

RESEARCH ARTICLE

Scaling species richness: When vegetation surveys don't match in plot area

Marco Barandun^{1,2}  | Caren M. Pauler³  | Gisela Lüscher⁴  | Antoine Guisan^{5,6}  |
Eva Knop^{2,7,8}  | Manuel K. Schneider^{1,3} 

¹Forage Production and Grassland Systems, Agroscope, Zurich, Switzerland; ²Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, Switzerland; ³Grazing Systems, Agroscope, Zurich, Switzerland; ⁴Agricultural Landscape and Biodiversity, Agroscope, Zurich, Switzerland; ⁵Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland; ⁶Institute of Earth Surface Dynamics, University of Lausanne, Lausanne, Switzerland; ⁷Plants and Plant Products, Agroscope, Zurich, Switzerland and ⁸Faculty of Biology and Pre-Clinical Medicine, University of Regensburg, Regensburg, Germany

Correspondence

Marco Barandun

Email: marco.barandun@agroscope.admin.ch**Funding information**

Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung, Grant/Award Number: 110000 and 31BD30_209612; European Union, Grant/Award Number: 227161; German Academic Scholarship Foundation; Agroscope; Biodiversa+, Grant/Award Number: 101052342

Handling Editor: Frederik De Laender**Abstract**

1. Differences in study plot area remain a central obstacle to the analysis of biodiversity data. Plot area (i.e. the surface assessed for a single observation) in vegetation surveys is rarely standardized and often varies across studies, complicating comparisons. The species–area relationship (SAR), often modelled as a power law, provides a simple means of scaling species richness across different plot areas. Yet, more than a century after its introduction, applying the power law to compare vegetation surveys conducted at different plot areas still lacks thorough evaluation and quantification of uncertainty.
2. Using a unique dataset with over 1000 nested-plot vegetation surveys from Swiss grasslands, with subplot sizes ranging from 1 to 100m², we assessed the scaling of species richness across plot areas. To do so, we compared SAR models with machine learning approaches with area parameters as covariates. Additionally, we examined how ecological indicator values (EIVs) vary with plot area, evaluating their potential integration into studies based on scaled species richness.
3. We found that species richness can be reliably scaled across surveys differing in plot area. SAR models using a constant accumulation coefficient offered a simple and effective approach, particularly for small differences in plot area between surveys. In contrast, machine learning models, including area parameters as covariates showed greater accuracy when extrapolating across surveys with widely differing plot areas.
4. EIVs differed only little across plot areas, with differences remaining well below the overall variability observed between surveys—even when comparing widely differing plot areas. This suggests that indicator values are largely independent of spatial grain and can therefore be used as stable covariates for characterizing vegetation, even when data are derived from differing plot areas.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *Journal of Ecology* published by John Wiley & Sons Ltd on behalf of British Ecological Society.

5. *Synthesis.* This study shows that species richness in grasslands can be reliably harmonized across surveys differing in plot area. It also highlights that EIVs remain largely stable across different plot areas, making them suitable for the use in data-rich ecological analyses. Together, these findings provide practical tools for harmonizing grassland datasets and offer new insights into how we can compare biodiversity assessments across space and time.

KEYWORDS

biodiversity scaling, determinants of plant community diversity and structure, ecological indicator values, grassland diversity, plot size, species–area relationship, vegetation survey standardization

1 | INTRODUCTION

One of the oldest and most fundamental questions in ecology is how biodiversity is distributed across space and what drives these patterns. Among the various metrics used to quantify biodiversity, species richness is the most widely adopted and easily understood. However, species richness is not a static measure; it is inherently dependent on the area (and time; not addressed here) at which it is observed. In particular, the way in which the number of species increases with the increase in the observed area represents a quasi-universal phenomenon (Matthews et al., 2021). The so-called species–area relationship (SAR) in the form of a power–law relationship was first postulated by Arrhenius (1921). This relationship serves as a conceptual baseline in ecological theory, has fundamentally shaped frameworks, such as the equilibrium theory of island biogeography (MacArthur & Wilson, 1967), and plays a critical role in conservation (Rosenzweig, 2004).

As biodiversity datasets grow exponentially (e.g. Chytrý et al., 2016), the heterogeneity in vegetation survey plot area—stemming from varying methodologies and research objectives—poses a major challenge for comparing biodiversity patterns across surveys, regions and time periods (Chytrý & Otýpková, 2003). A more systematic understanding of biodiversity scaling approaches is crucial for harmonizing datasets, enabling researchers to fully leverage the growing wealth of biodiversity data. This is especially important for integrating scarce data from underrepresented regions, where biodiversity is often high but sampling remains limited (Cazzolla Gatti et al., 2022; Chytrý et al., 2016). Improving data comparability is essential for advancing our understanding of biodiversity patterns at a time of unprecedented biodiversity decline and environmental change. SAR offers a powerful framework to address these challenges, providing a mathematical basis for standardizing species richness across plot areas.

The SAR has been described using various mathematical functions (Connor & McCoy, 1979; Tjørve, 2003), but the oldest and simplest model (Arrhenius, 1921) remains the one that is best understood and represents the accumulation of species over area best (Dengler, 2009). This model describes species richness (S) as a function of the area observed (A) through a power–law relationship:

$$S = cA^z$$

where c represents small-scale species richness, defined as the number of species present within a unit area (in this case 1 m^2), serving as a baseline measure of biodiversity at fine spatial scales. In contrast, the accumulation coefficient z quantifies the rate at which species richness increases as the sampled area expands. Higher z -values reflect steeper species–accumulation curves, indicating more heterogeneous species distributions driven by differences in species pool size, abundance, evenness or spatial aggregation (Chase & Knight, 2013; DeMalach et al., 2019; He & Legendre, 2002; Matthews et al., 2021; Tjørve et al., 2008).

There are few possible approaches to address differences in plot area between surveys: (1) restricting analysis to data of the most frequent plot area (e.g. Janišová et al., 2014), (2) including area parameters as an explanatory covariate in the model (Divíšek & Chytrý, 2018; Sabatini et al., 2022; Večeřa et al., 2019) or (3) using the SAR to scale richness (Arrhenius, 1921). While restricting to data of the most frequent plot area ensures consistency, it significantly reduces the usable dataset and may introduce bias. Including area parameters as a covariate allows broader data inclusion and may capture nonlinear effects in flexible models, such as random forests (RFs). However, these flexible models would still not explicitly represent the mathematical structure of the SAR and may fail to generalize its shape with high accuracy, especially in small datasets. In fact, models applied to environmental data layers—typically at much coarser resolutions (e.g. 1 km^2) than the survey plot area—are unlikely to accurately estimate the relationship between area and species richness at much finer spatial scales (e.g. 10 m^2). Conversely, the SAR has been described as one of the oldest and closest approximations to an ecological law and offers a simple approach to compare surveys of different plot areas (Arrhenius, 1921; Matthews et al., 2021). Much attention has been devoted to aspects, such as the functional form of SARs (e.g. Dengler, 2009; Dengler et al., 2020), their underlying drivers (e.g. DeMalach et al., 2019; Dembicz, Dengler, Gillet, et al., 2021) and the influence of spatial grain (e.g. Chase et al., 2018; Chase & Knight, 2013; Keil & Chase, 2019; Zhang et al., 2021), whereas their use as a framework for biodiversity scaling has received less

attention, with applications focusing on non-contingent plots (Portier et al., 2022) or on extrapolations from plot to biogeographic scales (e.g. Boussange et al., 2025; Kunin et al., 2018). In contrast, despite more than a century since its introduction, a systematic evaluation and error quantification of the SAR's original application (Arrhenius, 1921)—comparing vegetation surveys conducted at different vegetation survey plot areas—remains underexplored.

EIVs are semi-quantitative estimates of the environmental preferences of individual plant species. They are not direct measurements but expert-based scores derived from decades of field observations, ecological literature and experimental studies. EIV systems have been widely developed and applied across Europe (Ellenberg, 1974; Landolt, 1977; Tichý et al., 2023). In Switzerland, Landolt indicator values apply to nearly the entire native flora (Landolt et al., 2010). Each species is assigned an ordinal value along major ecological gradients, such as soil moisture, soil pH, nutrient availability, humus, light, temperature or continentality. These values capture central tendencies of a species' realized niche rather than precise physiological optima (Zolotova et al., 2022).

