Preslia 97: 541–566, 2025
doi: 10.23855/preslia.2025.541

Preslia

Spatial patterns of species richness in the Czech flora: effects of sampling intensity, environment and landscape history

Barbora Klímová[1,*], Jan Divíšek[1,2], Martin Večeřa[1], Milan Chytrý[1], Jiří Danihelka[1,3], Zdeněk Kaplan[3,4], Petr Novotný[5], Hana Skokanová[6], Jan Wild[3] & Irena Axmanová[1]

[1]Department of Botany and Zoology, Faculty of Science, Masaryk University, Kotlářská 2, CZ-61137 Brno, Czech Republic; [2]Department of Geography, Faculty of Science, Masaryk University, Kotlářská 2, CZ-61137 Brno, Czech Republic; [3]Institute of Botany, Czech Academy of Sciences, CZ-25243 Průhonice, Czech Republic; [4]Department of Botany, Faculty of Science, Charles University, Benátská 2, CZ-12801 Praha 2, Czech Republic; [5]Herbarium Collections, Faculty of Science, Charles University, Benátská 2, CZ-12800 Praha, Czech Republic; [6]Landscape Research Institute, Květnové nám. 391, CZ-25243 Průhonice, Czech Republic
[*]corresponding author: bklimova@mail.muni.cz

Abstract: Understanding the spatial patterns and drivers of species richness is crucial for biodiversity conservation. Using data from Pladias, a comprehensive botanical database of the Czech Republic, we mapped the richness of various plant species groups across grid cells in the country and examined the effects of current environmental conditions, current landscape structure, and historical landscape development. We also applied five methods to account for uneven sampling intensity and found that only rarefaction provided estimates of species richness independent of sampling intensity. Using spatial error models and a variation partitioning approach, we showed that plant richness at the country scale is predominantly driven by current environmental conditions. For overall species richness, as well as for native and threatened species richness, the most important factors were the proportion of carbonate bedrock and the level of climate moisture, while the heat sum in the growing season was crucial for naturalized alien species richness. Historical landscape development, especially the long-term continuity of forests and grasslands, significantly influenced the richness of all, native, and particularly threatened species. Human population density was positively related to all species groups, emerging as the most important variable for the richness of naturalized alien species. However, our study shows that uneven sampling intensity in the Pladias database may distort the effect of certain environmental factors, such as the heat sum in the growing season. These findings emphasize the importance of carefully considering the uneven intensity of sampling before analysing species richness and highlight the role of current environmental variables, current landscape structure, and historical landscape development in shaping plant species richness at the regional scale.

Keywords: Czech Republic, diversity, environment, historical landscape, landscape structure, rarefaction, sampling bias, species richness estimation, vascular plants

Received: 27 Jan 2025; Revised: 8 July 2025; Accepted: 9 July 2025; Published: 29 Aug 2025

Introduction

Species richness, the number of species within a defined region, is the most widely used and conceptually straightforward way of quantifying biodiversity. The spatial patterns and underlying factors influencing species richness across different temporal and spatial scales have been the subject of research for several decades (e.g. Gaston 2000, Slezák & Axmanová 2016, Večeřa et al. 2019, 2021, Němec et al. 2022) and are still of fundamental importance to disciplines such as biogeography, macroecology, community ecology, and nature conservation (Maestre 2004, Divíšek & Chytrý 2018), especially in the light of recently documented biodiversity decline (IPBES 2019).

Many studies have shown that environmental factors, such as climate, topography, and soil properties, play a crucial role in shaping plant species richness across landscapes (Chytrý et al. 2003, Field et al. 2008, Moeslund et al. 2013, Stein et al. 2014, Jonas et al. 2015). However, the relative importance of these factors depends strongly on the spatial scale and resolution of the study (Pearson & Dawson 2003, Zarzo-Arias et al. 2022). At global and continental scales and coarser spatial resolutions, climate is usually a major factor, while soil properties, disturbances, and biotic interactions become more influential at regional and local scales, particularly when finer spatial resolutions are considered (Meineri et al. 2012, Dubuis et al. 2013, Paniw et al. 2015, Huang et al. 2021). In addition, human activities significantly alter the environment and have a massive impact on species richness (Bergès et al. 2017, IPBES 2019), especially in regions with centuries of human influence, such as central Europe (Berglund 2011). Last but not least, numerous studies have shown that high species richness and the occurrence of threatened or dispersal-limited species are often associated with long-term continuity of land use and habitats in the landscape (e.g. Hájková et al. 2011, Dullinger et al. 2013, Bergès et al. 2017, Raduła et al. 2022). Nevertheless, few studies have directly investigated the combined effects of current environmental conditions, current landscape structure, and land-cover stability over time on plant species richness. The limited availability of historical land-cover maps has prevented such larger syntheses; therefore, only studies at the local scale have been conducted so far (e.g. Gustavsson et al. 2007, Zimmermann et al. 2010, Němec et al. 2022, Raduła et al. 2022), although historical land-cover dynamics can significantly influence plant richness at the regional scale (Pearson & Dawson 2003). One of the few examples of larger synthesis is the study by Midolo et al. (2025), who investigated the effect of historical land cover on grassland species richness in the borderland between Austria and the Czech Republic.

With the exponential increase in data availability on the occurrence of plant species (Wüest et al. 2019), there is a unique opportunity to analyse general patterns and mechanisms in the spatial distribution of plant species richness across different scales and habitat types (e.g. Dubuis et al. 2011, Divíšek & Chytrý 2018, Večeřa et al. 2019, 2021, Szymura et al. 2023). However, plant occurrence records have been collected for different purposes, focusing on flora mapping, nature conservation inventories or vegetation survey, using various methodologies and sampling intensities, which may potentially lead to various types of sampling bias (Engemann et al. 2015, Wild et al. 2019) and consequently to misinterpretations of species richness patterns (Yang et al. 2013, Wild et al. 2019, Zizka et al. 2020). Vegetation plots are often collected nonrandomly in space (Meineke & Daru 2021) and areas that have been surveyed more intensively tend to

appear richer in species than those surveyed less intensively. Consequently, a strong correlation between species richness and sampling intensity is often present, as demonstrated by Yang et al. (2013) and Engemann et al. (2015). To some extent, this pattern is natural and realistic, as species-rich areas tend to attract more survey effort. However, the overall effect of uneven sampling intensity on species richness patterns and the interpretation of their drivers remains poorly understood. Furthermore, potential sampling bias also depends on the spatial resolution, particularly on grid cell size when data from grid cell mapping are used in the analysis. According to Sousa-Baena et al. (2013) and Engemann et al. (2015), sampling bias tends to be stronger at finer spatial resolutions due to lower and more uneven data coverage compared to coarser spatial resolutions, and it is also more pronounced in undersampled regions (Yang et al. 2013).

The issue of varying sampling intensity and the need to correct for related biases is often inadequately addressed or oversimplified in species richness studies (e.g. Chytrý et al. 2021, Pyšek et al. 2022, Macek et al. 2023). Therefore, various approaches have been suggested to obtain more realistic estimates of species richness. These include: (i) nonparametric asymptotic estimators (e.g. Burnham & Overton 1978, 1979, Chao 1984, 1987, Wang & Lindsay 2005), which infer total species richness from the frequency of detected rare species; (ii) parametric asymptotic models (e.g. Soberón & Llorente 1993, Jiménez-Valverde & Lobo 2006), which extrapolate species richness based on sample-based accumulation curves; and (iii) parametric nonasymptotic rarefaction estimators (Sanders 1968), which estimate the expected number of species for a given sample size based on the number of individuals or samples. Performance of these estimators highly varies, and their accuracy depends on the shape of species-abundance data distribution and the sample size (Willie et al. 2012, Gwinn et al. 2015). For instance, Walther & Moore (2005) reviewed studies comparing the performance of species richness estimators and proposed that nonparametric estimators (mostly Chao and jackknife) outperform curve models or methods fitting species-abundance distribution, a conclusion later supported by Willie et al. (2012). However, most of these studies were conducted at the landscape scale and at fine spatial resolutions. In contrast, Engemann et al. (2015) tested six nonparametric and parametric methods (Margalef richness index, Chao1, second-order jackknife, bootstrapping resampling methods, Hill numbers, and rarefaction) to correct for the effect of sampling bias at the country scale and across various spatial resolutions, concluding that only rarefaction provided a suitable solution. Given the contradictory results and the limited number of studies focused on broader scales and coarser resolutions, this topic requires a cautious approach, and it is necessary to compare results from multiple methods when estimating species richness based on raw data.

The Czech Republic has a long tradition of botanical research (Danihelka et al. 2017), which has fostered the development of comprehensive national vegetation and floristic databases: the Czech National Phytosociological Database (CNPD; Chytrý & Rafajová 2003) and the Database of the Czech Flora and Vegetation (Pladias; Wild et al. 2019, Chytrý et al. 2021). At the same time, there is a good coverage of digitized maps of environmental variables (e.g. bedrock alkalinity; Chuman et al. 2014), information on fine-scale habitat distribution (Härtel et al. 2009, AOPK ČR 2024a) and historical topographic maps suitable for assessing past land-cover changes (spanning from the 1840s to the 2000s; TopoLandUse CZ database, Skokanová et al. 2012). This exceptional availability of data makes the Czech Republic an ideal model for conducting a comprehensive analysis

of the factors influencing plant species richness at a country-wide scale. However, although the Czech Republic is one of the best-surveyed countries in terms of plant species occurrence, several authors have noted that both floristic and vegetation records in the abovementioned databases are affected by uneven sampling intensity across the country (Chytrý & Rafajová 2003, Wild et al. 2019, Macek et al. 2023). If left unaddressed, this issue may result in distorted patterns of overall species richness and the representation of specific species groups, such as alien species.

In this study, we aim to (i) test different methods for reducing the effect of unequal sampling intensity on species richness patterns; (ii) compare the resulting richness patterns for all vascular plant species, and separately for native, threatened (Red List), and naturalized alien species; (iii) compare models of observed plant species richness in the Czech Republic, which is affected by unequal sampling intensity, with models of estimated richness, in which this effect has been reduced; and (iv) assess the relative importance of variables associated with current environmental conditions, current landscape structure, and historical landscape development that may influence species richness.

Methods

Study area

Our study area (Supplementary Fig. S1) covers the entire territory of the Czech Republic (78,867 km^2), with an altitudinal range from 115 to 1,603 m a.s.l., mean annual temperatures of 5.0–9.5 °C, and annual precipitation of 320–1,450 mm (Tolasz 2007). According to CORINE Land Cover (Copernicus 2018), 56.8% of the Czech Republic's land area in 2018 was covered by agricultural land, 32.9% by forests, 2.8% by open semi-natural vegetation types (e.g. natural grasslands, moors, heathlands and bare rock), 6.7% by artificial surfaces, 0.7% by water bodies, and 0.1% by wetlands.

Significant changes in land cover patterns in the study area began with the Industrial Revolution in the first half of the 19th century. The area of forests, which had been at its lowest proportion before this period, gradually increased. Cities expanded, and the extent of agricultural land peaked, although it began to steadily decline from the first half of the 20th century. During the communist regime (1948–1989), the industrialization of agriculture led to the consolidation of large blocks of arable land. The period after 1989 has been characterized by increased urbanization, but also by the transformation of less fertile arable land to meadows and pastures, and improvements in landscape management and nature conservation (Bičík et al. 2015, Sychrová et al. 2024).

