

Sampling and its effect on the composition of archaeobotanical assemblages from the Roman Period

Vzorkování a jeho vliv na složení archeobotanických souborů z doby římské

– Jana Apiar*, Peter Apiar –

KEY WORDS

Archaeobotany – sampling – plant macro-remains – statistical modelling – Roman Period

ABSTRACT

The subject of the presented study is taken from a dissertation project by one of the authors who focused on the processing of archaeobotanical assemblages from the Roman Period. The main aim of the research was the reconstruction of selected aspects of the subsistence strategy of the population in the given period based on the evaluation of archaeobotanical data from various chronological and cultural contexts in a designated region, available to author. The analysed sets were obtained during field excavations primarily conducted in the last third of the 20th century and the beginning of the 21st century. Uniform methods of archaeobotanical sampling were not applied in the acquisition of these assemblages. Source information on the origin of the samples was considerably heterogeneous and, in many cases, distinctly fragmentary. This was the impulse behind the investigation into the question as to whether, and to what extent, the method of sampling affects the interpretive value of the investigated dataset and what are the limitations of the analysis of such a dataset.

The principal aim of this study is not the archaeobotanical evaluation of samples, but rather to investigate a possible effect of their formal properties on the composition of archaeobotanical finds. The formal properties studied include the volume and the number of collected samples, and the spatial stratification of samples (context/feature). Intuitively, it would appear that the heterogeneous quality of this information may have a certain impact on the interpretive value of an archaeobotanical assemblage. We discuss the effect of the chosen method of sampling on the composition of macro-remains in archaeobotanical samples and assemblages with the use of statistical models.

Introduction

The study presented is based on the results of a dissertation project (Hlavatá 2017), which focused on the processing of archaeobotanical assemblages from the Roman Period. The main aim of the project was the reconstruction of selected paleo-economic aspects of the population in the given period, based on the evaluation of all the available archaeobotanical data from various chronological and cultural contexts in the designated region. The analysed assemblages were obtained primarily during rescue field excavations conducted in the last third of the 20th century and the beginning of the 21st century. Uniform methods of archaeobotanical sampling were not applied in the acquisition of these assemblages. This meant that the formal information was considerably heterogeneous and regarding the aim of the dissertation project, the use of samples was, for this reason, limited. This was the impulse behind the investigation of methodological questions as to whether and to what extent, the method of sample collection affects the informative value of the investigated dataset and to show the obstacles that can prevent archaeobotanical and statistical evaluation from using the full potential of archaeobotanical data (cf. Hlavatá 2017, 212–224). Under the formal properties of the studied samples investigated, we understand the volume of sediment collected, the number of samples collected and the source of the samples, i.e. the feature (context). The principal aim of this study is not to evaluate samples from the viewpoint of traditional archaeobotanical interpretation, but rather from the viewpoint of the effect of their formal properties on the composition of plant macro-remains and the identified species. From a large number of published archaeobotanical works, above all, it is well-known that the composition of macro-remains is dependent on the circumstances of their deposition in the place of their future recovery. Moreover, on the processes that plants and fruits undergo during growth, harvest, storage, handling before direct consumption and so on (e.g. Hillman 1984; Jones 1984; van der Veen 1992, 81–82). Nevertheless, the purpose of this study was primarily to verify the statement that the interpretation of an archaeobotanical assemblage can be related to the inconsistency of formal sample properties, rather than to the composition of the macro-remains in the samples.

The collection and processing of archaeobotanical samples are accompanied by a number of various factors. The most important are:

- a. the chosen method of sample collection during excavation (total, systematic interval, column sampling, i.e. sampling of profiles, judgmental, combined),
- b. archaeological and archaeobotanical field documentation, identification of archaeological structures and (post)deposition processes, storage of samples before extraction,

* Corresponding author – E-mail address: apiar@arub.cz

<https://doi.org/10.47382/pv0621-08>

Received 22 February 2021; received in revised form 21 May 2021.
Available online 30 June 2021.

Copyright © 2021 Czech Acad Sci, Inst Archaeology Brno, and the authors.
This is an open access article under the CC BY-NC-ND 4.0 license
(<https://creativecommons.org/licenses/by-nc-nd/4.0/>).
Competing interests: The authors have declared that no competing interests exist.

- c. floatation, microscopic analysis,
- d. the construction of an appropriate database and correct data selection,
- e. the use of statistical methods and interpretation of the results.

The data that archaeobotanists (and archaeozoologists) generally work with is not ideal and is already subject to reduction during the field excavation. Despite this, the data has great potential and irreplaceable informational value. However, it is important to recognise all, or as many as possible, of its limitations to better interpret the bygone reality that the data reflects. These limitations will be discussed below with the use of statistical analyses and models. This factor is frequently neglected during the realisation of field excavations – either consciously or as a result of the circumstances.

The basic dataset

The region of interest for this study is the area to the north of the River Danube, restricted largely to southwestern Slovakia, southern Moravia and central Bohemia (Fig. 1). This demarcation resulted from the available archaeobotanical material that comprised the original database of the dissertation project (Hlavatá 2017, 19–21, 23–26). The material comprising the subject of this study comes exclusively from the database of the given project and is used in line with the copyright of individual experts.¹

In chronological terms, the dataset belongs to the period from the beginning of the 1st century AD to the end of the 4th century or the beginning of the 5th century AD, which corresponds to the Roman Period and the beginning of the Migration Period in the chronology of the given region. Since information on the more detailed chronology of the contexts from which the individual archaeobotanical samples come was often ambiguous or vague, closer dating into individual stages was excluded from the analyses in this study (Hlavatá 2017, 21–22, 60).

For the purpose of the study, an *archaeobotanical sample* is a sample of sediment collected from the excavated archaeological situation for flotation and extraction of plant macro-remains for archaeobotanical analysis. The term *method of sampling* is understood here as the sampling itself, together with the documentation and flotation of samples. The available archaeobotanical samples came from a total of 54 sites. The original dataset was comprised of 1,000 archaeobotanical samples. In total,

	Volume (litres)	PMR	Taxa
Minimum	0.2	0	0
Maximum	80	8076	49
Sum	16764.5	60077.2	238
Median	15	4.5	3
Mean	17.33	62.1	4.7
Variance	173.3	124781,7	29.4
Standard deviation	13.2	353.2	5.4
Skewness	1.4	15.3	2.9
Kurtosis	2.8	301.1	12.9
Shapiro-Wilk test	0.9	0.1	0.7
Shapiro-Wilk – p-value	0	0	0

Tab. 1. Descriptive statistics for volume (in litres), “one plant macro-remain” (abbreviated below to PMR) and botanical species in samples.

Tab. 1. Deskriptivní statistika pro objem (v litrech), „jeden rostlinný makrozbytek“ (dále zkráceně RMZ) a botanické druhy ve vzorcích.

70,032 plant macro-remains (not taking charcoal into account) were identified in the samples, numbering 309 botanical species (Hlavatá 2017, 63–64, Príloha Mapa 1–3, Katalóg). Volumes were recorded for just 686 of the total number of samples (a total of 13,757.8 litres of collected sediment). Cases in which the volume of the sample was not recorded were excluded from the analyses in this study, meaning that the number of sites fell to 25 (Fig. 1). The number of macro-remains found fell to 36,460 and the number of botanical species to 238. The overall variance of the volumes of the samples in the studied set was from 0.2 to 80 litres and the variance of findings of macro-remains from 0 to 8,076 pcs (Tab. 1).

Archaeological sites are not given by name or otherwise specified because of authorial ethics (see above), except for the Jevišovka site, which was excavated under the supervision of B. Komoróczy within the activities of the authors' workplace, which is described here as an illustrative example. The set as a whole is understood here as a model set, for the reason that what is important from the viewpoint of its analysis are not specific localisations, but the properties of the samples from them.

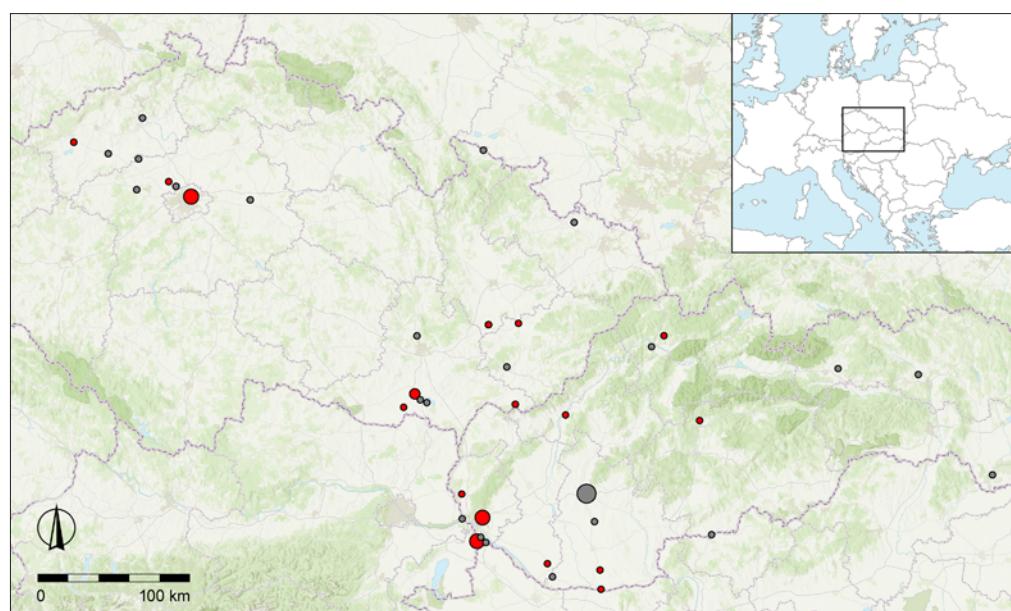


Fig 1. Map of archaeological sites where archaeobotanical samples were collected. Red – sites with available information about the volume of samples, grey – sites without information about the volume of samples. Author M. Vlach.

Obr 1. Mapa archeologických lokalit s odebíranými archeobotanickými vzorky. Červená – lokality s dostupnou informací o objemu vzorků; šedá – lokality bez informace o objemu vzorků. Autor M. Vlach.

Archaeobotanical samples

In contrast to the methodology for the collection of samples recommended in the archaeobotanical literature (cf. Barker 1975; Hajnalová 2012, 33; Jacomet et al. 1989, 36–41, 70–84; Jacomet, Kreuz 1999, 95–112; Jones, M. K. 1991, 53–62; Klečka 1934, 98; Kočář, Dreslerová 2010, 205; Lee 2012; Orton 2000; Lityńska-Zajac, Wasylikova 2005, 162–166; Pearsall 2000, 66–76; Renfrew, Bahn 2008, 276; Smith 2014; van der Veen 1982; 1984; 1992, 21, etc.) sediment from contexts has often been collected in a subjectively selective manner in archaeological practice, particularly in older excavations, and frequently with no predetermined theoretical or methodological concept (cf. Hajnalová, Varsík 2010; Hlavatá, Varsík 2019; Kočář, Dreslerová 2010). This effect is also manifested in the set studied here. Despite this, a certain number of systematically collected samples did appear among the analysed sets. Therefore, source information on the origin of the samples was considerably heterogeneous and often distinctly fragmentary.

Within the dataset, archaeobotanical samples were collected systematically (systematic interval sampling; cf. Jacomet, Kreuz 1999, 97–100; Pearsall 2000, 66–76) at four sites. Two sites were sampled in a combined manner – systematically and judgmentally (Hlavatá 2017, 28–29). In the case of the sites from which the samples were obtained in an unfloated/unprocessed state (Hlavatá 2017, 30), the only information on the method of sample collection was the descriptive information on the tags. These tags generally contained simple information on the feature number. In some cases, they contained information on the layer; in other cases, merely the serial number of the sample collected without any additional information. In most cases, additional information about sediment was not available so was not used for analysis at all. This method of sample description, without even a detailed sample catalogue accompanying the archaeobotanical material, makes the proper analysis and interpretation of archaeobotanical material difficult or, in certain cases, entirely impossible. It is precisely in the method of sampling and extraction of plant material from the samples taken, coming from diverse types of archaeological excavations (mostly rescue, partly science-driven), that the set displays the greatest variability. Disparity is also manifested in the archaeological method used within a single excavation (e.g. various excavation seasons in a single location, individual trenches, various excavation aims). Evidential value is also affected by the non-constant volume of individual samples or the volume of samples in general. Collecting sporadic easily visible findings of seeds (for direct analysis, generally without flotation, i.e. without sediment) falls into a category of its own. Samples collected in such a way are extremely difficult to use for the purposes of archaeobotanical analysis, such as the calculation of the density of plant macro-remains. The term “volume of sample” is not interchangeable with the term “volume of context”, unless the whole fill of the studied context is taken. We cannot simply replace the missing variable “volume of sample” with “volume of context”,² or vice versa, which also applies to density. This factor must be excluded from analysis in the absence of information on the volume of sample and replaced where necessary with another indicator, such as the total number of macro-remains or an index of “sample purity” (e.g. Simpson’s index D, *sensu lato* Wallace et al. 2019, 452). However, as stated by M. Wallace et al. (2019, 452), the number of macro-remains depends not merely on the density of the sample, but also on the size of the sample. In our case, not a single archaeological feature or context was totally sampled (in the sense that the entire filling was floated).

Archaeological source information

The studied set suffered from a lack of archaeological information. It was not possible to divide or compare all the results of archaeobotanical analysis by types of features or contexts. A clear methodological and terminological difference also appeared between samples from the Slovak, Moravian and Bohemian sites.

The lowest level of reference archaeological information was represented by an unspecified context for which descriptive information was stated vaguely as “feature”, “layer” or “sector” without further description. At the highest level, designation of the context was found by number, type, function and part (e.g. feature No. 1, pit, storage pit, western half etc.), followed by the designation and specification of the context (e.g. K8, destruction layer, etc.), with the archaeological feature also being situated spatially within the excavation area or trench (e.g. area P1, trench IV, sector 7/b, etc.). Spatial information about the actual samples, such as extraction depth, etc. was also sparse.

Selection, standardisation and processing of data

All the obtained carbonised plant macro-remains preserved in dry conditions (e.g. Hajnalová 1993, 11–12; 2001, 11; Jacomet, Kreuz 1999, 59–62; Lityńska-Zajac, Wasylikova 2005, 42–43), apart from findings of charcoals and woods, were included in the analysis. The term “one plant macro-remain” (abbreviated below to PMR) is understood to mean each whole grain, seed or fruit (*sensu lato*) that has retained a coherent shape, even if it has been largely destroyed. Fragments were converted in the standard manner (e.g. van der Veen 1992, 25; cf. Hajnalová 2012, 44; Hlavatá 2017, 46–48; Látková 2017, 29) to make their inclusion in the overall totals of the macro-remains possible. Fragments of straw, reed and carbonised organic porous material were left in units of fragments. The term “species” is understood as a botanical species (taxon). In the *species* variable, some uncertain determinations were also included. Specifically, these were the uncertain determination of species (cf.) or determination up to genus caused by serious degradation of a macro-remain. However, this approach was used only for samples where no other definitely determined species was present. For quantification of species abundance, ubiquity was used. This means each species was counted only once. In particular, given the species abundance in a feature, only the first occurrence of a species was recorded.

The analysed data was divided into four $m \times n$ data tables, in which m represents the number of observations (samples, features/contexts, sites) and n represents the variables (botanical species, PMR, volume, density). The variables *species*, *macro-remains* and *sample volume* were selected concerning their direct connection with the calculation of the density of macro-remains per litre of sediment and the calculation of ubiquity for archaeobotanical analysis, in general. In the same way, these variables are essential for further taphonomic, paleoeconomic and plant ecology analysis of archaeobotanical samples (e.g. Bogaard 2004; Hajnalová 2012, 95–110; van der Veen 1992, 73–90, 111–144).