As an example, the soil reaction indicator (*R*) in the Landolt system represents species' preferences along a pH gradient. Species are scored from 1 (extremely acid, pH 2.5–5.5) to 5 (alkaline, pH 6.5–8.5). Acidophilous species typical of, for example, raised bogs, such as *Drosera rotundifolia* L. (*R* = 1), receive low values, while calcicolous species occurring, for example, on limestone soils, such as *Dryas octopetala* L. (*R* = 5), receive high values (Landolt et al., 2010). Such ordinal scores allow botanists to summarize the ecological tendencies of thousands of species in a simple, comparable form.

At the community level, species EIVs are aggregated into site EIVs, usually as means or abundance-weighted means across all species in a vegetation survey plot. These site-level indicators integrate the ecological information of many species and provide robust proxies of long-term site conditions. Because they capture cumulative species–environment relationships, they can reflect variables that fluctuate strongly over time (such as soil moisture) more reliably than single field measurements. Site EIVs have been shown to reflect local conditions more accurately than coarse environmental predictor layers (Scherrer & Guisan, 2019) and are an important tool in plant ecology (Diekmann, 2003). They are widely used, for example, in global change research to detect vegetation responses to warming (Lenoir et al., 2010; Scherrer et al., 2017), as a tool in conservation to assess habitat quality (Oostermeijer & Van Swaay, 1998; Smart et al., 2003) and in spatial modelling (Descombes et al., 2020). Although site-level EIVs are widely applied to characterize vegetation and its environment, it remains unclear how strongly they depend on the spatial scale at which vegetation is sampled. This uncertainty raises important questions about how EIVs should be interpreted and compared across surveys of different plot areas. Clarifying these scale effects is essential if EIVs are to be applied consistently alongside other vegetation metrics in biodiversity research, conservation and monitoring. To date, however, a systematic assessment of the scaling

behaviour of EIVs with area is lacking, leaving researchers to establish their own practices on a case-by-case basis.

Nested-plot vegetation surveys—where smaller plots are embedded within larger ones—have long been employed to assess plant communities across spatial scales (Arrhenius, 1921; Stohlgren et al., 1995). This design is used in rapid plant diversity assessments (Kalkhan et al., 2007) and for studying scale-dependent patterns and drivers in plant communities (Dufour et al., 2006; Mod et al., 2020; Morzaria-Luna et al., 2004; Scherrer et al., 2019; Von Oppen et al., 2022). By capturing cumulative species lists with increasing subplot area, nested-plot vegetation surveys provide a unique opportunity for directly assessing the SAR, evaluating the performance of species richness scaling methods and examining how vegetation indices vary with changing plot area.

Here, we compiled a dataset of more than 1000 nested-plot grassland vegetation surveys from Switzerland. We quantify species richness scaling accuracy of different approaches across different plot areas. Furthermore, we investigate the effect of vegetation survey plot area scaling also on EIVs and highlight practical implications of standardizing vegetation surveys for area. To the authors' knowledge, no previous study has evaluated the accuracy of different approaches for accounting for plot area when harmonizing species richness in vegetation surveys—at survey scale—nor the effect of vegetation survey plot area on EIVs. We addressed three main questions: (1) To what extent can species richness be reliably up-scaled or down-scaled across different plot areas? (2) Which methods provide the most accurate estimates for this scaling? (3) To what extent do EIVs vary with plot area? Given the generality of the SAR (Arrhenius, 1921; Dengler, 2009; Matthews et al., 2021), we posit species richness to be scalable, especially at modest differences in plot area, and approaches explicitly accounting for the SAR to be most accurate. Average EIVs may exhibit minor variation with plot area. However, at the spatial scales studied, the surveyed plots typically capture a representative subset of the grassland plant community (Braun-Blanquet, 1928; Ewald, 2003), and therefore, variation in the averages is expected to be negligible.

2 | METHODS

We applied a six-step approach using data from Swiss grassland vegetation studies: (1) we obtained and cleaned nested-plot vegetation survey data of various studies; (2) supplemented them with EIVs as well as other environmental variables; (3) we fitted the power-law SAR to each nested-plot vegetation survey individually to estimate its accumulation coefficient (*z*); (4) we trained models to predict *z* based on environmental variables; (5) we compared the accuracy of species richness scaling for the different models; and (6) we compared, pairwise, the average EIVs of plots within the nested-plot surveys and related them to their coefficient of variation between surveys. The results were analysed using R v4.3.2 (R Core Team, 2023).

2.1 | Data preparation: Vegetation data

We compiled 1081 nested-plot vegetation surveys from Swiss grassland studies, sourced from various datasets (Table 1). No fieldwork permits were necessary. Each survey consisted of nested subplots of increasing area, each with a species list, enabling pairwise comparisons of subplots of differing area within single nested-plot vegetation surveys. Vegetation surveys were filtered to be present in the data layers (Table 1), include only subplots with areas between 1 and 100 m², with non-zero vascular plant species counts and with at least four subplots. After quality control and filtering, 1001 nested-plot vegetation surveys were retained for analysis. The dataset was restricted to vascular plants and taxonomically standardized according to the Flora Helvetica Checklist (Juillerat et al., 2017). To ensure consistency, taxa were aggregated to the highest taxonomic level defined in the checklist. Specifically, subspecies were consolidated at the species level, with their cover values summed. This same principle was applied to species contained in species aggregates. When an aggregated taxon appeared across multiple subplots, it was assigned to the smallest subplot in which it was recorded.

2.2 | Data preparation: Ecological indicator variables and other environmental variables

For each taxon, we extracted Ellenberg-type EIVs from the Flora Indicativa (Landolt et al., 2010). We selected nine indicators to characterize local conditions: soil nutrients (mainly nitrogen), soil pH, soil moisture availability, soil moisture variability, soil humus, soil aeration, light availability, temperature and continentality. Unweighted mean EIVs were calculated for each survey from the species present in the surveyed plant community. Undefined EIVs, that is if a plant species had particularly broad ecological preferences for certain parameters or if a taxon was identified at a higher level, such as genus, were not considered (6.57% of all values). EIVs were supplemented with topographic and bioclimatic variables derived from spatial data layers at 25 m resolution using each survey's geolocation (Table 2). To reduce multicollinearity among bioclimatic variables, we applied a correlation-based variable selection. Specifically, we identified and removed variables with pairwise Pearson correlation coefficients greater than 0.8. Highly correlated variables were identified using the *findCorrelation()* function from the R package *caret* (Kuhn, 2008) and subsequently excluded. This function iteratively excludes the variable with the largest average

absolute correlation, thereby retaining the variable that is overall less correlated with the remaining set. The retained variables are reported in Table 2. Together, these steps yielded a comprehensive set of potential drivers of the accumulation coefficient z , including edaphic, climatic and topographic parameters.

2.3 | Accumulation coefficient z

To the data of each nested-plot vegetation survey, we fitted a power-law SAR model ($S = cA^z$) to estimate the accumulation coefficient z . The fitting used nonlinear least squares implemented in the *nls()* function from the R base package *stats* (R Core Team, 2023) with initial estimates of $c = 1$ and $z = 0.1$. When species richness was constant across all subplots within a survey, we manually assigned a z -value of 0 ($n = 5$; no R^2 value). When the model failed to converge, the site was removed ($n = 1$).

2.4 | Scaling of species richness and model comparison

To evaluate model performance in upscaling and downscaling species richness, we tested two conceptually different harmonization approaches: (1) SAR-based models and (2) RF models that account for plot area by including area parameters as an additional covariate. SAR-based models explicitly incorporate the mathematical form of the relationship, whereas covariate-based models treat plot area as just one among several predictors, without capturing the explicit functional form inherent to the SAR. The nested-plot vegetation surveys were split into a training subset (80%) and a testing subset (20%), ensuring that each nested series was kept together and assigned entirely to either the training or the testing subset.

