Preslia 98: 241–263, 2026
doi: 10.23855/preslia.2026.241

Preslia

Predictors of vascular plant species richness in Poland

Tomasz H. Szymura[1,*], Henok Kassa[2], Dominika Chmolowska[3], Magdalena Szymura[4], Grzegorz Swacha[1], Zygmunt Kącki[1] & Adam Zając[5]

[1]Botanical Garden, University of Wrocław, Poland; [2]Department of Biology and Ecology, Technical University of Liberec, Czech Republic; [3]Institute of Systematics and Evolution of Animals, Polish Academy of Sciences, Kraków, Poland; [4]Institute of Agroecology and Plant Production, University of Environmental and Life Sciences, Wrocław, Poland; [5]Institute of Botany, Jagiellonian University in Kraków, Poland
[*]corresponding author: tomasz.szymura@uwr.edu.pl

Abstract: Using machine learning techniques, we modelled vascular plant species richness (SR) at a landscape scale (10 × 10 km squares) in Poland, for four plant groups: native species, red-list species, archaeophytes, and neophytes. As explanatory variables, we used environmental characteristics (soil, land relief, and climate), human footprint measures (land use, agricultural intensity, human population density, communication network), and palaeoecological factors (glaciation range). The influence of the individual factors on SR varied considerably, and the different species groups also differed in their reaction to particular drivers. The best predictors of SR, regardless of the species group examined, were potassium content and clay fraction in the soil. This is the result of Pleistocene geological processes: most soils in Poland are of glacial origin, and thus sandy and K-poor. The SR of native taxa increased almost linearly with increased K and clay content, up to ~150 $mg \cdot kg^{-1}$ K, and ~15% clay, and then stabilized, with further SR increases depending on other factors. Regarding these characteristics, the soils in Poland are more similar to those in Fennoscandia than to those in neighbouring countries like Czechia and Germany. The positive correlation of the SR of all species groups with population density results directly from anthropogenic influence for neophytes, but for native species, it results rather from humans intentionally choosing habitats promoting high plant SR when settling in prehistory. As a result, the environmental and palaeoecological variables explain native SR, while the anthropogenic variables were necessary for neophyte and archaeophyte richness. Besides the edaphic variables, a positive influence on native and red-list species richness was land relief diversity. The influence of pH variability, $CaCO_3$ content, water bodies presence, and other typical factors beneficial to biodiversity was rather minor. We did not find any direct effect of intensive agriculture on SR. It could be related to the data-set characteristics: the basic spatial units (100 km^2) are not suited for tracing the agricultural effect, and/or the SR data represented mainly the situation at the end of the 1990s and did not reflect the rapid agricultural intensification resulting from Poland's accession to the EU.

Keywords: biodiversity, biogeography, diversity-environment models, environmental filtering, environmental heterogeneity

Received: 2 Sep 2025; Revised: 29 Apr 2026; Accepted: 16 Jun 2026; Published: 29 Sep 2026

Introduction

Geographic patterns of species richness, along with the question of why some places contain more species than others, have intrigued ecologists for a long time (Barthlott et al. 2007, Kreft & Jetz 2007). Answering this question is also essential in the context of global change and its implications for biodiversity conservation (Brooks et al. 2006, Barthlott et al. 2007). The macroecology of biodiversity is currently advancing quickly due to the increasing availability of biodiversity and environmental data, as well as new analytical methods, which help us to understand the observed patterns (Pärtel et al. 2016).

The spatial pattern of plant species richness results from ecological and evolutionary processes acting at different spatial and temporal scales. In the context of biodiversity conservation, it is important to distinguish between local and regional effects on biodiversity (Urban 2015, Pärtel et al. 2016). At large (global, continental, and regional) spatial extents, evolutionary processes such as migration, speciation, and extinction, as well as geological and climatic history, are crucial, while at a smaller spatial scale, richness depends mostly on species dispersal, habitat filtering, and biotic interactions. Large-scale environmental factors influence local conditions, causing nested relationships between processes acting at different scales (Field et al. 2009, Sabatini et al. 2022). In addition to average values of climatic and edaphic factors, their variability is also considered important for promoting biodiversity, and positive correlations between spatial environmental heterogeneity and species richness are observed across different spatial scales (Stein et al. 2014). Human-derived influence, such as intensive agriculture and urbanization, has an important effect on observed species richness, especially in the case of non-native species (Firbank et al. 2008, de Albuquerque et al. 2011, Flohre et al. 2011, de Barros Ruas et al. 2022, Hou et al. 2023). Growing evidence confirms the effect of historical climate and/or geology on regional biodiversity (Jiménez-Alfaro et al. 2018, Divíšek et al. 2020); however, current richness patterns are typically interpreted as a legacy of landscape history (Divíšek et al. 2020), and palaeoecological processes shaping the distribution of species during the Holocene are typically omitted in biodiversity models.

Poland is a relatively large central-European country, with an area of about 312,000 km^2. Most of Poland's territory consists of flat lowlands and, to a lesser extent, uplands which are a part of the Great European Plain, while mountains, being a part of the Carpathian Mountains (Mts) and the Bohemian Massif, are located in the south (Fig. 1). The majority of Poland belongs to the Continental biogeographical region; a relatively small area in the Carpathian Mts is part of the Alpine region, and in the north-east Poland borders the Boreal region (Cervellini et al. 2020). The Polish flora consists of ~2,600, and 534 neophyte plant species, and ~160 archaeophytes (Tokarska-Guzik et al. 2012, Mirek et al. 2020). The flat land relief allows easy longitudinal plant migration (Szafer 1959), and plant species of different geographical regions occur here (Finnie et al. 2007, Zając & Zając 2011). On the other hand, about 45% of native species reach the limits of their geographical ranges in Poland: mostly northern (~20%), but also southern, western, and eastern (Zając & Zając 2011).

In Poland, generally, the north-east and central parts are species-poor, while the southern part is species-rich (Szymura et al. 2023). This overall trend is consistent with the coarse-grained pattern of vascular plant species richness in Europe (Fig. 1; Cai et al. 2023), but the distribution of regional hot- and cold-spots of species richness (Szymura et al.

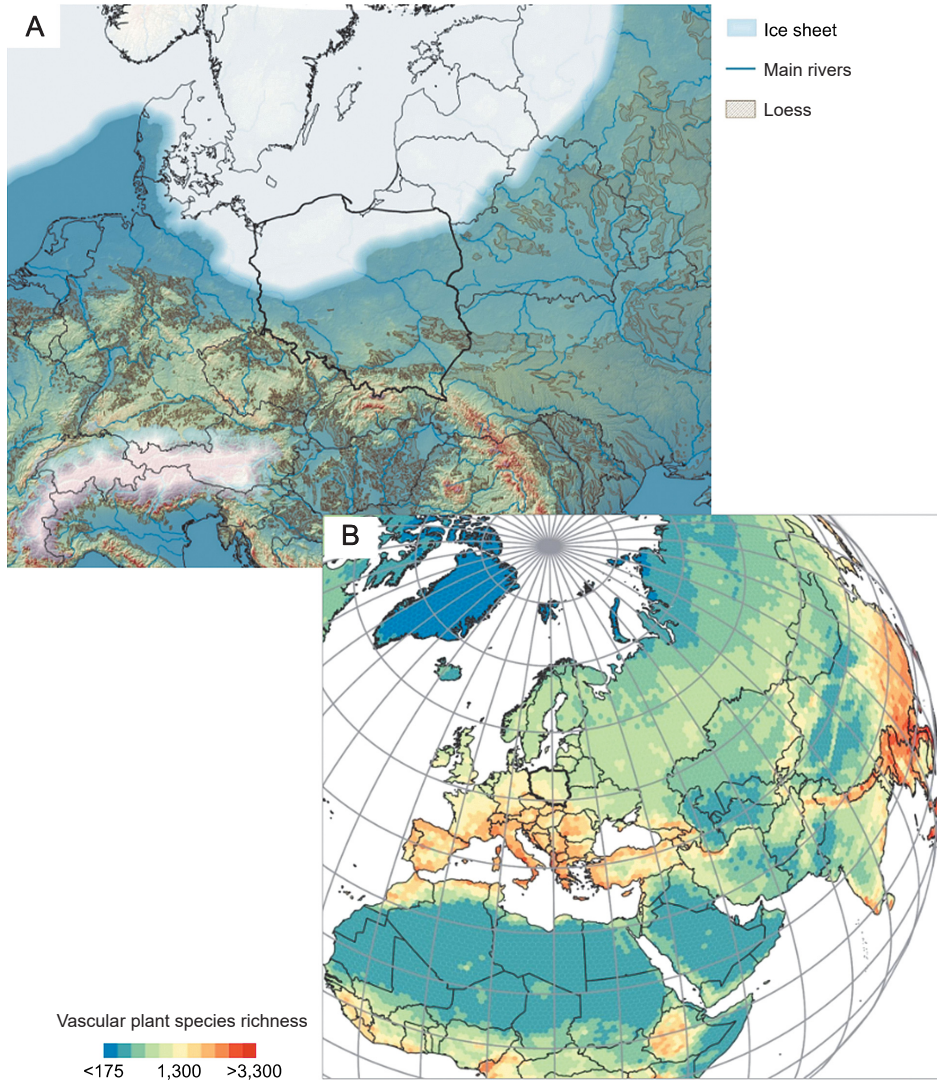

Fig. 1. Location of Poland (thicker lines) and nearby European countries against a background of shaded land relief, main river networks, the extent of the ice sheet during the Last Glacial Maximum (LGM), and the distribution of loess and its derivatives (A), along with the broad-scale pattern of vascular plant species richness (B). The land relief, elevation, and main river network data were obtained from the European Environment Agency; the distribution of loess and its derivatives is according to Lehmkuhl et al. (2021), while the ice sheet cover data follow European palaeoecological and historic atlas (EPHA) (Zentrum für Baltische und Skandinavische Archäologie 2019). The data on species richness (map B) are from Cai et al. (2023).

