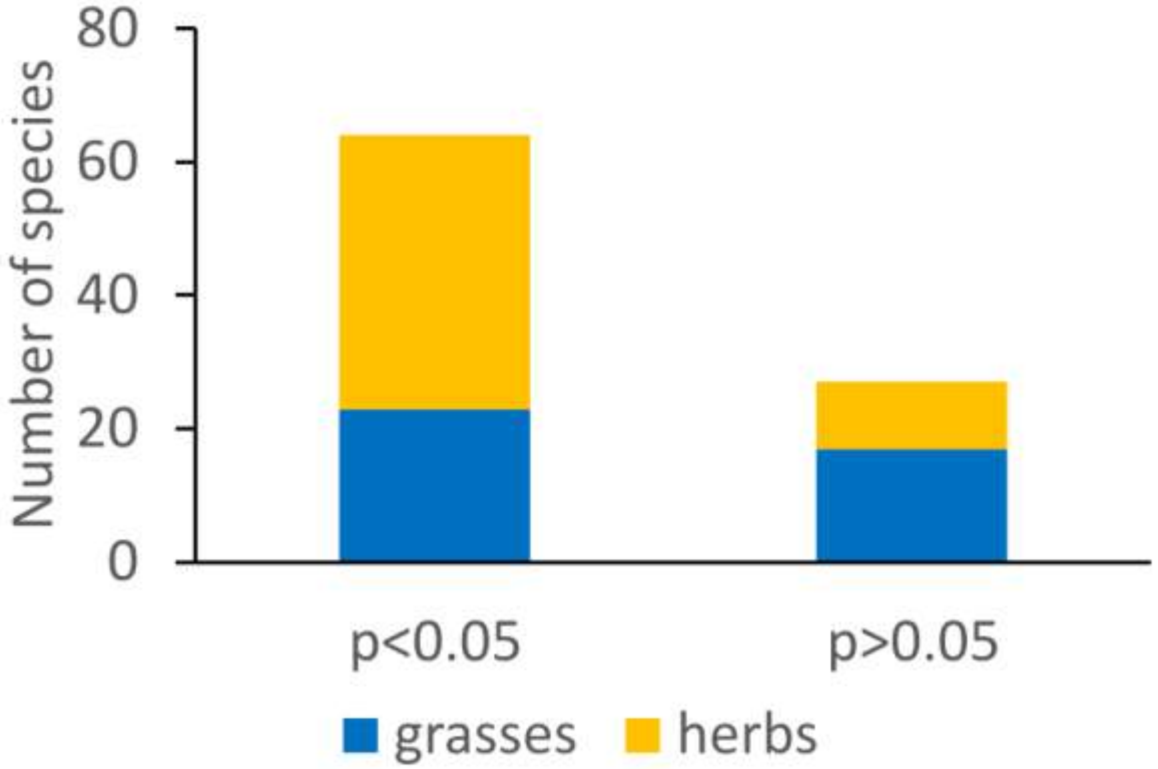


Figure 23: Proportion of significant (p<0.05) and non-significant (p>0,05) within-species correlations between SLA and LDMC.

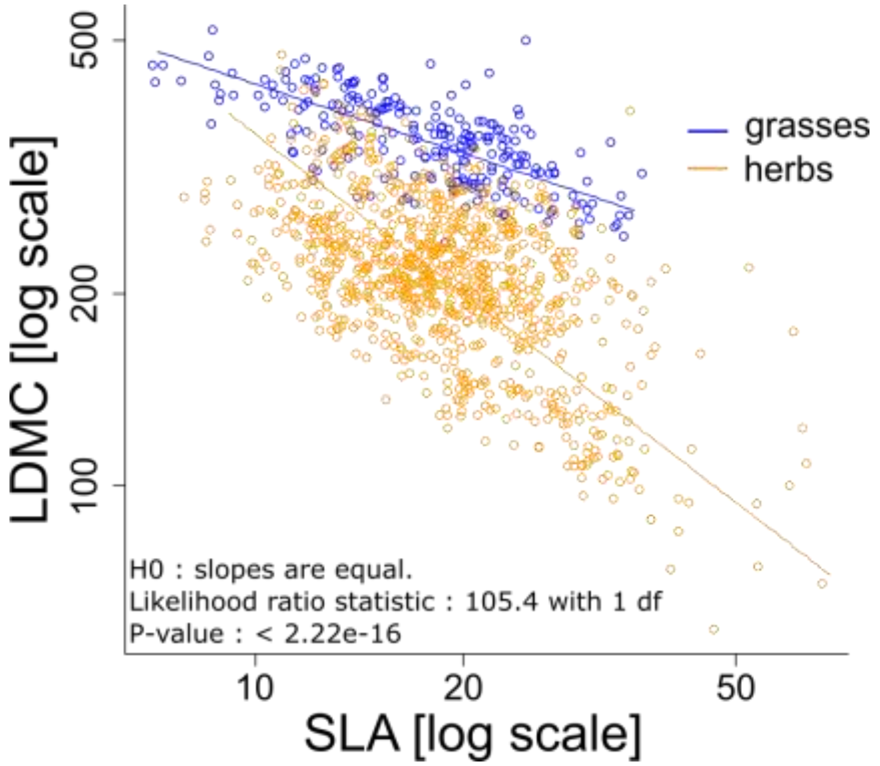
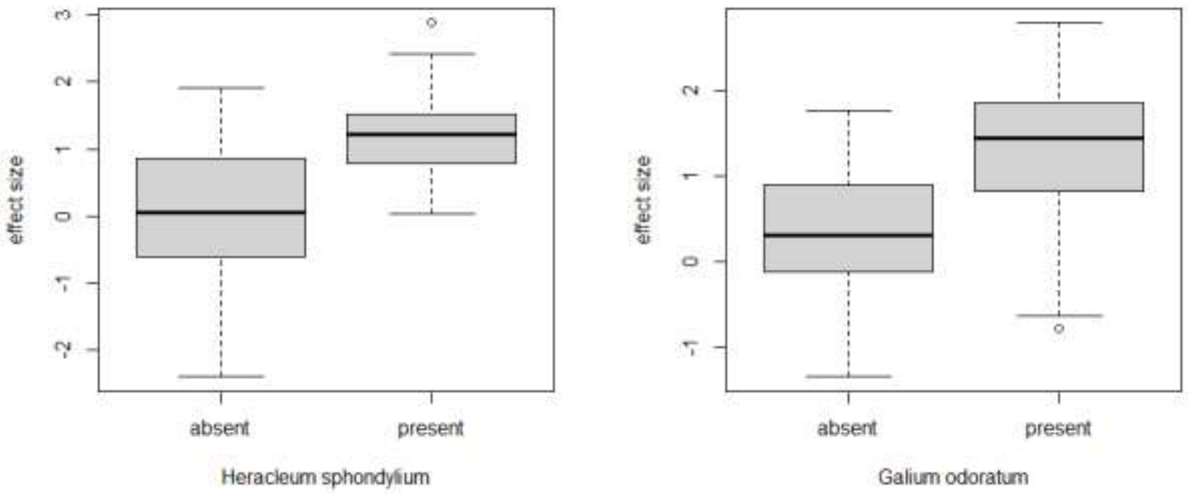


Figure 24: Type II regression between SLA and LDMC for the two main functional types, separately. Species with non-significant correlation are excluded

