

Pollen anthropogenic indicators revisited using large-scale pollen and archaeological datasets: 12,000 years of human-vegetation interactions in central Europe

Vojtěch Abraham^{1,2*}, Martin Macek³, Peter Tkáč^{2,4}, David Novák⁵, Petr Pokorný^{6,7}, Radka Kozáková⁸, Eva Jamrichová^{9,10}, Markéta Gabriela Soukupová⁹ & Jan Kolář^{2,4,11}

¹Department of Botany, Faculty of Science, Charles University, Benátská 2, CZ-12800 Prague, Czech Republic; ²Department of Vegetation Ecology, Institute of Botany of the Czech Academy of Sciences, Lidická 25/27, CZ-60200 Brno, Czech Republic; ³Department of Geoecology, Institute of Botany of the Czech Academy of Sciences, Zámek 1, CZ-25243 Průhonice, Czech Republic; ⁴Department of Archaeology and Museology, Faculty of Arts, Masaryk University, Arne Nováka 1, CZ-60200 Brno, Czech Republic; ⁵Department of Information Sources and Landscape Archaeology, Institute of Archaeology of the Czech Academy of Sciences, Prague, Letenská 123/4, CZ-11801 Prague, Czech Republic; ⁶Center for Theoretical Study, Charles University and the Czech Academy of Sciences, Jilská 1, CZ-11000 Prague, Czech Republic; ⁷Nuclear Physics Institute, Czech Academy of Sciences, Husinec 130, CZ-25068 Řež, Czech Republic; ⁸Department of Natural Sciences and Archaeometry, Institute of Archaeology of the Czech Academy of Sciences, Prague, Letenská 123/4, CZ-11801 Prague, Czech Republic; ⁹Department of Palaeoecology, Institute of Botany of the Czech Academy of Sciences, Lidická 25/27, CZ-60200 Brno, Czech Republic; ¹⁰Department of Botany and Zoology, Faculty of Science, Masaryk University, Kotlářská 2, CZ-61137 Brno, Czech Republic; ¹¹Institute of Archaeology, University College London, 31-34 Gordon Square, London, WC1H 0PY, United Kingdom

*corresponding author: vojtech.abraham@gmail.com

Abstract: Tracing human-vegetation interactions that occurred in the past has always been one of the key topics of paleoecology. Here we use the pollen and archaeological databases available for the Czech Republic to determine links between individual pollen taxa and archaeological data and search for the spatial scales of comparability. The datasets include 1,500 pollen samples and 65,000 archaeological components covering the period from 12,000 to 700 cal. BP, divided into time windows of 250 years. Spearman's rank correlation was used to measure the link between pollen and archaeological data at different sites. Using generalized additive models for the whole dataset, we explained the variance of pollen by archaeologically registered human activities and by two environmental variables. The first was the overall trend for each taxon in the Holocene representing the long-term dynamics of the species, the second was the elevation of pollen sites. Both factors affect species representation over the whole period studied or/and the area and cannot be statistically separated from human-induced changes. Both decrease the indicative strength of anthropogenic pollen; however, elevation did so more than the Holocene trend, since past human activities and elevation are strongly correlated and account for the first main gradient. The pollen taxa with a positive correlation with the level of past human activity, indicated by all methods, are: *Plantago lanceolata*, *Artemisia* and *Amaranthaceae*, resprouting edible trees that tolerate fire and pruning (*Quercus*) and pioneer trees (*Pinus*). Probability indicating the presence or absence of archaeological evidence when pollen of these species is present or absent is high (0.56–0.76). However, explained variability by the full model is

low (0.01–0.09). *Fagus*, *Carpinus* and *Abies* expand during the late-successional stages after human disturbance, therefore their relationships to past human activity are negative when considering a 250-year time window. *Secale* does not correlate at the level of individual sites due to its late appearance during the Holocene. We ascribe the weak relationship between archaeological data and pollen of *Cerealia* to inconsistent determinations. The radius of comparability of pollen and archaeological evidence is around tens of kilometres due to the spatial resolution of archaeology is the area of a parish, but lower for herbaceous plants (15–20 km) than for trees (30–40 km). This critical comparison delimits overlaps and gaps between widely-used assumptions and data-based evidence.

Keywords: Anthropocene, pollen analysis, archaeology, settlement history, anthropogenic indicators, human impact, *Cerealia*, woodland management, Czech Republic

Introduction

The increasing human impact on the environment during the Anthropocene (Boivin & Crowther 2021) requires a better understanding of the complex socioecological relationships that existed in the distant past. The transition from hunting and gathering to agropastoral economies included significant land-use transformations. Management of anthropogenic habitats such as fields, gardens and pastures in open landscapes over-ruled the climatic drivers of environmental change (Marquer et al. 2017) and resulted in an increase in biodiversity already in the Neolithic (Colombaroli et al. 2013, Šizling et al. 2016, Kuneš et al. 2019a, Lamentowicz et al. 2019, Roleček et al. 2021, Woodbridge et al. 2021). The increased biodiversity due to the human land use was accompanied by the introduction of species such as domesticated plants and weeds (Pokorná et al. 2018, Macek et al. 2022). Anthropogenic disturbances favoured the spontaneous spread of weeds, which originated in the same areas as the wild predecessors of cultivated plants or in disappearing native open habitats and so survived the Holocene forest optimum (Pokorný et al. 2015, Hájek et al. 2016). Human activities, however, did not only actively spread the herbs. Depopulation and abandoning of previously settled and managed habitats for various reasons (e.g. migration, high mobility, famines, violent conflicts) provided niches for new tree taxa (Gillson et al. 2008). The spread of *Fagus* (Küster 1997), *Carpinus* (Ralska-Jasiewiczowa et al. 2003) and *Abies* (Kozáková et al. 2011) during Late Holocene was triggered by human activity (Giesecke et al. 2017). Anthropogenic disturbance influenced the species composition of natural habitats such as deciduous forests and became a key factor driving vegetation development in central Europe.

Finding evidence of past human settlements was a key topic of archaeology until the beginning of the 20th century. The latter development of pollen analysis introduced another way of searching for signals of anthropogenic vegetation that indicates the presence of humans in the past (Firbas et al. 1939). Pollen assemblages document changes in land cover over time, which is indicative of land use at a particular time (Behre 1981). The interpretation of anthropogenic pollen indicators greatly enriched the outputs of pollen analysis (Gaillard 2007). Pollen of crops have the strongest indicative value, as they provide unambiguous evidence of human agricultural activity and are therefore often called “primary anthropogenic pollen indicators” (Lang 1994). “Secondary anthropogenic pollen indicators” are assumed to originate from open or woodland pastures, clearings,

paths, settlements or meadows, fallow – i.e. habitats that are either human-induced or a result of natural disturbances (animals, windthrows, wildfires). In addition to the indirect ecological link between some secondary anthropogenic species and human activity, the low taxonomic resolution of pollen analysis decreases the indicative strength of anthropogenic pollen indicators. Refining the selection of indicator pollen within an area of interest can partly overcome these problems. Gaillard et al. (1992) and Hjelle (1999) measured and quantified the strength of the relationship between such indicators and recent land use using a set of sites in the present vegetation of southern Sweden and western Norway, respectively.

Extending this approach to the southern part of central Europe provided slightly different sets of anthropogenic indicators, and differences could be also expected between archaeological periods (Brun 2011). Moreover, the relationship between pollen and land use can be obscured by steep precipitation and altitudinal gradients if the results from sites in the lowlands and mountains are analysed together (Court-Picon et al. 2005). In summary, temporal and spatial scales and the variations of environmental factors may strongly influence the reliability of secondary anthropogenic indicators. Syntheses of human impact on past vegetation at large spatial scales (Morrison et al. 2021) might require regionally more specific anthropogenic indicators independent of considerations of environmental variables.

Vegetation during the Holocene was affected by both climate (Beckage et al. 2020) and human societies. Therefore, next to paleoclimatic data, good quality archaeological data is crucial for the study of Holocene vegetation. A synchronous signal of particular pollen types and archaeological evidence of human activity is frequently reported for sites with palaeoecological archives (e.g. Dreslerová et al. 2013). Comparisons that challenge pollen anthropogenic indicators at broad spatial and temporal scales require large-scale archaeological databases covering thousands of years, however they are not available for most countries or they have only a limited spatial extent (cf. Richards et al. 2021). Some studies use rough estimates of human population sizes and technological advances (e.g. Deza-Araujo et al. 2020, Milligan et al. 2020). Nevertheless, as Riede (2018) stresses, data-driven archaeological approaches are the most useful for studying human-environment interactions.

The most widespread proxy for human population size or the intensity of human activity is the summed probability distribution of radiocarbon dates of archaeological evidence (Hinz et al. 2012, Lechterbeck et al. 2014, Whitehouse et al. 2014, Kolář et al. 2022), which has been criticized (Contreras & Meadows 2014, Crombé & Robinson 2014, Torfing 2015) and also corrected (Timpson et al. 2014, Edinborough et al. 2017, Crema & Bevan 2021). Alternatively, models based on numbers of archaeological sites have been successfully used for similar purposes (Feaser & Furrholt 2014, Kolář et al. 2018). Results of both these data-driven approaches combined with knowledge of human behaviour (e.g. subsistence strategies, social organization or woodland management practices), have a significant research advantage (e.g. Kolář et al. 2022).

The aim of this paper is to explore long-term ecological interactions between vegetation, represented by individual pollen taxa, and prehistoric human populations. Considering the relationship with human activity, the pollen taxa were arranged in a spectrum ranging from cultivated to wild plants. It is hypothesized that anthropogenic indicators should show a positive relationship with archaeologically-detected human activity.

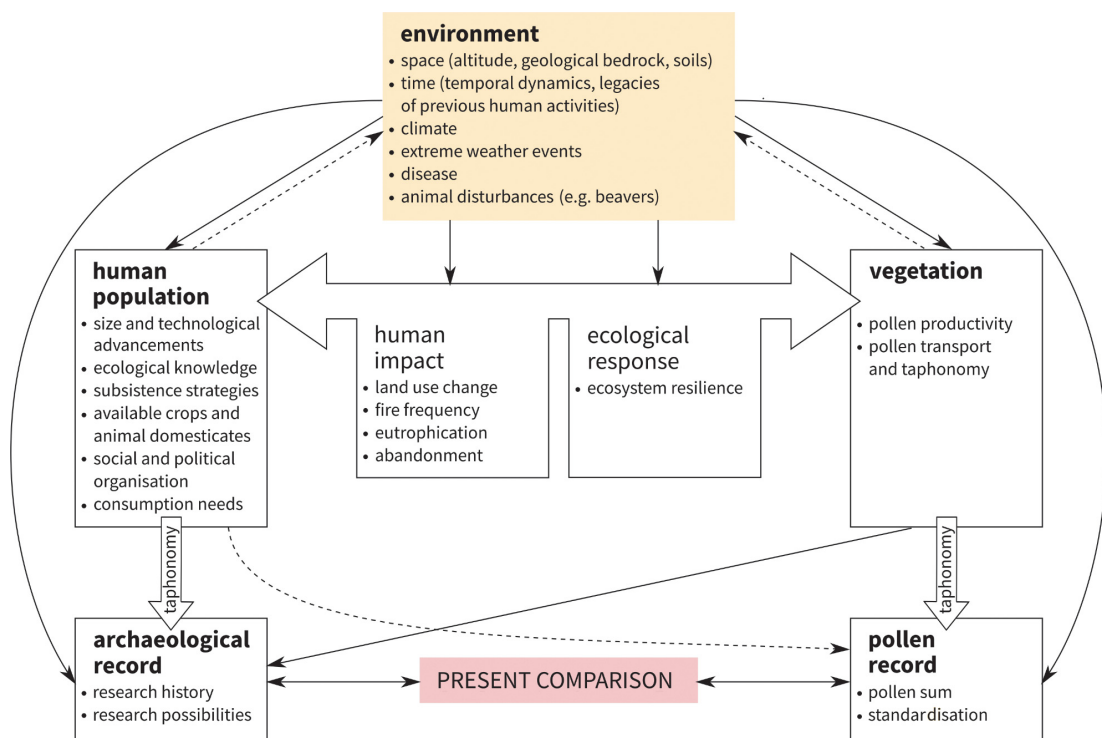


Fig. 1. The physical and abstract entities (bold) and factors (normal font) in the relationship between pollen and archaeological records.

