6

7 **Trait-based numerical classification of mesic and wet grasslands in Poland**

8

9 Attila Lengyel[1,2], Grzegorz Swacha[1], Zoltán Botta-Dukát[2], Zygmunt Kącki[1]

10

11 [1]Botanical Garden, University of Wrocław, ul. Sienkiewicza 23, 50-335 Wrocław, Poland

12 [2]Centre for Ecological Research, Institute of Ecology and Botany, Alkotmány u. 2-4, 2163
13 Vácrátót, Hungary

14 A.L.: corresponding author, lengyel.attila@okologia.mta.hu, ORCID: 0000-0002-1712-6748

15 G.S.: grzegorz.swacha@uwr.edu.pl, ORCID: 0000-0002-6380-2954

16 Z.B.D.: botta-dukat.zoltan@okologia.mta.hu, ORCID: 0000-0002-9544-3474

17 Z.K.: zygmunt.kacki@uni.wroc.pl, ORCID: 0000-0002-2241-1631

18

21

22 **Running title**

23 Trait-based classification of grasslands

24

25 **Abstract**

26 *Questions*: What vegetation types can be distinguished on the basis of plant functional traits

27 using numerical classification? How do they match syntaxonomical units?

28 *Location*: Poland

29 *Methods*: 6985 vegetation plots representing mesic and wet grasslands (*Molinio-*

30 *Arrhenatheretea, Polygono-Poetea*) were retrieved from the Polish Vegetation Database.

31 Plant functional trait data were assembled from the LEDA and Clo-Pla databases for most

32 species occurring in the data set. Community-weighted mean for five traits were calculated

33 for each plot: specific leaf area, canopy height, seed mass, bud bank index and clonality

34 index. Plots were classified using Ward's method and iterative relocation based on silhouette

35 widths. The clusters were interpreted and characterized on the basis of species and trait

36 composition, functional diversity, functional redundancy, Ellenberg indicator values, and

37 geographical distribution. The similarity between the trait-based classification and the

38 syntaxonomical assignment of plots is evaluated both statistically and by expert knowledge.

39 *Results*: Twelve clusters were distinguished. The classification mirrored the main gradients

40 structuring grasslands in Poland, although, some vegetation types with the strong dominance

41 of functionally unique species appeared more distinct than they are treated in syntaxonomy.

42 Clusters did not differ significantly in functional diversity and redundancy. The differences of

43 clusters in species and trait composition and environmental background are discussed.

44 *Conclusion*: The application of trait data and numerical methods is a promising approach for

45 obtaining vegetation classifications. Such classifications can be in closer relationship with the

46 most important ecosystem processes than floristic classifications because communities

47 comprising different species but similar functional trait distribution are not separated. Trait-

48 based classifications match phytosociological units to a variable degree. Functional

49 uniqueness and variation of abundance determines how individual species influence the

50 delimitation of vegetation types using our approach.

51

Introduction

Due to its central role in ecosystem processes, vegetation characteristics are frequently used as general descriptors of ecosystems or habitat types for the purposes of nature conservation, land-use planning, and landscape mapping. Vegetation-plot databases are widely used for establishing classifications, very often with the application of statistical methods (De Cáceres et al., 2015). Such databases contain tens or hundreds of thousands of species by site records collected during the long history of phytosociology, often with additional data on vegetation physiognomy, geographical location and environmental background (Dengler et al., 2011). These data sources make it possible to answer questions about vegetation variation on scales as broad as countries or continents. As the essential type of data which is recorded in these plots with the highest consistency is species occurrence (and often some form of abundance), most typically these analyses use species as variables and sites as objects. In consequence, the classifications reflect patterns in species composition, together with all the possible mechanisms which influence community assembly, including selection, speciation, dispersal, and drift (Vellend, 2010). However, some of these processes, e.g. random drift, may not be interesting from the viewpoint of the potential user of the classification. Limited dispersal of species has strong consequences on classification results. If the sample includes areas with different site history, which is a common situation, species may not have had enough time to colonise all habitat patches which would have been suitable for them. In this case, a classification based on species composition will reflect not only environmental gradients but differences in regional species pools. When the geographical extent of the study is very large, and the effect of site history is strong, it can become impossible to reach a vegetation classification reflecting environmental gradients, which would be valid over the entire study area. This might be a primary reason for the high level of idiosyncrasy in national vegetation classifications. Patterns of speciation are mainly relevant on biogeographic scales in time and space; although, in specific studies differences in the phylogenetic structure of communities may be important (Lososová et al. 2015). Nevertheless, most vegetation classification studies seek answers for questions about what types of communities exist, and how their occurrence is related to environmental gradients, and ecosystem functions – i.e., the process of selection. However, it is increasingly recognised that patterns in species identities are not always tightly related with ecosystem properties, instead, traits of species are more relevant from this perspective (Díaz & Cabido, 2001; Díaz et al., 2004). Species respond to biotic and abiotic factors by their traits (response traits), as well as they form their environment by them (effect

traits; Lavorel & Garnier, 2002). Species which have similar traits may substitute each other without significantly altering ecosystem functioning – a phenomenon called functional redundancy (Hooper et al., 2005). Integrating the trait-based approach should improve the relevance of vegetation classification with respect to ecosystem functioning, and enhance the generalizability of results. Hérault and Honnay (2007) already presented a classification where instead of species, groups of species sharing similar traits called 'emergent groups' were applied as variables. Hence, it was possible to differentiate two types of riverine forests in Luxembourg different in life-form spectra, dispersal modes, and conservation relevance. However, this study restricted its scope on a specific vegetation type of a rather narrow area, while the most typical challenge of recent vegetation classification works is providing relevant and generalizable results over broad sample coverage in space and along ecological gradients.

The functional approach of ecology has been in an intensive research period now for more than a decade, partially shifting the focus away from the study of patterns of species-level composition and diversity (Carlow, 1987; Tilman et al., 1997; McGill, Enquist, Weiher, & Westoby, 2006). A major outcome of this field is the emergence of plant trait databases, which provide trait measurements for thousands of taxa and hundreds of traits (Kattge et al., 2011). Technically they are readily connectable with vegetation-plot databases, providing avenues for the same types of analyses of trait composition which has only appeared at the level of species yet. A fundamental question of trait-based ecology is the distinction between processes which impedes the co-existence of functionally similar organisms (i.e. functional divergence) and those which promotes it (i.e. functional convergence; Lhotsky et al., 2016). Competitive exclusion is known to increase divergence according to the theory of limiting similarity, while environmental (or niche) filtering (and, for traits increasing competitive vigour, also competitive exclusion) supports functional convergence (Weiher & Keddy, 1995). Considering the interest of vegetation classification studies in the response of vegetation to environment, trait convergence should be a key phenomenon in the construction of functionally relevant classifications. Moreover, competitive exclusion is subordinated to environmental filtering according to the filter model of community assembly (Keddy 1992), which is a likely reason why trait convergence is more frequently detected than divergence.

In this paper we present a trial for integrating the functional approach into the context of vegetation classifications using multivariate statistical methods. Our aim is to classify semi-natural grasslands of Poland in a way that resulting groups are relatively similar in their trait

composition with no respect to species composition. We discuss the environmental

background, trait composition, functional diversity, and redundancy of the clusters

distinguished. We assess the similarity between the trait-based classification and the

syntaxonomical system. We expect the resulting classification to be generalizable over the

entire study area, while showing a strong relationship with ecosystem processes.

Materials and Methods

9725 phytosociological relevés representing temperate semi-natural grasslands were retrieved

from the Polish Vegetation Database (Kącki & Śliwiński, 2012; GIVD identifier: EU-PL-

001). In the syntaxonomical system according to Kącki, Czarniecka and Swacha (2013), these

grasslands are classified to *Molinio-Arrhenatheretea* class and comprise three orders, called

Potentillo-Polygonetalia (temporarily flooded and heavily grazed and trampled vegetation on

nutrient-rich soils), *Arrhenatheretalia elatioris* (lowland and montane mesic grasslands), and

Molinietalia caeruleae (wet grasslands and tall-forb vegetation). We included into the data set

also *Polygono arenastri-Poetea annuae* with one order *Polygono arenastri-Poetalia annuae*

(therophyte-rich dwarf-herb vegetation of trampled habitats) because it strongly resembles

heavily trampled pastures of the *Cynosurion* alliance (*Arrhenatheretalia*). *Polygono-Poetea*

annuae was treated as part of the *Molinio-Arrhenatheretea* in previous syntaxonomical

overviews in Poland. Plot size was restricted to 10 to 100 m^2. Moss and lichen species were

removed from the data set due to their uneven data availability across plots. Where species

covers were recorded on ordinal scales (e.g. Braun-Blanquet scales, which is ca. 90% of all

plots) cover categories were transformed to their respective mid-point percentages using the

JUICE software (Tichý 2002). Relevés with >5% cover of trees and shrubs were excluded.