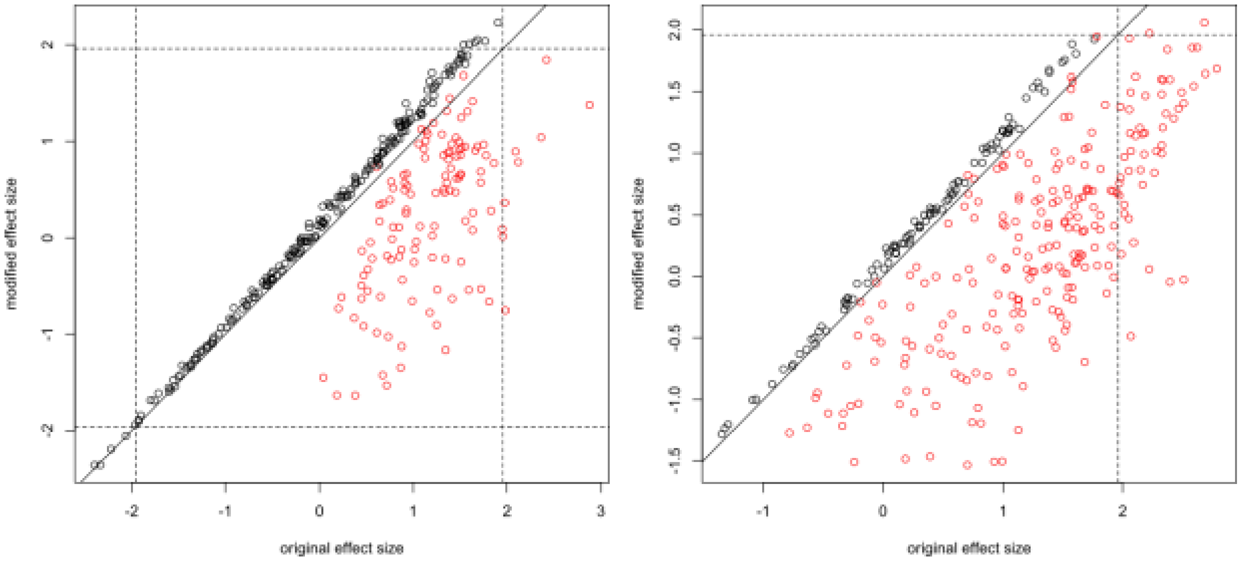
without considering the outlier species gives the same results. Methods may differ in sensitivity to the presence of outliers. We hypothesized that log-transformation and replacing the weighted mean with the weighted median could decrease the sensitivity.

To explore the role of outliers, three leaf traits (leaf size, SLA, LDMC) of the Várhegy dataset were used. This dataset was also used in Botta-Dukát et al. (2022), see this paper for details of sampling. For each trait, the species with the highest value was selected for the subsequent analysis: *Heracleum sphondylium* for leaf size, *Inula ensifolia* for LDMC, and *Galium odoratum* for SLA. The effect sizes are significantly higher (Wilcoxon rank test) in plots where these species are present (Figure 25).



a) c)

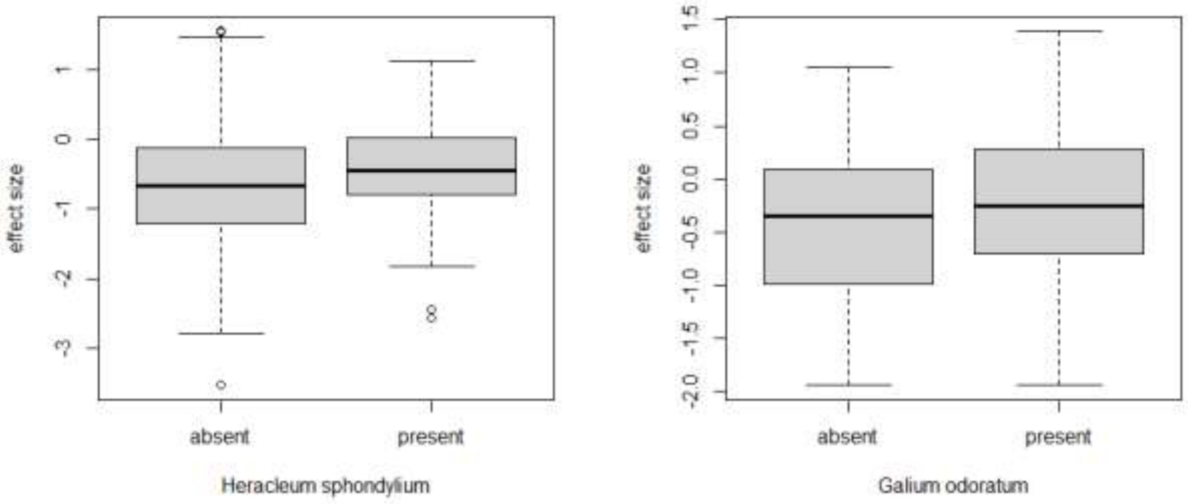
Figure 25: Effect sizes calculated from leaf size (a), LDMC (b) and SLA (c) in plots where species with outlier trait values are present/absent.



a) leaf size b) SLA

Figure 26: Relationship between effect sizes calculated from the whole dataset (x-axis) and with excluding the outlier species (y-axis). Plots where the outlier species presents are indicated by red dots. Dashed lines show the critical values at the 5% significance level.

Excluding the outlier species only slightly modified the effect size of plots where the species was absent. However, excluding the outlier species strongly modified the results for plots where these species are present, sometimes changing patterns from significant to non-significant or from divergent to convergent (Figure 26).

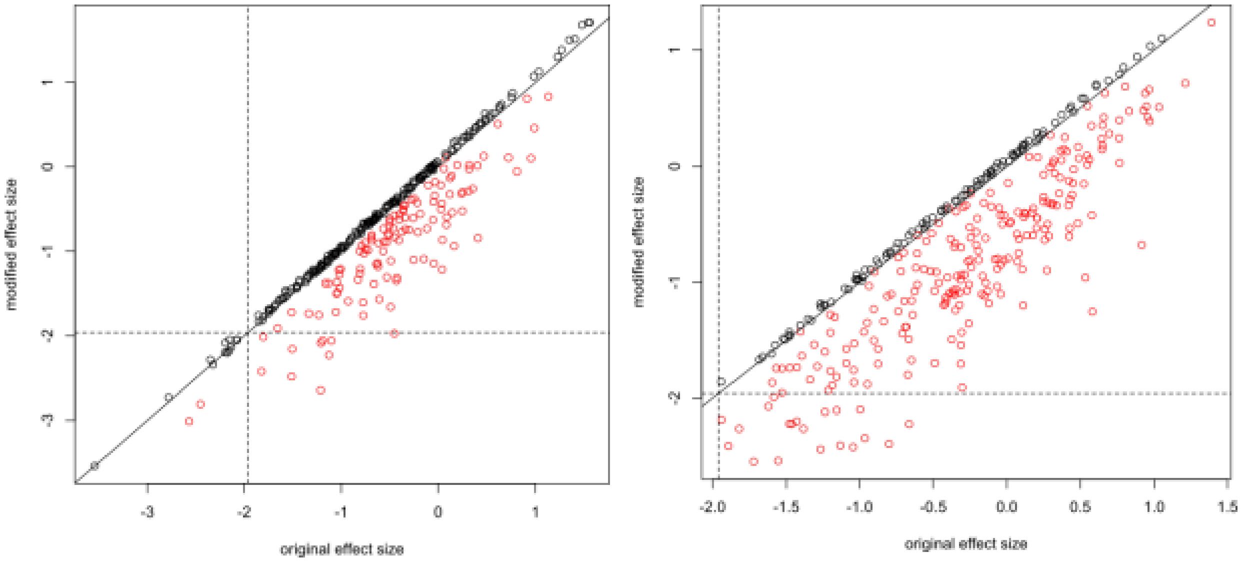


a) b)

Figure 27: The same as Figure 25, but calculated from log-transformed traits.

