A MULTIVARIATE APPROACH TO PEBBLE MORPHOMETRY: A CASE STUDY OF THE WITÓW SECTION (SOUTHERN POLAND)

Szymon MOL, Krzysztof NINARD *, Agata KUŹMA, Anastasiia DERYK & Piotr ŁAPCIK

Institute of Geological Sciences, Jagiellonian University,
Gronostajowa 3a, 30-387 Kraków, Poland;
e-mails: szymon.mol@student.uj.edu.pl, krzysztof.ninard@uj.edu.pl,
aga.kuzma@student.uj.edu.pl, anastasiia3977.deryk@student.uj.edu.pl,
piotr.lapcik@uj.edu.pl
** Corresponding author*

Mol, S., Ninard, K., Kuźma, A., Deryk, A. & Łapcik, P., 2024. A multivariate approach to pebble morphometry: a case study of Witów section (southern Poland). *Annales Societatis Geologorum Poloniae*, 94: 273–286.

Abstract: Gravel-dominated Neogene – Early Pleistocene braided river deposits of the Witów Series occur in the Carpathian Foredeep, about 20 km north of the front of the Polish Outer Carpathians, between the Szreniawa and Vistula rivers. For the first time, these deposits were subjected to numerical analyses, based on the morphometry and mass of almost 1,500 pebbles sampled at 10 cm intervals along the 4.4-m-high section in the Witów Quarry. In contrast to the traditional approach to pebble morphometry, multivariate statistics was utilised. This enables the examination of various aspects of the dataset holistically and simultaneously. A multivariate method, called principal component analysis (PCA), is widely used in the life sciences, but the employment of PCA for pebble morphometry has not yet been described. Here, PCA was applied to reveal the interrelations between pebble size, mass, lithological composition and stratigraphic height. Most notably, some differences in the distribution of morphometric features between different pebble lithotypes are displayed. Even though the morphometric features and petrological composition of pebbles remain similar in the section as a whole, overall upward-decreasing trends of stream-bed velocity proxies were recognized with the aid of PCA results and were validated, using standard bivariate correlation methods. This approach to the multivariate analysis of large quantitative and qualitative datasets should be considered as a possible part of the integrated sedimentological research of coarse-grained deposits. The consistency between results among the multiple indicators studied reduces the uncertainty of the sedimentological interpretations, presented in this work.

Key words: Multivariate analysis, principal component analysis, pebble morphometry, Carpathian Foredeep.

Manuscript received 23 February 2024, accepted 4 July 2024

INTRODUCTION

Among the Cenozoic deposits of southern Poland, the gravel-dominated fluvial sedimentary succession of the so-called Witów Series represents one of the most comprehensively studied cases. Existing research encompasses the stratigraphy and general sedimentological characteristics of deposits, exposed in the Witów section. Nonetheless, this succession has not yet been subjected to a well-established approach of grain dimension ratio-based pebble morphometric analysis yet. For more than 60 years, such analyses were aimed mainly at comparing different sites laterally and only rarely were conducted to examine vertical variation in a single sedimentary succession. Still, such a traditional approach to pebble morphometry has long served as a valuable tool for inferring on weathering, transport, provenance, and depositional processes (e.g., Sneed and Folk, 1958; Glover, 1975; Russell, 1980; Barudžija *et al.*, 2020; Gale, 2021).

Ordination analysis is a subset of multivariate statistical analysis methods, aimed at data exploration in search of relationships between variables. It also serves as a tool to summarise the main characteristics of the dataset, visualize possible clusters and generate indexes (Bialik *et al.*, 2021). In contrast to inferential statistics, ordination techniques do not allow the verification of preexisting hypotheses but can lead to the formulation of new hypotheses. Ordination analysis, most notably principal component analysis (PCA), is the cornerstone of modern morphometric studies within the scope of life sciences and palaeontology (Hammer and Harper, 2006). Nonetheless, the popularity of that approach is not reflected in morphometric studies of coarse-grained sedimentary rocks, perhaps owing to the considerable amount of work, needed to collect a sufficiently large set of multivariate data and the lack of well-established know-how among investigators, concerning ordination methods (Bialik *et al.*, 2021).

Within the geosciences, ordination methods are increasingly frequently applied in the exploration of palaeontological and geochemical data, but still to a significantly lesser extent in sedimentological investigations. Among previous implementations of PCA in research concerning grain size, cases based on Folk and Ward (1957) granulometric statistics predominate (e.g., Babej *et al.*, 2016; Oyedotun, 2022). However, even though PCA is well-suited to the analysis of pebble morphometric datasets, this research area has remained unexplored (see a summary of PCA applications in geosciences in Bialik *et al.*, 2021). The use of the analytical tools mentioned can be a step towards a more quantitative approach in geological studies. This improvement in data quality allows the more precise comparison of rock records, independently of study area or age of the deposits. In contrast to univariate or bivariate statistics, ordination analysis would allow the simultaneous investigation of both the quantitative (dimensions and mass) and qualitative (lithological composition) properties of clasts, together with their changes in the stratigraphic domain. The advantage of PCA is that it is able to analyse the entire multidimensional dataset at once, instead of merely looking at pairwise relations, as in the bivariate approach. Here, the present authors apply PCA to explore the multivariate morphometric dataset, sampled from the deposits of the Witów Series, which consists of grain dimensions, complemented by grain mass and lithological composition. The aim of the present study is: 1. to demonstrate the interdependence between lithological composition, dimensions and mass of pebbles; 2. to show vertical changes, collectively recorded as all aforementioned parameters; and 3. to present the utility of PCA in the morphometric analysis of gravels as a part of quantitative sedimentological studies, with reference to the example of the classic, well-studied Witów section.

GEOLOGICAL SETTING

Deposits of the Witów Series, originally distinguished by Łyczewska (1948), occur as an elongated sedimentary body, oriented W-E in the area between Nowe Brzesko and Witów, in the interfluve of the Szreniawa and Vistula rivers (Fig. 1). These deposits are located in the Carpathian Foredeep, about 20 km north of the front of Polish Outer Carpathians. The Witów Series was deposited on the Middle Miocene Krakowiec Clay (Lindner and Nowakowski, 1996; Rauch-Włodarska *et al.*, 2006). The lower part of the Witów Series consists of fine-grained clay-rich deposits that gradually coarsen upwards to fine-grained sandstone. The upper part consists of sandy-gravelly deposits, rarely interbedded with silty mudstone. The entire sedimentary succession reaches a thickness up to 40 m and is characterised by high lateral variability of facies, due to the presence of discontinuous lenticular bodies of various lithologies, as well as syn- and post-depositional faults and joints (Rauch-Włodarska *et al.*, 2006). The top of the profile is covered by Pleistocene tills and fluvioglacial gravels, with a loess cover (e.g., Rutkowski, 1987a, b; Lindner and Siennicka, 1994).

Fig. 1. Study area with an overlay of a hillshade digital terrain model (https://baza.pgi.gov.pl/cbdg/geoportal). The outcrop studied is marked by a red arrow.