2024) suggests that there are fine-scale drivers modifying the broad, continental species richness patterns. Traditionally, altitudinal differentiation has been considered a main factor shaping plant species distribution in Poland (Szafer 1959). However, neither this hypothesis nor any other hypotheses explaining the variation in species richness have been formally

tested. It is a significant gap, since distinguishing between local and regional effects is crucial for understanding the macroecology of biodiversity (Mittelbach & Schemske 2015, Pärtel et al. 2016). In this article, we explore the role of potential drivers of landscape-scale vascular plant species richness (SR) in Poland (gamma diversity), considering different species groups (native species, species with high conservation value, archaeophytes, and neophytes). Particularly, this study addresses three main questions: (i) How do environmental, socioeconomic, and palaeoecological factors shape the spatial distribution of species richness? (ii) What is the relative importance of individual drivers, including soil, topography, climate, human footprint, land-use systems, and glacial legacy? (iii) Do these drivers affect the richness of native species and neophytes differently?

Methods

Environment of Poland

The altitudinal differentiation of Poland ranges from –1.8 m a.s.l. in a depression located in the north, near the Baltic Sea coast, up to 2,499 m a.s.l. in the Carpathian Mts. However, most of Polish territory (~90%) consists of lowlands with an altitude below 100 m a.s.l. The climate is temperate, warm, and transitional, with the influence of the continental climate increasing from west to east (Błażejczyk 2006). The mean annual temperature is 8.3 °C, while the yearly sum of precipitation is 613 mm. During the Last Glacial Maximum (LGM), the territory of Poland was partially covered by a continuous ice sheet extending from the north, with permafrost soils covering the remaining territory of Poland. Only outside the ice sheet, in the southern part of Poland, did loess sediments accumulate. As a result, most of the soils in Poland are of glacial origin, formed from components that were strongly washed and sorted by glacial waters. They are acidic and poor in fine soil fractions. Soils of loess and alluvial origin cover a small area (Fotyma et al. 2013). In the Palaeolithic Age, the loess areas in the southern part of Poland were preferentially inhabited (Moskal-del Hoyo 2021), as well as rather small, isolated areas in the west, north, and north-west of Poland on more fertile soils (Nowak 2013). The palaeoecological and historical settlement patterns generally correlate with present human population density (Karpińska-Kołaczek et al. 2014, Demján et al. 2022). From 1772 to 1918, the territory of Poland was partitioned by the Habsburg monarchy, the Kingdom of Prussia, and the Russian Empire. This caused socioeconomic differences resulting in, among others, larger areas of private farms in areas governed by Prussia/Germany, compared to those under the Habsburg Empire (Rudnicki et al. 2016). After WWII, the local population from some areas considered non-Polish was evicted or replaced, and in the depopulated areas, large collective farms were created during the communist period (1945–1989), which further enhanced the differentiation in farm size between different regions of Poland.

Species distribution data

The species distribution data were derived from a new data set presenting total species richness, as well as the richness of species groups, mapped across Poland using a grid of 10 × 10 km squares (hereafter referred to as ‘squares’) covering the entire country. In

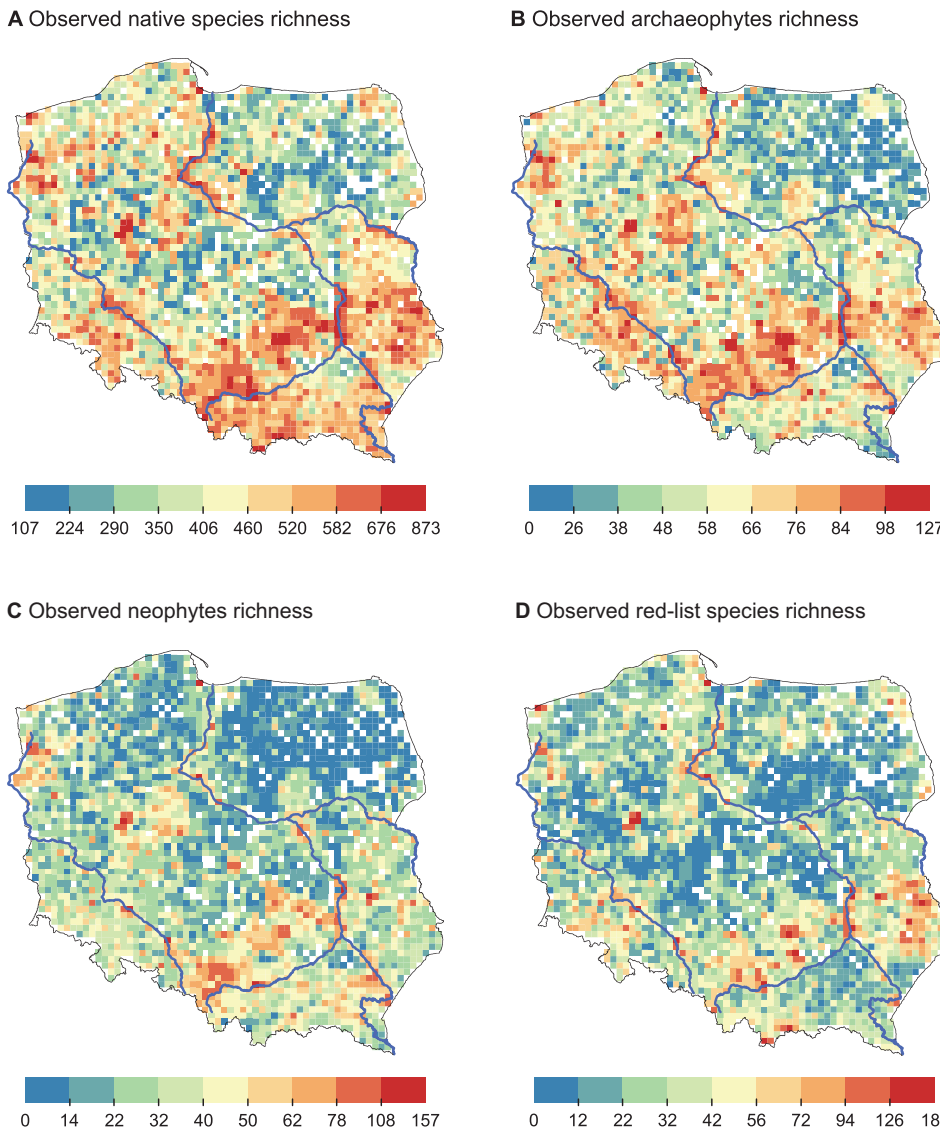

Fig. 2. Spatial pattern of vascular plant species richness in Poland: (A) native species, (B) archaeophytes, (C) neophytes, and (D) red-list species. The white areas indicate the locations of undersampled squares excluded from modelling.

brief, the data come from two sources: the Atlas of Distribution of Vascular Plants in Poland (ATPOL), a project designed for mapping species distribution in Poland using the cartogram method (Zając 1978), and the Polish Vegetation Database (PVD), which collects vegetation-plot data from Poland (Kącki & Śliwiński 2012). The main part of the data from ATPOL was published in 2001, with an appendix in 2019 (Zając & Zając 2001, 2019). The data from PVD come from the years 1925 to 2020, with the highest number of

records from the pre-2000 period. The taxonomical nomenclature was harmonized according to the Euro+Med PlantBase concept, and some aggregates were created if necessary. The status of a species (native, archaeophyte, and neophyte) was checked following Tokarska-Guzik et al. (2012), while conservation value (red-list species) was based on Kaźmierczakowa et al. (2016). The data set is freely available on the Zenodo repository (doi: 10.5281/zenodo.8150757), with a detailed description available in Szymura et al. (2023). Here we examined: native species richness (native), archaeophyte species richness (archaeo), neophyte species richness (neo), and the richness of species of high conservation value (red-list species) (Fig. 2). The last group includes both natives and archaeophytes. For the modelling, we excluded squares with potentially underestimated species richness and squares only partially located within Polish territory (Szymura et al. 2023). Since the data obtained from ATPOL, a main data source, are not accompanied by information about record year and sampling effort (i.e. number of records in a square), we applied a simple procedure, relying on a 'benchmark' approach (Maes & van Swaay 1997, Hill 2012), to detect potentially undersampled squares. The 20 most frequent species in the dataset were determined; next, the species were checked for their geographical ranges and ecological niches. Because the species were common and occurred over the entire territory of Poland, we considered them as a 'wish' list of species which should be recorded in each square. Then, we detected and excluded from the analysis squares where three or more species from the wish list were missing, considering them as undersampled. Finally, 2,137 species across 2,866 squares were considered in the modelling. The average number of native species per square in the data set used for modelling was 433, ranging from 107 to 873; for red-list species: 32 (range 0–181); for archaeophytes: 58 (range 0–127); and for neophytes: 32 (range 0–157).

Explanatory variables

The explanatory variables were segregated into five broad groups: (i) soil properties, (ii) climate, (iii) topography, (iv) anthropogenic variables, including land use pattern and agricultural intensity, and (v) the extent of ice-sheet cover during LGM, which represents a palaeoecological factor that potentially may have influenced recent SR. The topsoil characteristics were calculated from the LUCAS database (Ballabio et al. 2016, 2019). In the case of nitrogen and phosphorus, the distribution of these elements is influenced not only by natural factors but also by fertilization and pollution (Ballabio et al. 2019). The climate variables were obtained from CHELSA 2.1 (Karger et al. 2017), while the topographic variables were calculated from a digital elevation model for Europe (EU-DEM) at 30 × 30 m resolution. The anthropogenic variables were combined using: the CORINE 2018 database (land use), the Central Statistical Office of Poland (data on agriculture and human population density), and the Polish Geographical Objects Database (BDOO) (communication and river networks). The data on the distribution of loess sediments and their derivatives were taken from Lehmkuhl et al. (2021), while ice sheet cover during the Last Glacial Maximum followed the European palaeoecological and historic atlas (EPHA) (Zentrum für Baltische und Skandinavische Archäologie 2019). The details of the data sources are shown in Supplementary Table S1. For each 10 × 10 km square, the average value and coefficient of variation, as a heterogeneity measure, were calculated using SAGA GIS (version 7.8) and QGIS (version 3.2) software.