Generalized additive models were used to determine the association between spatial and temporal variability in the abundance of individual pollen taxa and past settlement intensity based on archaeological evidence. Dealing with an extensive, but heterogeneous palynological dataset and archaeological evidence posed several methodological problems. First, we had to find the optimal spatial scale at which archaeologically-detected human activity fits vegetation documented in the fossil pollen record. Secondly, how to account for differences in the pollen productivity and sampling intensity, and their effect on the recording of rare pollen types (Fig. 1). Finally, we had to consider the confounding environmental effects and temporal trends that affected the Holocene vegetation development and anthropogenic pollen indicator abundances.

The main questions addressed were: (i) What is the optimum spatial scale for comparing pollen and archaeological data? (ii) Can archaeological proxies explain the variability in individual pollen taxa when considering the overall Holocene trends in pollen types and the elevation of pollen sites? (iii) Is there a positive relationship between anthropogenic pollen indicators, established in the literature, and the archaeological record for the Czech Republic?

Material and methods

Study area

The spatial and temporal relationships between the fossil pollen record and evidence of past human settlements in the Czech Republic, central Europe (Fig. 2), for which there are digital large-scale datasets, were studied. This country is dominated by the temperate forest biome, but the lowlands also contain biomes of hemiboreal taiga and forest-steppe biomes; and in the mountains encompass vegetation of coniferous spruce forests, and limited areas of alpine tundra at the highest elevations (Chytrý 2012). Currently, forests cover 34% of the area, 57% is covered by agricultural land and 6% is covered by urban areas.

Postglacial vegetation history of the area studied was shaped by the spread of trees, at the beginning of the Holocene by pine and birch, around 10,300 cal. BP by hazel, and subsequently by species of mixed oak woodlands (oak, lime, ash and elm). Spruce appeared during the Early Holocene, spread in the mountains from the east to the west of the Czech Republic and colonized the lowlands in the Late Holocene. With the onset of the Late Holocene around 4,200 cal. BP, fir, beech and hornbeam started to spread (Kuneš & Abraham 2017).

The prehistory and early historic periods of the Czech Republic were traditionally studied separately for two historic lands: Bohemia (the western two-thirds of the country) and Moravia including Czech Silesia (the eastern third of the country). The prehistory and early history of settlements and demographic development was only recently studied for the entire Czech Republic (Demján et al. 2022, Kolář et al. 2022). The first five thousand years of the Holocene overlap with the presence of Mesolithic hunter-gatherers. From 7,350 cal. BP onwards, the fertile lowlands were used by farmers. The local Linearbandkeramik (LBK) farmers were part of a large-scale pan-European transition to agro-pastoral economies, which resulted in a great increase in the number of permanent settlements, which indicates a possible increase in the population. Eneolithic societies, although probably experiencing a significant decrease in population between 6,000 and 5,000 cal. BP, were characterized by technological innovations (plough, metallurgy, wheel and animal traction) and the beginning of regular burials. The Early Bronze Age in the Czech Republic (starting ca. 5,250 cal. BP) most likely resulted in an increase in the population, connected again to social transformations and technological innovations. Prehistoric archaeological evidence is most abundant between 3,450 cal. BP and 1,950 cal. BP, i.e. during the Late Bronze Age, Hallstatt and La Tène Periods. Humans colonized areas at high altitudes and societies became more hierarchical. The building of large hillforts and oppida (fortified settlements), indicates the beginning of urbanization. The subsequent Roman and Migration Periods saw a significant decrease in the number of settlements and the abandonment/depopulation of some areas (e.g. south Bohemia). The area of the current Czech Republic was just outside the Roman Empire, but still closely connected. This is especially true for Moravia with its rich evidence of military camps and permanent forts. The Early Mediaeval period after ca. 1,350 cal. BP was the formation period for the Great Moravian Empire (with large proto-urban centres in south Moravia) and Přemyslid Bohemia (with its early centres in central Bohemia and Prague; Podborský et al. 1993, Jirán & Venclová 2008).

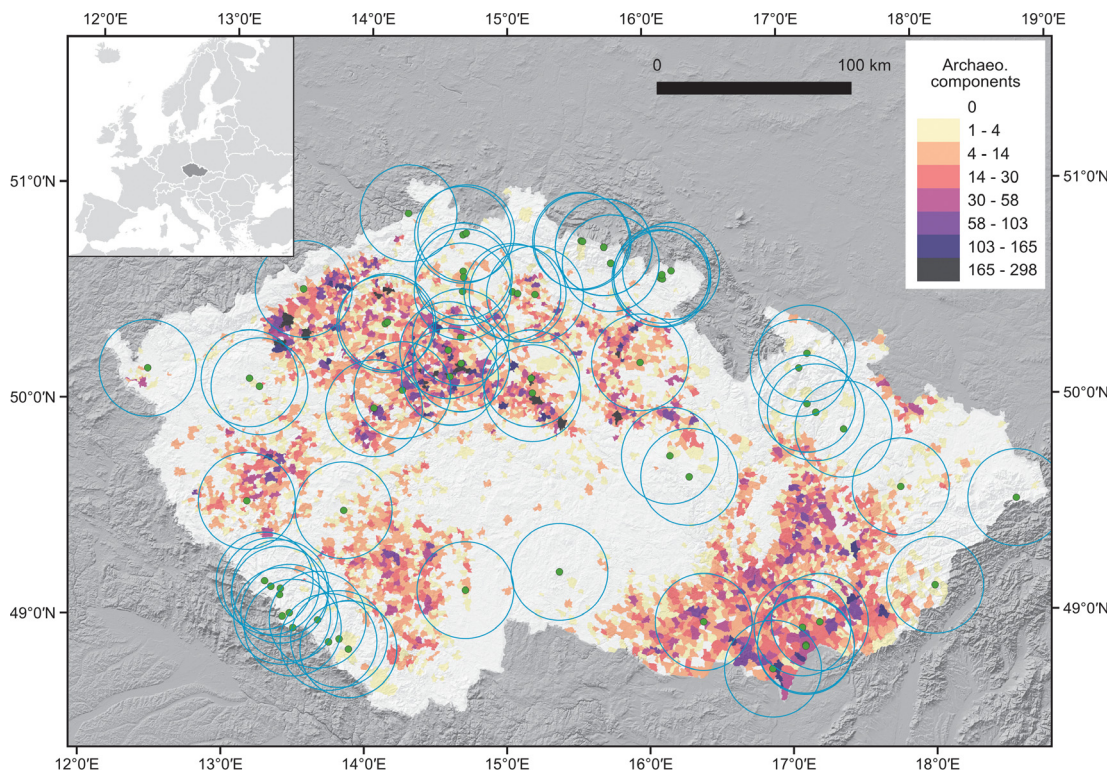


Fig. 2. Map showing the locations of pollen profiles (green symbols) within the region studied (Czechia) and the influence area represented by the 25 km buffer zone around pollen profiles (blue circles). Colour gradient depicts the total number of archaeological components dated between 12,000 and 700 cal. BP summarized at the civil parish level. Hillshade based on SRTM DEM, geographic projection: UTM 33N.

Fossil pollen data (8,075–825 cal. BP)

Pollen sequences were obtained from the PALYCZ database (<https://botany.natur.cuni.cz/palycz>; Kuneš et al. 2009) with a chronological control based on at least two radiocarbon dates and the geographic location restricted to the area studied (Czech Republic). Following Abraham et al. (2016), samples biased by unknown taphonomy and pollen-based chronologies were excluded. Since classical and Bayesian depth-age models, in which the error bars have different meanings were used, the criteria for chronological control were density of radiometric dates per temporal span covered by each sequence. Each 2,000 years of sedimentation must be confirmed by at least one radiocarbon date. These selection criteria resulted in 75 pollen sequences (Supplementary Table S1). Finally, all the sequences in 250-year time windows were aggregated.

The comparison with archaeological data was based on the proportional pollen counts of target taxa in the sample total, which does not provide independent values of taxa included in the assemblage. Only the pollen of plants from terrestrial habitats were considered, since wetland species tend to be heavily over-represented in the palaeoecological

record due to their local occurrence at the site sampled (Rybníčková & Rybníček 1971). It is known that taxa with high pollen productivity confound results because their abundance in the pollen record is higher than their abundance within the landscape (Odgaard 1999). As a consequence, if the pollen sum is low, these species decrease the evenness of the pollen spectra and decrease the detection probability of less abundant pollen types, which includes many anthropogenic indicators (Hellman et al. 2009). To control for these effects, we created three pollen data sets. The first set was composed of unadjusted relative pollen counts of terrestrial species (APNAP). The second contained the relative pollen counts of herbaceous plants and trees adjusted by an estimate of their pollen productivity (APNAPpe) based on their average pollen productivity values recorded in the Northern Hemisphere (Supplementary Table S2; Wiczorek & Herzsuh 2020). This approach can reduce the effect of tree pollen in forest samples, so that they provide a signal comparable to that of rare herbaceous plant pollen in samples from open habitats (Roleček et al. 2021). The third included only the relative pollen counts of trees (AP); the separate modelling of tree taxa pollen allowed the assessment of relative changes in the species composition of forest trees, irrespective of changes in the extent of the forested area.

The names of pollen taxa were unified according to Beug (2004). The only exceptions are the genera *Centaurea* and *Rumex*, whose pollen taxonomy follows Punt & Hoen (2009) and van Leeuwen et al. (1988). For further analysis only taxa occurring in at least 30 samples were included. Along with the relative proportions of individual pollen taxa, the models used the joint proportions of anthropogenic indicators and trees in the pollen spectra. For this purpose, herbaceous plants and shrub taxa are commonly considered as anthropogenic indicators (e.g. by Kozáková et al. 2015): such as *Cerealia*, *Secale cereale*, *Cannabaceae*, *Agrostemma githago*, *Arctium*, *Centaurea cyanus* type, *Hyoscyamus*, *Matricaria*-Typ, *Polygonum persicaria*-Typ, *Polygonum aviculare*-Typ, *Solanum nigrum*, wild grasses, *Artemisia*, *Amaranthaceae*, *Urtica*, *Plantago lanceolata*-Typ, *P. major-media*-Typ, and *Rumex acetosa* type. Following the findings of Kozáková et al. (2015) and Svoboda et al. (2018), *Calluna vulgaris*, *Melampyrum*, *Sambucus nigra*-Typ and *Trifolium* were added.

Archaeological data

The data on past human activity in the Czech Republic came from two large-scale databases. Firstly, the archaeological database of the Institute of Botany of the Czech Academy of Sciences provided the dataset covering Moravia and Czech Silesia (Kolář et al. 2016b). Secondly, the Archaeological Map of the Czech Republic (<https://digiarchiv.aiscr.cz>; Kuna et al. 2015) was used, which covers the western part of the Czech Republic (Bohemia). The temporal coverage of both datasets was limited to between 12,000 and 950/700 cal. BP. These datasets were merged into one dataset covering the whole country, while records with insufficiently accurate dating (dating stretching over several periods or assigned only vaguely to prehistory) were omitted. The basic analytical unit in the final dataset are the archaeological components, which are functionally defined (settlement, burial ground, hillfort, etc.) and spatially localized to a civil parish or a site, and temporally assigned to an archaeological period, chronological phase, or archaeological culture. The whole dataset used in this study includes 65,473 archaeological components,

which are, of course, not evenly distributed in space and time, but concentrated (Fig. 2), indicating hotspots of archaeological records of human activities.

To link archaeological evidence with the pollen record, the spatially-weighted probability of human settlement was associated with the location of pollen profiles in 250-year time windows using a spatial-temporal modelling approach (Kolář et al. 2016a). Here, a permutation of 1,000 replications was used for randomly assigning a year to each archaeological component located within the buffer zone the values of which depended on the archaeological period, chronological phase, or archaeological culture. In the next step, the proportional area of the parishes within the buffer area regarded to be occupied in the given time period was evaluated (i.e. parishes containing at least one component). Finally, the mean proportional area of the occupied parishes (“occupancy”) from the 1,000 permutations was determined for each buffer zone and time window, which was used in subsequent analyses.

Because human societies affect the wider surroundings of their settlements and the pollen source area may differ considerably according to pollen dispersal ability and taphonomy, the relationship between fossil pollen and human occupancy at multiple spatial scales, represented by buffers zones around pollen profiles of 5, 10, 15, 20, 25, 30, 40, and 50 km radius was evaluated. For each buffer zone and time frame, the mean occupancy as described above, was determined.