For SAR-based models, we tested two variants: (1a) using the median z -value derived from the training set or (1b) predicted survey-specific z -values to account for site differences. Each subplot (corresponding to a distinct plot area within a nested-plot vegetation survey) was modelled separately to account for plot area effects. To predict survey-specific z -values, we used a RF model (500 trees, fivefold cross-validation repeated three times, with default *mtry* tuning via the *caret* package; Kuhn, 2008) based on species richness, plot area and all additional predictors listed in Table 2 (e.g. mean EIVs).

RF-based models predicted scaled species richness (i.e. species richness at the scaled plot area) as a function of environmental

Dataset	Number of surveys	Plot area range (m ²)	Number of plots within nested-plot survey	Data source
P	102	1–25	5	Pauler et al. (2019)
B	63	4–100	5	Lüscher et al. (2016)
G	333	1–100	6–9	Unpublished
M	504	1–64	4	Randin et al. (2009), see also Engler et al. (2009)

TABLE 1 Summary of the nested-plot vegetation survey datasets included in this study. For the published datasets, only the largest surveyed area of the nested-plot vegetation surveys has been included in the indicated publication.

TABLE 2 Description and source of predictor variables used in the analysis. Ecological indicator values (EIVs) are based on Flora Indicativa (Landolt et al., 2010).

Variable	Variable short name	Data source
Vegetation cover. Total vascular plant plot cover (maximum 100%). If not directly reported, calculated as the sum of individual plant covers, capped at 100%.	cover_veg	Vegetation survey (Landolt et al., 2010)
Mean ecological indicator value for soil nutrients (mainly nitrogen, unweighted).	eiv_n	Vegetation survey (Landolt et al., 2010)
Mean ecological indicator value for pH (unweighted).	eiv_r	Vegetation survey (Landolt et al., 2010)
Mean ecological indicator value for soil moisture (unweighted).	eiv_f	Vegetation survey (Landolt et al., 2010)
Mean ecological indicator value for soil moisture variability (unweighted).	eiv_w	Vegetation survey (Landolt et al., 2010)
Mean ecological indicator value for soil humus (unweighted).	eiv_h	Vegetation survey (Landolt et al., 2010)
Mean ecological indicator value for soil aeration (unweighted).	eiv_d	Vegetation survey (Landolt et al., 2010)
Mean ecological indicator value for light availability (unweighted).	eiv_l	Vegetation survey (Landolt et al., 2010)
Mean ecological indicator value for temperature (unweighted).	eiv_t	Vegetation survey (Landolt et al., 2010)
Mean ecological indicator value for continentality (unweighted).	eiv_k	Vegetation survey (Landolt et al., 2010)
Elevation above sea level (in metres; 25 m resolution).	elevation	SWECO25 (Külling et al., 2024)
Slope in degrees (25 m resolution).	slope	SWECO25 (Külling et al., 2024)
Index of terrain shading based on solar angle (25 m resolution).	hillshade	SWECO25 (Külling et al., 2024)
Beer's transformed aspect (cosine-transformed orientation; 25 m resolution).	aspect	SWECO25 (Külling et al., 2024)
Isothermality (diurnal/annual temp. range; 1981–2010 average; 25 m resolution)	chbio3	SWECO25 (Külling et al., 2024)
Mean temperature of driest quarter (1981–2010 average; 25 m resolution).	chbio9	SWECO25 (Külling et al., 2024)
Precipitation of wettest month (1981–2010 average; 25 m resolution).	chbio13	SWECO25 (Külling et al., 2024)
Precipitation of driest month (1981–2010 average; 25 m resolution).	chbio14	SWECO25 (Külling et al., 2024)
Precipitation seasonality (1981–2010 average; 25 m resolution).	chbio15	SWECO25 (Külling et al., 2024)

variables (Table 2), species richness at the original plot area, the original plot area (A_1) and the scaled plot area (A_2). Two model types were tested: (2a) using untransformed richness and area values, and (2b) using log-transformed species richness and area. For model training and cross-validation, we included all pairwise combinations of subplots within each survey, capturing the full range of spatial relationships and enabling the model to independently learn the association between species richness and plot area. We applied a RF model with 500 trees, using fivefold cross-validation repeated three times and default *mtry* tuning via the *caret* package (Kuhn, 2008).

To evaluate model robustness and generalization, predictions were assessed on an independent testing dataset consisting of surveys not used during training. Prediction accuracy was quantified as the absolute value of the relative error, calculated as $|(observed - predicted)/observed|$, across all pairwise plot comparisons. To contextualize model performance, we assessed prediction accuracy using a symmetric scaling factor, defined as

$$\text{Scaling factor} = e^{|\ln(A_2/A_1)|}$$

Because the function takes the absolute value of the logarithm, the scaling factor yields the same result regardless of the order in which the plot areas are compared. It thus expresses how many times larger one plot area is relative to the other (e.g. a scaling factor of 2 indicates that one plot is twice the size of the other).

To assess whether prediction errors differed meaningfully between models, we calculated Cohen's *d* effect sizes (Cohen, 2013) using the *cohen.d()* function of the package *effsize* (Torchiano, 2020) for all pairwise comparisons of absolute relative errors on the independent testing dataset. Cohen's *d* expresses the difference between the means of two groups relative to their pooled standard deviation, providing a standardized, unitless measure of effect size:

$$d = \frac{\bar{x}_1 - \bar{x}_2}{s_p}, \text{ with } s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

where $\bar{x}_1$, $\bar{x}_2$ are group means, s_1 , s_2 are standard deviations, and n_1 , n_2 are sample sizes.

2.5 | Sensitivity of EIVs to plot area

To assess how much plots of different areas differ in their EIVs, we conducted pairwise comparisons within each survey of the mean of the nine ecological indicators. Similarly to species richness, we quantified their difference by taking the absolute of the relative error—calculated as $|(observed - predicted)/observed|$ —across all pairwise comparisons of plots and examined them in relation to the scaling factor. To contextualize this within-site variability, we compared it to the coefficient of variation in EIVs observed between sites at the

plot area of 4 m², which was consistently present across all included vegetation survey datasets.

3 | RESULTS

3.1 | Analysis of accumulation coefficients in nested-plot surveys

Overall, the power-law SAR $S = cA^z$ displayed a high fit to the species richness progression across subplots with 88% of the models having an $R^2 > 0.9$ (median $R^2 = 0.97$, $SD = 0.06$; $n = 1001$; [Figure 1](#)). The median z -value was of 0.17 ($SD = 0.10$; $n = 1001$; [Figure 1](#); [Figure S1](#)).

3.2 | Modelling of accumulation coefficients

In the second variant of the SAR-based models (1b), we predicted site-specific z -values to account for site-specific variation in the shape of the SAR. Overall, the models exhibited only moderate ability to capture variation in z across sites, with a coefficient of determination of $R^2 = 0.34$ ([Figure S2](#)). Vegetation cover, that is the proportion of a plot covered by vascular plants, emerged as the

most influential predictor by a considerable margin, resulting in the largest drop in R^2 when permuted. All other predictors contributed only minimally to predicting the accumulation coefficient. No single variable aside from vegetation cover accounted for a dominant share of the explained variance; all other variables contributed only marginally to predicting the accumulation coefficient ([Figure S3](#)). A simple linear regression revealed a significant negative relationship between vegetation cover and the accumulation coefficient z ($\beta = -0.0016 \pm 0.00009$, $t = -18.63$, $p < 0.001$; [Figure S4](#)), indicating that higher vegetation cover is consistently associated with a flatter SAR. The linear model explained only a quarter of the variance in z -values ($R^2 = 0.258$; $F(1, 999) = 347.2$, $p < 0.001$).