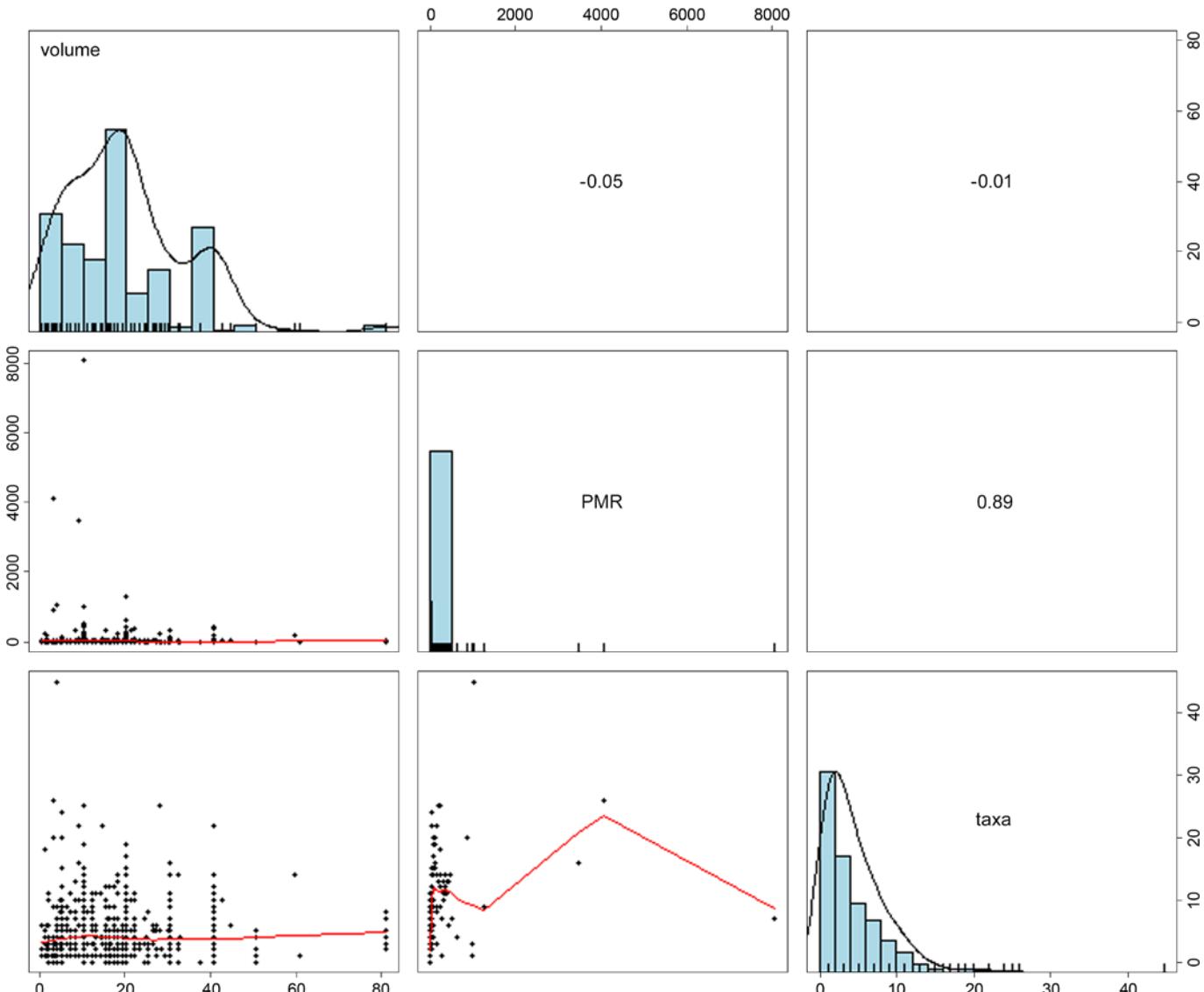
Archaeobotanical data is generally adjusted for statistical analyses. This usually involves eliminating samples with a small number of findings and botanical species (taxa) that occur in a small number of samples (cf. Bogaard 2004; van der Veen 1992; in summary Reed 2016, 212–214, and further literature). In general, only those samples that contain 10 or more findings and taxa that occur in three or more samples, or samples that contain at least 50 or perhaps 100 findings of crop seeds and at least 10 findings of weed seeds, are used (in line with Bogaard 2004, 60–64, and further literature; Hajnalová 2012, 122;

Hajnalová, Varsik 2010, 210; Jones, G. E. M. 1991, 67–69; Látková 2017; van der Veen 1992). M. Kuna et al. (2013, 130) argue that such adjustment of archaeobotanical data leads to the loss of a large spectrum of findings. In short, data standardisation alters the dataset in favour of rich/richer samples, while poorer samples are excluded from analysis even though they may often contain relevant and valuable findings and/or information. A large number of poorer samples may be more representative than a small number of rich samples (Kuna et al. 2013, 130; Orton 2000, 167). This form of dataset processing has greater justification in view of a more comprehensive interpretation of archaeobotanical data and may reveal taphonomic processes or differences in the management of cultivated crops, for example (Kuna et al. 2013, 130).

This situation came about primarily in the case of the sum totals of the macro-remains (PMR) in individual samples that showed a larger dispersion in statistical models (below) and often caused overdispersion or underdispersion. This did not

change, even after the removal of outliers. The solution would be either to remove around 10% of the values from the top and bottom or to transform the data. However, this would lead to the loss of a large amount of information and also exclude null values, and last but not least, the transformed data needs to be handled with more caution (O'Hara, Kotze 2010; Field et al. 2012, 191–193). The data was not reduced because information on the absence of macro-remains in a sample is also relevant data. An essential part of the analysis was the observation of raw data in its original state with the smallest possible interventions in view of the assumption that this data more faithfully reflects the true method of sampling.

Combining samples according to independent sampled elements (*sensu stricto* Lee 2012, 650 – elements/units of population; cf. Orton 2000, 30–39, 165–168) or stratigraphically and contextually indistinguishable samples is necessary for further archaeobotanical analysis since these do not create independent entities (Jones, G. E. M. 1991, 67–69; Smith 2014, 182, and



Graph 1. Graph of the correlation matrix displaying relationships between volume in litres, PMR (plant macro-remains) and botanical species (taxa). Bottom left – correlation graphs; diagonally – histograms with kernel density estimations; top right – Spearman's correlation coefficient.

Graf 1. Graf korelační maticy zobrazující vztahy mezi objemem (volume) v litrech, RMZ (PMR – plant macro-remains) a botanickými druhy (taxa). Levá dolní část – korelační grafy; diagonálně – histogramy s jádrovými odhady hustoty (kernel density); pravá horní část – Spearmanův korelační koeficient.

further literature; cf. Kuna et al. 2013; Látková 2017). However, it was possible to combine only five samples in this way, which contained sufficient archaeological information and also came from the same context or its part.

As far as the selection of features and their types are concerned, these were included in the analysis to reflect the true state of our set. This means that samples from identified contexts or features were analysed the same as samples with a general designation (e.g. "layer", without more specific description, etc.).

Given the issue considered here, non-reduced data, i.e. almost all samples³ that held information about the volume of sediment collected, were used for analysis. Samples that did not contain any findings of macro-remains or contained just a small amount of PMR (e.g. 1 PMR, etc.) were also analysed. The aim was to discern possible elements within the material that would not be included in statistical evaluation or that could escape undetected during the interpretation of the results in the case of "traditional" standardisation.

Statistical models

A closer investigation of the characteristics of the volumes of the samples, the quantity of macro-remains and the number of species identified in them were required to answer the question posed as to whether sampling affects archaeobotanical interpretation. With the help of statistical models, we decided to investigate the mutual relationships between the given variables from the viewpoint of individual samples (Models 1 to 5), from the viewpoint of features/contexts (Models 6 to 11) and the viewpoint of sites (Models 12 and 13). We used the RStudio (RStudio Team 2020) program to create statistical models.

The data (Apiar 2021) was first subjected to descriptive statistics and basic visual representation before creating statistical models to reveal its fundamental properties (Tab. 1, Graph 1). As can be seen from Table 1, the skewness and kurtosis values are large, particularly for species and PMR, and they clearly dominate. Additional indicators are large differences between means, dispersions and standard deviations. In the case of PMR in particular, this difference points to overdispersion – a common phenomenon in ecological and environmental data that may have different causes (O'Hara, Kotze 2010; Zuur et al. 2009, 270–271; Martin et al. 2005; in summary Hilbe 2011, Chapter 4). Similarly, the value of the *Shapiro-Wilk* test and its p-value (Tab. 1, Shapiro-Wilk) points to the fact that none of the variables meets the conditions of normal distribution (e.g. Shennan 2006, 156–159).

Graph 1 depicts raw data in which any search for a normal distribution would be worthless. Histograms and their distributions are displayed diagonally accompanied by kernel density estimation curve (Baxter 2015, 37–49). The data on the volumes of individual samples show marks of a bimodal distribution, while data on species is visibly skewed to the left. The distribution of PMR is influenced profoundly by the larger number of low and null values in combination with extremely high values (cf. Tab. 1). These stretch the X axis and compress values into a single interval (bin). Scatterplots along with a trendline are displayed in the bottom left of Graph 1.⁴ Almost no dependency can be observed in the graph for volume with PMR and volume with species. The relationship between botanical species and the quantity of macro-remains cannot be unambiguously described from the correlation graph.

The top right of Graph 1 displays the degree of association according to Spearman's correlation coefficient.⁵ It shows a strong positive correlation between PMR and species, indicating that the number of identified species increases along with the increase in the number of PMR extracted. At first glance, it would

appear logical but this assertion is not necessarily causally conditioned. In effect, this correlation is mainly caused by the fact that each macro-remain basically represents a botanical species and vice versa. The almost null correlation between volume and PMR/species (Graph 1), which also corresponds with the results in the scatterplots, also appears counterintuitive. This would indicate that the number of PMR and species found does not depend on the amount of sediment collected. Greater attention will be devoted to this issue below.

A search for a statistical model that would best represent the processed data to enable further statistical assessment was conducted based on the results of descriptive statistics and basic data analysis (Orton 2000, 9–10; 1980, 20–21). Since the response variables did not meet the preconditions of normal distribution, a negative binomial distribution that best approximated the studied data was used. This is actually an extension of the Poisson distribution. It contains an extra parameter for overdispersion (Hilbe 2011, 253) and is often used in ecological and environmental studies (e.g. Zuur et al. 2009, 225–230).

Generalised additive models (GAM),⁶ which are an extension of generalised linear models (GLM), although GAM do not presume a normal distribution or linearity, resp. a parametric form of data (Faraway 2016, Chapter 15; Wood 2017; Zuur et al. 2009), were used primarily in modelling.

The main difference from linear models lies in the fact that GAM use non-parametric smooth functions in place of β parameters (Wood 2017).

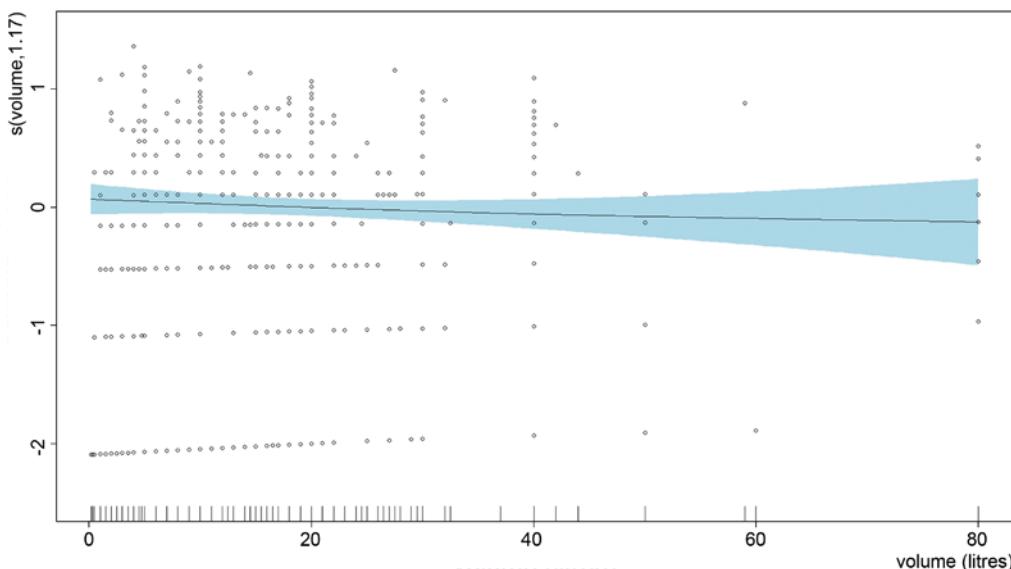
A total of 13 simple models were created, with each model represented by a single graph (Graph 2–14) with the corresponding resultant values in Table 2. There were two reasons for the larger number of models: 1) an endeavour to illustrate the analytic approach and observe the individual variables separately, and 2) macro-remains (PMR), in combination with other variables, often caused overdispersion or underdispersion due to the nature of the data. Multiple variables were introduced into the model where this was possible and meaningful. The words "influence" and "effect" (referred to as edf)⁷ are often used in the description. These can be taken as synonyms to a certain extent, although effect refers to the statistical aspect of the model, while influence refers to the archaeobotanical aspect.

In the case of models calculated using a tensor product interaction⁸ (multi-coloured graphs), we reduced the penalisation with a gamma function from the recommended value of 1.4 to an initial (default) value of 1 to observe the finer nuances of the data in the model (Wood 2017, 227; Zuur et al. 2009).

Models 1 to 5

An $m \times n$ matrix, in which m represents the number of samples (683) and n represents variables – volume of the sample, number of identified macro-remains (PMR) in the sample, number of botanical species found in the sample, density of macro-remains in a litre of the sample, was used for Models 1 to 5 (Graphs 2–6).

Model 1 (Graph 2) depicts the relationship between botanical species and sample volume. This relationship is not statistically significant; its effect on the number of species is negligible and explains just 0.24% of total variability. The samples were then divided into four volume categories by the number of litres (A: 1–9 litres, B: 10–29 litres, C: 30–40 litres, D: 41–80 litres) for verification. The statistically most significant and largest effect (edf 2.22) was seen in category A, i.e. samples with the smallest number of litres. The other categories essentially did not have any effect on the number of species found in the samples. From the viewpoint of the interpretation of this model, this indicates that the volume of the sample itself plays almost no role in how



Graph 2. Model 1 displaying the relationship of volume to the number of botanical species in samples. X axis – volume in litres, Y axis – smooth function of volume with edf value (effective degrees of freedom).

Graf 2. Model 1 zobrazující vztah objemu k počtu botanických druhů ve vzorcích. Osa X – objem v litrech (volume); osa Y – smooth funkce objemu (volume) s hodnotou edf (effective degrees of freedom).

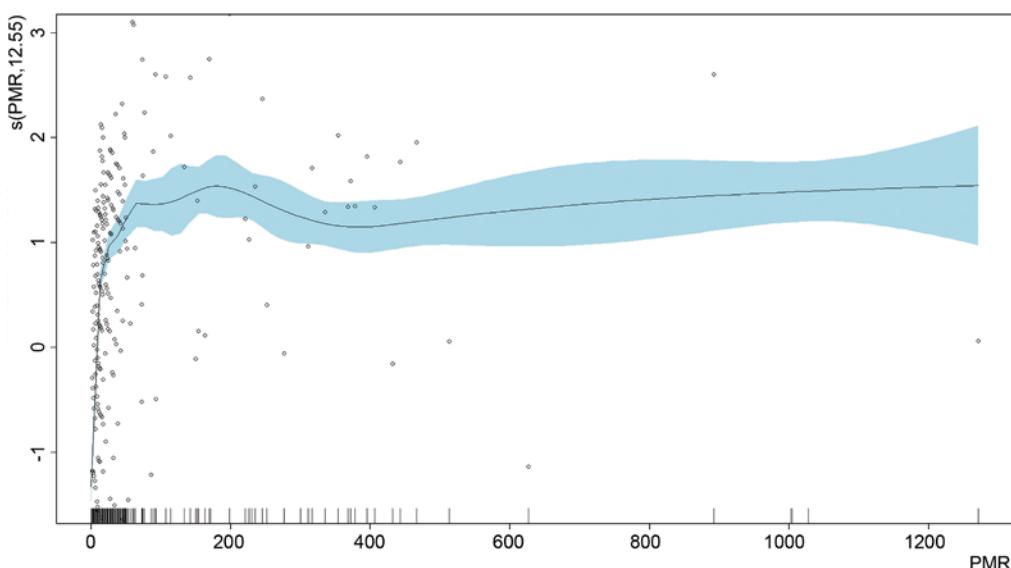
many species occur in the sample, although there is a visible effect of smaller samples on the species composition recorded from the viewpoint of volume categories. It is an accepted fact that any archaeobotanical or archaeological interpretation need not necessarily be so direct; see below.

Model 2 (Graph 3) studied the relationship between botanical species and the number of macro-remains (PMR). As previously mentioned, the values of PMR are problematic. This was manifested in the model primarily by underdispersion, which is frequently caused by the clustering of low values or by adapting the model to multiple extreme values (Hilbe 2014, 136; Zuur et al. 2009, 225). The combination of these factors can be considered in the presented example. This is graphically illustrated in the corresponding graph, in which the number of species shows a relationship primarily to the lower numbers of the macro-remains, and a pronounced break in the curve caused by outlying values can be observed at an amount of around 50 to 100 macro-remains. When the numbers of macro-remains are higher, their relationship to the number of species appears “flat” and we cannot refer to their mutual dependence in the sense that larger volumes of samples would result in richer archaeobotanical assemblages. A model from the GLM family was created with the use of negative binomial regression for verification of the mutual

dependence of this relationship.⁹ The resultant coefficient for PMR ($\log 0.0032$) may be positive, but it is essentially negligible. The chance that the acquisition of one macro-remain more would also result in the acquisition of another species is in this case almost zero. The GLM outcome corresponds with the results of Model 2, that a larger number of species cannot necessarily be expected from a larger number of macro-remains (PMR).