Species records

The vascular-plant occurrences were obtained from the Pladias database (https://www.pladias.cz; Wild et al. 2019, Chytrý et al. 2021), which contains 13.6 million plant occurrence records assigned to 2,716 grid cells with a size of 5 × 3 minutes (~ 6.0 km × 5.5 km, i.e. 33 km^2). The Pladias database contains data from both floristic recordings and vegetation plots.

Taxonomic concepts and nomenclature follow Kaplan et al. (2019). For our analyses, we only considered taxa at the species level but merged some large and taxonomically

complicated groups into aggregates (e.g. *Rubus fruticosus* agg.). This was done to ensure that the number of taxa we analysed (hereafter 'species' for simplicity) was not influenced by variations in the level of taxonomic resolution across grid cells. For each grid cell, we calculated the number of (i) all species, (ii) native species (including native threatened species, see further), (iii) threatened species included in the Red List of Vascular Plants of the Czech Republic (categories C1–C4; Grulich 2017), and (iv) naturalized alien species, including invasive species (Pyšek et al. 2022). Casual alien species (Richardson et al. 2000), which do not form self-sustaining populations and whose distribution is mostly driven by local introductions or escapes from cultivation, were not considered in our analyses. Grid cells localized at the borders of the Czech Republic, with less than 50% of their area within the country (348 cells), were excluded to avoid potential bias, leaving 2,368 grid cells for the analyses.

For each grid cell, we also calculated the total number of records (the sum of individual records of all species occurring in the grid cell) as a proxy for sampling intensity, which was used to estimate species richness.

Selection of species richness estimator

To obtain estimates of species richness that are independent of sampling intensity (Fig. 1A), we used the following three approaches:

(i) Nonparametric asymptotic estimators. We used three nonparametric asymptotic estimators: Chao1, jackknife, and the penalized nonparametric maximum likelihood estimator under a Poisson mixture model (PNPMLE), all available in the SPECIES R package version 1.2 (Wang 2011). These estimators use information about species frequency in the sample and offer different approaches for considering rare species. Chao1 (Chao 1984) provides lower-bound estimates of species richness, offering a minimum expected value for the community's richness. This estimator uses rare species, specifically singletons (species represented by exactly one record in the grid cell) and doubletons (species represented by two records in the grid cell), to infer the number of unobserved species. The jackknife estimator (Burnham & Overton 1978, 1979) reduces estimator bias by iteratively removing subsets of data and recalculating the estimator (here, the number of species per grid cell) with the reduced sample. We utilized the default jackknife order specified by the argument k equal to 5 and set the confidence interval coverage to 0.95. The PNPMLE estimator (Wang & Lindsay 2005) is a nonparametric maximum likelihood-based approach that improves estimate's stability by incorporating a quadratic penalty function into the conditional likelihood. The PNPMLE estimator is calculated for several subsets selected by a multinomial-based bootstrap, which is also used to construct a confidence interval. We kept the parameter t, defining the threshold for relatively less abundant species in the Poisson mixture estimation, at its default value ($t = 15$), and set C to 1 to calculate the confidence interval.

(ii) Parametric linear regression. We employed ordinary least squares (OLS) regression to model the relationship between the log-transformed number of species and the log-transformed number of records per grid cell. Using this model, we predicted the number of all species for 209 records per grid cell, which was the minimum number of records across grid cells in our data, excluding outlier values.

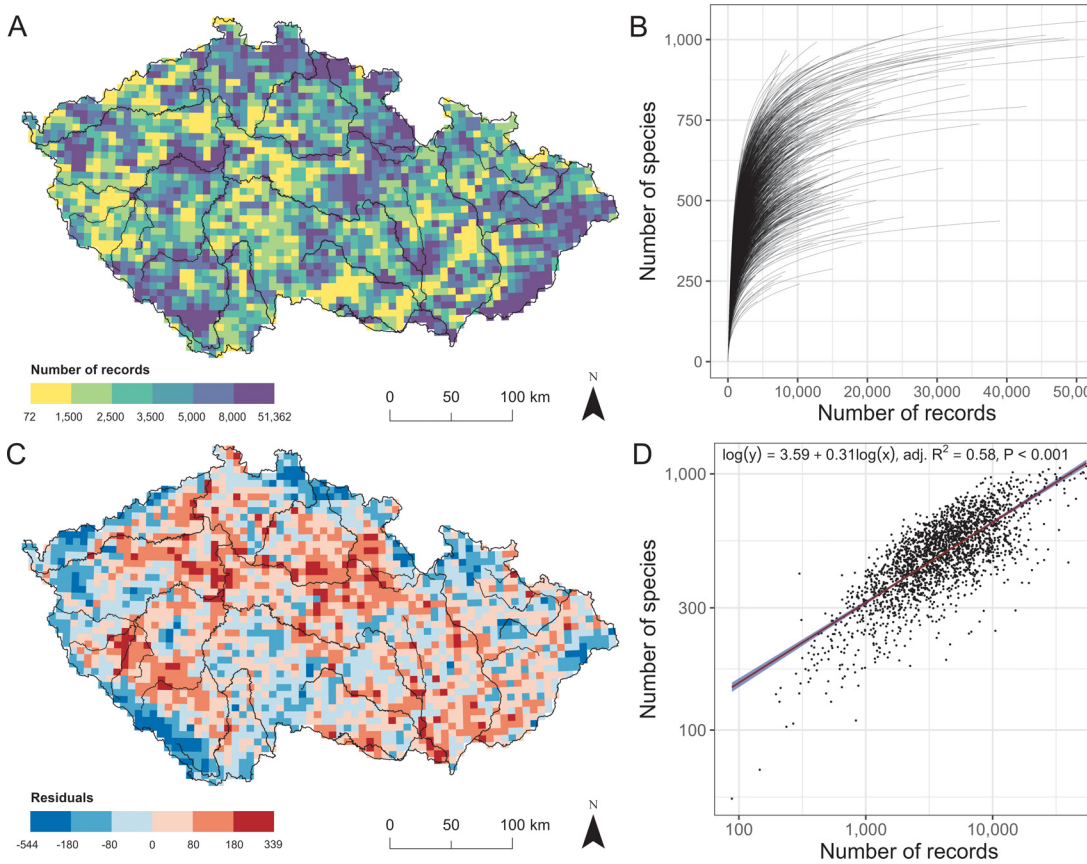

Fig. 1. (A) The number of records of vascular plant species per grid cell (~6.0 km × 5.5 km) in the Pladias database; (B) Rarefaction curves for each grid cell in the Pladias database; (C) Residuals from the linear regression of the log-transformed number of records and the log-transformed number of species for each grid cell in the Pladias database. Positive residuals (red) indicate a higher number of observed species than expected from the sampling intensity based on the regression model, while negative residuals (blue) indicate a lower number of observed species than expected; (D) Relationship between the log-transformed number of records (x-axis) and the log-transformed number of all species (y-axis), with a fitted linear trend line and 95% confidence interval bands.

(iii) Rarefaction. We also used the rarefaction method (Sanders 1968), a statistical interpolation technique involving the thinning of a reference sample (here, the number of records per grid cell) by randomly selecting a subset of individual records and subsequently calculating species richness. In each grid cell, we conducted 1,000 random selections of records (209 records for an overall richness estimate, Fig. 1B; 173 for native, 100 for threatened, and 100 for naturalized alien species) to calculate the mean number of observed species. These thresholds corresponded to the number of records available in the grid cell with the lowest number of records for each species group (excluding outliers).

To evaluate the performance of the above-mentioned methods, we used Spearman rank correlation (rho) and tested the relationship between the estimated number of species and the number of records. We acknowledge that some correlation between species

richness and sampling intensity is naturally expected, as species-rich areas tend to attract more intensive survey effort. Nevertheless, we selected the method with the weakest correlation in order to compare two contrasting approaches for analysing how spatial patterns of species richness and their drivers can be influenced by sampling intensity. We considered: (i) observed species richness, which is expected to be strongly affected by sampling intensity, and (ii) estimated richness, for which this effect has been substantially reduced. Therefore, only the estimation method that showed no statistically significant ($P < 0.05$) correlation with sampling intensity was included in further analysis of factors affecting plant species richness.

Drivers of spatial patterns of species richness

We selected variables that we hypothesized to significantly influence the spatial pattern of vascular plant species richness across the country, reflecting current environmental conditions, current landscape structure, and historical landscape development. Spearman correlation (Supplementary Fig. S2) was used to assess the relationship between each pair of variables. If two variables exhibited a correlation of rho > 0.8, only the one that was expected to have an ecologically more meaningful effect on regional differences in species richness was retained. This selection resulted in 12 variables used for further analysis (see below and Supplementary Table S1, and Supplementary Fig. S3).

To describe current environmental conditions (Supplementary Fig. S3A), we selected (A) the climate moisture index ($kg \cdot m^{-2} \cdot month^{-1}$; cmi) and (B) the heat sum in the growing season, also known as the growing degree days heat sum above 10 °C (°C; gdd10), from the CHELSA-BIOCLIM+ (Brun et al. 2022) at a spatial resolution of 30 arc seconds (~1 km). These variables serve as proxies of the overall energy available for plant growth and water availability. We also included (C) soil pH from the LUCAS database (Ballabio et al. 2019) at a spatial resolution of 500 m. For all these variables (A–C), we calculated the mean value in each grid cell. Moreover, we incorporated (D) the proportion of carbonate bedrock in each grid cell (%), obtained from the bedrock alkalinity map of the Czech Republic (Chuman et al. 2014) at a scale of 1:50,000.

The current landscape structure (Supplementary Fig. S3B) was described by two variables based on habitat mapping: (E) the proportion of natural habitats (%; aggregated at the second highest hierarchical level of the national habitat classification, e.g. A4 – subalpine tall-herb vegetation or L1 – alder carrs; Chytrý et al. 2010), considering only habitat patches where representativeness in terms of species composition was assessed as V (distinct, undoubtedly classifiable) or P (transitional habitat with a significant occurrence of characteristic species from two or more natural habitats; Lustyk 2016), and (F) the Shannon index (Shannon 1948) for natural habitats (aggregated at the same level as in E), reflecting the representation of different habitats within each grid cell. These variables were calculated based on the habitat mapping layer of the Czech Republic at a scale of 1:10,000 (AOPK ČR 2024a). Additionally, we calculated (G) the Shannon index (Shannon 1948) for land cover categories, based on the Czech topographic base map from 2002–2006, using the TopoLandUse CZ database (Skokanová et al. 2012). This database distinguishes eight land cover categories (arable land, grasslands, orchards, vineyards and hop fields, forests, water bodies, built-up areas, and other areas, including mining areas or dump sites). We also incorporated (H) human population density ($people \cdot km^{-2}$)

that we calculated as the total human population per each grid cell, based on the data from the 2021 Census of Population, Houses and Apartments (Czech Statistical Office 2021; at a 250 m spatial resolution), relative to the area of the grid cell.