3.3 | Comparison of species richness scaling models

All models were assessed using an independent test dataset. Cohen's d effect sizes comparing models' relative prediction errors aggregated across all scaling factors were consistently modest ($|d| < 0.1$; [Table S1](#)). However, while prediction errors increased with scaling factor in all models, they differed in their initial error levels and in the rate at which errors increased ([Figure 2](#); [Figure S5](#)). At scaling factors below 5, SAR

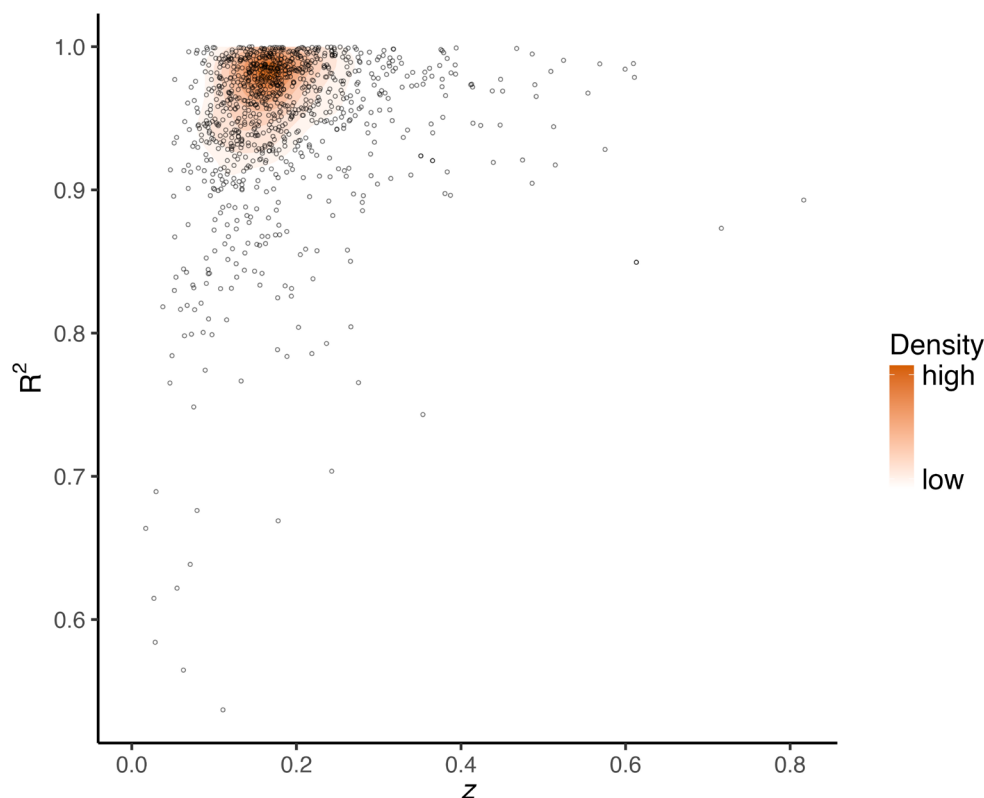


FIGURE 1 Relationship between accumulation coefficient z and power-law model fit. Density distribution of power-law model fits (R^2) of the species–area relationships of the nested-plot surveys, shown in relation to their corresponding accumulation coefficient (z). A simple power-law model of the form $S = cA^z$ was fitted to each survey using nonlinear least squares. Colour illustrates the distribution of fitted values, with darker shading indicating areas of higher point density. Most models (88%) achieved a fit of $R^2 > 0.9$, while the median z -value was 0.17.

models produced lower errors than RF models, with absolute relative errors ranging from 4.1% to 10.9% for SAR models and from 6.3% to 10.7% for RF models. At scaling factors above 5, RF models yielded lower error rates than SAR-based methods, with absolute relative errors ranging from 9.6% to 26.9% for RF models and from 9.9% to 38% for SAR models (Table S2). At small scaling factors, the SAR-based approach using a constant z -value (Model 1a) performed similarly to more complex alternatives, while its performance was markedly worsened at large scaling factors. The inclusion of environmentally fitted, site-specific z -values (Model 1b) led to modest improvements over the constant- z SAR baseline, especially at high scaling factors. Overall, differences among approaches were limited, with SAR models performing slightly better at small scaling factors and RF models showing a relative advantage at larger scaling factors.

3.4 | Sensitivity of EIVs to plot area

Average EIVs showed low within-site variation across plots of different areas (Figure 3). The absolute of the relative differences increased gradually with increasing scaling factor, starting around 1%

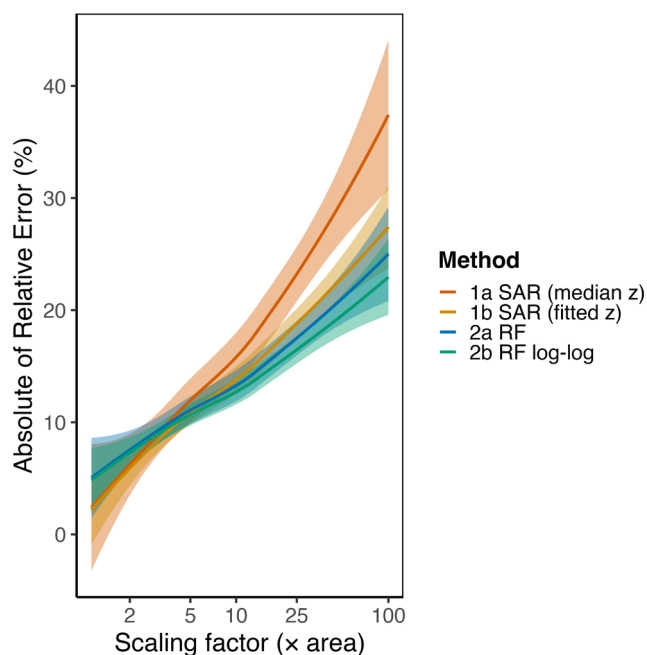


FIGURE 2 Species richness scaling across scaling factors with different methods. Absolute relative error (%) in species richness predictions in dependence of scaling factor, comparing different methods. This shows how scaling error increases with the scaling factor between the scaled area (A_2) and the original plot area (A_1). Scaling factors are expressed as absolute fold changes, regardless of whether an area was scaled up or down. Lines represent the trend in scaling error for four approaches: species–area relationship-based on a median z -value and fitted z -values, and random forest-based models, including the area parameters as covariate untransformed and logarithmically transformed. LOESS-smoothed curves (span = 1.2) show the mean trend, with shaded ribbons indicating standard error.

at small scaling factors (1–2) and reaching values between 3.4% and 7.9% at large scaling factors (25–100). Differences varied by ecological indicator. Soil aeration and temperature showed the largest increases in relative difference with increasing scaling factors, with 7.9%. In contrast, continentality and light availability exhibited the smallest within-site variation, remaining below 3.5% across all scaling levels. Overall, variation in mean indicator values due to differences in subplot area within a survey was markedly smaller than the variation observed between surveys at different sites at the common plot area of 4 m² (Table S3). Coefficients of variation of mean EIVs between sites ranged from 7.1% for continentality, 9.3% for moisture availability and 9.9% for light to 14.5% for moisture variability, 14.8% for pH, 18.1% for humus, 18.4% for nutrients and 20.9% for temperature, with soil aeration showing the highest variability (30.0%). The shape of the response curves was similar across indicators, showing a gradual increase in the relative difference in average EIVs between the plots within a survey as the scaling factor increased.

4 | DISCUSSION

Leveraging a comprehensive dataset of over 1000 nested-plot vegetation surveys from Swiss grasslands, spanning a 100-fold difference in plot area, we conducted the first comprehensive analysis aimed at the harmonization of vegetation indices of vegetation surveys of different plot area since the original publication of Arrhenius (1921). Our findings

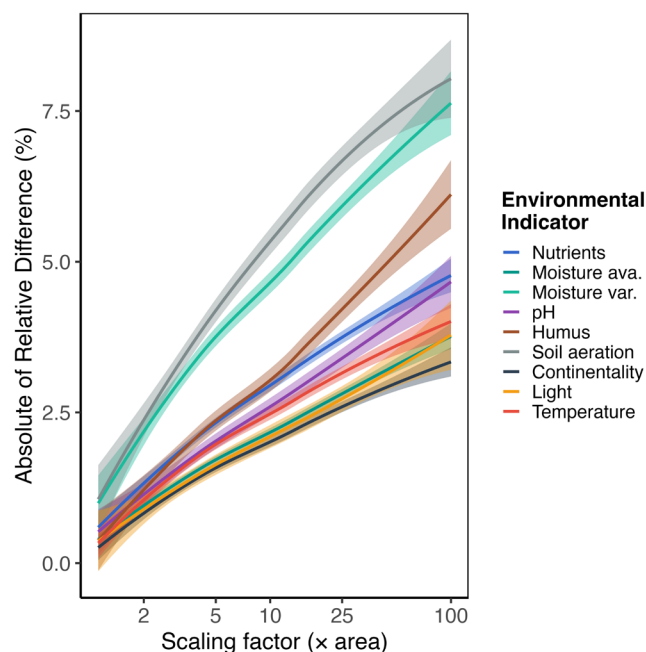


FIGURE 3 Ecological indicator values (EIVs) scaling across scaling factors. Absolute values of the relative differences in EIVs between pairs of vegetation plots, plotted as a function of the scaling factor. Each curve represents one ecological indicator. For each indicator, pairwise differences are calculated between two plot areas at the same site. LOESS-smoothed curves (span = 1.2) show the trend, with shaded ribbons indicating the standard error.

show that species richness can be reliably scaled, with small differences between approaches. At scaling factors below 5, error rates were relatively low, and SAR-based models achieved lower error rates than RF-based models. In contrast, at scaling factors above 5, errors increased overall, and RF-based models performed better than SAR-based models. Moreover, average EIVs remained robust across varying plot area, enabling a consistent characterization of the vegetation. Collectively, our results highlight that datasets containing vegetation survey of non-standardized plot areas can be effectively harmonized, and we propose a simple method to facilitate this process.