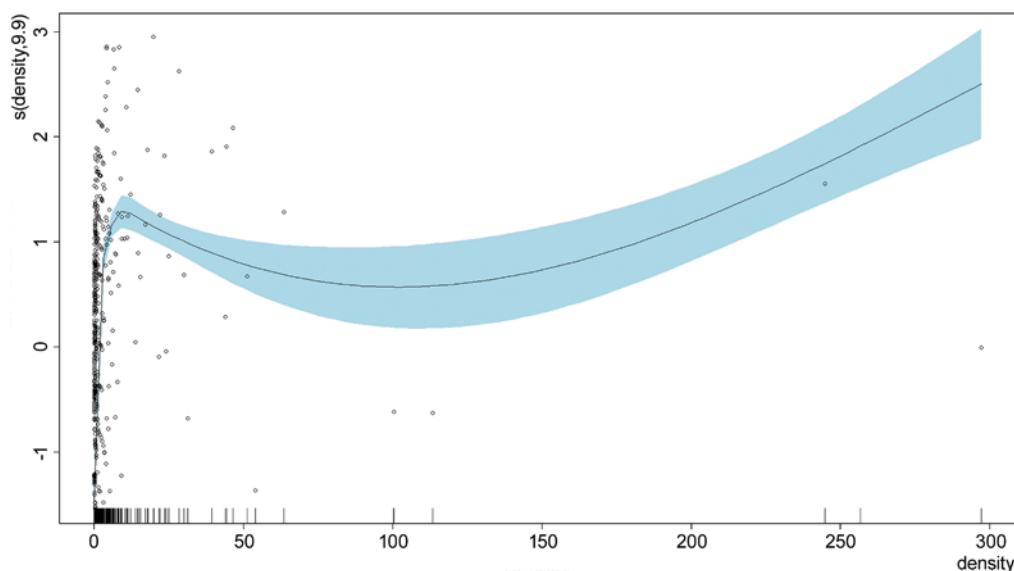
Model 3 (Graph 4) is presented to illustrate the relationship between the number of species and the density of macro-remains in the sample. The density of macro-remains is calculated as the ratio between the number of macro-remains and the volume of the sample from which they come. This gives us the density or concentration of macro-remains per litre of sediment in the given sample.

A comparison of Models 2 and 3 shows pronounced similarities, both in terms of the shape of the function (it shows similar behaviour to PMR) and in respect of the values (Tab. 2). However, density is implicitly contained in the variable PMR (resp. volume) in all the models, for which reason it will be excluded from the other models. It is presented here mainly as a comparison with other models and because of its frequent use in archaeobotanical analyses and studies (e.g. Lee 2012, with further literature; Látková 2017; Hajnalová 2012; Kuna et al. 2013, etc.).



Graph 3. Model 2 displaying the relationship between PMR and the number of botanical species in samples. X axis – počet RMZ (PMR), Y axis – smooth funkce RMZ (PMR) s hodnotou edf (effective degrees of freedom).

Graf 3. Model 2 zobrazující vztah RMZ k počtu botanických druhů ve vzorcích. Osa X – počet RMZ (PMR); osa Y – smooth funkce RMZ (PMR) s hodnotou edf (effective degrees of freedom).



Graph 4. Model 3 displaying the relationship between density and the number of botanical species in samples. X axis – density of macro-remains per litre, Y axis – smooth function of density with edf value (effective degrees of freedom).

Graf 4. Model 3 zobrazující vztah hustoty k počtu botanických druhů ve vzorcích. Osa X – hustota makrozbytků na jeden litr (density); osa Y – smooth funkce hustoty (density) s hodnotou edf (effective degrees of freedom).

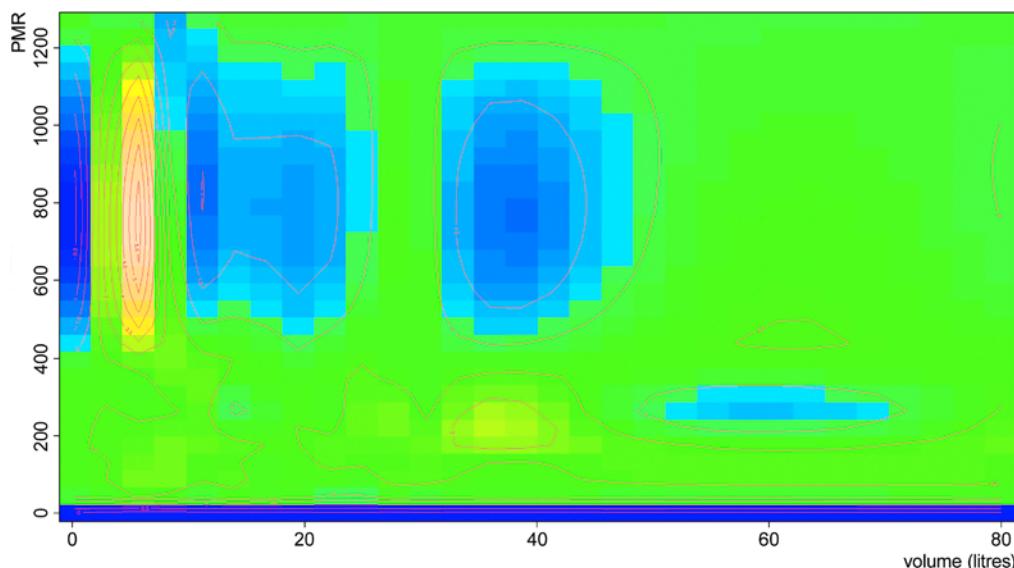
Model 4 (Graph 5) depicts a very similar relationship to that seen for the density of PMR (Graph 4), although with the difference that it is calculated here with the use of tensor product interaction. This models the mutual “pure” interaction between the studied variables and does not take into consideration the main effect of the individual variables. These variables also need not be recorded in the same units to be modelled. The diagram is also more readable, as the X and Y axes are displayed in the original units. As can be seen, Model 4 (Graph 5) condenses the results of Models 2 and 3, with the greatest effect observed at lower volumes and larger numbers of macro-remains. A slightly smaller effect can be seen in the bottom part of the graph. This reflects the situation when larger samples were collected in which a larger number of various species was identified.

Model 5 (Graph 6) shows a closer analysis of the relationship between the number of macro-remains and the volume of individual samples. This model was influenced by overdispersion, again due to the properties of the PMR. Yet, the calculated function (curve) is statistically significant and substantiated by the large effect of the sample volume on the amount of macro-remains obtained from the sample. However, the relationship between these variables is markedly non-linear, for which reason whether we also obtain a larger number of PMR with higher

sample volumes and vice versa cannot be simply predicted. Meanwhile, the model explains a mere 6.5% of the variability, and the volume is not, therefore, sufficient on its own to reveal further hidden structures in the data. Slightly better information was obtained during modelling with the use of volume categories. A large effect (edf 6.3) and statistical significance was exhibited by category B (10–29 litres), which has the greatest potential from the statistical viewpoint during the acquisition of a larger number of PMR.

Models 6 to 11

Models 6 to 11 (Graphs 7–14) investigated the same types of relationships as the previous models, although from the viewpoint of features or contexts. This provides a rather more intuitive view of the influence of the sampling of archaeological situations. It was also possible to monitor the given relationships by the numbers of samples taken from individual features. However, it must be stated that no more than two samples were retrieved from as many as 80% of these features. This is a rather unrepresentative dataset from the archaeobotanical viewpoint but from the statistical viewpoint, this is a relatively “evenly” sampled set, although influenced to a large extent by the inconsistency of the sediment volume taken from features or contexts.



Graph 5. Model 4 displaying the interaction (effect) between volume, PMR and botanical species in samples. X axis – volume in litres, Y axis – number of PMR. Yellow to white = strongest interaction, dark blue = weakest interaction.

Graf 5. Model 4 zobrazující interakci (efekt) objemu a RMZ k botanickým druhům ve vzorcích. Osa X – objem v litrech (volume); osa Y – počet RMZ (PMR). Žlutá až bílá = nejsilnější interakce; tmavě modrá = nejslabší interakce.

Model 6 (Graph 7) observes the relationship between volume and the number of species. The samples are concentrated largely on the left-hand side of the graph in the area of lower total volumes per feature, while the function (curve) grows most to around 250 litres taken from the feature and then veers towards outlier values. According to the data in Table 2 (m6), the total volume of sediment taken from the feature has a clear effect on the number of species found in the feature, particularly at lower volumes (up to around 250 litres). This is rather a different situation concerning Model 1, suggesting that the combined volume from all samples within a feature can affect the species composition of such a feature, which is crucially important when taken into account. This does not, of course, mean that 250 litres of sediment must be collected from each feature to obtain a larger, i.e. a more representative amount, of archaeobotanical material. These figures serve rather as an aid to decision-making relating to a future strategy for the collection of samples.

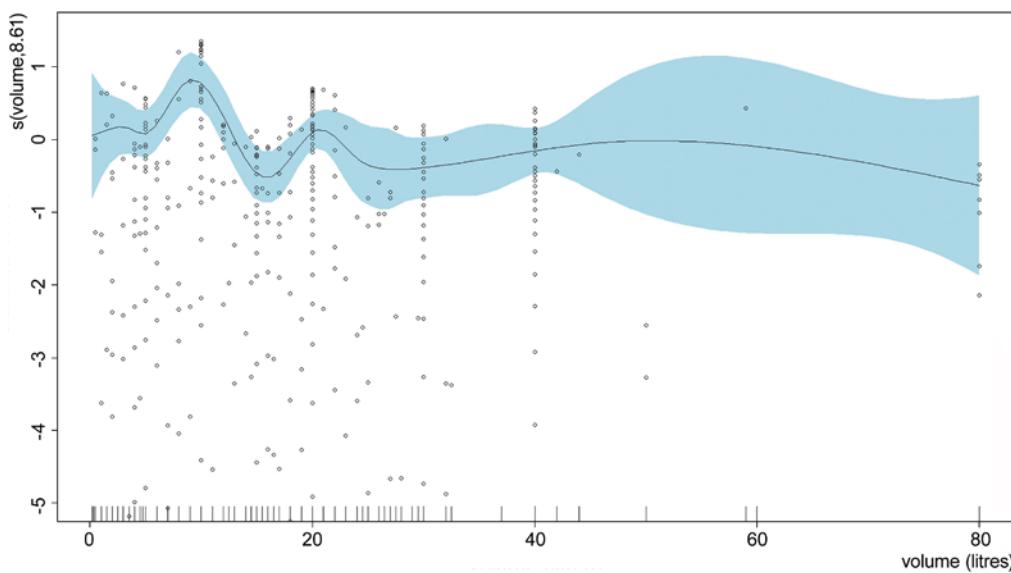
An even more pronounced effect of the numbers of the samples on the number of the species can be seen in **Model 7** (Graph 8), in which the function (curve) reaches a peak at around 20 samples from a single feature, after which the result is again influenced by outliers. The numbers of the samples explain twice as much of the variability (21.1%) as the volume itself and also according to the values in Table 2 (a lower AIC value than Model 6) it is a better-fitted model. A larger amount of samples taken from individual features makes for an unequivocally greater contribution towards obtaining a richer composition of species from features than a larger quantity of sediment taken from a feature in a single sample.

Model 8 (Graph 9) presents the interaction between the volume and the number of collected samples, which is not statistically significant (Tab. 2, m8) since the principal effect is again represented by the number of samples (edf 3.03) to the detriment of the volume (edf 1).¹⁰ This can also be seen in the percentage of explained variability, which is almost the same as in Model 7.

Model	Response	Function	k	EDF	k-index	p-value	R ² adj	Dev. expl. (%)	REML	AIC	Dispersion	gamma	Matrix
m1	Taxa	s(volume)	15	1.17	0.39	0.322	0.000645	0.24	1260.9	3524.66	1.14	1.4	Taxa
m2	Taxa	s(PMR)	15	12.55	0.75	2.00E-16	0.641	76.1	972.58	2618.91	0.64	1.4	
m3	Taxa	s(density)	15	9.9	0.73	2.00E-16	0.527	64.5	1049.8	2842.44	0.78	1.4	
m4	Taxa	ti(volume; PMR)	15; 15	22.7	0.82	2.00E-16	0.794	81.1	1315.4	2533.83	0.62	1	
m5	PMR	s(volume)	15	8.61	0.29	9.80E-06	0.013	6.51	1948.8	5419.53	4.02	1.4	
m6	Taxa	s(volume)	15	2.7	0.85	0.000194	0.157	10.3	476.51	1328.03	1.04	1.4	Features/ contexts
m7	Taxa	s(sample_count)	15	3.09	0.61	2.00E-16	0.362	21.1	466.52	1299.02	0.98	1.4	
m8	Taxa	ti(volume; sample_count)	10; 10	0.0005	0.85	0.86	0.364	22.6	646.98	1296.52	0.99	1	
m9	PMR	s(volume)	15	1	0.54	0.288	-0.00285	0.48	779.27	2185.64	4.9	1.4	
m10	PMR	s(sample_count)	15	2.39	0.41	0.03	-0.00683	4.13	777.29	2178.25	5.3	1.4	
m11	PMR	ti(volume; sample_count)	10; 10	8.55	0.63	1.28E-05	-0.0219	23.9	1075.3	2140.93	3	1	
m12	Taxa	s(volume)	10	1	1.24	0.93	0.777	55.9	174.06	353.51	1.04	1	Sites
		s(sample_count)	10	2.67	0.84	0.003							
		ti(volume; sample_count)	5; 5	2.8	0.79	0.55							
m13	PMR	s(volume)	10	1	0.78	0.98	-0.118	10.1	325.39	668.57	1.8	1	
		s(sample_count)	10	1.16	0.61	0.09							
		ti(volume; sample_count)	5; 5	1.63	0.57	0.48							

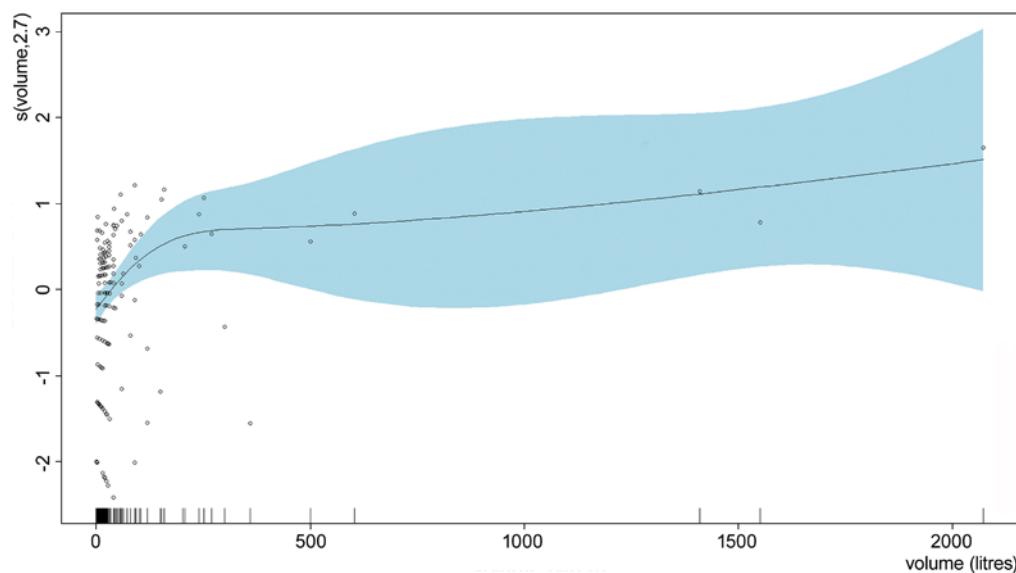
Tab. 2. Results of individual models. Response: explained variable; Function: smooth function of studied variable; k: node values for basis function; EDF: effective degrees of freedom; R² (adj): coefficient of determination; Dev. Expl (%): null deviance (percentage of explained variability); p-value: value of probability; REML: values of marginal likelihood maximisation method; AIC: Akaike information criterion; Dispersion: value of model dispersion; gamma: increases or decreases the degree of the smoothing function.

Tab. 2. Výsledky jednotlivých modelů. Response: Vysvětlovaná proměnná; Function – smooth funkce zkoumané proměnné; k – uzlové hodnoty pro základní funkci (basis function); EDF – efektivní stupně volnosti; R² (adj) – koeficient determinace; Dev. Expl (%) – nulová deviance (procento vysvětlené variabilita); p-value – hodnota pravděpodobnosti; REML – hodnoty tzv. marginal likelihood maximization metody; AIC – Akaikeovo informační kritérium; Dispersion – hodnota disperze modelu; gamma – zvyšuje nebo snižuje míru vyhlazení smooth funkce.



