VEGETATION CLASSIFICATION AND SURVEY

International Association for Vegetation Science (IAVS)

Vegetation Classification and Survey 6: 57–67
doi: 10.3897/VCS.134800

VCS METHODS

Mean ecological indicator values: use EIVE but no cover-weighting

Grzegorz Ostrowski[1], Severin Aicher[2], Agnieszka Mankiewicz[3], Olha Chusova[2,4], Iwona Dembicz[5], Stefan Widmer[2], Jürgen Dengler[2,6]

1 *Institute of Evolutionary Biology, Biological and Chemical Research Centre, Faculty of Biology, University of Warsaw, Warsaw, Poland*
2 *Vegetation Ecology Research Group, Institute of Natural Resource Sciences (IUNR), Zurich University of Applied Sciences (ZHAW), Wädenswil, Switzerland*
3 *College of Inter-faculty Individual Studies in Mathematics and Natural Sciences (MISMaP), University of Warsaw, Warsaw, Poland*
4 *Department of Geobotany and Ecology, M.G. Kholodny Institute of Botany of the National Academy of Sciences of Ukraine, Kyiv, Ukraine*
5 *Institute of Environmental Biology, Faculty of Biology, University of Warsaw, Warsaw, Poland*
6 *Plant Ecology, Bayreuth Center for Ecology and Environmental Research (BayCEER), University of Bayreuth, Bayreuth, Germany*

Corresponding author: Jürgen Dengler (juergen.dengler@zhaw.ch)

Academic editor: Jorge Capelo ♦ **Linguistic editor:** Michael Glaser
Received 15 August 2024 ♦ **Accepted** 20 February 2025 ♦ **Published** 26 March 2025

Abstract

Aims: To test the predictive power of mean ecological indicator values (EIVs) based on different EIV systems and weighting approaches. **Study area:** Preda in Grisons, Switzerland. **Methods:** We used three regional datasets of vegetation plots accompanied with measured soil pH values or mean annual near soil air temperature. We calculated mean EIVs for each plot with four EIV systems that cover the region, namely “Ellenberg”, “Landolt”, “Tichý” and the Ecological Indicator Values for Europe (EIVE), combined with four weighting approaches (unweighted, cover-weighted, square-root cover-weighted, inverse niche width weighted). We correlated the mean EIVs of each combination with the measured environmental variables and compared the mean Pearson r values. **Results:** No cover-weighting (0.78) was slightly better than square-root cover-weighting (0.75) and clearly better than full cover-weighting (0.68). In the two EIV systems providing a niche width measure (EIVE and “Landolt”), inverse niche weighting gave similar results than no weighting. Mean EIVE values (0.76) had significantly stronger correlations than “Tichý” (0.73), while the differences to “Ellenberg” (0.75) and “Landolt” (0.71) were not significant. **Conclusions:** The results suggest that even within the definition areas of two long-established EIV systems (“Ellenberg”, “Landolt”), EIVE gives at least as good predictions as these and significantly better than “Tichý”. Likewise, any type of cover-weighting reduces the predictive power of mean EIVs. Both findings could be the consequence of the statistical principle “wisdom of the crowd”, according to which the average estimate of several sources (be it regional scientists or plant species) is usually better than the estimate of one or few experts. Accordingly, EIVE currently is the best choice for mean EIVs in Europe, and no cover-weighting should be applied. We recommend that similar studies should be undertaken in other regions and for other niche dimensions.

Syntaxonomic reference: Mucina et al. (2016).

Abbreviations: EIV = ecological indicator value; EIVE = Ecological Indicator Values for Europe 1.0; R = ecological indicator value for soil reaction; T = ecological indicator value for temperature.

Keywords

bioindication, calibration, cover weighting, Ecological Indicator Values for Europe (EIVE), Ellenberg indicator value, inverse niche width weighting, Landolt indicator value, methods comparison, pH, Switzerland, temperature, wisdom of the crowd

Introduction

Ecological indicator values (EIVs) of plants are a widely used tool in applied and fundamental vegetation ecology in Europe (Diekmann 2003). The first such EIV systems have been independently developed by Ramensky et al. (1956) for the European part of the former Soviet Union and Ellenberg (1974) for Central Europe. The principle of such EIV systems is that the species of a certain region are placed on ordinal scales by expert knowledge representing the main niche dimensions (ecological factors) of plants, such as soil moisture, soil nutrients or temperature. In practice, the indicator values of one niche dimension are averaged for all species occurring in one site (i.e. vegetation plot). The average value is then used as a proxy for an ecological factor. The correlation of mean EIVs and measured ecological factors has been demonstrated many times to be good, with monotonous, but not always linear relationships (reviews by Ellenberg et al. 1991; Diekmann 2003). Based on these properties, mean EIVs have been used among others to quantify vegetation change over time and its drivers (Diekmann et al. 2019; Scherrer et al. 2024), compare different management methods (Chytrý et al. 2009; Reutimann et al. 2023), characterise vegetation types comparatively (Chytrý 2007; Vassilev et al. 2024), calculate regional species pool of plant communities (Pärtel et al. 1996) or for palaeoecological reconstructions (Blaus et al. 2020). Following the mentioned pioneering works, about 30 additional EIV systems were developed for different parts of Europe but remained geographically restricted and largely incompatible due to different scales (see overview in Dengler et al. 2023).

Only in the year 2023, two international author teams have overcome this limitation by combining several or even all available EIV systems at that time into pan-European systems, the "Ellenberg-type indicator values" by Tichý et al. (2023) and the "Ecological Indicator Values for Europe" (EIVE; Dengler et al. 2023). The first tried to be as close as possible to Ellenberg et al. (1991), including retaining the original scaling, which is different for different niche dimensions, while the second approach created a new uniform scaling of niche position from 0 to 10 and for the first time also a comprehensive assessment of niche widths for all taxon × indicator combinations on a continuous scale, too. The availability of these European systems in the short time since their publication has already prompted numerous studies to use them at the continental scale (e.g. De Pauw et al. 2024) or in regions without regional systems (e.g. Reczyńska et al. 2024; Vassilev et al. 2024). However, what is missing up to date are (a) a systematic calibration of the two new systems against measured environmental variables and (b) a comparative test of the performance of these two systems among each other and compared to regional systems.