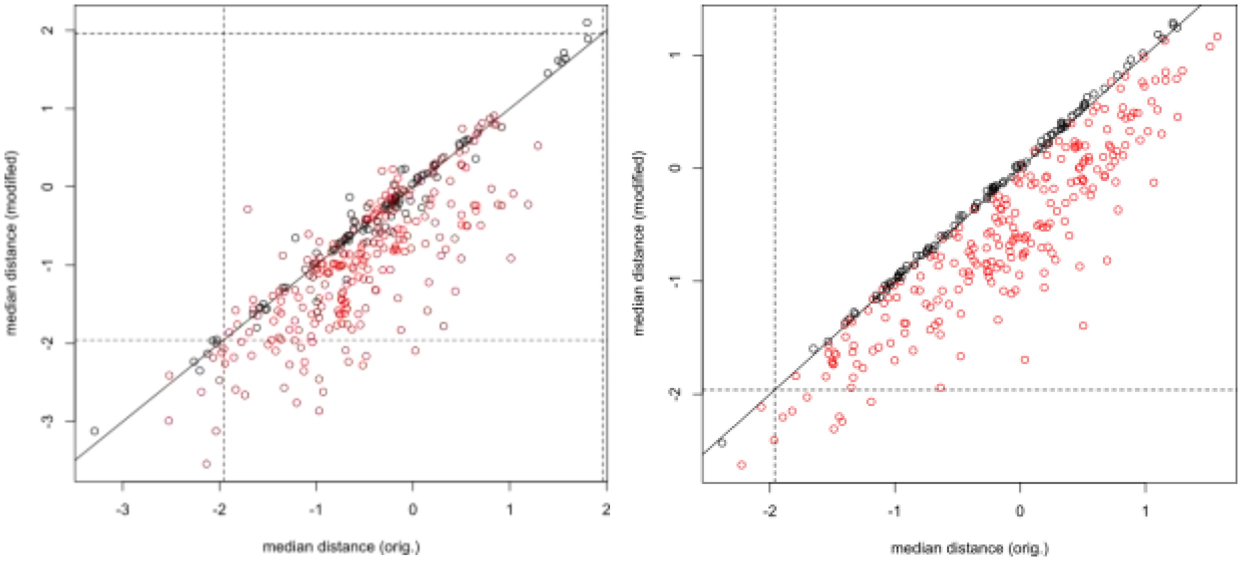
Log-transformation decreased the difference in effect size between plots where the outlier species was present/absent (Figure 27); for leaf size, the difference remained significant, while for SLA it was only marginally significant (p=0.06). After log transformation, the effect of excluding the outlier species is smaller, but still not negligible (Figure 28).

Replacing Rao's quadratic entropy (i.e. the weighted mean of trait differences) with the weighted median of trait differences also decreased but did not completely remove the effect of outliers (not shown). Combining log transformation and weighted median proved to be the satisfying solution. Using this approach, the effect of outliers was negligible (Figure 29).



a) leaf size b) SLA

Figure 28: The same as Figure 26, but calculated from log-transformed traits.



a) leaf size b) SLA

Figure 28: The same as Figure 26, but calculated from log-transformed traits using weighted median instead of weighted mean (i.e. Rao’s quadratic entropy).

8 Literature

Botta-Dukát, Z. 2012. Co-occurrence-based measure of species’ habitat specialization: robust, unbiased estimation in saturated communities. - J. Veg. Sci. 23: 201–207.

Botta-Dukát, Z. 2018. Cautionary note on calculating standardized effect size (SES) in randomization test. - Community Ecol. 19: 77–83.

Botta-Dukát, Z. and Czúcz, B. 2016. Testing the ability of functional diversity indices to detect trait convergence and divergence using individual-based simulation. - Methods Ecol. Evol. 7: 114–126.

Botta-Dukát, Z. et al. 2022. Oversimplified models underestimate the role of local environmental filtering. - J. Veg. Sci. in press.

Carmona, C. P. et al. 2016. Traits Without Borders: Integrating Functional Diversity Across Scales. - Trends Ecol. Evol. 31: 382–394.

Carmona, C. P. et al. 2019a. Trait hierarchies and intraspecific variability drive competitive interactions in Mediterranean annual plants (N Barber, Ed.). - J. Ecol. 107: 2078–2089.

Carmona, C. P. et al. 2019b. Trait probability density (TPD): measuring functional diversity across scales based on TPD with R. - Ecology 100: e02876.

Götzenberger, L. et al. 2016. Which randomizations detect convergence and divergence in trait-based community assembly? A test of commonly used null models. - J. Veg. Sci. 27: 1275–1287.

Hothorn, T. et al. 2006. Unbiased Recursive Partitioning: A Conditional Inference Framework. - J. Comput. Graph. Stat. 15: 651–674.

Lhotsky, B. et al. 2016. Changes in assembly rules along a stress gradient from open dry grasslands to wetlands. - J. Ecol. 104: 507–517.

Messier, J. et al. 2017. Trait variation and integration across scales: is the leaf economic spectrum present at local scales? - Ecography 40: 685–697.

Munoz, F. et al. 2018. ecolottery: simulating and assessing community assembly with environmental filtering and neutral dynamics in R. - Methods Ecol. Evol. 9: 693–703.

Shipley, B. 2014. Measuring and interpreting trait-based selection versus meta-community effects during local community assembly. - J. Veg. Sci. 25: 55–65.

Shipley, B. et al. 2006. From Plant Traits to Plant Communities: A Statistical Mechanistic Approach to Biodiversity. - Science 314: 812–814.

Siefert, A. et al. 2015. A global meta-analysis of the relative extent of intraspecific trait variation in plant communities. - Ecol. Lett. 18: 1406–1419.

Sonnier, G. et al. 2010. Plant traits, species pools and the prediction of relative abundance in plant communities: a maximum entropy approach. - J. Veg. Sci. 21: 318–331.

Violle, C. et al. 2012. The return of the variance: intraspecific variability in community ecology. - Trends Ecol. Evol. 27: 244–252.

Warton, D. I. et al. 2015. CATS regression – a model-based approach to studying trait-based community assembly. - Methods Ecol. Evol. 6: 389–398.

9 Appendix: Manuscript submitted to Ecology and Evolution

comsimitv: an R package for simulating trait-based community assembly with intraspecific trait variation

Zoltán Botta-Dukát
Centre for Ecological Research, Alkotmány u. 2-4., Vácrátót, 2163, Hungary
botta-dukat.zoltan@ecolres.hu

Abstract

Process-based models are useful tools for ecological research to check the statistical power and validity of methods and for doing experiments in silico. Trait-based analyses considering intraspecific trait variation (ITV) are becoming more and more popular, so process-based simulations that consider ITV have to be developed to check the validity of applied methods.

I present an R package that provides a flexible framework for such simulations. The flexibility of the package comes from the implementation of each simulation step in a separate function that could be replaced by user-written functions. Functions for simulating the sampling of trait and abundance data are also included.

Simple in silico experiments with the new package illustrate that the incorporation of ITV may result in unexpected patterns, such as a disappearing relationship between the dispersal rate and gamma diversity at high ITV or overestimation of trait variation by global sampling when a trait is filtered by the local environment. High type I error rate of T-statistics highlight the hidden assumptions of these tests.