To integrate the historical landscape development (Supplementary Fig. S3C) into the species richness models, we used the TopoLandUse CZ database (Skokanová et al. 2012), which was created by digitizing historical topographic maps for five time periods: 1840s (2nd Austrian Military Survey maps), 1870s (3rd Austrian Military Survey maps), 1950s, 1990s (Czechoslovak military topographic maps from 1952–1956 and 1988–1995), and 2000s (Czech topographic base map from 2002–2006). The period of the 2000s was used to depict the current landscape structure (see previous paragraph). Based on this data, we calculated the stable area of (I) grasslands, (J) forests, and (K) built-up areas as the proportion (%) of land in each grid cell that has not changed its land cover type (i.e. grassland, forest, or built-up area) since the 1840s. Finally, we considered (L) landscape dynamics, which was calculated using the Ružička index (Ružička 1958; a quantitative version of the Jaccard index). For each grid cell, we created a matrix of five time periods × eight land cover categories, where the values represented the proportion of each land cover category in a given period (% of total area). Applying the Ružička index to this data resulted in a dissimilarity matrix. The dissimilarities between each pair of consecutive periods were averaged to obtain a single value representing the historical landscape dynamics in each grid cell (the higher the value, the higher the historical landscape dynamics). All calculations and data processing were performed using ArcGIS 10.8 (ESRI 2020) and R 4.2.2 software (R Core Team 2024).

Statistical analyses

To assess the effects of the 12 variables described above (Supplementary Table S1) on the spatial patterns of both the observed and estimated species richness, we employed spatial error models (SEMs), implemented with the errorsarlm function of the spdep R package version 1.3.8 (Bivand 2022). We chose this method because our data exhibited spatial autocorrelation, and preliminary analyses using ordinary least squares (OLS) regression indicated that a statistically significant ($P < 0.05$) spatial autocorrelation persisted in the residuals of these models (results not shown). For all species and each species group (i.e. native, threatened, and naturalized alien species), we constructed models in which the dependent variable (y_i) was represented by either observed or estimated species richness, and the independent variables (x_i) included all 12 predictors simultaneously. To ensure comparability between the SEMs for observed and estimated richness, we included only those grid cells for which a rarefied number of species was calculated: 2,364 grid cells for all and native species; 1,895 for threatened species; and 1,925 for naturalized alien species. Additionally, we created an alternative set of models of observed species richness in which we accounted for uneven sampling intensity by including the log-transformed number of records as an additional explanatory variable. Results of these models are available in Supplementary Fig. S5. In each model, human population density, the proportion of carbonate bedrock, stable area of forests, grasslands, and built-up areas were log-transformed using the log1p function (R Core Team 2024), which calculates the natural logarithm of each predictor plus 1. All explanatory variables were standardized to zero mean and unit variance. The SEMs use the following equation:

$$y_i = \beta_0 + x_i\beta + \lambda\omega_i\xi_i + \varepsilon_i$$

where β represents the coefficients to be estimated, x_i is the predictor variable, $\lambda\omega_i\xi_i$ is a spatially dependent error term, and ε_i indicates a spatially uncorrelated error term. The parameter λ determines the degree of correlation among the errors based on the spatial weight matrix ω_i.

The spatial weight matrix, used to account for spatial autocorrelation in SEMs, was constructed based on Moran's correlograms calculated for each dependent variable using the correlog function in the pgirmess R package version 2.0.2 (Giraudoux 2023). In each correlogram, the distance at which the spatial autocorrelation decreased to zero was recorded. Grid cells within this distance were considered neighbouring when constructing the spatial weight matrix using the dnearneigh and nb2listw functions from the spdep R package version 1.3.8 (Bivand 2022). For each model, we extracted summary statistics, including R^2, using the glance function from the broom R package version 1.0.3 (Robinson et al. 2023). Finally, the residuals of these models were tested for spatial autocorrelation using the global Moran's index (Moran 1948).

To identify the relative importance of the three groups of explanatory variables (i.e. current environmental conditions, current landscape structure, and historical landscape development) on both observed and estimated species richness, we employed the variation partitioning approach (Borcard et al. 1992) as implemented in the varpart function from the vegan R package version 2.6.4 (Oksanen et al. 2022). To ensure comparability, we included the same grid cells for both observed and estimated richness in the variation partitioning models (see above for the number of grid cells for different species groups). To account for potential nonlinear relationships, we also incorporated quadratic and cubic functions of each explanatory variable. All polynomial functions were retained in the variation partitioning models without further selection to capture the full range of possible nonlinear effects. The results were visualized using Venn diagrams that show the percentage of variation explained by each group of explanatory variables. When interpreting these diagrams, it is important to note that negative shared variation can occur when a group of variables has close-to-zero correlation with the dependent variable but is at the same time correlated with another group of variables that is strongly correlated with the dependent variable (Azen & Budescu 2003). Negative shared variation can also arise when two strongly correlated groups of variables exert opposing effects on the dependent variable, one positive and the other negative (Legendre & Legendre 1998).

Results

Species richness patterns and sampling intensity

Using the number of records per grid cell in the Pladias database (Fig. 1A) as a measure of sampling intensity, we found that grid cells within protected areas (e.g. Český kras Protected Landscape Area (PLA), Krkonoše National Park (NP), Křivoklátsko PLA, Šumava NP, Bílé Karpaty PLA; AOPK ČR 2024b, c) and those situated near large cities (e.g. Prague and Brno) had higher sampling intensity (see Supplementary Fig. S1 for the locations of these areas). In contrast, grid cells that are less attractive in terms of natural values and further away from large cities were sampled less intensively (e.g. the Dyjsko-svratecký úval lowland, parts of the Krušné Mts, and the Plzeňská pahorkatina uplands).

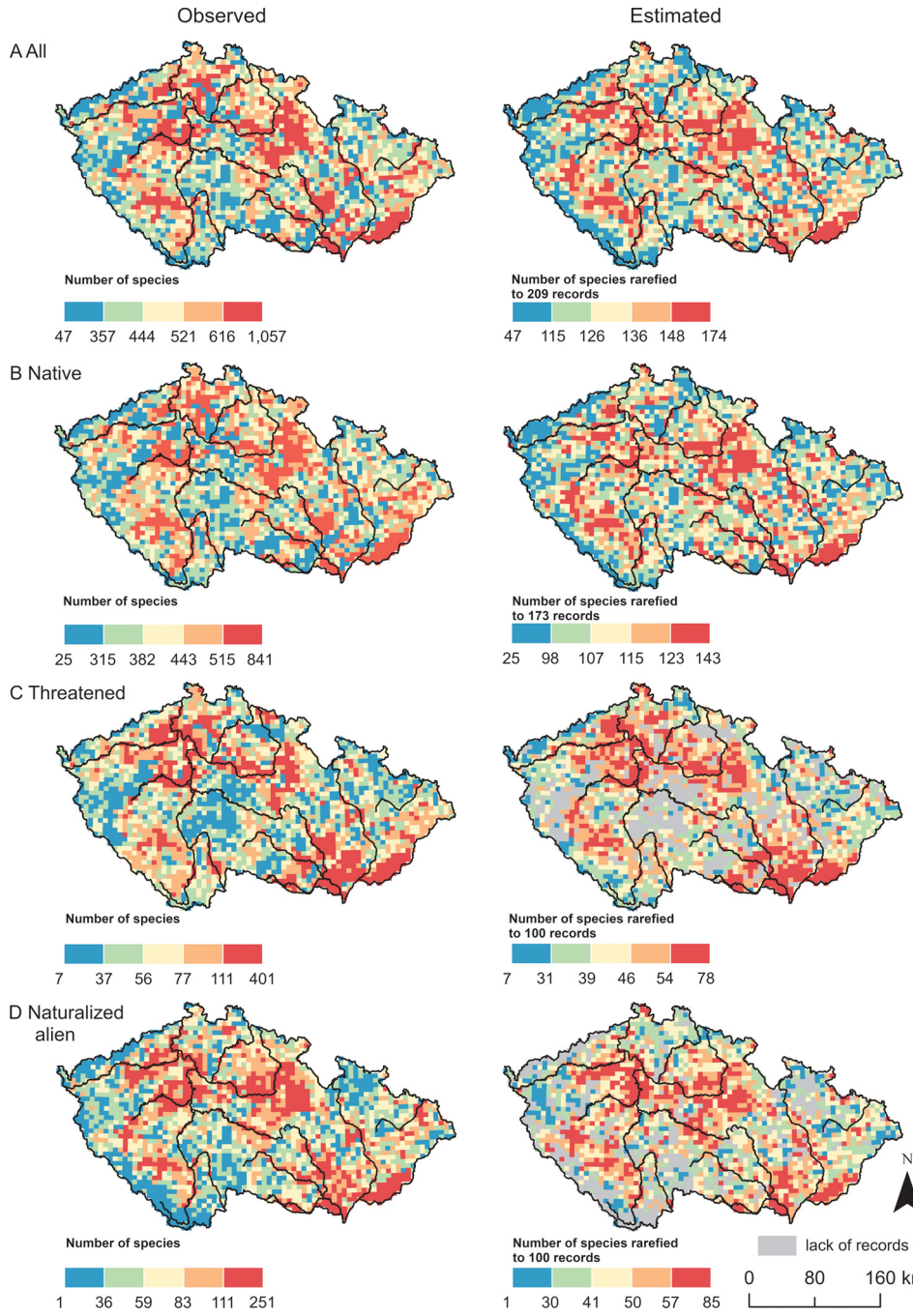

Fig. 2. Spatial patterns of species richness of (A) all, (B) native, (C) threatened (Red List), and (D) naturalized alien species in the Czech Republic. Observed richness (left column) represents the number of species without accounting for sampling intensity; estimated species richness (right column) was calculated using rarefaction curves for 209 (all species), 173 (native species) and 100 (threatened and naturalized alien species) records per grid cell (~6.0 km × 5.5 km). Grey colour indicates grid cells with a lack of records for rarefaction. The number of species per grid cell was classified by quantiles, ensuring that each class contains an equal number of grid cells.

The observed species richness showed a significantly positive Spearman correlation (rho = 0.72) with the number of records, as also evidenced by a visual comparison of the maps of sampling intensity (Fig. 1A) and species richness patterns (Fig. 2 or Supplementary Fig. S4, column 'Observed'). The linear regression of the log-transformed number of species and records per grid cell ($R^2_{Adj} = 0.58$; Fig. 1D) also confirmed this finding. The analysis of residuals from this model revealed positive spatial autocorrelation across the Czech Republic (Moran's I = 0.51, P < 0.001; Fig. 1C). Residuals with high positive values, indicating areas where more species were recorded than would be expected based on sampling intensity, were found close to large cities (e.g. Brno, České Budějovice, Hradec Králové, Prague), in the foothills of the Šumava Mts, Orlické Mts, and in some protected areas (e.g. České středohoří PLA, Podyjí NP, Pálava PLA, Železné hory PLA). The highest observed species richness of all, native, threatened, and naturalized alien species (Fig. 2 or Supplementary Fig. S4, column 'Observed') was recorded in the PLAs Český kras, České středohoří, Křivoklátsko, Moravský kras, Pálava, Bílé Karpaty, and Železné hory, and grid cells in or near Prague and Brno. Conversely, species richness coldspots for all these groups were found in the Krušné Mts and in the PLAs Jizerské hory, Český les, and partly in Šumava NP, Novohradské Mts, Jeseníky PLA, as well as in agricultural lowlands such as the Dyjsko-svratecký úval.