The two widespread approaches to indicate site conditions are to calculate (i) unweighted or (ii) cover-weighted means of the EIVs of all occurring species for the niche dimension of interest (Ellenberg et al. 1991; Diekmann 2003). While Ellenberg et al. (1991) favour unweighted means (without clear reasoning), Diekmann (2003) suggests that most plant ecologists are using weighted means, with differences between the two approaches being negligible in species rich communities, while in species poor communities the weighted average is more informative. Intuitively, one could interpret higher cover of the same species in a way that the site conditions are more favourable for this species, justifying giving the indicator values more weight in the calculation of the mean EIV. On the other hand, the cover of species not only depends on the favourability of the habitat, but also species-specific traits like size and growth form. For example, it is hard to imagine a very low cover of *Fagus sylvatica* of < 1%, while a very high cover of > 50% probably never can be found for a species like *Linum catharticum*; nevertheless, its presence can have high indicative value. These considerations have prompted some researchers to adopt an intermediate solution between no weighting and full cover weighting, i.e. using the square root of cover for weighting (Reutimann et al. 2023). Despite knowing the best of these three approaches would be very beneficial, there are hardly any empirical studies on this topic. We only know of Hájek et al. (2020) who found that cover-weighting in one dataset outperformed non-weighting, while it was the other way round in another dataset, and of Tölgyesi et al. (2014) who found non-weighting to be superior to cover-weighting. Hájek et al. (2020) also proposed a fourth averaging approach, weighting by inverse niche width. They found that this averaging approach alone or combined with cover-weighting outperformed other approaches. Inverse niche width weighting requires a measure for niche width, called "tolerance" in Hájek et al. (2020). While these authors published "tolerance" values for moisture in wetland plants, some of the regional EIV systems provide niche width information across all indicators, e.g. Landolt et al. (2010) on a three-step ordinal scale and Didukh (2011) instead of a niche position always a minimum and maximum. However, such niche width information has only rarely been used for the calculation of mean EIVs (e.g. Gillet et al. 2016). With the publication of EIVE (Dengler et al. 2023), for the first time all included species have for five indicators both a continuous value for niche position and for niche width. While this information in itself is valuable for ecologists, it remains to be tested how useful it is when calculating mean EIVE values.

To provide guidelines for the optimal use of ecological indicator value systems of plants, we made use of three sets of vegetation plots sampled in Preda in the Swiss Alps. They cover wide ecological gradients and are combined with *in situ* measured environmental variables (soil pH and mean annual temperature). Using these datasets, we addressed two questions:

- Which of the four weighting approaches (unweighted, cover-weighted, square root-cover-weighted, inverse-niche-width-weighted) provides the best predictions of the actual environmental conditions?

- Which of the four EIV systems that cover the Swiss Alps (Ellenberg et al. 1991; Landolt et al. 2010; Dengler et al. 2023; Tichý et al. 2023) provides the best predictions of the local environmental conditions?

Regarding the weighting approach, we had no *a priori* hypothesis which of the four solutions would yield the best results as there were arguments or examples for any of these. Among the EIV systems, by contrast, we hypothesised that EIVE outperforms the other three systems based on previous findings by Moeys (2020) and Dengler et al. (2023: table 3).

Study area

The study was conducted in Preda, canton of Grisons in Switzerland, within the regional nature park "Parc Ela" (Figure 1). Our sampling included the valley bottom of the Albula valley as well as the northern side valley Val Zavretta and the southern side valley Val Mulix. Our plots were distributed from 46.56° to 46.60° northern latitude, from 9.73° to 9.80° eastern longitude and from 1738 to 2636 m a.s.l. The plots comprise a wide range of different natural and semi-natural habitats (forests, grasslands, heathlands, mires, screes) of the subalpine to alpine belts, both on limestone and on siliceous bedrock (see Table 1).

Methods

Vegetation-plot and environmental data

We used three unpublished datasets collected for other purposes, which differed in plot size and measured environmental variables (Table 1). Vegetation was sampled in precisely delimited square plots of 10 m^2 and 1 m^2, respectively (Figure 1, Table 1). Dataset #1 was established to represent the full diversity of natural and semi-natural habitat types in the communal nature reserve and an alluvial plain of national importance around the Sonnenhof Preda. Sampling involved one transect across the alluvial plain (20 plots) and 44 additional, subjectively located plots. Dataset #3 is based on an elevational transect of 13 EDGG Biodiversity Plots (Dengler et al. 2016), with one 100-m^2 plot placed approximately every 100 m of elevation along the hiking path in Val Mulix (mostly acidic). Vegetation was sampled and pH measured in two opposite 10 m^2 corners (Dengler et al. 2016), while the temperature sensor was placed in the centre of the 100 m^2 plot and its data used for both 10 m^2 plots. Due to a technical issue with the vegetation recording app "FlorApp" (Info Flora 2025), we lost four of the plots, leaving 22 for analysis. In dataset #2, the plots of dataset #3 were augmented with additional 12 plots sampled along an elevational transect in Val Zavretta (mostly limestone).

All vascular plants were recorded with the shoot-presence system (Dengler 2008), and their cover estimated in

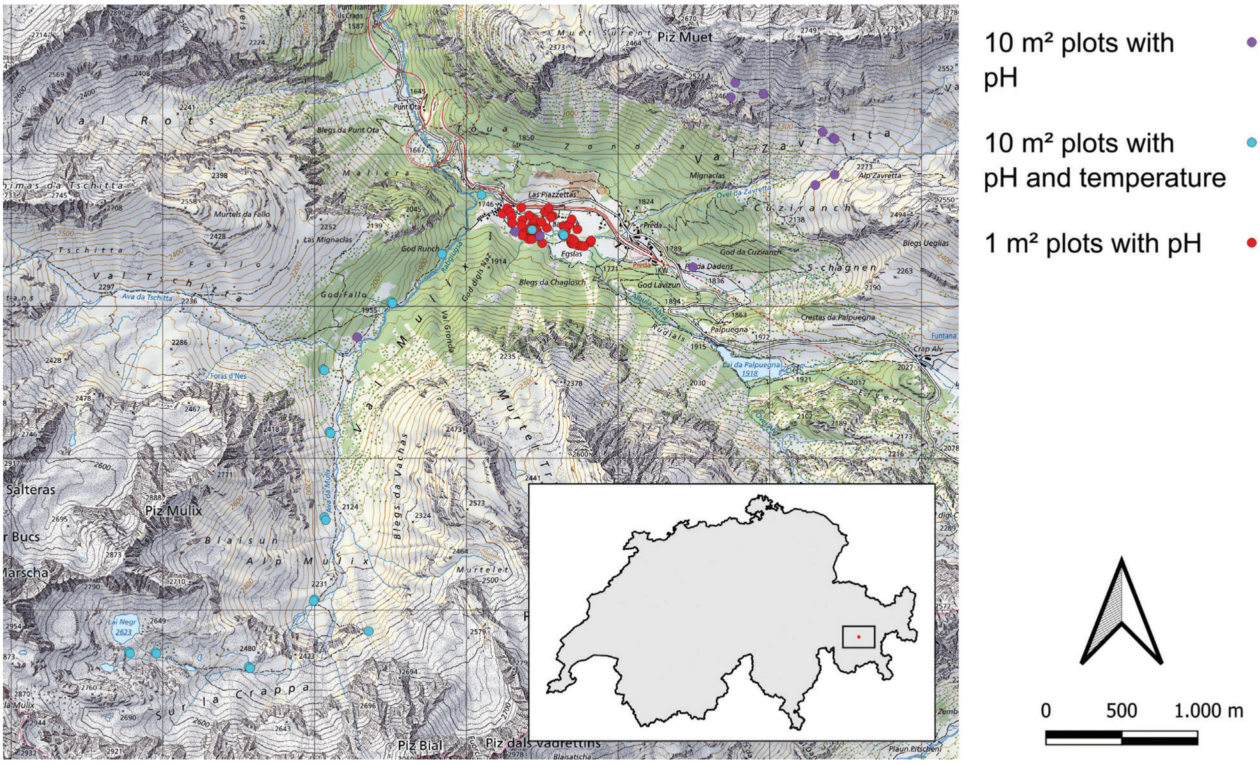

Figure 1. Sampling area around Preda (Parc Ela, Grisons, Switzerland) in the Albula valley as well as its side valleys, Val Mulix to the South and Val Zavretta to the North. The coloured dots indicate the location of the plots of the three datasets used. Note that some points are so close together at this scale that they are not visible as separate entities. Source of the map: swisstopo.ch.

percent (Dengler and Dembicz 2023) separately for tree, shrub and herb layer. If a species occurred in more than one layer, we added the values, reflecting the higher importance of a species when present several times, but the cumulative values for woody species mostly remained low and only once reached 100%.