Statistical modelling

To identify pollen taxa sensitive to human presence within the landscape and the role of context dependency, two statistical modelling approaches were used. Pollen signals at individual sites should be considered in the context of adjacent records in order to avoid overinterpretation of the fuzzy character of the pollen record. However, consideration of random effects in the preliminary runs of the generalized additive models (GAM) accounted for most of the variability, so an alternative approach employing site-by-site correlation, in addition to the GAM analysis, was used.

Firstly, the relationships between pollen proportions and archaeological occupancy for each species at each site was determined using Spearman rank correlation test, which is suitable for situations in which there are outliers and nonlinear percentage data, and Spearman’s rho values were recorded. Then, for each taxon, whether the observed number of positive and negative correlations for individual sites deviated from random was tested using a binomial test. The significant prevalence of sites with positive or negative correlations at the alpha level of 0.05 were visualized.

Secondly, a generalized additive modelling approach (GAM), was used to develop complex models of pollen proportion data for each taxon, which take into account nonlinear relations with environmental proxies (site elevation) and inherent temporal trends driven by external forces (climate, delayed migration), using the ‘mgcv’ R package (Wood 2013). Restricted maximum likelihood (REML) was used for optimizing the parameters. To deal with proportional data, the beta error family was used.

To evaluate the robustness of the results when sampling intensity varied, an additional eight pollen sets were obtained from the original dataset, using resampling without replication to obtain equal target sums of 10, 50, 100, 150, 200, 300, 400 or 500 pollen grains per sample (pollen grains from the same locality and time-window). Resampling was

repeated 100 times, and pollen richness in each run was recorded. A replicate with median richness was used in subsequent analyses (Roleček et al. 2021). Samples that contained less than 500 pollen grains were not included in the resampling. Resampling always results in loss of information, therefore we primarily present results based on original set with no resampling, denoted as “Inf”.

For each taxon, pollen set and buffer zone, six GAM models using different combinations of predictors for the proportion of the pollen of the target taxa were fitted. The full model contained local archaeological occupancy in the selected buffer zone around the pollen profile as a linear parametric predictor, as well as smooth terms for site elevation and time as years BP/1,000 (model AEY; Archaeology-Elevation-Year). Smooth terms used cubic regression splines and a base dimension set to six. Further, reduced models containing only a single predictor of archaeological occupancy, elevation or time as a single predictor (models A, E or Y), or archaeological occupancy in combination with smooth terms for elevation (model AE; Archaeology-Elevation) or time (model AY; Archaeology-Year) were fitted. To force comparability of the models, the most complex model was fitted first with all predictors and extracted smoothing parameters, which were then used in the reduced models. Finally, adjusted R^2 of the models and the difference between hierarchical sets of models, with and without temporal or environmental covariates, were calculated. The significant taxa, which remained in the Archaeology-Elevation-Year model using the difference in the adjusted R^2 of the Elevation-Year and Archaeology-Elevation-Year models, were explored. The relationship between the three predictors in Non-Metric Multidimensional Scaling was visualized (Oksanen et al. 2020).

Reliable anthropogenic pollen indicators in landscapes should correlate with human presence at individual sites and also independently of other confounding factors. Spurious correlations may result from the confounding effects of collinear predictors, in this case collinearity between archaeological occupancy and altitude, or archaeological occupancy and time. Next, the significance of the linear correlation between archaeological occupancy and the pollen proportion of target taxa was evaluated. Finally, the unique contribution of archaeological occupancy to the explained variability by comparison with the reduced model without a term for archaeological data was calculated. The archaeological and pollen datasets and R codes that were used are available at the following link via the repository Gitlab: https://gitlab.com/macek_mm.

Coverage of the datasets

Pollen sites were equally distributed along the elevation gradient when elevation classes are defined by regular steps. When each elevation class covers an equal proportion of the area studied, mountains (680–1,600 m a.s.l.) and lowlands (48–230 m a.s.l.) are slightly oversampled at the expense of middle elevation classes (230–580 m a.s.l., Fig. 3A). The number of archaeological components generally decreases at high elevations within each period (Fig. 3B). More specifically, 350 m a.s.l. separates areas of high and low occupancy.

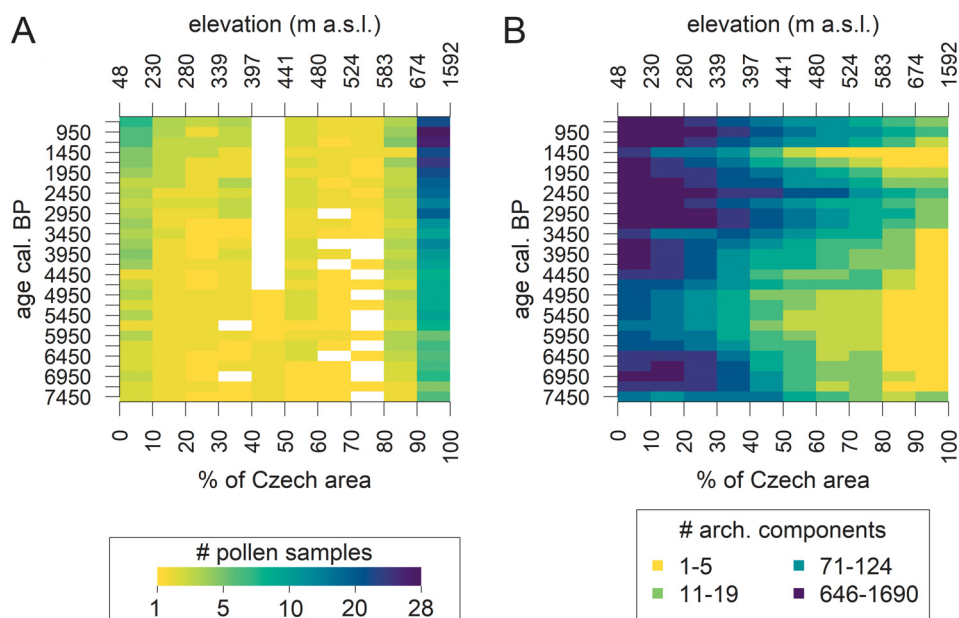


Fig. 3. Elevational and temporal coverage of pollen (A) and archaeological (B) data in this study. “#” means “number of”. Note: Each altitudinal class (columns) represents an equal proportion of the study area.

Results

The role of environmental variables

The relationship of the environmental variables with variability in the pollen record is shown in Fig. 4. Species at the extremes of the multidimensional space include many herbaceous plant taxa (*Senecio*-Typ, *Mentha*-Typ, *Caryophyllaceae*, *Cerealia*) and *Picea*, on the left and right side of the x-axis; *Abies* and *Ulmus* are at the bottom and top of the y-axis. Archaeological evidence decreases from left to right and elevation increases in a similar direction, so both variables are correlated with the main gradient in the data. Time increases from top to bottom and is the second main gradient, because the isoline 4,000 cal. BP is almost perpendicular to the isolines of archaeology (0.05) and elevation (400 m a.s.l.).

Most of the taxa have a negative relationship with elevation based on the Archaeology-Elevation-Year model (Supplementary Fig. S1), whereas the relationship of *Fagus* with elevation is positive and pollen proportion of *Plantago lanceolata*, *Abies* and *Carpinus* do not change substantially with elevation. The maximum pollen proportion for *Ranunculus acris*-Typ (hump-shaped relationship) and the minimum for *Pinus* and *Amaranthaceae* (U-shaped relationship) occur at middle altitudes. Long-term trend corresponds to the mean pollen proportions across all sites, thus the trend for taxa dominating during the early or middle Holocene (*Quercus*) decreases, whereas for the late Holocene it increases (*Abies*).

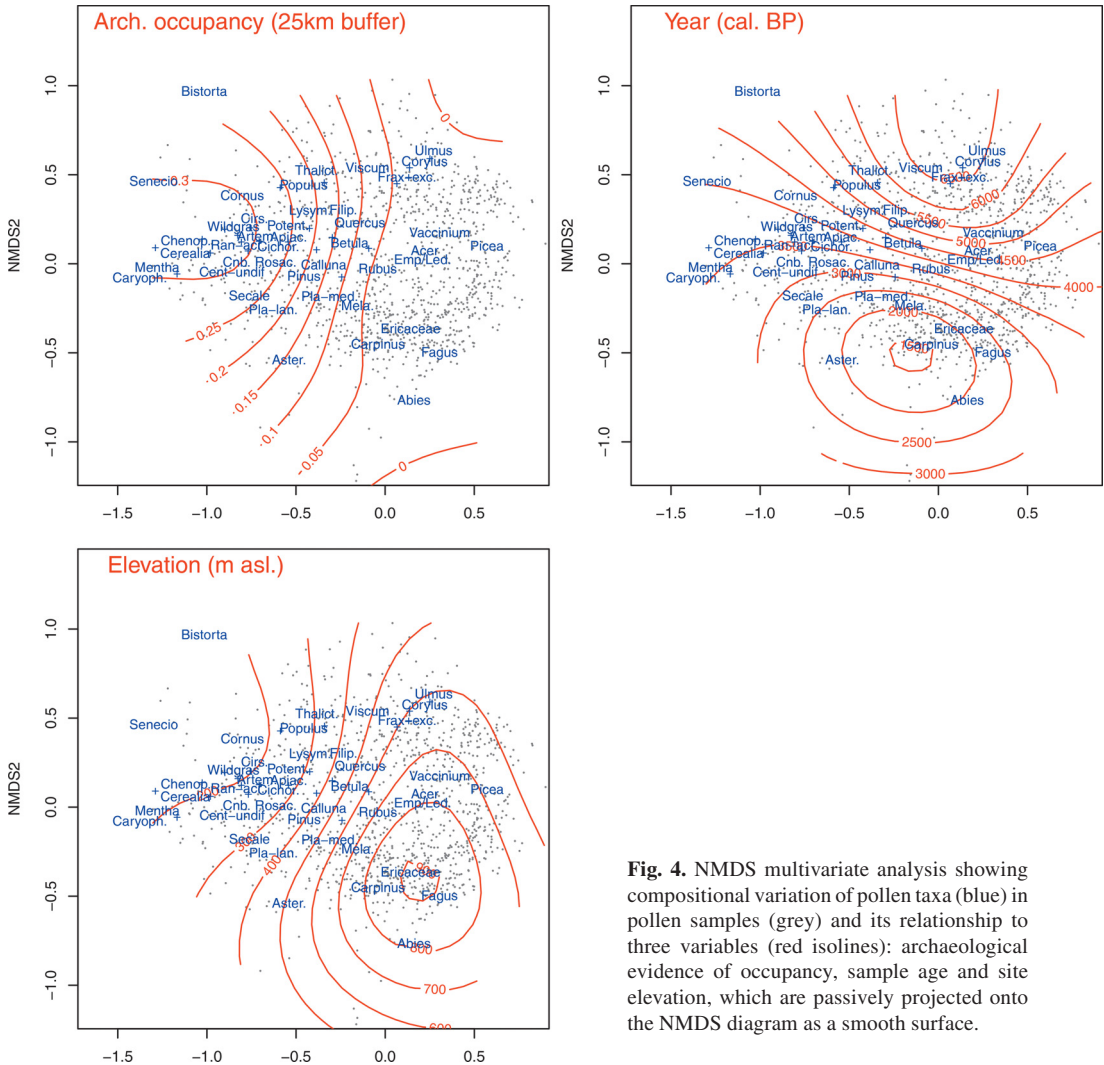


Fig. 4. NMDS multivariate analysis showing compositional variation of pollen taxa (blue) in pollen samples (grey) and its relationship to three variables (red isolines): archaeological evidence of occupancy, sample age and site elevation, which are passively projected onto the NMDS diagram as a smooth surface.