The data set was subjected to geographical stratification and heterogeneity constrained

random resampling (Knollová, Chytrý, Tichý, & Hájek, 2005; Lengyel, Chytrý, & Tichý,

2011) using Bray-Curtis index calculated on square-root transformed abundance data. Strata

were 6'×10' in size. From each stratum the number of plots to select was determined as 5 +

(n-5) × d, where n is the total number of plots in the stratum and d is the mean pairwise

dissimilarity among plots within the stratum. This method down-weighted the contribution of

oversampled areas only if their beta-diversity was low, while intensively sampled but diverse

regions kept their high share (Wiser & De Cáceres 2013). No resampling was done in strata

containing five or less plots. The stratified resampling reduced the data set to 6985 plots.

A key decision in any trait-based study is the selection of traits to involve in the analysis. According to Westoby (1998) and Westoby, Falster, Moles, Vesk, and Wright (2002), the traits of the leaf economics spectrum (Wright et al., 2004), the height, and the seed represent the major dimensions of plant variability along the most typical ecological gradients (the so-called leaf-height-seed or LHS system). Leaf traits, especially specific leaf area (SLA; Cornelissen et al., 2003) are in close connection with resource acquisition and relative growth rate, and thus related to the productivity of the habitat (Wilson & Tilman, 1993; Lhotsky et al., 2016). Canopy height is positively correlated with the ability to outcompete other species in productive habitats, where light is the major limiting factor (e.g. Borer et al., 2014), while seed mass corresponds to reproduction strategy (e.g. Moles & Westoby, 2004). Besides LHS traits, there is growing evidence that clonal growth and bud bank are important in the adaptation of plants to regular disturbance, hence they have central role in the response of herbaceous vegetation to environment (Klimešová, Tackenberg, & Herben, 2016). Clonal growth enables plants to avoid disturbance, while bud bank is key in regeneration after minor damages. Five plant traits were included in the analysis, which can be regarded as the response traits that play the most fundamental role in the adaptation of plants to the environment and management regime. Specific leaf area (SLA), canopy height and seed mass were retrieved from the LEDA database (Kleyer et al., 2008). The 'bud bank' and the 'clonality' index were introduced according to Johansson, Cousins and Eriksson (2011) and E.-Vojtkó et al. (2016). Bud bank index is the rank sum of above- and belowground bud bank categories, while clonality index is the rank sum of lateral spread and total number of offspring per parent categories. These data were obtained from the Clo-Pla database (Klimešová, Danihelka, Chrtek, de Bello, & Herben, 2017). All measurements were subjected to a semi-automated outlier exclusion and averaging procedure for each species by traits. First, the mean and standard deviation of all measurements from a given species and given trait was calculated. Those measurements which differed by >2*SD from the mean were excluded. The remaining measurements were subjected to averaging weighted by the square-root of the number of replications given for each measurement in the public database (typically, the number of measured individuals for a given record). For species which lacked no more than two out of the five trait values, Bayesian Hierarchical Matrix Factorization (Schrodt et al., 2015) was used to fill the gaps in the trait table. Species with more than two missing trait measurements were rejected from the analysis resulting in 885 species in the final matrix. Plots where the relative cover of such rejected species was higher than 5% were excluded before the stratified resampling. Species-level mean trait values were checked for

189 normality by quantile-quantile plots. Since LHS traits proved to be right-skewed, they were

190 log-transformed. Then, all traits were standardized to mean = 0 and standard deviation = 1.

191 These species-level standardized means were used for calculating community-weighted

192 means (CWM; Garnier et al., 2004), which is considered a satisfactory indicator of niche

193 filtering along large-scale environmental gradients (Kleyer et al., 2012; De Bello et al., 2013).

194 The CWM is constrained by the relative dominance of the most dominant trait value in the

195 plot; therefore we expect species forming monodominant vegetation stands to have a high

196 influence on the classification (De Bello, Lepš, Lavorel, & Moretti, 2007). This conforms the

197 mass ratio hypothesis by Grime (1998) stating that ecosystem functioning is mainly

198 determined by traits of the dominant species.

199 Hierarchical classification was carried out by Ward's agglomerative method (Podani, 2000)

200 on the matrix of CWM values. Ward's method relies on Euclidean distances in the trait space

201 between plots. The upper 20 hierarchical levels of the classification were evaluated by several

202 cluster validity indices (see Appendix S1); however, they suggested different numbers of

203 clusters as optimal making it impossible to decide on a single 'best' solution. To overcome

204 this, on the one hand, we considered also the biological interpretation of the clusters and the

205 resolution desired typically in such large-scale classifications. Hence, the dendrogram was cut

206 at a particular level and then it was improved using iterative relocation methods (Roberts

207 2015). Most iterative relocation methods proposed by Roberts (2015) are computationally

208 very demanding; therefore, we applied the REMOS2 algorithm (Lengyel, Roberts, & Botta-

209 Dukát 2019). This procedure uses the silhouette width index (Rousseeuw 1987) to identify

210 misclassified objects, which are then re-assigned to their closest neighbour cluster. After re-

211 assignment, silhouette widths are updated, and misclassified plots are relocated again to their

212 closest neighbour cluster, until the classification cannot be further optimized. This has

213 changed the assignment of 31.58% of all plots. However, we did not change the hierarchical

214 relations of the basic clusters.

215 Delimited clusters were interpreted as biologically relevant units using expert-based

216 knowledge and we attempted to find correspondence to already known and well-defined

217 vegetation units in the traditional (floristic) syntaxonomical approach (Kącki et al., 2013). For

218 this purpose, phytosociological relevés were assigned to syntaxa at the class, order, and

219 alliance levels using their formal definitions. Relevés were classified to respective syntaxa

220 based on explicit definitions of vegetation units in the way that the relevé matched by the

221 definition of alliance must also match the definitions of the superior syntaxonomical units, i.e.

order and class. This classification system was created using combination of sociological

species groups (Bruelheide, 1997), and total cover of individual species or group of species

(Kočí, Chytrý, & Tichý, 2003; Dengler et al., 2006; Landucci, Tichý, Šumberová, & Chytrý,

2015), and is part of an ongoing project of vegetation classification in Poland. The outcome of

this classification is presented in the shortened synoptic table (Appendix S2). Assignments at

order and alliance levels were compared with the trait-based classification using the

symmetric version of Goodman and Kruskal's lambda index (Goodman & Kruskal, 1954). To

assess the strength of similarity, the observed lambda values were compared to a reference

distribution obtained by re-calculating the index after permuting the trait-based classification

9999 times. In total, five tests were performed: 1) class-level assignment vs. trait-based

classification with random permutation; 2) order-level assignment vs. trait-based

classification with random permutation; 3) order-level assignment vs. trait-based

classification with restricted random permutation using class-level assignment as strata; 4)

alliance-level assignment vs. trait-based classification with random permutation; 5) alliance-

level assignment vs. trait-based classification with restricted random permutation using order-

level assignment as strata. We report the P-values of the null hypothesis stating that the

similarity between the trait-based classification and the syntaxonomical assignment is as high

as we can observe due to chance. We also show the standardized effect sizes of the observed

values calculated by probit transformation (Botta-Dukát, 2018).

For each level of the hierarchical classification until reaching the level of basic clusters and

for each trait a Wilcoxon test was carried out to test the difference between the two clusters to

be merged in the respective fusion. Two-tailed P-values were calculated by using permutation

tests (Hothorn, Hornik, van de Wiel, & Zeileis, 2006). Bonferroni-corrected P-values and

standardized test statistics (W_{st}) were used for ranking traits by support to the tested fusions.