The age and depositional environment of the Witów Series have been the subject of numerous studies over the past 75 years, leading to different interpretations. Originally, it was considered to be marine littoral deposits of the Carpathian Foredeep, micropalaeontologically dated as Upper Tortonian (Łyczewska, 1948). Subsequently, these deposits were identified as fluvial deposits of the Early Pleistocene (Rühle, 1957), which were later reinterpreted as fluvioglacial deposits, formed during the interstadial of the Nidanian Glaciation. This reinterpretation was based on the rare occurrence of northern-sourced pebbles and the recognition of Miocene foraminifera as redeposited material (Gradziński and Unrug, 1959). In further studies, it was shown that the resistant heavy mineral assemblages differed from those in the deposits with erratic boulders (Kucia-Lubelska, 1966), and the presumed erratic material represented redeposited exotic components from the Polish Outer Carpathians (Dżułyński *et al.*, 1968). The interpretation of the Witów Series as fluvial deposits was maintained, but its age was once again changed to being no younger than the Early Pleistocene (Kucia-Lubelska, 1966), also supported by palynological evidence (Dżułyński *et al.*, 1968). On the other hand, Tyczyńska (1978) interpreted the Witów Series as Miocene marine fan-delta deposits, partially representing a subaerial delta plain. In the 1980s and 1990s, there was a prevailing view of the fluvial nature of the Witów Series, but two competing concepts regarding their age emerged. The first concept supported an Early Pleistocene age, based in part on paleomagnetic studies (Rutkowski, 1987a, b, 1995, 1998; Nawrocki and Wójcik, 1990; Radzki *et al.*, 1992). The second concept suggested a return to the age of the interstadial of the Nidanian Glaciation, supported by thermoluminescence dating (Lindner, 1988; Lindner and Siennicka, 1994; Lindner and Siennicka-Chmielewska, 1995, 1998; Lindner and Nowakowski, 1996). Subsequent palynological and palaeobotanical studies (Brud and Worobiec, 2003) and tectonic analysis (Rauch-Włodarska *et al.*, 2006) indicated sedimentation in braided rivers as part of the freshwater Carpathian molasse during the late Miocene to Pliocene. The authors agree with the interpretation of the Witów Series as a record of deposition in a river floodplain environment. However, new data do not allow for a definitive determination of its age. Therefore, for the purposes of this study, the present authors lean towards interpreting the age of the Witów Series as late Miocene to Pliocene, in accordance with the largest amount of published data (see Brud and Worobiec, 2003; Rauch-Włodarska *et al.,* 2006).

The study was carried out in a former gravel pit, located in the easternmost part of the Witów Series on the left bank of the Vistula River bend, south of Witów village. The study site is a classic outcrop of the Witów Series, which has been studied previously by number of geoscientists (see Rauch-Włodarska *et al.,* 2006), and references therein). The studied deposits are poorly lithified conglomerates and pebbly sandstones with ferruginous-clay-sandy, rarely calcareous cement (Fig. 2). The sources for the gravel are various deposits of the Polish Outer Flysch

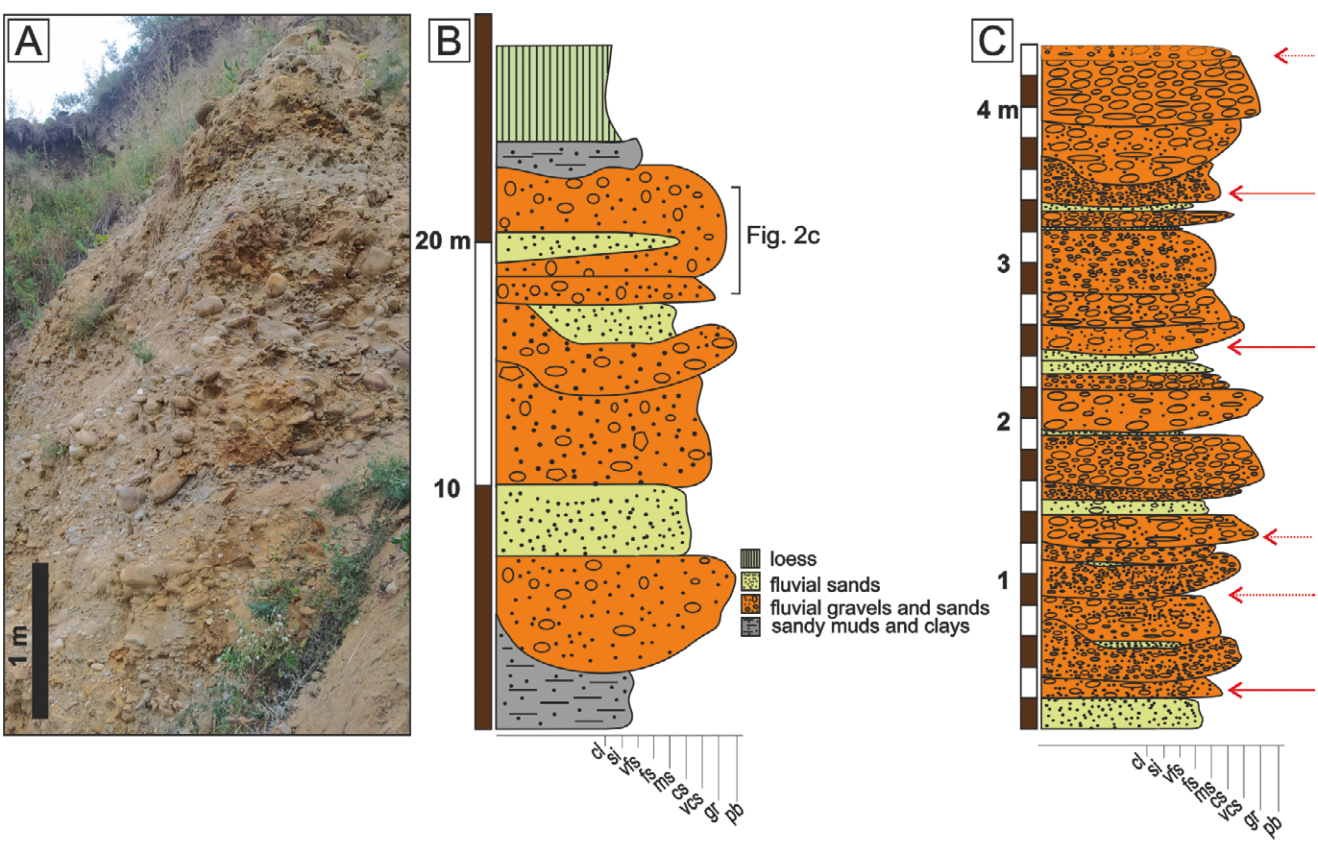

Fig. 2. A Witów Quarry. **A.** Sampled section. **B.** Simplified lithostratigraphic log of the entire Witów Series after Brud and Worobiec (2003). **C.** Sedimentary log of the sedimentary succession studied; episodes of well-documented more energetic flow (red solid arrows), and less evident more energetic flow episodes (red dotted arrows), according to PCA results.

Carpathians (Gradziński and Unrug, 1959; Dżułyński *et al.*, 1968). Sandy intercalations consist mostly of coarse-grained to fine-grained, mostly cross-laminated beds (Gradziński, 1957).

MATERIAL AND METHODS

Sample collection and grain measurements

A 4.4-m-thick and ca. 0.5–1-m-wide section in the upper part of the Witów Series was examined (Fig. 2). The investigated succession was subdivided into contiguous 10-cm-thick interval, 44 in total. The 10 cm thickness of an interval was a compromise between grain size, the desired resolution of results, and labour-intensity. For simplicity, each interval was assigned a single number, starting with the signature 0.0, which denotes a 10-cm interval, starting at a particular height, relative to the bottom of the measured section. Each interval was examined in terms of its grain size and lithological composition. Sampling was carried out for each interval, taking into account lithological diversity. Only grains with *b* dimension size, equal to or greater than 8 mm (medium pebbles and larger, according to Wentworth's 1922 nomenclature), were chosen for further investigation. Thus, 15–48 gravel grains per interval and 1,432 gravel grains in total were measured and weighed. Three linear axes of the particles were obtained using a metric vernier caliper: length *a*, width *b*, and thickness *c*, where *a* is the longest dimension, *b* is the longest dimension perpendicular to *a*, and *c* is the longest dimension, perpendicular to both *a* and *b*.

Petrological composition

The lithotype (quartz, granite, lydite, sandstone) of each clast was identified. On account of the dominance of quartz arenites, sandstone pebbles were subdivided according to their textural properties, i.e., very fine-grained sandstone (vf), fine-grained (f), medium-grained (m), coarse-grained (c), and very coarse-grained (vc). Basic quantitative statistical analysis of petrological composition was performed for the entire sampled population (1,432 grains), as well as for each individual sampling horizon (44 subpopulations). The distribution of pebbles mass was studied using boxplots split by lithotype classes. Moreover, a statistically insignificant admixture of Jurassic limestones and crystalline rocks occur (e.g., Gradziński and Unrug, 1959; Lindner and Nowakowski, 1996).