We found contrasting results among the methods estimating species richness while reducing the effect of sampling intensity. Except for rarefaction, all estimates, whether based on nonparametric asymptotic estimators or linear regression, showed statistically significant correlations (P < 0.05) with the number of records (Fig. 3). The results of all estimators, especially the nonparametric ones, were significantly positively correlated with each other. The highest correlations were found between sampling intensity and the nonparametric estimates of species richness, namely PNPMLE (rho = 0.64) and Chao1 (rho = 0.63). The species richness estimated through OLS exhibited a weak but statistically significant negative correlation with sampling intensity (rho = –0.19). Only rarefaction showed estimates independent of sampling intensity (rho = 0.07, P = 0.056). Therefore, we selected rarefied species richness for further analysis and for comparisons with observed patterns. Although the spatial patterns of species richness estimated by rarefaction differed the most from the observed patterns, most hotspots still overlapped with those of observed species richness. The Bílé Karpaty PLA remained a biodiversity hotspot for all, native and threatened species, while the pattern changed for naturalized alien species (Supplementary Fig. S4) after rarefaction. The regions with the lowest estimated number of species across all groups (Fig. 2 or Supplementary Fig. S4, column 'Estimated') were the Krušné Mts, Šumava NP, and Jizerské hory PLA.

Factors influencing plant species richness

To identify the factors with the strongest effects on the distribution of species richness across different species groups, we used spatial error models (SEMs) with (i) observed (Fig. 4A) and (ii) rarefaction-estimated (Fig. 4B) species richness as the dependent variable. The three most important variables in the model for observed richness of all and native species were the stable area of forests (positive effect), the proportion of carbonate bedrock (positive effect), and human population density (positive effect). For threatened species, soil pH (positive effect) emerged as the third most important variable, following

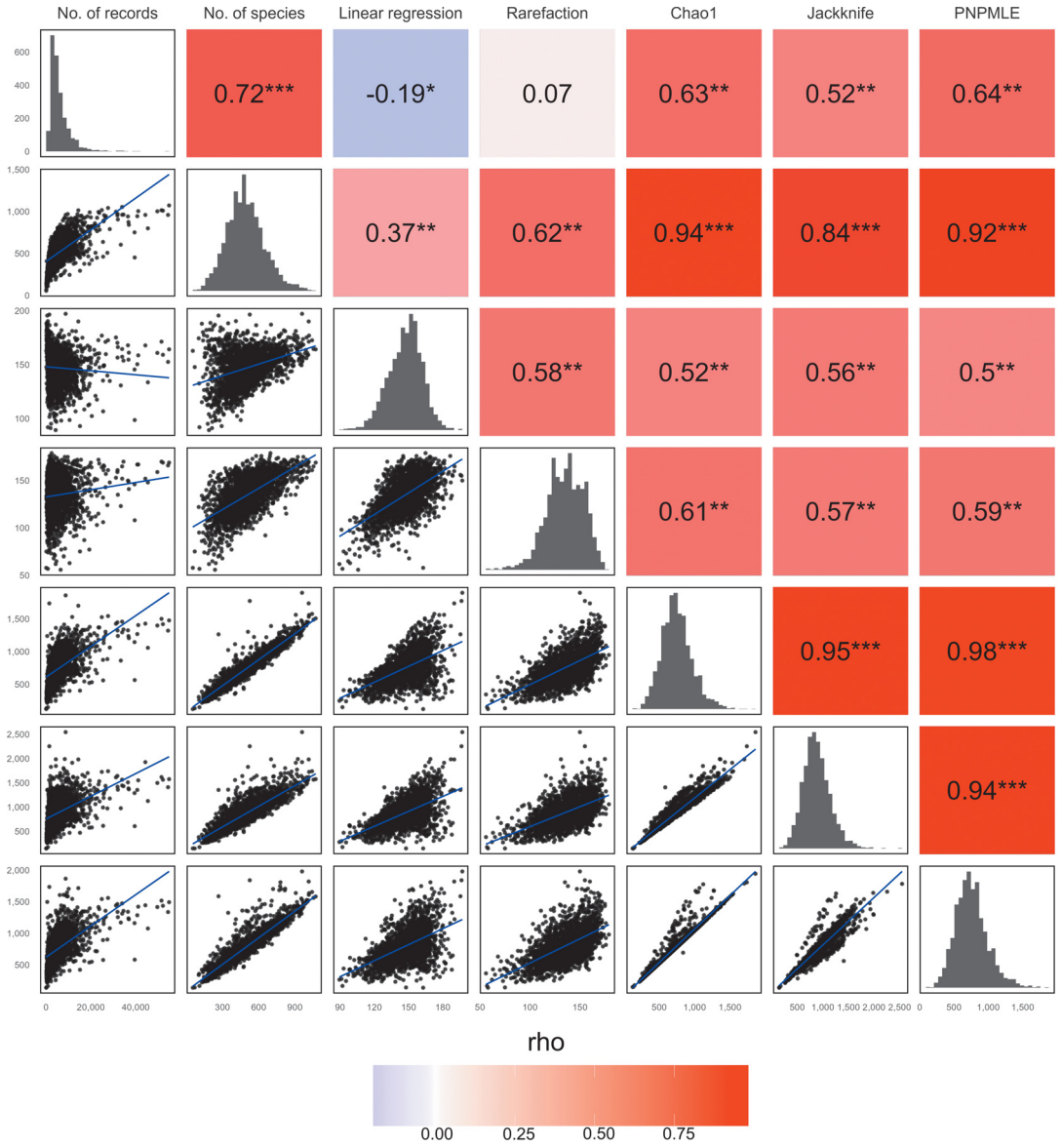

Fig. 3. Spearman rank correlation coefficients between species richness estimates and the number of records, along with histograms showing their distributions and scatterplots illustrating their relationships. Only the relationship between the number of records and the rarefied number of species was not statistically significant.

the stable area of forests and the proportion of carbonate bedrock (both with positive effects). In the model for observed richness of naturalized alien species, human population density (positive effect) was the key variable, followed by heat sum in growing season (positive effect) and stable area of forests (positive effect).

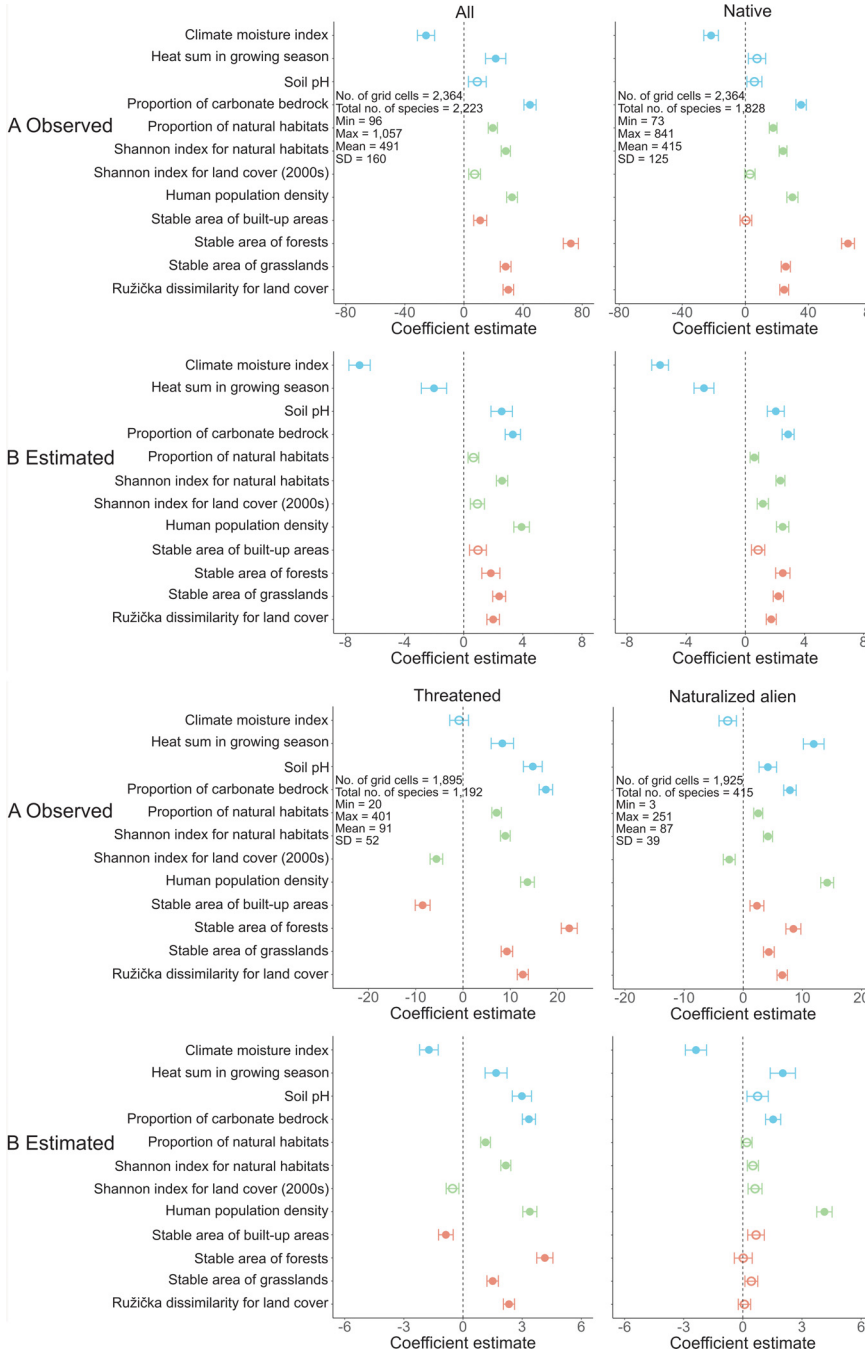

Fig. 4. Coefficients of variables used in spatial error models to predict (A) observed and (B) estimated numbers of all, native, threatened (Red List), and naturalized alien species in the Czech Republic. The dots represent coefficient estimates, and the error bars indicate the standard error on each side. Statistically significant ($P < 0.05$) variables are depicted by filled dots. Environmental variables are shown in blue, current landscape variables in green, and historical landscape variables in red. For the models of observed richness, summary statistics are provided, including the number of grid cells used in the analysis, the total number of species across these grid cells, and the minimum, maximum, mean, and standard deviation of the number of species within the grid cells.

For the estimated richness (Fig. 4B) of all and native species, the climate moisture index emerged as a key predictor, showing a negative relationship with species richness. Human population density and the proportion of carbonate bedrock were also among the most influential variables, showing a positive relationship. In the case of threatened species, the stable area of forests proved to be the most important variable, followed by human population density, the proportion of carbonate bedrock, soil pH, the Shannon index for natural habitats, and the landscape dynamics (all positively related). In the model of naturalized alien species richness, human population density emerged as the key variable (positive effect), followed by the climate moisture index (negative effect) and the heat sum in the growing season (positive effect).

In general, the importance and ranking of individual predictors varied between the SEM models for observed and estimated species richness, depending on the species group. The most pronounced differences were found in the models of observed and estimated richness of all and native species. While the stable area of forests was the most important variable for observed richness, it was not among the top three variables for estimated richness. The heat sum in the growing season exhibited a nonsignificant positive effect on observed richness of all species but showed a significant negative effect on estimated richness. For native species, the heat sum in the growing season had an nonsignificant positive relationship with observed richness but a significant negative relationship with estimated richness. Soil pH was statistically nonsignificant in the SEMs for observed richness of both all and native species, while it had a statistically significant positive effect in the SEMs for estimated richness. In the case of threatened species, the ranking of the most important variables was very similar in the models of both the observed and estimated species richness. In contrast, for naturalized alien species, the climate moisture index was not statistically significant in the model for observed richness but showed a significant negative effect in the model for estimated richness. Another factor, the Shannon index for current land cover (2000s), had a statistically significant negative relationship with observed richness but was not significant for estimated richness. Moreover, variables representing historical landscape development were not statistically significant in the model for estimated richness, whereas they were significant in the model for observed richness.