Distribution of CWMs across clusters are shown on 'boxes-and-whiskers' plots. Clusters are

also compared on the basis of functional vulnerability measured by Rao Q diversity index

(Botta-Dukát, 2005), and functional redundancy (Ricotta et al., 2016). We also provide

synoptic tables containing diagnostic, constant and dominant species, Ellenberg indicator

values (EIV; Ellenberg et al., 1992), and geographical distribution of the clusters as

Appendices S3-S5. Statistical comparison of clusters on the basis of variables dependent on

compositional information (e.g. diversity indices, aggregated species attributes) often result in

false positive tests due to the so-called 'similarity issue' (Zelený & Schaffers, 2012; Zelený,

2018) and other structural biases (Hawkins et al., 2017) if the same occurrence information

had been used for the definition of clusters. To avoid false conclusions we refrain from formal statistical tests in the case of between-cluster comparisons and present only boxes-and-whiskers plots.

To measure the influence of individual species on the trait-based (especially, CWM-based) classification we used a simple formula, that we call 'influence index'. That species affects CWM of a single community (e.g., a single relevé) the most, which has high abundance and highly different trait value from the other species in the community. If the variation of CWM across several communities is examined, species with highly variable abundance and unique trait values are supposed to be the most influential. Thus, we calculated the influence index for the kth species as follows:

$$I_k = \sqrt{D_{.k} \times SD(\mathbf{a_k})}$$

where $D_{.k}$ is the distance of the kth species from the unweighted mean of trait values of all species in the data set ('uniqueness'), and $SD(\mathbf{a_k})$ is the standard deviation of the abundance vector of species k ('variation of abundance'). Thus, the influence index is the geometric mean of two components, functional uniqueness and variation of abundance. We recommend to re-scale both $D_{.k}$ and $SD(\mathbf{a_k})$ by division by the maximum, respectively.

Calculations were carried out by the R software (R Core Team, 2017) using vegan (Oksanen et al., 2018), cluster (Maechler, Rousseeuw, Struyf, Hubert, & Horni, 2017), FD (Laliberté, Legendre, & Shipley, 2014), rapportools (Blagotić & Daróczi, 2015), and coin (Hothorn, Hornik, van de Wiel, & Zeileis, 2008) packages. Nomenclature of plants follow the Euro+Med PlantBase (last accessed on 27 Sep 2018), syntaxon names are according to Kącki et al. (2013).

Results

Most cluster validity measures indicated a peak value between 2 and 5, as well as another peak near 10 clusters (Appendix S1). After considering non-formal criteria, we chose the level of twelve basic clusters for the interpretation because it provided a reasonable compromise between details and conciseness; however, coarser solutions can easily be assembled from this fine-scale classification by merging low-level clusters according to the dendrogram fusions (Fig. 1). Distribution of CWM values across clusters is shown on Fig. 2. Synoptic tables and

textual descriptions of the clusters are presented in Appendix S3, EIVs in Appendix S4, and geographic distributions in Appendix S5.

Clusters 1 to 4 represent different kinds of grazed or frequently cut, often trampled grasslands. Cluster 1 consisted of plots of heavily grazed and trampled, or frequently cut grasslands from various soil types and moisture levels. This cluster is characterised by the lowest canopy among all clusters, high SLA and high clonality index. Cluster 2 represents mostly mesic grasslands on nutrient-poor and acidic soils which are mown or occasionally grazed. In this cluster SLA is slightly above and canopy height is slightly below the sample-wise average, while clonality is similarly high as in Cluster 1. Cluster 3 represents a small and distinct group of relevés dominated by *Agrostis stolonifera* and *Alopecurus geniculatus*. They occur on trampled, sometimes slightly alkaline, irregularly inundated habitats with nutrient-rich soils. In this cluster SLA is highest, canopy is second lowest, seed mass and bud bank are the lowest across all clusters, while clonality is highest among them. Similarly to Cluster 2, Cluster 4 contains plots mostly from extensively grazed mesic grasslands. SLA, seed mass and bud bank values of this cluster are slightly above the average of all clusters, canopy height is below average, clonality is intermediate.

Clusters 5 and 6 are two large, heterogeneous clusters containing several mesic and wet meadow types from lowland to montane sites. They are characterized by above-average seed mass and bud bank. Canopy height is lower in Cluster 5 than in Cluster 6.

Clusters 7 and 8 are two, distinct types with monodominant graminoid species with high canopy. Cluster 7 contains common lowland and montane wet grasslands, mostly with the dominance of *Scirpus sylvaticus*. This cluster has the highest canopy on average, highest bud bank on average, low seed mass and high clonality. Cluster 8 contains mesic meadows of ruderal character with the dominance of *Arrhenatherum elatius* mostly on post-arable lands converted to grasslands. In this cluster SLA, canopy, seed mass and clonality are high.

From Cluster 9 to 12 herbaceous communities of mostly wet habitats are found. Cluster 9 contains a variety of wet and mesic communities including intermittently wet meadows with *Molinia caerulea*, wet tall-forb vegetation, and nutrient-rich mesic meadows. This cluster has high seed mass, high bud bank, low clonality index and intermediate values for the other two traits. Cluster 10 contains montane meadows, and degraded wet meadows with *Deschampsia caespitosa* and to lesser extent *Juncus* species. This cluster was characterised by low SLA and clonality, high bud bank, and intermediate values for the other two traits. Cluster 11 contained

a variety of wet meadows with constant presence of tall forbs. This cluster had low SLA and clonality, and high values for the other three traits. Cluster 12 comprised relevés dominated by *Juncus* species, most frequently *Juncus effusus*, occasionally *J. subnodulosus* or *J. conglomeratus*. These stands occur mostly on nutrient-poor, waterlogged and acidic soils, which are sometimes managed by grazing. This cluster has high canopy and low values for all the other traits.

The cross-tabulation of the trait-based and the syntaxonomical classification is shown on Tab. 1. Permutation tests with Goodman and Kruskal's lambda index rejected the null hypothesis stating that trait-based classification and syntaxonomical assignment are as similar as expected by chance alone. Observed lambda values were higher than any element of the reference distribution using either the class-level (lambda = 0.007), order-level assignments (lambda = 0.313) or the alliance-level assignments (lambda = 0.236). In all cases $P < 0.001$ which gave SES = 3.719 after probit-transformation. However, the matching between syntaxonomical and trait-based classifications was not perfect. *Potentillion anserinae* (Clusters 2 and 3), *Juncion effusi* (Clusters 10 and 12), and *Polygono-Poetalia* (Clusters 1 and 4) were the few syntaxa which concentrated on a relatively limited number of trait-based clusters, while the majority of other units were more broadly dispersed across several clusters.

At the highest hierarchical level (i.e., two clusters), clonality showed a difference between the merged clusters which was the most extreme not only among all traits at that level but also across all levels ($W_{st} = 66.56$; Table 2). From the three-cluster level onwards, we could found no difference of this magnitude; although, with <7 clusters all tests showed significant difference between the merged clusters. The bud bank showed the second largest difference on absolute scale at the four-cluster level ($W_{st} = -38.04$). Apart from those mentioned above, we could recognize no pattern in the contribution of individual traits to the merging of clusters.

With some minor inequalities attributable to the unbalanced distribution of relevés, all clusters were distributed over almost the entire country, none of them was obviously restricted geographically.

In terms of functional diversity, clusters showed high overlap (Fig. 3). The highest median Rao Q was detected in Cluster 12, while the lowest in Cluster 1. Cluster 12 showed also the lowest functional redundancy together with Cluster 3. The other clusters resembled each other very much also in this aspect.

349   On Fig. 4 we show the distribution of species in the space of $D_{.k}$ and $SD(a_k)$ and Table 3

350   shows the ten species with the highest scores. Three of the first four species are those, which

351   form monodominant and distinct vegetation types (*Arrhenatherum elatius*, *Scirpus sylvaticus*,

352   *Juncus effusus*), while the rest species also occur as typical dominants of certain clusters

353   (Appendix S3).

354

355   **Discussion**

356   In our paper we present the numerical classification of plots representing semi-natural

357   grasslands of Poland, based on plant trait data, more specifically, on community-weighted

358   trait means of phytosociological relevés.