Grain shape

The dimension measurements were plotted on the Zingg (1935) and Sneed and Folk (1958) diagrams obtained, using custom code in Python with matplotlib (Hunter, 2007) and seaborn (Waskom, 2021) libraries. Additionally, the same types of diagram were applied to each main lithotype separately in order to study the differences and similarities between classes (see Fig. A1 for details). Separate Sneed and Folk (1958) diagrams were also created for 5 larger intervals (0–1 m, 1–2 m, 2–3 m, 4–4.4 m) of the sampled log, with each 10 cm interval plotted (see Appendix Fig. A2).

Principal Component Analysis (PCA)

Measured *a*, *b* and *c* clast dimensions, clast mass (wt), lithotype (lithological composition; ltype) and corresponding sedimentary succession interval above the reference level (stratigraphic height; hgt), together represent the multivariate dataset within the scope of the present study. Comprehension of the relationship between multivariate populations and the interpretation of the processes controlling their distribution can be facilitated by dimensionality reduction of such data. Principal component analysis (PCA) reduces dimensionality with a minimal loss of original variance (e.g., Swan and Sandilands, 1995; Hammer and Harper, 2006), allowing the projection of multidimensional data into two-dimensional space. The normal distribution of all of the data populations studied was rejected, using the Shapiro-Wilk test. Even though the performance of PCA is optimal in the case of symmetrically-distributed, preferably normally-distributed variables, it can also be successfully implemented to not-normally-distributed data (Bialik *et al.*, 2021). Most importantly, all analysed variables (*a, b, c* dimensions and mass) consist of non-compositional data, therefore being suitable for PCA without pre-processing, other than the removal of outliers and normalisation.

While outliers are an intrinsic part of natural datasets, they influence the estimation of principal components (Maronna *et al.*, 2006), obscuring the large-scale tendencies, which were sought in the present study. Therefore, 58 outliers (4% of the dataset) in total, within *a*, *b* and *c* clast dimension variables, were detected using the generalised Extreme Studentized Deviate test (Rosner, 1983) and removed prior to PCA. Thus, correlation matrix-based PCA was carried out for 1,374 four dimensional vectors in PAST 4.14 software (Hammer *et al.*, 2001). Data normalisation, being an inherent part of such a routine, rendered the original variables dimensionless, which allowed direct comparison of the variables measured in different units. Lithotype was adopted as a grouping variable and interval variable of stratigraphic height (hgt) served as a supplementary variable, which was plotted directly in the principal component (PC) reference frame, instead of being included in ordination. Envelopes of data point clouds for each of the lithotypes are represented by convex hulls, with abbreviated lithotype names, projected in the centroid of a particular convex hull polygon. A scatter plot of data point clouds, grouped by lithotype, is provided in Figure A3. In a standard PCA approach, the relative significance of PCs is represented with their normalised eigenvectors. Therefore, it might reflect the importance of individual features.

Detailed analysis of selected features

Mass and *b* dimension size were chosen as the most representative features, which potentially can give an insight into the vertical variability of flow conditions in the studied log. Moreover, a more detailed study of their interrelations with stratigraphic height was motivated by the PCA results. In order to perform such an analysis, mass and *b* dimension size were plotted against stratigraphic height, the best-fit line was determined, and correlation coefficients (Pearson-r, Spearman-rho and Kendall-tau) were calculated. For better visualisation of feature distribution in relation to vertical position, series of boxplots were utilised. Also, *b* dimension size was studied in greater detail as a traditional average grain size proxy (Blott and Pye, 2008). The arithmetic mean and standard deviation of *b* dimension size were calculated for the whole sample, as well as for each interval, separately.

RESULTS

Petrological composition

The dominant lithotype in the samples collected is fine-grained sandstone (Tab. 1). Together with the two next most common classes (medium-grained sandstone and quartz), they constitute more than 93% of pebbles in the entire population. Other lithotypes (very coarse-, coarse-, and very fine-grained sandstone, lydite, granite) are poorly represented in the sample.

The distribution of lithotypes in each interval is roughly uniform along the entire log, even though the subsample (one horizon) size varies from 15 to 48 elements (Fig. 3).

Table 1

Lithotypes of the sampled pebbles

Lithotype	Number of pebbles	Percentage in the sample (%)
Very coarse-grained sandstone (vc)	3	0.21
Coarse-grained sandstone (c)	51	3.56
Medium-grained sandstone (m)	359	25.07
Fine-grained sandstone (f)	763	53.28
Very fine-grained sandstone (vf)	24	1.68
Quartz (q)	216	15.08
Lydite (l)	15	1.05
Granite (g)	1	0.07
Total	1432	100.00

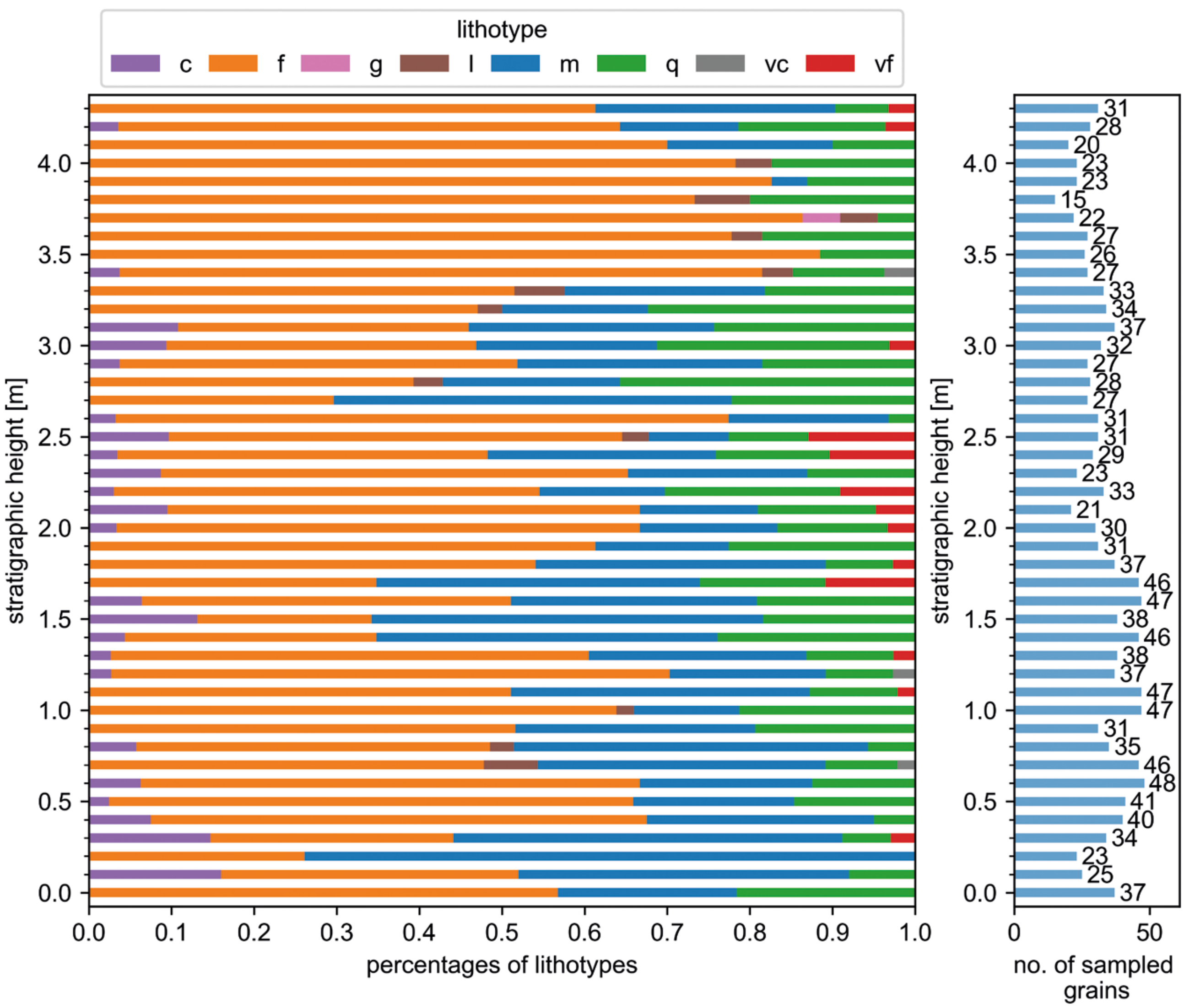

Fig. 3. Percentages of lithotypes and number of sampled grains in each horizon.