To measure the pH, mixed soil samples were collected from the top 10 cm of the soil in five randomly selected locations within the 10-m^2 and four locations within the 1-m^2 plots. The soil samples were then air-dried and, once dry, sieved to extract the particle fraction < 2 mm (fine soil). For each plot, 10 g of the sieved soil was mixed with 25 g of the distilled water. Once shaken, the mixture in the test tube was left still for the next hour, after which the soil pH was measured with a pH electrode (HANNA instruments, model HI991300).

To measure the temperature relevant for plants, we installed temperature loggers (iButtons) in the plots in August 2019. These measured the temperature in 30-min intervals both 10 cm below and 10 cm above the soil surface. For our study, we took the average near-surface air temperature for the 12-month period from August 2019 to August 2020.

Plant nomenclature and matching to the EIV systems

Plant nomenclature was harmonised to Euro+Med (2025) with additional aggregates as defined in Euro+Med augmented (Dengler et al. 2023), mapping taxonomic concepts, not names (Jansen and Dengler 2010). We kept all records at the taxonomic resolution they had after the identification, that is, as fine as possible, be it as subspecies, species or aggregate in the taxon view of Euro+Med augmented. Taxa determined with some uncertainty ("cf.") were treated like the respective taxon without cf., based on the assumption that cf. is commonly used when the likelihood that a sample belongs to a certain taxon is high.

In the study we considered the following four EIV systems with the given area of validity:

- Ellenberg et al. (1991) for Central Europe (further "Ellenberg")
- Landolt et al. (2010) for Switzerland and the entire Alps (further "Landolt")
- Dengler et al. (2023) for entire Europe (further EIVE)
- Tichý et al. (2023) for nemoral Europe and Italy (further "Tichý")

The values of the first three systems were taken from Dengler et al. (2023: suppl. materials 2 and 8) where they are harmonized to Euro+Med augmented, while the data from Tichý et al. (2023) were taken from the source, where they are also largely matched to the Euro+Med standard. All cases of non-matching names were manually double-checked and corrected if necessary. In case of non-matches in Ellenberg et al. (1991) and Landolt et al. (2010), we also checked the original printed publications to ensure that our treatment was not biased by missing or erroneous assignments of taxonomic concepts in Dengler et al. (2023). This led to a few corrections, particularly in the case of Landolt et al. (2010), which were reported to the EIVE team to be adjusted in the next release of EIVE.

Statistical analyses

All statistical analyses were performed in R (R Core Team 2024). We calculated mean EIVs with the package 'FD' (Laliberté et al. 2014) in four weighting approaches: cover-weighted, square-root cover-weighted, unweighted and inverse niche width weighted. The latter was possible only for "Landolt" and EIVE. In "Landolt", we replaced the niche width value I with 1 and II with 2 (see also Gillet et al. 2016).

This resulted in 42 combinations of the 3 datasets × 4 EIV systems × 3–4 weighting approaches. For each of these we calculated the Pearson and Spearman correlations with the respective measured environmental variable and a linear regression of the measured environmental variable against the corresponding mean indicator value. The resulting Pearson correlation coefficients (r) were then compared among datasets, EIV systems and weighting approaches using linear mixed effects models. We did this with the function 'lmer' from the package 'lmerTest' (Kuznetsova et al. 2017). We note that the sampling design did not allow to construct a factorial model with EIV system × weighting approach × indicator × plot size because

Table 1. Descriptive statistics of the three datasets used in this study.

Dataset ID	Indicator	Number of plots	Plot size [m^2]	Elevational range [m a.s.l.]	Vegetation classes
#1	pH	64	1	1743–1764	*Elyno-Seslerietea*, *Erico-Pinetea*, *Festuco-Brometea*, *Juncetea trifidi*, *Loiseleurio procumbentis-Vaccinietea*, *Molinio-Arrhenatheretea*, *Montio-Cardaminetea*, *Rhododendro hirsuti-Ericetea carneae*, *Salicetea purpureae*, *Scheuchzerio palustris-Caricetea fuscae*, *Thlaspietea rotundifolii*, *Trifolio-Geranietea sanguinei*, *Vaccinio-Piceetea*
#2	pH	34	10	1738–2636	*Elyno-Seslerietea*, *Erico-Pinetea*, *Juncetea trifidi*, *Loiseleurio procumbentis-Vaccinietea*, *Rhododendro hirsuti-Ericetea carneae*, *Thlaspietea rotundifolii*, *Vaccinio-Piceetea*
#3	Temperature	22	10	1738–2636	*Elyno-Seslerietea*, *Erico-Pinetea*, *Juncetea trifidi*, *Loiseleurio procumbentis-Vaccinietea*, *Vaccinio-Piceetea*

of missing combinations, insufficient sample size and the fact that the three datasets differed in more aspects than just indicator and plot size, e.g. length of the gradient and diversity of habitats included (Table 1). We thus “combined” indicator and plot size in the factor “dataset”.

First, we excluded inverse niche-width weighting (available for only two EIV systems) and ran one model for weighting approach and one for EIV system:

lmer(pearson ~ dataset + (1 | weighting_approach/EIV_system))

lmer(pearson ~ EIV_system + (1 | dataset/weighting_approach))

Second, we compared all four weighting approaches, but restricted to “Landolt” and EIVE:

lmer(pearson ~ weighting_approach + (1 | dataset/EIV_system))

Third, we tested for the effect of the EIV system with a modified approach as we had hypothesised that EIVE would perform better than any of the other three systems (“Tichý”, “Ellenberg”, “Landolt”). Thus, we were not interested in all possible comparisons between the four systems, but conducted pair-wise comparisons of each of the latter three systems with EIVE and report Bonferroni-corrected p-values.

In the first and second case, we applied a posthoc test with the function ’glht’ from the package ‘multcomp’ (Hothorn et al. 2008).

Results

Taxonomic matching and coverage

Disregarding identifications at genus level or higher (which do not have EIV assignments in any of the four EIV systems), we had 371 valid taxa (subspecies, species or aggregates) each for the pH indicator (datasets #1 and #2 combined) and the temperature indicator (dataset #3). In EIVE, 100% of these taxa had definitive indicator values. For the six remaining combinations of indicator and EIV system, the fraction of valid taxa without definitive indicator value varied between 12.1% and 40.4% (Table 2). The fractions were lowest in “Landolt”, intermediate in “Tichý” and highest in “Ellenberg”. Completely missing taxa were not an issue in Tichý, while one and 13 taxa were missing in “Landolt” and “Ellenberg”, respectively (Table 2). Taxa represented in the source EIV system only at an inferior or superior level contributed between 5.7% and 7.3% of the non-matches. Taxa were considered as indifferent in 4.6% to 31.8%, which accounted for the majority of NA’s in five out of six combinations (Table 2).