Keywords: assembly rules, intraspecific variation, process-based model, simulation, trait,

1 | INTRODUCTION

In the last decade, functional traits have become an important part of community ecologists' toolbox (Götzenberger et al., 2012; McGill et al., 2006). Case studies apply various null models (Götzenberger et al., 2012; Hardy, 2008) and test statistics, which are most often functional diversity indices (Pavoine & Bonsall, 2011) or descriptors (e.g. mean, variance, skewness, kurtosis) of trait distribution (e.g. Gross et al., 2021; Kraft et al., 2008). In a reliable statistical test, there is a low probability of both type I error (detecting an effect when it does not exist) and type II error (not detecting an existing effect) (Wilson, 1995). The performance of statistical models can be checked by analyzing virtual data created by a process-based simulation (Connolly et al., 2017). Previous attempts in trait-based ecology have shown the importance of such checks, which many suggested methods have failed (e.g. Botta-Dukát & Czúcz, 2016; Götzenberger et al., 2016).

The intraspecific variation of traits (ITV) is usually neglected both in the collection and analysis of trait data. Its reason is that measuring trait means needs much less effort than representing ITV (means for species often available in databases, like TRY; Kattge et al., 2020). However, neglecting this form of variation may result in biased conclusions (Jung et al., 2010; Violle et al., 2012). For example, Bricca et al. (2022) have found convergence in interspecific, but divergence in intraspecific trait variations at the lower part of an elevation gradient, while the pattern of inter- and intraspecific trait variation was not so strongly different at higher elevations. It illustrates that the bias caused by neglecting ITV is context-dependent (Catford et al., 2022).. Process-based simulations could help to explore situations when neglecting ITV causes large bias.

Incorporating ITV into analyses requires new methods (e.g. Carmona et al., 2016, 2019; Enquist et al., 2015; Laughlin et al., 2012; Mammola & Cardoso, 2020; Violle et al., 2012), which have untested behavior. Some of the new methods (e.g. Laughlin et al., 2012; Violle et al., 2012) can be applied only if the information on ITV is available. To check the type I error rate and statistical power of these new methods, also a process-based simulation model that considers ITV is needed.

This paper presents the comsimitv R package that provides a flexible, process-based simulation framework that could simulate an assembly of plant communities where functional traits may vary within species. There are two R packages for a similar purpose: ecolottery (Munoz et al., 2018) on CRAN and VirtualCom (Münkemüller et al., 2015) on R-Forge, but only the former can simulate intraspecific trait variation. However, there are conceptual and technical differences between ecolottery and comsimitv packages discussed in a separate chapter after the overview of the simulation framework of *comsimitv*.

2 | OVERVIEW OF THE SIMULATION FRAMEWORK

The simulation framework is a meta-community model where individuals interact only within local communities, but produced propagules can migrate among localities. Each individual is characterized by the values of numerical traits. In the default version, three traits are modeled: trait A is related to habitat matching, trait B regulates resource acquisition, while trait C is neutral, however, it is possible to define additional traits. The three default traits represent the three hypothetical roles of traits in population dynamics: (1) a trait may influence survival and/or reproduction in a given environment; (2) a trait may be related to resource competition; and (3) a trait may be neutral, i.e. its value has not any effect on the dynamics. Of course, using these three traits is only the simplest scenario: both habitat matching and resource acquisition may be related to several traits and the same trait may influence both of them. The simulation framework allows to model these more complicated scenarios; it allows for defining correlation among traits and applying user-defined modules (see below).

Individuals compete for space and resources, and competition for space is strict: for each local community, the total number of individuals is limited, and a new individual can enter only after another resident has died. The strength of competition for resources between all pairs of individuals in the local community is calculated from their traits related to resource acquisition by a competition kernel function (Appendix 2). The strength of the competition summed up for each focal individual determines the expected value of produced propagules, while the probability of growing up from seed to adult is related to the adaptedness determined by matching trait A to the local environments. At the level of individuals, the model is not spatiality explicit, so the spatial position of the individuals within the local community does not influence competition, but the position of the localities may be explicit and may influence propagule exchange, however spatially explicit propagule exchange is not implemented in the recent version.

The simulation consists of a community initialization followed by an iterative simulation of a "disturbance–regeneration" cycle. The following steps are involved:

Step 1: Initial communities are created

Step 2: In each local community, a single randomly selected individual dies. It models a situation where mortality is caused by fine-scale disturbance.

Step 3: Surviving individuals may produce seeds. The number of seeds produced depends on the strength of competition for resources (i.e. similarity to other individuals in trait B). Competition between individuals can be symmetric or asymmetric and it depends on their traits related to resource use, thus the same function models intra- and inter-specific competition. ITV is created in this step: trait values of the offspring may differ from their mothers' (see details in Appendix 1).

Step 4: Each produced seed either remains in the local community where it is produced or spreads to another one.

Step 5: In this last step, a lottery competition takes place among seedlings. The probability of winning is calculated from trait A of the seedling and the local environment. The lottery competition is a stochastic process, so the best-adapted individual may not necessarily win. In each turn, only one seedling can survive, and all other propagules are lost. The lucky champion is supposed to grow up until the next iteration and become a fully functional individual capable of reproduction in full competition with all other individuals of the local community.

Iteration cycles: After completing step 1 once, steps 2-5 are repeated many times for each local community so that the final composition becomes independent from the colonization process.

Since changes are slower in larger communities, the number of cycles was set to the simulation length parameter (*sim.length*) multiplied by the number of individuals in the local community (*J*).

2.1 | Simulation modules and available functions

The main module of the *comsimitv* R package is the *comm.simul* function, which calls the other modules. Its parameters are names of the functions for other modules for generating initial communities, calculating survival probability, the competition kernel, the number of produced propagules, determining trait values of seeds, and the spread of propagules, as well as the general settings for the whole simulation. Beginners could make simulations by calling only this function and using built-in modules, while advanced users can write their functions for each module. The following demonstrates the input and output requirements for each module and the available built-in functions.

The module for generating initial communities has to generate species identity and trait values for *J* individuals in each local community. The only recently available function is the *Gener.species.pool*, which first creates a regional species pool consisting of S species. Each species is characterized by probability distributions of numeric traits (A, B, and C in the default setting). Any distribution for which the density, quantile, and random-number-generating function are available in R can be chosen. In this step, a single value of each trait is generated for each species (i.e. there is no ITV). Traits may be independent, or their correlation can be given in form of a covariance matrix (see details in Appendix 1).

After generating the initial regional pool of species, the local communities are filled with species from this species pool until the predefined total number of individuals is reached. Propagules are assumed to arrive from the regional pool at the same rate for all species, while their survival is considered to depend on the suitability of the local conditions for the species.

The competition kernel module calculates a matrix of pairwise competition coefficients between individuals in a local community. Currently, two competition kernel functions are available: *Gaussian.competition.kernel* and *asymmetric.competition.kernel*. In both functions, the strength of competition depends on the values of the resource acquisition trait (*B*) and an additional parameter that influences the shape of the kernel (see details in Appendix 2).

The seed production module determines the number of seeds produced depending on the strength of competition of the mother plant (see Appendix 2 for more details).