Fine- and medium-grained sandstone pebbles predominate in each individual horizon. Quartz grains occur almost in the entire log (except at log 0.2 m horizon). The distribution of coarse-grained sandstones displays four clusters around the intervals 0.1–0.8, 1.2–1.6, 2.0–2.6 and 2.9–3.4 m above the base of the log. Most of the very fine-grained sandstone pebbles are clustered in the log interval 1.7–2.5 m. The occurrences of lydite are concentrated in intervals 0.7–1.0 and 3.2–4.0. The rest of the very fine-grained sandstones, the rest of the lydites, as well as other lithotypes (granites and very-coarse sandstones), are scattered throughout the log, without any specific pattern.

The median masses of very fine-, fine-, medium-grained sandstone, quartz and lydite pebbles are comparable (Fig. 4). Coarse- and very-coarse grained sandstone pebbles have much higher median masses than the rest of the lithotypes. Regarding mass dispersion, quartz has a smaller interquartile range than other classes (except granite).

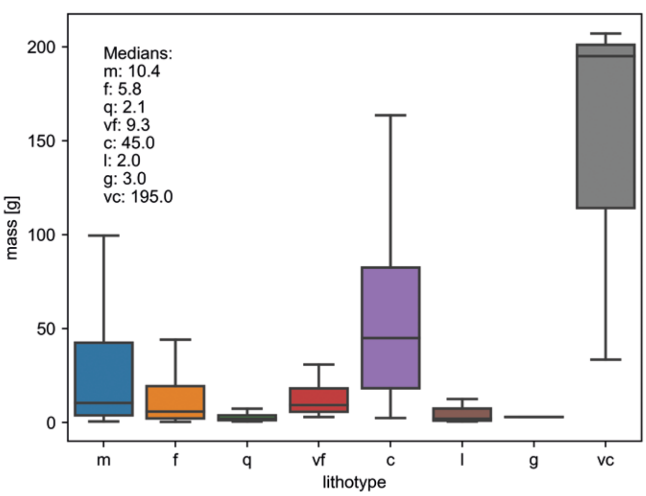

Fig. 4. Boxplot (median with quartiles) of mass distribution in recognised lithologies.

Shape of the measured pebbles

The measurements clearly cluster around the centre of the Sneed and Folk (1958) diagram and its near vicinity, clearly becoming more dispersed towards the boundary shapes at the corners of the diagram (Fig. 5). Over 80% of the total measured pebbles are bladed (27.85%), compact-bladed (14.14%), elongate (11.97%), platy (10.51%), compact elongate (9.28%), and very bladed (8.12%).

The measurements were plotted onto the Sneed and Folk (1958) diagram also separately for each lithotype (Fig. 6). Some variance in dominant grain shapes is visible between the lithotypes. For quartz pebbles, the compact-bladed shape is dominant, whereas the bladed shape, which is dominant in sandstones, is poorly represented. It is also the case for Zingg diagrams, where quartz pebbles are significantly more spherical than other lithologies (Fig. A1). Generally, quartz pebbles are more compact and more spherical, compared to sandstones. Very fine-grained sandstones are more elongated than the rest of the lithologies.

All of the 5 intervals, plotted on the Sneed and Folk (1958) diagrams (0.0-0.9, 1.0-1.9, 2.0-2.9, 3.0-3.9, 4.0-4.4 m; Fig. A2), show a similar distribution of grain shape, which corresponds to the distribution for the whole sample (Fig. 5).

PCA

88% of the total variance is explained by the first principal component (PC) and almost 10% by PCs 2 and 3 together (Tab. A1). In contrast to stratigraphic height, all textural variables (three dimensions and mass) have positive loadings on PC1 (see Tab. A2). Inferring from the acute angles between the eigenvectors, all three clast dimensions are positively interrelated with mass, which by visual inspection of the angles displays a strongly positive interrelation with PC1 (Fig. 7) and a negative one with PC3 (Fig. 8). On the basis of the PC1 to PC2 plot in turn, stratigraphic height is negatively interrelated with both clast mass and PC1,

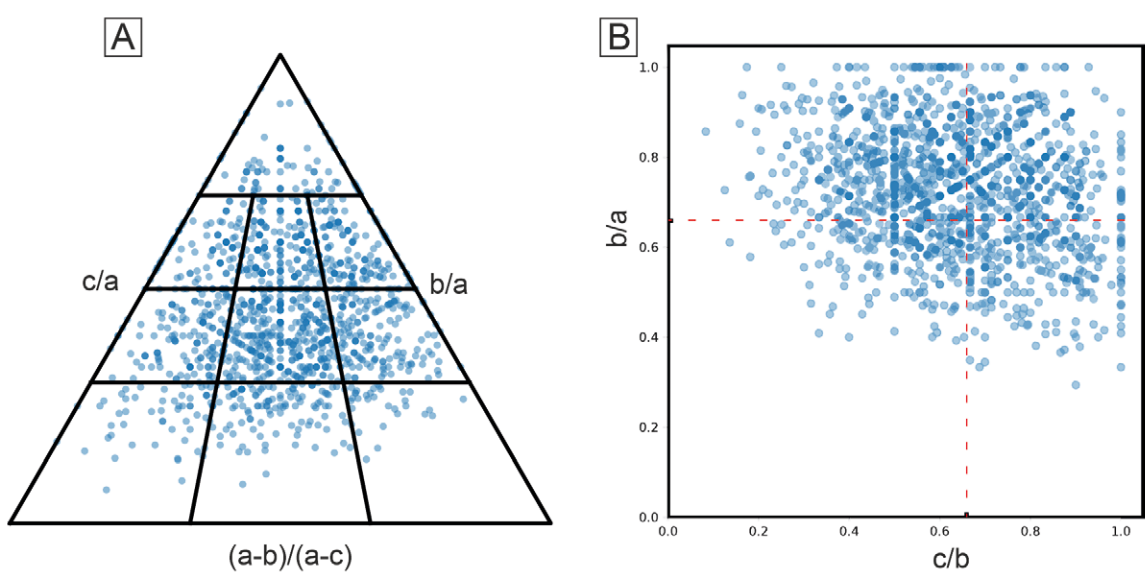

Fig. 5. Results of the grain size morphometry. **A.** Long, intermediate and short orthogonal axes of each measured grain plotted on particle shape diagram after Sneed and Folk (1958). **B.** Three axes of each measured grain plotted on the Zingg (1935) diagram.