Statistical analyses

Among the 42 combinations of dataset × EIV system × averaging method, mean EIVs were significantly positively correlated with the respective measured environmental variable in all cases except one both for Pearson’s r and Spearman’s r_S (Suppl. material 1). Pearson’s r values ranged from 0.34 to 0.84 (mean: 0.74) and Spearman’s r_S values from 0.21 to 0.84 (mean: 0.67) (Suppl. material 1). The correlation results with Pearson’s r and Spearman’s r_s were consistent with a high correlation between both values (r = 0.88). Therefore, all further results are shown only for Pearson’s r. The slopes and intercepts of linear regressions predicting the actual environmental conditions from mean EIVs are also provided in Suppl. material 1, while their visualisation can be found in Suppl. material 2. The regression functions in the case of EIVE are pH = 2.28 + 0.70 EIVE-R (mean of all functions) or pH = 2.38 + 0.72 EIVE-T (highest R^2) and MAT (°C) = 1.70 + 0.82 EIVE-T (mean) or MAT (°C) = 1.59 + 0.86 EIVE-T (highest R^2). Some of the best-fitting regressions for the three datasets are shown in Figure 2.

Among the three datasets, r values varied significantly (p = 0.003). They were highest in the 10-m² pH dataset (mean: 0.79), followed by the 1-m² pH dataset (mean: 0.73) and lowest in the 10-m² temperature dataset (mean: 0.69), albeit the differences were only

Table 2. Frequency and reasons of taxa without definitive indicator values in the three EIV systems by Tichý et al. (2023), Ellenberg et al. (1991) and Landolt et al. (2010). In the assessments of both indicators, we had 371 valid taxa, but the taxon lists were not completely identical as the plots differed.

Indicator	**R (pH)**			**T (temperature)**		
EIV system	**Tichý et al. (2023)**	**Ellenberg et al. (1991)**	**Landolt et al. (2010)**	**Tichý et al. (2023)**	**Ellenberg et al. (1991)**	**Landolt et al. (2010)**
Taxa indifferent	40	62	17	53	118	33
Taxa completely missing	0	13	1	0	12	1
Taxa only treated at superior or inferior level	22	21	27	22	20	27
All NA’s	62	96	45	75	150	61
Fraction indifferent	10.8%	16.7%	4.6%	14.3%	31.8%	8.9%
Fraction completely missing	0.0%	3.5%	0.3%	0.0%	3.2%	0.3%
Fraction other level	5.9%	5.7%	7.3%	5.9%	5.4%	7.3%
Fraction NA’s	16.7%	25.9%	12.1%	20.2%	40.4%	16.4%

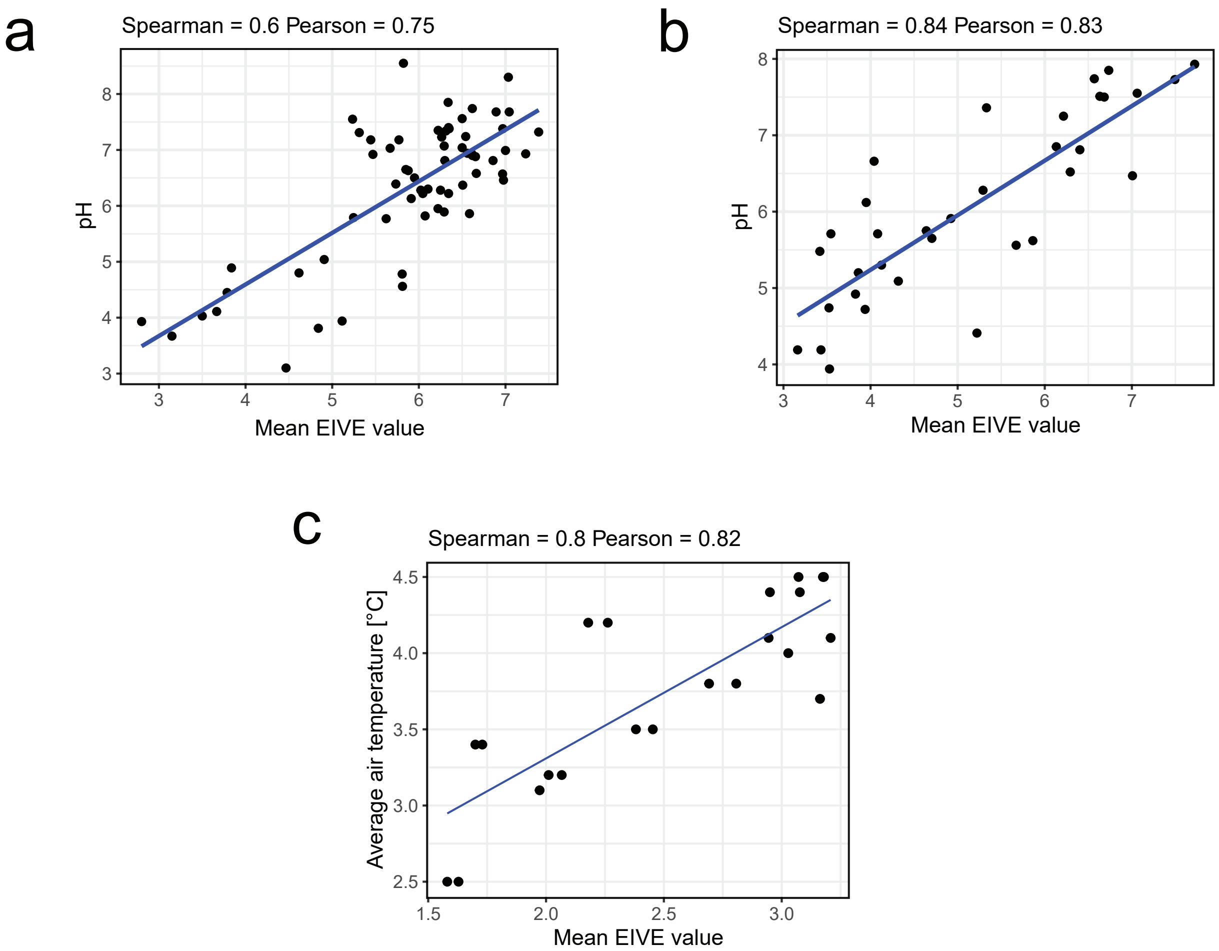

Figure 2. Linear regressions for environmental variables (pH, mean annual temperature) vs. the mean indicator values (R, T) for the three datasets: (a) 1 m² R, (b) 10 m² R and (c) 10 m² T. We show the versions for unweighted mean EIVE values, which (together with inverse-niche width weighted EIVE values) was always the best combination of EIV system and weighting approach. Above the figures, the corresponding Spearman and Pearson correlation coefficients are given.