359   Using the emergent group approach, Hérault and Honnay (2007) showed that the involvement

360   of trait data into classification could provide typologies which reflect certain ecosystem

361   properties better than what would be achieved using only species composition. The main

362   difference between Hérault and Honnay's approach and ours lays in how we took into account

363   trait information. Hérault and Honnay classified species on the basis of their trait values into

364   'emergent groups', which were used as variables instead of species. The power of the

365   emergent group approach is that it accounts for functionally redundant species explicitly,

366   since emergent groups consist of species possessing the same trait syndrome and thus having

367   very similar ecological functions. On the other hand, classification of species into discrete

368   groups requires subjective decisions from the researcher regarding the clustering algorithm,

369   similarity measure, and number of emergent groups. Moreover, even objective algorithms

370   produce non-intuitive classifications due to methodological constraints, e.g. certain methods

371   tend to prefer clusters with specific size or shape. Our approach avoided this pitfall by using

372   trait information as continuous variables to calculate CWMs which were input for

373   classification.

374   We divided the sample into 12 clusters based on biological interpretability; although, several

375   cluster validity indices had higher values at lower numbers of clusters. It might suggest that

376   the trait-based classification approach recognized coarser vegetation units than we found

377   relevant and well separable. Nevertheless, classification studies are often aimed at providing

378   vegetation typologies at different hierarchical levels, enabling practitioners to choose the most

379   suitable resolution for a given application. The finer cluster resolution discussed here allows a

380   more detailed overview of the whole sample with reduced within-cluster heterogeneity;

381 however, for specific purposes it is still possible to merge lower-level clusters, e.g., according

382 to the fusions of the dendrogram. Therefore, our view of vegetation classification and

383 typology suggested here is flexible to a degree.

384 The trait-based classification mirrored the most significant gradients shaping grassland

385 vegetation of Poland, i.e. soil nutrient supply, soil moisture, and management. At the highest

386 classification level, mostly nutrient-rich and mesic types (Clusters 1 to 8, except Clusters 3

387 and 7) were separated from communities of nutrient-poor and wet habitats (Clusters 9 to 12).

388 Clusters 1 to 4 form a separate group at the four-cluster level. Their separation at high

389 hierarchical level is notable, since these trampled and grazed, highly specialized grasslands

390 include plots which differ in species composition very much but they are rather similar in

391 terms of physiognomy and traits with characteristically low canopy and high SLA.

392 Goodman and Kruskal's lambda with a permutation test rejected the independency between

393 the trait-based and the syntaxonomical classification. This is not very surprising given that

394 both formal definitions and CWM values rely on the species composition of relevés; however,

395 we were not able to design a formal test of similarity with higher practical relevance since

396 there is no standard threshold for 'tolerable difference' determining whether two

397 classifications can be considered the same or not. With expert-based evaluation of the clusters

398 we could point out several mismatches between the trait-based classification and the

399 syntaxonomical system. A striking example can be seen in form of clusters which were

400 dominated by functionally unique species, e.g. *Scirpus sylvaticus* (Cluster 7), *Arrhenatherum*

401 *elatius* (Cluster 8), or *Juncus* spp. (Cluster 12). These vegetation types are either

402 differentiated at the association (e.g., *Scirpetum sylvatici*) or alliance level (e.g., *Juncion*

403 *effusi*), or not differentiated unequivocally (e.g., grasslands dominated by *Arrhenatherum*

404 *elatius*) in the syntaxonomical system, while in the trait-based classification they appeared as

405 very distinct clusters standing alone sometimes even at high hierarchical levels. Obviously,

406 functionally unique and monodominant types also defined as separate syntaxa increase

407 matching between syntaxonomic and trait-based classification, while syntaxonomically

408 undefined types decrease it. Since monodominant communities are often species-poor, their

409 distinct occurrence in the trait-based classification might be viewed as an artefact attributable

410 to differences in species richness, considering that the more species are selected from the total

411 species pool, the less likely it is to obtain an extreme community-weighted mean.

412 Nevertheless, we consider differences in dominance structure as a relevant aspect of the

413 biological phenomenon we study which mirrors environmental stress, disturbance, or specific

site-history that should not be removed from the analysis. The influence index accurately
identified those species which appeared as dominants of certain clusters; therefore, we
recommend its application for estimating the influence of individual species on CWM-based
classifications. On the other hand, several alliances with more balanced dominance structure,
higher species richness, and higher functional similarity between species did not separate well
in the trait-based classification. For example, most meadow alliances, including
Arrhenatherion and *Polygono-Trisetion* in the *Arrhenatheretalia* order, and *Calthion*,
Cnidion, and *Molinion* in *Molinietalia*, similarly occurred in Clusters 5, 6, and 9. Considering
the five traits we selected for our analysis, there is a high functional overlap between these
meadow types despite being assigned to different orders in the syntaxonomical system.
Importantly, the inclusion of other traits may explain specific functional differences between
these alliances and orders.

Clonality index was the trait having the highest influence on the classification, which is in line
with the findings of Klimesová et al. (2008, 2016) and E.-Vojtkó et al. (2016). Bud bank also
seemed to have a relatively strong impact. There are several possible reasons for the
efficiency of vegetative traits in revealing patterns in herbaceous vegetation. One reason is
that grasslands in the temperate zone are usually maintained by some form of biomass
removal, typically grazing or mowing. Plants adapt to such disturbances through avoidance or
regeneration using clonal and bud bank traits (Klimešová et al., 2016); although, affecting
other traits due to developmental trade-offs (Rusch, Wilmann, Klimešová, & Evju, 2011;
Herben, Šerá, & Klimešová, 2015). Differences in the form and timing of management may
be at least as significant as abiotic variation among the vegetation types included in this
analysis (i.e., mesic and wet, semi-natural grasslands without extreme conditions in abiotic
environment). Another potential explanation for the high influence of vegetative traits is that
the ability of clonal growth as expressed on the relatively coarse scale applied in the Clo-Pla
database shows lower levels of intraspecific variation due to stronger phylogenetic constraints
and less measurements error. We consider all these explanations similarly likely, and agree
that clonal and bud bank traits should be given high attention in the study of functional
responses of vegetation to environmental and management gradients.

We did not find striking difference between clusters in terms of functional diversity and
redundancy. Only Cluster 12 showed higher functional diversity and lower redundancy than
the others in median values, which can be explained also by the functional uniqueness and
high dominance of *Juncus* species. However, this may not be a reliable indication of the

447 vulnerability or conservation importance of this vegetation type, since *Juncus*-dominated

448 stands, especially with *Juncus effusus*, are very common on nutrient-poor, disturbed or

449 successional wetlands.

450 We had to apply simplifications during our analyses which could have limited us in revealing

451 certain patterns. We retrieved trait data from LEDA and Clo-Pla database with no respect to

452 the geographical and environmental origin of the records, thus neglecting an amount of

453 variation in trait values attributable to population-level adaptation to local conditions. With

454 the application of community-weighted mean in the description of plot-level trait values, and

455 using their Euclidean distances as dissimilarity measure, we neglected the role of intraspecific

456 variation, despite growing evidence on its significant role in community assembly and

457 response to environmental gradients (Bolnick et al., 2011; Violle et al., 2012; Siefert et al.,

458 2015). Phylogenetic constraints may also bias the relationship between CWM values and an

459 environmental gradient (Duarte, Debastiani, Carlucci, & Diniz-Filho, 2018). Since there is no

460 obvious implementation of phylogenetic correction into a classification framework yet, we

461 neglected this effect. Despite the greatest and honourable efforts of database curators,

462 differences in measurement protocols or technicalities may have caused an amount of

463 variation between data sets coming from different providers. Nevertheless, we believe that

464 these sources of bias do not compromise our results at the scale of the classification discussed

465 here.

466 In any trait-based study, the choice of the trait determines all the potential results and

467 conclusions. We included LHS, clonal and bud bank traits with equal weight because there is

468 growing evidence of their ability to describe major dimensions of plant variability and

469 response of plants to environmental and management gradients (Westoby et al., 2002;

470 Klimešová et al., 2016). We believe that these five response traits describe the most important

471 vegetation gradients appropriately in the analysed data set. However, different sets of traits, or

472 different weights attributed to them, may have resulted in fundamentally different

473 classifications. For specific studies, it is straightforward to select traits which are relevant for

474 the ecosystem property under study.