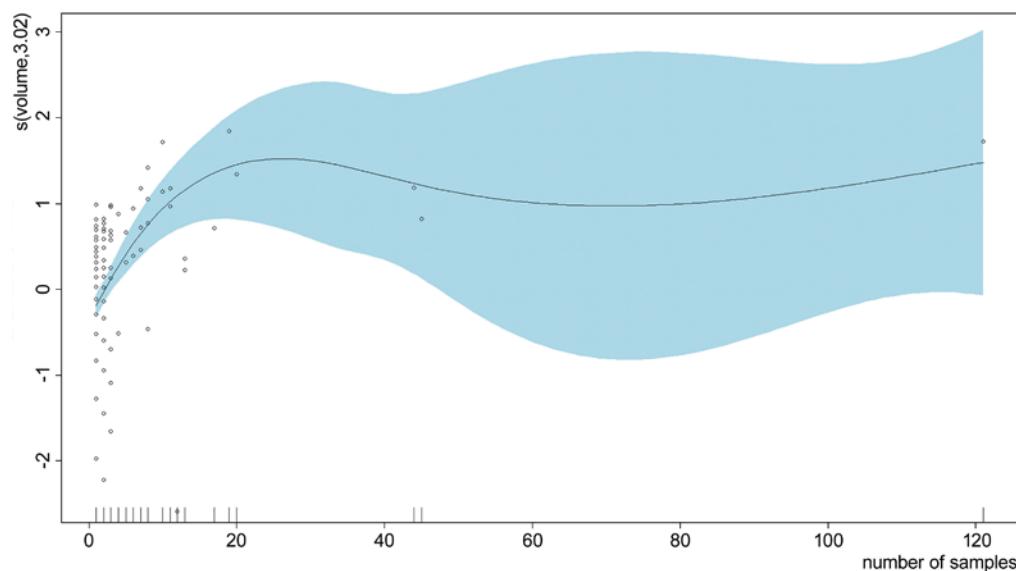
Graph 6. Model 5 displaying the relationship between volume and PMR in samples. X axis – volume in litres, Y axis – smooth function of volume with edf value (effective degrees of freedom).

Graf 6. Model 5 zobrazující vztah objemu k RMZ ve vzorcích. Osa X – objem v litrech (volume); osa Y – smooth funkce objemu (volume) s hodnotu edf (effective degrees of freedom).



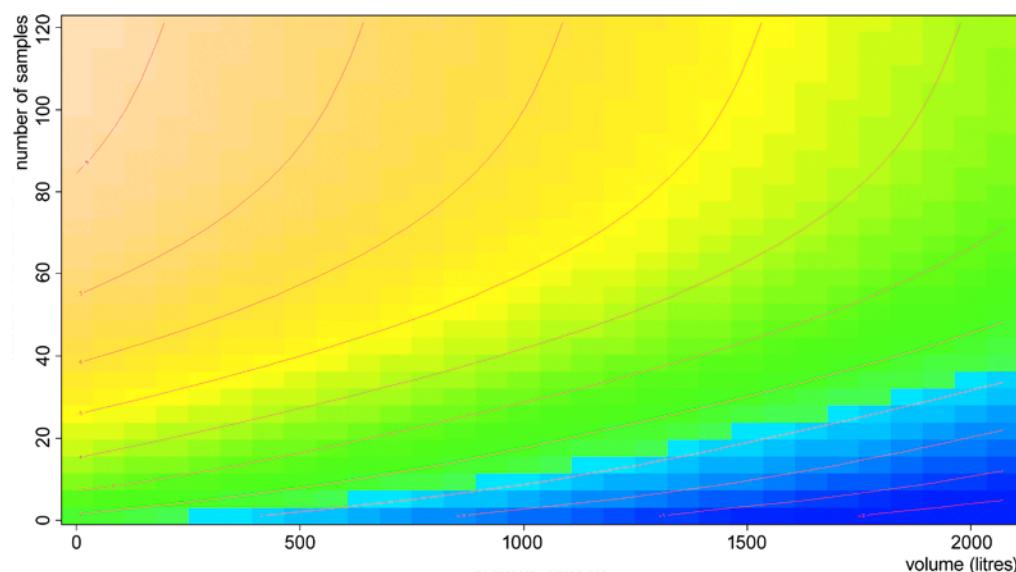
Graph 7. Model 6 displaying the relationship between volume and number of botanical species in archaeological features.
X axis – volume in litres,
Y axis – smooth function of volume with edf value (effective degrees of freedom).

Graf 7. Model 6 zobrazující vztah objemu k počtu botanických druhů v archeologických objektech.
Osa X – objem v litrech (volume);
osa Y – smooth funkce objemu (volume) s hodnotou edf (effective degrees of freedom).



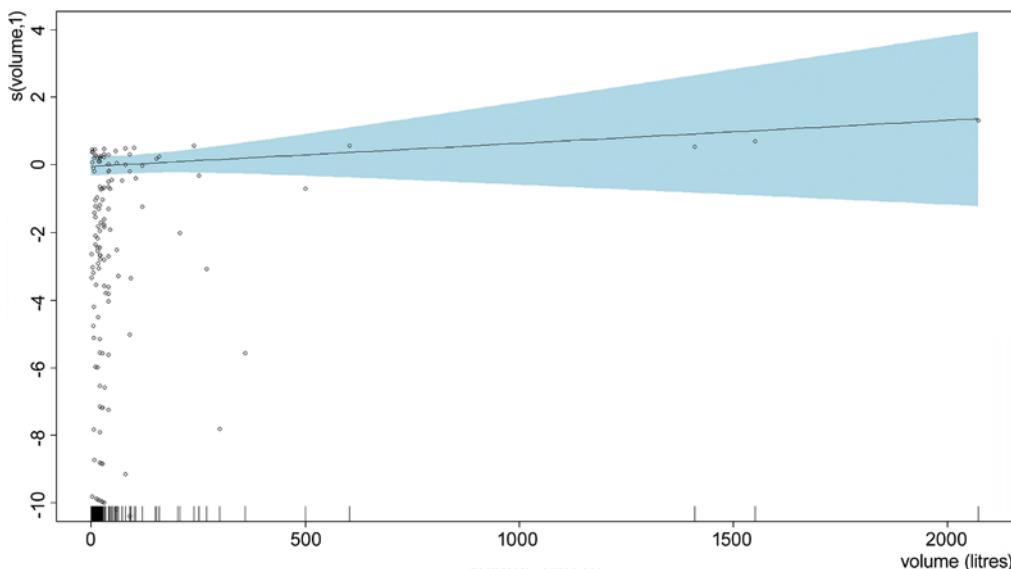
Graph 8. Model 7 displaying the relationship between the number of samples and botanical species in archaeological features.
X axis – number of samples,
Y axis – smooth function of number of samples with edf value (effective degrees of freedom).

Graf 8. Model 7 zobrazující vztah počtu vzorků k botanickým druhům v archeologických objektech.
Osa X – počet vzorků (number of samples); osa Y – smooth funkce počtu vzorků (samples) s hodnotou edf (effective degrees of freedom).



Graph 9. Model 8 displaying the interaction (effect) between volume and numbers of samples and botanical species in archaeological features.
X axis – volume in litres,
Y axis – number of samples. Yellow to white = strongest interaction, dark blue = weakest interaction.

Graf 9. Model 8 zobrazující interakci (efekt) objemu a počtu vzorků k botanickým druhům v archeologických objektech.
Osa X – objem v litrech (volume); osa Y – počet vzorků (number of samples). Žlutá až bílá = nejsilnější interakce; tmavě modrá = nejslabší interakce.



Graph 10. Model 9 displaying the relationship between volume and PMR in archaeological features. X axis – volume in litres, Y axis – smooth function of volume with edf value (effective degrees of freedom).

Graf 10. Model 9 zobrazující vztah objemu k RMZ v archeologických objektech. Osa X – objem v litrech (volume); osa Y – smooth funkce objemu (volume) s hodnotou edf (effective degrees of freedom).

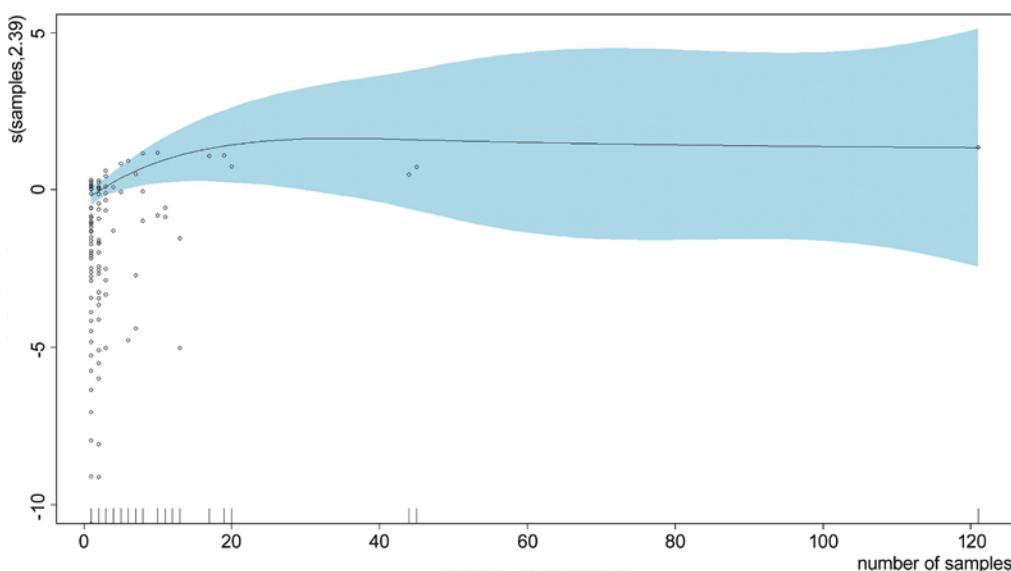
This is also clear from Graph 9, where the number of species is more strongly inclined to larger numbers of collected samples.

Model 9 depicts the relationship between PMR and volume. According to the result (Tab. 2, m9) and Graph 10, the volume itself does not affect the quantity of macro-remains obtained and is not statistically significant. This is the model with the worst result from the viewpoint of the set of features (cf. AIC values). This is in agreement with Model 1, in which the effect of the volume may have been pronounced but was influenced to a large degree by outlier values and was also the weakest of all the models (1 to 5). The fact that the volumes have weak evidentiary value in explaining the amount of macro-remains is rather interesting.

Model 10 depicts PMR explained by the numbers of samples. As can be seen on the left-hand side of Graph 11 (up to a value of approximately 20 samples), we can observe a slight effect, although on the boundary of significance, and the results are marked by overdispersion as in the case of Model 9 (Tab. 2, m9, m10). Regardless, once again there is a very poor contribution of the combined volume (per archaeological feature) to the overall explanation of PMR.

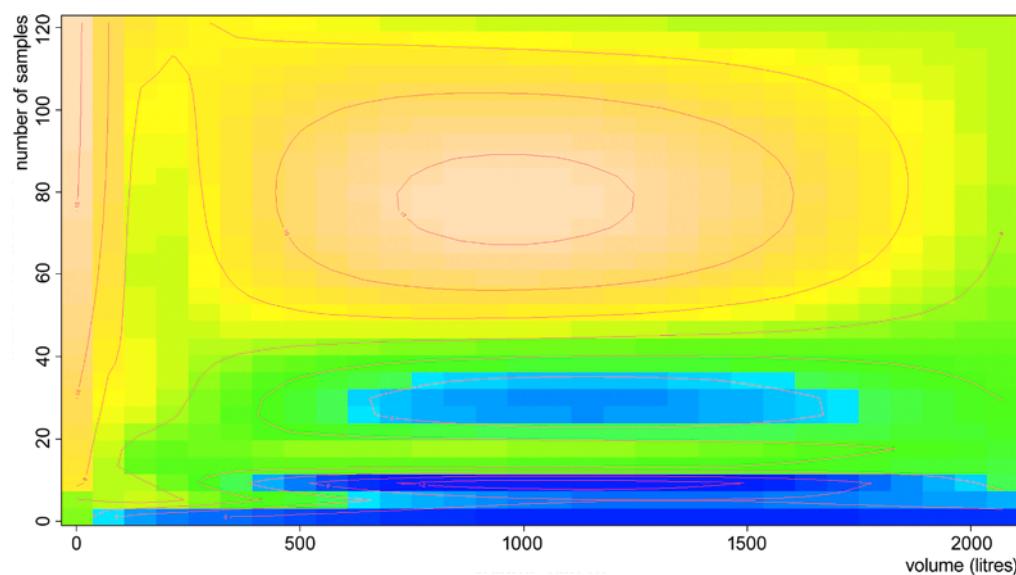
The situation is rather more interesting in **Model 11** (Graph 12), which depicts the interaction between PMR and the numbers of samples. In contrast to Models 9 and 10, a pronounced effect (Tab. 2, m11) and a high degree of statistical significance can be observed. This is the most informative model from the viewpoint of PMR with the highest percentage of explained variability. Two principal effects are evident in Graph 12. The first, on the left-hand side, points to the close relationship between PMR and a higher number of samples and smaller volumes, while the second is associated with a higher volume (ca between 300 and 1,800 litres) and larger numbers of samples (ca between 40 and 110 samples). Model 11 reflects, first and foremost, the method of feature sampling in the studied set (see the chapter Introduction). It also points to the fact that it is necessary to choose a balance between a sufficient number of samples taken and their volume to capture a representative number of macro-remains from a sampled situation.

The relationship between botanical species (Graph 13) and PMR (Graph 14) with the volume and the number of samples in a modified $m \times n$ matrix, where m stands for archaeological sites



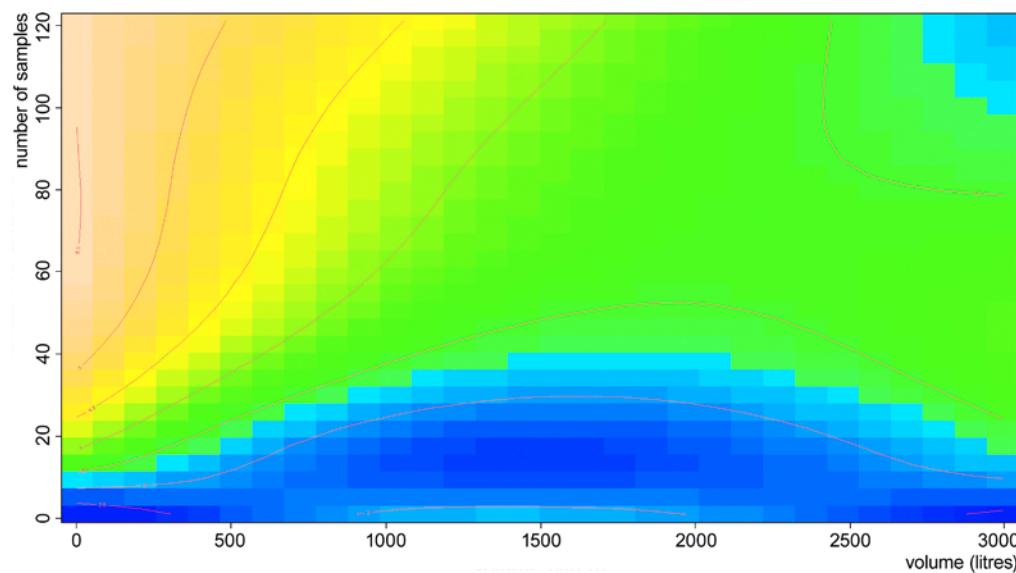
Graph 11. Model 10 displaying the relationship between the number of samples and PMR in archaeological features. X axis – počet vzorků (number of samples); Y axis – smooth funkce počtu vzorků (samples) s hodnotou edf (effective degrees of freedom).

Graf 11. Model 10 zobrazující vztah počtu vzorků k RMZ v archeologických objektech. Osa X – počet vzorků (number of samples); osa Y – smooth funkce počtu vzorků (samples) s hodnotou edf (effective degrees of freedom).