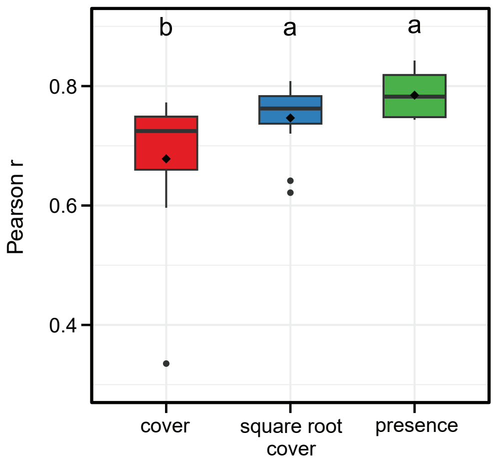

Figure 3. Boxplots of the Pearson *r* values for the correlations between mean ecological indicator values (EIVs) and measured environmental variables for the three weighting approaches available across all four EIV systems. Arithmetic means are indicated as rhombi. The overall effect of the weighting approach was significant in the mixed effects model (p = 0.003) (see text for details). The lower-case letters refer to homogenous groups according to a post-hoc test.

significant between the 10-m² pH and the 10-m² temperature datasets (not shown).

When comparing the *r* values for the three weighting approaches available for all four EIV systems, the differences were highly significant (p = 0.003), with presence (mean: 0.78) and square-root cover (mean: 0.75) performing significantly better than cover (mean: 0.68) (Figure 3). When comparing the *r* values for all four weighting approaches across the two EIV systems that allow inverse niche width weighting (EIVE, "Landolt"), the differences were significant (p = 0.014), with presence (mean: 0.79) and inverse niche width (mean: 0.79) performing best, square-root cover (mean: 0.75) intermediate and cover (mean: 0.65) worst (Figure 4).

Regarding the performance of the different EIV systems, the overall differences in *r* values were relatively small, ranging from EIVE (mean: 0.76) via "Ellenberg" (mean: 0.75) and "Tichý" (mean: 0.73) to "Landolt" (mean: 0.71) (Figure 5). We could confirm our a priori hypothesis that EIVE performs better than the other EIV systems in the case of "Tichý" (p = 0.006 after Bonferroni correction), but not for "Ellenberg" (p = 0.466) and "Landolt" (p = 0.164).

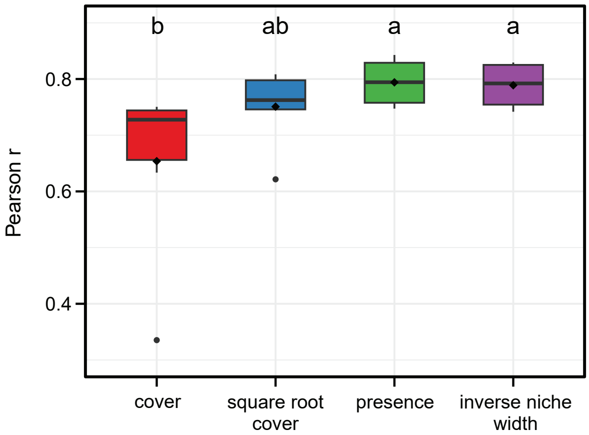

Figure 4. Boxplots of the Pearson *r* values for the correlations between mean ecological indicator values (EIVs) and measured environmental variables for all four weighting approaches, restricted to the two EIV systems that allow the use of inverse niche width weighting (EIVE and "Landolt"). Arithmetic means are indicated as rhombi. The overall effect of the weighting approach was significant in the mixed effects model (p = 0.014) (see text for details). The lower-case letters refer to homogenous groups according to a post-hoc test.

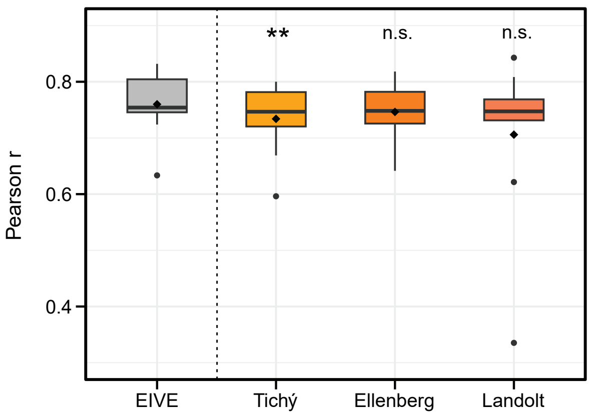

Figure 5. Boxplots of the Pearson *r* values for the correlations between mean ecological indicator values (EIVs) and measured environmental variables, depending on the EIV system. The arithmetic means are indicated as rhombi. The differences compared to EIVE were tested with Bonferroni-corrected pairwise *t*-tests and the results displayed about the respective boxes. For details, see text.

Discussion

Taxonomic matching and coverage

In our examples, the indicator value coverage was 100% in EIVE, but significantly lower for the three other systems. As expected in Central Europe, which is the centre of the suggested validity of the three other systems, there were only few completely missing taxa. A total of 13 missing taxa in Ellenberg et al. (1991) demonstrates that this system has some knowledge gaps in the alpine areas of Central Europe (despite having Central Europe in the title). Unexpectedly, also one taxon (*Hieracium pilosum*) is not included in Landolt et al. (2010) under this name or any known synonym. A bigger issue were taxa that are essentially present in the source systems, but at a higher or lower taxonomic resolution. Only in EIVE this problem did not occur as the EIVE workflow populated all the valid taxa below the genus with indicator values (Dengler et al. 2023). In the three other EIV systems, such cases could be resolved by an approximative assignment by hand, for example, by applying the indicator value of the species for its subspecies, but this is tedious work and comes with loss of information. The situation is pronounced in Tichý et al. (2023), whose authors refrained providing indicator values to any subspecies even if subspecies of one species differ substantially in their niches, often leading to an assessment as indifferent at the higher taxonomic level. The most frequent reasons for NA's in the three systems (besides EIVE) however was that their authors considered the amplitude of a species as too wide to be given a definitive indicator value.

Overall performance and differences between the three datasets

Not surprisingly for almost all our calculation methods and datasets the correlations between mean EIVs and measured environmental variables were significant. This coincides with the findings of numerous studies on the suitability of mean EIVs for biological indication (reviewed in Ellenberg et al. 1991; Diekmann 2003). Moreover, the high correlation between Pearson's r and Spearman's r_S values combined with the generally higher r values indicate that the relationship between mean R values and soil pH and mean T values and mean annual temperature are almost linear. This was also found in other studies (Wamelink et al. 2002 for R; Dengler et al. 2023 for T), while for other niche dimensions such relationships might only be monotonous, but not linear.