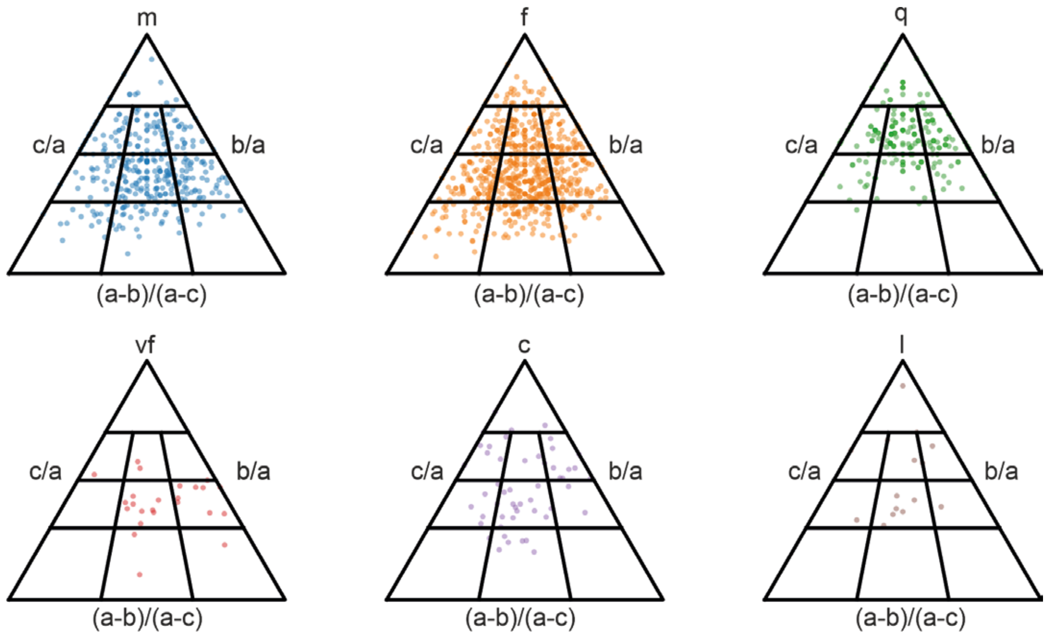

Fig. 6. Long, intermediate and short orthogonal axes of each measured grains plotted on particle shape diagram after Sneed and Folk (1958) for different lithotypes separately (m – medium-grained-sandstone, f – fine-grained sandstone, q – quartz, vf – very fine-grained sandstone, c – coarse-grained sandstone, l – lydite).

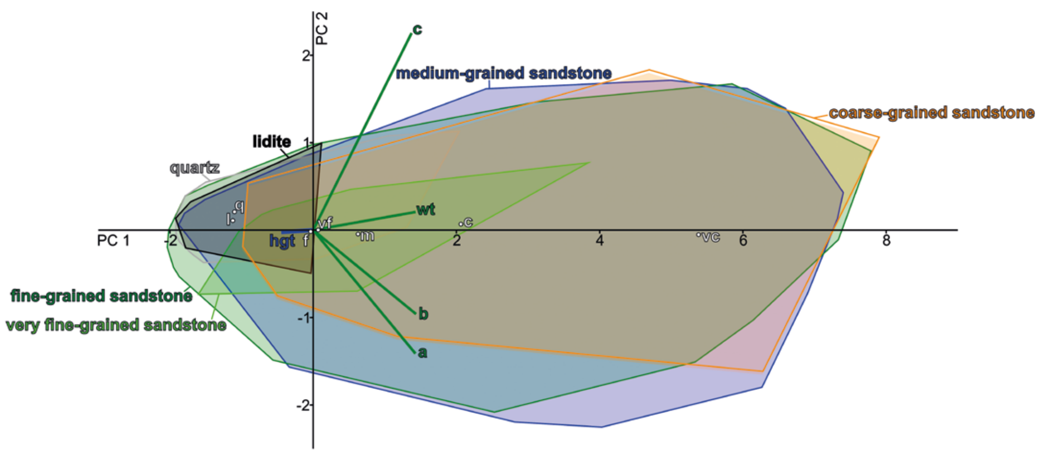

Fig. 7. Biplot of PC1 and PC2 with lithotypes convex hulls and centroids.

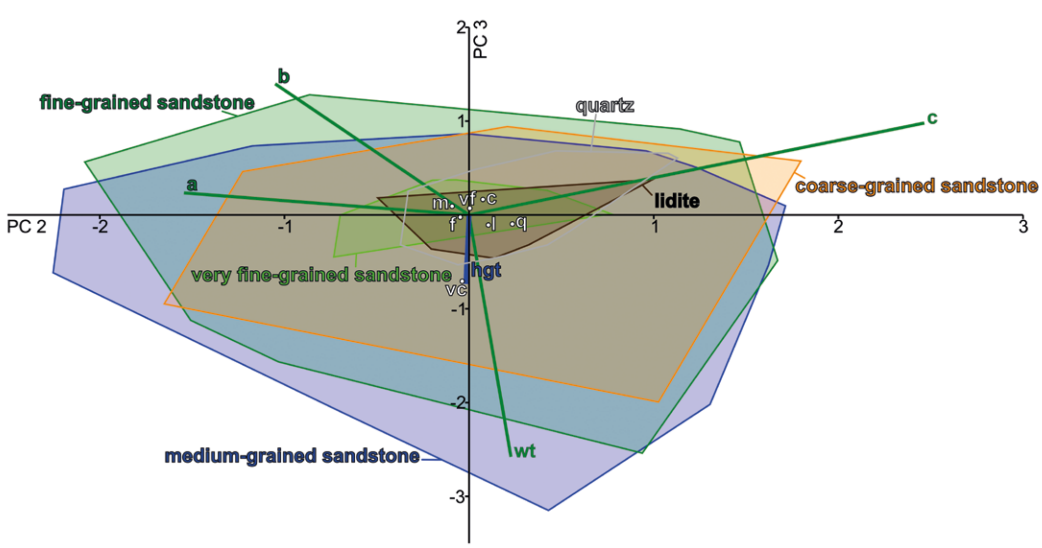

Fig. 8. Biplot of PC2 and PC3 with lithotypes convex hulls and centroids.

because their eigenvectors and PC1 unit vector are nearly antiparallel. Clast dimensions *a* and *b*, while mutually related, are nearly orthogonal to dimension *c*. Clast dimensions *a* and *c* display some relationship to PC2. Even though in the scattergram there is no clear separation between the point clouds representing particular lithotypes, the centroids of respective convex hull polygons display a linear array along PC1. Ordering, according to the texture of clasts (from very coarse-, through coarse-, medium-, very fine- and fine-grained sandstone, to lydite and microcrystalline quartz, left to right), is clearly visible.

Vertical trends

Mass distribution in the sample displays a very slight negative correlation with stratigraphic height (Pearson-r = -0.04, see Fig. A4). In general, the median mass is more or less similar in each sampled horizon except a few of them, where it is greater (Fig. 9). The value of 7.15 g (70th centile) was arbitrarily chosen as a boundary value for median mass in the intervals studied. Such a value enables objective differentiation of the intervals with heavier grains. Pebbles from intervals 0.1–0.4, 3.4–3.5 and 4.3 are clearly of greater mass. There are also intervals, such as 0.8–0.9, 1.2, 2.4, 3.7, and 4.1, which have a minor increase in median mass, compared to others. The Pearson-r coefficient equal to -0.11 shows that *b* dimension has a weak negative correlation with stratigraphic height (see Fig. A5 for details). Median of *b* dimension size is moderately varied along the profile (Fig. 10). Most of the individual medians are in the 10–20 mm range. There are only 3 groups of intervals, where the median *b* dimension size exceeds 20 mm: 0.1–0.4, 3.4–3.5, 4.3. There are also additional, less prominent peaks with higher median *b* dimension sizes than the immediate vicinity, e.g., 0.8–0.9, 1.1–1.2, 2.3–2.5. The arithmetic mean of *b* dimension size for the entire sample is 20.21 mm and the standard deviation is 13.91 mm. The majority of *b* dimension size arithmetic means for individual horizons fall in the range 15–25 mm (Fig. 11). Note that these figures are greater than for analogous values, related to medians.