In the variation partitioning analysis for observed richness (Fig. 5, column ‘Observed’), the pure effect of environmental variables was the highest: 17% for all species, 15% for native species, 29% for threatened species, and 16% for naturalized alien species. For estimated richness (Fig. 5, column ‘Estimated’), the analysis revealed that the highest portion of explained variation in the richness for all, native and naturalized alien species was attributed to the shared effect of all predictor groups (environmental, historical, and current landscape). Specifically, 16% of the variation was explained for all species, 12% for native species and 21% for naturalized alien species. The pure effect of environmental variables explained the most variation in the estimated richness of threatened species (19%) and accounted for more variation than historical and current landscape variables alone across all species groups. The pure effect of variables describing current landscape explained less variation or was equally influential as that of historical landscape variables across all species groups, except for naturalized alien species, for which the current landscape explained more variation than the historical landscape in both observed and estimated richness.

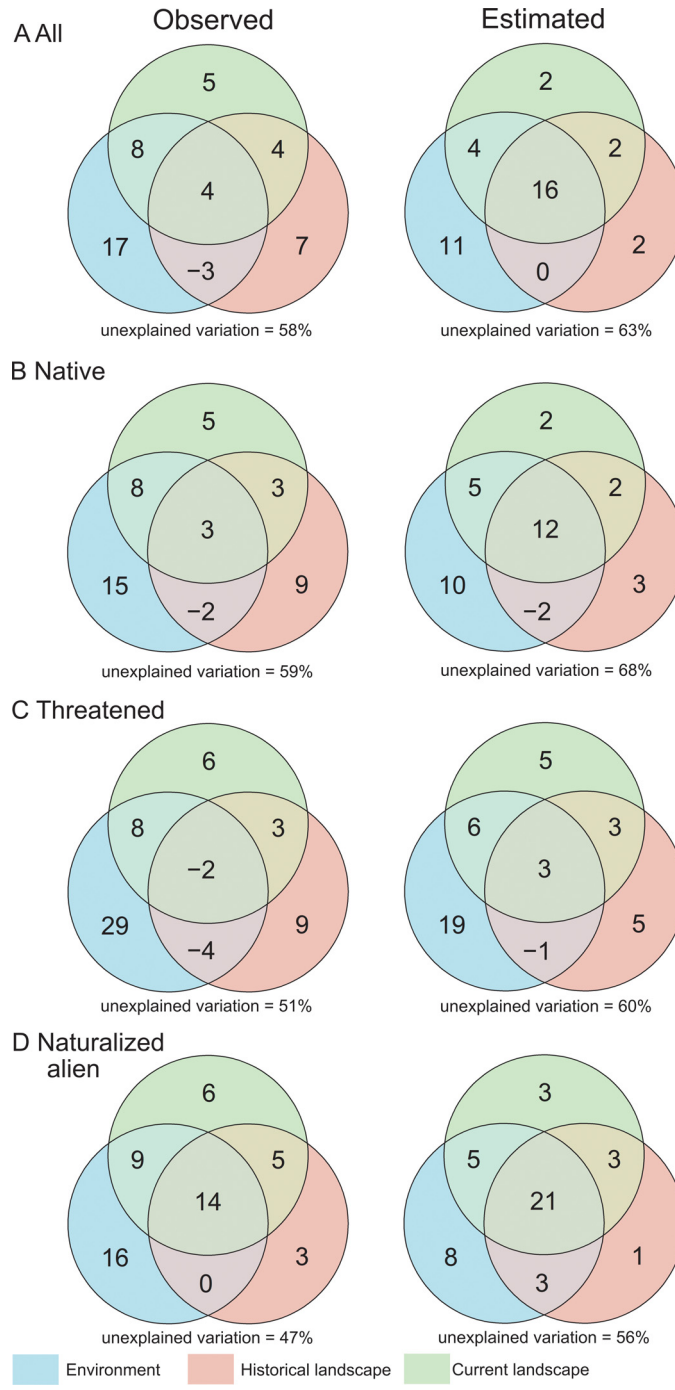

Fig. 5. Variation partitioning of observed (left column) and estimated (right column) species richness of (A) all, (B) native, (C) threatened (Red List), and (D) naturalized alien species. Venn diagrams show the percentage of explained variation (R^2_{Adj}, %) in species richness attributed to environmental variables, historical landscape variables, current landscape variables, and their shared effects.

In addition, we fitted SEMs in which we used the number of records per grid cell as an additional predictor (Supplementary Fig. S5). It emerged as the most influential factor, and the total variation explained by the model (R^2 = 0.789, for all species) was substantially higher compared to the models for estimated (R^2 = 0.374) and observed (R^2 = 0.373) species richness, the latter without including the number of records as an additional predictor.

Discussion

Spatial patterns of species richness

In our study, we found a significant positive relationship between the number of species and sampling intensity within the Czech flora. A similar pattern has also been found in other studies that used extensive vegetation and floristic databases, often revealing even stronger positive relationships (e.g. Yang et al. 2013, Engemann et al. 2015). Some studies demonstrated that high sampling intensity, often associated with high plant richness, was typically found in easily accessible locations (near cities and roads) or protected areas (Engemann et al. 2015, Wild et al. 2019, Hughes et al. 2021). However, species-rich areas tend to attract more survey effort. Therefore, not only the intensity of sampling affects the number of species, but also the knowledge of species-rich areas influences the number of records collected there. A certain degree of correlation between the number of species and the intensity of sampling is therefore natural. This is consistent with the spatial patterns of the number of species and number of records observed in our study. However, the limitations in using such data for mapping and studying spatial patterns of species richness and underlying factors have already been recognized (Wild et al. 2019, Macek et al. 2023). Therefore, we used five different methods for estimating the distribution of species richness independent of sampling intensity, compared selected estimates with the observed species richness and tested how they are influenced by environmental and landscape predictors

The species richness estimates produced by nonparametric estimators showed relatively high positive correlations with sampling intensity and were also strongly positively correlated with each other. According to Poulin (1998), inaccurate and biased estimates of richness by these methods can be attributed to the presence of many singleton or doubleton species in the community. This could be the case in our study, where more than half of the grid cells contained a proportion of rare species (singletons and doubletons) greater than 50%. We found the most significant differences between the observed (uncorrected) richness and the richness estimated by nonparametric estimators in grid cells with low sampling intensity and high representation of singletons and doubletons.

Only the results of the OLS model and rarefaction showed a low correlation with sampling intensity. However, only the rarefaction estimates showed no statistically significant relationship with sampling intensity. Therefore, we analysed rarefied species richness as a contrasting approach to analysing observed richness. The lack of a statistically significant relationship between species richness estimated by rarefaction and sampling intensity is consistent with the findings of Engemann et al. (2015), who also conducted their study at the country scale using grid cells of similar spatial resolution to ours. The spatial patterns of our rarefied richness estimates closely matched the spatial patterns

based on observed richness. Most hotspots and coldspots remained unchanged, except for hotspots of all, native, and threatened species near large cities such as Brno and Prague, which were diminished following rarefaction. This suggests that rarefaction largely preserves high species richness in the areas that are both botanically attractive (i.e. with a high number of records) and represent true richness hotspots, such as the Bílé Karpaty PLA. Importantly, rarefaction does not underestimate species richness in such areas. The only notable exception in our study was the Křivoklátsko PLA, where a hotspot identified in the observed richness did not persist after rarefaction.

The spatial patterns of estimated species richness across the Czech Republic for all, native, threatened, and naturalized alien species in our study are very similar to those identified in previous studies conducted in this country, despite the differences in methods and data sources. For example, Divíšek & Chytrý (2018) used random forest models calibrated on vegetation plots, while Chytrý et al. (2021) and Pyšek et al. (2022) calculated the relative number of species using data from the Pladias database. The most notable differences between the spatial patterns of estimated richness for both all and native species in our study and those reported in the studies by Divíšek & Chytrý (2018) and Chytrý et al. (2021) were identified between the Úhlava and Radbuza rivers near the city of Plzeň. While these studies reported low species richness for both species groups in this area, our findings showed the opposite. These grid cells exhibited a high proportion of stable grasslands and forests, along with a high Shannon index for current land cover and natural habitats, supporting a higher number of estimated species. While Divíšek & Chytrý (2018) used data from vegetation plots of a few (or a few dozen) square metres for model calibration and focused solely on grassland and forest vegetation, we used data from the Pladias database, aggregated into grid cells with a spatial resolution ~ 6.0 km × 5.5 km, covering various habitat types. Furthermore, we found that the sampling intensity in this area was low, and the residuals from our linear regression suggested an above-average number of species relative to the number of records. This indicates that sampling bias may play a significant role in this region.

Factors influencing overall, native and threatened species richness

Although our results show only small differences in the spatial patterns between observed and estimated species richness, the factors influencing richness differed more markedly between the spatial error models (SEMs) for observed and estimated richness, with differences depending on the species group. The most pronounced differences were found in the models for all and native species, where the heat sum in the growing season had a positive effect on observed richness but a negative effect on estimated richness. This may be explained by a shift in the hotspots of species richness from the lowlands with a high heat sum in the growing season, in the case of observed richness, to mid-altitudes in the case of estimated richness. Moreover, the stable area of forests was the most important variable in the SEMs for observed richness of all and native species, whereas it was less important for estimated richness. This may be due to the effect of sampling intensity, as many botanically attractive and/or protected areas are located in regions with long-term continuity of forests. For threatened species, the effects of variables were very similar for both observed and estimated richness, with the most important variables being the same. These findings suggest that the species richness of threatened species may be less

sensitive to sampling bias. However, this may partly reflect the exclusion of a relatively large number of grid cells with a low number of records, which could reduce the influence of sampling bias in the models for threatened species. Variables that show similar effects in both types of SEMs can be considered to have a robust influence on species richness, whereas caution is needed when interpreting the effects of other variables, as these may be biased by sampling intensity.

Climate variables are often considered fundamental for explaining variation in species richness at the regional scale (Bhattarai 2018). In our study, the climate moisture index, expressing the difference between mean annual precipitation and potential evapotranspiration (Hogg 1997), was the most important climate variable, showing a negative relationship with the observed and estimated species richness of all, native and threatened (Red List) species. This means that grid cells in drier lowlands harbour higher species richness of these groups than grid cells at higher elevations. This result contrasts with the general expectation that precipitation has a positive effect on plant richness (e.g. Adler & Levine 2007, Engemann et al. 2015), which typically applies to regions with a high energy supply where water is a limiting factor. This suggests that water is not yet a limiting factor for plants in the Czech lowlands, although this may change in the future due to climate warming (Konapala et al. 2020). The negative relationship identified in our study may, at least partly, be linked to other factors associated with higher precipitation in the Czech Republic, such as low temperatures, short growing season, and acidic soils with nutrient deficiencies in the mountains. It could also be related to the high proportion of species in the Czech flora that are adapted to average or lower soil moisture (Chytrý et al. 2018). Another crucial environmental variable with a robust positive effect on the richness of all, native and threatened species was the presence of carbonate bedrock, on which relatively base-rich soils usually develop, characterized by high base saturation, high pH, and low C:N ratio (Oades 1988). According to Chytrý et al. (2003) and Ewald (2003), the higher plant richness on carbonate bedrocks can be attributed to the larger species pool of plants that are historically adapted to base-rich or calcareous soils in the central-European flora. This is likely due to the disproportionately higher extinction of acidophilous species during the Pleistocene, when habitats with acidic soils were rare. Another reason for the lower plant richness on highly acidic soils is the lower availability of nutrients for plants. In addition, the mobility of potentially toxic elements such as aluminium, manganese, cadmium, zinc, and nickel increases in such soils, and the activity of nitrogen-fixing microbes is slowed down (Bian et al. 2013, Plaster 2013).