We found some differences regarding correlation strength between the three datasets. First, mean R values were more strongly correlated to soil pH in the 10-m² than in the 1-m² dataset, but insignificantly so. This tendency was not expected at first glance as one should assume lower within-plot variability in soil properties and thus a stronger relationship within smaller plots. While this reasoning is probably not false, the fact that the mean indicator values in the larger plots are based on more species on average and thus more reliable, seems to be the more relevant factor. Second, we found clearly stronger correlations in 10-m² plots for mean R values than for mean T values. The most likely reason for this difference in the correlation strength for T and R in our study is the different gradient length covered. Our mean EIVE-R values ranged from 3 to 8, while our mean EIVE-T values covered only values of 1.5 to 3.5 (Figure 2). With 10 EIVE units being defined as the difference between the lowest and highest value of a variable in a plant-inhabited site in Europe (Dengler et al. 2023), we covered approximately 50% of the European gradient of soil reaction, but only 20% of the temperature gradient. It is a well-known statistical phenomenon that the longer an observed gradient is, the more significant a pattern becomes.

Differences among the four weighting approaches

We found that the less influence species cover is given in the calculation of mean EIVs, the higher their predictive power is. Across EIV systems and datasets, unweighted means were better than square-root cover-weighted means (with this difference not being significant) and these again better than fully cover-weighted means. Up to now there was little evidence on the effect of different weighting approaches, and even more rarely originating from comparative studies (but see Käfer and Witte 2004; Hájek et al. 2020). In his comprehensive review, Diekmann (2003) concluded that the results of the weighting vs. non-weighting "with few exceptions do not differ much". Our study now suggests that while the overall relation remains the same, the strength of the relation is higher if cover is disregarded. This agrees with the findings of Käfer and Witte (2004) for mean EIVs for moisture. Prior to our study we would have expected that the intermediate solution (square-root cover-weighting) outperforms either non-weighting or full cover weighting (for reasoning, see Introduction). But our results did not support this expectation, and instead the number of species that entered into the calculation of the mean was more decisive. While all species in a system usually have been estimated by the same expert or group of experts, each of the species EIVs comes with an estimation error. Some are under- and others overestimated. Thus, the larger the number of averaged species, the more reliable the resulting mean. Using presence-absence data, effectively more species enter into the calculation of the mean compared to situations when some species are rated differently.

Finally, we found that inverse niche width-weighting essentially produced results identical to no weighting. This finding contrasts with the logical assumption that species with wider niches should have lower indicative value. This idea had motivated some EIV systems (e.g. Landolt et al. 2010; Böhling et al. 2002; Hájek et al. 2020) as well as EIVE (Dengler et al. 2023) to provide a measure of niche width in addition to niche position. Hájek et al. (2020) showed that under certain circumstances using niche width (they call it "tolerance") for weighting yielded the strongest correlations with measured environmental variables, but not consistently so. Our counterintuitive finding could suggest that the effect of niche width is indeed small or that the niche width measure in EIVE needs further improvement.

Differences among the four indicator value systems

We found differences between the four systems to be relatively small, with mean *r* values ranging from 0.76 for EIVE to 0.71 for "Landolt". Based on previous studies we had hypothesised that EIVE generally should outperform all other systems. Indeed, the mean *r* values had been highest for EIVE but there only was a significant difference compared to "Tichý". This is astonishing as the motivation of developing the two European systems (Dengler et al. 2023; Tichý et al. 2023) was not to replace the existing regional systems, but to provide a solution for supra-regional to continental analyses. Tichý et al. (2023) even explicitly wrote "It is likely that most regional systems of indicator values provide more accurate estimates of site conditions in their region than the European data set, which is based on averaging indicator values from different regions." Already during the preparation of EIVE 1.0 (Dengler et al. 2023), this assumption formerly also shared by the EIVE authors was challenged. First, Moeys (2020) in her methodologically similar Master thesis reported for a beta version of EIVE (far less developed than EIVE 1.0) that it outperformed regional EIV systems in most cases. Dengler et al. (2023) then found that the correlation of the T values of individual species to their climatic niche (derived from distribution data) was strongest in EIVE, closely followed by Tichý et al. (2023) while it was much weaker for most of the regional systems. Our results now are similar, except that Tichý et al. (2023) was even behind the two applicable regional EIV systems (Ellenberg et al. 1991; Landolt et al. 2010). How could this result arise even though there is little doubt that niches of many species vary in space? It seems that EIVs at species level are better, the more independent assessments were used for their creation, which refers to the statistical principle "wisdom of the crowd" (Galton 1907; Surowiecki 2004). According to this principle the mean of the estimates of a continuous variable by many non-experts is often more precise than a single estimate by the best expert. Surowiecki (2004) highlights five prerequisites for the wisdom of the crowd to work, the first two being "diversity of opinion" and "independence of people's opinion". This would nicely explain why EIVE was superior to "Tichý": EIVE 1.0 is based on 28 systems for R and 23 for T, respectively, "Tichý" on seven systems for both reaction and temperature. This means that "Tichý" for the two niche dimensions considered had only between one fourth and one third of the sources that EIVE 1.0 used, and, perhaps more importantly, Tichý et al. (2023) according to their methods description used only sources that were conceptually very similar to "Ellenberg", thus likely hardly independent from each other, while Dengler et al. (2023) included all available systems they were aware of, how "eccentric" they might be. Our study suggests that the approach of Tichý et al. (2023) to "tie" themselves to "Ellenberg" prevented them from becoming better than "Ellenberg" despite combining several systems. Lastly, the fact that "Ellenberg" and "Landolt" were not significantly different from EIVE suggests that there the two drivers of quality, regionality of the assessment and number of independent assessments, cancelled themselves out.

Conclusions and outlook

Our results suggest that (full) cover weighting is clearly not recommendable when calculating mean ecological indicator values, in line with Käfer and Witte (2004), but contrary to the assumption of Diekmann (2003) that in species-rich vegetation (as in our case) this weighting approach is not relevant. While our small dataset does not provide a clear answer regarding square-root cover weighting, there was a tendency that the results with this intermediate solution were worse than without any weighting. By contrast, inverse niche width weighting produced almost identical results to no weighting, suggesting that with improved niche width measures from future editions of EIVE it might be possible toachieve a predictive power better than with no weighting. One could also think of other approaches to involve niche width in the predictions such as the amplitude overlap method successfully applied by Peppler-Lisbach (2008) or even multiplying the response curves of all occurring species. Currently, the latter approach is impossible to implement at large scales as response curves exist only for a limited number of species and few niche dimensions, based on relatively small datasets. Large vegetation-plot databases like EVA (Chytrý et al. 2016) and GrassPlot (Dengler et al. 2018) might offer a chance to overcome this limitation.

Regarding different EIV systems, our study clearly disproves the assumption of Tichý et al. (2023) that, within their region of definition, well-established regional EIV systems perform better than European EIV systems – as there was no significant difference between EIVE and the two regional systems, but a tendency of EIVE to perform better – in line with the findings of Moeys (2020). By contrast, we found a small but highly significant superiority of the European EIVE system (Dengler et al. 2023) vs. the European "Ellenberg-type" system by Tichý et al. (2023). From the perspective of practitioners these findings suggest that in pan-European studies and in European regions without well-established regional system, EIVE should be the EIV system of choice. By contrast, in regions with a well-established EIV system one can either select the regional system or EIVE depending on whether comparisons of the results are intended to be done regionally or internationally. Beyond the reliability of the predictions, we found that there is another practical aspect that makes the use of EIVE easier and better reproducible than that of any of the other systems. Only EIVE systematically provides EIV values for all accepted taxonomic ranks (subspecies, species, aggregates), which allows direct usage without either losing information on many taxa recorded at a different taxonomic level or requiring tedious and necessarily imperfect "matching" with the next-lower or next-higher taxon.