A set of 64 explanatory variables was assessed for collinearity using a Spearman rank correlation test conducted before modelling. Pairs of environmental variables with correlation coefficients exceeding 0.7 were identified, and one variable from each highly correlated pair was excluded from the analysis. Particularly strongly correlated were indices of soil texture; e.g. soils with a low clay fraction typically are also sandy. The correlation matrix among the explanatory variables is shown in Supplementary Fig. S1. The spatial pattern of distribution of chosen explanatory variables is shown in Fig. 3, and Supplementary Figs S2–S5. Finally, a total of 40 candidate environmental variables were used for the modelling (Table 1), and their descriptive statistics are shown in Supplementary Table S2.

Statistical modelling

The relationships between the species richness and explanatory variables were modelled using Random Forest (RF) with the ranger package (Wright & Ziegler 2017) in the R environment (version 4.2.1, R Core Team 2022). Random Forest (RF) is a machine learning ensemble method with random feature selection according to the bagging idea, able to solve regression problems (Breiman 2001). In brief, multiple trees are constructed on random samples of cases by the bootstrapping technique, and each split, in each tree, is selected from the random sample of predictors. This procedure ensures that a part of the observations is left out in the fitting process and used to calculate the prediction error. Finally, the results of particular trees are averaged. In our examination, we constructed models based on 5,000 trees (num.trees parameter) and used the permutation technique (aka Mean Decrease Accuracy) to calculate variable importance (Breiman 2001). The remaining model parameters were set as defaults.

The significance of each explanatory variable in the models was tested. To obtain a parsimonious model, the RF models were initially run with the candidate variables, followed by a second model run using exclusively variables found as significant in the initial modelling. Significance was checked using the Altmann et al. approach built into the ranger package (Altmann et al. 2010, Wright & Ziegler 2017). We used the 0.05 P-level, and no additional adjustments for multiple comparisons. The variable importance in the main article body was expressed as a percentage to enable comparisons among models. The raw values of variable importance and associated the P-values for initial models are shown in Supplementary Table S3. The values of R^2 and root mean squared error (RMSE) used for model evaluation were calculated on out-of-bag (OOB) data, while root mean squared percentage error was calculated using the entire data set. To elucidate the direct effects of human activity on SR, two models were prepared for each species group: (i) with all types of variables (full model), and (ii) without anthropogenic variables.

The ecological implication of the models was assessed using variable importance and partial dependence plots for each explanatory variable. The variable importance of individual variables was scaled so that the sum of all influences equals 100%, with higher numbers indicating a stronger influence on the response. The partial dependence plots show a modelled effect of a variable on a given response after accounting for the average effects of all other variables in the model (Friedman 2001, Friedman & Meulman 2003). In the context of this modelling, explanatory variables with higher variable importance are related to a greater change in the number of species along the gradient represented by the variable, and the magnitude and course of the change are visible on partial dependence plots.

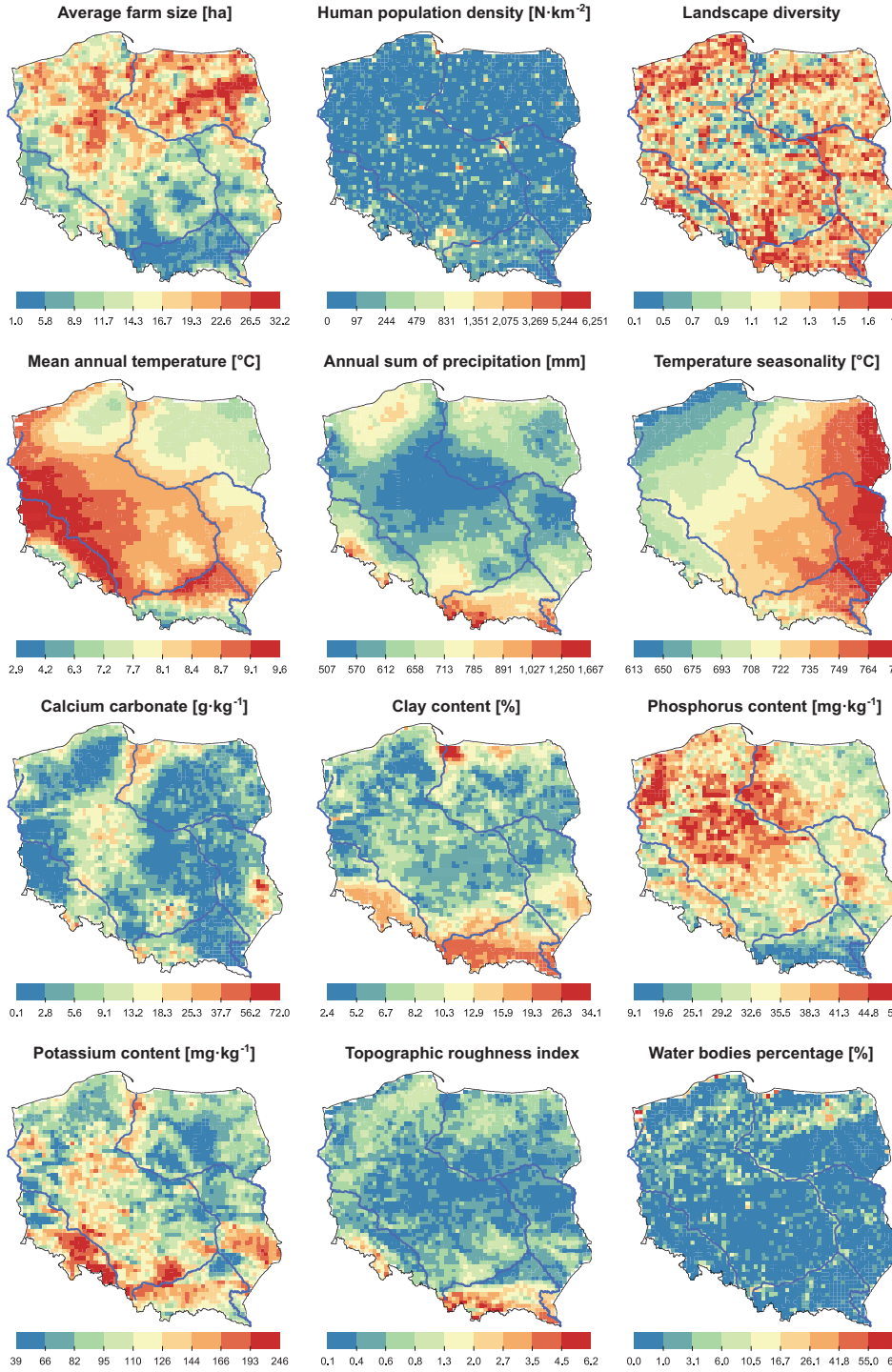

Fig. 3. Distribution of average farm size, human population density, landscape diversity, mean annual temperature, annual sum of precipitation, temperature seasonality, calcium carbonate, clay content, phosphorus content, potassium content, topographic roughness index, and the percentage of water bodies in Poland.

Table 1. Explanatory variables used for final modelling.

Explanatory variables/group	units	Explanatory variables/group	units
Anthropogenic (anthropo):		Soil:	
agricultural tractors density	$N \cdot ha^{-1}$	bulk density	$kg \cdot m^{-3}$
average farm size	ha	calcium carbonate	$g \cdot kg^{-1}$
broadleaf and mixed forests	%	clay content	%
communication routes length	km	coarse grain	%
human population density	$N \cdot km^{-2}$	coarse grain CV	%
human population density CV	%	loess sediments	%
landscape diversity	–	nitrogen content	$g \cdot kg^{-1}$
		nitrogen content CV	%
Climate (climat):		pH	
annual sum of precipitation	mm	pH CV	%
annual sum of precipitation CV	%	phosphorus content	$mg \cdot kg^{-1}$
mean annual temperature	°C	potassium content	$mg \cdot kg^{-1}$
temperature seasonality	°C	potassium content CV	%
Palaeoecological (paleo):		Topographic (topo):	
ice-sheet cover during the	%	sum of rivers length	km
Last Glacial Maximum		topographic roughness index	–
		topographic wetness index CV	%
		water bodies percentage	%

Results

The full models were fitted best for neophytes ($R^2 = 0.61$), followed by archaeophytes ($R^2 = 0.53$), native species ($R^2 = 0.50$), and red-list species ($R^2 = 0.46$). Among the models without anthropogenic variables, the highest explanatory power was for archaeophytes ($R^2 = 0.48$) and native species ($R^2 = 0.48$), while the weakest fit was for red-list species ($R^2 = 0.39$) and neophytes ($R^2 = 0.44$). With the exception of native species, the models including anthropogenic variables performed better, especially for neophytes (Fig. 4). The details of model evaluation are presented in Supplementary Table S4, while the plots and maps of observed vs. modelled SR are in Supplementary Figs S6–S7.

The soil variables explained most of the SR variability in the models for archaeophytes, red-list species, and native species, while for neophytes, the anthropogenic variables, when included in models, explained more SR variability than the soil variables (Fig. 5). The climate variables significantly improved models for all species groups. The fraction of SR variability explained by topographic variables was also significant for all species groups, but relatively small for archaeophytes and neophytes. The ice sheet extent during LGM explained a rather low amount of variability but significantly contributed to models for all groups of species, except neophytes. The exclusion of anthropogenic variables from the models typically increased the fraction of variance explained by soil and, to a smaller extent, by topography (Fig. 5).