475

476 **Conclusions**

477 We prepared a classification system of a broad vegetation unit, semi-natural mesic and wet

478 grasslands of Poland, relying on plot-based numerical classification of community-weighted

means of LHS and vegetative plant traits. The classification mirrors differences in management, moisture, as well as types dominated by functionally unique species (*Scirpus sylvaticus*, *Arrhenatherum elatius*, *Juncus* spp.). Among all traits, clonal index had the strongest influence on the classification. Although, the matching between the trait-based and the syntaxonomical classification was closer than the randomized references applied here, it varied across vegetation types. Syntaxa with high dominance of functionally unique species, typically occurring under more stressed environmental conditions or specific site-history, appeared distinct also in the trait-based classification. In contrast, syntaxa with typically more balanced dominance structure and higher functional overlap between species did not separate well in the trait-based classification. Despite some discrepancies with the traditional species-based classification approach, functional trait-based classification provides biologically interpretable clusters. It must be, however, noted that our classification was performed on a type of vegetation highly dependent on management type and intensity. Classification of less disturbed vegetation types may bring sharper delimitation of vegetation units and different importance of individual traits.

Author contributions

A.L. set the idea, carried out data analysis and led the writing. G.S. and Z.K. prepared the data and added interpretation to the results. Z.B.D. gave suggestions on the methodology. All Authors critically revised the manuscript.

Data accessibility

Vegetation plot data are accessible from the Polish Vegetation Database (Kącki & Śliwiński, 2012; GIVD identifier: EU-PL-001), trait data are available from the LEDA (Kleyer et al., 2008) and Clo-Pla (Klimešová et al., 2017) databases.

509

References

Blagotić, A., & Daróczi, G. (2015). Rapport. R package version 1.0. Retrieved from https://cran.r-project.org/web/packages/rapport/rapport.pdf

Bolnick, D. I., Amarasekare, P., Araújo, M.S., Bürger, R., Levine, J. M., Novak, M., … Vasseur, D. A. (2011). Why intraspecific trait variation matters in community ecology. Trends in Ecology and Evolution, 26(4), 183–192. https://doi.org/10.1016/j.tree.2011.01.009

Borer, E. T., Seabloom, E. W., Gruner, D. S., Harpole, W.S., Hillebrand, H., Lind, E. M., … Yang, L. H. (2014). Herbivores and nutrients control grassland plant diversity via light limitation. Nature, 508, 517–520. https://doi.org/10.1038/nature13144

Botta-Dukát, Z. (2005). Rao's quadratic entropy as a measure of functional diversity based on multiple traits. Journal of Vegetation Science, 16(5), 533–540. https://doi.org/10.1111/j.1654-1103.2005.tb02393.x

Botta-Dukát, Z. (2018). Cautionary note on calculating standardized effect size (SES) in randomization test. Community Ecology, 19(1), 77–83. https://doi.org/10.1556/168.2018.19.1.8

Carlow, P. (1987). Towards a definition of functional ecology. Functional Ecology, 1(1), 57–61. https://doi.org/10.2307/2389358

Cornelissen, J. H. C., Lavorel, S., Garnier, E., Díaz, S., Buchmann, N., Gurvich, D. E., … Poorter, H. (2003). A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. Australian Journal of Botany, 51(4), 335–380. https://doi.org/10.1071/BT02124

De Bello, F., Lavorel, S., Lavergne, S., Albert, C. H., Boulangeat, I., Mazel, F., & Thuiller, W. (2013). Hierarchical effects of environmental filters on the functional structure of plant communities: a case study in the French Alps. Ecography, 36(3), 393–402. https://doi.org/10.1111/j.1600-0587.2012.07438.x

De Bello, F., Lepš, J., Lavorel, S., & Moretti, M. (2007). Importance of species abundance for assessment of trait composition: an example based on pollinator communities. Community Ecology, 8(2), 163–170. https://doi.org/10.1556/ComEc.8.2007.2.3

538 De Cáceres, M., Chytrý, M., Agrillo, E., Attorre, F., Botta-Dukát, Z., Capelo, J., … Wiser, S.
539 K. (2015). A comparative framework for broad-scale plot-based vegetation classification.
540 Applied Vegetation Science, 18(4), 543–560. https://doi.org/10.1111/avsc.12179

541 Dengler, J., Jansen, F., Glöckler, F., Peet, R. K., De Cáceres, M., Chytrý, M., … Spencer, N.
542 (2011). The Global Index of Vegetation-Plot Databases (GIVD): a new resource for
543 vegetation science. Journal of Vegetation Science, 22(4), 582–597.
544 https://doi.org/10.1111/j.1654-1103.2011.01265.x

545 Dengler, J., Solvita, R., Steffen, B., Bruun, H.H., Diekmann, M., Klaus, D., Dolnik, C.,
546 Dupré, C., Golub, V.B., Grytnes, J.-A., Helm, A., Ingerpuu, N., Löbel, S., Pärtel, M.,
547 Rašomavičius, V., Tyler, G., Znamenskiy, S.R., & Zobel, M. (2006). Working Group on Dry
548 Grasslands in the Nordic and Baltic Region - Outline of the Project and First Results for the
549 Class Festuco-Brometea. Annali di Botanica nuova serie, 6, 1–28.

550 Díaz, S., & Cabido, M. (2001). Vive la différence: plant functional diversity matters to
551 ecosystem processes. Trends in Ecology and Evolution, 16(11), 646–655.
552 https://doi.org/10.1016/S0169-5347(01)02283-2

553 Díaz, S., Hodgson, J., Thompson, K., Cabido, M., Cornelissen, J., Jalili, A., … Zak, M.
554 (2004). The plant traits that drive ecosystems: evidence from three continents. Journal of
555 Vegetation Science, 15(3), 295–304. https://doi.org/10.1111/j.1654-1103.2004.tb02266.x

556 Duarte, L. D., Debastiani, V. J., Carlucci, M. B., & Diniz-Filho, J. A. (2018). Analyzing
557 community-weighted trait means across environmental gradients: should phylogeny stay or
558 should it go? Ecology, 99(2), 385–398. https://doi.org/10.1002/ecy.2081

559 E.-Vojtkó, A., Freitag, M., Bricca, A., Martello, F., Compañ, J. M., Küttim, M., …
560 Götzenberger, L. (2017). Clonal vs leaf-height-seed (LHS) traits: which are filtered more
561 strongly across habitats? Folia Geobotanica, 52(3-4), 269–281.
562 https://doi.org/10.1007/s12224-017-9292-1

563 Ellenberg, H., Weber, H. E., Düll, R., Wirth, V., Werner, W., & Paulißen, D. (1992).
564 Zeigerwerte von Pflanzen in Mitteleuropa. Scripta Geobotanica, 18, 3–258.

565 Euro+Med (2018, September 27). Euro+Med PlantBase – the information resource for Euro-
566 Mediterranean plant diversity. Retrieved from http://ww2.bgbm.org/EuroPlusMed/query.asp

567   Garnier, E., Cortez, J., Billès, G., Navas, M., Roumet, C., Debussche, M., … Toussaint, J.

568   (2004). Plant functional markers capture ecosystem properties during secondary. Ecology,

569   85(9), 2630–2637. https://doi.org/10.1890/03-0799

570   Grime, J. P. (1998). Benefits of plant diversity to ecosystems: immediate, filter and founder

571   effects. Journal of Ecology, 86(6), 902–910. https://doi.org/10.1046/j.1365-

572   2745.1998.00306.x

573   Hawkins, B. A., Leroy, B. , Rodríguez, M. Á., Singer, A., Vilela, B., Villalobos, F. , … &

574   Zelený, D. (2017). Structural bias in aggregated species-level variables driven by repeated

575   species co-occurrences: a pervasive problem in community and assemblage data. Journal of

576   Biogeography, 44(6), 1199–1211. https://doi.org/10.1111/jbi.12953

577   Hérault, B., & Honnay, O. (2007). Using life-history traits to achieve a functional

578   classification of habitats. Applied Vegetation Science, 10(1), 73–80.

579   https://doi.org/10.1111/j.1654-109X.2007.tb00505.x

580   Herben, T., Šerá, B., & Klimešová, J. (2015). Clonal growth and sexual reproduction:

581   tradeoffs and environmental constraints. Oikos, 124(4), 469–476.