One function called *randomITV* is available for determining the trait values of seeds. Sources of intraspecific trait variation are heritable differences among individuals and phenotypic plasticity (Violle et al., 2012). Since „genetic variability and phenotypic plasticity cause similar effects in ecological communities" (Violle et al., 2012; see further references therein), processes creating ITV were not mechanistically modeled, rather a simple phenomenological model was applied. It supposes that traits follow the same distribution within- and between-species and the ratio of the two variance components are the same for all traits (see mathematical details in Appendix 1). The reason for choosing this approach was its simplicity, not its reality. For example, the constant ratio of within- and between-species variation contradicts field experiences. The correlation among traits is also fixed and the same correlation matrix is applied for simulating within- and between-species variation, however, correlation structure may vary among species (e.g. Messier et al., 2018) or even among populations of the same species (Boucher et al., 2013). On the other hand, it should be emphasized that ITV measured in field studies has already been shaped by selection, while we aim to model potential ITV that is still not influenced by selection.

The propagule spreading model decides whether a produced propagule remains in the local community where it was born or spreads to another local community. Currently, only the *MetaCom.Dispersal* function is available. This function implements a classic meta-community where propagules spread into each of the other communities with equal probability independently from the spatial position of local communities and traits of species. The survival probability module calculates the probability that a seedling survives and grows into an adult. Currently, only one function called *Gaussian.tolerance* is available. It assumes that the probability of survival (*s*) decreases as the distance between the habitat trait value and the local environment value increases according to a Gaussian curve.

2.2 | Flexibility and limitations of the simulation framework

Due to its modular structure, the presented framework has high flexibility. Users can choose the most suitable function for each module or write their function if neither built-in function fulfills the requirements. When the main function calls the functions of modules, it uses all additional parameters. Thus, user-defined functions have freedom in terms of the number and names of parameters. However, the use of the same parameter names in different functions has to be avoided except if they are supposed to have the same value (see parameters related to trait distributions in the *Gener.species.pool* and *randomITV* functions).

Of course, there are limits of flexibility. The steps of the disturbance-regeneration cycle are fixed in the main module as its modification would result in a completely different simulation framework. Spatially implicit dynamics within local populations also cannot be changed.

2.3 | Sampling from the simulated communities

The *comsimitv* R package contains two functions for the simulated field estimation of abundances and trait values. They are useful for showing how sampling effort influences the results. The function *comm.sampling* randomly selects a predefined individual from each local community without replacement and counts species abundances in these samples. It can also be used to transform simulation results into a site-by-species matrix. The function *trait.sampling* simulates different trait measurement scenarios. If the parameter *ITV* is equal to *FALSE*, a predefined number of individuals from each species are chosen and measured irrespectively of the sub-communities. Otherwise, trait measurements are done in each sub-community separately. This may result in raw measurement or aggregates them by calculating mean values.

3 | Comparing ecolottery and comsimitv packages

There are two conceptual differences between ecolottery and comsimitv. The *ecolottery* package simulates the species composition of a (habitat) island, while *comsimitv* simulates the dynamics of a (closed) meta-community. In *ecolottery*, there is no emigration, and immigrants arrive from the constant external species pool. Although this external species pool could be interpreted as the rest of the meta-community, this design does not allow modeling dispersal among sub-communities. In comsimitv, the external species pool is used only for creating initial communities, then the meta-community is closed, i.e. there is no immigration from and emigration to the outside of the meta-community, however, there is dispersal between local communities.

Another conceptual difference is handling intra-specific variation. In ecolottery, each individual is identical to its parent. The only source of ITV is the random trait variation in the external pool. The advantage of this approach is that details of heredity do not have to be modeled. On the other hand, this approach narrows the range of processes that can be modeled by excluding any process related to the inheritance of trait values (e.g. probability of mutations, phenotypic variation, maternal effect).

The *ecolottery* package applies two simulation approaches: coalescent and forward simulation, while in the *comsimitv* package only the forward simulation is available. The main advantage of coalescent simulation is that computation time is much lower, especially for large community sizes (see Figure 4 in Munoz et al., 2018). Its disadvantage is that the selection rule is independent of the actual community composition, thus the effect of limiting similarity cannot be modeled. In forward simulation by ecolottery package, the probability of death depends on traits of the focal individual and the actual trait composition of the community, allowing simulation effect of limiting similarity. Gaussian overlap is built-in in the simulation, other competition kernels cannot be directly used, however user-defined trait distance function can be applied that allows some flexibility. The comsimitv package contains three competition kernel functions (Appendix 2) and allows using any further user-defined kernel.

There is a trade-off between the simplicity of models and their ability to quantitatively predict the composition of local communities (Evans et al., 2013). Despite the above-discussed differences, *ecolottery* and *comsititv* are very similar in their complexity/realism. Both can be categorized as “strategic models based on the phenomenological description” (Evans et al., 2013). They are “strategic models” because their aim is qualitative description and understanding of the role of processes rather than making quantitative predictions for given conditions, however, *ecolottery* allows estimating parameters of the best fitting model. On the other hand, they are “based on the phenomenological description”, because the exact form of built-in functions (e.g. competition kernels) cannot be derived from first principles. Neither of them aims to make predictions for specific traits. It makes connecting traits in simulations and traits measured in the field very difficult that often “soft” traits are measured that may be related to both environmental filtering and limiting similarity processes.

4 | EXAMPLES

Beyond checking the behavior of statistical tests, process-based models are suitable for virtual experiments where the effect of otherwise hardly manipulated variables on the community composition can be studied (Connolly et al., 2017). Simulated community and trait data can be used to optimize the sampling effort as illustrated by the first example. The second example is a virtual experiment which aims to illustrate such type of application of simulation models rather than do exhaustive research on the topic. Finally, the third example checks the statistical behavior of T-statistics and illustrates that checks by mechanistic models are more reliable than tests based on artificially constructed trait distributions.

4.1. Example 1

In this example, global mean and local mean sampling designs (Carmona et al., 2015) are compared. The simulation of a meta-community was done using default parameters except for the simulation length, which was set to 100, and *ITV.ratio*, which was set to 0.1 or 0.5 (lower *ITV.ratio* means lower intra-specific variation; for more details see Appendix 1.) In global sampling, each species was characterized by the mean trait value of 10 individuals selected randomly irrespectively of the locality. In local sampling, one individual was measured for each occurring species in each local community. Next, the community-weighted mean (CWM) and variance (CWV), which are two widely used characteristics of trait composition (Ricotta & Moretti, 2011), were estimated for each community using both globally and locally measured trait values. The departures of estimates from the real values (i.e. the mean and variance of all individuals’ trait values) were measured by the relative bias for each local community:

$$B_{rel} = \frac{\hat{\theta} - \theta}{\theta}$$

where $\hat{\theta}$ denotes the estimated value of each local community using global mean or local mean sampling designs, while θ denotes the real value (i.e. value calculated after measuring all individuals).

We expected that in CWM, both estimates would be unbiased (i.e. relative bias values spread around zero), while both samplings would underestimate CWV as a result of neglecting total ITV (global sampling) or the within-site component of ITV (local sampling). As expected, the relative bias values of CWM were spread around zero (Figure 1, Appendix S3.1). The difference between the two sampling designs was negligible in mean bias. However, sometimes the variation of bias was higher in global sampling.

In the case of CWV, we obtained the expected pattern for the neutral and resource use traits: global sampling underestimated the variance, while local sampling performed better (Figure 2, Appendix S3.2, S3.3). Surprisingly, global sampling greatly overestimated CWV of the habitat preference trait. The explanation for this is the local adaptation of species: trait values of locally occurring individuals depart from species' global means toward the local optimum trait values, which is close to the local CWM. By applying local sampling, we remove this source of bias, so the expected moderate underestimation is obtained.