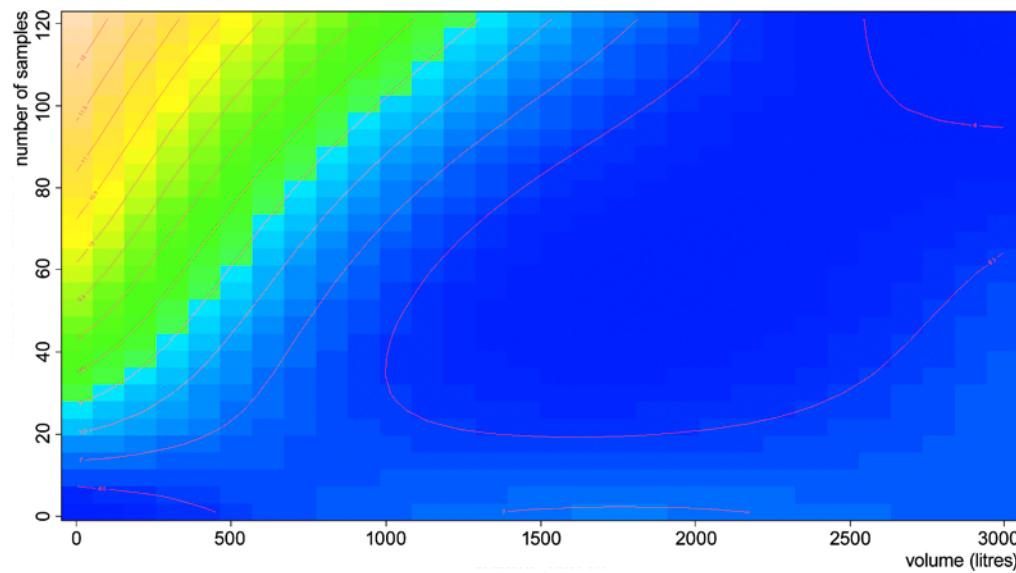
Graph 12. Model 11 displaying the interaction (effect) between volume and numbers of samples and PMR in archaeological features.
X axis – volume in litres,
Y axis – number of samples. Yellow to white = strongest interaction, dark blue = weakest interaction.

Graf 12. Model 11 zobrazující interakci (efekt) objemu a počtu vzorků k RMZ v archeologických objektech.
Osa X – objem v litrech (volume);
osa Y – počet vzorků (number of samples). Žlutá až bílá = nejsilnější interakce; tmavě modrá = nejslabší interakce.



Graph 13. Model 12 displaying the interaction (effect) between volume and numbers of samples and botanical species at archaeological sites.
X axis – volume in litres,
Y axis – number of samples. Yellow to white = strongest interaction, dark blue = weakest interaction.

Graf 13. Model 12 zobrazující interakci (efekt) objemu a počtu vzorků k botanickým druhům na archeologických lokalitách.
Osa X – objem v litrech (volume);
osa Y – počet vzorků (number of samples). Žlutá až bílá = nejsilnější interakce; tmavě modrá = nejslabší interakce.



Graph 14. Model 13 displaying the interaction (effect) between volume and numbers of samples and PMR at archaeological sites.
X axis – volume in litres,
Y axis – number of samples.
Yellow to white = strongest interaction, dark blue = weakest interaction.

Graf 14. Model 13 zobrazující interakci (efekt) objemu a počtu vzorků k RMZ na archeologických lokalitách.
Osa X – objem v litrech (volume);
osa Y – počet vzorků (number of samples). Žlutá až bílá = nejsilnější interakce; tmavě modrá = nejslabší interakce.

and n the variables botanical species, macro-remains (PMR), volumes taken within a site, and the overall number of samples collected from the site, was studied in the final two models – **Models 12 and 13**. The principal effects and their interactions were studied in the models. The graphs and the data in Table 2 (m12, m13) underline what has gradually taken shape in the previous models, i.e. that the number of samples taken again plays a determining role in how representative material can be obtained from a site. To summarise, it can be inferred from the presented models that the smaller the number of collected samples from the feature, the lower the certainty that the species composition would represent the whole feature.

Jevišovka

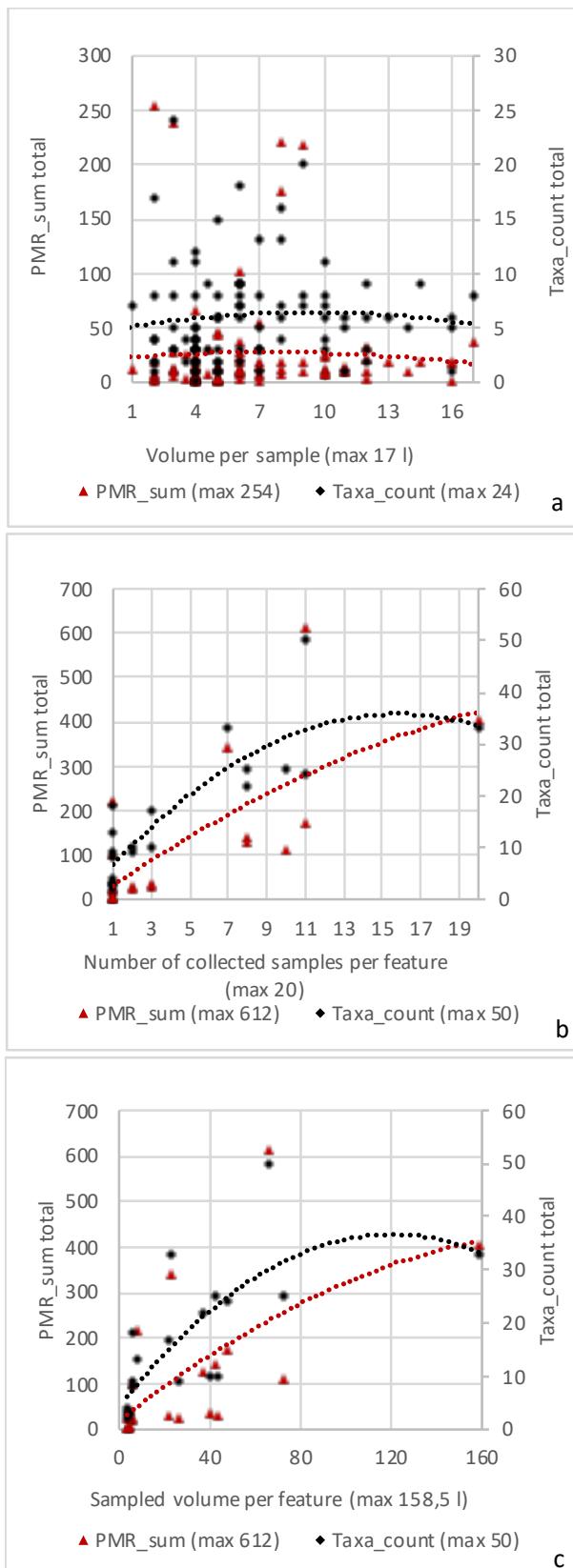
The archaeobotanical material obtained from the Jevišovka site was also used in all the above models and is presented here purely as an illustrative example of how the results from one site correspond with the results from the overall models.¹¹

The site Jevišovka (Břeclav District, Czech Republic) is a polycultural settlement uncovered during a rescue excavation and the archaeobotanical assemblage comes from the features dated to the Roman Period. A total of 94 archaeobotanical samples collected from 20 archaeological features were preliminarily analysed within the previously mentioned dissertation project (Hlavatá 2017, 176–182, Príloha). Sampled archaeological features included pit-houses, storage pits, unspecified settlement features and postholes. The information about the volume of samples was recorded; therefore, all the samples were used for statistical modelling. The overall volume of all collected samples was 625.75 litres, in which 2,432 macro-remains of 96 species were identified.

The archaeological information of individual samples consisted of an archaeological feature number, an uncovered arbitrary layer number, and separate information if it is a posthole or not, eventually with the sediment taken as a column sample (from the feature cross-section). Within the excavated pit-houses, 50% of the samples were collected from interior postholes. The remaining samples were taken from other interior pit-house layers and in one case from a grave deposited inside a pit-house. The proportion of pit-houses and other features at the Roman Period settlement was 1 to 3. However, three times more samples were collected from the pit-houses compared to the other archaeological features. There is a clearly visible partial subjectivity of sampling due to preferences (respectively the assumed potential) of the features to be sampled. Still, the archaeobotanical analysis did not show a generally higher abundance of macro-remains in individual samples taken from pit-houses than in those collected from other features. At the same time, it cannot be stated that the samples collected from interior postholes are more abundant than samples taken from other interior layers of the pit-houses.

Regarding the analysis of volume categories (see above Model 1), the samples from Jevišovka mainly belonged to category A (up to 9 litres, 78% of samples), and category B (10 to 29 litres, 21% of samples). The remaining 1% of samples fell into category C (30 to 40 litres, 1 sample). Category A volume mainly consisted of samples collected from pit-houses and postholes. In the case of postholes, the impossibility of taking bulky samples was expected to some extent (e.g. only 5 litres). Yet, this approach was also used in the case of some of the other archaeological interior layers of the pit-houses (e.g. 1 sample per 1 layer with a volume of around 5 litres).

Regarding the volume categories, the dataset from the Jevišovka site is displayed on the left-hand side of Graphs 2 to 6



Graf 15. Archaeological site Jevišovka. Scatterplots displaying the relationship of a) number of macro-remains (PMR), number of species (taxa) and volume per sample; b) number of macro-remains (PMR), number of species (taxa) and number of samples collected per feature samples; c) number of macro-remains (PMR), number of species (taxa) and sampled volume per feature. The polynomial trendline is applied in all cases.

Graf 15. Lokalita Jevišovka. Rozptylové diagramy zobrazující vztah: a) počtu makrobytků (PMR) a počtu druhů (taxa) s objemem jednotlivých vzorků; b) počtu makrobytků (PMR) a počtu druhů (taxa) s počtem odebraných vzorků z jednoho objektu; c) počtu makrobytků (PMR) a počtu druhů (taxa) s celkovým odebraným objemem z jednoho objektu. Spojnice trendů jsou ve všech případech polynomické.

(Models 1 to 5) in the area of lower sample volumes. From the viewpoint of individual Jevišovka samples (Graph 15a), as described for Models 1 to 5, there is also the readable volume itself, which does not have a significant effect on the composition of macro-remains and botanical species. In Models 6 to 11 (Graph 7–12), in which a relationship between the same variables was observed, but with respect to the features, the dataset from Jevišovka corresponds with those results. That is that the number of samples taken from archaeological features has a significant effect on the composition of macro-remains and botanical species (Graph 15b), i.e. with the higher number of collected samples, the number of macro-remains and botanical species has a tendency to increase. Also consistent with the results of Model 6 is that the overall volume of sediment taken per feature from Jevišovka affects the composition of the finds. Graph 15c displays a trend whereby increasing the volume, the number of determined macro-remains also increases. The interval of the sample volumes from Jevišovka varied from 0.5 to 32 litres, with the predominance of 10 litre samples. In addition, the volumes of the samples in the whole dataset are between 0.2 and 80 litres per sample.

Summary and discussion

The relationships between the volume, the number of botanical species, the PMR and the density from the viewpoint of the individual samples can be observed in **Models 1 to 5**. After investigating the relationship between the number of macro-remains (PMR) and density, we finally excluded the variable density from further analyses as in essence it copied the variable PMR and there may have been a risk that it would have a negative effect on the results of the models due to concurvity (Wood 2021c). The tensor product interaction function was used instead. The relationship between the volume of the sample and the number of species and macro-remains (PMR) implies that the volume has almost no influence on their quantity when each sample is studied separately. However, from the viewpoint of volume categories, small (up to 10 litres) and medium-large (up to 29 litres) samples are seen to be more significant from the statistical viewpoint and show a larger effect or provide a greater chance for a more representative number of findings to be obtained from them. Nevertheless, this result has to be seen in complexity. In other words, the most significant effect represented by small volumes (up to 10 litres, see Model 1) means, at the same time, the most significant influence of sampling on the composition of macro-remains obtained from the excavated situation. If the low-volume samples are combined with the low number of samples taken, there is a major risk that the overall interpretation will be negatively influenced by the sampling method itself. The small volume of the samples needs to be compensated with a higher number of samples collected from the same context (if there is a sufficient volume of sediment to be taken). Accordingly, this claim cannot be based merely on the amount of sediment taken within a sample and applied across the board to other assemblages from different preservation conditions. In this case, it illustrates just a small part of the issue of sampling and represents the first step towards further analyses and interpretations. It is important to realise that the volume samples up to 29 litres represented the largest proportion of the studied set (515 of 683 in which 157 samples had volumes of 0.2 to 9 litres). From this perspective, it can be stated that the effect of sampling was expressed in the results of the statistical analysis.

Models 6 to 11 shed further light on the influence of sampling on the resultant archaeobotanical assemblage. First and

foremost, incorporating a variable number of samples into the models provided new information about how the composition of botanical species can be influenced by larger numbers of samples and vice versa. A larger number of samples taken would appear to be the better alternative to gain a more representative archaeobotanical assemblage. As far as the amount of excavated sediment within the features is concerned, the results are in partial agreement with the corresponding Models 1 to 5, and the volume is not in itself a determining factor for the acquisition of a larger number of findings. An appropriately selected combination of a larger number of smaller samples (10 to 30 litres) could produce reliable results. The volume of the sample depends to a large extent on the goals of the excavations and also on the chronological dating of the excavated site.¹²

In any case, it can be assumed that rather than a one hundred-litre sample from one spot, it is more appropriate to take five twenty-litre samples covering a specified area within the excavated archaeological unit. Such samples can always be combined into a single element within any analysis (*sensu stricto* Lee 2012). Unfortunately, spatial information pertaining to the individual samples or archaeological units was inadequate in the investigated set, which was why it was not incorporated into the models. In particular, it was not possible to separate feature samples (e.g. a pit, without described layers) from context samples (e.g. a layer inside the feature) entirely reliably, since some features were represented by a single sample from one context, contexts (layers) were not stated at all for some features, and only contexts (layers) were stated for some samples. For these reasons, they were analysed together. There is a great probability that more detailed spatial information would help explain the high percentage of variability and reveal additional hidden structures within the data.

The models presented here were not always able to explain certain aspects of the data reliably, and this was frequently manifested in a low percentage of explained variability. To summarise, it can be said that they raised more questions than they answered although this need not be seen in an exclusively negative light.

The aim of this study was not to offer a manual on how to select the “correct” sampling method. Rather, it was intended to provoke a more active discussion about the influence of such a method and how it might be possible to increase the information potential of an investigated archaeological site (situation) to take in less conspicuous findings. It was also important to substantiate our assertions with data.

It is probable that the addition of further variables to our models, such as spatial (stratigraphic or coordinate) information, other types of findings extracted from samples, the definition of post-depositional processes, the type and properties of sediment collected, a more precise chronology, etc., would also affect the results and new statistically significant relationships and structures would appear. Therefore, what exactly is the answer to the tested question – how can the sample formal properties influence the analysis, respectively, the interpretation of archaeobotanical assemblage? In the case of the analysed dataset, it was the high heterogeneity and variable quality of archaeological information, in many cases even its absence, which meant the samples could not be sufficiently comparable. Thus, the only properties, which were suitable for the comparison of the samples were the species and macro-remains abundance and the sample volume, when recorded. Certainly, these are archaeobotanical samples, hence, the information of species and macro-remains abundance is, from an archaeobotanical point of view, essential. Yet, it cannot be the only information.

Regarding the statistical modelling results and the focus on the analysed dataset, the influence of chosen sampling method is markedly visible in the case of archaeological features with the minimum amount of samples collected (1 or 2 samples). We believe that an archaeobotanical interpretation of these features or contexts is uncertain and, as was shown above, the smaller number of samples collected can also cause the smaller number of macro-remains and botanical species found in the feature fill. In other words, there is still a chance that collecting other samples from the same feature can affect the archaeobotanical and probably the archaeological interpretation as well. As displayed in Graphs 8, 9 (Models 7, 8), the number of collected samples per feature has a noticeable statistical effect on the botanical species ubiquity and the number of macro-remains found. At the same time, there is a significant trend rising to the amount of 20 collected samples per feature; furthermore, this effect gradually disappears (Graph 8). Concerning the analysed dataset, only a few archaeological features were sampled more abundantly than 20 samples per feature. In view of this, there is no certainty as to whether the effect change represents the mentioned fact.

Nevertheless, as for the presented dataset, by collecting a minimum of 15–20 samples per feature, the degree of uncertainty whether the species and macro-remains composition are representative for the sampled feature decreases significantly. However, this result needs to be verified in the future by comparing it with the features from which larger quantities of archaeobotanical samples were taken (at least 30).