582   https://doi.org/10.1111/oik.01692

583   Hooper, D. U., Chapin, F. S., Ewel, J. J., Hector, A., Inchausti, P., Lavorel, S., … Wardle, D.

584   A. (2005). Effects of biodiversity on ecosystem functioning: A consensus of current.

585   Ecological Monographs, 75(1), 3–35. https://doi.org/10.1890/04-0922

586   Hothorn, T., Hornik, K., van de Wiel, M. A., & Zeileis, A. (2006). A Lego System for

587   Conditional Inference. The American Statistician, 60(3), 257–263.

588   https://doi.org/10.1198/000313006X118430

589   Hothorn, T., Hornik, K., van de Wiel, M.A., Zeileis, A. (2008). Implementing a Class of

590   Permutation Tests: The coin Package. Journal of Statistical Software, 28(8), 1–23.

591   https://doi.org/10.18637/jss.v028.i08

592   Johansson, V. A., Cousins, S. A. O., & Eriksson, O. (2011). Remnant populations and plant

593   functional traits in abandoned semi-natural grasslands. Folia Geobotanica, 46(2-3), 165–179.

594   https://doi.org/10.1007/s12224-010-9071-8

595   Kącki, Z., & Śliwiński, M. (2012). The Polish Vegetation Database: structure, resources and

596   development. Acta Societatis Botanicorum Poloniae, 81(2), 75–79.

597   https://doi.org/10.5586/asbp.2012.014

598   Kącki, Z., Czarniecka, M., & Swacha, G. (2013). Statistical determination of diagnostic,

599   constant and dominant species of the higher vegetation units of Poland. Monographiae

600   Botanicae, 103, 1–267. https://doi.org/10.5586/mb.2013.001

601   Kattge, J., Díaz, S., Lavorel, S., Prentice, I.C., Leadley, P., Bönisch, G., … Wirth, C. (2011).

602   TRY – a global database of plant traits. Global Change Biology, 17(9), 2905–2935.

603   https://doi.org/10.1111/j.1365-2486.2011.02451.x

604   Keddy, P.A. (1992). Assembly and response rules–2 goals for predictive community ecology.

605   Journal of Vegetation Science,3(2), 157–164. https://doi.org/10.2307/3235676

606   Kleyer, M. , Dray, S. , Bello, F. , Lepš, J. , Pakeman, R. J., Strauss, B. , … Lavorel, S. (2012).

607   Assessing species and community functional responses to environmental gradients: which

608   multivariate methods? Journal of Vegetation Science, 23: 805–821.

609   https://doi.org/10.1111/j.1654-1103.2012.01402.x

610   Kleyer, M., Bekker, R. M., Knevel, I. C., Bakker, J. P, Thompson, K., Sonnenschein, M., …

611   Peco, B. (2008). The Leda Traitbase: a database of life-history traits of Northwest European

612   flora. Journal of Ecology, 96(6), 1266–1274. https://doi.org/10.1111/j.1365-

613   2745.2008.01430.x

614   Klimešová J., Danihelka J., Chrtek J., de Bello F., & Herben T. (2017). CLO-PLA: a database

615   of clonal and bud bank traits of Central European flora. Ecology, 98(4), 1179.

616   http://dx.doi.org/10.1002/ecy.1745

617   Klimešová, J., Tackenberg, O., & Herben, T. (2016). Herbs are different: clonal and bud bank

618   traits can matter more than leaf–height–seed traits. New Phytologist, 210(1), 13–17.

619   https://doi.org/10.1111/nph.13788

620   Knollová, I. , Chytrý, M. , Tichý, L., & Hájek, O. (2005). Stratified resampling of

621   phytosociological databases: some strategies for obtaining more representative data sets for

622   classification studies. Journal of Vegetation Science, 16(4), 479–486.

623   https://doi.org/10.1111/j.1654-1103.2005.tb02388.x

624 Kočí, M., Chytrý, M., & Tichý, L. (2003). Formalized reproduction of an expert-based

625 phytosociological classification: A case study of subalpine tall-forb vegetation. Journal of

626 Vegetation Science, 14, 601–610.

627 Laliberté, E., Legendre, P., & Shipley, B. (2014). FD: measuring functional diversity from

628 multiple traits, and other tools for functional ecology. R package version 1.0-12. Retrieved

629 from https://cran.r-project.org/web/packages/FD/FD.pdf

630 Landucci, F., Tichý, L., Šumberová, K., & Chytrý, M. (2015). Formalized classification of

631 species-poor vegetation: a proposal of a consistent protocol for aquatic vegetation. Journal of

632 Vegetation Science, 26, 791–803.

633 Lavorel, S., & Garnier, E. (2002). Predicting changes in community composition and

634 ecosystem functioning from plant traits: revisiting the Holy Grail. Functional Ecology, 16,

635 545–556. https://doi.org/10.1046/j.1365-2435.2002.00664.x

636 Lengyel, A., Roberts, D.W. & Botta-Dukát, Z. (2019). Comparison of silhouette-based

637 reallocation methods for vegetation classification. Retrieved from

638 https://www.biorxiv.org/content/10.1101/630384v2

639 Lengyel, A., Chytrý, M., & Tichý, L. (2011). Heterogeneity-constrained random resampling

640 of phytosociological databases. Journal of Vegetation Science, 22(1), 175–183.

641 https://doi.org/10.1111/j.1654-1103.2010.01225.x

642 Lhotsky, B., Kovács, B., Ónodi, G., Csecserits, A., Rédei, T., Lengyel, A., … Botta-Dukát, Z.

643 (2016). Changes in assembly rules along a stress gradient from open dry grasslands to

644 wetlands. Journal of Ecology, 104(2), 507–517. https://doi.org/10.1111/1365-2745.12532

645 Lososová, Z. , Šmarda, P. , Chytrý, M. , Purschke, O. , Pyšek, P. , Sádlo, J. … Winter, M.

646 (2015). Phylogenetic structure of plant species pools reflects habitat age on the geological

647 time scale. Journal of Vegetation Science, 26(6), 1080–1089. https://doi:10.1111/jvs.12308

648 Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., & Hornik, K. (2017). Cluster: Cluster

649 Analysis Basics and Extensions. R package version 2.0.6. Retrieved from https://cran.r-

650 project.org/web/packages/cluster/cluster.pdf

651 McGill, B. J., Enquist, B. J., Weiher, E., & Westoby, M. (2006). Rebuilding community

652 ecology from functional traits. Trends in Ecology and Evolution, 21(4), 178–185.

653 https://doi.org/10.1016/j.tree.2006.02.002

654 Moles, A. T., & Westoby, M. (2004). Seedling survival and seed size: a synthesis of the

655 literature. Journal of Ecology, 92(3), 372–383. https://doi.org/10.1111/j.0022-

656 0477.2004.00884.x

657 Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P.. McGlinn, D., … Wagner,

658 H. (2018). Vegan: Community Ecology Package. R package version 2.4-6. Retrieved from

659 https://CRAN.R-project.org/package=vegan

660 R Core Team. (2017). R: A language and environment for statistical computing. R Foundation

661 for Statistical Computing, Vienna, Austria. Retrieved from https://www.R-project.org/

662 Ricotta, C. , Bello, F. , Moretti, M. , Caccianiga, M., Cerabolini, B. E., Pavoine, S., & Peres-

663 Neto, P. (2016). Measuring the functional redundancy of biological communities: a

664 quantitative guide. Methods in Ecology & Evolution, 7(11), 1386–1395.

665 http://dx.doi.org/10.1111/2041-210X.12604

666 Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of

667 cluster analysis. Computational and Applied Mathematics, 20, 53–65.

668 https://doi.org/10.1016/0377-0427(87)90125-7

669 Rusch, G. M., Wilmann, B., Klimešová, J., & Evju, M. (2011). Do clonal and bud bank traits

670 vary in correspondence with soil properties and resource acquisition strategies? Patterns in

671 alpine communities in the Scandian Mountains. Folia Geobotanica, 46(2-3), 237–254.