For all groups of studied species, regardless of the inclusion of anthropogenic variables, potassium and clay content were consistently among the explanatory variables with the largest relative influence (Fig. 6). Species richness was low at low values of these predictors, increased almost linearly with increased soil K and clay content up to the average values for Poland, and then stabilized at a certain level, showing no further increase

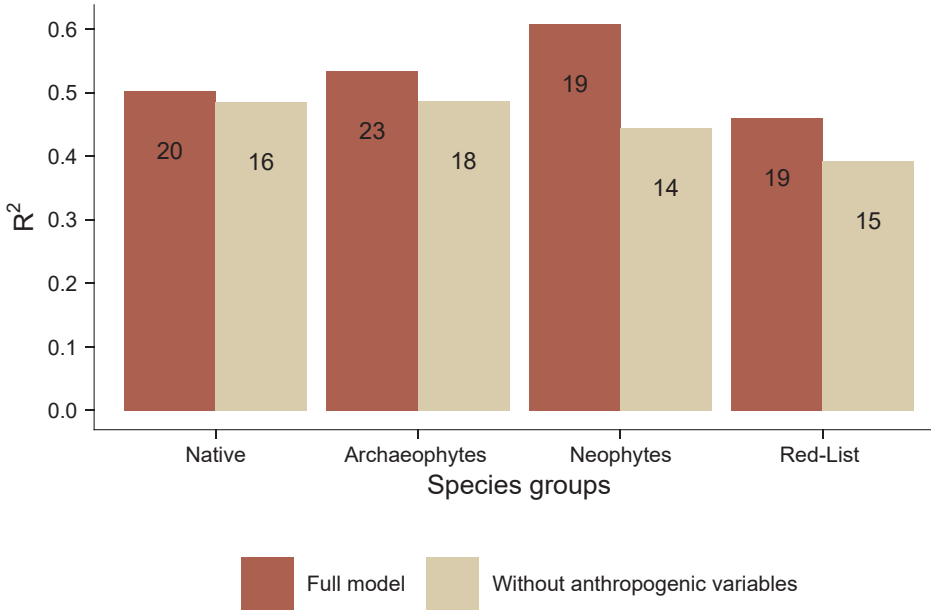

Fig. 4. Values of R^2 for full models and models excluding anthropogenic variables for each species group. The numbers correspond to the number of explanatory variables in each model.

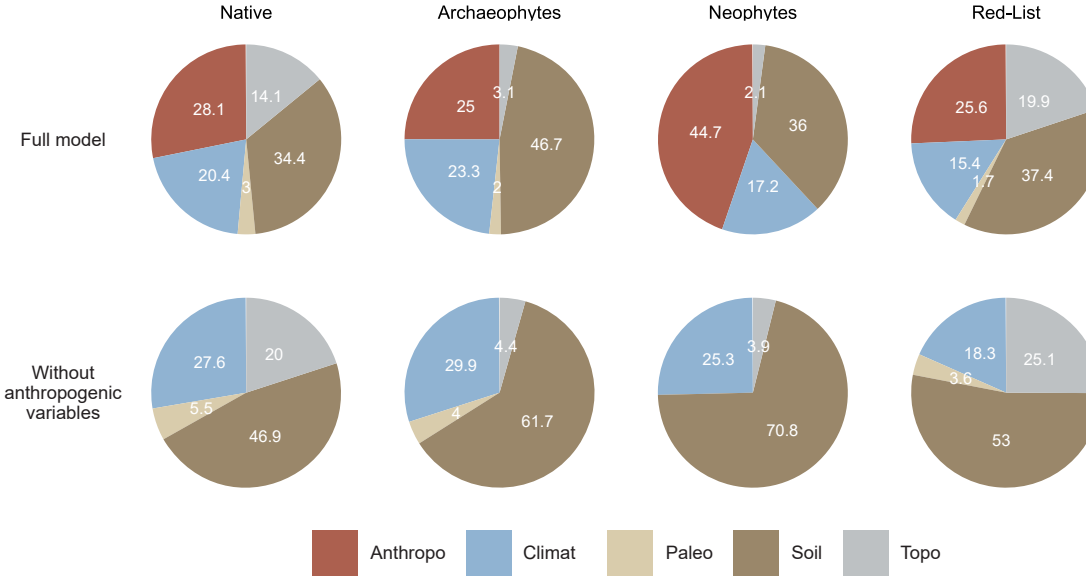

Fig. 5. Sum of scaled variables importance (in %) in different groups of explanatory variables: anthropogenic (anthropo), climatic (climate), palaeoecological (paleo), soil (soil), and topographic (topo) for species groups (columns) and model types (rows).

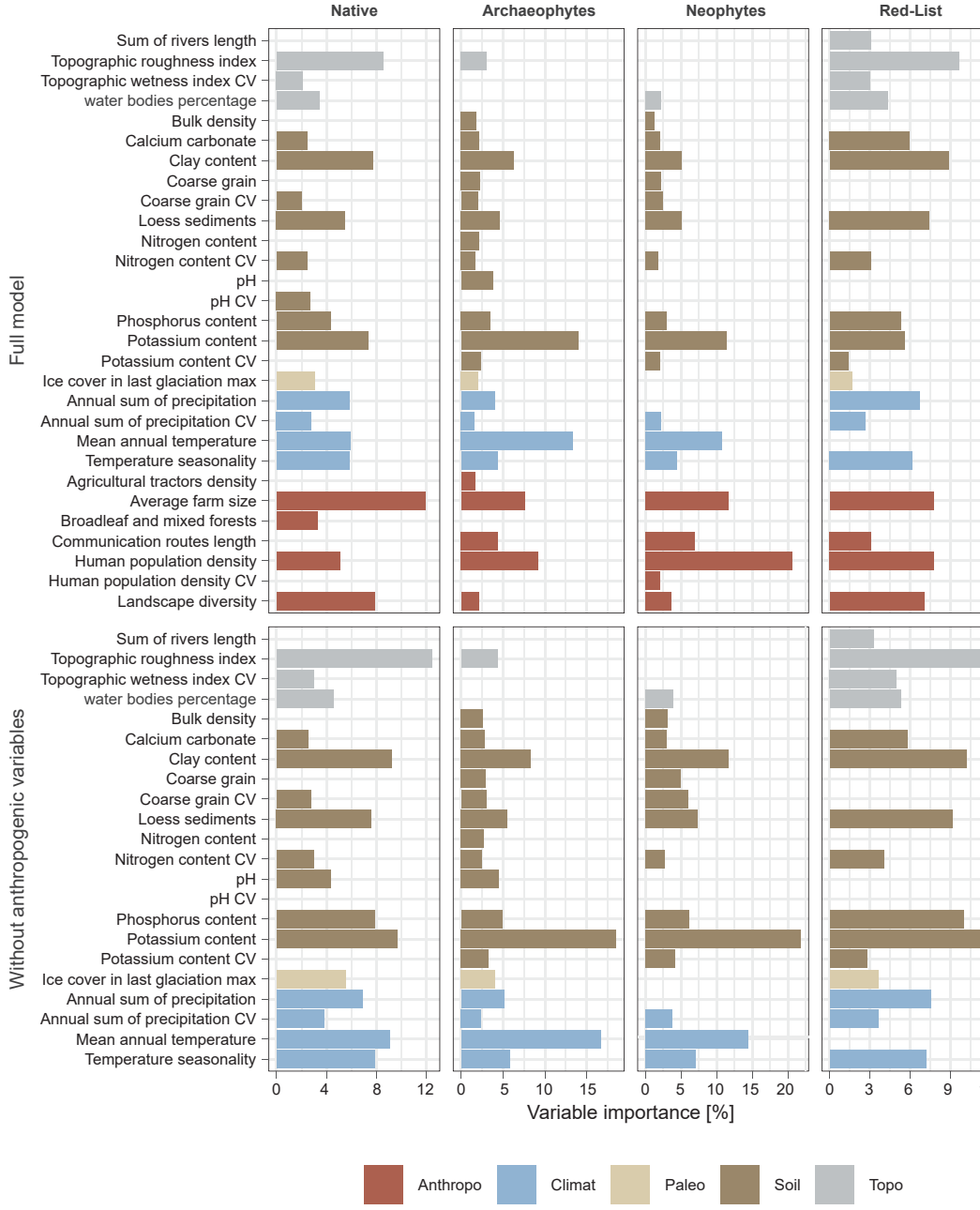

Fig. 6. Relative importance of variables for species groups (native, archaeophytes (archaeo), neophytes (neo), and red-list species – columns) and models (full model – upper panel, and model without anthropogenic variables – lower panel). Different colours represent groups of explanatory variables (anthropo, climate, paleo, soil, topo). Note the different scales on the x-axis among species groups. The absence of a bar indicates non-significant variables in the model for a particular species group. The explanatory variables are sorted alphabetically within each variable type.