The archaeobotanical set analysed here was not ideal as far as its quality was concerned although it did present a “typical” data sample of the kind often encountered in archaeobotany. Therefore, it must be added that the results presented here cannot be extrapolated uncritically to all periods or sites because of the unique nature of archaeological sites. It is well documented that the numbers of preserved macro-remains and species in samples in a way reflect their handling before their deposition or the manner of their deposition since both the origin of the sampled population (when referring to “multiple event contexts”; cf. Hajnalová 2012, 33, 95; Kuna et al. 2013, 71–74; cf. Lityńska-Zajac, Wasylkova 2005, 160–162) and the preservation of remains¹³ have an effect on the resultant concentration of macro-remains. Of course, those events cannot be influenced, but what a researcher can influence, and what also greatly contributes to the quality of the evaluation and interpretation of archaeobotanical material, is how and to what extent this material will be obtained and documented. In our case, the samples taken, with a few exceptions,¹⁴ represented only a small part of the excavated features. If we exclude the postholes, then in general approximately 1% to 5% of the filling per feature was sampled. In such a case, it cannot be stated with certainty that the number of species obtained from the samples faithfully reflects the original population of macro-remains or even a function of a feature¹⁵ (especially when such samples were collected from only one specific spot within a context). To rely on taking only a very few samples from a feature (or layer) would mean to expect that every other sample taken from the same feature (layer) should consist of a very similar species composition, which, has not yet been proven in Jevišovka.

These are further reasons for this study being more of an exploratory nature than predictive or normative, and why it does not provide ready solutions, but it may serve as an aid in decision-making relating to the selected sampling strategy.

It should not be forgotten that thought must naturally be given to financial, logistical and human resources during sampling (and during many other excavation activities). Smaller samples (10 up to 30 litres), which are easier to handle in the

field and during flotation, sieving and the extraction of macro-remains, etc., would appear to be more appropriate, particularly from the logistical perspective. Not least, a larger number of smaller samples of a constant volume covering a larger area of the archaeological situation provides a far more representative set during analysis and evaluation. This is not merely from the archaeological or archaeobotanical perspective, but first and foremost from the statistical perspective, with the resultant analyses or models reflecting a more realistic situation at the site or within features (cf. Hlavatá, Apiar 2020). We believe that the influence of the chosen methodology on the results of (not only) archaeobotanical analysis can be suppressed only by consistent sampling. When it is not possible to take the same sample volumes, it is necessary to attempt to take at least approximately the same amount of sediment and samples from individual features (or layers within a feature) so these can be comparable. It is also important to choose the appropriate percentage of representativeness per one feature or situation and persist with it with the rest of the features. If the archaeological situation allows it (respectively depending on the nature of the research questions), it is necessary to attempt to take a constant number of samples from features or contexts across the entire research area. If it is not possible to sample all features in the site, then choose to sample the selected features proportionally and depending on the volume of their fill. If we decide to sample a group of the features defined in advance, according to the assumed function, e.g. all pit-houses, then the informative value of the assemblage obviously predominantly represents these features, and it cannot be directly related to the features of a different function at the same site. In general, to decide about a suitable sampling methodology followed by consultation with an archaeobotanist is recommended before (and during) an excavation.

Conclusion

This study aimed to answer the question as to whether, and to what extent, an archaeobotanical set may be influenced by the method of sampling and by the formal properties of the samples. The results of the study are based on a considerably heterogeneous archaeobotanical dataset in which the dispersion of sample volumes was from 0.2 to 80 litres and the dispersion of the findings of macro-remains from 0 to 8.076. The dataset was analysed and the relationships between volume (amount of sediment collected in litres), the samples count, and the number of identified botanical species and macro-remains were explained using generalised additive models (Faraway 2016; Wood 2017; Zuur et al. 2009). These relations were investigated at multiple levels – from the viewpoint of samples (Models 1 to 5), archaeological features (Models 6 to 11) and sites (Models 12 and 13).

A relatively convincing pattern gradually took shape at all levels, according to which the volume itself plays a marginal role in almost all cases, with large-volume samples not necessarily resulting in a larger potential for the acquisition of a richer composition of species or larger numbers of macro-remains. Collecting smaller samples is shown to be the more suitable alternative based on the model set. However, it is not possible to specify the ideal size of a “smaller” sample based on the investigated set. Nevertheless, volumes from 10 to 30 litres of sediment can be considered according to Models 1 to 5. The specific intentions of the excavation and the investigated archaeological situation must be taken into consideration. Although volume samples of up to 10 litres (157 samples) were shown to be significant from the viewpoint of the number of species, we assume that a large proportion of these were obtained subjectively (particularly those smaller than 5 litres), and/or it was objectively not possible

to take more sediment (e.g. small postholes). For this reason, their potential must be further investigated in the future. It is also necessary to realise that this estimate of the “ideal” volume could differ if, for example, the model set contained more samples with dispersion from 30 to 150 litres. However, it is not possible to interpret these results in such a way that taking several small samples instead completely replaces the same number of large-volume samples. As shown in Model 1, the small sample volumes need to be compensated with a higher number of collected samples per feature.

According to the models, the greatest influence on the number of botanical species and macro-remains obtained is undoubtedly by using a sampling strategy involving a larger number of samples taken from the investigated situation. The chance of excavating a more representative archaeobotanical assemblage is shown to increase with an increasing number of samples. This is in line with previous findings in the literature (e.g. Hajnalová 2012, 33; Hald 2008, 224; Jacomet et al. 1989, 70–84; Jacomet, Kreuz 1999, 95–112, Abb. 5.5.; Kreuz 2004, 111). When it is the other way round, i.e. the smaller the number of collected samples from the feature, then the lower the certainty that the species composition represents the whole feature, even if one sample did indeed represent the precise species composition (see Summary and discussion).

It is important to emphasise that each archaeological situation (feature, context, etc.) should have the same chance of being sampled, regardless of whether it shows a “potential” to be sampled. In this way, it is possible to improve the quality of the data.

A clear reason or purpose why and how the samples were taken was quite often missing. This “phenomenon” can be observed across many analysed excavations, from the older to the most recent ones. On the one side, there is a clear tendency that as a greater number of samples are taken from a feature, the more abundant and, in the first place, a more “reality-matching” assemblage can be expected. On the other side, there were a lot of features, from which were taken, at best, two samples, often without more detailed information. The overall degree of representativeness, as well as the credibility of the acquired knowledge, then fluctuates. Furthermore, if subjective sampling is added to this, one needs to consider to what extent to rely on interpretations, for example, of the subsistence strategies or the ecological and environmental models based on such data. The presented analysis as well as the results of the models showed that there is still much unexplained variability in our data, most probably caused by a lack of information while many structures and patterns remain hidden within the data. Yet, many of the mentioned issues can be prevented, or at least minimised, if the sampling strategy/method is consulted with an expert in advance and during excavation. It is a pity when an archaeologist puts the effort into the sampling but the evaluation of the archaeobotanical assemblage fails due to an absence of information.

Notes

- 1 This material comes from the analyses produced by the author, from the expert reports of other authors and from the literature. For the dissertation project analysis, the original unpublished data with the permission of the following authors was used: R. Daňo, K. Elschek, G. Fusek, E. Hajnalová, M. Hajnalová, J. Hečková, D. Hulínek, L. Chmelo, P. Ježínek, D. Kendrala, P. Kočár, R. Kočárová, T. Kolník, B. Komoróczy, D. Krčová, K. Kuzmová, M. Lamiová-Schmidlová, J. Mihályiová, E. Mirošayová, V. Plachá, J. Rajtár, M. Ruttkay, J. Schmidlová, B. Šebesta, O. Šedo, M. Šútor, P. Valent, V. Varsik, M. Vrablec, F. Žák Matyasowszky. We would like

to thank them all, including on this occasion. Data obtained from archaeobotanical expert reports are used in this study to the extent necessary, i.e. only in the form of aggregate numbers without specific findings being stated, without original specimen numbers or the circumstances pertaining to their finding. This for the purpose of the copyright protection of the individual experts (Hlavatá 2017, 12–13, 63–64, Katalóg). The dissertation was approved by the dissertation committee via the oral defense at the Constantine the Philosopher University in Nitra, in 2017.

- 2 For example, when the approximate volume of a context is calculated.
- 3 Only three samples of mass findings that showed extreme outlying values (more than 4,000 macro-remains in a sample of small volume) were excluded from the analysis.
- 4 The trendline is represented by the LOESS curve (see, e.g. Zuur et al. 2009, 39–42; Lepš, Šmilauer 2003, 123–124).
- 5 Highly similar results are also obtained with the use of Kendall's Tau coefficient.
- 6 With the use of functions from the mgcv package (Wood 2017). For details on GAM viz Faraway 2016; Wood 2017; Zuur et al. 2009. A great deal of useful information and examples can also be found on the internet, e.g. Wood 2021a.
- 7 The value of edf (effective degrees of freedom) represents the size of the effect or (non-)linearity, where the value 1 is equivalent to a linear relationship. The higher the value, the more non-linear the relationship (Zuur et al. 2009, 52–53).
- 8 This involves the function “tensor product smooth interaction” (cf. Wood 2017, 187, 326; 2021b).
- 9 Using the function *nbinomial* in the *msme* package (Hilbe, Robinson 2018).
- 10 Not included in Table 2.
- 11 More detailed results will be published as a case study.
- 12 Generally, for prehistoric sites, especially for those of the Neolithic Age, the volume of 100 litres of sediment, at least, should be taken per context or feature. That is also due to the usual preservation manner of carbonised macro-remains in dry or mineral saturated conditions of the temperate climate zone (cf. Jacomet, Kreuz 1999, 102–107; Kühn, Antolín 2016, 20, 40, and further literature).
- 13 Of the order of tens to hundreds, fewer remains are preserved in a dry and mineralised environment than in a waterlogged environment – see, for example, Antolín et al. 2015; Kühn, Antolín 2016, 20, 40; Jacomet et al. 1989, 70–84.
- 14 For instance, at the settlement site in Velký Meder (Slovakia), the two Germanic pit-houses were sampled systematically at intervals (Hajnalová, Varsik 2010, 192). However, this is only one of the few cases where several dozen samples came from the features of the analysed dataset and, at the same time, in large volumes.
- 15 This does not refer to the contexts originated from e.g. accidental burn of the grain pit etc. (cf. van der Veen, Jones 2006).

Acknowledgement

This article was supported by institutional support RVO: 68081758 of The Czech Academy of Sciences, Institute of Archaeology, Brno.

References

- Antolín, F., Steiner, B., Vach, W., Jacomet, S. 2015: What is a litre of sediment? Testing volume measurement techniques for wet sediment and their implications in archaeobotanical analyses at the Late Neolithic lake-dwelling site of Parkhaus Opéra (Zürich, Switzerland). *Journal of Archaeological Science* 61, 36–44.

- Apiar, J. 2021:** *d754bb5* [dataset]. GitHub, Mar 10.2021 [cit. 2021-03-10]. Available from: <https://github.com/JanaApiar/DataSet>.
- Barker, G. 1975:** „To sieve or not to sieve“. *Antiquity* 49(193), 61–63.
- Baxter, M. 2015:** *Notes on Quantitative Archaeology and R* [online]. Unpublished study. ResearchGate ©2008–2021. May 2015. [cit. 2021-03-07]. Available from: https://www.researchgate.net/publication/277931925_Notes_on_Quantitative_Archaeology_and_R.
- Bogaard, A. 2004:** *Neolithic Farming in Central Europe. An archaeobotanical study of crop husbandry practices*. Oxon: Routledge.
- Faraway, J. J. 2016:** *Extending the Linear Model with R. Generalized Linear, Mixed Effects and Nonparametric Regression Models*. Second edition. Texts in Statistical Science. Boca Raton: Chapman and Hall, CRC Press.
- Field, A., Miles, J., Field, Z. 2012:** *Discovering statistics using R*. Los Angeles: SAGE Publications.
- Hajnalová, E. 1993:** *Obilie v archeobotanických nálezoch na Slovensku*. Nitra: Archeologický ústav Slovenskej akadémie vied.
- Hajnalová, E. 2001:** *Ovocie a ovocinársvo v archeobotanických nálezoch na Slovensku*. Nitra: Archeologický ústav Slovenskej akadémie vied.
- Hajnalová, M. 2012:** *Archeobotanika doby bronzovej na Slovensku. Štúdie ku klíme, prírodnému prostrediu, poľnohospodárstvu a paleoekonómii*. Nitra: Univerzita Konštantína Filozofa, Filozofická fakulta.
- Hajnalová, M., Varsik, V. 2010:** Kvádske roľníctvo na Slovensku z pohľadu archeológie a archeobotaniky. In: J. Beljak, G. Březinová, V. Varsik (eds.): *Archeológia barbarov* 2009. Hospodárstvo Germánov. Sídliskové a ekonomicke štruktúry od neskorej doby laténskej po včasný stredovek. *Archaeologica Slovaca Monographiae. Communicationes*, Tomus X. Nitra: Archeologický ústav SAV Nitra, 181–224.
- Hald, M. M. 2008:** The use of archaeobotanical assemblages in palaeoeconomical reconstructions. In: N. Marchetti, I. Thuesen (eds.): *ARCHAIA. Case studies on research planning, Characterisation, Conservation and Management of Archaeological Sites*. BAR International Series 1877. Oxford: Archaeopress, 223–229.
- Hilbe, J. M. 2011:** *Negative Binomial Regression*. Second edition. Cambridge: Cambridge University Press.
- Hilbe, J. M. 2014:** *Modeling count data*. First edition. Cambridge: Cambridge University Press.
- Hilbe, J., Robinson, A. 2018:** *msme: Functions and Datasets for “Methods of Statistical Model Estimation”*. R package version 0.5.3. 2018-03-18. [cit. 2021-02-07]. Available from: <https://CRAN.R-project.org/package=msme>.
- Hillman, G. 1984:** Interpretation of archaeological plant remains: The application of ethnographic models from Turkey. In: W. van Zeist, W. A. Casparie (eds.): *Plants and Ancient Man*. Rotterdam: A. A. Balkema, 1–41.
- Hlavatá, J. 2017:** *Rastlinné potraviny, ich producia, úprava a skladovanie v germánskom a rímsko-provinciálnom prostredí na území stredného Dunaja* [online]. Unpublished doctoral dissertation. Univerzita Konštantína Filozofa. Filozofická fakulta. Katedra archeológie. Stored in: Centrálny register závarečných a kvalifikačných prác. [cit. 2021-02-28]. Available from: <https://opac.crzp.sk/?fn=detailBiblioForm&csid=7521CD21230D5DBE37BF890A9CE&seo=CRZP-detail-kniha>.
- Hlavatá, J., Apiar, P. 2020:** Archäobotanischer Bericht. Ergebnisse der Schnellanalyse im Auftrag der Pro Arch Prospektion und Archäologie GmbH. In: S. Hornung, J. Gilhaus, B. Glunz-Hüsken: *Rituell oder profan? Ein bronzezeitlicher Fundplatz in der bayerischen Donau-Aue. Berichte über die archäologischen Untersuchungen im Gesamtprojekt der Gas-loopleitung von Forchheim nach Finsing, Trassenabschnitt 26 bei Gaden, Gde. Pförring, Lkr. Eichstätt*. Archäologische Quellen, Band 4. Stuttgart: DGUF-Verlag, 85–108.
- Hlavatá, J., Varsik, V. 2019:** Sídisko autochtonnej panónskej populácie v Rusovciach: prvé výsledky archeobotanických analýz. In: N. Beljak Pažinová, D. Repka (red.): *Sedem kruhov Jozefa Bujnu*. Studia Historica Nitriensia. Supplementum 2. Nitra: Univerzita Konštantína Filozofa v Nitre, 427–448.
- Jacomet, S., Brombacher, Ch., Dick, M. 1989:** *Archäobotanik am Zürichsee. Ackerbau, Sammelwirtschaft und Umwelt von neolithischen und bronzezeitlichen Seeffersiedlungen im Raum Zürich*. Ergebnisse von Untersuchungen pflanzlicher Makreste der Jahre 1979–1988. Zürich: Orell Füssli Verlag.
- Jacomet, S., Kreuz, A. 1999:** *Archäobotanik. Aufgaben, Methoden und Ergebnisse vegetations- und agrargeschichtlicher Forschung*. Stuttgart: Verlag Eugen Ulmer GmbH & Co.
- Jones, G. E. M. 1984:** Interpretation of archaeological plant remains: Ethnographic models from Greece. In: W. van Zeist, W. A. Casparie (eds.): *Plants and Ancient Man*. Rotterdam: A. A. Balkema, 43–61.
- Jones, G. E. M. 1991:** Numerical analysis in archaeobotany. In: W. van Zeist, K. Wasylkowa, K.-E. Behre (eds.): *Progress in Old World Palaeoethnobotany. A retrospective view on the occasion of 20 years of the International Work Group for Palaeoethnobotany*. Rotterdam: A. A. Balkema, 63–80.
- Jones, M. K. 1991:** Sampling in palaeoethnobotany. In: W. van Zeist, K. Wasylkowa, K.-E. Behre (eds.): *Progress in Old World Palaeoethnobotany. A retrospective view on the occasion of 20 years of the International Work Group for Palaeoethnobotany*. Rotterdam: A. A. Balkema, 53–62.
- Klečka, A. 1934:** Rostlinná produkce v našem pravěkém zemědělství. *Věstník československého zemědělského muzea* 3, 97–98.
- Kočár, P., Dreslerová, D. 2010:** Archeobotanické nálezy pěstovaných rostlin v pravěku České republiky. *Památky archeologické* CI, 203–242.
- Kreuz, A. 2004:** Landwirtschaft im Umbruch? Archäobotanische Untersuchungen zu den Jahrhunderten um Christi Geburt in Hessen und Meinfranken. *Berichte der Römisch-Germanischen Kommission* 85, 97–292.
- Kühn, M., Antolín, F. 2016:** *Einführung in die Archäobotanik* [online]. Unpublished study material. Uloženo: Institut für Prähistorische und Naturwissenschaftliche Archäologie. Universität Basel. [cit. 2021-02-28]. Available from: https://duw.unibas.ch/fileadmin/user_upload/duw/IPNA/PDF_s/PDF_s_in_use/Einfuehrung_Archaeobotanik2016.pdf.
- Kuna, M., Hajnalová, M., Kovačíková, L., Lisá, L., Novák, J., Bureš, M., Cílek, V., Hošek, J., Kočár, P., Majer, A., Makowiecki, D., Scott Cummings, L., Sůrová, Z., Světlík, I., Vandenberghe, D., Van Nieuland, J., Yost, Ch. L., Zabilska-Kunek, M. 2013:** Raně středověký areál v Roztokách z pohledu ekofaktů. *Památky archeologické* CIV, 59–147.
- Látková, M. 2017:** *The Archaeobotany of Mikulčice. Food Supply to the Early Medieval Stronghold*. Studien zum Burgwall von Mikulčice XI. Spisy archeologického ústavu AV ČR Brno 55. Brno: Institute of Archaeology of the Academy of the Czech Republic, Brno.
- Lee, G. 2012:** Taphonomy and sample size estimation in paleoethnobotany. *Journal of Archaeological Science* 39, 648–655.
- Leplš, J., Šmilauer, P. 2003:** *Multivariate analysis of ecological data using CANOCO TM*. Cambridge: Cambridge University Press.
- Lityńska-Zajęc, M., Wasylkowa, K. 2005:** *Przewodnik do badań archeobotanicznych*. Poznań: Sorus.
- Martin, T. G., Wintle, B. A., Rhodes, J. R., Kuhnert, P. M., Field, S. A., Low-Choy, S. J., Tyre, A. J., Possingham, H. P. 2005:** Zero tolerance ecology: improving ecological inference by modelling the source of zero observations. *Ecology Letters* 8, 1235–1246.
- O’Hara, R. B., Kotze, D. J. 2010:** Do not log-transform count data. *Methods in Ecology and Evolution* 1, 118–122.