672 https://doi.org/10.1007/s12224-010-9072-7

673 Schrodt, F., Kattge, J., Shan, H., Fazayeli, F., Joswig, J., Banerjee, A., … Reich, P. B. (2015).

674 Gap-filling in trait databases. Global Ecology and Biogeography, 24(9), 1510–1521.

675 https://doi.org/10.5194/bg-15-2601-2018

676 Siefert, A., Violle, C., Chalmandrier, L., Albert, C. H., Taudiere, A., Fajardo, A., … Chase, J.

677 (2015). A global meta-analysis of the relative extent of intraspecific trait variation in plant

678 communities. Ecology Letters, 18(12), 1406–1419. https://doi.org/10.1111/ele.12508

679 Tichý, L. (2002). JUICE, software for vegetation classification. Journal of Vegetation

680 Science, 13(3), 451-453. doi:10.1111/j.1654-1103.2002.tb02069.x

681 Tilman, D., Knops, J., Wedin, D., Reich, P., Ritchie, M., & Siemann, E. (1997). The influence

682 of functional diversity and composition on ecosystem processes. Science, 277(5330), 1300–

683 1302. https://doi.org/10.1126/science.277.5330.1300

684 Vellend, M. (2010). Conceptual synthesis in community ecology. The Quarterly Review of
685 Biology, 85(2), 183–206. https://doi.org/10.1086/652373

686 Violle, C., Enquist, B. J., McGill, B. J., Jiang, L., Albert, C. H., Hulshof, C., … Messier, J.
687 (2012). The return of the variance: intraspecific variability in community ecology. Trends in
688 Ecology and Evolution, 27(4), 244–252. https://doi.org/10.1016/j.tree.2011.11.014

689 Westoby, M. (1998). A leaf-height-seed (LHS) plant ecology strategy scheme. Plant and Soil,
690 199(2), 213–227. https://doi.org/10.1023/A:100432722

691 Westoby, M., Falster, D. S., Moles, A. T., Vesk, P. A., & Wright, I. J. (2002). Plant ecological
692 strategies: some leading dimensions of variation between species. Annual Review of Ecology
693 and Systematics, 33, 125–159. https://doi.org/10.1146/annurev.ecolsys.33.010802.150452

694 Wilson, S. D., & Tilman, D. (1993). Plant competition and resource availability in response to
695 disturbance and fertilization. Ecology, 74(2), 599–611. https://doi.org/10.2307/1939319

696 Wiser, S. K. & De Cáceres, M. (2013). Updating vegetation classifications: an example with
697 New Zealand's woody vegetation. Journal of Vegetation Science, 24: 80–93.
698 https://doi.org/10.1111/j.1654-1103.2012.01450.x

699 Wright I. J., Reich P. B., Westoby, M., Ackerly D. D., Baruch Z., Bongers F., … Villar R.
700 (2004). The worldwide leaf economics spectrum. Nature, 428, 821–827.
701 https://doi.org/10.1038/nature02403

702 Zelený, D. (2018). Which results of the standard test for community weighted mean approach
703 are too optimistic? Journal of Vegetation Science, 29, 953–966.
704 https://doi.org/10.1111/jvs.12688

705 Zelený, D., & Schaffers, A. P. (2012). Too good to be true: pitfalls of using mean Ellenberg
706 indicator values in vegetation analyses. Journal of Vegetation Science, 23(3), 419–431.
707 https://doi.org/10.1111/j.1654-1103.2011.01366.x

708

709 **Supporting information**

710 **Appendix S1:** Evaluation of the hierarchical classification using six validity indices.

711 **Appendix S2:** Shortened synoptic tables of the syntaxonomic orders and alliances.

712 **Appendix S3:** Shortened synoptic table and textual description of the discussed clusters.

713   **Appendix S4:** Ellenberg indicator values of the discussed clusters.

714   **Appendix S5:** Geographical distribution of the discussed clusters.

715   **Figure 1.** Dendrogram of the upper 12 clusters in the hierarchy
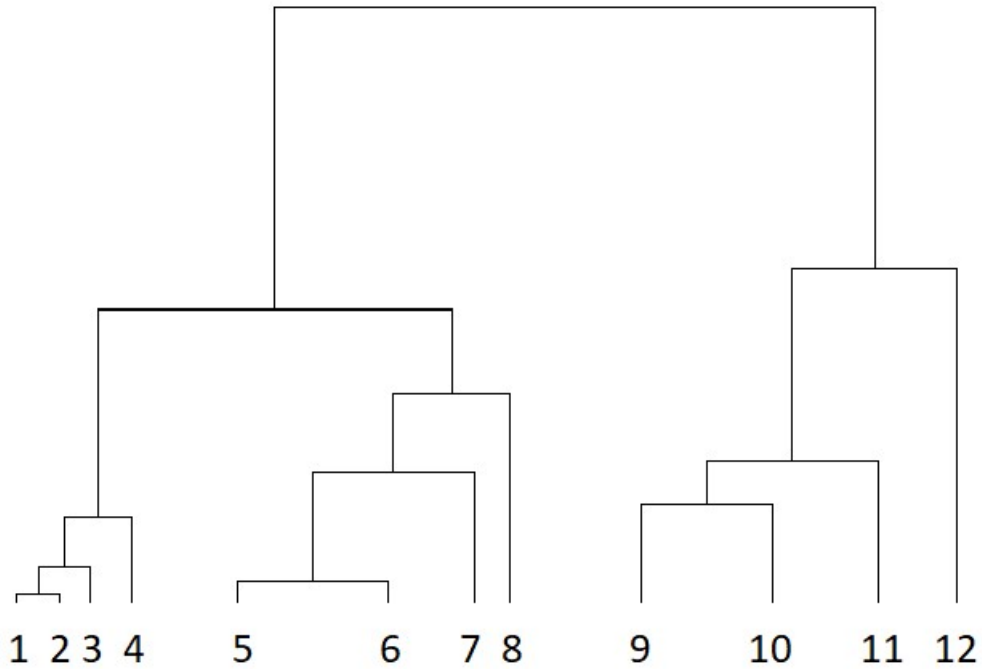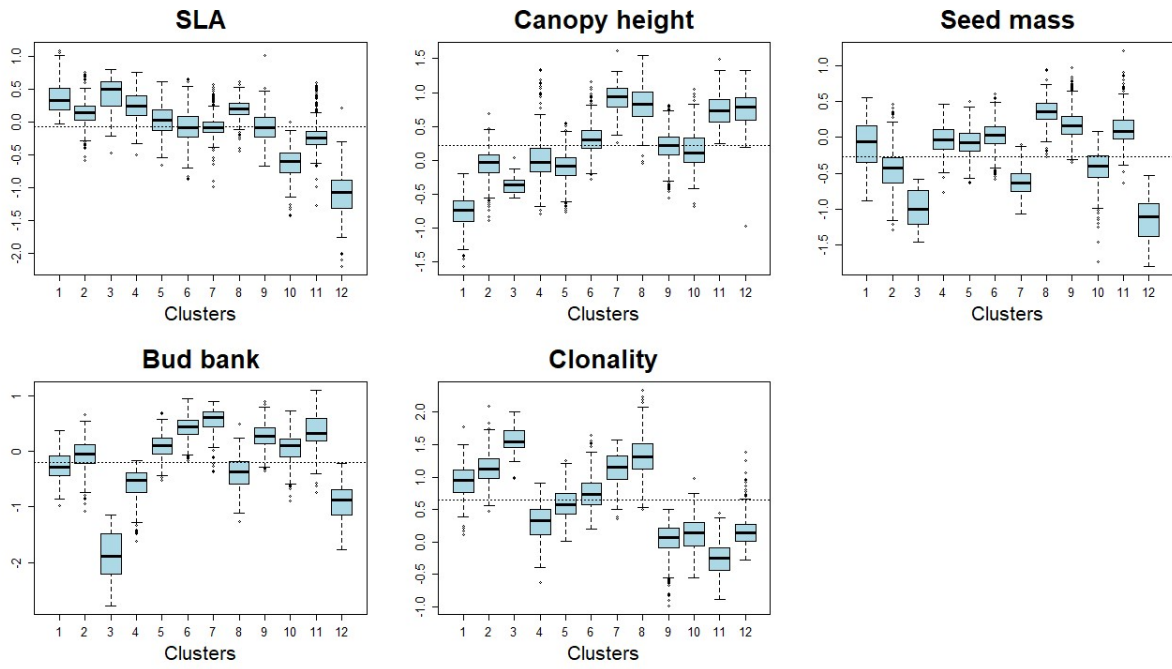

716

717

718

719   **Figure 2.** Boxes-and-whiskers plots comparing community-weighted means of traits across

720   the twelve clusters. Boxes show upper and lower quartiles; whiskers show minimum and