Relationship between pollen and archaeological data at individual sites

Twenty-nine taxa have a positive or negative relationship with archaeological occupancy in at least in one buffer zone, based on correlations for individual sites (Fig. 5A). Established anthropogenic indicators are all positively correlated with archaeological evidence of human activity. There are tree and herbaceous plant taxa that are both positively and negatively correlated. The strongest signal is for *Abies*, *Fagus* and *Carpinus*, whose significant positive relationship is recorded at more than 65% of the sites. *Artemisia*, *Calluna vulgaris*, *Plantago lanceolata*, *Sambucus nigra*-Typ, *Thalictrum* and *Juniperus* and the sum of anthropogenic indicators have a clear relationship with archaeological occupancy in more than four buffer zones. The rest of the 41 taxa do not have a clear relationship

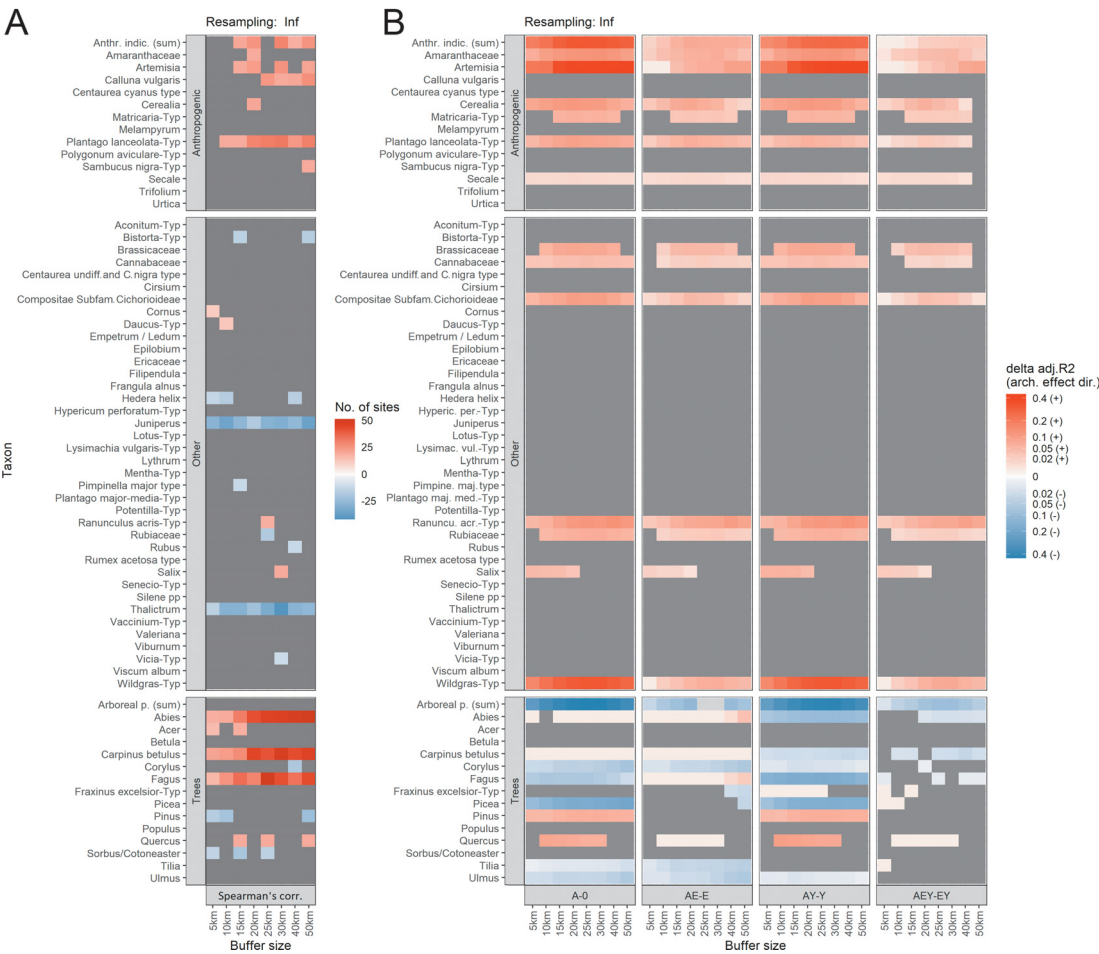


Fig. 5. Relationship of pollen taxa to archaeological evidence across spatial scales (buffer sizes). Number of sites with a significantly prevailing sign of Spearman's correlation (A) and delta adjusted R^2 (A-0, AE-E, AY-Y, and AEY-EY) multiplied by the sign of the regression coefficient from the model using archaeological occupancy, and individual taxa from the model using original pollen counts. The capital letters mean: E – altitude of sites and Y – the overall trend of each taxon. Scenario APNAP without resampling (see Methods). Positive correlations are red, negative blue and non-significant grey. Full names of taxa abbreviated in (B) are shown on the same line in (A).

with archaeologically detected human activity at any distance. This means that for the sites where a given taxon was present: approximately half of the sites were positive and half were negative.

Relationship between pollen and archaeological data in the whole datasets

In the first step, the relationship between the original pollen counts (APNAP scenario) and human occupancy based on the archaeological data in four principal models (Fig.

5B) based on the sign, significance and adjusted R^2 was considered. The basic model that includes only human occupancy (model Archaeology) indicates a significant relationship for 23 taxa, of which 14 herbaceous plants and shrub taxa also have remain relationships in the Archaeology-Elevation-Year model; however, adding environmental variables decreases the number of tree taxa or changes the sign of the relationship. The inclusion of an elevational gradient (model Archaeology-Elevation) removes the correlation for *Pinus* and *Picea* in most of the buffer zones, changes the correlation of *Fagus* to negative and weakens the relationships for *Artemisia*, wild grasses, *Quercus*, the sum of anthropogenic indicators and sum of trees. Incorporation of the long-term trend (model Archaeology-Year) removes the correlation for *Tilia* and changes the correlations for *Abies* and *Carpinus* to negative.

When elevation and long-term trend was added to the model already containing the evidence of past human activity (model Archaeology-Elevation-Year), archaeological evidence of human occupancy is a significant variable for four tree taxa: based on four consecutive buffer zones, *Quercus* is positive, while *Abies*, *Carpinus* and the overall arboreal pollen sum are negative.

The correlations of proportions from pollen counts adjusted by estimates of pollen productivity (APNAPppe scenario) also include similar taxa (*Amaranthaceae*, *Artemisia*, *Cereal*ia, *Plantago lanceolata*-Typ, *Secale*), although *Senecio*-Typ additionally has a positive correlation, while *Matricaria*-Typ does not (Supplementary Fig. S2). The correlations of the proportions of pollen of particular trees within the sum of arboreal pollen (AP scenario) show a similar pattern to that of the APNAP scenario. However, *Abies* and *Carpinus* are negative, while *Quercus*, *Pinus* and *Picea* are positive at a few distances (Supplementary Fig. S3). Resampling the APNAP scenario using lower pollen sums results in a lower number of taxa with a significant relationship with archaeological evidence of human occupancy (Supplementary Fig. S4).

Spatial scale and indicative strength

The numbers of taxa with significant correlations at all distances and groups of taxa (herbs APNAP, trees APNAP, trees AP), based on the Archaeology-Elevation-Year model were summarized (Fig. 6). Histograms for herbaceous plants and shrubs indicate that the maximum number of significant correlations with archaeological occupancy are recorded at a smaller radius (15–20 km) than for trees based on the AP scenario (30–40 km). Trees in the APNAP scenario do not show a clear pattern.

Twelve taxa with significant correlation to archaeological occupancy were recorded by both methods, Spearman and GAM (Table 1). The level of alpha in the binomial test and number of sites determined the probability with which their occurrence indicates an archaeological record, which is 0.62–0.85. This means that they are positively correlated at 46–62 sites and negatively correlated at 11–28 sites. In this case the buffer zones with no archaeological data at sites with pollen were not included and when included the probability decreases to 0.43–0.84. Archaeological records accounts for between 0.01 and 0.09 of variability in the results for 8 taxa. The pollen patterns of the other four taxa (*Abies*, *Carpinus*, *Fagus*, and *Quercus*) significantly correlate with archaeological occupancy-based on the Archaeology-Year model, however, the overall variability accounted for is lower than that based on the Elevation-Year model.

Table 1. Significant anthropogenic indicators based on Spearman’s correlation and all GAM models. Sign of correlation – list of models with a negative correlation, the rest of them are positive; buff. – number of buffer zones with significant correlation; prob. p – probability of the indicating presence of archaeology when pollen is present; prob. p/a – probability indicating presence/absence of archaeology when pollen is present/absent; ΔR^2 AEY-EY – variability which can be attributed to archaeology after subtracting the variability due to the temporal and elevational trends.

Taxon	Sign of correlation	buff.	prob. p	prob. p/a	ΔR^2 AEY-EY
<i>Abies</i>	AY & AEY negative	5	0.78–0.85	0.77–0.83	<0
<i>Amaranthaceae</i>	all positive	1	0.63	0.61	0.04
Sum of anthropogenic indicators	all positive	5	0.62–0.69	0.60–0.69	0.01–0.03
<i>Artemisia</i>	all positive	4	0.63–0.67	0.60–0.67	0.01–0.09
<i>Carpinus betulus</i>	AY & AEY negative	6	0.68–0.84	0.57–0.81	<0
<i>Cerealia</i>	all positive	1	0.64	0.60	0.06
<i>Fagus</i>	A, AY & AEY negative	4	0.67–0.84	0.43–0.84	<0
<i>Pinus</i> AP	negative	3	0.63–0.73	0.63–0.73	0.01–0.01
<i>Plantago lanceolata</i> -Typ	all positive	6	0.64–0.72	0.56–0.71	0.01–0.03
<i>Quercus</i>	all positive	2	0.63–0.64	0.59–0.63	0
<i>Ranunculus acris</i> -Typ	all positive	1	0.63	0.60	0.08
<i>Rubiaceae</i>	Spearman negative	1	0.63	0.64	0.02

Discussion

Relationship between pollen and archaeological data

An ideal anthropogenic indicator should be closely linked to a (the presence of) human population and scale quantitatively with human presence in the landscape, regardless of the environmental setting. Such species should not exhibit temporal or spatial trends unrelated to human presence in the landscape, i.e. caused by their ecological response to fluctuations in climate, post-glacial migration, affinity to a certain bedrock, etc. In reality, plant species occupy a particular environmental niche (Chase & Leibold 2009) and therefore their distribution and relative abundance in the pollen record is determined by these environmental drivers, which may result in spurious spatial and temporal correlations with human presence, due to collinearity issues. Unsuitable environmental conditions or lagged post-glacial migration may result in the absence of an indicator species in a landscape populated by humans, while some natural habitats that are frequently disturbed and have a high soil-nutrient status, such as animal resting places or floodplains (Lang 1994), may be colonized by species indicating human presence. In this study, elevation was used as a surrogate for the most prominent environmental gradient. The temporal dynamics of each species, which synchronously affects the local species abundance in the area studied, was also accounted for. Either way, large fluctuations in the abundance of species may also be associated with human activity for many reasons, including human-aided introduction of species outside of their natural distribution range, or more recently due to human-induced climate change, however, these effects, which affect an area greater than that studied, cannot be statistically separated in a regression analysis of their natural dynamics.

The proportions of the pollen of the most widely-used pollen anthropogenic indicators are positively correlated with local archaeological records. Their indicative strength increases and more revelatory taxa are recorded when spatially-correlated environmental factors are considered, for example, the local bedrock or elevation. This indicates that

pollen analysis can detect a variety of anthropogenic signals in the vegetation, however, specific sets of taxa should be used for different environments. For example, since geological data was not included as an explanatory variable, *Melampyrum*, which is considered to be a good anthropogenic pollen indicator in sandstone landscapes with poor soils (Svoboda et al. 2018), is not associated with archaeological occupancy. However, correlations between pollen and archaeological records may be biased by collinearity and temporal autocorrelations between archaeological evidence and external drivers, such as spatial and temporal variability in climate, plant ecological preferences, or migration history, which may coincide with human population dynamics, but are not causally linked. The full model (Archaeology-Elevation-Year), accounts for these potentially confounding effects, revealed a significant positive association with archaeological records for fewer indicator species than the basic model (Archaeology). The possible collinearity of influencing factors and existing biases therefore question the general validity of anthropogenic indicators in pollen analysis.