- Orton, C.** 1980: *Mathematics in Archaeology*. Cambridge: Cambridge University Press.
- Orton, C.** 2000: *Sampling in Archaeology*. Cambridge: Cambridge University Press.
- Pearsall, D. M.** 2000: *Palaeoethnobotany: a Handbook of Procedures*. San Diego: Academic Press.
- Reed, K.** 2016: Archaeobotanical analysis of Bronze Age Feudvar. In: H. Kroll, K. Reed: *Die Archäobotanik. Feudvar III*. Würzburg: Julius-Maximilians Universität Würzburg, 197–292.
- Renfrew, C., Bahn, P.** 2008: *Archaeology. Theories, Methods and Practice*. Fifth edition. London: Thames & Hudson, Ltd.
- RStudio Team** 2020: *RStudio: Integrated Development Environment for R*. Boston, MA: RStudio, PBC [online]. © 2021. [cit. 2021-03-07]. Available from: <http://www.rstudio.com/>.
- Shennan, S.** 2006: *Quantifying Archaeology*. Second edition. Reprinted 2006. Edinburgh: Edinburgh University Press.
- Smith, A.** 2014: The Use of Multivariate Statistics within Archaeobotany. In: J. M. Marston, J d'Alpoim Guedes, C. Warinner (eds.): *Method and Theory in Palaeoethnobotany*. Boulder: University Press of Colorado.
- van der Veen, M.** 1982: Sampling seeds. *Journal of Archaeological Science* 9, 287–298.
- van der Veen, M.** 1984: Sampling for Seeds. In: W. van Zeist, W. A. Casparie (eds.): *Plants and Ancient Man*. Rotterdam: A. A. Balkema, 193–199.
- van der Veen, M.** 1992: *Crop Husbandry Regimes. An Archaeobotanical Study of Farming in northern England 1000 BC – AD 500*. Sheffield Archaeological Monographs 3. Sheffield: Sheffield Academic Press.
- van der Veen, M., Jones, G. E. M.** 2006: A re-analysis of agricultural production and consumption: implications for understanding the British Iron Age. *Vegetation History and Archaeobotany* 15, 217–228.
- Wallace, M., Jones, G., Charles, M., Forster, E., Stillman, E., Bonhomme, V., Livarda, A., Osborne, C.P., Rees, M., Frenck, G., Preece, C.** 2019: Re-analysis of archaeobotanical remains from pre- and early agricultural sites provides no evidence for a narrowing of the wild plant food spectrum during the origins of agriculture in southwest Asia. *Vegetation, History and Archaeobotany* 28, 449–463.
- Wood, S. N.** 2017: *Generalized Additive Models. An Introduction with R*. Second Edition. Boca Raton: Chapman and Hall, CRC Press.
- Wood 2021a:** *mgcv: Mixed GAM Computation Vehicle with Automatic Smoothness Estimation*. Version 1.8-36. RDRR, June 2021. [cit. 2021-03-07]. Available from: <https://rdrr.io/cran/mgcv/>.
- Wood 2021b:** *gam: Generalized additive models with integrated smoothness...* Version 1.8-35. RDRR, June 2021. [cit. 2021-03-07]. Available from: <https://rdrr.io/cran/mgcv/man/gam.html>.
- Wood 2021c:** *concurvity: GAM concurvity measures*. Version 1.8-34. RDocumentation. June 2021. [cit. 2021-03-07]. Available from: <https://www.rdocumentation.org/packages/mgcv/versions/1.8-34/topics/concurvity>.
- Zuur, A. F., Ieno, E. N., Walker, N. J., Savaliev, A. A., Smith, G. M.** 2009: *Mixed effects models and extensions in ecology with R*. New York: Springer-Verlag.

Resumé

Předložená studie vychází z výsledků disertačního projektu autorky (Hlavatá 2017), jehož předmětem bylo zpracování rozsáhlých archeobotanických souborů z doby římské. Hlavním záměrem zmíněného projektu byla rekonstrukce vybraných aspektů ekonomických poměrů obyvatelstva v daném období, která se opírala o vyhodnocení v té době všech dostupných archeobotanických dat pocházejících z různých chronologických a kulturních kontextů ve vymezeném regionu. Formální informace o vzorcích byly velmi různorodé a použití vzorků vzhledem

k zmíněnému cíli disertačního projektu bylo proto omezené. Tato skutečnost byla impulzem pro stanovení metodologických otázek, zda a do jaké míry ovlivňuje metodika odběru vzorků výpovědní hodnotu zkoumaného souboru dat a jaké z toho plynou omezení při zpracovávání takového souboru. Z publikované archeobotanické literatury je zřejmé, že na poměry a koncentrace nálezů ve vzorcích mají vliv především okolnosti jejich depozice v místě nalezení a procesy, kterými rostlinky a jejich plody procházejí během svého růstu, sklizně, skladování a úpravy před konzumací (např. Hillman 1984; Jones 1984; van der Veen 1992, 81–82). Pro potřeby této studie však bylo důležité otestovat také tvrzení, že pokud vzorkování a zaznamenané informace o vzorcích nejsou konzistentní, tak výsledky analýz mohou odrážet něco jiného, než jsou depoziční poměry a koncentrace makrozbytků. Tato omezení jsou přiblížena níže prostřednictvím výsledků statistické analýzy a statistického modelování.

Základní soubor

Zájmovým regionem studie je území severně od Dunaje omezené především na jihozápadní Slovensko, jižní Moravu a střední Čechy (obr. 1). Toto vymezení bylo dáno existencí dostupného archeobotanického materiálu, který tvořil výchozí databázi disertačního projektu (Hlavatá 2017, 19–21, 23–26). Materiál, který je předmětem této studie, pochází výhradně z databáze zmíněného projektu a je používán v souladu s ochranou autorství jednotlivých expertů (viz poznámka 1).

Z hlediska chronologie soubor náleží do období od zlomu letopočtu do konce 4. století, resp. začátku 5. století, což v chronologii daného regionu odpovídá době římské a počátku stěhování národů. Jelikož byly informace o podrobnější chronologii kontextů, ze kterých pochází jednotlivé archeobotanické vzorky, často nejednoznačné či vágní, byla bližší datace do jednotlivých stupňů vyloučena z analýz této studie (Hlavatá 2017, 21–22, 60).

Archeobotanickým vzorkem se v této studii rozumí vzorek sedimentu odebraný ze zkoumané archeologické situace za účelem proplavení a získání rostlinných makrozbytků pro archeobotanickou analýzu. Pod metodou vzorkování zde rozumíme nejen samotný odběr vzorku z terénu, ale také dokumentaci a flotaci.

Základní soubor tvořilo 1 000 archeobotanických vzorků, které byly odebrány z celkem 54 lokalit. Dohromady bylo identifikováno 70 032 rostlinných makrozbytků (bez zohlednění uhlíků), čítající 309 botanických druhů (Hlavatá 2017, 63–64, Príloha Mapa 1–3, Katalóg). Z celkové sumy vzorků byl pouze u 686 zaznamenán objem (celkem 13 757,8 litrů odebraného sedimentu). Případy bez záznamu objemu vzorku byly z analýz této studie vyloučeny a počet lokalit tak klesl na 25 (obr. 1). Rovněž i množství nalezených makrozbytků klesl na 36 460 a 238 botanických druhů. Celkový rozptyl objemu vzorků ve zkoumaném souboru se vešel do intervalu od 0,2 do 80 litrů a rozptyl nálezů makrozbytků do rozmezí od 0 do 8 076 ks (tab. 1).

Archeologické lokality nejsou jmenovitě uváděny a specifikovány z důvodu autorské etiky (viz výše), s výjimkou lokality Jevišovka, která byla pod vedením B. Komoróczyho zkoumána pracovištěm autorů a která je zde použita jako příkladová. Celý soubor je vnímán jako modelový, tudiž z hlediska jeho analýzy nejsou důležité konkrétní lokalizace, ale vlastnosti vzorků z nich pocházejících.

Archeobotanické vzorky

Zejména při výzkumech staršího data byl sediment z uloženin často odebíráno subjektivně selektivním způsobem, často bez předem stanoveného teoretického či metodického konceptu (srov. Hajnalová, Varsík 2010; Hlavatá 2017, Katalóg; Hlavatá, Varsík 2019; Kočář, Dreslerová 2010). Tato skutečnost se projevila i ve

zkoumaném souboru. Vstupní informace o původu vzorků byly tudíž značně heterogenní a mnohdy i výslovně torzovité.

V případě některých lokalit, ze kterých byly vzorky získány v neproplaveném/nezpracovaném stavu (Hlavatá 2017, 30), byly jedinou informací o způsobu odběru popisky ze štítků, které obsahovaly většinou číslo objektu, někdy údaje o vrstvě, jindy jen pořadové číslo odebraného vzorku; u většiny proplavených vzorků nebyly dostupné ani údaje o typu sedimentu. Takový způsob popisu, kdy s archeobotanickým materiélem není dodán podrobný příručkový katalog vzorků, ztěžuje, a v některých případech dokonce znemožňuje řádnou analýzu a interpretaci archeobotanického materiálu. Výpočední hodnotu ovlivnil též nekonstantní objem jednotlivých vzorků, resp. objem vzorků vůbec. V našem případě ani jeden objekt z dostupných lokalit nebyl vzorkován totálně (nebyla proplavena jeho celá výplň).

Vstupní archeologické informace

Zkoumaný soubor trpěl nedostatkem archeologických údajů o nálezových okolnostech; nejnižší úroveň představoval neurčený kontext nebo uloženina, na nejvyšší úrovni se nacházelo určení kontextu/uloženiny podle čísla, typu, funkce a části, následovalo označení a specifikace uloženiny a zároveň byl archeologický objekt situován prostorově.

Selekce, standardizace a úprava dat

Do analýzy byly zařazeny všechny získané zuhelnatěné rostlinné makrobytky konzervované v suchém prostředí kromě nálezů uhlíků a dřev. Pod pojmem „jeden rostlinný makrobytek“ (dále zkráceně RMZ) se rozumí každé celé zrno, semeno nebo plod (*sensu lato*), pokud si zachovalo celistvý tvar i v případě, že bylo velmi destruované. Pod pojmem „druh“ se rozumí botanický druh (*taxon*).

Analyzovaná data byla rozdělena do čtyř datových tabulek $m \times n$, kdy m reprezentuje počet pozorování (vzorky, objekty/uloženiny, lokality) a n představuje proměnné (botanické druhy, RMZ, objem, hustota). Proměnné druhy, makrobytky a objem byly zvoleny z důvodu jejich přímé souvislosti s výpočty koncentrace (hustoty) makrobytků v litru sedimentu a výpočty četnosti druhů (*ubiquity*). Rovněž jsou tyto proměnné zásadní při tafonomické, ekonomické a ekologické analýze archeobotanických vzorků (např. Bogaard 2004; Hajnalová 2012, 95–110; van der Veen 1992, 73–90, 111–144).

Vzhledem k řešené problematice byla pro analýzu použita nereduovaná data, tedy všechny vzorky, které disponovaly informací o objemu odebraného sedimentu. Analyzovány byly rovněž i vzorky, které neobsahovaly žádné nálezy makrobytků, anebo obsahovaly jen malé množství RMZ (např. 1 ks apod.). Cílem bylo postřehnout v rámci materiálu případné elementy, které by se „tradičně“ standardizací do statistického hodnocení nedostaly, anebo by mohly uniknout při interpretaci výsledků. Vyloučeny z analýzy byly pouze tři vzorky hromadných nálezů, které vykazovaly velmi odlehle hodnoty (více jak 4 000 makrobytků ve vzorku s malým objemem).

Statistické modely

V modelech 1 až 5 (graf 2–6) lze pozorovat vzájemné vztahy mezi objemem, počtem botanických druhů, RMZ a hustotou jednotlivých vzorků. Po prozkoumání vztahu počtu makrobytků (RMZ) a hustoty jsme proměnnou hustotu nakonec vyloučili z dalších analýz, protože v podstatě kopírovala proměnnou RMZ a mohlo by hrozit, že negativně ovlivní výsledky modelů kvůli tzv. souběžnosti (Wood 2021c). Místo toho byla použita funkce tenzorového produktu interakce. Ze vztahu objemu vzorku k počtu

druhů a makrobytků (RMZ) vyplývá, že objem nemá téměř žádný vliv na jejich množství, pokud zkoumáme každý vzorek samostatně. Z hlediska kategorií objemů (A: 1–9 litrů, B: 10–29 litrů, C: 30–40 litrů, D: 41–80 litrů) se však ukazuje, že malé (do 10 litrů) a středně velké vzorky (do 29 litrů) jsou ze statistického hlediska signifikantnější a výkazují vyšší efekt, neboli poskytují o něco vyšší šanci k tomu, aby z nich bylo získáno i reprezentativnější množství nálezů. Nicméně tento výsledek musí být chápán komplexně. Jinými slovy, nejsignifikantnější statistický efekt reprezentovaný malými objemy (do 10 litrů, model 1) rovněž znamená nejsignifikantnější vliv vzorkování na skladbu zachycených makrobytků ze zkoumané situace. Když se nízkoobjemové vzorky zkombinují také s nízkým počtem odebraných vzorků, značně tím riskujeme, že i celková interpretace nálezů bude negativně ovlivněna samotnou metodou vzorkování. V tomto smyslu předložené závěry nemohou být postavené pouze na množství odebraného sedimentu ve vzorku a generalizované na jiné souory pocházející například z jiných podmínek konzervování.

V tomto případě představuje vzorkování jen malou část problematiky a spíše jen první stupeň k dalším analýzám a interpretacím. Důležité je uvědomit si, že ve zkoumaném souboru právě tyto vzorky (do 29 litrů) tvořily nejpočetnější část. Z tohoto hlediska lze tvrdit, že se projevil vliv vzorkování na výsledky statistické analýzy. Není totiž zřejmé, zda by tahle skupina vzorků byla reprezentativní i v případě, kdyby v souboru převažovalo většinové nebo aspoň poloviční zastoupení objemnějších vzorků (40 litrů a více).