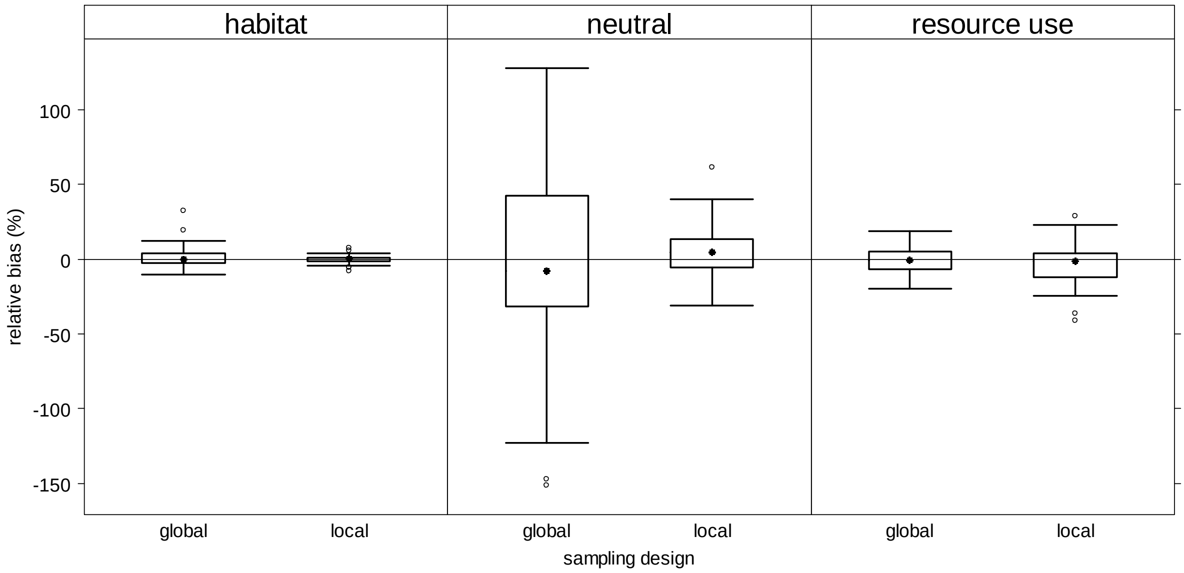


Figure 1: Relative bias of community-weighted mean (CWM) in different traits under different sampling designs at the low level of ITV (*ITV.ratio*=0.1).

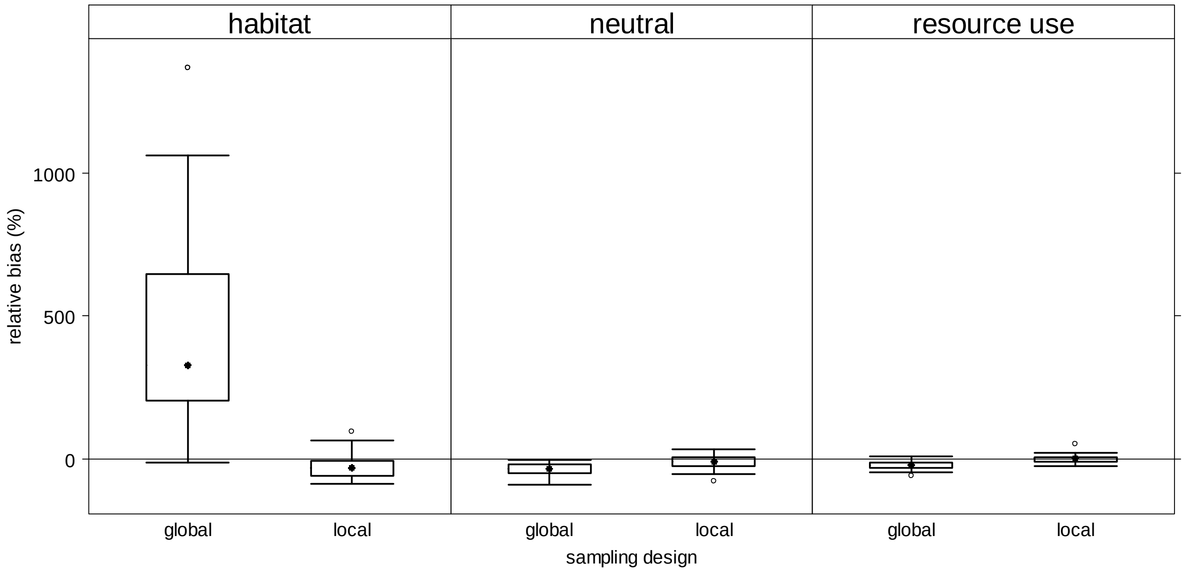
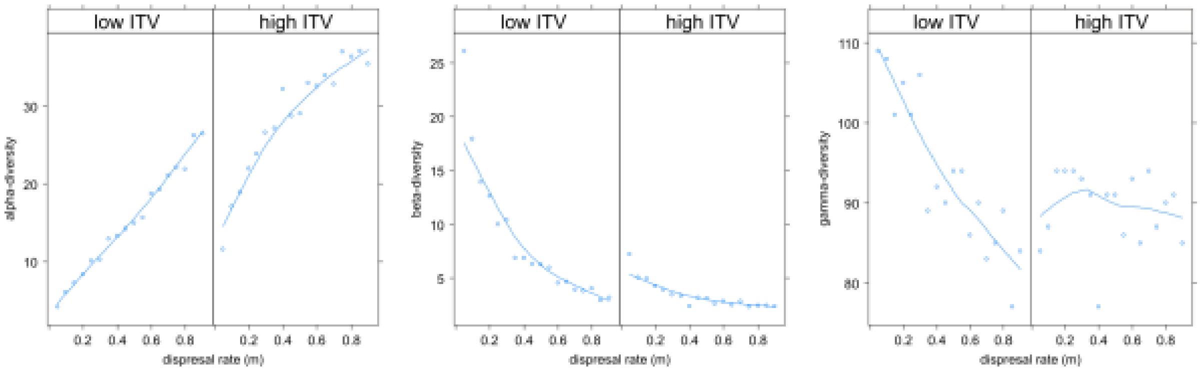


Figure 2: Relative bias of community-weighted variance (CWV) in different traits under different sampling designs at the low level of ITV (*ITV.ratio*=0.1).

.2. Example 2: effect of dispersal and ITV on species richness

Meta-communities were simulated with dispersal rate (*m*) changing from 0.1 to 0.9 in steps of 0.05, *ITV.ratio* of 0.1 (low ITV) or 0.9 (high ITV), and default settings for the other parameters. The number of species in the meta-community (gamma diversity), the mean species richness of sub-communities (alpha diversity), and their ratio (beta diversity; Jost, 2007) were calculated for each meta-community. It is a fully explorative study (Tredennick et al., 2021): we have no testable hypotheses a priori.

At low ITV, increasing dispersal resulted in higher alpha but lower beta and gamma diversity, as expected (Figure 3). When ITV is high, alpha diversity is higher, but its dependence on the dispersal rate remains almost the same. Gamma diversity, however, does not show a clear trend, while the decrease of beta diversity is less steep. This simulation highlights that the effect of ITV on diversity patterns is worth detailed studies and that process-based simulations are valuable tools for such studies.



a) b) c)

Figure 3: Dependence of alpha (mean richness at sub-community level), beta (ratio of gamma and alpha diversity), and gamma diversity (species richness at meta-community level) on dispersal rate and ITV.