Primary pollen indicators

Both primary and secondary indicators are significantly correlated in all models that analysed the whole dataset, including all samples independent of the pattern recorded at individual sites, however, site-by-site Spearman's correlation with archaeology results in weak mutual relationship. The ratio of the number of sites with positive and negative correlations is significant for *Cerealia* only at one distance, 25 km, and marginally significant for *Secale*. This low indicative ability might seem surprising for both primary pollen indicators, whose presence in a landscape is undoubtedly linked to the planting of crops by humans. This might be due to several factors. First, the late introduction of *Secale* into the study area during the Early Bronze Age and its massive use in the Early Medieval Period are unfavourable prerequisites for a universal pollen indicator (Kočár & Dreslerová 2010, Kozáková et al. 2015). Second, the determination of cereal pollen and its separation from wild grasses is difficult. Differences between *Cerealia* and wild grass pollen (size, pore and annulus) do not indicate discrete groups (Beug 2004). Moreover, glycerine, the most common mounting medium causes the palynomorphs to inflate and changes in grain size depending on how long they been mounted and the dilution. The micro-ornamentation on cereal pollen is a very distinctive feature, but it requires 1000× magnification and phase contrast (Beug 2004), which are not used in routine counting. Moreover, having a dataset compiled by many authors further compounds this human factor (Mander et al. 2014). Similarly, the pollen syntheses from Fennoscandia (Josefsson et al. 2014) raise methodological problems that might have resulted in the inconsistent timing of the appearance of *Cerealia* in pollen data used in the current study. In several sequences signals of *Cerealia* are recorded occurring in the Early Holocene (Supplementary Fig. S5), misdetermined wild grasses, hamper the use of *Cerealia* pollen as a quantitative indicator of human presence in the landscape.

Third, prehistoric farming might not have resulted in large quantities of cereal pollen being deposited in wetland sites. This is most likely the case for Neolithic cultivation that occurred in landscapes with more woodland and at a scale similar to gardening (Bogaard 2004). Therefore, the effects of Neolithic and Eneolithic farming recorded in the pollen record might have been masked by these factors.

Secondary pollen indicators

The best proxy to validate the contact of past human populations with plant species are plant remains found in archaeological contexts. All the positively correlated herbaceous pollen taxa (*Amaranthaceae*, *Artemisia*, *Plantago lanceolata*, *Ranunculus acris*-Typ, and *Rubiaceae*) have corresponding species in the Czech archaeobotanical record appearing since the Neolithic: *Artemisia vulgaris*, *Atriplex oblongifolia*, *Galium aparine*, *G. mollugo*, *G. spurium*, *G. verum* agg., *Chenopodium album*, *Ch. hybridum*, *Ch. murale*, *Ch. polyspermum*, *Ch. urbicum*, *Ranunculus repens*, and – since the Eneolithic: *Plantago lanceolata* (Pokorná et al. 2018).

Among the herbaceous plant taxa, *Plantago lanceolata* occurred in the maximum number of buffer zones and has the maximum probability of being an indicator species and the high variability can be attributed to archaeology (Table 1). This corresponds with the results of previous studies in Ireland (Whitehouse et al. 2014) and Norway (Hjelle et al. 2006), which are based on regional syntheses of the effect of human activity on this taxon and its good indicative ability. Similarly strong results were obtained using the sum of anthropogenic indicators, commonly used in pollen diagrams. This measure can overcome problems with regionally different environments in larger syntheses of the effect of human activity and different taxa can substitute for each other. However, one must select taxa carefully and keep in mind that in some regions anthropogenic indicators can also occur in wetlands.

Amaranthaceae includes plants that are edible as a staple food and also as a vegetable. We can assume their intentional cultivation (Mueller-Bieniek et al. 2019). However, archaeobotanical findings do not fully support this (Kočár & Dreslerová 2010). Carbonized seeds of *Amaranthaceae* have only been found at Mesolithic sites (Svoboda 2003) and even then at a similar frequency as weeds or ruderals. Nonetheless, the strong relationship of *Amaranthaceae* pollen with archaeological evidence of human occupancy can also be explained by their association with disturbed terrestrial habitats, together with *Artemisia*. These pollen taxa are generally attributed to the presence of ruderal and trampled communities or arable fields (Behre 1981). Another habitat rarely considered as the main source of *Amaranthaceae* and *Artemisia* pollen is wetland vegetation consisting of nitrophilous annuals, such as *Bidentetea tripartitae*. This class of wetland vegetation occurs in landscapes affected by anthropogenic eutrophication due to periodic inundation, accompanied by the accumulation of salt, nutrients and nitrogen (Danihelka et al. 2022). The scattered distribution of *Bidentetea tripartitae* includes site Olbramovice in southern Moravia (Svobodová 1997), which was included in the current study. *Amaranthaceae* from *B. tripartitae* leave a local pollen signal, *Amaranthaceae* from dry ruderal habitats leave a regional pollen signal, and both habitats can originate from human activity.

Similarly, the two other taxonomically wide indicators, *Ranunculus acris*-Typ and *Rubiaceae*, which can be represented by species that occur in wetlands (e.g. *R. sceleratus* and *Galium uliginosum*) or in ruderal habitats (e.g. *R. repens* and *G. aparine*). However, both wetland species are common in the local vegetation at most of the pollen sites studied, which indicates natural rather than an anthropogenic origin of *Ranunculus acris*-Typ and *Rubiaceae* pollen at the sites studied. Moreover, archaeological evidence seems to be important for both taxa in only one buffer zone based on the pollen archive, therefore, both are considered to be weak indicators.

Trees

Tree pollen relationship with archaeological data changed most dramatically when environmental variables are considered. The strong links between arboreal pollen and archaeological evidence indicated by the basic comparison (model Archaeology) are weaker when the complete model (Archaeology-Elevation-Year) is considered. In spite of the significance of the Archaeology-Elevation-Year model, the variability accounted for archaeology is very low. Still, there is an association between the pattern in tree pollen and human behaviour. The inclusion of the temporal collinearity of the factors studied (models Archaeology-Year and Archaeology-Elevation-Year) changed the correlation between the pollen of trees and archaeological occupancy from positive to negative for *Abies* and *Carpinus*. The late Holocene spread of these taxa is associated with an increase in human activity (Giesecke et al. 2017). Their relationships, however, is negative, because they are late-successional species, so there is likely to have been a considerable delay between the abandonment of anthropogenically-disturbed areas and their colonization by late succession species. For example, succession on the prehistoric hillfort Vladař in western Bohemia lasted 350 years, during which the incidence of *Abies* pollen increased (Pokorný et al. 2006). Similarly, there is a gap of approximately 500 years between the maximum peaks of non-arboreal pollen and that of *Carpinus* in northern Poland (Ralska-Jasiewiczowa et al. 2003). The time window of 250 years used in the current study is shorter than the duration of succession recorded in these palynological studies, thus the negative correlation with archaeological evidence of occupancy recorded in the current study; however, there is a clear link between some tree taxa and human activity, an anthropogenic indication of late post-disturbance stage.

Fagus belongs to the same ecological group of Late Holocene taxa that spread after disturbances during late succession and thus also have a negative correlation with archaeology due the temporal collinearity (Archaeology-Year and Archaeology-Elevation-Year models). Unlike *Carpinus* and *Abies*, *Fagus* is negatively correlated in the basic model (Archaeology), when archaeology is included, because it grows mainly in the mountains, where the archaeological evidence for settlements is generally low (Fig. 4C). Spearman's correlation and the Archaeology-Elevation model predictions are not dependent on altitude; thus, the result is positive due to the temporal collinearity, as in *Carpinus* and *Abies*.

The role of humans in the spread of late Holocene trees is underlined by findings of fragments of charcoal of these taxa in archaeo-anthracological assemblages from the Neolithic in Moravia (Novák et al. 2020). The long coexistence of *Abies*, *Carpinus* and *Fagus* with human populations was, after approximately 2,500–3,000 years, followed by the rapid spread of these trees and an increase in their dominance in the rest of the landscape. Compared to the currently invasive trees (e.g. *Ailanthus altissima*, *Robinia pseudoacacia*, *Quercus rubra*), *Abies*, *Carpinus* and *Fagus* spread 10 times more slowly in the landscape (Křivánek et al. 2006). This is most likely due to a very different scale and character of current human activity than that in prehistory. Moreover, the change in the rate of spread of trees in the late Holocene might also be due to other factors (soil degradation, climate change; Pokorný & Kuneš 2005).

Quercus and *Pinus* are taxa that are positively correlated with archaeological evidence of human occupancy. *Pinus* is an early successional tree colonizing areas immediately after disturbance. Moreover, the fire-prone character of *Pinus* stands increases the incidence

of this type of disturbance (Adámek et al. 2018). *Quercus* is closely associated with human settlements, since it dominates most of the archaeological charcoal assemblages (Novák et al. 2020). This positive relationship is possible, firstly, due to its tolerance of fire, use of acorns as a source of food and ability to survive woodland grazing and periodic woodland coppicing (Szabó 2010). Secondly, the strong association with human occupancy and the abundance of *Quercus* in prehistory is due to their spatial overlap, which is documented in several case studies (Kuneš et al. 2015, Jamrichová et al. 2017). There was an abundance of prehistoric agro-pastoral communities in the central-European lowlands along with a high natural occurrence of *Quercus*. The spatial collinearity between the dominance of *Quercus* and archaeological evidence in lowlands is indicated by the weak delta R^2 between AE-E and AEY-EY models (Fig. 5), although still the relationship is significant.

Spatial factor in the relationship between pollen evidence and archaeological settlement

The radius that maximizes the best fit between pollen and archaeological data is greater for trees than for herbaceous plants (Fig. 6). This result is in accordance with the prediction of the model of dispersal of pollen, which is that pollen released at great heights will disperse over greater distances (Kuparinen et al. 2007). However, the distances of 15–20 and 30–40 km, resulted from the AEY model, greatly exceed the empirical estimates of the source area for herbaceous plants, tens of metres (Shaw & Whyte 2020) and for tree pollen, 4 km (Matthias & Giesecke 2014). Looking at the mixed assemblages of herbaceous plants and trees recorded in contemporary pollen-vegetation studies in Czechia, pollen analysis reveals the vegetation pattern within 1.5 km (Abraham & Kozáková 2012) and 0.8 km, respectively (Kuneš et al. 2019b).

Similarly, the vegetation recorded in the above pollen-vegetation studies, both distances of best fit (15–20 and 30–40 km) are dependent on the density and spatial precision of the archaeological data, i.e. the mean area of a parish. Owing to the sparsity of archaeological evidence, that for the smallest radius of 5 km, which only consists of a few parishes, may be less reliable than the archaeological records for a greater radius. Moreover, effect of human activity is not restricted to areas close to settlements, from where most of the archaeological evidence is obtained, because a wider range of landscapes would have been used and managed. Human activities that have a major effect on vegetation (for example, agriculture, grazing and logging) would have affected the surroundings of the settlements at a diverse range of spatial scales. Fields and gardens are most likely to occur in closest vicinity, whereas pastures for domestic herbivores could have been located at more distant places. Managed woodlands and logging areas were probably the most distant. The most striking example is perhaps Iron Age fortified sites (e.g. hillforts or, latter, oppida). Logging for building their complex fortifications not only occurred close by, but probably also over a much wider area. Several thousand standardized logs were needed in just one construction phase and these logs originated from tens or hundreds of hectares of forests (e.g. oppidum Závist, 37.5 ha, oppidum Manching, 370 ha; Venclová et al. 2008). In addition, these constructions would have been regularly repaired, resulting in a much wider deforestation or even the transport of much-needed logs from more distant places (Opravil 1989). Moreover, as these calculations do not include woodland areas being used for other purposes (for example, charcoal production

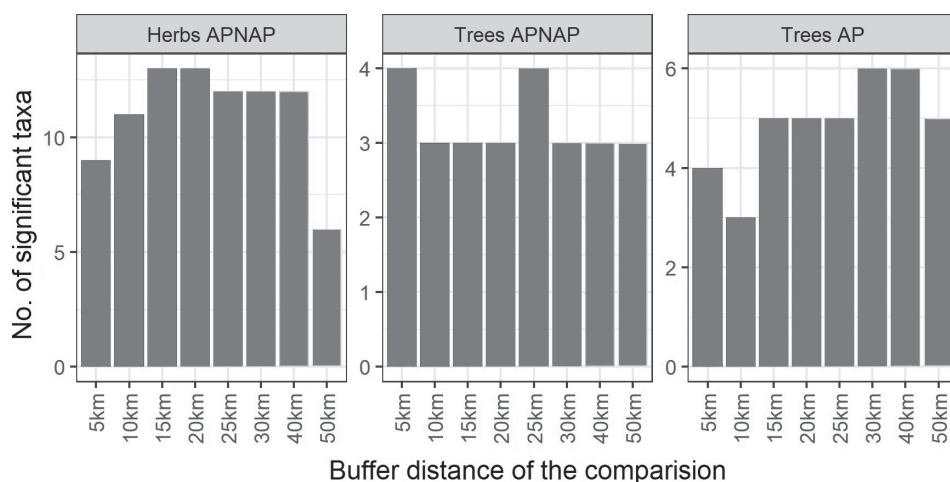


Fig. 6. Number of taxa with a significant correlation between pollen proportion and archaeological evidence in the AEY model.

for metallurgy, wood for pottery production, leaf fodder, wood for household cooking, wood for constructing houses), it is likely that large areas of wooded landscape, probably extending over tens of kilometres from settlements, were affected by human activities.