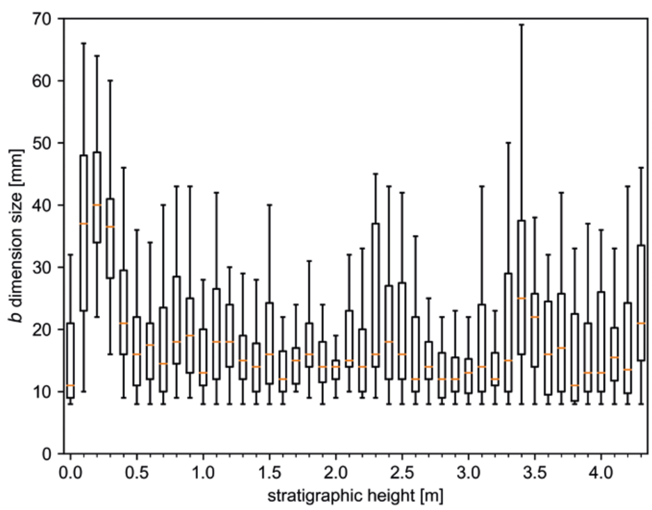

Fig. 9. Mass distribution boxplots (median with quartiles, without outliers) in relation to the stratigraphic height.

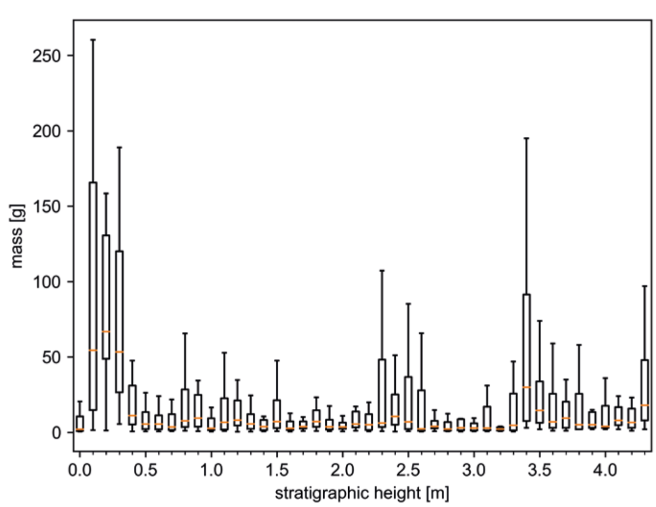

Fig. 10. Distribution boxplots of b dimension size (median with quartiles, without outliers) in relation to the stratigraphic height.

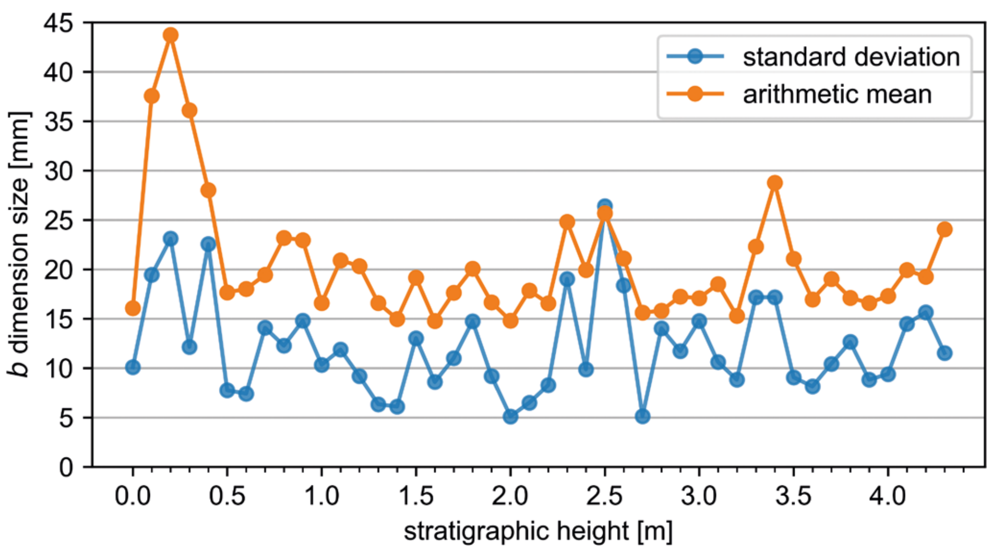

Fig. 11. Arithmetic mean and standard deviation of *b* dimension size in relation to the stratigraphic height.

Only horizons 0.1–0.3, 2.3, 2.5 and 3.4 display arithmetic means greater than 25 mm. It appears that for mean *b* dimension size, there is a kind of cyclicity with a period of ~0.3–0.4 m, especially between horizons from 0.5 to 2.0. The curve of standard deviation appears to be synchronised with the curve of arithmetic mean, i.e., usually peaks correspond to peaks and lows to lows.

DISCUSSION

Petrological compositions

Owing to the poor representation of granites and very coarse-grained sandstones in the sample (Tab. 1), any quantitative results regarding these two classes are insignificant. Marginal content of granites (<0.1%) in the sample supports the observations of Gradziński and Unrug (1959), regarding the petrological composition of the Witów Series. Similar median masses of certain lithotypes indicate that individual grains in the very fine-, fine-, and medium-grained sandstone, and quartz and lidyte subpopulations were transported and deposited in comparable flow conditions. The fact that coarse- and very coarse-grained sandstone pebbles are in general heavier than grains belonging to other lithotypes indicates their transport under conditions of greater stream-bed velocity. Nevertheless, the possibility of reduced availability of such grains for the most part in the sedimentary system cannot be excluded. A small interquartile range of quartz grains mass implies that they are more or less uniform in mass (Fig. 4). Assuming they have similar density, they should also have similar volume. Combining this with the shape of quartz grains, which has a tendency to be more spherical (Fig. A1) and more compact (Fig. 6), it might be inferred that in the sample, most of quartz pebbles can be characterised as grains, roughly similar to one another in both shape and size. The overall trends in the shape of quartz pebbles might indicate that they are not of a primary deposit, but had undergone several redeposition events, which led to increased degrees of sphericity and compactness (cf. Gradziński and Unrug, 1959).

In general, the dominant lithotypes (fine- and medium-grained sandstone, quartz) in the entire sample remain dominant in every interval studied. Thus, the entire succession has similar characteristics, in terms of petrological composition, regardless of the stratigraphic height, which is in agreement with a previous study (Rutkowski, 1987b, 1998). With the lack of visible hiatuses in the log, such a phenomenon can be interpreted as a roughly uniform deposition of fine- and medium-grained sandstone and quartz material in a fluvial environment. However, some additional lithotypes occur in a minority in many horizons. Their clustering in a vertical direction (Fig. 3) can be explained by changes in the textural properties of eroded sediments at the source to those with a higher proportion of coarse- or very fine-grained sandstones, respectively (cf. Gradziński and Unrug, 1959). This is supported by the varied composition of different lithostratigraphic units of the Polish Outer Flysch Carpathians (Ślączka *et al.*, 2006). However, a detailed analysis of grain provenance is beyond the scope of this study. Additionally, in the case of coarse-grained sandstone pebbles, which are more massive in comparison to other classes (Fig. 4), their distribution can be related to the occurrence of more powerful flows. Nevertheless, the relationship between clusters and proxies of stream-bed velocity (such as *b* dimension size and mass) is not clear. Some clusters are not directly linked to increased values of these features (Fig. 12).