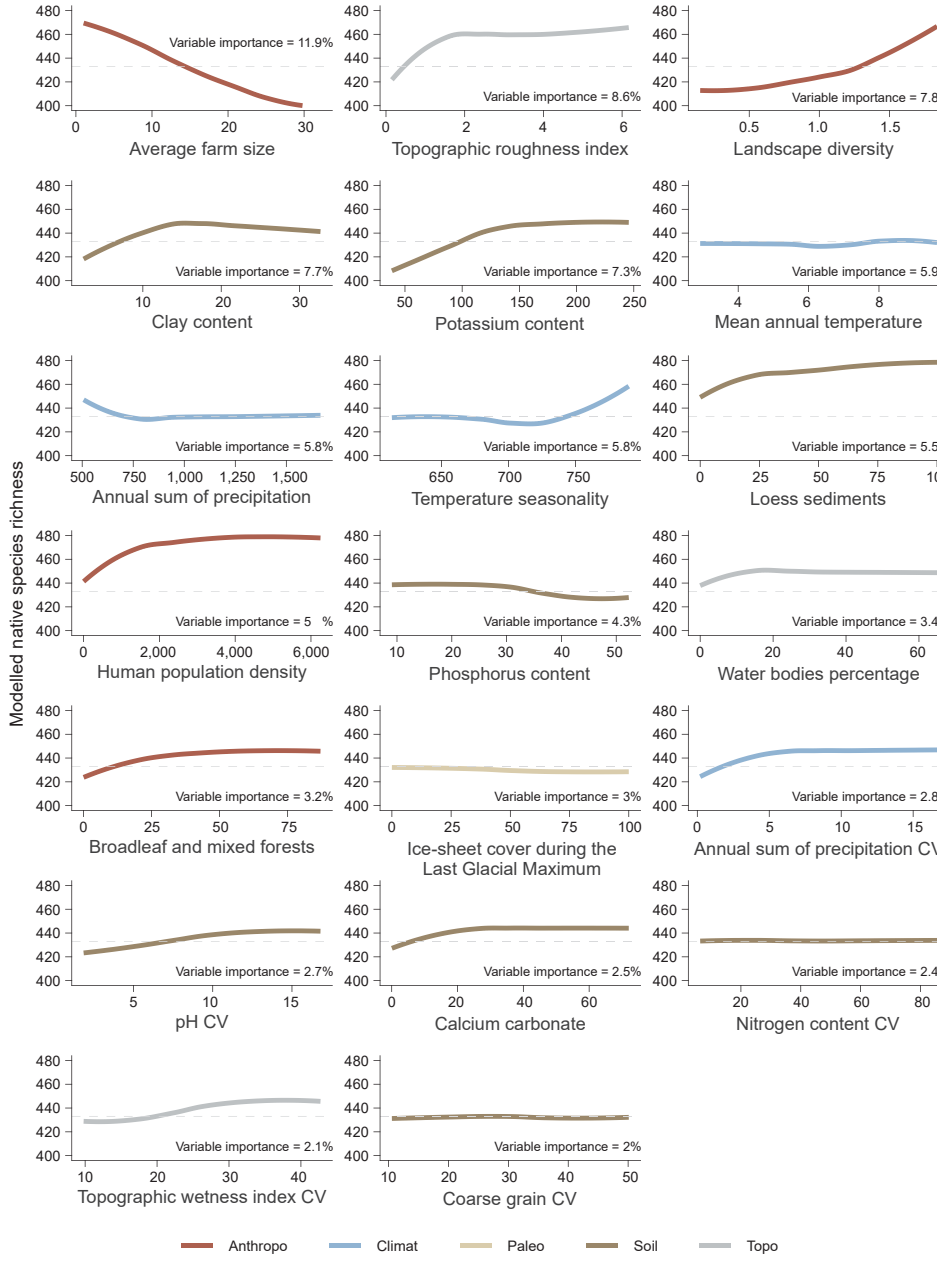

Fig. 7. The partial dependence plot (lines) and relative importance of variables for the full model of native species richness. The explanatory variables are sorted in decreasing order of their relative importance. The dashed grey horizontal line depicts the average species richness (SR).

with higher K and clay content (Fig. 7 and Supplementary Figs S8–S15). For all species groups, especially red-list species, the presence of loess sediments and their derivatives increased SR. Interestingly, the influence of $CaCO_3$ on the models was relatively low,

with the highest effect on the red-list species group. Generally, an increase of $CaCO_3$ in topsoil increased species richness in a manner resembling logarithmic growth. In models without anthropogenic variables, the effect of topsoil phosphorus was relatively strong: above some threshold of extreme phosphorus content, species richness decreased (Fig. 7 and Supplementary Figs S12–S15). In full models, population density and farm size were also highly influential for all species groups (Fig. 6). The species richness of all groups increased sharply at low population density and then levelled off, with a slight decline at the highest densities, and SR decreased almost linearly with increased average farm size (Fig. 7 and Supplementary Figs S8–S11). Among anthropogenic variables, we also observed a relatively large influence of landscape diversity, but only in the case of native and red-list species, for which increasing landscape diversity increased species richness (Fig. 7 and Supplementary Figs S8–S11). Among the topographic variables, the effect of the topographic roughness index was consistent, except for neophytes. Species richness increased along with increased topographic roughness index values in a logarithmic-like curve (Fig. 7). Finally, ice-sheet cover during the LGM had a negative effect on species richness. It was observed for all species groups, but not for neophytes (Fig. 7 and Supplementary Figs S8–S15).

Discussion

The results show that species richness (SR) patterns of different groups of vascular plants across Poland are shaped by numerous interacting variables, primarily soil conditions, but also topography, climate, and human influence. Furthermore, the importance of explanatory variables differs among species groups. In particular, anthropogenic variables were essential for modelling neophyte richness and, to a lesser extent, archaeophyte richness, but were not crucial for native species. The importance of anthropogenic factors in explaining the richness of red-listed species likely reflects the presence of numerous archaeophytes within this group.

Soil properties

Variation in soil properties is among the key drivers of biodiversity patterns (for a review see Hulshof & Spasojevic 2020). Potassium (K), nitrogen (N) and phosphorus (P) are among the main nutrients limiting plant growth (Sardans & Peñuelas 2015), while the clay fraction is crucial for increasing water holding capacity, enhancing nutrient retention, and reducing leaching in sandy soil (Newman 1984). The main soil K source is the geological parent material (Mengel & Kirkby 1980), and its low amount is observed in sandy soils (Zawadzki 1999); thus, in European soils, K content is positively correlated with clay, producing quite similar spatial patterns of distribution (Ballabio et al. 2019). From a continental perspective, Poland, along with Nordic countries and north-east Germany, is part of a large area of K-poor soils, which resulted from the effect of glaciation on soil formation (Ballabio et al. 2016, 2019). In Poland, about 70% of soils are of glacial origin (Fotyma et al. 2013), and average soil K content in Poland is about 108 $mg \cdot kg^{-1}$, compared to 171 $mg \cdot kg^{-1}$ in Germany and 186 $mg \cdot kg^{-1}$ in the Czech Republic (own calculation) based on data from Ballabio et al. (2019). A similar pattern is observed in the case of clay content. It could be assumed that the country-scale differences in K and clay content in

soils are the reason why in a study for Czechia, both K and clay were not among the most influential drivers of SR (Klímová et al. 2025), contrary to our results. Due to the glaciation spreading from the north, the distributions of soil K and clay follow a distinct spatial pattern: soils in the lowlands of northern and central Poland (with some exceptions) are sandy, K-poor, and have a low clay content, while in the mountains and highlands located in the south, the soils are less sandy and richer in K and clay (Fig. 3). The logarithmic-like shape of the modelled relation between K, clay, and SR suggests that only a restricted number of species are adapted to grow on sandy, K-poor soils. However, above certain threshold values of soil K and clay content (~150 $mg \cdot kg^{-1}$ for K and 15% clay in the case of native species), further increases did not promote an increase in SR. Thus, low clay and K content in topsoil restricted SR and could be an indicator of low SR, but their high content did not promote high SR, and this pattern was observed for all species groups examined here.

The effect of the remaining topsoil variables is less pronounced and not consistent across species groups. A majority of vascular plant taxa constituting the central-European species pool are restricted to base-rich and calcareous soils; thus, over central Europe, calcium carbonate content in soils, along with soil pH, are among the most important drivers of species richness (Chytrý et al. 2003, Ewald 2003). In Czechia, the proportion of carbonate bedrock in a landscape is among the most influential factors positively influencing SR (Klímová et al. 2025). We believe that the limited positive effect of topsoil calcium content on species richness (SR) may result from the predominance of glacial-origin soils, whose properties, especially pH, are more strongly determined by material deposited and modified by glacial and postglacial processes than by the local presence of calcium-rich bedrock. The stronger effect of landscape-scale topsoil pH variability on species richness (SR), compared with the effect of mean pH, may result from the correlation between these two variables: grid cells with high pH variability also tend to have high mean pH values. The species group for which the positive influence of calcium carbonate was relatively highest was red-list species. It appears to be the result of the preferential protection of species typical of calcareous substrates in Poland. Since the area of calcareous soils in Poland is rather low, species with distributions restricted to such substrates are relatively rare; thus, many of them are protected. Interestingly, there is a negative correlation between P content and the SR of native and red-list species. Since P content in soils is influenced by fertilization (Ballabio et al. 2019), it could be hypothesized that this reflects excessive P fertilization of agricultural land, especially given the positive correlation between P content and NPK fertilizer application in Poland (Supplementary Fig. S1).

The positive effect of the presence of loess sediments and their derivatives on SR in Poland seems to be rather complex. The sediments have specific chemical and physical properties, and soils developed from them are very productive compared with those developed from other parent materials (Busacca & Sweeney 2005). As a result, in the territory of Poland, loess areas were preferred for human settlements in the Neolithic period. Besides the timing of Neolithic settlement development, its character is also important: Neolithic farming communities occupied fertile lands mainly in southern Poland, whereas hunter-gatherer communities prevailed mainly in the north and north-east (Nowak 2013). As a result, large areas in the south of contemporary Poland have been continuously deforested since the Neolithic period (5000–4000 BC), while less productive areas in the north-east part, e.g. some areas in the Masurian Lakes, were not significantly changed by

humans until the late Middle Ages (Wacnik et al. 2012, 2016, Ryzner & Owczarek 2020, Kabala et al. 2025). Therefore, the increase of SR for all species groups on loess areas, besides the physico-chemical soil properties, is likely correlated with human settlement patterns, both prehistorical and recent. Such correlations among species richness, settlement history, environment, and current land-use were also found in the Czech Republic (Macek et al. 2023).

Climate

In the context of Poland, most of the variability in mean annual temperature and annual sum of precipitation results from altitudinal differentiation (Błażejczyk 2006). The most pronounced effect appears to be that of mean annual temperature, especially visible in the case of archaeophytes, and the shape of the dependence plot underlines a lack of archaeophytes at higher altitudes, where the temperature does not exceed ~6 °C. Similarly, the effect of the annual sum of precipitation, where values around 750 mm separate the mountains and some areas close to the Baltic Sea from most of the lowlands and highlands, can be observed. The influence of temperature seasonality is also relatively high in the case of red-list species. Temperature seasonality reflects the west-to-east continentality gradient and displays a V-shaped pattern, although it shows asymmetry towards higher values (i.e. eastern Poland). The partial dependence plots indicate the lowest species richness in the centre of Poland, compared to the western and, to a larger extent, eastern parts of Poland.

Anthropogenic variables

Among the variables representing human footprint, the positive relationship between human population density and SR is striking. Similar positive correlations were also observed in Czechia for native, threatened, and naturalized alien SR (Klímová et al. 2025). For neophytes, it resulted from higher propagule pressure in densely populated, urbanized areas (Hulme et al. 2009, Pyšek et al. 2010). For native species, such a broad-scale positive correlation suggests that humans intentionally settled in areas with high biodiversity, and additionally contributed to biodiversity through habitat diversification and/or by introducing archaeophytes (Luck 2007, Pautasso 2007). In the European context, the flora of cities is naturally rich (Kühn et al. 2004), and present-day regional diversity of archaeophytes reflects the intensity of past human settlements (Macek et al. 2023). Because recent low settlement density reflects similar low interest in the past (Karpińska-Kołaczek et al. 2014), the positive relationship between population density and archaeophyte richness, observed here, seems to be causal.