In addition, prehistoric agricultural settlements were most likely much more mobile than conventionally thought (cf. Graeber & Wengrow 2021). Archaeological research in areas with better preservation of organic materials and a chronology of settlement based on dendrochronology (e.g. Hofmann et al. 2016) indicate that communities and their economic areas could have been relocated elsewhere after just 10 or 20 years. Thus, even larger areas of vegetation would be affected by ongoing or previous activities.

Human activity also directly affected wetlands (again possibly far from settlements), by directly affecting a potential pollen site (for example, the use of wet meadows for pastures documented by the presence of spores of coprophilous fungi and the presence of ritual places that are regularly visited for depositing offerings so the wooden access paths had to be maintained). Changes in local vegetation can be caused by eutrophication due to the erosion of the watershed or a change in the sedimentation from organic mire to clays, or the creation of pools or ponds. A different interpretative framework must be used when humans also interact with the taphonomy of pollen and when dealing with the archaeopalynological record, for example, when the core originates from a cistern in the centre of a hillfort (Pokorný et al. 2006) or when the site was used for hemp retting, as indicated by the presence of *Cannabis* pollen and seeds (Jamrichová et al. 2018). However, identifying such direct human influence on the sequence can be difficult in the absence of additional information from other proxies, so all kinds of sites were included in the dataset in order to obtain more robust results.

The scale and character of the effect of human activity on vegetation depend upon population size, social organization and advances in technology. High occupancy during the Neolithic period resulted in fewer pollen signals of anthropogenic indicators than the same level of occupancy in the Iron Age. This is due to the spatial scale of Neolithic farming,

which was similar to gardening, and that of the large scale Iron Age agriculture, with more domesticated plants and the use of iron ploughs. In addition, the mobility of farming communities could have significantly affected vegetation dynamics, colonizing abandoned and reclaimed anthropogenic habitats. Changes in technology, especially such affecting landscape management or energy consumption (based on wood and fire), changed land use strategies and enabled the creation of new habitats by transforming existing ones. Species pools for a specific period determined the resilience of the habitat to changes in land use and influenced the ecology of the vegetation.

Constraints on studying past human-vegetation interactions

Studying past human-vegetation relationships is difficult due to the many factors that affect the indicators in the fossil record (Fig. 1). Regarding the factors in the fossil record that indicate the effect on vegetation are land use practices that affect the flowering of plants. Early and frequent mowing or felling of trees results in low pollen signals (Theuerkauf et al. 2015). Secondly, plants must produce sufficient pollen or grow close to the deposition site, for them to be in the fossil record. On the other hand, high pollen producers in the Czech Republic do not affect fossil records of rare types of pollen types, as the patterns based on unadjusted and adjusted datasets are similar (Supplementary Fig. S2, Fig. 5).

Thirdly, low pollen counts also considerably decrease the probability of the inclusion of rare pollen types and result in discrete distributions of pollen proportions with many pseudoabsences. Changes in the presence of a single pollen taxon in a sample is dependent upon the total pollen count (Odgaard 2007). Results based on complete samples were compared with results based on subsampled samples consisting of the same number of pollen grains. Low pollen sums (below 100 grains) resulted in a dramatic decline in the number of pollen taxa present in the dataset and number of taxa indicating a significant relation with archaeological evidence of human occupancy (Supplementary Fig. S4). Therefore, the analysis of pollen sequences with low pollen sums is not suitable for studying human-vegetation interactions. Trees, for which there are usually large pollen counts, are less affected by the selection of the target pollen sum compared to herbaceous plants (Supplementary Fig. S4).

Fourthly, the standardization of different pollen sums when using large databases, either by rarefaction (Birks & Line 1992) or resampling as in the current study inevitably results in loss of information, both by not omitting a sufficient number of samples or randomly deleting pollen that is above the threshold value for the total sample. To avoid this trade-off between comparability of pollen proportions and sample size, two different approaches were adopted when considering taxa at individual sites and within the database as a whole. By using a beta error family with information on the pollen count for the modelled taxa and the total sample, instead of just proportions, it was possible to correctly fit models of the log-likelihood function for small samples. Finally, a great advantage of the GAM models is the inclusion of the random effect of those sites with different pollen proportions, which enabled the joint analysis undertaken in the current study.

Conclusions

Environmental factors confound the relationship between fossil pollen assemblages and the intensity of past human settlements recorded in the archaeological database. The indicative strength of anthropogenic signals in the form of fossil pollen decreases more along the elevational gradient than in the long-term trend, because the first main gradient in the dataset is either archaeology or elevation. The temporal changes in the conditions, such as climate or soil development, affect the migration history of plants. The temporal dynamics of each taxon synchronously affects local abundance in the area studied and cannot be statistically separated from human-induced changes. Taking this into consideration changes the correlation of *Abies* and *Carpinus* to archaeology from positive to negative.

The strongest anthropogenic indicator, *Plantago lanceolata* grows within various types of open habitats other than wetlands. *Amaranthaceae* and *Artemisia* are ruderal herbaceous plants indicating disturbance and a high level of nitrogen in the soil and more rarely anthropogenic eutrophication of wetlands by periodic inundation and the accumulation of salt, nutrients and nitrogen. The tree taxa were placed in one of two groups based on the human activity indication. The first group has a positive relationship with archaeology due to their ability to resprout, produce acorns and tolerance of fire (*Quercus*) or pioneer trees that rapidly spread within 250 years after human disturbance (*Pinus*). The second group of trees has a negative relationship with archaeology, because they spread during late-successional stages after human disturbance (*Abies*, *Carpinus*, and *Fagus*). The taxa with positive relationships have high probability, 0.56–0.76, of indicating the presence or absence of an archaeological record when pollen is present or absent, but the variability accounted for by the full model is low, 0.01–0.09. *Secale* does not correlate at level of individual sites due to its late appearance during the Holocene. The weak relationship between the archaeological data and the pollen of *Cerealia* is attributed to the inconsistency of pollen determinations. The radius for comparability of pollen and archaeological data is around tens of kilometres due to the spatial resolution of the archaeology for areas of the size of a parish and is lower for herbaceous plants (15–20 km) than for trees (30–40 km).

Pollen taxa were identified by their fluctuations in abundance in the pollen record that indicate changes in the size of local human populations at explicit temporal and spatial scales. They are indicators of a high human presence in the landscape, although the relation may be indirect and unintentional from the point of view of humans. The quality of archaeological evidence and the number of available pollen profiles in the area studied are among the best datasets available worldwide. Given the spatial extent of this study (approximately 80,000 km²), the list of anthropogenic indicators is likely to be the same throughout entire central Europe. It must be admitted, however, that the list of indicators is quite short, but fully data-driven and statistically strong.

Supplementary material

Table S1. List of the pollen sites included in the analysis.

Table S2. Estimates of mean pollen productivity (Wieczorek & Herzschuh 2020) used to adjust the pollen counts.

Fig. S1. Default output from the GAM AEY APNAP model showing relationships of pollen to the three environmental variables: elevation, temporal trend and archaeology.

Fig. S2. Delta adjusted R^2 multiplied by the sign of the regression coefficient from the model using archaeological occupancy and individual taxa from the model using original pollen counts. Scenario APNAPpe without resampling.

Fig. S3. Delta adjusted R^2 multiplied by the sign of the regression coefficient from the model using archaeological occupancy and individual taxa from the model using original pollen counts. Scenario AP without resampling.

Fig. S4. Number of significant taxa for different sums within the three scenarios of pollen counts.

Fig. S5. Comparison of *Cerealia* and *Secale* pollen with archaeological occupancy within 25 km radius at selected sites.

Supplementary materials are available at www.preslia.cz

Acknowledgements

We would like to thank all contributors who provided their records to the archaeological database and Czech Quaternary Palynological Database. Without their long-term hard work, this research would not be possible. We are grateful to Péter Szabó, Steve Ridgill and James Preston for proof-reading. We thank two anonymous reviewers for their fruitful comments. This work was supported as a long-term research development project of the Czech Academy of Sciences (RVO 67985939) and is part of the project “Land use, social transformations and woodland in central-European Prehistory: Modelling Approaches to Human-Environment Interactions”, supported by the Czech Science Foundation (19-20970Y). The participation of D. Novák was supported by the research infrastructure project “Archaeological Information System of the Czech Republic – second generation” (EF16_013/0001439) financed by MEYS and the European Union. P. Pokorný was supported by OP RDE, MEYS, under the project “Ultra-trace isotope research in social and environmental studies using accelerator mass spectrometry”, reg. no. CZ.02.1.01/0.0/0.0/16_019/0000728.

References

- Abraham V. & Kozáková R. (2012) Relative pollen productivity estimates in the modern agricultural landscape of Central Bohemia (Czech Republic). – Review of Palaeobotany and Palynology 179: 1–12.
- Abraham V., Kuneš P., Petr L., Svitavská-Svobodová H., Kozáková R., Jamrichová E., Švarcová M. G. & Pokorný P. (2016) A pollen-based quantitative reconstruction of the Holocene vegetation updates a perspective on the natural vegetation in the Czech Republic and Slovakia. – Preslia 88: 409–434.
- Adámek M., Jankovská Z., Hadincová V., Kula E. & Wild J. (2018) Drivers of forest fire occurrence in the cultural landscape of Central Europe. – Landscape Ecology 33: 2031–2045.
- Beckage B., Lacasse K., Winter J. M., Gross L. J., Fefferman N., Hoffman F. M., Metcalf S. S., Franck T., Carr E., Zia A. & Kinzig A. (2020) The Earth has humans, so why don't our climate models? – Climatic Change 163: 181–188.
- Behre K.-E. (1981) The interpretation of anthropogenic indicators in pollen diagrams. – Pollen et Spores 23: 225–245.
- Beug H.-J. (2004) Leitfaden der Pollenbestimmung in Mitteleuropa und angrenzenden Gebiete. – Verlag Dr. Friedrich Pfeil, München.
- Birks H. J. B. & Line J. M. (1992) The use of rarefaction analysis for estimating palynological richness from Quaternary pollen-analytical data. – The Holocene 2: 1–10.
- Bogaard A. (2004) Neolithic farming in Central Europe: an archaeobotanical study of crop husbandry practices. – Taylor & Francis, Abingdon.
- Boivin N. & Crowther A. (2021) Mobilizing the past to shape a better Anthropocene. – Nature Ecology and Evolution 5: 273–284.
- Brun C. (2011) Anthropogenic indicators in pollen diagrams in eastern France: a critical review. – Vegetation History and Archaeobotany 20: 135–142.
- Chase J. M. & Leibold M. A. (2009) Ecological niches: linking classical and contemporary approaches. – University of Chicago Press, Chicago.
- Chytrý M. (2012) Vegetation of the Czech Republic: diversity, ecology, history and dynamics. – Preslia 84: 427–504.
- Colombaroli D., Beckmann M., van der Knaap W. O., Curdy P. & Tinner W. (2013) Changes in biodiversity and vegetation composition in the central Swiss Alps during the transition from pristine forest to first farming. – Diversity and Distributions 19: 157–170.