4.1 | A specific accumulation coefficient z for temperate grasslands

The classic power-law SAR model has a remarkably high fit across the majority of our surveys. With a median R^2 of 0.97 and 88% of models exceeding $R^2 > 0.9$, this finding confirms the robustness of the power-law formulation in capturing species accumulation patterns within nested-plot surveys. The median z -value of 0.17 (SD=0.10; $n=1001$) is lower than the 'canonical' $z=0.262$ proposed by Preston (1962) and reaffirmed by Wilson et al. (2012). However, this is consistent with empirical observations from non-isolated, species-rich habitats, such as European grasslands, where higher overall richness tends to flatten the SAR slope (Johnson et al., 1968; MacArthur & Wilson, 1967; Matthews et al., 2021). Similarly, a recent large-scale study of Palearctic grasslands by Dembicz, Dengler, Steinbauer, et al. (2021) found a median z of 0.21 (SD=0.10), but if limited to comparable plot areas and nesting configurations (1–100 m² and ≥ 4 subplots), the median z -value declined (median=0.18). Collectively, our findings align with the broader ecological theory that z tends to decline with increasing richness.

4.2 | Simple models reliably scale species richness

Our comparison of harmonization approaches revealed that at scaling factors below 5, all methods—whether SAR-based or using RF—produced accurate predictions, with average relative errors ranging between 4.1% and 10.9%. This suggests that even a simple SAR model using a fixed z -value might provide adequate harmonization in this range. In fact, while z -values showed some variability across sites (SD=0.10; Figure S1), our findings suggest that species richness scaling is relatively insensitive to these differences at low scaling factors. This further supports the practical utility of SAR-based approaches for harmonization, particularly when only small area adjustments are required. However, at scaling factors above 5 and especially above 10, error rates increased more steeply for SAR-based models, especially for the simple SAR model using a fixed z -value. In contrast, RF-based approaches maintained greater predictive accuracy. Interestingly, incorporating environmentally predicted z -values into SAR models (Model 1b) yielded only marginal improvements over a constant z -value. This suggests that although z varies across

sites, its prediction from ecological covariates is still constrained by unexplained variability or noise in the underlying relationships.

4.3 | EIVs are insensitive to plot area

We found consistently low within-site variation in average EIVs across different plot areas, indicating that EIVs are robust to changes in spatial grain. Relative differences increased only modestly with plot area—rarely exceeding 5%—supporting the use of EIVs as reliable descriptors of site conditions. This is particularly relevant for vegetation databases integrating surveys with differing plot areas, suggesting that variation in sampling grain does not introduce substantial bias into EIV-based site comparisons, at least within the tested range (1–100 m²). In contrast, between-site variation was almost an order of magnitude greater than within-site differences at small scaling factors (Table S3), emphasizing that landscape-scale ecological gradients exert a much stronger influence on community composition than sampling grain. These results are consistent with Ewald (2003) and Boch et al. (2022), who found that mean EIVs are largely robust to differences in survey completeness and observer effects. Similarly, Scherrer et al. (2017) and Rumpf et al. (2025) showed that EIVs capture long-term environmental trends more reliably than species composition. Although EIVs were historically applied primarily in Central Europe (Ellenberg, 1974; Landolt, 1977), their use has been shown to be robust across large geographic ranges (Hedwall et al., 2019). Moreover, recent efforts to extend and harmonize Ellenberg-type EIVs at the European scale (Tichý et al., 2023) further highlight their value for comparison of vegetation surveys across large spatial extents.

We note, however, that this robustness may not hold at larger spatial scales or in environments with pronounced small-scale heterogeneity. In such contexts, variation in abiotic conditions, management or microhabitat diversity can generate stronger within-site differences in species composition and, consequently, in EIVs. For example, larger plots may integrate multiple vegetation types or environmental gradients, while highly heterogeneous grasslands can host sharp transitions in soil moisture, fertility or disturbance regimes over just a few metres. Under these conditions, average EIVs may not provide a robust descriptor of site conditions. Nevertheless, within the spatial scales tested here (1–100 m²), EIVs proved generally insensitive to plot area, suggesting that they capture underlying environmental conditions in a way that is robust to typical vegetation survey design. Taken together, EIVs are not only robust to temporal noise and species turnover but also to spatial grain, reinforcing their utility for ecological monitoring, retrospective analysis and cross-study comparisons.

4.4 | Practical considerations for scaling of species richness

There are several important considerations when interpreting these results. Because species richness is inherently scale-dependent, we advise careful consideration when harmonizing richness values

across large differences in plot area—particularly when scaling spans one or more orders of magnitude (scaling factor >10). Such harmonization approaches are most appropriate for comparing vegetation surveys that differ only modestly in plot area. To ensure the reliability of analyses and assessments, we recommend constraining richness scaling efforts to a narrow range of commonly used plot areas within the vegetation type of interest (e.g. 9–28 m²), where SARs are generally more stable and predictable. We selected an accumulation coefficient of $z=0.17$, aligning with both the median value observed in our surveys and the z -value corresponding to the most frequently occurring plot area (i.e. 25 m²; see Figure S6) in prominent vegetation survey databases, such as the National Data and Information Center on the Swiss Flora (Info Flora, <https://www.infoflora.ch>), the European Vegetation Archive (Chytrý et al., 2016) or ReSurveyGermany (Jandt et al., 2022). This selection facilitates species richness comparisons at a common scale and minimizes the need for more extreme scaling. Moreover, we strongly encourage the inclusion of at least two different plot areas in survey designs which requires minimal additional effort but offers substantial benefits, enabling ecological insights and facilitating consistent data harmonization across studies.

4.5 | Area-based harmonization is most practical

While our study focuses on standardizing the effects of plot area on species richness and EIVs, this is only one of several possible approaches to compare surveys, especially relating to their species diversity. Individual-based rarefaction, which standardizes richness by the number of individuals sampled (Gotelli & Colwell, 2001), and coverage-based rarefaction, which standardizes by sample completeness rather than sample size (Chao & Jost, 2012; Roswell et al., 2021), are widely used alternatives. Both are conceptually powerful, but their application to vegetation surveys is limited because they depend on clearly defined and countable individuals—something rarely feasible and/or available for perennial and clonal plants. Using cover or biomass estimates as proxies could, in principle, extend these methods, but such measures are not equivalent to individuals: they distort abundance structures (e.g. a single tussock may equal many seedlings in cover) and do not reflect detectability. Coverage-based approaches are further complicated by observer bias, since completeness estimates rely heavily on rare species—the very component most inconsistently recorded.

Rarefaction corrects for uneven sampling effort only under the assumption that all individuals have an equal chance of being recorded. This works where differences among samples arise from random variation in sample size, but vegetation surveys differ in two critical respects. First, area is not simply equivalent to effort: larger plots encompass not only more individuals but also more microhabitats, heterogeneity and spatial turnover. As a result, SARs reflect ecological processes beyond sampling intensity, a point also emphasized in broader discussions showing that standardized sampling alone cannot fully capture biodiversity patterns across scales

(Chase et al., 2018; Chase & Knight, 2013). Second, observer bias in vegetation data is systematic rather than random. Inconspicuous species may be consistently overlooked, the abundance of dominant species misestimated, and difficult groups inconsistently identified. Rarefaction cannot correct these errors, since resampling a biased dataset merely reproduces the bias.