The "wisdom of the crowd" principle might play an unexpected role for the predictive power of mean EIVs: This single principle could explain why EIVE performs better than Tichý et al. (2023), why no weighting of species is better than cover-weighting and finally why the predictive power tended to be higher for 10-m^2 than for 1-m^2 plots. In the first case it is about averaging more independent expert opinions, in the latter two cases about more species that effectively or actually enter into the calculation. Currently, "wisdom of the crowd" is just a hypothesis that would explain the named patterns at once and thus parsimoniously. However, whether this is the true and sole reason for the findings needs to be addressed in future studies specifically devoted to this topic.

While the current study provided some answers, it should be emphasised that it comes from a narrow geographic region with three limited datasets. Therefore, similar studies are also needed in other regions or with pan-European datasets to assess the generality of our findings and to provide better-founded translation functions to physical variables. It also remains to be studied whether similar principles apply for the other three current niche dimensions in EIVE 1.0 (moisture, nitrogen, light). Hájek et al. (2020), for example, found that in the case of soil moisture, depending on the specific setting, different weighting approaches might be best. Lastly, also the question how the prediction quality of mean indicator values depends on plot size is of great theoretical and practical interest but requires for proper testing high quality nested-plot datasets where environmental variables were measured separately for each size.

Data availability

The correlation coefficients underlying the statistics presented are provided in Suppl. material 1; the raw plot data can be requested from the corresponding author.

Author contributions

SA, AM and GO performed the underlying study as a student project in the trinational Master Summer School "Biodiversity Monitoring" (coordinators: JD and ID) under the supervision of JD, OC and SW in 2023. SA, AM, GO and JD collected and harmonized the data, GO led the analyses and visualisations in R, with contributions by AM, JD, SW and ID, while SA prepared the map. JD led the writing of the manuscript, which was improved and approved by the rest of the authors.

Acknowledgements

The data were collected during Summer Schools in the years 2019, 2020, 2021 and 2023. We thank all other students and teachers who were involved in vegetation sampling and measurement of abiotic variables. We are grateful to François Gillet and one anonymous reviewer for very detailed and helpful feedback on an earlier version of this manuscript and to Michael Glaser for the linguistic improvement of our manuscript.

References

Blaus A, Reitalu T, Gerhold P, Hiiesalu I, Massante JC, Veski S (2020) Modern pollen–plant diversity relationships inform palaeoecological reconstructions of functional and phylogenetic diversity in calcareous fens. Frontiers in Ecology and Evolution 8: Article 207. https://doi.org/10.3389/fevo.2020.00207

Böhling N, Greuter W, Raus T (2002) Zeigerwerte der Gefäßpflanzen der Südägäis (Griechenland). Braun-Blanquetia 32: 1–108.

Chytrý M (Ed.) (2007) Vegetation of the Czech Republic – 1. Grassland and heathland vegetation [in Czech, with English summary]. Academia, Praha, CZ, 526 pp.

Chytrý M, Hejcman M, Hennekens SM, Schellberg J (2009) Changes in vegetation types and Ellenberg indicator values after 65 years of fertilizer application in the Rengen Grassland Experiment, Germany. Applied Vegetation Science 12: 167–176. https://doi.org/10.1111/j.1654-109X.2009.01011.x

Chytrý M, Hennekens SM, Jiménez-Alfaro B, Knollová I, Dengler J, Jansen F, Landucci F, Schaminée JHJ, Aćić S, … Yamalov S (2016) European Vegetation Archive (EVA): an integrated database of European vegetation plots. Applied Vegetation Science 19: 173–180. https://doi.org/10.1111/avsc.12191

De Pauw K, Depauw L, Calders K, Cousins SAO, Decocq G, De Lombaerde E, Diekmann M, Frey D, Lenoir J, … De Frenne P (2024) Nutrient-demanding and thermophilous plants dominate urban forest-edge vegetation across temperate Europe. Journal of Vegetation Science 35: e13236. https://doi.org/10.1111/jvs.13236

Dengler J (2008) Pitfalls in small-scale species-area sampling and analysis. Folia Geobotanica 43: 269–287. https://doi.org/10.1007/s12224-008-9014-9

Dengler J, Dembicz I (2023) Should we estimate plant cover in percent or on ordinal scales? Vegetation Classification and Survey 4: 131–138. https://doi.org/10.3897/VCS.98379

Dengler J, Boch S, Filibeck G, Chiarucci A, Dembicz I, Guarino R, Henneberg B, Janišová M, Marcenò C, … Biurrun I (2016) Assessing plant diversity and composition in grasslands across spatial scales: the standardised EDGG sampling methodology. Bulletin of the Eurasian Dry Grassland Group 32: 13–30.

Dengler J, Wagner V, Dembicz I, García-Mijangos I, Naqinezhad A, Boch S, Chiarucci A, Conradi T, Filibeck G, … Biurrun I (2018) GrassPlot – a database of multi-scale plant diversity in Palaearctic grasslands. Phytocoenologia 48: 331–347. https://doi.org/10.1127/phyto/2018/0267

Dengler J, Jansen F, Chusova O, Hüllbusch E, Nobis MP, Van Meerbeek K, Axmanová I, Bruun HH, Chytrý M, … Gillet F (2023) Ecological Indicator Values for Europe (EIVE) 1.0. Vegetation Classification and Survey 4: 7–29. https://doi.org/10.3897/VCS.98324

Didukh YP (2011) The ecological scales of the species of the Ukrainian flora and their use in synphytoindication. Phytosociolcentre, Kiyv, UA, 175 pp.

Diekmann M (2003) Species indicator values as an important tool in applied plant ecology – a review. Basic and Applied Ecology 4: 493–506. https://doi.org/10.1078/1439-1791-00185

Diekmann M, Andres C, Becker T, Bennie J, Blüml V, Bullock JM, Culmsee H, Fanigliulo M, Hahn A, … Wesche K (2019) Patterns of long-term vegetation change vary between different types of semi-natural grasslands in Western and Central Europe. Journal of Vegetation Science 30: 187–202. https://doi.org/10.1111/jvs.12727

Ellenberg H (1974) Zeigerwerte mitteleuropäischer Gefäßpflanzen. Scripta Geobotanica 9: 1–97.

Ellenberg H, Weber HE, Düll R, Wirth V, Werner W, Paulißen D (1991) Zeigerwerte von Pflanzen in Mitteleuropa. Scripta Geobotanica 18: 1–248.