- Contreras D. A. & Meadows J. (2014) Summed radiocarbon calibrations as a population proxy: a critical evaluation using a realistic simulation approach. – *Journal of Archaeological Science* 52: 591–608.
- Court-Picon M., Buttler A. & de Beaulieu J.-L. (2005) Modern pollen–vegetation relationships in the Champsaur valley (French Alps) and their potential in the interpretation of fossil pollen records of past cultural landscapes. – *Review of Palaeobotany and Palynology* 135: 13–39.
- Crema E. R. & Bevan A. (2021) Inference from large sets of radiocarbon dates: software and methods. – *Radiocarbon* 63: 23–39.
- Crombé P. & Robinson E. (2014) 14C dates as demographic proxies in Neolithisation models of northwestern Europe: a critical assessment using Belgium and northeast France as a case-study. – *Journal of Archaeological Science* 52: 558–566.
- Danihelka J., Chytrý K., Harásek M., Hubatka P., Klinkovská K., Kratoš F., Kučerová A., Slachová K., Szokala D., Prokešová H., Šmerdová E., Večeřa M. & Chytrý M. (2022) Halophytic flora and vegetation in southern Moravia and northern Lower Austria: past and present. – *Preslia* 94: 13–110.
- 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.
- Deza-Araujo M., Morales-Molino C., Tinner W., Henne P. D., Heitz C., Pezzatti G. B., Hafner A. & Conedera M. (2020) A critical assessment of human-impact indices based on anthropogenic pollen indicators. – *Quaternary Science Reviews* 236: 106291.
- Dreslerová D., Waldhauser J., Abraham V., Kočár P., Křivánek R., Meduna P. & Sádlo J. (2013) Bezděžsko – Dokesko v pravěku a laténské sídliště v Oknech [The Bezděz – Doksy region (Northern Bohemia) in pre-history and the La Tène settlement at Okna]. – *Archeologické rozhledy* 65: 535–573.
- Edinborough K., Porčić M., Martindale A., Brown T. J., Supernant K. & Ames K. M. (2017) Radiocarbon test for demographic events in written and oral history. – *Proceedings of the National Academy of Sciences of United States of America* 114: 12436–12441.
- Feeser I. & Furholt M. (2014) Ritual and economic activity during the Neolithic in Schleswig-Holstein, northern Germany: an approach to combine archaeological and palynological evidence. – *Journal of Archaeological Science* 51: 126–134.
- Firbas F., Losert H. & Broihan Fr. (1939) Untersuchungen zur jüngeren Vegetationsgeschichte im Oberharz. – *Planta* 30: 422–456.
- Gaillard M.-J. (2007) Pollen methods and studies. Archaeological applications. – In: Scott A. E. (ed.), *Encyclopedia of Quaternary Science*, p. 2570–2595. Elsevier, Oxford.
- Gaillard M.-J., Birks H. J. B., Emanuelsson U. & Berglund B. E. (1992) Modern pollen/land-use relationships as an aid in the reconstruction of past land-uses and cultural landscapes: an example from south Sweden. – *Vegetation History and Archaeobotany* 1: 3–17.
- Giesecke T., Brewer S., Finsinger W., Leydet M. & Bradshaw R. H. W. (2017) Patterns and dynamics of European vegetation change over the last 15,000 years. – *Journal of Biogeography* 44: 1441–1456.
- Gillson L., Ekblom A., Willis K. J. & Froyd C. (2008) Holocene palaeo-invasions: the link between pattern, process and scale in invasion ecology? – *Landscape Ecology* 23: 757–769.
- Graeber D. & Wengrow D. (2021) *The dawn of everything. A new history of humanity*. – Straus and Giroux, New York.
- Hájek M., Dudová L., Hájková P., Roleček J., Moutelíková J., Jamrichová E. & Horsák M. (2016) Contrasting Holocene environmental histories may explain patterns of species richness and rarity in a Central European landscape. – *Quaternary Science Reviews* 133: 48–61.
- Hellman S., Bunting M. J. & Gaillard M.-J. (2009) Relevant source area of pollen in patchy cultural landscapes and signals of anthropogenic landscape disturbance in the pollen record: a simulation approach. – *Review of Palaeobotany and Palynology* 153: 245–258.
- Hinz M., Feeser I., Sjögren K.-G. & Müller J. (2012) Demography and the intensity of cultural activities: an evaluation of Funnel Beaker Societies (4200–2800 cal BC). – *Journal of Archaeological Science* 39: 3331–3340.
- Hjelle K. L. (1999) Modern pollen assemblages from mown and grazed vegetation types in western Norway. – *Review of Palaeobotany and Palynology* 107: 55–81.
- Hjelle K. L., Hufthammer A. K. & Bergsvik K. A. (2006) Hesitant hunters: a review of the introduction of agriculture in western Norway. – *Environmental Archaeology* 11: 147–170.
- Hofmann D., Ebersbach R., Doppler T. & Whittle A. (2016) The life and times of the house: multi-scalar perspectives on settlement from the Neolithic of the northern Alpine foreland. – *European Journal of Archaeology* 19: 596–630.

- Jamrichová E., Gálová A., Gašpar A., Horsák M., Frodlová J., Hájek M., Hajnalová M. & Hájková P. (2018) Holocene development of two calcareous spring fens at the Carpathian-Pannonian interface controlled by climate and human impact. – *Folia Geobotanica & Phytotaxonomica* 53: 243–263.
- Jamrichová E., Hédl R., Kolář J., Tóth P., Bobek P., Hajnalová M., Procházka J., Kadlec J. & Szabó P. (2017) Human impact on open temperate woodlands during the middle Holocene in Central Europe. – Review of Palaeobotany and Palynology 245: 55–68.
- Jiráň L. & Venclová N. (2008) Archeologie pravěkých Čech [Archaeology of prehistoric Bohemia] 1–8. – Archeologický ústav AV ČR, Praha.
- Josefsson T., Ramqvist P. H. & Hörnberg G. (2014) The history of early cereal cultivation in northernmost Fennoscandia as indicated by palynological research. – *Vegetation History and Archaeobotany* 23: 821–840.
- Kočár P. & Dreslerová D. (2010) Archeobotanické nálezy pěstovaných rostlin v pravěku České republiky [Archaeobotanical finds of cultivated plants in the prehistory of the Czech Republic]. – *Památky archeologické* 51: 203–242.
- Kolář J., Kuneš P., Szabó P., Hajnalová M., Svobodová H. S., Macek M. & Tkáč P. (2018) Population and forest dynamics during the Central European Eneolithic (4500–2000 BC). – *Archaeological and Anthropological Sciences* 10: 1153–1164.
- Kolář J., Macek M., Tkáč P., Novák D. & Abraham V. (2022) Long-term demographic trends and spatio-temporal distribution of past human activity in Central Europe: comparison of archaeological and palaeoecological proxies. – *Quaternary Science Reviews* 297: 107834.
- Kolář J., Macek M., Tkáč P. & Szabó P. (2016a) Spatio-temporal modelling as a way to reconstruct patterns of past human activities. – *Archaeometry* 58: 513–528.
- Kolář J., Tkáč P., Macek M. & Szabó P. (2016b) Archaeology and historical ecology: the archaeological database of the LONGWOOD ERC project. – *Archäologisches Korrespondenzblatt* 46: 539–554.
- Kozáková R., Pokorný P., Peša V., Danielisová A., Čuláková K. & Svobodová H. S. (2015) Prehistoric human impact in the mountains of Bohemia. Do pollen and archaeological data support the traditional scenario of a prehistoric “wilderness”? – Review of Palaeobotany and Palynology 220: 29–43.
- Kozáková R., Šamonil P., Kuneš P., Novák J., Kočár P. & Kočárová R. (2011) Contrasting local and regional Holocene histories of *Abies alba* in the Czech Republic in relation to human impact: evidence from forestry, pollen and anthracological data. – *The Holocene* 21: 431–444.
- Křivánek M., Pyšek P. & Jarošík V. (2006) Planting history and propagule pressure as predictors of invasion by woody species in a temperate region. – *Conservation Biology* 20: 1487–1498.
- Kuna M., Hasil J., Křiváňková D., Novák D. & Lečbychová O. (2015) Archaeological Map of the Czech Republic. – In: Kuna M. (ed.), Structuring archaeological evidence. The Archaeological Map of the Czech Republic and related information systems, p. 25–68, Institute of Archaeology CAS, Prague.
- Kuneš P. & Abraham V. (2017) History of Czech vegetation since the Late Pleistocene. – In: Chytrý M., Danihelka J., Kaplan Z. & Pyšek P. (eds), *Flora and vegetation of the Czech Republic*, p. 193–227, Springer International Publishing, Cham.
- Kuneš P., Abraham V. & Herben T. (2019a) Changing disturbance-diversity relationships in temperate ecosystems over the past 12000 years. – *Journal of Ecology* 107: 1678–1688.
- Kuneš P., Abraham V., Kovářik O. & Kopecký M. (2009) Czech Quaternary Palynological Database (PALYCZ): review and basic statistics of the data. – *Preslia* 81: 209–238.
- Kuneš P., Abraham V., Werchan B., Plesková Z., Fajmon K., Jamrichová E. & Roleček J. (2019b) Relative pollen productivity estimates for vegetation reconstruction in central-eastern Europe inferred at local and regional scales. – *The Holocene* 29: 1708–1719.
- Kuneš P., Svobodová-Svitavská H., Kolář J., Hajnalová M., Abraham V., Macek M., Tkáč P. & Szabó P. (2015) The origin of grasslands in the temperate forest zone of east-central Europe: long-term legacy of climate and human impact. – *Quaternary Science Reviews* 116: 15–27.
- Kuparinen A., Markkanen T., Riikonen H. & Vesala T. (2007) Modeling air-mediated dispersal of spores, pollen and seeds in forested areas. – *Ecological Modelling* 208: 177–188.
- Küster H. (1997) The role of farming in the postglacial expansion of beech and hornbeam in the oak woodlands of central Europe. – *The Holocene* 7: 239–242.
- Lamentowicz M., Kołaczek P., Mauquoy D., Kittel P., Łokas E., Słowiński M., Jassey V. E. J., Niedziółka K., Kajukalo-Drygalska K. & Marcisz K. (2019) Always on the tipping point – A search for signals of past societies and related peatland ecosystem critical transitions during the last 6500 years in N Poland. – *Quaternary Science Reviews* 225: 105954.
- Lang G. (1994) Quartäre Vegetationsgeschichte Europas. – Gustav Fischer, Jena.