While it is commonly assumed that an increase in human population density is a threat to biodiversity, our study found that human population density has a robust positive effect on the richness of all, native and threatened (Red List) species. According to a review by Luck (2007), the positive relationship is generally found at broader spatial scales, and there are various explanations for this relationship. Several studies suggested that both high human population density and high plant richness tend to occur in regions with high productivity and suitable conditions for settlement (Kühn et al. 2004, Evans & Gaston 2005). This is also true for the Czech Republic, where highly productive regions, such as the warm lowlands and uplands, are densely populated and harbour the highest numbers of plant species. Urbanized ecosystems also provide a more diverse mosaic of habitats and microenvironments supporting high plant richness (Lososová et al. 2024). Moreover, many threatened species are associated with habitats that were historically maintained by

traditional management (e.g. pasture grazing, grass cutting, and forest coppicing) in lowland and mid-altitude areas, where human population density is higher compared to mountainous regions (Douda et al. 2016).

Among the variables representing the historical development of the landscape, long-term forest continuity, expressed here as forest area that has not changed since the 1840s, proved to be the most important explanatory variable for both the observed and estimated richness of threatened species. This can be primarily attributed to the dispersal limitation of forest plant specialists, especially perennials with heavy seeds, geophytes, mycotrophs and myrmecochorous species (Ehrlen & Eriksson 2000, Hérault & Honnay 2005). In contrast, recently established forest plantations have different soil properties and host limited number of generalists, while more demanding and characteristic species are missing (Slezák & Axmanová 2016, Bergès et al. 2017, Yang et al. 2021).

Factors influencing the naturalized alien species richness

The SEMs for observed and estimated richness of naturalized alien species differed most notably in the effects of historical variables. While all of these variables were significant in the models for observed richness, they were insignificant in the models for estimated richness. This can be supported by the study of McDonald et al. (2008), who also indicated that past land cover had no causal relationship with alien species richness. We found that human population density had a robust positive effect on the richness of naturalized alien species. This finding is in line with previous studies, such as Spear et al. (2013), Dimitrakopoulos et al. (2017), and Szymura et al. (2018). Areas with high human population density often experience increased introduction of alien species through human activities, further facilitated by favourable habitat conditions provided by human settlements, which promote the naturalization of alien species (McKinney 2006).

Among the environmental variables, the heat sum in the growing season emerged as the crucial driver of both observed and estimated naturalized alien species richness, exhibiting a positive effect. The higher richness of naturalized alien species in areas with a higher heat sum during the growing season is attributed to the fact that, in central Europe, most naturalized alien species originate from warmer climatic regions (Pyšek et al. 2022). Therefore, these species may be better adapted to a limited water supply and high temperatures, giving them a competitive advantage over native species in warmer and drier climates (Pyšek 1998). Moreover, in regions with a lower heat sum in the growing season, the reduced richness of naturalized alien species may be due to a shorter history of human colonization compared to the dry and warm lowlands. These lowland areas have experienced greater and longer propagule pressure, at least since the Neolithic, along with a higher frequency of anthropogenic disturbances (Pyšek et al. 2005). Consequently, naturalized alien species had only a limited time to establish themselves in regions with colder climates.

Specific characteristics of the Czech Republic: effect of spatial resolution and data coverage

The significance of factors influencing plant species richness can vary depending on the spatial extent and resolution of the study (Stein et al. 2014). Our findings indicate that plant species richness in the Czech Republic is influenced by both historical landscape development and current landscape structure. However, variation partitioning revealed

that environmental variables played the most important role in explaining species richness across all selected groups of variables. This can be attributed to the spatial scale and resolution of our study, which is consistent with the conclusions of Pearson & Dawson (2003) for similar spatial scales and resolutions.

Unlike countries with sparse data coverage or regions lacking species records, the Czech Republic benefits from comprehensive data coverage and a long tradition of flora recording. This ensures that areas with fewer records generally correspond to areas of lower species richness, and vice versa (Danihelka et al. 2017). However, sampling bias also plays a significant role in the Czech Republic, mostly evident in the differences in effects of variables between SEMs for observed and estimated richness, as well as in the effect of the number of records tested in SEMs for observed richness, which emerged as the most important predictor across all selected species groups. This highlights the need for careful consideration of uneven sampling intensity in the database data, particularly when using absolute species counts.

Although rarefaction was applied as the primary method because it most effectively reduced the effect of uneven sampling intensity and produced the most contrasting results compared to observed richness, we are aware that it is not always the best method and may not completely eliminate sampling bias. It has several disadvantages, such as a reduction in the variation of richness between grid cells due to the selection of a threshold corresponding to the lowest number of records in the grid cell, and the potential loss of cells with a low number of records. Despite these limitations, we believe that using rarefaction for estimating species richness unaffected by sampling intensity remains a valuable approach and can help gain a better understanding of the factors influencing species richness across different groups of species.

Conclusions

Our results emphasize that rarefaction can be an effective approach for reducing the effect of uneven sampling intensity, which may otherwise distort species richness patterns and influence the relative importance of variables in models. By comparing models for observed and estimated richness, rarefaction can help identify variables with a robust, unbiased effect on plant richness. However, rarefaction also has limitations, highlighting the need for ongoing efforts to develop more precise and reliable methods for estimating species richness in order to further refine our understanding of biodiversity patterns and the factors driving them.

This study also underlines the importance of integrating variables that represent current environmental conditions, as well as current and historical landscape structures, as key drivers of plant species richness at regional and national scales. Current environmental conditions, such as climate moisture, carbonate bedrock, and heat sum in the growing season, strongly influenced species richness, while historical landscape development, particularly long-term forest continuity, played an important role in the species richness of native and threatened species. For naturalized alien species, human population density was essential. To further advance our understanding and conservation efforts, it is important to interpret and digitize historical maps, enhance the spatial resolution and accuracy of environmental variables, and improve the quality of unbiased species records. These

approaches will help identify and protect biodiversity hotspots and develop effective strategies for mitigating the impacts of environmental change in our rapidly changing world.

Supplementary materials

Fig. S1. Topographic map of the study area, Czech Republic, and its location within Europe.
Fig. S2. Spearman rank correlation matrix of dependent and independent variables used in the spatial error models, along with geographical coordinates.
Fig. S3. Maps of selected variables.
Fig. S4. Maps of species richness hotspots and coldspots in grid cells across the Czech Republic for all, native, threatened, and naturalized alien species.
Table S1. Descriptive statistics for mean values of selected variables per grid cell.

Supplementary materials are available at https://www.preslia.cz.

Acknowledgements

We would like to thank all botanists who collected data and contributed to the Pladias database, as well as the creators of the TopoLandUse CZ database. We also thank the reviewers for their insightful comments, which helped improve the manuscript. BK, JDi, MV, MC, HS, JW, and IA were supported by the project Centre for Landscape and Biodiversity (DivLand), funded by the Technology Agency of the Czech Republic (SS02030018). ZK a JW were supported by the long-term research development project no. RVO 67985939 of the Czech Academy of Sciences.

References

Adler P. B. & Levine J. M. (2007) Contrasting relationships between precipitation and species richness in space and time. – Oikos 116: 221–232.
AOPK ČR (2024a) Habitat mapping layer and its updates. – URL: https://data.nature.cz/ds/21.
AOPK ČR (2024b) Large-scale specially protected areas. – URL: https://data.nature.cz/ds/3.
AOPK ČR (2024c) Small-scale specially protected areas. – URL: https://data.nature.cz/ds/1.
Azen R. & Budescu D. V. (2003) The dominance analysis approach for comparing predictors in multiple regression. – Psychological Methods 8: 129–148.
Ballabio C., Lugato E., Fernández-Ugalde O., Orgiazzi A., Jones A., Borrelli P., Montanarella L. & Panagos P. (2019) Mapping LUCAS topsoil chemical properties at European scale using Gaussian process regression. – Geoderma 355: 113912.
Bergès L., Feiss T., Avon C., Martin H., Rochel X., Dauffy-Richard E., Cordonnier T. & Dupouey J.-L. (2017) Response of understorey plant communities and traits to past land use and coniferous plantation. – Applied Vegetation Science 20: 468–481.
Berglund B. E. (2011) How humans have shaped European nature. – Nature and Culture 6: 97–99.
Bhattarai K. R. (2018) Variation of plant species richness at different spatial scales. – Botanica Orientalis 11: 49–62.
Bian M., Zhou M., Sun D. & Li C. (2013) Molecular approaches unravel the mechanism of acid soil tolerance in plants. – The Crop Journal 1: 91–104.
Bičík I., Kupková L., Jeleček L., Kabrda J., Štych P., Janoušek Z. & Winklerová J. (2015) Land use changes in the Czech Republic 1845–2010. – Springer, Cham.
Bivand R. (2022) R Packages for analyzing spatial data: a comparative case study with areal data. – Geographical Analysis 54: 488–518.
Borcard D., Legendre P. & Drapeau P. (1992) Partialling out the spatial component of ecological variation. – Ecology 73: 1045–1055.
Brun P., Zimmermann N. E., Hari C., Pellissier L. & Karger D. N. (2022) CHELSA-BIOCLIM+ A novel set of global climate-related predictors at kilometre-resolution. – EnviDat.
URL: https://www.doi.org/10.16904/envidat.332.
Burnham K. P. & Overton W. S. (1978) Estimation of the size of a closed population when capture probabilities vary among animals. – Biometrika 65: 625–633.