Modely 6 až 11 (graf 7–12) o něco víc osvětlují vliv vzorkování na výsledný archeobotanický soubor. Především začlenění proměnné „počet vzorků“ do modelů přineslo nové informace o tom, jak může být skladba botanických druhů ovlivněna vyššími počty odebraných vzorků, resp. *vice versa*. Více odebraných vzorků se v tomto případě jeví lepší alternativou pro získání reprezentativnější skladby archeobotanických nálezů. Co se týče množství odebraného sedimentu v rámci objektů, výsledky částečně korespondují s příslušnými modely 1 až 5, a tedy objem sám o sobě není určující pro získání hojnějšího množství nálezů. Právě vhodně zvolená kombinace vyššího počtu menších vzorků (10 až 30 litrů) by mohla přinést spolehlivé výsledky. Do velké míry však objem vzorku závisí na cílech výzkumu a také na chronologickém datování zkoumané archeologické lokality. Všeobecně, v případě pravěkých archeologických lokalit, především těch neolitických, by mělo být odebráno minimálně 100 litrů sedimentu z jednoho kontextu nebo objektu. To je ovlivněno především způsobem konzervace makrobytků, které se v oblasti mírného klimatu a v suchých půdních podmínkách nejčastěji zachovávají v karbonizovaném (zuhelnatěném) stavu (srov. Jacomiet, Kreuz 1999, 102–107; Kühn, Antolín 2016, 20, 40, s další literaturou). Každopádně je možné domnívat se, že místo jednoho 100litrového vzorku je vhodnější vyzvednout pět 20litrových vzorků prostorově pokrývajících určitou plochu zkoumané archeologické jednotky. Takovéto vzorky se vždy dají v rámci analýzy spojit do jednoho elementu (*sensu stricto* Lee 2012). Bohužel prostorové informace příslušející jednotlivým vzorkům či archeologickým jednotkám byly ve zkoumaném souboru nedostatečné, a proto nebyly začleněny do modelů. Především nebylo možné úplně spolehlivě oddělit vzorky z objektů od vzorků z uloženin, přičemž část objektů byla reprezentována jen jedním vzorkem z jedné uloženiny, u některých objektů nebyly uloženiny uvedeny vůbec, a u několika vzorků byly naopak uvedeny pouze uloženiny. Z těchto důvodů byly analyzovány společně. Blížší prostorová informace by však s velkou pravděpodobností pomohla vysvětlit větší procento variability a poohlít další skryté struktury v datech.

Jevišovka

Archeobotanický materiál získaný z lokality Jevišovka byl rovněž použit v prezentovaných modelech. Jedná se o polykulturní sídlištění lokalitu zkoumanou v režimu záchranného archeologického výzkumu, a pro nás účel byla analyzována data z objektů datovaných do doby římské. Předběžné analýze bylo v disertační práci podrobeno 94 archeobotanických vzorků, které pocházely z 20 objektů (Hlavatá 2017, 176–182, Príloha). Všechny vzorky nesly informaci o objemu, a proto bylo možné použít je pro modelování v celém rozsahu. Celkový objem těchto vzorků dosahoval 626,75 litrů, celkem se v nich nacházelo 2 432 makrobytků patřících 96 druhům.

Z archeologických informací byly k vzorkům uvedeny informace o čísle objektu, odkryvané technické úrovni, kúlové jámě, případně, že se jednalo o vzorek z profilu. V rámci polozemnic pocházelo 50 % odebraných vzorků z interiérových kúlových jam, zbyvající byly odebrány z interiérových uloženin, v jednom případě z hrobu v interiéru. Celkově na lokalitě z doby římské byly odkryty obytné a jiné sídlištění objekty v poměru 1 : 3, nicméně z polozemnic se odebralo třikrát více vzorků než ze zbyvajících. Tady je zřetelná částečná subjektivita vzorkování vzhledem k preferenci (resp. předpokládanému potenciálu) vzorkovaných situací. Přitom z archeobotanické analýzy nevyplynulo, že jednotlivé vzorky z polozemnic by byly paušálně bohatší než vzorky z jiných objektů. Rovněž není možné generalizovat, že by byly vzorky z kúlových jam bohatší než vzorky z interiérových vrstev polozemnic (Hlavatá 2017, CD Príloha).

Ve smyslu analýzy objemů (model 1, výše), patřily vzorky především do objemové kategorie A (do 9 litrů, 78 % vzorků). Šlo převážně o vzorky z polozemnic a kúlových jam. V případě kúlových jam lze do jisté míry očekávat, že nebude vždy možné odebrat objemnější vzorky. Odběr vzorků kategorie A byl však dodržen i v souvislosti s objemnějšími interiérovými vrstvami polozemnic.

Když se podíváme na složení jednotlivých vzorků z Jevišovky (graf 15a), korespondují do velké míry s modelem 1 až 5 (graf 2–6). I zde se ukazuje, že samotný objem vzorků nemá prokazatelný vliv na počet makrobytků a druhů v nich získaných. Avšak počet z objektů odebraných vzorků má signifikantní vliv na počet makrobytků i na počet druhů v nich zjištěných (graf 15b), což se také prokázalo v modelech 6 až 11. Je zde zřetelný trend, že s více vzorky získáme i více makrobytků a rostlinných druhů. Konzistentně s výsledky modelu 6 i tady se jeví, že celkový objem odebraný z objektu má vliv na archeobotanické nálezy. Na grafu 15c je pozorovatelný celkem jasný trend, kdy při odebrání více litrů sedimentu z objektu bylo získáno více makrobytků i více rostlinných druhů. Rozptyl objemů vzorků z Jevišovky zaujímal interval 0,5–32 litrů. V případě celkového souboru dosahoval rozptyl objemů vzorků vyšších hodnot (od 0,2 do 80 litrů).

Shrnutí a diskuse

Zde analyzovaný archeobotanický soubor nebyl, co se týče kvality, ideální, ale představoval „typický“ vzorek dat, se kterými se v archeobotanice potýkáme. I proto je třeba dodat, že zde prezentované výsledky nelze nekriticky přebírat na každé období či lokalitu, je třeba také respektovat jedinečnost archeologických lokalit. Z publikované literatury je zřejmé, že v počtech druhů a makrobytků ve vzorcích se odráží způsob nakládání s nimi nebo způsob jejich depozice. Na výslednou koncentraci makrobytků působí jak původ vzorkované populace (např. pokud mluvíme o tzv. *multiple event contexts*; srov. Hajnalová 2012, 33, 95; Kuna et al. 2013, 71–74; cf. Litynska-Zajac, Wasylkova 2005,

160–162), tak i samotná konzervace zbytků. V suchém a mineralizovaném prostředí se zachovává rádově o desítky až stovky zbytků méně než v zamokřeném (k tomu např. Antolín et al. 2015; Kühn, Antolín 2016, 20, 40; Jacomet et al. 1989, 70–84). To samozřejmě nelze ovlivnit, avšak co může badatel usměrnit, a co také velkou měrou přispívá ke kvalitě vyhodnocení a interpretaci archeobotanického materiálu, je, jakým způsobem a v jakém rozsahu bude materiál získán a zdokumentován. V našem případě odebrané vzorky až na výjimky reprezentovaly malou část zkoumaných objektů. Pokud vynecháme kúlové jámy, jednalo se přibližně o 1 až 5 % z výplně objektu. Proto nelze s jistotou tvrdit, že počet získaných druhů ve vzorcích věrně odráží skutečnou druhovou skladbu zkoumané populace nebo dokonce funkci objektu. Pak by měl mít každý odebraný vzorek ze stejného objektu (vrstvy) velmi podobnou skladbu, což se však zatím například v Jevišovce neprokázalo. Zde nemluvíme o uloženinách, které vznikly například náhodným shořením zásobní jámy, apod. (srov. van der Veen, Jones 2006).

Co je tedy odpovídá na předloženou otázku, jak vzorkování a formální vlastnosti vzorků ovlivňují analýzu, resp. interpretaci archeobotanického souboru? V případě zde analyzovaného souboru byly použitelné pouze informace o zastoupení botanických druhů a makrobytků, jakož i o počtu a objemu vzorků.

Vzhledem k statistickému modelování byl vliv zvolené metody vzorkování zřetelně čitelný v případě objektů, ze kterých pocházelo minimální množství vzorků (1 nebo 2). Myslíme si proto, že archeobotanická interpretace těchto objektů může být nejistá. Jak se ukázalo výše, čím menší množství vzorků z jednoho objektu odebereme, tím menší množství makrobytků můžeme zachytit. Jinými slovy, existuje šance, že odebrání dalších vzorků z téhož objektu může ovlivnit jeho archeobotanickou a možná i archeologickou interpretaci. Jak bylo možné vidět na grafech 8, 9 (modeley 7, 8), počet odebraných vzorků na objekt má jednoznačný statistický efekt odrážející zastoupení botanických druhů a makrobytků v objektu. Rovněž tenhle signifikantní trend je pozorovatelný po hranici 20 odebraných vzorků na objekt, a za touto hranicí se efekt postupně vytrácí (graf 8). Odebráním minimálně 15 až 20 vzorků z jednoho objektu se výrazně snižuje nejistota, že získaná skladba archeobotanických nálezů nebude pro daný objekt reprezentativní. Ve sledovaném souboru však bylo jen z několika objektů odebráno více než 20 vzorků. Vzhledem k tomu si nemůžeme být jistí, zda se ve výsledku neodráží zmíněný fakt. Tenhle závěr musí být v budoucnu ověřen srovnáním s objekty, ze kterých pocházejí vyšší počty archeobotanických vzorků (nejméně 30).

Jsme toho názoru, že vliv zvolené metody na výsledky analýzy souborů je možné potlačit jenom konzistentním vzorkováním. Pokud to archeologická situace dovoluje (resp. v závislosti na charakteru výzkumných otázek), je třeba se pokusit odebírat konstantní množství vzorků z objektů nebo kontextů napříč celou zkoumanou plochou. Také je důležité zvolit vhodné procento reprezentativnosti na jeden objekt/situaci a toho se držet i při zbylých objektech. Z jednotlivých objektů je třeba odebírat alespoň přibližně stejné objemy vzorků, což umožní optimalizovat srovnatelnost výsledků analýz. Jestliže není možné ovzorkovat všechny objekty/vrstvy na lokalitě, tak zvolit vzorkování vybraných objektů proporceně v závislosti na objemu jejich výplně. Rozhodneme-li se vzorkovat podle předem předpokládané funkce definovanou skupinou objektů, např. všechny obytné objekty, výpočetní hodnota souboru samozřejmě reprezentuje především tyto objekty, a nelze ji přímo vztahovat k objektům odlišné funkce na téže lokalitě.

Závěr

Cílem studie bylo zodpovězení otázky, zda a do jaké míry může být archeobotanický soubor ovlivněn i způsobem neboli metodou vzorkování. Výsledky studie vycházejí ze značně heterogenního archeobotanického souboru, ve kterém činil rozptyl objemů vzorků od 0,2 do 80 litrů a rozptyl nálezů makrozbytků od 0 do 8 076 RMZ. Soubor byl archaeobotanicky analyzován a pomocí tzv. generalizovaných aditivních modelů (Faraway 2016; Wood 2017; Zuur et al. 2009) byly modelovány vztahy mezi objemem (množstvím odebraného sedimentu v litrech), počtem vzorků a počtem identifikovaných botanických druhů a makrozbytků. Tyto vztahy byly sledovány na více úrovních, a to z hlediska vzorků (model 1 až 5), archeologických objektů (model 6 až 11) a lokalit (model 12 a 13).

Z dosažených výsledků je možné usuzovat, že zmíněné modely kromě sledovaných vztahů odrážejí také samotnou metodu vzorkování. To se do velké míry podepsalo při analýze a vyhodnocení zde zpracovávaného souboru. Ten musel být podstatně zredukován jak v počtu vzorků, tak z hlediska použitých informací o vzorcích, které často nebyly k dispozici, resp. vůbec nebyly zaznamenány nebo se nedochovaly. Do jaké míry by se výsledky modelů změnily, kdybychom tyto nezaznamenané nebo chybějící informace použili, nelze v téhle chvíli říct.

Kromě informací o vlastnostech vzorků často chyběl i záměr, s nímž byla daná archeologická situace vzorkována a jakým způsobem. Zjednodušeně řečeno, archeobotanické vzorkování postrádalo v terénu výzkumnou otázku. Tuto skutečnost je možné sledovat od výzkumu staršího data až po ty současné. Ačkolи na jedné straně je možné pozorovat tendenci, kdy s vyšším počtem vzorků lze očekávat reprezentativnější a hlavně realitě více odpovídající soubor, na druhé straně je pak k dispozici mnoho objektů a vrstev, ze kterých byl odebrán maximálně 1 nebo 2 vzorky. Celková míra reprezentativnosti jakož i věrohodnosti získaných poznatků pak kolísá. Dojde-li i k subjektivnímu vzorkování, je nutné se zamýšlet nad mírou spolehlivosti interpretací (např. subsistenčních strategií či různých ekologických a environmentálních modelů), vytvořených na základě podobných dat. I proto je vhodné řešit strategii vzorkování v předstihu s odborníkem, a to i z dlouhodobého hlediska. Analýza zde zpracovávaného souboru, jakož i výsledky modelů ukazují, že zde kvůli absenci potřebných informací zůstává velké procento nevysvětlené variability v datech, velmi pravděpodobně ukrývající další skryté struktury a vzorce, které nám prozatím unikají.

Co se týče výsledků zde prezentovaných modelů z hlediska bohatosti archeobotanického souboru získaného z jednotlivých objektů či lokalit, jako vhodnější alternativa se ukazuje odebírat větší počet vzorků, nejlépe o konstantním objemu (10 až 30 litrů). Tenhle poznatek je však v našem případě částečně ovlivněn i vysokým počtem odebraných malých vzorků (do 10 litrů). To však nikterak nezpochybňuje výsledky modelů jako takových, ale především odráží charakter souboru.

Navzdory tomu, že vzorky o objemu do 10 litrů (157 vzorků) se ukázaly jako signifikantní z hlediska počtu druhů, předpokládáme, že velká část z nich byla získána subjektivně (především ty menší jako 5 litrů), nebo šlo o vzorky, které z objektivních důvodů nemohly mít větší objem (například malé kůlové jámy). Proto je potřeba jejich potenciál v budoucnu ještě prozkoumat. Dále je třeba si uvědomit, že i tento odhad objemu by se mohl různit, pokud by v modelovém souboru byly například vzorky s rozptylem od 30 do 150 litrů (viz výše). Není však možné tyto výsledky vykládat způsobem, že odběr několika malých vzorků nahradí stejný počet velkoobjemových vzorků. Jak se ukázalo ve výsledcích modelu 1, malé objemy vzorků by měly být kompenzovány vyšším počtem odebraných vzorků na jeden objekt.

Největší vliv na získané množství botanických druhů a makrozbytků má podle modelů nepochybně ta strategie vzorkování, při které se odebere větší počet vzorků ze zkoumané situace. Zde se ukazuje, že s více vzorky roste i šance vyzvednout reprezentativnější soubor archeobotanických nálezů, což se projevilo na stoupajícím počtu zachycených botanických druhů. To souhlasí s dosavadními zjištěními známými z literatury (např. Hajnalová 2012, 33; Hald 2008, 224; Jacomet et al. 1989, 70–84; Jacomet, Kreuz 1999, 95–112, Abb. 5.5.; Kreuz 2004, 111).

Je důležité myslit i na to, že každá archeologická situace (objekt, uloženina apod.) by měla mít stejnou šanci být ovzorkovaná, bez ohledu na to, zda se v terénu jeví jako potenciálně vhodná, nebo ne. Jenom tak je možné zvýšit kvalitu a věrohodnost dat. Samozřejmě, že každý archeologický výzkum je jiný, a neexistuje tedy ideální situace nebo ideální počet vzorků, který je třeba odebrat. Je potřeba brát v úvahu konkrétní výzkumný záměr, možnosti logistiky, jednotlivé odkryté situace a jejich specifika, stejně jako způsob dokumentace apod. V mnoha případech se (nejen) archeobotanické vzorky zpracovávají až po několika letech a v mnoha případech už k nim nelze potřebné informace dohledat, nedají se tak zachránit. Je škoda, když si archeolog dá tu námahu a archeologické situace i vhodně ovzorkuje, avšak vyhodnocení souboru ztroskotá např. na nedostatečné (nedochované) dokumentaci. I tomu se ale dá předejít, pokud se příprava na odběr, odběr samotný i cíle archeobotanického výzkumu konzultují s odborníkem nejlépe předem nebo i přímo během výzkumu.

Contacts

Jana Apiar

Archeologický ústav AV ČR, Brno, v. v. i.
Čechyňská 19
CZ-602 00 Brno
apiar@arub.cz

Peter Apiar

Archeologický ústav AV ČR, Brno, v. v. i.
Čechyňská 19
CZ-602 00 Brno
papiar@arub.cz