The negative relationship between average farm size and the richness of different plant species groups is also pronounced, but this relationship is rather complex. In Poland, farm size is largely shaped by socioeconomic legacies: farms tend to be large in areas where, after WWII, vast areas were taken over by the communist government and converted into collective farms. After the fall of the communist system, the large amount of state-owned land enabled the establishment of large private farms (Głębocki 2014). In Poland, small farms are typical of the south-eastern highlands and mountains, with a well-preserved traditional ownership structure and a characteristic regional landscape, the so-called Polish strip fields (Meeus 1995), a kind of ribbon-farm system. It is believed that this region, with traditional, extensive land-use and a distinctive landscape, supports high biodiversity

(Gwiazdowicz 2010). Conversely, the majority of the largest farms are located in the northern part of Poland, in relatively unproductive areas where most of the inhabitants were replaced after WWII and huge collective farms were created. Thus, we do not interpret the correlation between species richness and farm size as causal.

In Europe, the negative effect of intensive agriculture on species richness is well recognized (Firbank et al. 2008, Flohre et al. 2011). Also, in Poland, it was found that the number of arable weeds decreased in areas with intense agriculture, while remaining stable in areas with extensive management (Dąbkowska et al. 2017). Nonetheless, our models did not confirm a pronounced influence of any direct measure of agricultural intensity (e.g. amount of applied NPK fertilisers, number of tractors per ha) on species richness. We can relate this to species richness data characteristics: they best reflect the situation at the end of the 20th century (Szymura et al. 2024), while it is believed that a rapid increase in agricultural intensification in Poland occurred after the fall of the communist system in the 1990s (Gwiazdowicz 2010). Thus, it is likely that we did not capture the cumulative effect of recent intensive agriculture in these data. Moreover, it is unclear to what extent the effect of intensive agriculture was captured by species counts within relatively large, and likely environmentally heterogeneous, areas with a square size of 10 × 10 km.

Topography and effect of environmental heterogeneity

Spatial environmental heterogeneity promotes species diversity mainly through an increase in the number of habitat types, resources, and structural complexity, which in turn increases the available niche space and allows more species to coexist (Stein et al. 2014). In our examination, the effect of heterogeneity expressed by the topographic roughness index was among the single most influential SR factors. The positive effect of the topographic roughness index on the SR of native and red-list species can be interpreted in the context of low spatial heterogeneity on flat, homogeneous lowlands, which results in restricted SR. In the case of landscape diversity and soil moisture spatial variability (topographic wetness index CV), the environmental heterogeneity effect overrides the influence of any single land use type and average values of water conditions. Besides the above factors, variability in soil properties (nitrogen and potassium content CVs and coarse grain CV), climate (annual sum of precipitation CV), and human population density CV also enhanced SR, in addition to the effects of average values of these factors.

Effect of ice sheet cover during Last Glacial Maximum

We observed low species richness in areas covered by the ice sheet during the LGM; however, this palaeoecological factor did not influence neophyte richness, since this group of species was introduced relatively recently (after 1492). It was hypothesized that the SR of glaciated areas is restricted because species from glacial refugia did not have enough time to settle and adapt there (Firbas 1949). Results of recent SDM models (e.g. Normand et al. 2011) confirmed the migration lag time effect on species richness. Nonetheless, palynological studies from central Europe suggest that the rapid increase of SR in the late Holocene resulted mostly from human land-use changes (Giesecke et al. 2019), while in Fennoscandia, the regional species pool increased in the early and mid-Holocene, then levelled off in the late Holocene (Rijal et al. 2021). Therefore, in Poland, it is rather unlikely that observed lower SR in areas covered by ice sheets during the LGM

resulted from a migration lag time effect (Szymura et al. 2024), especially since the width of the area covered by a continuous ice sheet ranged from 70 to 270 km in the contemporary land area of Poland (Fig. 1A), which is a relatively short range for migration. Thus, the low SR in areas previously covered by an ice sheet likely resulted from other factors related to soil and land-relief characteristics, especially since the effects of the last glaciation are still visible in Poland's landscape (Marks et al. 2024).

Shortcomings in the data and methodological approach

The main problem is the quantification of the effect of sampling effort on observed SR in our species data set. As noted above, the character of the data restricted the possibility of quantifying sampling bias. Results from Czechia suggest that observed SR is biased due to differences in sampling effort, but the unbiased pattern obtained by rarefaction methods correlates well with the observed SR (Klímová et al. 2025). Therefore, it could be assumed that the data analyzed here also represent general trends in the spatial pattern of SR in Poland (Szymura et al. 2024), especially since the most likely undersampled squares were excluded from the analysis. We did not decide to use the modelled SR values to assess bias in our data set, as was done by, e.g. Klímová et al. (2025). This results from the modelling methodology applied here: the RF approach averages predictions across trees, which results in overall low values of errors (observed vs. estimated values), but also causes some over- and underestimation for the lowest and highest predicted values, respectively (Breiman 1999, Hastie et al. 2009). Methods exist to correct this bias (Belitz & Stackelberg 2021), but applying them would 'force' the modelled SR to approach the lowest and highest values in the original data set, thus limiting the applicability of modelled SR to assess sampling bias in undersampled squares.

Regarding the explanatory variables, landscape-scale modelling with a resolution of 100 km^2 is quite coarse regarding the spatial variability of some environmental variables, e.g. topography and soil properties, which seem to be finer (Siefert et al. 2012). As a result, averaging environmental variable values at a 100 km^2 square extent may have decreased the models' explanatory potential. A lack of information regarding temporal changes can also decrease the explanatory power of the models. We do not have data on temporal changes in either the response variable (species richness) or explanatory variables such as land-use patterns and agricultural intensification, and quantitative data on early human settlement distributions are also lacking. The results of previous studies in central Europe highlighted the role of temporal changes in land-use on SR (Klímová et al. 2025), as well as the correlation among prehistorical settlements, recent landscapes, and SR (Demján et al. 2022, Macek et al. 2023). Overall, the effect of explanatory variable averaging and the lack of data about very recent SR in areas with high farming intensity can be a reason why intensive agriculture was not among the factors negatively influencing SR. Finally, the study relies on correlative models, while most environmental variables are somehow intercorrelated (e.g. the content of nutrients in soils correlates also with fertilization); thus, it is hard to elucidate the effect of each individual factor.

The importance for nature conservation

The presented models help us better understand the ecological conditions for the distribution of vascular plant species richness hotspots in Poland and indicate the challenges of

protecting them at a landscape scale. Generally, the models underline the interaction of high soil productivity with diversified landscapes and topography as drivers of high SR for almost all species groups. The results indicate the necessity of conserving diversified land use, which is challenging since more productive lands are demanded for agriculture. Moreover, more productive lands, often placed in loess areas in Poland, have been relatively densely populated since the Neolithic period up to now. On the one hand, it increased the SR by, e.g. the presence of archaeophytes; on the other hand, it challenges landscape conservation as well as protection against invasive species, which also prefer such landscapes. Altogether, it provides a scientific rationale for the protection of plants in urbanized and agricultural landscapes. Considering the locations of 'blue spots' (areas with low species richness) reveals their large-scale ecological context: poor soils in areas less attractive to farmers. In such landscapes, the SR of vascular plants was never high, and low SR should not be a reason for omitting nature conservation.

Supplementary materials

Fig. S1. Spearman rank correlations among the explanatory data.
Fig. S2. Spatial distribution of anthropogenic explanatory variables included in the final models. The maps were produced using the Fisher method.
Fig. S3. Spatial distribution of explanatory variables representing climate included in the final models.
Fig. S4. Spatial distribution of soil explanatory variables included in the final models.
Fig. S5. Spatial distribution of topographic explanatory variables included in the final models.
Fig. S6. Observed vs. predicted scatterplot.
Fig. S7. Spatial distribution of observed and predicted species richness for native species, archaeophytes, neophytes, and red-list species.
Fig. S8. The partial dependence plot and relative influence of variables for the full model of native species richness.
Fig. S9. The partial dependence plot and relative influence of variables for the full model of archaeophytes richness.
Fig. S10. The partial dependence plot and relative influence of variables for the full model of neophytes richness.
Fig. S11. The partial dependence plot and relative influence of variables for the full model of red-list species richness.
Fig. S12. The partial dependence plot and relative influence of variables for the reduced model of native species richness.
Fig. S13. The partial dependence plot and relative influence of variables for the reduced model of archaeophyte richness.
Fig. S14. The partial dependence plot and relative influence of variables for the reduced model of neophyte richness.
Fig. S15. The partial dependence plot and relative influence of variables for the reduced model of red-list species richness.
Table S1. Description of the explanatory variables.
Table S2. Descriptive statistics of the explanatory variables used for modelling.
Table S3. The raw values of variable importance and associated P-values for initial models with all candidates' explanatory variables.
Table S4. Descriptive statistics of the full models and models without anthropogenic variables.

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

Acknowledgements

This work was supported by the National Science Centre, Poland, project number 2019/35/B/NZ8/00273, and funds from the Botanical Garden, University of Wrocław. We would like to thank Milan Chytrý and an anonymous reviewer for comments which helped us to improve this submission.

References

Altmann A., Toloşi L., Sander O. & Lengauer T. (2010) Permutation importance: a corrected feature importance measure. – Bioinformatics 26: 1340–1347.

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.

Ballabio C., Panagos P. & Montanarella L. (2016) Mapping topsoil physical properties at European scale using the LUCAS database. – Geoderma 261: 110–123.

Barthlott W., Hostert A., Kier G., Küper H., Kreft H., Mutke J., Rafiqpoor M. D. & Sommer J. H. (2007) Geographic patterns of vascular plant diversity at continental to global scales. – Erdkunde 61: 305–315.