Burnham K. P. & Overton W. S. (1979) Robust estimation of population size when capture probabilities vary among animals. – Ecology 60: 927–936.
Chao A. (1984) Nonparametric estimation of the number of classes in a population. – Scandinavian Journal of Statistics 11: 265–270.
Chao A. (1987) Estimating the population size for capture-recapture data with unequal catchability. – Biometrics 43: 783–791.
Chuman T., Gürtlerová P., Hruška J. & Adamová M. (2014) Geochemical reactivity of rocks of the Czech Republic. – Journal of Maps 10: 341–349.
Chytrý M., Danihelka J., Kaplan Z., Wild J., Holubová D., Novotný P., Řezníčková M., Rohn M., Dřevojan P., Grulich V., Klimešová J., Lepš J., Lososová Z., Pergl J., Sádlo J., Šmarda P., Štepánková P., Tichý L., Axmanová I., Bartušková A., Blažek P., Chrtek J. Jr., Fischer F. M., Guo W.-Y., Herben T., Janovský Z., Konečná M., Kühn I., Moravcová L., Petřík P., Pierce S., Prach K., Prokešová H., Štech M., Těšitel J., Těšitelová T., Večeřa M., Zelený D. & Pyšek P. (2021) Pladias Database of the Czech Flora and Vegetation. – Preslia 93: 1–87.
Chytrý M., Kučera T., Kočí M., Grulich V. & Lustyk P. (2010) Katalog biotopů České republiky [Habitat catalogue of the Czech Republic]. Ed. 2. – Agentura ochrany přírody a krajiny České republiky, Praha.
Chytrý M. & Rafajová M. (2003) Czech National Phytosociological Database: basic statistics of the available vegetation-plot data. – Preslia 75: 1–15.
Chytrý M., Tichý L., Dřevojan P., Sádlo J. & Zelený D. (2018) Ellenberg-type indicator values for the Czech flora. – Preslia 90: 83–103.
Chytrý M., Tichý L. & Roleček J. (2003) Local and regional patterns of species richness in Central European vegetation types along the pH/calcium gradient. – Folia Geobotanica 38: 429–442.
Copernicus (2018) CORINE Land Cover. – URL: https://land.copernicus.eu/pan-european/corine-land-cover.
Czech Statistical Office (2021) Statistický atlas. Sčítání lidu, domů a bytů 2021 [Statistical atlas: Census of population, houses and apartments 2021]. – URL: https://geodata.csu.gov.cz/as/atlas.
Danihelka J., Chytrý M., Kučera J. & Palice Z. (2017) History of botanical research in the Czech Republic. – In: Chytrý M., Danihelka J., Kaplan Z. & Pyšek P. (eds), Flora and vegetation of the Czech Republic, p. 27–87, Springer, Cham.
Dimitrakopoulos P. G., Koukoulas S., Galanidis A., Delipetrou P., Gounaridis D., Touloumi K. & Arianoutsou M. (2017) Factors shaping alien plant species richness spatial patterns across Natura 2000 Special Areas of Conservation of Greece. – Science of the Total Environment 601–602: 461–468.
Divíšek J. & Chytrý M. (2018) High-resolution and large-extent mapping of plant species richness using vegetation-plot databases. – Ecological Indicators 89: 840–851.
Douda J., Boublík K., Doudová J. & Kyncl, M. (2016) Traditional forest management practices stop forest succession and bring back rare plant species. – Journal of Applied Ecology 54: 761–771.
Dubuis A., Giovanettina S., Pellissier L., Pottier J., Vittoz P. & Guisan A. (2013) Improving the prediction of plant species distribution and community composition by adding edaphic to topo-climatic variables. – Journal of Vegetation Science 24: 593–606.
Dubuis A., Pottier J., Rion V., Pellissier L., Theurillat J.-T. & Guisan A. (2011) Predicting spatial patterns of plant species richness: a comparison of direct macroecological and species stacking modelling approaches. – Diversity and Distributions 17: 1122–1131.
Dullinger S., Essl F., Rabitsch W., Erb K. H., Gingrich S., Haberl H., Hülber K., Jarošík V., Krausmann F., Kühn I., Pergl J., Pyšek P. & Hulme P. E. (2013) Europe's other debt crisis caused by the long legacy of future extinctions. – Proceedings of the National Academy of Sciences of the United States of America 110: 7342–7347.
Ehrlen J. & Eriksson O. (2000) Dispersal limitation and patch occupancy in forest herbs. – Ecology 81: 1667–1674.
Engemann K., Enquist B. J., Sandel B., Boyle B., Jørgensen P. M., Morueta-Holme N., Peet R. K., Violle C. & Svenning J.-C. (2015) Limited sampling hampers "big data" estimation of species richness in a tropical biodiversity hotspot. – Ecology and Evolution 5: 807–820.
ESRI (2020) ArcGIS Desktop. – Environmental Systems Research Institute, Redlands.
Evans K. L. & Gaston K. J. (2005) People, energy and avian species richness. – Global Ecology and Biogeography 14: 187–196.
Ewald J. (2003) The calcareous riddle: why are there so many calciphilous species in the Central European flora? – Folia Geobotanica 38: 357–366.
Field R., Hawkins B. A., Cornell H. V., Currie D. J., Diniz-Filho J. A. F., Guégan J.-F., Kaufman D. M., Kerr J. T., Mittelbach G. G., Oberdorff T., O'Brien E. M. & Turner J. R. G. (2008) Spatial species-richness gradients across scales: a meta-analysis. – Journal of Biogeography 36: 132–147.

Gaston K. J. (2000) Global patterns in biodiversity. – Nature 405: 220–227.
Giraudoux P. (2023) pgirmess: spatial analysis and data mining for field ecologists. R package version 2.0.2. – URL: https://CRAN.R-project.org/package=pgirmess.
Grulich V. (2017) Červený seznam cévnatých rostlin ČR [The Red List of vascular plants of the Czech Republic]. – Příroda 35: 75–132.
Gustavsson E., Lennartsson T. & Emanuelsson M. (2007) Land use more than 200 years ago explains current grassland plant diversity in a Swedish agricultural landscape. – Biological Conservation 138: 47–59.
Gwinn D. C., Allen M. S., Bonvechio K. I., Hoyer M. V. & Beesley L. S. (2015) Evaluating estimators of species richness: the importance of considering statistical error rates. – Methods in Ecology and Evolution 7: 294–302.
Hájková P., Roleček J., Hájek M., Horsák M., Fajmon K., Polák M. & Jamrichová E. (2011) Prehistoric origin of the extremely species-rich semi-dry grasslands in the Bílé Karpaty Mts (Czech Republic and Slovakia). – Preslia 83: 185–204.
Härtel H., Lončáková J. & Hošek M. (eds) (2009) Mapování biotopů v České republice: východiska, výsledky, perspektivy [Habitat mapping in the Czech Republic: foundations, results, perspectives]. – Agentura ochrany přírody a krajiny České republiky, Praha.
Hérault B. & Honnay O. (2005) The relative importance of local, regional and historical factors determining the distribution of plants in fragmented riverine forests: an emergent group approach. – Journal of Biogeography 32: 2069–2081.
Hogg E. H. (1997) Temporal scaling of moisture and the forest-grassland boundary in western Canada. – Agricultural and Forest Meteorology 84: 115–122.
Huang E., Chen Y., Fang M., Zheng Y. & Yu S. (2021) Environmental drivers of plant distributions at global and regional scales. – Global Ecology and Biogeography 30: 697–709.
Hughes A. C., Orr M. C., Ma K., Costello M. J., Waller J., Provoost P., Yang Q., Zhu C. & Qiao H. (2021) Sampling biases shape our view of the natural world. – Ecography 44: 1259–1269.
IPBES (2019) Global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. Brondízio E. S., Settele J., Díaz S. & Ngo H. T. (eds). – IPBES secretariat, Bonn. URL: https://doi.org/10.5281/zenodo.3831673.
Jiménez-Valverde A. & Lobo J. M. (2006) Establishing reliable spider (Araneae, Araneidae and Thomisidae) assemblage sampling protocols: estimation of species richness, seasonal coverage and contribution of juvenile data to species richness and composition. – Acta Oecologica 30: 21–32.
Jonas J. L., Buhl D. A. & Symstad A. J. (2015) Impacts of weather on long-term patterns of plant richness and diversity vary with location and management. – Ecology 96: 2417–2432.
Kaplan Z., Danihelka J., Chrtek J. Jun., Kirschner J., Kubát K., Štech M. & Štepánek J. (eds) (2019) Klíč ke květeně České republiky [Key to the flora of the Czech Republic]. Ed. 2. – Academia, Praha.
Konapala G., Mishra A. K., Wada Y. & Mann M. E. (2020) Climate change will affect global water availability through compounding changes in seasonal precipitation and evaporation. – Nature Communications 11: 3044.
Kühn I., Brandl R. & Klotz S. (2004) The flora of German cities is naturally species rich. – Evolutionary Ecology Research 6: 749–764.
Legendre P. & Legendre L. (1998) Numerical ecology. Ed. 2. – Elsevier, Amsterdam.
Lososová Z., Danihelka J., Dřevojan P., Hájek O., Kalusová V., Večeřa M., Chytrý K., Chytrý M., Čeplová N., Filippov P., Jiroušek M., Kadaš D., Kalníková V., Knollová I., Macků M., Niederle J., Novák P., Rohel J., Rotreklová O., Řepka R., Řezníčková M., Šmerdová E., Šumberová K., Veselý P., Vymazalová M., Wirth T. & Tichý L. (2024) Flora of the city of Brno, Czech Republic. – Preslia 96: 123–147.
Luck G. W. (2007) A review of the relationships between human population density and biodiversity. – Biological Reviews 82: 607–645.
Lustyk P. (2016) Metodika aktualizace vrstvy mapování biotopů [Methodology for updating the habitat mapping layer]. – Agentura ochrany přírody a krajiny České republiky, Praha.
Macek M., Abraham V., Tkáč P., Novák D. & Kolář J. (2023) The diversity and distribution of introduced plant species reflect 8000 years of settlement history. – Journal of Ecology 111: 787–798.
Maestre F. T. (2004) On the importance of patch attributes, environmental factors and past human impacts as determinants of perennial plant species richness and diversity in Mediterranean semiarid steppes. – Diversity and Distributions 10: 21–29.
McDonald R. I., Motzkin G. & Foster D. R. (2008) Assessing the influence of historical factors, contemporary processes, and environmental conditions on the distribution of invasive species. – The Journal of the Torrey Botanical Society 135: 260–271.