In contrast, biodiversity metrics are strongly dependent on spatial grain and extent (Chase et al., 2019). Harmonizing surveys by area therefore directly addresses a dominant source of variation and, importantly, enables integration of the large body of vegetation survey data collected across broad spatial and temporal scales, allowing the inclusion of many historical surveys and/or those lacking abundance data. Observer bias, by contrast, must be mitigated through careful field protocols, observer training, paired resurveys to quantify pseudoturnover, and statistical controls such as including observer identity as a random effect. While alternative frameworks may eventually be adapted for plants, area-based scaling—combined with explicit attention to observer effects—currently represents the most practical and broadly applicable route for comparative analyses of plant diversity.

5 | CONCLUSIONS

Prior to our study, efforts to harmonize vegetation survey data across varying plot areas were limited and lacked quantitative validation. Our results address this gap by providing empirical evidence that vegetation indices derived from homogeneous vegetation types can be effectively standardized across different plot areas using relatively simple methodologies. This facilitates comparisons and integrations of ecological data collected at diverse spatial scales. In the future, the assessment of the harmonization potential in other ecosystems and of other vegetation indices, particularly those based on species abundance, should be tested. Collectively, our findings offer practical tools for harmonizing vegetation survey datasets, enabling more consistent biodiversity assessments across spatial and temporal scales. By mitigating biases arising from traditional or regional differences in plot survey areas, we aim to contribute to more comprehensive and comparable ecological studies across space and time.

AUTHOR CONTRIBUTIONS

Marco Barandun, Eva Knop and Manuel K. Schneider conceived the study. Marco Barandun conducted the analysis and led the manuscript writing with critical input of all co-authors. Marco Barandun, Caren M. Pauler, Gisela Lüscher and Antoine Guisan provided the field data. Marco Barandun led the manuscript writing with critical input of all co-authors.

ACKNOWLEDGEMENTS

We sincerely thank Tobia Stöckli and Linus Koch for the help in the field, Dirk N. Karger, Niklaus E. Zimmermann, Pascal A. Niklaus, David Storch, Victor Boussange, Jeanne Portier and Sarina Daniöth for

insightful comments, the two anonymous reviewers for their detailed and constructive suggestions, the Associate Editor for the careful handling of the manuscript and all farmers for granting access to their land. This research was funded by the project 'G4B: Grasslands for biodiversity: supporting the protection of the biodiversity-rich grasslands and related management practices in the Alps and Carpathians' funded by Biodiversa+, the European Biodiversity Partnership under the 2021–2022 BiodivProtect joint call for research proposals, co-funded by the European Commission (GA N°101052342) and with the funding organizations SNSF, DFG, NCN, PROV BZ, SAS and UEFISCDI through the Swiss National Science Foundation grant to M. K. Schneider (number: 31BD30_209612). Open Access funding was provided by Agroscope via the Consortium of Swiss Academic Libraries. Substantial parts of the dataset used were collected in the context of a previous Swiss National Science Foundation grant to A. Guisan (number: 110000), of a fellowship of the German Academic Scholarship Foundation to C. M. Pauler and in the FP7 project BioBio funded by the European Union (227161).

CONFLICT OF INTEREST STATEMENT

Manuel K. Schneider is leading the German section of the Swiss Grassland Society (non-profit organization to promote sustainable grassland use). There are no further conflicts of interest.

PEER REVIEW

The peer review history for this article is available at <https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/1365-2745.70232>.

DATA AVAILABILITY STATEMENT

All data and code used to run the models, conduct the analyses, and reproduce the results and figures presented in this study are openly available on Zenodo at <https://doi.org/10.5281/zenodo.17778077> (Barandun et al., 2025).

ORCID

Marco Barandun  <https://orcid.org/0000-0003-4951-665X>

Caren M. Pauler  <https://orcid.org/0000-0002-7872-2711>

Gisela Lüscher  <https://orcid.org/0000-0001-8515-7880>

Antoine Guisan  <https://orcid.org/0000-0002-3998-4815>

Eva Knop  <https://orcid.org/0000-0001-9402-2216>

Manuel K. Schneider  <https://orcid.org/0000-0002-3842-2618>

REFERENCES

- Arrhenius, O. (1921). Species and area. *The Journal of Ecology*, 9, 95.
- Barandun, M., Pauler, C., Lüscher, G., Guisan, A., Knop, E., & Schneider, M. K. (2025). Species richness at different plot areas in Swiss grasslands [data set]. *Zenodo*. <https://doi.org/10.5281/zenodo.17778077>
- Boch, S., Küchler, H., Küchler, M., Bedolla, A., Ecker, K. T., Graf, U. H., Moser, T., Holderegger, R., & Bergamini, A. (2022). Observer-driven pseudoturnover in vegetation monitoring is context-dependent but does not affect ecological inference. *Applied Vegetation Science*, 25, e12669.
- Boussange, V., Wuyts, B., Brun, P., Malle, J. T., Midolo, G., Portier, J., Sanchez, T., Zimmermann, N. E., Axmanová, I., Bruehlheide, H., Chytrý, M., Kambach, S., Lososová, Z., Večeřa, M., Biurrun, I., Ecker, K. T., Lenoir, J., Svenning, J.-C., & Karger, D. N. (2025). *Multi-scale species richness estimation with deep learning*. arXiv, 2507.06358.
- Braun-Blanquet, J. (1928). *Pflanzensoziologie*. Springer Berlin Heidelberg.
- Cazzolla Gatti, R., Reich, P. B., Gamarra, J. G. P., Crowther, T., Hui, C., Morera, A., Bastin, J. F., de-Miguel, S., Nabuurs, G. J., Svenning, J. C., Serra-Diaz, J. M., Merow, C., Enquist, B., Kamenetsky, M., Lee, J., Zhu, J., Fang, J., Jacobs, D. F., Pijanowski, B., ... Liang, J. (2022). The number of tree species on earth. *Proceedings of the National Academy of Sciences*, 119, e2115329119.
- Chao, A., & Jost, L. (2012). Coverage-based rarefaction and extrapolation: Standardizing samples by completeness rather than size. *Ecology*, 93, 2533–2547.
- Chase, J. M., & Knight, T. M. (2013). Scale-dependent effect sizes of ecological drivers on biodiversity: Why standardised sampling is not enough. *Ecology Letters*, 16, 17–26.
- Chase, J. M., McGill, B. J., McGlinn, D. J., May, F., Blowes, S. A., Xiao, X., Knight, T. M., Purschke, O., & Gotelli, N. J. (2018). Embracing scale-dependence to achieve a deeper understanding of biodiversity and its change across communities. *Ecology Letters*, 21, 1737–1751.
- Chase, J. M., McGill, B. J., Thompson, P. L., Antão, L. H., Bates, A. E., Blowes, S. A., Dornelas, M., Gonzalez, A., Magurran, A. E., Supp, S. R., Winter, M., Bjorkman, A. D., Bruehlheide, H., Byrnes, J. E. K., Cabral, J. S., Elahi, R., Gomez, C., Guzman, H. M., Isbell, F., ... O'Connor, M. (2019). Species richness change across spatial scales. *Oikos*, 128, 1079–1091.
- Chytrý, M., Hennekens, S. M., Jiménez-Alfaro, B., Knollová, I., Dengler, J., Jansen, F., Landucci, F., Schaminée, J. H. J., Aćić, S., Agrillo, E., Ambarli, D., Angelini, P., Apostolova, I., Attorre, F., Berg, C., Bergmeier, E., Biurrun, I., Botta-Dukát, Z., Brisse, H., ... Yamalov, S. (2016). European vegetation archive (EVA): An integrated database of European vegetation plots. *Applied Vegetation Science*, 19, 173–180.
- Chytrý, M., & Otýpková, Z. (2003). Plot sizes used for phytosociological sampling of European vegetation. *Journal of Vegetation Science*, 14, 563–570.
- Cohen, J. (2013). *Statistical power analysis for the behavioral sciences* (2nd ed.). Routledge.
- Connor, E. F., & McCoy, E. D. (1979). The statistics and biology of the species-area relationship. *The American Naturalist*, 113, 791–833.
- DeMalach, N., Saiz, H., Zaady, E., & Maestre, F. T. (2019). Plant species-area relationships are determined by evenness, cover and aggregation in drylands worldwide. *Global Ecology and Biogeography*, 28, 290–299.
- Dembicz, I., Dengler, J., Gillet, F., Matthews, T. J., Steinbauer, M. J., Bartha, S., Campos, J. A., de Frenne, P., Dolezal, J., García-Mijangos, I., Guarino, R., Güler, B., Kuzemko, A., Naqinezhad, A., Noroozi, J., Peet, R. K., Terzi, M., & Biurrun, I. (2021). Fine-grain beta diversity in Palaearctic open vegetation: Variability within and between biomes and vegetation types. *Vegetation Classification and Survey*, 2, 293–304.
- Dembicz, I., Dengler, J., Steinbauer, M. J., Matthews, T. J., Bartha, S., Burrascano, S., Chiarucci, A., Filibeck, G., Gillet, F., Janišová, M., Palpurina, S., Storch, D., Ulrich, W., Aćić, S., Boch, S., Campos, J. A., Cancellieri, L., Carboni, M., Ciaschetti, G., ... Biurrun, I. (2021). Fine-grain beta diversity of Palaearctic grassland vegetation. *Journal of Vegetation Science*, 32, e13045.
- Dengler, J. (2009). Which function describes the species-area relationship best? A review and empirical evaluation. *Journal of Biogeography*, 36, 728–744.
- Dengler, J., Matthews, T. J., Steinbauer, M. J., Wolfrum, S., Boch, S., Chiarucci, A., Conradi, T., Dembic, I., Marcenò, C., García-Mijangos, I., Nowak, A., Storch, D., Ulrich, W., Campos, J. A., Cancellieri, L., Carboni, M., Ciaschetti, G., de Frenne, P., Dolezal, J., ... Biurrun, I.