Belitz K. & Stackelberg P. E. (2021) Evaluation of six methods for correcting bias in estimates from ensemble tree machine learning regression models. – Environmental Modelling & Software 139: 105006.

Błażejczyk K. (2006) Climate and bioclimate of Poland. – In: Degórski M. (ed.), Natural and human environment of Poland. A geographical overview, p. 31–48, Stanisław Leszczycki Institute of Geography and Spatial Organization, Polish Academy of Sciences, Polish Geographical Society, Warsaw.

Breiman L. (1999) Using adaptive bagging to debias regressions. Technical Report 547. – Statistics Department, University of California, Berkeley.

Breiman L. (2001) Random forests. – Machine Learning 45: 5–32.

Brooks T. M., Mittermeier R. A., da Fonseca G. A. B., Gerlach J., Hoffmann M., Lamoreux J. F., Mittermeier C. G., Pilgrim J. D. & Rodrigues A. S. L. (2006) Global biodiversity conservation priorities. – Science 313: 58–61.

Busacca A. J. & Sweeney M. R. (2005) Loess. – In: Hillel D. (ed.), Encyclopedia of soils in the environment, p. 364–373, Elsevier/Academic Press, Oxford.

Cai L., Kreft H., Taylor A., Denelle P., Schrader J., Essl F., van Kleunen M., Pergl J., Pyšek P., Stein A., Winter M., Barcelona J. F., Fuentes N., Inderjit, Karger D. N., Kartesz J., Kuprijanov A., Nishino M., Nickrent D., Nowak A., Patzelt A., Pelser P. B., Singh P., Wieringa J. J. & Weigelt P. (2023) Global models and predictions of plant diversity based on advanced machine learning techniques. – New Phytologist 237: 1432–1445.

Cervellini M., Zannini P., Di Musciano M., Fattorini S., Jiménez-Alfaro B., Rocchini D., Field R., Vetaas O. R., Irl S. D. H., Beierkuhnlein C., Hoffmann S., Fischer J.-C., Casella L., Angelini P., Genovesi P., Nascimbene J. & Chiarucci A. (2020) A grid-based map for the Biogeographical Regions of Europe. – Biodiversity Data Journal 8: e53720.

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.

Dąbkowska T., Grabowska-Orządała M. & Łabza T. (2017) The study of the transformation of segetal flora richness and diversity in selected habitats of southern Poland over a 20-year interval. – Acta Agrobotanica 70: 1712.

de Albuquerque F. S., Castro-Díez P., Rodríguez M. Á. & Cayuela L. (2011) Assessing the influence of environmental and human factors on native and exotic species richness. – Acta Oecologica 37: 51–57.

de Barros Ruas R., Costa L. M. S. & Bered F. (2022) Urbanization driving changes in plant species and communities: a global view. – Global Ecology and Conservation 38: e02243.

Demján P., Dreslerová D., Kolář J., Chuman T., Romportl D., Trnka M. & Lieskovský T. (2022) Long time-series ecological niche modelling using archaeological settlement data: tracing the origins of present-day landscape. – Applied Geography 141: 102669.

Divíšek J., Hájek M., Jamrichová E., Petr L., Večeřa M., Tichý L., Willner W. & Horsák M. (2020) Holocene matters: landscape history accounts for current species richness of vascular plants in forests and grasslands of eastern Central Europe. – Journal of Biogeography 47: 721–735.

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. (2009) Spatial species-richness gradients across scales: a meta-analysis. – Journal of Biogeography 36: 132–147.

Finnie T. J. R., Preston C. D., Hill M. O., Uotila P. & Crawley M. J. (2007) Floristic elements in European vascular plants: an analysis based on Atlas Florae Europaeae. – Journal of Biogeography 34: 1848–1872.

Firbank L. G., Petit S., Smart S., Blain A. & Fuller R. J. (2008) Assessing the impacts of agricultural intensification on biodiversity: a British perspective. – Philosophical Transactions of the Royal Society B 363: 777–787.

Firbas F. (1949) Spät- und nacheiszeitliche Waldgeschichte Mitteleuropas nördlich der Alpen. Bd. 1. Allgemeine Waldgeschichte. – Gustav Fischer, Jena.

Flohre A., Fischer C., Aavik T., Bengtsson J., Berendse F., Bommarco R., Ceryngier P., Clement L. W., Dennis C., Eggers S., Emmerson M., Geiger F., Guerrero I., Hawro V., Inchausti P., Liira J., Morales M. B., Oñate J. J., Pärt T., Weisser W. W., Winqvist C., Thies C. & Tscharntke T. (2011) Agricultural intensification and biodiversity partitioning in European landscapes: comparing plants, carabids, and birds. – Ecological Applications 21: 1772–1781.

Fotyma M., Ochal P. & Łabętowicz J. (2013) Potassium in soils of glacial origin. – In: Hernandez Soriano M. C. (ed.), Soil processes and current trends in quality assessment, p. 205–228, IntechOpen.

Friedman J. H. (2001) Greedy function approximation: a gradient boosting machine. – Annals of Statistics 29: 1189–1232.

Friedman J. H. & Meulman J. J. (2003) Multiple additive regression trees with application in epidemiology. – Statistics in Medicine 22: 1365–1381.

Giesecke T., Wolters S., van Leeuwen J. F. N., van der Knaap P. W. O., Leydet M. & Brewer S. (2019) Postglacial change of the floristic diversity gradient in Europe. – Nature Communications 10: 5422.

Głębocki B. (ed.) (2014) Zróżnicowanie przestrzenne rolnictwa [Spatial diversification of agriculture]. Powszechny Spis Rolny 2010. – Główny Urząd Statystyczny, Warszawa.

Gwiazdowicz M. (2010) Środowisko przyrodnicze na obszarach wiejskich – zagrożenia i szanse [The natural environment in rural areas – threats and opportunities]. – Studia BAS 4: 247–272.

Hastie T., Tibshirani R. & Friedman J. (2009) The elements of statistical learning: data mining, inference, and prediction. Ed. 2. – Springer, New York.

Hill M. O. (2012) Local frequency as a key to interpreting species occurrence data when recording effort is not known. – Methods in Ecology and Evolution 3: 195–205.

Hou Y., Li J., Li G. & Qi W. (2023) Negative effects of urbanization on plants: a global meta-analysis. – Ecology and Evolution 13: e9894.

Hulme P. E., Pyšek P., Nentwig W. & Vilà M. (2009) Will threat of biological invasions unite the European Union? – Science 324: 40–41.

Hulshof C. M. & Spasojevic M. J. (2020) The edaphic control of plant diversity. – Global Ecology and Biogeography 29: 1634–1650.

Jiménez-Alfaro B., Girardello M., Chytrý M., Svenning J.-C., Willner W., Gégout J.-C., Agrillo E., Campos J. A., Jandt U., Kącki Z., Šilc U., Slezák M., Tichý L., Tsiripidis I., Turtureanu P. D., Ujházyová M. & Wohlgemuth T. (2018) History and environment shape species pools and community diversity in European beech forests. – Nature Ecology & Evolution 2: 483–490.

Kabala C., Krupski M., Mackiewicz M., Przybył A., Dudek M. & Bednik M. (2025) Stages and chronology of Late Holocene transformation of Chernozems into Luvisols in the loess belt of south-west Poland. – CATENA 250: 108725.

Kącki Z. & Śliwiński M. (2012) The Polish Vegetation Database: structure, resources and development. – Acta Societatis Botanicorum Poloniae 81: 75–79.

Karger D. N., Conrad O., Böhner J., Kawohl T., Kreft H., Soria-Auza R. W., Zimmermann N. E., Linder H. P. & Kessler M. (2017) Climatologies at high resolution for the Earth's land surface areas. – Scientific Data 4: 170122.

Karpińska-Kołaczek M., Kołaczek P. & Stachowicz-Rybka R. (2014) Pathways of woodland succession under low human impact during the last 13,000 years in northeastern Poland. – Quaternary International 328–329: 196–212.

Kaźmierczakowa R., Bloch-Orłowska J., Celka Z., Cwener A., Dajdok Z., Michalska-Hejduk D., Pawlikowski P., Szczęśniak E. & Ziarnek K. (2016) Polska czerwona lista paprotników i roślin kwiatowych [Polish red list of pteridophytes and flowering plants]. – Instytut Ochrony Przyrody Polskiej Akademii Nauk, Kraków.

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.

Kreft H. & Jetz W. (2007) Global patterns and determinants of vascular plant diversity. – Proceedings of the National Academy of Sciences of the United States of America 104: 5925–5930.

Kühn I., Brandl R. & Klotz S. (2004) The flora of German cities is naturally species rich. – Evolutionary Ecology Research 6: 749–764.

Lehmkuhl F., Nett J. J., Pötter S., Schulte P., Sprafke T., Jary Z., Antoine P., Wacha L., Wolf D., Zerboni A., Hošek J., Marković S. B., Obreht I., Sümegi P., Veres D., Zeeden C., Boemke B., Schaubert V., Viehweger J. & Hambach U. (2021) Loess landscapes of Europe: mapping, geomorphology, and zonal differentiation. – Earth-Science Reviews 215: 103496.

Luck G. W. (2007) A review of the relationships between human population density and biodiversity. – Biological Reviews 82: 607–645.

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.

Maes D. & van Swaay C. A. M. (1997) A new methodology for compiling national Red Lists applied to butterflies (Lepidoptera, Rhopalocera) in Flanders (N-Belgium) and the Netherlands. – Journal of Insect Conservation 1: 113–124.

Marks L., Bitinas A., Błaszkiewicz M., Börner A., Guobyte R., Rinterknecht V. & Tylmann K. (2024) Glacial landscape evolution during the Holocene in Northern Central Europe. – In: Palacios D., Hughes P., Jomelli V. & Tanarro L. M. (eds), European glacial landscapes: the Holocene, p. 295–314, Elsevier, Amsterdam.

Meeus J. H. A. (1995) Pan-European landscapes. – Landscape and Urban Planning 31: 57–79.

Mengel K. & Kirkby E. A. (1980) Potassium in crop production. – Advances in Agronomy 33: 59–110.