McKinney M. L. (2006) Urbanization as a major cause of biotic homogenization. – Biological Conservation 127: 247–260.
Meineke E. K. & Daru B. H. (2021) Bias assessments to expand research harnessing biological collections. – Trends in Ecology & Evolution 36: 1071–1082.
Meineri E., Skarpaas O. & Vandvik V. (2012) Modeling alpine plant distributions at the landscape scale: do biotic interactions matter? – Ecological Modelling 231: 1–10.
Midolo G., Skokanová H., Clark A. T., Vymazalová M., Chytrý M., Dullinger S., Essl F., Šibík J. & Keil P. (2025) Nineteenth-century land use shapes the current occurrence of some plant species, but weakly affects the richness and total composition of Central European grasslands. – Landscape Ecology 40: 22.
Moeslund J. E., Arge L., Bøcher P. K., Dalgaard T. & Svenning J.-C. (2013) Topography as a driver of local terrestrial vascular plant diversity patterns. – Nordic Journal of Botany 31: 129–144.
Moran P. A. P. (1948) The interpretation of statistical maps. – Journal of the Royal Statistical Society, Ser. B (Methodological) 10: 243–251.
Němec R., Vymazalová M. & Skokanová H. (2022) The impact of fine-scale present and historical land cover on plant diversity in Central European National Parks with heterogeneous landscapes. – Land 11: 814.
Oades J. M. (1988) The retention of organic matter in soils. – Biogeochemistry 5: 35–70.
Oksanen J., Simpson G., Blanchet F., Kindt R., Legendre P., Minchin P., O'Hara R., Solymos P., Stevens M., Szoecs E., Wagner H., Barbour M., Bedward M., Bolker B., Borcard D., Carvalho G., Chirico M., De Caceres M., Durand S., Evangelista H., FitzJohn R., Friendly M., Furneaux B., Hannigan G., Hill M., Lahti L., McGlinn D., Ouellette M., Ribeiro Cunha E., Smith T., Stier A., ter Braak C. & Weedon J. (2022) vegan: community ecology package. R package version 2.6-4. – URL: https://CRAN.R-project.org/package=vegan.
Paniw M., Salguero-Gómez R. & Ojeda F. (2015) Local-scale disturbances can benefit an endangered, fire-adapted plant species in Western Mediterranean heathlands in the absence of fire. – Biological Conservation 187: 74–81.
Pearson R. G. & Dawson T. P. (2003) Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? – Global Ecology and Biogeography 12: 361–371.
Plaster E. J. (2013) Soil science and management. Ed. 6. – Delmar Cengage Learning, Clifton Park.
Poulin R. (1998) Comparison of three estimators of species richness in parasite component communities. – The Journal of Parasitology 84: 485–490.
Pyšek P. (1998) Alien and native species in Central European urban floras: a quantitative comparison. – Journal of Biogeography 25: 155–163.
Pyšek P., Jarošík V., Chytrý M., Kropáč Z., Tichý L. & Wild J. (2005) Alien plants in temperate weed communities: prehistoric and recent invaders occupy different habitats. – Ecology 86: 772–785.
Pyšek P., Sádlo J., Chrtek J. Jr., Chytrý M., Kaplan Z., Pergl J., Pokorná A., Axmanová I., Čuda J., Doležal J., Dřevojan P., Hejda M., Kočár P., Kortz A., Lososová Z., Lustyk P., Skálová H., Štajerová K., Večeřa M., Vítková M., Wild J. & Danihelka J. (2022) Catalogue of alien plants of the Czech Republic (3rd edition). – Preslia 94: 447–577.
Raduła M. W., Szymura T. H., Szymura M. & Swacha G. (2022) Macroecological drivers of vascular plant species composition in semi-natural grasslands: a regional study from Lower Silesia (Poland). – Science of the Total Environment 833: 155151.
R Core Team (2024) R: a language and environment for statistical computing. – R Foundation for Statistical Computing, Vienna. URL: https://www.r-project.org/.
Richardson D. M., Pyšek P., Rejmánek M., Barbour M. G., Panetta F. D. & West C. J. (2000) Naturalization and invasion of alien plants: concepts and definitions. – Diversity and Distributions 6: 93–107.
Robinson D., Hayes A. & Couch S. (2023) broom: convert statistical objects into tidy tibbles. R package version 1.0.3. – URL: https://CRAN.R-project.org/package=broom.
Ružička M. (1958) Anwendung mathematisch-statistischer Methoden in der Geobotanik (Synthetische Bearbeitung von Aufnahmen). – Biológia 13: 647–661.
Sanders H. L. (1968) Marine benthic diversity: a comparative study. – The American Naturalist 102: 243–282.
Shannon C. E. (1948) A mathematical theory of communication. – Bell System Technical Journal 27: 379–423.
Skokanová H., Havlíček M., Roman B., Demek J., Eremiášová R., Chrudina Z., Mackovčin P., Rysková R., Slavík P., Stránská T. & Svoboda J. (2012) Development of land use and main land use change processes in the period 1836–2006: case study in the Czech Republic. – Journal of Maps 8: 88–96.
Slezák M. & Axmanová I. (2016) Patterns of plant species richness and composition in deciduous oak forests in relation to environmental drivers. – Community Ecology 17: 61–70.
Soberón M. J. & Llorente B. J. (1993) The use of species accumulation functions for the prediction of species richness. – Conservation Biology 7: 480–488.

Sousa-Baena M. S., Garcia L. C. & Peterson A. T. (2013) Completeness of digital accessible knowledge of the plants of Brazil and priorities for survey and inventory. – Diversity and Distributions 20: 369–381.

Spear D., Foxcroft L. C., Bezuidenhout H. & McGeoch M. A. (2013) Human population density explains alien species richness in protected areas. – Biological Conservation 159: 137–147.

Stein A., Gerstner K. & Kreft H. (2014) Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. – Ecology Letters 17: 866–880.

Sychrová M., Skokanová H., Musil M. & Divíšek J. (2024) Landscape heterogeneity shows contrasting spatial patterns but similar temporal changes since the 1840s. – Applied Geography 172: 103431.

Szymura T. H., Kassa H., Swacha G., Szymura M., Zając A. & Kącki Z. (2023) Spatial patterns of vascular plant species richness in Poland - a data set. – Scientific Data 10: 542.

Szymura T. H., Szymura M., Zając M. & Zając A. (2018) Effect of anthropogenic factors, landscape structure, land relief, soil and climate on risk of alien plant invasion at regional scale. – Science of the Total Environment 626: 1373–1381.

Tolasz R. (ed.) (2007) Atlas podnebí Česka. Climate atlas of Czechia. – Český hydrometeorologický ústav, Praha, and Univerzita Palackého, Olomouc.

Večeřa M., Axmanová I., Padullés Cubino J., Lososová Z., Divíšek J., Knollová I., Aćić S., Biurrun I., Boch S., Bonari G., Campos J. A., Čarni A., Carranza M. L., Casella L., Chiarucci A., Ćušterevska R., Delbosc P., Dengler J., Fernández-González F., Gégout J.-C., Jandt U., Jansen F., Jašková A., Jiménez-Alfaro B., Kuzemko A., Lebedeva M., Lenoir J., Lysenko T., Moeslund J. E., Pielech R., Ruprecht E., Šibík J., Šilc U., Škvorc Ž., Swacha G., Tatarenko I., Vassilev K., Wohlgemuth T., Yamalov S. & Chytrý M. (2021) Mapping species richness of plant families in European vegetation. – Journal of Vegetation Science 32: e13035.

Večeřa M., Divíšek J., Lenoir J., Jiménez-Alfaro B., Biurrun I., Knollová I., Agrillo E., Campos J. A., Čarni A., Crespo Jiménez G., Ćuk M., Dimopoulos P., Ewald J., Fernández-González F., Gégout J., Indreica A., Jandt U., Jansen F., Kącki Z., Rašomavičius V., Řezníčková M., Rodwell J. S., Schaminée J. H. J., Šilc U., Svenning J.-C., Swacha G., Vassilev K., Venanzoni R., Willner W., Wohlgemuth T. & Chytrý M. (2019) Alpha diversity of vascular plants in European forests. – Journal of Biogeography 46: 1919–1935.

Walther B. A. & Moore J. L. (2005) The concepts of bias, precision and accuracy, and their use in testing the performance of species richness estimators, with a literature review of estimator performance. – Ecography 28: 815–829.

Wang J.-P. (2011) SPECIES: an R package for species richness estimation. – Journal of Statistical Software 40: 1–15.

Wang J.-P. Z. & Lindsay B. G. (2005) A penalized nonparametric maximum likelihood approach to species richness estimation. – Journal of American Statistical Association 100: 942–959.

Wild J., Kaplan Z., Danihelka J., Petřík P., Chytrý M., Novotný P., Rohn M., Šulc V., Brůna J., Chobot K., Ekrt L., Holubová D., Knollová I., Kocián P., Štech M., Štepánek J. & Zouhar V. (2019) Plant distribution data for the Czech Republic integrated in the Pladias database. – Preslia 91: 1–24.

Willie J., Petre C. A., Tagg N. & Lens L. (2012) Evaluation of species richness estimators based on quantitative performance measures and sensitivity to patchiness and sample grain size. – Acta Oecologica 45: 31–41.

Wüest R. O., Zimmermann N. E., Zurell D., Alexander J. M., Fritz S. A., Hof C., Kreft H., Normand S., Cabral J. S., Szekely E., Thuiller W., Wikelski M. & Karger D. N. (2019) Macroecology in the age of Big Data – Where to go from here? – Journal of Biogeography 47: 1–12.

Yang W., Ma K. & Kreft H. (2013) Geographical sampling bias in a large distributional database and its effects on species richness – environment models. – Journal of Biogeography 40: 1415–1426.

Yang Q., Zhang H., Wang L., Ling F., Wang Z., Li T. & Huang J. (2021) Topography and soil content contribute to plant community composition and structure in subtropical evergreen-deciduous broadleaved mixed forests. – Plant Diversity 43: 264–274.

Zarzo-Arias A., Penteriani V., Gábor L., Šímová P., Grattarola F. & Moudrý V. (2022) Importance of data selection and filtering in species distribution models: a case study on the Cantabrian brown bear. – Ecosphere 13: e4284.

Zimmermann P., Tasser E., Leitinger G. & Tappeiner U. (2010) Effects of land-use and land-cover pattern on landscape-scale biodiversity in the European Alps. – Agriculture Ecosystems & Environment 139: 13–22.

Zizka A., Antonelli A. & Silvestro D. (2020) *sampbias*, a method for quantifying geographic sampling biases in species distribution data. – Ecography 44: 25–32.

Prostorové rozmístění druhového bohatství české flóry: vliv intenzity sběru dat, přírodních podmínek a historie krajiny

Pro ochranu biodiverzity je klíčová znalost jejího prostorového rozmístění a faktorů ovlivňujících rozdíly v druhovém bohatství mezi jednotlivými regiony. S využitím databáze Pladias, která obsahuje údaje o flóře a vegetaci České republiky, jsme zmapovali druhovou bohatost cévnatých rostlin v celé zemi a zkoumali, jak ji ovlivňují současné podmínky prostředí, struktura krajiny a její historický vývoj. Abychom zohlednili vliv nerovnoměrného sběru dat, použili jsme pět různých metod, které odhadly počet druhů v kvadrantech síťového mapování. Pouze metoda rarefakce poskytla odhad nezávislý na intenzitě sběru dat. Pomocí modelů prostorové chyby (spatial error models) a rozkladu variance jsme zkoumali vliv výše uvedených faktorů na distribuci pozorované a odhadované druhové bohatosti. Potvrdili jsme, že druhová bohatost rostlin v ČR je ovlivněna především současnými podmínkami prostředí. Pro diverzitu všech, původních i ohrožených druhů byly nejdůležitějšími faktory podíl karbonátových hornin v podloží a index klimatické vlhkosti. Pro zdomácnělé nepůvodní druhy byla klíčová suma teplot vzduchu během vegetačního období. Historický vývoj krajiny, zejména dlouhodobá kontinuita lesů a trávníků, pozitivně ovlivnil především počet ohrožených druhů, ale vliv byl průkazný i na počet všech a původních druhů. Hustota zalidnění měla pozitivní vliv na celkový počet druhů, a to původních, ohrožených i zdomácnělých. Naše studie ukazuje, že vliv nerovnoměrného sběru floristických dat je poměrně významný a může zkreslovat vliv určitých faktorů prostředí jako např. sumy teplot ve vegetační sezóně. Je proto zásadní jej zohlednit, aby bylo možné korektně posoudit význam jednotlivých faktorů ovlivňujících druhovou bohatost. Ty zahrnují nejen současné environmentální podmínky, strukturu krajiny a vliv člověka, ale také její historický vývoj, jenž je v podobných studiích často opomíjen kvůli špatné dostupnosti a kvalitě dat.

How to cite: Klímová B., Divíšek J., Večeřa M., Chytrý M., Danihelka J., Kaplan Z., Novotný P., Skokanová H., Wild J. & Axmanová I. (2025) Spatial patterns of species richness in the Czech flora: effects of sampling intensity, environment and landscape history. – Preslia 97: 541–566.

Preslia, a journal of the Czech Botanical Society

https://www.preslia.cz