PCA

A typical behaviour of PCA results, based on the measured morphometric features of natural objects, is that PC1 can be associated with the general 'size' of study objects (Hammer and Harper, 2006). Accordingly, in the present study, PC1 reflects a combination of clast dimensions and mass, therefore being directly representative of the minimum stream-bed velocity (Rubey, 1938). The negative relationship between

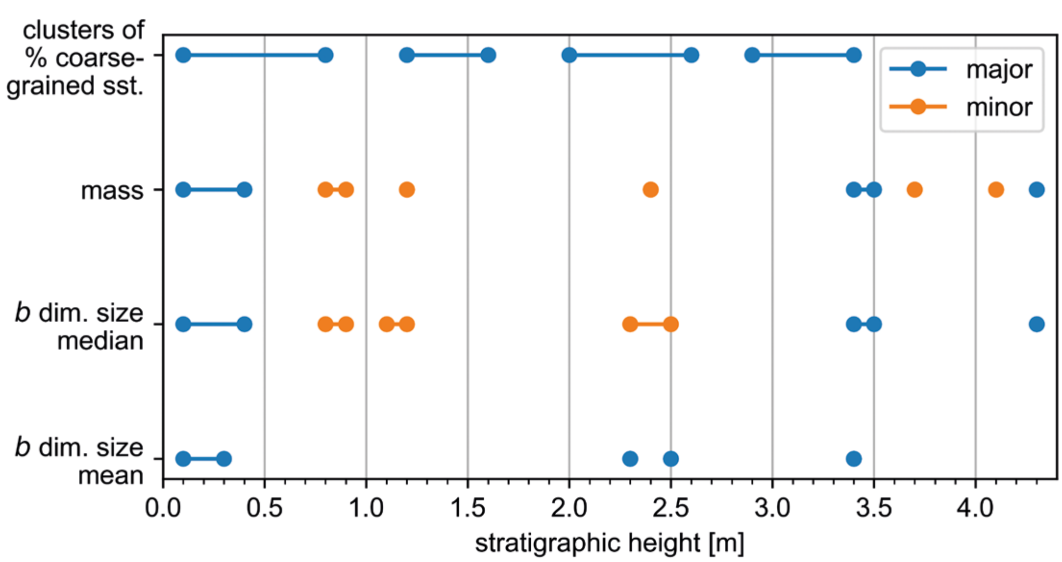

Fig. 12. Comparison of horizons with higher-than-typical values of features studied. Major and minor increases are marked with different colours, as appropriate.

PC1 (approximating clast size) and stratigraphic depth (Fig. 7) can demonstrate the general trend of decrease in average grain size up the studied section. Concurrently, lithotype centroids represent the averaged multivariate populations, associated with each lithotype, as reduced to single points, allowing examination of the general relationship between lithological composition, size and mass of grains. Linkage between the size and lithological composition of a clast can be inferred from the evident distribution of lithotype centroids along PC1. Although convex hulls occupy more or less the same space, centroids are separated. Such behaviour indicates that in various lithotypes, the distribution of certain features can be slightly different. The ordering of mass medians by lithotype (Fig. 4) fits surprisingly well to the ordering of lithotype centroids in the PCA biplot (Fig. 7). This phenomenon can be considered as a good example of the distributions of different feature among lithological classes. Whether such a delicate difference is sufficient to correctly distinguish the lithotypes of individual grains should be a matter for future studies.

Vertical trends

PCA results indicate that there are negative relationships between *b* dimension size and stratigraphic height as well as between mass and stratigraphic height. Both of them are confirmed by the negative Pearson-r coefficients for mass (-0.04) and *b* dimension size (-0.11), although only the latter correlation is clearly statistically significant (see Figs A4, A5). These values of correlation coefficients indicate that for *b* dimension size, the correlation is more negative than for mass. However, the angles in the PCA biplot (Fig. 7) imply that the negative relationship with stratigraphic height should be stronger for mass than for *b* dimension size. In this work, such a discrepancy can be considered as a disadvantage of the quantitative interpretation of PCA. However, the qualitative results are validated, using the standard bivariate statistical approach.

B dimension size means, *b* dimension size medians and mass medians are characterized by moderate dispersion with respect to stratigraphic height. Each of these features individually can be interpreted as a proxy of stream-bed velocity. Thus, it can be inferred that flow conditions were more or less similar in the entire sampled log. Moreover, such a conclusion is supported by the multiple indicators studied.

Nevertheless, each of the features, studied quantitatively in detail (*b* dimension size and mass), displays some horizons, where the measured values are greater than typical figures. The stratigraphic height of such horizons is not consistent for all of the features (Fig. 12). However, it seems that clusters 0.1–0.3, ~2.4 and 3.4 above the reference level occur in the mass median, *b* size median and *b* size mean. There are also less evident clusters 0.8–0.9, 1.2 and 4.3, which are common for mass median and *b* size median. Collectively, it can be inferred that the log records 3 episodes of well-documented more energetic flow. Additionally, there are probably 3 more such episodes, where a minor increase in stream-bed velocity occurred, but apparently their sedimentological record in the *b* dimension size means is weaker. However, it should be noted that it is possible for grains of different sizes to be transported in similar flow conditions. Vertical variations in the section rather are related to local, lateral migrations of fluvial facies than to major changes in the entire sedimentary system.

Overall, *b* dimension size arithmetic means for individual horizons are greater than medians. This reflects the fact that in general, outliers in each subsample are represented by larger grains. Some cyclicity in the *b* dimension size mean curve can be related to roughly regular changes in the flow conditions. Such a situation is plausible in a braided river, which is the sedimentary environment of the Witów Series. In general, the standard deviation of a grain size distribution is interpreted as a proxy of sediment sorting (Folk, 1974). Synchronization of *b* dimension size arithmetic mean, and *b* dimension size standard deviation curves reflects a tendency in the log, i.e., the coarser the sediment, the poorer its sorting.

CONCLUSIONS

The numerical and graphical results of PCA comprehensively reflect the interrelations between all measured pebble dimensions and stratigraphic height. A general decrease in grain size and mass up the section is indicated, albeit not very clearly, by standard bivariate plotting and correlation analysis. However, PCA indicates this trend clearly, collectively taking into consideration all grain dimensions and mass. Compared to the traditional bivariate approach, PCA allowed the registration of trends with higher sensitivity and facilitated the interpretation of results by providing an integrated image of interdependence between all variables at once. The points representing particular pebble lithotypes are arrayed along PC1, according to the texture of the clasts. Such ordering evidently reflects the relationship between the lithology and size/mass of the pebbles and shows that such morphometric data are adequate for discrimination between lithotypes, using PCA. Furthermore, the same methodology-based studies in different sedimentary successions would be advantageous to verify if a similar phenomenon persists in the presence of other pebble lithotypes, such as crystalline or carbonate rocks.

Sedimentary conditions were roughly similar during the entire time of deposition of the studied section. However, there were some episodes with deposition of larger grains and possibly a general slightly negative trend in grain size. The similarity is both in terms of petrological composition and flow conditions. In the entire dataset, as well as in all individual horizons studied, fine- and medium-grained sandstone and quartz pebbles are predominant. The values of the proxies of stream-bed velocity (*b* dimension, size and mass) do not change much with respect to stratigraphic position. The level of uncertainties in the interpretations presented is reduced because of consistent results for the multiple indicators studied.

The general tendency of quartz grains to be more spherical and more compact than grains of other lithotypes indicates that they underwent several depositional cycles. There is a synchronisation of the degree of coarsening and the degree of sorting. There are some intervals in the studied section, in which a kind of cyclicity in stream-bed velocity can be recognised.

Acknowledgements

This work was partly funded by the "Excellence Initiative – Research University" programme at the Jagiellonian University. We are deeply grateful to Anna Mickiewicz for her tremendous help with fieldwork and laboratory measurements. We also thank Artur Galicki, Amelia Janas and Józef Brzózka, who supported us at various stages of this project. We would like to express appreciation to Michal Janocko, an anonymous reviewer, as well as editor Michał Warchoł for their valuable comments on the manuscript.

REFERENCES

Babej, J., Máčka, Z., Ondrejka, P. & Peterova, P., 2016. Surface grain size variation within gravel bars: a case study of the River Opava, Czech Republic. *Geografia Fisica e Dinamica Quaternaria*, 39: 3–12.

Barudžija, U., Velić, J., Malvić, T., Trenc, N. & Matovinović Božinović, N., 2020. Morphometric characteristics, shapes and provenance of Holocene pebbles from the Sava River gravels (Zagreb, Croatia). *Geosciences*, 10: 92.