Mirek Z., Piękoś-Mirkowa H., Zając A. & Zając M. (2020) Vascular plants of Poland: an annotated checklist. – W. Szafer Institute of Botany, Polish Academy of Sciences, Kraków.

Mittelbach G. G. & Schemske D. W. (2015) Ecological and evolutionary perspectives on community assembly. – Trends in Ecology & Evolution 30: 241–247.

Moskal-del Hoyo M. (2021) Open canopy forests of the loess regions of southern Poland: a review based on wood charcoal assemblages from Neolithic and Bronze Age archaeological sites. – Quaternary International 593–594: 204–223.

Newman A. C. D. (1984) The significance of clays in agriculture and soils. – Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences 311: 375–389.

Normand S., Ricklefs R. E., Skov F., Bladt J., Tackenberg O. & Svenning J.-C. (2011) Postglacial migration supplements climate in determining plant species ranges in Europe. – Proceedings of the Royal Society B 278: 3644–3653.

Nowak M. (2013) Neolithisation in Polish territories: different patterns, different perspectives, and Marek Zvelebil's ideas. – Interdisciplinaria Archaeologica Natural Sciences in Archaeology 4: 85–96.

Pärtel M., Bennett J. A. & Zobel M. (2016) Macroecology of biodiversity: disentangling local and regional effects. – New Phytologist 211: 404–410.

Pautasso M. (2007) Scale dependence of the correlation between human population presence and vertebrate and plant species richness. – Ecology Letters 10: 16–24.

Pyšek P., Jarošík V., Hulme P. E., Kühn I., Wild J., Arianoutsou M., Bacher S., Chiron F., Didžiulis V., Essl F., Genovesi P., Gherardi F., Hejda M., Kark S., Lambdon P. W., Desprez-Loustau M.-L., Nentwig W., Pergl J., Poboljšaj K., Rabitsch W., Roques A., Roy D. B., Shirley S., Solarz W., Vilà M. & Winter M. (2010) Disentangling the role of environmental and human pressures on biological invasions across Europe. – Proceedings of the National Academy of Sciences of the United States of America 107: 12157–12162.

R Core Team (2022) R: A language and environment for statistical computing. – R Foundation for Statistical Computing, Vienna.

Rijal D. P., Heintzman P. D., Lammers Y., Yoccoz N. G., Lorberau K. E., Pitelková I., Goslar T., Murguzur F. J. A., Salonen J. S., Helmens K. F., Bakke J., Edwards M. E., Alm T., Bråthen K. A., Brown A. G. & Alsos I. G. (2021) Sedimentary ancient DNA shows terrestrial plant richness continuously increased over the Holocene in northern Fennoscandia. – Science Advances 7: eabf9557.

Rudnicki R., Kluba M. & Biczkowski M. (2016) Struktura ciągników według mocy a struktura obszarowa gospodarstw rolnych w Polsce – analiza porównawcza wyników Powszechnego Spisu Rolnego 2010 [Structure of tractors according to their power and area structure of agricultural holdings in Poland – a comparative analysis based on National Agricultural Census 2010]. – Roczniki Naukowe Ekonomii Rolnictwa i Rozwoju Obszarów Wiejskich 103: 112–125.

Ryzner K. & Owczarek P. (2020) Outline of the development of research on the impact of Neolithic settlements on the transformation on loess landscapes in Southern Poland. – Environmental & Socio-economic Studies 8: 32–44.

Sabatini F. M., Jiménez-Alfaro B., Jandt U., Chytrý M., Field R., Kessler M., Lenoir J., Schrodt 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., Peet R. K., Petřík P., Pillar V. D., Sandel B., Schmidt M., Tang Z., van Bodegom P., Vassilev K., Violle C., Alvarez-Davila E., Davidar P., Dolezal J., Hérault B., Galán-de-Mera A., Jiménez J., Kambach S., Kepfer-Rojas S., Kreft H., Lezama F., Linares-Palomino R., Monteagudo Mendoza A., N'Dja J. K., Phillips O. L., Rivas-Torres G., Sklenář P., Speziale K., Strohbach B. J., Vásquez Martínez R., Wang H. F., Wesche K. & Bruelheide H. (2022) Global patterns of vascular plant alpha diversity. – Nature Communications 13: 4683.

Sardans J. & Peñuelas J. (2015) Potassium: a neglected nutrient in global change. – Global Ecology and Biogeography 24: 261–275.

Siefert A., Ravenscroft C., Althoff D., Alvarez-Yépiz J. C., Carter B. E., Glennon K. L., Heberling J. M., Jo I. S., Pontes A., Sauer A., Willis A. & Fridley J. D. (2012) Scale dependence of vegetation–environment relationships: a meta-analysis of multivariate data. – Journal of Vegetation Science 23: 942–951.

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.

Szafer W. (ed.) (1959) Szata roślinna Polski [Vegetation of Poland]. – Państwowe Wydawnictwo Naukowe, Warszawa.

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., Tegegne H. & Szymura M. (2024) Spatial patterns of vascular plant species richness in Poland: relations among species group richness and hot spot distribution. – Acta Societatis Botanicorum Poloniae 93: 192897.

Tokarska-Guzik B., Dajdok Z., Zając M., Zając A., Urbisz A., Danielewicz W. & Hołdyński C. (2012) Rośliny obcego pochodzenia w Polsce ze szczególnym uwzględnieniem gatunków inwazyjnych [Alien plants in Poland with particular reference to invasive species]. – Generalna Dyrekcja Ochrony Środowiska, Warszawa.

Urban M. C. (2015) Accelerating extinction risk from climate change. – Science 348: 571–573.

Wacnik A., Goslar T. & Czernik J. (2012) Vegetation changes caused by agricultural societies in the Great Mazurian Lake District. – Acta Palaeobotanica 52: 59–104.

Wacnik A., Tylmann W., Bonk A., Goslar T., Enters D., Meyer-Jacob C. & Grosjean M. (2016) Determining the responses of vegetation to natural processes and human impacts in north-eastern Poland during the last millennium: combined pollen, geochemical and historical data. – Vegetation History and Archaeobotany 25: 479–498.

Wright M. N. & Ziegler A. (2017) Ranger: A fast implementation of random forests for high dimensional data in C++ and R. – Journal of Statistical Software 77: 1–17.

Zając A. (1978) Atlas of distribution of vascular plants in Poland (ATPOL). – Taxon 27: 481–484.

Zając A. & Zając M. (eds) (2001) Atlas rozmieszczenia roślin naczyniowych w Polsce [Distribution atlas of vascular plants in Poland]. – Laboratory of Computer Chorology, Institute of Botany, Jagiellonian University, Cracow.

Zając A. & Zając M. (eds) (2019) Atlas rozmieszczenia roślin naczyniowych w Polsce: dodatek [Distribution atlas of vascular plants in Poland: appendix]. – Instytut Botaniki Uniwersytetu Jagiellońskiego, Kraków.

Zając M. & Zając A. (2011) Directional northern element in the flora of vascular plants of Poland. – Biodiversity Research and Conservation 21: 19–26.

Zawadzki S. (ed.) (1999) Gleboznawstwo [Soil science]. – Państwowe Wydawnictwo Rolnicze i Leśne (PWRiL), Warszawa.

Zentrum für Baltische und Skandinavische Archäologie (2019) EPHA: European paleoecological and historic atlas. Last Glacial Maximum. – URL: https://zbsa.eu/en/last-glacial-maximum.

Predikторy druhové bohatosti polské flóry

Modelovali jsme druhovou bohatost cévnatých rostlin v Polsku ve čtvercové síti o rozměrech 10 × 10 km; celé území pokrývá celkem 2 866 čtverců. Druhová bohatost byla vyhodnocena samostatně pro původní druhy, archeofyty, neofyty a druhy uvedené v červeném seznamu. Předpokládali jsme, že lidská činnost a prehistorické faktory ovlivňují různé skupiny druhů v různé míře. Zahrnutí antropogenních proměnných podstatně zlepšilo vypovídací schopnost modelů pro neofyty a archeofyty, nikoli však pro původní druhy. Odebrání antropogenních proměnných z modelů zvýšilo rozptyl vysvětlený půdou a topografií, což naznačuje vliv prostředí na antropogenní aktivitu. Obsah draslíku (K) a podíl jílu v ornici měly konzistentně největší vliv na všechny skupiny druhů. Druhová bohatost byla nízká při nízkých hodnotách K a jílu, ale téměř lineárně se zvyšovala až do dosažení ~150 $mg \cdot kg^{-1}$ K a ~15 % jílovité frakce, poté došlo ke stabilizaci. Uhličitan vápenatý měl relativně nízký, ale zřetelně pozitivní vliv na druhovou bohatost, zejména u druhů uvedených v červeném seznamu. Rozmanitost krajiny pozitivně ovlivnila počet původních druhů a druhů uvedených v červeném seznamu. Přítomnost sprašových sedimentů je spojena s vysokým počtem druhů všech skupin, zatímco oblasti pokryté ledovými příkrovy během posledního zalednění vykazovaly sníženou bohatost pro všechny druhové skupiny kromě neofytů. Půdní charakteristiky v Polsku sledují jasný severojižní gradient formovaný historií zalednění a souvisejícími procesy tvorby půdy. Silný pozitivní vztah mezi lidskou populací a druhovou bohatostí je dán větším přísunem diaspor nepůvodních druhů v urbanizovaných oblastech a tendencí lidí usazovat se v oblastech s přirozeně vysokou diverzitou původních druhů. Podobný dopad má i zavlékání archeofytů; průměrná velikost farmy negativně korelovala s druhovou bohatostí, ačkoli tento vztah v Polsku spíše odráží některé zvláštnosti socioekonomického dědictví vlastnictví půdy a zemědělských systémů než přímé kauzální účinky.

How to cite: Szymura T. H., Kassa H., Chmolowska D., Szymura M., Swacha G., Kącki Z. & Zając A. (2026) Predictors of vascular plant species richness in Poland. – Preslia 98: 241–263.

Preslia, a journal of the Czech Botanical Society

https://www.preslia.cz