721   maximum values; circles show outliers

Figure 3. Boxes-and-whiskers plots comparing Rao's functional diversity and functional redundancy (Ricotta et al., 2016) across the twelve clusters. Boxes show upper and lower quartiles; whiskers show minimum and maximum values; circles show outliers

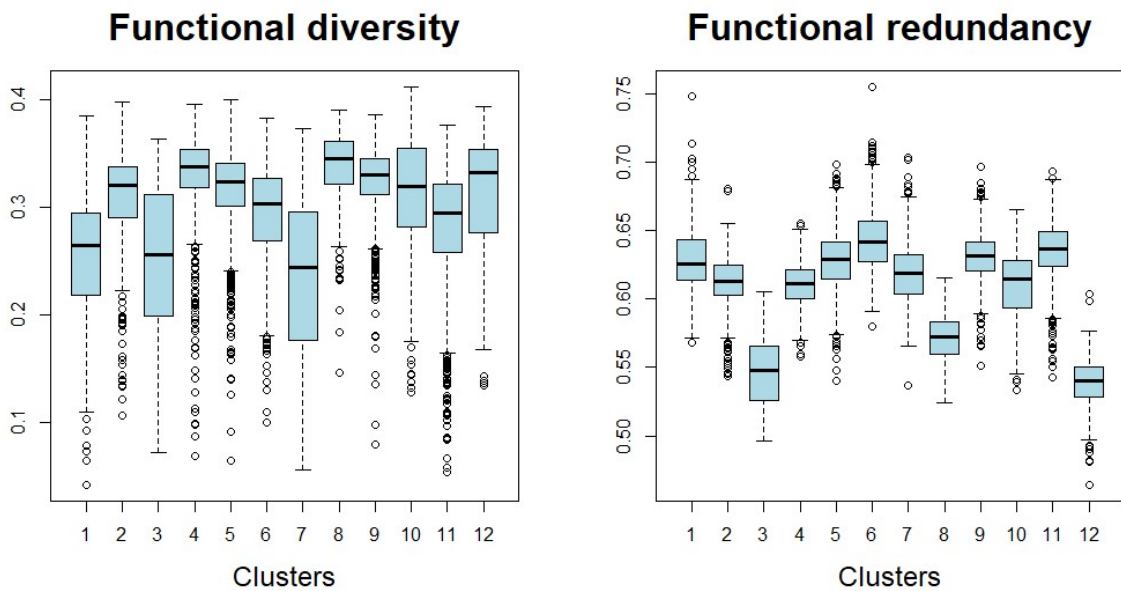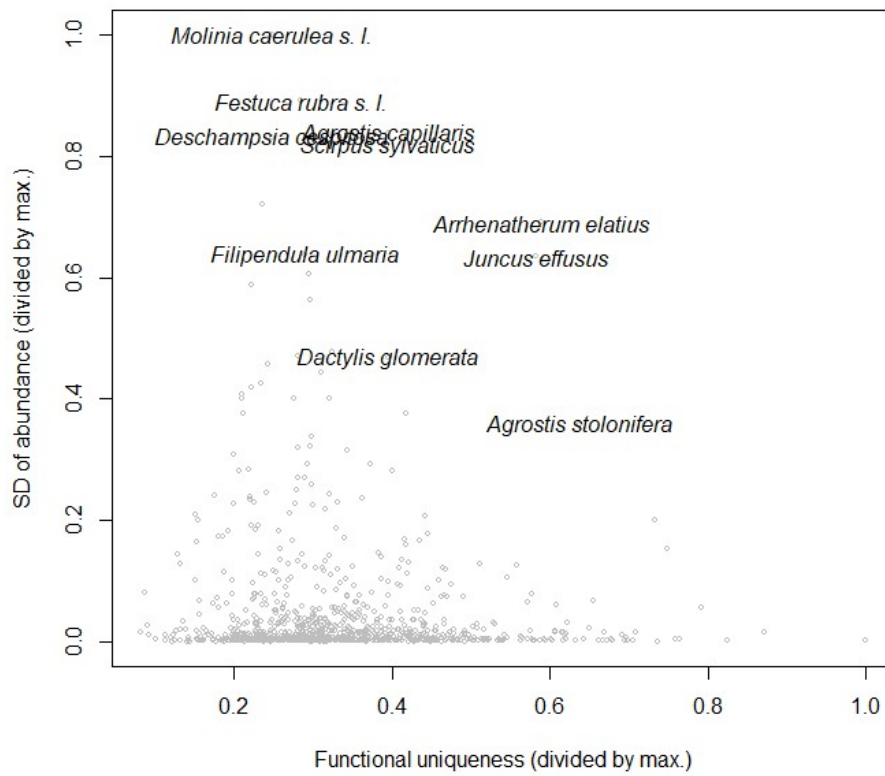730 **Figure 4.** Distribution of species in dimensions of functional uniqueness and SD of

731 abundance (both divided by maximum). Only names of the ten species with the highest

732 influence index values are shown

733

734

Table 1. Cross-tabulation of the trait-based and the syntaxonomical classification

	CLASSES	ORDERS	ALLIANCES	Clusters of the trait-based classification											
				1	2	3	4	5	6	7	8	9	10	11	12
Syntaxonomical classification	Molinio-Arrhenatheretea	Arrhenatheretalia	Arrhenatherion	1	70	0	1	120	200	4	251	99	0	8	0
			Cynosurion	83	30	0	127	223	25	0	0	99	2	0	0
			Poion alpinae	0	0	0	1	0	0	0	0	1	0	0	0
			Polygono-Trisetion	0	11	0	0	34	24	0	0	24	1	2	0
			Unidentified at alliance level	9	287	0	39	446	452	1	11	190	7	4	0
		Molinietalia	Calthion	0	7	0	41	165	193	203	4	545	92	214	12
			Cnidion	0	0	0	7	35	55	1	0	49	28	4	14
			Juncion effusi	0	2	0	2	0	0	0	0	0	36	0	111
			Molinion	0	0	0	2	23	69	0	0	240	38	320	0
			Unidentified at alliance level	0	9	0	4	112	109	20	12	201	196	113	21
		Potentillo-Polygonetalia	Potentillion anserinae	1	11	11	0	1	0	0	0	0	2	0	0
			Unidentified at alliance level	16	13	23	1	7	13	0	1	1	1	0	0
		Unidentified at order level		14	41	0	13	153	260	11	10	33	7	2	2
	Polygono-Poetea	Polygono-Poetalia	Unidentified at alliance level	79	7	0	36	10	0	0	0	0	0	0	1
	Dubious assignment			5	0	0	0	1	0	1	0	0	1	0	0

Table 2. Standardized test statistic of Wilcoxon tests of CWM-s between clusters to be merged in each fusion level of the hierarchical classification. Significance levels after Bonferroni adjustment (two-tailed tests): *** $p<0.001$, ** $p<0.01$, * $p<0.05$, ns = not significant

No. Clusters	SLA	Canopy height	Seed mass	Bud bank	Clonality
2	34.65***	-19.67***	-17.82***	-9.48***	66.56***
3	20.62***	-15.13***	20.93***	20.81***	-4.74***
4	20.53***	-29.81***	-19.41***	-38.04***	-6.83***
5	-15.69***	-22.37***	-22.22***	25.05***	-20.93***
6	3.37***	-36.72***	-12.05***	-17.44***	25.98***
7	2.22*	-25.42***	23.24***	-15.14***	-17.91***
8	28.88***	2.58**	31.22***	10.05***	-0.74[ns]
9	3.19**	-15.86***	-15.15***	10.13***	24.96***

10	-5.5***	1.43^{ns}	8.52***	9.56***	-8.44***
11	18.81***	-17.4***	-0.79^{ns}	-30.07***	15.83***
12	8.47***	-18.19***	17.84***	-0.63^{ns}	-2.76**

742

743 **Table 3.** Species with the highest influence index

Species	Influence index
Arrhenatherum elatius	0.407
Juncus effusus	0.369
Agrostis capillaris	0.331
Scirpus sylvaticus	0.322
Festuca rubra s. l.	0.253
Molinia caerulea s. l.	0.248
Agrostis stolonifera	0.229
Deschampsia cespitosa	0.207
Dactylis glomerata	0.186
Filipendula ulmaria	0.184

744