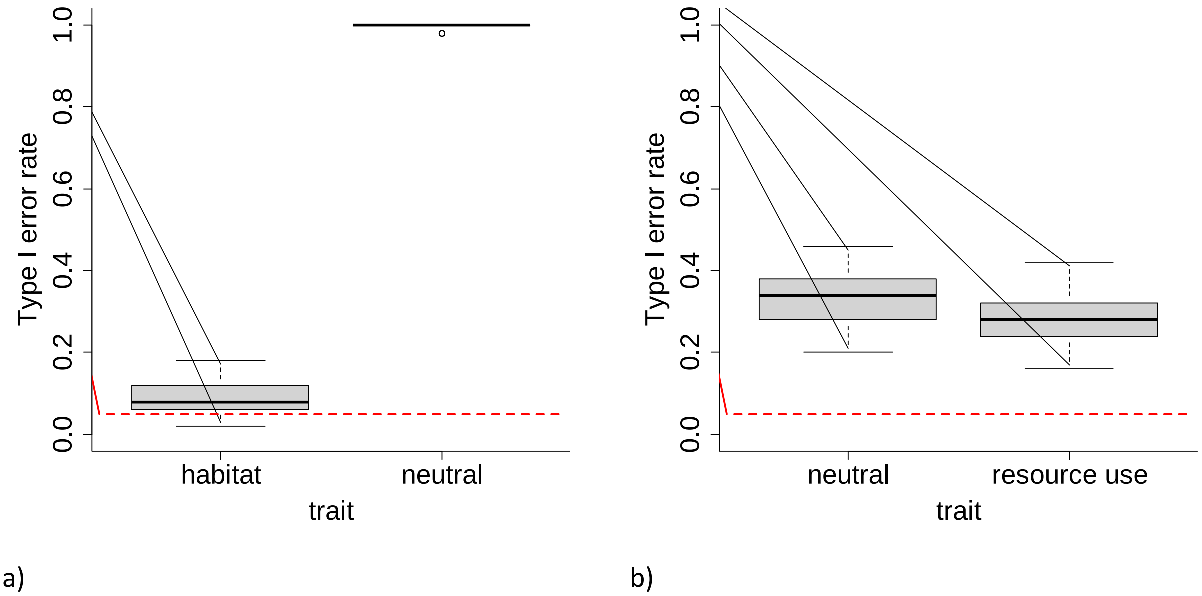
4.3. Example 3: checking Type I error rate of T-statistics

Violle et al. (2012) proposed that variance ratios, called T-statistics, could be used to detect the effect of external (i.e. habitat conditions) and internal filtering (i.e. limiting similarity) in trait distribution. Taudiere and Violle (2016) in their Appendix 4 have already checked the type I and type II error rates of T-statistics. However, they applied trait-based rules to create simulated communities, not a process-based simulation. According to my experience (Botta-Dukát & Czúcz, 2016), checking the error rate based on process-based simulation may reveal problems in methods that passed the previous checks. Thus, in spite of a detailed study by Taudiere and Violle (2016), checking error rates using process-based simulation may give new results.

I have focused on the two statistics that use individual-level trait values: $T_{IP/IC}$ and $T_{IC/IR}$. The former can be used to detect internal filtering; if co-existing species have to differ in their traits (cf. limiting similarity; MacArthur & Levins, 1967), $T_{IP/IC}$ should be lower than random expectation. On the other hand, if $T_{IC/IR}$ is lower than expected, it can be interpreted as the effect of external (habitat) filtering. For appropriate null-models see Taudiere and Violle (2016).

Simulations were done using all possible combinations of three levels of dispersal rate (m = 0.05; 0.1; 0.5) and *ITV.ratio* (0.01; 0.1; 0.25) and default setting of all other parameters. Five replicate meta-communities were simulated with each of the nine parameter combinations, resulting in total 45 meta-communities. One-sided test of external and internal filtering was done by the *Tstats* function of the *cati* package (Taudiere & Violle, 2016), using $T_{IC/IR}$ and $T_{IP/IC}$ as test statistics, respectively. The proportion of communities, where the null-hypothesis was rejected at α=5% significance level, was calculated for each trait in each meta-community. It is an estimate of the type I error rate if the null-hypothesis should not be rejected for the given trait. External filtering is related to the "habitat" trait, thus null hypotheses should be rejected only for this trait, but should not be rejected for the other two traits (i.e. "neutral" and "resource use"). Similarly, internal filtering is related to the "resource use" trait, thus the type I error was committed when the null-hypothesis was rejected for the other two traits (i.e. "neutral" and "habitat") in this test.

In general, the type I error rate is higher than the used significance level (Figure 4). In the testing of internal filtering, there is a big difference in the type I error rate between neutral and habitat traits, while in the testing of external filtering the difference is negligible between the two traits.



a) b)

Figure 4: Type I error rate in testing internal filtering by $T_{IP/IC}$ (a) and external filtering by $T_{IC/IR}$ (b) calculated for two traits whose distribution was not influenced by the tested filter. The dashed red line indicates the applied significance level (5%).

Contrary to these results, Taudiere and Violle (2016) have found that type I error rate does not differ considerably from the significance level. The reason for this contradiction is that they analyzed trait distributions created according to the null-hypothesis. In this study, I analyzed the distribution of traits that were not affected by the tested filter. The difference between the results of the two approaches indicates that the distribution of traits not affected by the tested filter does not necessarily follow the null-hypothesis.

Testing internal filtering by $T_{IP/IC}$ assumes that if there was no competition among coexisting species (or it had no effect on the studied trait), species' trait distributions would be the same, and only the limiting similarity can create the differences between them. It implicitly assumes that there are no among-species differences due to phylogeny. Although it would be hard (if not impossible) to falsify this implicit assumption, it is highly implausible. The simulation starts with communities without intraspecific trait variation. ITV emerges during the simulation due to differences between mothers and their offspring and among offspring of the same mother. It may mean that phylogenetic determination of the trait values is stronger in the simulation than the nature, thus this simulation may over-estimate the type I error of this test. However, regarding the implausibility of the assumption behind the test, I think the type I error rate always exceeds the significance level.

Testing external filtering by $T_{IC/IR}$ assumes that without habitat filtering the community-level trait distribution would be the same in all localities. If the trait was affected by the habitat filtering, the null hypothesis is almost always rejected by this test. Thus it is clear that habitat filtering creates differences among community-level trait distributions. But the high type I error rates indicate that another mechanism is also able to create such differences. In the simulation drift is the only mechanism that affects all traits, including the neutral one. Although offspring do not have the same as traits of their mother, they are similar to it. Since most of the produced seeds remain in the locality where the mother plant grows, drift could create differences among localities detectable by this test.

5 | CONCLUSION

Simulation of community assembly that considers intraspecific trait variation and results in not only community structure, but also data on ITV, can be a valuable tool for community ecology. The R package presented in this paper has two advantages: it allows modeling ITV and due to its modular structure, it is easy to include new features. As a maintainer of the R package, I am willing to include functions developed by other ecologists and develop new functionalities upon request. The most important development to be done in near future by myself are: (1) including species-specific dispersal rate and spatially explicit meta-community; (2) allowing loading of user-defined (field-measured) traits; (3) including the number of seeds in the competition-free environment as a new trait; and (4) allowing age-dependent mortality and adding the number of died individuals (strength of fine-scale disturbance) as a new parameter.

Although these developments will allow more realistic simulation scenarios, the presented examples (see also Appendix 4) illustrate that package has already been useful for several purposes: doing *in silico* experiments, optimizing sampling design, and checking the statistical behavior of new statistical procedures.