- I. (2020). Species–area relationships in continuous vegetation: Evidence from Palaearctic grasslands. *Journal of Biogeography*, 47, 72–86.
- Descombes, P., Walthert, L., Baltensweiler, A., Meuli, R. G., Karger, D. N., Ginzler, C., Zurell, D., & Zimmermann, N. E. (2020). Spatial modelling of ecological indicator values improves predictions of plant distributions in complex landscapes. *Ecography*, 43, 1448–1463.
- Diekmann, M. (2003). Species indicator values as an important tool in applied plant ecology – A review. *Basic and Applied Ecology*, 4, 493–506.
- Divišek, J., & Chytrý, M. (2018). High-resolution and large-extent mapping of plant species richness using vegetation-plot databases. *Ecological Indicators*, 89, 840–851.
- Dufour, A., Gadallah, F., Wagner, H. H., Guisan, A., & Buttler, A. (2006). Plant species richness and environmental heterogeneity in a mountain landscape: Effects of variability and spatial configuration. *Ecography*, 29, 573–584.
- Ellenberg, H. (1974). Zeigerwerte der Gefäßpflanzen Mitteleuropas [Indicator values of vascular plants in Central Europe]. *Scripta Geobotanica*, 9, e13168.
- Engler, R., Randin, C. F., Vittoz, P., Czaka, T., Beniston, M., Zimmermann, N. E., & Guisan, A. (2009). Predicting future distributions of mountain plants under climate change: Does dispersal capacity matter? *Ecography*, 32(1), 34–45.
- Ewald, J. (2003). The sensitivity of Ellenberg indicator values to the completeness of vegetation relevés. *Basic and Applied Ecology*, 4, 507–513.
- Gotelli, N. J., & Colwell, R. K. (2001). Quantifying biodiversity: Procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters*, 4, 379–391.
- He, F., & Legendre, P. (2002). Species diversity patterns derived from species–area models. *Ecology*, 83, 1185–1198.
- Hedwall, P., Brunet, J., & Diekmann, M. (2019). With Ellenberg indicator values towards the north: Does the indicative power decrease with distance from Central Europe? *Journal of Biogeography*, 46, 1041–1053.
- Jandt, U., Bruelheide, H., Berg, C., Bernhardt-Römermann, M., Blüml, V., Bode, F., Dengler, J., Diekmann, M., Dierschke, H., Doerfler, I., Döring, U., Dullinger, S., Härdtle, W., Haider, S., Heinken, T., Horchler, P., Jansen, F., Kudernatsch, T., Kuhn, G., ... Wulf, M. (2022). ReSurveyGermany: Vegetation-plot time-series over the past hundred years in Germany. *Scientific Data*, 9, 631.
- Janišová, M., Michalcová, D., Bacaro, G., & Ghisla, A. (2014). Landscape effects on diversity of semi-natural grasslands. *Agriculture, Ecosystems & Environment*, 182, 47–58.
- Johnson, M. P., Mason, L. G., & Raven, P. H. (1968). Ecological parameters and plant species diversity. *The American Naturalist*, 102, 297–306.
- Juillerat, P., Bäumler, B., Bornand, C., Gyga, A., Jutzi, M., & Möhl, A. (2017). *Flora Helvetica Checklist 2017: Der Gefäßpflanzen der Schweiz = de la flore vasculaire de la Suisse = della flora vascolare della Svizzera*.
- Kalkhan, M. A., Stafford, E. J., & Stohlgren, T. J. (2007). Rapid plant diversity assessment using a pixel nested plot design: A case study in beaver meadows, Rocky Mountain National Park, Colorado, USA. *Diversity and Distributions*, 13, 379–388.
- Keil, P., & Chase, J. M. (2019). Global patterns and drivers of tree diversity integrated across a continuum of spatial grains. *Nature Ecology & Evolution*, 3, 390–399.
- Kuhn, M. (2008). Building predictive models in R using the caret package. *Journal of Statistical Software*, 28, 1–26.
- Külling, N., Adde, A., Fopp, F., Schweiger, A. K., Broennimann, O., Rey, P.-L., Giuliani, G., Goicolea, T., Petitpierre, B., Zimmermann, N. E., Pellissier, L., Altermatt, F., Lehmann, A., & Guisan, A. (2024). SWECO25: A cross-thematic raster database for ecological research in Switzerland. *Scientific Data*, 11, 21.
- Kunin, W. E., Harte, J., He, F., Hui, C., Jobe, R. T., Ostling, A., Polce, C., Šizling, A., Smith, A. B., Smith, K., Smart, S. M., Storch, D., Tjørve, E., Ugland, K. I., Ulrich, W., & Varma, V. (2018). Upscaling biodiversity: Estimating the species–area relationship from small samples. *Ecological Monographs*, 88, 170–187.
- Landolt, E. (1977). *Ökologische Zeigerwerte zur Schweizer Flora [Ecological indicator values for the Swiss flora]* (p. 64). Veröffentlichungen des Geobotanischen Instituts der Eidgenössischen Technischen Hochschule.
- Landolt, E., Bäumler, B., Ehrhardt, A., Hegg, O., Klötzli, F., Lämmli, W., Nobis, M., Schleppi, P., Tonteri, T., & Wolf, E. (2010). *Flora indicativa: Ökologische Zeigerwerte und biologische Kennzeichen zur Flora der Schweiz und der Alpen* (Vol. 2). völlig neu bearbeitete und erweiterte Auflage.
- Lenoir, J., Gégout, J. C., Dupouey, J. L., Bert, D., & Svenning, J.-C. (2010). Forest plant community changes during 1989–2007 in response to climate warming in the Jura Mountains (France and Switzerland): Plant community response to climate warming in the Jura Mountains. *Journal of Vegetation Science*, 21, 949–964.
- Lüscher, G., Amari, Y., Andriets, A., Angelova, S., Arndorfer, M., Bailey, D., Balázs, K., Bogers, M., Bunce, R. G. H., Choisis, J. P., Dennis, P., Díaz, M., Dyman, T., Eiter, S., Fjellstad, W., Fraser, M., Friedel, J. K., Garchi, S., Geijendorffer, I. R., ... Zanetti, T. (2016). Farmland biodiversity and agricultural management on 237 farms in 13 European and two African regions. *Ecology*, 97(6), 1625.
- MacArthur, R. H., & Wilson, E. O. (1967). *The theory of Island biogeography*. Princeton University Press.
- Matthews, T. J., Triantis, K. A., & Whittaker, R. J. (2021). *The species–area relationship: Theory and application* (1st ed.). Cambridge University Press.
- Mod, H. K., Chevalier, M., Luoto, M., & Guisan, A. (2020). Scale dependence of ecological assembly rules: Insights from empirical datasets and joint species distribution modelling. *Journal of Ecology*, 108, 1967–1977.
- Morzaria-Luna, L., Callaway, J., Sullivan, G., & Zedler, J. (2004). Relationship between topographic heterogeneity and vegetation patterns in a Californian salt marsh. *Journal of Vegetation Science*, 15, 523–530.
- Oostermeijer, J. G. B., & Van Swaay, C. A. M. (1998). The relationship between butterflies and environmental indicator values: A tool for conservation in a changing landscape. *Biological Conservation*, 86, 271–280.
- Pauler, C. M., Isselstein, J., Braunbeck, T., & Schneider, M. K. (2019). Influence of highland and production-oriented cattle breeds on pasture vegetation: A pairwise assessment across broad environmental gradients. *Agriculture, Ecosystems & Environment*, 284, 106585.
- Portier, J., Zellweger, F., Zell, J., Alberdi Asensio, I., Bosela, M., Breidenbach, J., Šebeň, V., Wüest, R. O., & Rohner, B. (2022). Plot size matters: Toward comparable species richness estimates across plot-based inventories. *Ecology and Evolution*, 12, e8965.
- Preston, F. W. (1962). The canonical distribution of commonness and rarity: Part I & II. *Ecology*, 43, 410–432.
- R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
- Randin, C. F., Jaccard, H., Vittoz, P., Yoccoz, N. G., & Guisan, A. (2009). Land use improves spatial predictions of mountain plant abundance but not presence-absence. *Journal of Vegetation Science*, 20(6), 996–1008.
- Rosenzweig, M. L. (2004). Applying species–area relationships to the conservation of species diversity. In *Frontiers of biogeography: New directions in the geography of nature* (pp. 325–343). Sinauer Associates.
- Roswell, M., Dushoff, J., & Winfree, R. (2021). A conceptual guide to measuring species diversity. *Oikos*, 130, 321–338.
- Rumpf, S. B., Buri, A., Grand, S., Randin, C. F., Tesson, S., Cianfrani, C., & Guisan, A. (2025). Independent trends of mountain vegetation and soil properties over 40 years of environmental change. *Journal of Vegetation Science*, 36, e70006.