Euro+Med (2025) Euro+Med PlantBase – the information resource for Euro-Mediterranean plant diversity. http://www.europlusmed.org [accessed on 21 January 2025]

Galton F (1907) Vox populi. Nature 75: 450–451. https://doi.org/10.1038/075450a0

Gillet F, Machamp L, Badot PM, Mouly A (2016) Recent change in mountain grasslands: a vegetation resampling study. Ecology and Evolution 6: 2333–2345. https://doi.org/10.1002/ece3.1987

Hájek M, Dítě D, Horsáková V, Mikulášková E, Peterka T, Navrátilová J, Jiménez-Alfaro B, Tichý L, Horsák M (2020) Towards the pan-European bioindication system: Assessing and testing updated hydrological indicator values for vascular plants and bryophytes in mires. Ecological Indicators 116: Article 106527. https://doi.org/10.1016/j.ecolind.2020.106527

Hothorn T, Bretz F, Westfall P (2008) Simultaneous inference in general parametric models. Biometrical Journal 50: 346–363. https://doi.org/10.1002/bimj.200810425

Info Flora (2025) Vegetationsaufnahmen in FlorApp. https://docs.infoflora.ch/de/FlorApp/Vegetationsaufnahmen_Florapp.html [accessed 14 Jan 2025]

Jansen F, Dengler J (2010) Plant names in vegetation databases – a neglected source of bias. Journal of Vegetation Science 21: 1179–1186. https://doi.org/10.1111/j.1654-1103.2010.01209.x

Käfer J, Witte JPM (2004) Cover-weighted averaging of indicator values in vegetation. Journal of Vegetation Science 15: 647–652. https://doi.org/10.1111/j.1654-1103.2004.tb02306.x

Kuznetsova A, Brockhoff PB, Christensen RHB (2017) lmerTest package: Tests in linear mixed effects models. Journal of Statistical Software 82: 1–26. https://doi.org/10.18637/jss.v082.i13

Laliberté E, Legendre R, Shipley B (2014) FD: measuring functional diversity from multiple traits, and other tools for functional ecology. R package version 1.0-12.3. https://cran.r-project.org/web/packages/FD/index.html [accessed 13 Aug 2024]

Landolt E, Bäumler B, Erhardt A, Hegg O, Klötzli F, Lämmler W, Nobis M, Rudmann-Maurer K, Schweingruber FH, … Wohlgemuth T (2010) Flora indicativa – Ökologische Zeigerwerte und biologische Kennzeichen zur Flora der Schweiz und der Alpen. 2nd ed. Haupt, Bern, CH, 378 pp.

Moeys K (2020) Testing a new European system of Ecological Indicator Values for plants. Master thesis, Faculty of Bioscience Engineering, KU Leuven, Leuven, BE, 62 pp.

Mucina L, Bültmann H, Dierßen K, Theurillat JP, Raus T, Čarni A, Šumberová K, Willner W, Dengler J, … Tichý L (2016) Vegetation of Europe: Hierarchical floristic classification system of vascular plant, bryophyte, lichen, and algal communities. Applied Vegetation Science 19(Supplement 1): 3–264. https://doi.org/10.1111/avsc.12257

Pärtel M, Zobel M, Zobel K, van der Maarel E (1996) The species pool and its relation to species richness: evidence from Estonian plant communities. Oikos 75: 111–117. https://doi.org/10.2307/3546327

Peppler-Lisbach C (2008) Using species-environment amplitudes to predict pH values from vegetation. Journal of Vegetation Science 19: 437–444. https://doi.org/10.3170/2008-8-18394

R Core Team (2024) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, AT. http://www.r-project.org/

Ramensky LG, Tsatsenkin IA, Chizhikov ON, Antipin NA (1956) Ekologicheskaya otsenka kormovykh ugodiy po rastitel'nomu pokrovu [Ecological assessment of fodder lands based on vegetation traits]. VNIIK, Moscow, RU, 471 pp.

Reczyńska K, Novák P, Majerová M, Świerkosz K (2024) The current state of relic rocky grasslands with Sesleria tatrae and Calamagrostis varia in the Eastern Sudetes (Śnieżnik Massif, SW Poland). Acta Societatis Botanicorum Poloniae 93: Article 181645. https://doi.org/10.5586/asbp/181645

Reutimann P, Billeter R, Dengler J (2023) Effects of grazing versus mowing on the vegetation of wet grasslands in the Northern Pre-Alps, Switzerland. Applied Vegetation Science 26: e12706. https://doi.org/10.1111/avsc.12706

Scherrer M, Lüthi R, Bugmann H, Burnand J, Wohlgemuth T, Rudow A (2024) Impact of climate warming, pollution, and management on the vegetation composition of Central European beech forests. Ecological Indicators 160: 111888. https://doi.org/10.1016/j.ecolind.2024.111888

Surowiecki J (2004) The wisdom of crowds. Doubleday, New York, US, 336 pp.

Tichý L, Axmanová I, Dengler J, Guarino R, Jansen F, Midolo G, Nobis MP, Van Meerbeek K, Attorre F, … Chytrý M (2023) Ellenberg-type indicator values for European vascular plant species. Journal of Vegetation Science 34: e13168. https://doi.org/10.1111/jvs.13168

Tölgyesi C, Bátori Z, Erdős L (2014) Using statistical tests on relative ecological indicator values to compare vegetation units – Different approaches and weighting methods. Ecological Indicators 26: 441–446. https://doi.org/10.1016/j.ecolind.2013.09.002

Vassilev K, Bergmeier E, Boch S, Pedashenko H, Sopotlieva D, Tsiripidis I, Apostolova I, Fotiadis G, Ganeva A, … Dengler J (2024) Classification of the high-rank syntaxa of the Balkan dry grasslands with a new hierarchical expert system approach. Applied Vegetation Science 27: e12779. https://doi.org/10.1111/avsc.12779

Wamelink GWW, Joosten V, van Dobben HF, Berendse F (2002) Validity of Ellenberg indicator values judged from physico-chemical field measurements. Journal of Vegetation Science 13: 269–278. https://doi.org/10.1111/j.1654-1103.2002.tb02047.x

E-mail and ORCID

Grzegorz Ostrowski (g.ostrowski@uw.edu.pl), ORCID: https://orcid.org/0000-0001-5453-0670
Severin Aicher (aicherseverin@gmail.com), ORCID: https://orcid.org/0009-0000-7875-1937
Agnieszka Mankiewicz (am.mankiewic@student.uw.edu.pl), ORCID: https://orcid.org/0009-0009-1252-2686
Olha Chusova (olgachusova28@gmail.com), ORCID: https://orcid.org/0000-0002-8081-9918
Iwona Dembicz (iwona.dembicz@uw.edu.pl), ORCID: https://orcid.org/0000-0002-6162-1519
Stefan Widmer (stefan.widmer@zhaw.ch), ORCID: https://orcid.org/0000-0002-4920-5205
Jürgen Dengler (Corresponding author, juergen.dengler@zhaw.ch), ORCID: https://orcid.org/0000-0003-3221-660X

Supplementary material

Supplementary material 1
Full correlation and regression results for all combinations of dataset × EIV system × weighting approach (*.xlsx).
Link: https://doi.org/10.3897/VCS.134800.suppl1

Supplementary material 2
Visualisation of the linear regressions for all combinations of dataset × EIV system × weighting approach (*.pdf).
Link: https://doi.org/10.3897/VCS.134800.suppl2