References

Botta-Dukát, Z., & Czúcz, B. (2016). Testing the ability of functional diversity indices to detect trait convergence and divergence using individual-based simulation. *Methods in Ecology and Evolution*, *7*(1), 114–126. https://doi.org/10.1111/2041-210X.12450

Boucher, F. C., Thuiller, W., Arnoldi, C., Albert, C. H., & Lavergne, S. (2013). Unravelling the architecture of functional variability in wild populations of *Polygonum viviparum* L. *Functional Ecology*, *27*(2), 382–391. https://doi.org/10.1111/1365-2435.12034

Bricca, A., Musciano, M. D., Ferrara, A., Theurillat, J.-P., & Cutini, M. (2022). Community assembly along climatic gradient: Contrasting pattern between- and within- species. *Perspectives in Plant Ecology, Evolution and Systematics*, *56*, 125675. https://doi.org/10.1016/j.ppees.2022.125675

Carmona, C. P., Bello, F., Mason, N. W. H., & Lepš, J. (2019). Trait probability density (TPD): Measuring functional diversity across scales based on TPD with R. *Ecology*, *100*(12), e02876. https://doi.org/10.1002/ecy.2876

Carmona, C. P., de Bello, F., Mason, N. W. H., & Lepš, J. (2016). Traits Without Borders: Integrating Functional Diversity Across Scales. *Trends in Ecology & Evolution*, *31*(5), 382–394. https://doi.org/10.1016/j.tree.2016.02.003

Carmona, C. P., Rota, C., Azcárate, F. M., & Peco, B. (2015). More for less: Sampling strategies of plant functional traits across local environmental gradients. *Functional Ecology*, *29*(4), 579–588. https://doi.org/10.1111/1365-2435.12366

Connolly, S. R., Keith, S. A., Colwell, R. K., & Rahbek, C. (2017). Process, Mechanism, and Modeling in Macroecology. *Trends in Ecology & Evolution*, *32*(11), 835–844. https://doi.org/10.1016/j.tree.2017.08.011

Evans, M. R., Grimm, V., Johst, K., Knuuttila, T., de Langhe, R., Lessells, C. M., Merz, M., O'Malley, M. A., Orzack, S. H., Weisberg, M., Wilkinson, D. J., Wolkenhauer, O., & Benton, T. G. (2013). Do simple models lead to generality in ecology? *Trends in Ecology & Evolution*, *28*(10), 578–583. https://doi.org/10.1016/j.tree.2013.05.022

Götzenberger, L., Botta-Dukát, Z., Lepš, J., Pärtel, M., Zobel, M., & de Bello, F. (2016). Which randomizations detect convergence and divergence in trait-based community assembly? A test of commonly used null models. *Journal of Vegetation Science*, *27*(6), 1275–1287. https://doi.org/10.1111/jvs.12452

Götzenberger, L., de Bello, F., Bråthen, K. A., Davison, J., Dubuis, A., Guisan, A., Lepš, J., Lindborg, R., Moora, M., Pärtel, M., Pellissier, L., Pottier, J., Vittoz, P., Zobel, K., & Zobel, M. (2012). Ecological assembly rules in plant communities—Approaches, patterns and prospects. *Biological Reviews*, *87*(1), 111–127. https://doi.org/10.1111/j.1469-185X.2011.00187.x

Gross, N., Le Bagousse-Pinguet, Y., Liancourt, P., Saiz, H., Violle, C., & Munoz, F. (2021). Unveiling ecological assembly rules from commonalities in trait distributions. *Ecology Letters*. https://doi.org/10.1111/ele.13789

Hardy, O. J. (2008). Testing the spatial phylogenetic structure of local communities: Statistical performances of different null models and test statistics on a locally neutral community. *Journal of Ecology*, *96*(5), 914–926. https://doi.org/10.1111/j.1365-2745.2008.01421.x

Jung, V., Violle, C., Mondy, C., Hoffmann, L., & Muller, S. (2010). Intraspecific variability and trait-based community assembly. *Journal of Ecology*, *98*(5), 1134–1140. https://doi.org/10.1111/j.1365-2745.2010.01687.x

Kattge, J., Bönisch, G., Díaz, S., Lavorel, S., Prentice, I. C., Leadley, P., Tautenhahn, S., Werner, G. D. A., Aakala, T., Abedi, M., Acosta, A. T. R., Adamidis, G. C., Adamson, K., Aiba, M., Albert, C. H., Alcántara, J. M., Alcázar C, C., Aleixo, I., Ali, H., … Wirth, C. (2020). TRY plant trait database – enhanced coverage and open access. *Global Change Biology*, *26*(1), 119–188. https://doi.org/10.1111/gcb.14904

Kraft, N. J. B., Valencia, R., & Ackerly, D. D. (2008). Functional Traits and Niche-Based Tree Community Assembly in an Amazonian Forest. *Science*, *322*(5901), 580–582. https://doi.org/10.1126/science.1160662

Laughlin, D. C., Joshi, C., van Bodegom, P. M., Bastow, Z. A., & Fulé, P. Z. (2012). A predictive model of community assembly that incorporates intraspecific trait variation. *Ecology Letters*, *15*(11), 1291–1299. https://doi.org/10.1111/j.1461-0248.2012.01852.x

Mammola, S., & Cardoso, P. (2020). Functional diversity metrics using kernel density *n* -dimensional hypervolumes. *Methods in Ecology and Evolution*, *11*(8), 986–995. https://doi.org/10.1111/2041-210X.13424

McGill, B. J., Enquist, B. J., Weiher, E., & Westoby, M. (2006). Rebuilding community ecology from functional traits. *Trends in Ecology & Evolution*, *21*(4), 178–185. https://doi.org/10.1016/j.tree.2006.02.002

Medrano, M., Herrera, C. M., & Bazaga, P. (2014). Epigenetic variation predicts regional and local intraspecific functional diversity in a perennial herb. *Molecular Ecology*, *23*(20), 4926–4938. https://doi.org/10.1111/mec.12911

Münkemüller, T., Gallien, L., & Travis, J. (2015). VirtualCom: A simulation model for eco-evolutionary community assembly and invasion. *Methods in Ecology and Evolution*, *6*, 735–743.

Munoz, F., Grenié, M., Denelle, P., Taudière, A., Laroche, F., Tucker, C., & Violle, C. (2018). ecolottery: Simulating and assessing community assembly with environmental filtering and neutral dynamics in R. *Methods in Ecology and Evolution*, *9*, 693–703. https://doi.org/10.1111/2041-210X.12918

Pavoine, S., & Bonsall, M. B. (2011). Measuring biodiversity to explain community assembly: A unified approach. *Biological Reviews*, *86*(4), 792–812. https://doi.org/10.1111/j.1469-185X.2010.00171.x

Ricotta, C., & Moretti, M. (2011). CWM and Rao’s quadratic diversity: A unified framework for functional ecology. *Oecologia*, *167*(1), 181–188. https://doi.org/10.1007/s00442-011-1965-5

Violle, C., Enquist, B. J., McGill, B. J., Jiang, L., Albert, C. H., Hulshof, C., Jung, V., & Messier, J. (2012). The return of the variance: Intraspecific variability in community ecology. *Trends in Ecology & Evolution*, *27*(4), 244–252. https://doi.org/10.1016/j.tree.2011.11.014

Wilson, J. B. (1995). Null Models for Assembly Rules: The Jack Horner Effect Is More Insidious Than the Narcissus Effect. *Oikos*, *72*(1), 139–144. https://doi.org/10.2307/3546047

Supplementary materials:

- Appendix 1: Generating trait values and modeling intraspecific trait variation in comsimitv package
- Appendix 2: Modelling Competition and Seed production in comsimitv package
- Appendix 3: Additional figures
- Appendix 4: R scripts and data for re-creating figures