- Sabatini, F. M., Jiménez-Alfaro, B., Jandt, U., Chytrý, M., Field, R., Kessler, M., Lenoir, J., Schrod, F., Wiser, S. K., Arfin Khan, M. A. S., Attorre, F., Cayuela, L., de Sanctis, M., Dengler, J., Haider, S., Hatim, M. Z., Indreica, A., Jansen, F., Pauchard, A., ... Bruehlheide, H. (2022). Global patterns of vascular plant alpha diversity. *Nature Communications*, 13, 4683.
- Scherrer, D., & Guisan, A. (2019). Ecological indicator values reveal missing predictors of species distributions. *Scientific Reports*, 9, 3061.
- Scherrer, D., Massy, S., Meier, S., Vittoz, P., & Guisan, A. (2017). Assessing and predicting shifts in mountain forest composition across 25 years of climate change. *Diversity and Distributions*, 23, 517–528.
- Scherrer, D., Mod, H. K., Pottier, J., Litsios-Dubuis, A., Pellissier, L., Vittoz, P., Götzenberger, L., Zobel, M., & Guisan, A. (2019). Disentangling the processes driving plant assemblages in mountain grasslands across spatial scales and environmental gradients. *Journal of Ecology*, 107, 265–278.
- Smart, S. M., Clarke, R. T., van de Poll, H. M., Robertson, E. J., Shield, E. R., Bunce, R. G., & Maskell, L. C. (2003). National-scale vegetation change across Britain; an analysis of sample-based surveillance data from the countryside surveys of 1990 and 1998. *Journal of Environmental Management*, 67, 239–254.
- Stohlgren, T. J., Falkner, M., & Schell, L. (1995). A modified-Whittaker nested vegetation sampling method. *Vegetatio*, 117, 113–121.
- Tichý, L., Axmanová, I., Dengler, J., Guarino, R., Jansen, F., Midolo, G., Nobis, M. P., van Meerbeek, K., Ačić, S., Attorre, F., Bergmeier, E., Biurrun, I., Bonari, G., Bruehlheide, H., Campos, J. A., Čarni, A., Chiarucci, A., Čuk, M., Čušterevska, R., ... Chytrý, M. (2023). Ellenberg-type indicator values for European vascular plant species. *Journal of Vegetation Science*, 34, e13168.
- Tjørve, E. (2003). Shapes and functions of species–area curves: A review of possible models. *Journal of Biogeography*, 30, 827–835.
- Tjørve, E., Kunin, W. E., Polce, C., & Calf Tjørve, K. M. (2008). Species–area relationship: Separating the effects of species abundance and spatial distribution. *Journal of Ecology*, 96, 1141–1151.
- Torchiano, M. (2020). *effsize: Efficient effect size computation*. <https://cran.r-project.org/web/packages/effsize/effsize.pdf>
- Večeřa, M., Divišek, J., Lenoir, J., Jiménez-Alfaro, B., Biurrun, I., Knollová, I., Agrillo, E., Campos, J. A., Čarni, A., Crespo Jiménez, G., Čuk, M., Dimopoulos, P., Ewald, J., Fernández-González, F., Gégout, J. C., Indreica, A., Jandt, U., Jansen, F., Kącki, Z., ... Chytrý, M. (2019). Alpha diversity of vascular plants in European forests. *Journal of Biogeography*, 46, 1919–1935.
- Von Oppen, J., Assmann, J. J., Bjorkman, A. D., Treier, U. A., Elberling, B., Nabe-Nielsen, J., & Normand, S. (2022). Cross-scale regulation of seasonal microclimate by vegetation and snow in the Arctic tundra. *Global Change Biology*, 28, 7296–7312.
- Wilson, J. B., Peet, R. K., Dengler, J., & Pärtel, M. (2012). Plant species richness: The world records. *Journal of Vegetation Science*, 23, 796–802.
- Zhang, J., Gillet, F., Bartha, S., Alatalo, J. M., Biurrun, I., Dembiczy, I., Grytnes, J. A., Jaunatre, R., Piech, R., van Meerbeek, K., Vynokurov, D., Widmer, S., Aleksanyan, A., Bhatta, K. P., Campos, J. A., Czortek, P., Dolezal, J., Essl, F., García-Mijangos, I., ... Dengler, J. (2021). Scale dependence of species–area relationships is widespread but generally weak in Palaearctic grasslands. *Journal of Vegetation Science*, 32, e13044.
- Zolotova, E., Ivanova, N., & Ivanova, S. (2022). Global overview of modern research based on Ellenberg indicator values. *Diversity*, 15, 14.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Figure S1. Distribution of z-values across nested-plot vegetation surveys.

Figure S2. Comparison of observed and predicted z values.

Figure S3. Variable importance in determining z.

Figure S4. Relationship between vegetation cover and accumulation coefficient z.

Figure S5. Scale dependence of differences in model performance.

Figure S6. Accumulation coefficient (z) across subplot areas.

Table S1. Pairwise Cohen's d effect sizes among the four scaling models.

Table S2. Mean absolute relative errors (%) of the four investigated models across scaling factor ranges.

Table S3. Mean absolute relative differences (%) of ecological indicator values (EIVs) across scaling factor ranges.

How to cite this article: Barandun, M., Pauler, C. M., Lüscher, G., Guisan, A., Knop, E., & Schneider, M. K. (2026). Scaling species richness: When vegetation surveys don't match in plot area. *Journal of Ecology*, 114, e70232. <https://doi.org/10.1111/1365-2745.70232>