- Lechterbeck J., Edinborough K., Kerig T., Fyfe R., Roberts N. & Shennan S. (2014) Is Neolithic land use correlated with demography? An evaluation of pollen-derived land cover and radiocarbon-inferred demographic change from Central Europe. – *The Holocene* 24: 1297–1307.
- Macek M., Abraham V., Tkáč P., Novák D. & Kolář J. (2022) The diversity and distribution of introduced plant species reflects eight thousand years of settlement history. – *Journal of Ecology* 111: 787–798.
- Mander L., Baker S. J., Belcher C. M., Haselhorst D. S., Rodriguez J., Thorn J. L., Tiwari S., Urrego D. H., Wesseln C. J. & Punyasena S. W. (2014) Accuracy and consistency of grass pollen identification by human analysts using electron micrographs of surface ornamentation. – *Applications in Plant Sciences* 2: 1400031.
- Marquer L., Gaillard M.-J., Sugita S., Poska A., Trondman A.-K., Mazier F., Nielsen A. B., Fyfe R. M., Jönsson A. M., Smith B., Kaplan J. O., Alenius T., Birks H. J. B., Bjune A. E., Christiansen J., Dodson J., Edwards K. J., Giesecke T., Herzschuh U., Kangur M., Koff T., Latalowa M., Lechterbeck J., Olofsson J. & Seppä H. (2017) Quantifying the effects of land use and climate on Holocene vegetation in Europe. – *Quaternary Science Reviews* 171: 20–37.
- Matthias I. & Giesecke T. (2014) Insights into pollen source area, transport and deposition from modern pollen accumulation rates in lake sediments. – *Quaternary Science Reviews* 87: 12–23.
- Milligan G., Bradshaw R. H. W., Clancy D., Żychalak K. & Spencer M. (2020) Effects of human land use and temperature on community dynamics in European forests. – *Quaternary Science Reviews* 247: 106458.
- Morrison K. D., Hammer E., Boles O., Madella M., Whitehouse N., Gaillard M.-J., Bates J., Linden M. V., Merlo S., Yao A., Popova L., Hill A. C., Antolin F., Bauer A., Biagetti S., Bishop R. R., Buckland P., Cruz P., Dreslerová D., Dusseldorp G., Ellis E., Filipovic D., Foster T., Hannaford M. J., Harrison S. P., Hazarika M., Herold H., Hilpert J., Kaplan J. O., Kay A., Goldewijk K. K., Kolář J., Kyazike E., Laabs J., Lancelotti C., Lane P., Lawrence D., Lewis K., Lombardo U., Lucarini G., Arroyo-Kalin M., Marchant R., Mayle F., McClatchie M., McLeester M., Mooney S., Hoyo M. M., Navarrete V., Ndiema E., Neves E. G., Nowak M., Out W. A., Petrie C., Phelps L. N., Pinke Z., Rostain S., Russell T., Sluyter A., Styring A. K., Tamanaha E., Thomas E., Veerasamy S., Welton L. & Zanon M. (2021) Mapping past human land use using archaeological data: a new classification for global land use synthesis and data harmonization. – *PLoS ONE* 16: e0246662.
- Mueller-Bieniek A., Bogucki P., Pyzel J., Kapcia M., Moskal-del Hoyo M. & Nalepka D. (2019) The role of *Chenopodium* in the subsistence economy of pioneer agriculturalists on the northern frontier of the Linear Pottery culture in Kuyavia, central Poland. – *Journal of Archaeological Science* 111: 105027.
- Novák J., Kočárová R., Kočár P. & Abraham V. (2020) Long-term history of woodland under human impact, archaeoanthracological synthesis for lowlands in Czech Republic. – *Quaternary International* 593–594: 195–203.
- Odgaard B. V. (1999) Fossil pollen as a record of past biodiversity. – *Journal of Biogeography* 26: 7–17.
- Odgaard B. V. (2007) Pollen methods and studies: Reconstructing Past Biodiversity Development. – In: Scott A. E. (ed.), *Encyclopedia of Quaternary Science*, p. 2508–2514, Elsevier, Oxford.
- Oksanen J., Blanchet F. G., Friendly M., Kindt R., Legendre P., McGlinn D., Minchin P. R., O'Hara R. B., Simpson G. L., Solymos P., Stevens M. H. H., Szoecs E. & Wagner H. (2020) vegan: Community Ecology Package – URL: <https://CRAN.R-project.org/package=vegan>.
- Opravit E. (1989) Přírodní poměry laténské hradě Závist [Natural conditions of the La-Tène hillfort Závist]. – *Památky archeologické* 80: 116–120.
- Podborský V., Čížmář M., Dvořák P., Erhart A., Janák V., Medunová-Benešová A., Nekvasil J., Ondráček J., Pavelčík J., Salaš M., Stuchlík S., Stuchlíková J., Šebela L., Šmíd M., Štrof A., Tejral J. & Valoch K. (1993) Praveké dějiny Moravy [Prehistory of Moravia]. – Muzejní a vlastivědná společnost v Brně, Brno.
- Pokorná A., Kočár P., Novák J., Šálková T., Žáčková P., Komárková V., Vaněček Z. & Sádlo J. (2018) Ancient and Early Medieval man-made habitats in the Czech Republic: colonization history and vegetation changes. – *Preslia* 90: 171–193.
- Pokorný P., Boenke N., Chytráček M., Nováková K., Sádlo J., Veselý J., Kuneš P. & Jankovská V. (2006) Insight into the environment of a pre-Roman Iron Age hillfort at Vladař, Czech Republic, using a multi-proxy approach. – *Vegetation History and Archaeobotany* 15: 419–433.
- Pokorný P., Chytrý M., Juříčková L., Sádlo J., Novák J. & Ložek V. (2015) Mid-Holocene bottleneck for central European dry grasslands: did steppe survive the forest optimum in northern Bohemia, Czech Republic? – *The Holocene* 25: 716–726.
- Pokorný P. & Kuneš P. (2005) Holocene acidification process recorded in three pollen profiles from Czech sandstone and river terrace environments. – *Ferrantia* 44: 101–107.
- Punt W. & Hoen P. P. (2009) The Northwest European Pollen Flora, 70: *Asteraceae – Asteroideae*. – *Review of Palaeobotany and Palynology* 157: 22–183.

- Ralska-Jasiewiczowa M., Nalepka D. & Goslar T. (2003) Some problems of forest transformation at the transition to the oligocratic/*Homo sapiens* phase of the Holocene interglacial in northern lowlands of central Europe. – *Vegetation History and Archaeobotany* 12: 233–247.
- Richards J. D., Jakobsson U., Novák D., Štular B. & Wright H. (2021) Digital archiving in archaeology: the state of the art. Introduction. – *Internet Archaeology* 58. URL: <https://doi.org/10.11141/ia.58.23>
- Riede F. (2018) Deep pasts – deep futures. A palaeoenvironmental humanities perspective from the Stone Age to the Human Age. – *Current Swedish Archaeology* 26: 11–28.
- Roleček J., Abraham V., Vild O., Svobodová Svitavská H., Jamrichová E., Plesková Z., Pokorný P. & Kuneš P. (2021) Holocene plant diversity dynamics show a distinct biogeographical pattern in temperate Europe. – *Journal of Biogeography* 48: 1366–1376.
- Rybníčková E. & Rybníček K. (1971) The determination and elimination of local elements in pollen spectra from different sediments. – *Review of Palaeobotany and Palynology* 11: 165–176.
- Shaw H. & Whyte I. (2020) Interpretation of the herbaceous pollen spectra in paleoecological reconstructions: a spatial extension of Indices of Association and determination of individual pollen source areas from binary data. – *Review of Palaeobotany and Palynology* 279: 104238.
- Šizling A. L., Pokorný P., Juříčková L., Horáčková J., Abraham V., Šizlingová E., Ložek V., Tjørve E., Tjørve K. M. C. & Kunin W. (2016) Can people change the ecological rules that appear general across space? – *Global Ecology and Biogeography* 25: 1072–1084.
- Svoboda J. (2003) Mezolit severních Čech, komplexní výzkum skalních převisů na Českolipsku a Děčínsku, 1978–2003 [Mesolithic of northern Bohemia, complex excavation of rockshelters in the Česká Lípa and Děčín areas, 1978–2003]. – Institute of Archeology CAS, Brno.
- Svoboda J., Pokorný P., Horáček I., Sázelová S., Abraham V., Divišová M., Ivanov M., Kozáková R., Novák J., Novák M., Šída P. & Perri A. (2018) Late Glacial and Holocene sequences in rockshelters and adjacent wetlands of Northern Bohemia, Czech Republic: correlation of environmental and archaeological records. – *Quaternary International* 465: 234–250.
- Svobodová H. (1997) Die Entwicklung der Vegetation in Südmähren (Tschechien) während des Spätglazials und Holozäns – eine palynologische Studie. – *Verhandlungen der zoologisch-botanischen Gesellschaft in Österreich* 134: 317–356.
- Szabó P. (2010) Driving forces of stability and change in woodland structure: a case-study from the Czech lowlands. – *Forest Ecology and Management* 259: 650–656.
- Theuerkauf M., Dräger N., Kienel U., Kuparinen A. & Brauer A. (2015) Effects of changes in land management practices on pollen productivity of open vegetation during the last century derived from varved lake sediments. – *The Holocene* 25: 733–744.
- Timpson A., Colledge S., Crema E., Edinborough K., Kerig T., Manning K., Thomas M. G. & Shennan S. (2014) Reconstructing regional population fluctuations in the European Neolithic using radiocarbon dates: a new case-study using an improved method. – *Journal of Archaeological Science* 52: 549–557.
- Torring T. (2015) Neolithic population and summed probability distribution of 14C-dates. – *Journal of Archaeological Science* 63: 193–198.
- van Leeuwen P., Punt W. & Hoen P. P. (1988) *Polygonaceae*. – *Review of Palaeobotany and Palynology* 57: 81–151.
- Venclová N., Drda P., Michálek J., Militký J., Salač V., Sankot P. & Vokolek V. (2008) Archeologie pravěkých Čech 7. Doba laténská [Archaeology of prehistoric Bohemia 7. La-Tène]. – Institute of Archeology CAS, Praha.
- Whitehouse N. J., Schulting R. J., McClatchie M., Barratt P., McLaughlin T. R., Bogaard A., Colledge S., Marchant R., Gaffrey J. & Bunting M. J. (2014) Neolithic agriculture on the European western frontier: the boom and bust of early farming in Ireland. – *Journal of Archaeological Science* 51: 181–205.
- Wieczorek M. & Herzschuh U. (2020) Compilation of relative pollen productivity (RPP) estimates and taxonomically harmonised RPP datasets for single continents and Northern Hemisphere extratropics. – *Earth System Science Data* 12: 3515–3528.
- Wood S. (2013) mgcv: Mixed GAM Computation Vehicle with Automatic Smoothness Estimation. – Comprehensive R Archive Network (CRAN), URL: <https://CRAN.R-project.org/package=mgcv>.
- Woodbridge J., Fyfe R., Smith D., Pelling R., de Vareilles A., Batchelor R., Bevan A. & Davies A. L. (2021) What drives biodiversity patterns? Using long-term multidisciplinary data to discern centennial-scale change. – *Journal of Ecology* 109: 1396–1410.

Pylové antropogenní indikátory přehodnocené pomocí rozsáhlých souborů pylových a archeologických dat: 12 tisíc let interakcí mezi člověkem a vegetací ve střední Evropě.

Sledování interakcí mezi člověkem a vegetací v minulosti bylo vždy jedním z klíčových témat paleoekologie. Zde využíváme velké soubory pylových a archeologických dat z České republiky pro testování vztahů mezi jednotlivými pylovými taxony a archeologickými údaji a hledáme prostorovou škálu, na které si tato data nejlépe odpovídají. Oba soubory dat, zahrnující 1 500 pylových vzorků a 65 000 archeologických položek a pokrývající období od 12 000 až po 700 let před současností, byly rozděleny do časových oken o délce 250 let. Vztah mezi pylem a archeologickými daty na jednotlivých lokalitách jsme měřili Spearmanovou pořadovou korelací. Varianci pylu v celém souboru dat jsme za použití zobecněných aditivních modelů (GAM) vysvětlovali pomocí archeologických dat a dalších dvou proměnných prostředí. První je holocenní trend zastoupení pylových taxonů, druhou je nadmořská výška pylové lokality. Obě ovlivňují abundanci druhů v celém studovaném období a území a nelze je statisticky oddělit od změn způsobených člověkem. Obě také snižují indikační hodnotu antropogenních pylových indikátorů, nadmořská výška více než holocenní trend, protože hustota archeologických nálezů a nadmořská výška spolu silně korelují a tvoří hlavní gradient. Pylové taxony s pozitivní korelací k intenzitě lidské činnosti jsou *Plantago lanceolata*, *Artemisia* a *Amaranthaceae*, *Quercus* (vegetativně zmlazující dřevina s jedlými semeny a částečnou odolností proti požárům) a *Pinus* (pionýrská dřevina, expandující do 250 let od lidského narušení). Indikační pravděpodobnost pro přítomnost archeologie je u těchto taxonů vysoká (0,63–0,85), ale variabilita vysvětlená archeologií v úplném modelu je nízká (0,01–0,09). *Fagus*, *Carpinus* a *Abies* expandují během pozdních sukcesních fází po lidském narušení, a proto je jejich vztah k minulé lidské činnosti negativní, uvažujeme-li 250leté časové okno. *Secale* nekoreluje na úrovni jednotlivých lokalit kvůli pozdnímu výskytu (od vrcholného středověku). Slabý vztah mezi archeologickými daty a pylem obilí připisujeme nekonzistentnímu určování. Rádus nejsilnější korelace pylu a archeologických dat se pohybuje kolem desítek kilometrů kvůli prostorovému rozlišení archeologických dat danému rozlohou katastru, ale stále je nižší u bylin (15–20 km) než u stromů (30–40 km).

How to cite: Abraham V., Macek M., Tkáč P., Novák D., Pokorný P., Kozáková R., Jamrichová E., Soukupová M. G. & Kolář J. (2023) Pollen anthropogenic indicators revisited using large-scale pollen and archaeological datasets: 12,000 years of human-vegetation interactions in central Europe. – *Preslia* 95: 385–411.

Preslia, a journal of the Czech Botanical Society

© Česká botanická společnost / Czech Botanical Society, Praha 2023

www.preslia.cz

This is an open access article published under a CC BY license, which permits use, distribution and reproduction in any medium, provided the original work is properly cited (Creative Commons Attribution 4.0 International License, <http://creativecommons.org/licenses/by/4.0>).