Bialik, O. M., Jarochowska, E. & Grossowicz, M., 2021. Ordination analysis in sedimentology, geochemistry and palaeoenvironment—Background, current trends and recommendations. *The Depositional Record*, 7: 541–563.

Blott, S. J. & Pye, K., 2008. Particle shape: a review and new methods of characterization and classification. *Sedimentology,* 55: 31–63.

Brud, S. & Worobiec, G., 2003. Results of investigations on plant macroremains found in deposits of the Witów Series (southern Poland). *Przegląd Geologiczny,* 51: 392–401. [In Polish, with English summary.]

Centralna Baza Danych Geologicznych Geoportal, 2024. Polish Geological Institute – National Research Institute. https://baza.pgi.gov.pl/cbdg/geoportal [In Polish; 09.01.2024].

Dżułyński, S., Krysowska-Iwaszkiewicz, M., Oszast, J. & Starkel, L., 1968. On Lower Quaternary gravels in the Sandomierz Basin. *Studia Geomorphologica Carpatho-Balcanica,* 2: 63–74. [In Polish, with English summary.]

Folk, R. L., 1974. *Petrology of Sedimentary Rocks*. Hemphill Publishing Co., Austin, 170 pp.

Folk, R. L. & Ward, W. C., 1957. Brazos River bar [Texas]; a study in the significance of grain size parameters. *Journal of Sedimentary Research*, 27: 3–26.

Gale, S. J., 2021. The shape of fluvial gravels: Insights from Fiji's Sabeto River. *Geosciences*, 11: 161.

Glover, B. K., 1975. A morphometric analysis of terrace gravels in Santa Ynez basin, Santa Barbara County, California. *Sedimentary Geology*, 13: 109–124.

Gradziński, R., 1957. Notes on the sedimentation of the Miocene in the environs of Proszowice. *Rocznik Polskiego Towarzystwa Geologicznego*, 26: 3–27. [In Polish, with English summary.]

Gradziński, R. & Unrug, R., 1959. Origin and age of the „Witów series" near Cracow. *Rocznik Polskiego Towarzystwa Geologicznego*, 29: 181–195. [In Polish, with English summary.]

Hammer, Ø. & Harper D. A. T., 2006. *Paleontological Data Analysis*. Blackwell, Oxford, 368 pp.

Hammer, Ø., Harper D. A. T. & Ryan P. D., 2001. PAST: Paleontological statistics software package for education and data analysis. *Palaeontologia Electronica,* 4. [9 pp.]

Hunter, J. D., 2007. Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering*, 9: 90–95.

Kucia-Lubelska, M., 1966. The age of Witów series based on the heavy mineral assemblages. *Rocznik Polskiego Towarzystwa Geologicznego*, 36: 303–313. [In Polish, with English summary.]

Lindner, L., 1988. Jednostki glacjalne i interglacjalne w plejstocenie Wyżyny Miechowskiej i Niecki Nidziańskiej. *Przegląd Geologiczny*, 36: 140–147. [In Polish, with English summary.]

Lindner, L. & Nowakowski, A., 1996. Problem pochodzenia otoczaków granitu w osadach serii witowskiej w świetle badań petrograficznych. *Przegląd Geologiczny*, 44: 950–952. [In Polish.]

Lindner, L. & Siennicka, A. E., 1994. Osady czwartorzędowe w strefie północnej krawędzi doliny Wisły na wschód od Brzeska Nowego, Wyżyna Miechowska. *Przegląd Geologiczny*, 42: 105–112. [In Polish.]

Lindner, L. & Siennicka-Chmielewska, E., 1995. Loesses and their bedrocks in the Southeastern Part of the Miechów Upland (S Poland). *Annales Universitas Mariae Curie-Skłodowska, sectio B*, 50: 75–89.

Lindner, L. & Siennicka-Chmielewska, E., 1998. The Witów series and the problem of the Tertiary/Quaternary boundary in South-Eastern Poland. *Mededelingen Nederlands Instituut voor Toegepaste Geowetenschappen TNO*, 60: 287–292.

Łyczewska, J., 1948. Report on the geological investigations in the north-western corner of Brzesko Nowe map 1:100 000 (Carpathian foreland). *Biuletyn Państwowego Instytutu Geologicznego,* 42: 46–75. [In Polish, with English summary.]

Maronna, R, Martin, R. D. & Yohai, V. J., 2006. *Robust Statistics: Theory and Methods.* John Wiley, New York, 436 pp.

Nawrocki, J. & Wójcik, A., 1990. On age of the "Witów series" in the light of palaeomagnetic studies. *Studia Geomorphologica Carpatho-Balcanica*, 24: 3–9.

Oyedotun, T. D. T., 2022. Compositional and multivariate statistical analyses for grain-size characterisation of intertidal sedimentary facies in an estuarine environment. *Geology, Ecology, and Landscapes*, 6: 224–230.

Radzki, P., Łopusiński, L. & Widz, D., 1992. *Objaśnienia do Szczegółowej mapy geologicznej Polski w skali 1:50 000, arkusz Borzęcin.* Wydawnictwa Geologiczne, Warszawa, 33 pp. [In Polish.]

Rauch-Włodarska, M., Zuchiewicz, W., Brud, S. & Galicia T. Group, 2006. Tectonics of Miocene-Pliocene fresh-water molasses in the Carpathian Foredeep (Witów Series, South Poland). *Journal of Geodynamics,* 41: 369–384.

Rosner, B., 1983. Percentage points for a generalized ESD many-outlier procedure. *Technometrics,* 25: 165–172.

Rubey, W. W., 1938. The force required to move particles on a stream bed. *United States Geological Survey Professional Papers*, 189-E: 121–141.

Russell, T., 1980. Use of clast shape in determining the sedimentary history of the Late Devonian Keepit Conglomerate, Australia. *Sedimentary Geology*, 25: 277–290.

Rutkowski, J., 1987a. O niektórych trzecio- i czwartorzędowych żwirach rejonu Krakowa. In: Rutkowski, J. (ed.), *Trzecio- i staroczwartorzędowe żwiry Kotliny Sandomierskiej. Materiały Sympozjum Komitetu Badań Czwartorzędu PAN, 22–24 czerwca, AGH, Kraków*. Wydawnictwa Akademii Górniczo-Hutniczej, Kraków, pp. 24–26.

Rutkowski, J., 1987b. Uwagi o żwirach witowskich Krakowa. In: Rutkowski, J. (ed.), *Trzecio- i staroczwartorzędowe żwiry Kotliny Sandomierskiej. Materiały Sympozjum Komitetu Badań Czwartorzędu PAN, 22–24 czerwca, AGH, Kraków.* Wydawnictwa Akademii Górniczo-Hutniczej, Kraków, pp. 33–39. [In Polish.]

Rutkowski, J., 1995. Petrographic composition of the Quaternary gravels of the Carpathians and their foreland. *Studia Geomorphologica Carpatho-Balcanica*, 29: 77–88.

Rutkowski. J., 1998. O staroczwartorzędowych żwirach z Witowa. In: Gorczyca, E. (ed.), *Współczesne procesy morfologiczne i ewolucja rzeźby progu Karpat i ich przedpola. Warsztaty Geomorfologiczne w Łazach k/Bochni, wrzesień 1998.* Instytut Geografii Uniwersytetu Jagiellońskiego, Kraków, pp. 90–97. [In Polish.]

Rühle, E., 1957. *Mapa utworów czwartorzędowych Polski w skali 1:2 000 000. Biuletyn Instytutu Geologicznego,* 118. [In Polish.]

Sneed, E. D. & Folk, R. L., 1958. Pebbles in the lower Colorado River, Texas a study in particle morphogenesis. *The Journal of Geology*, 66: 114–150.

Swan A. R. H. & Sandilands, M., 1995. *Introduction to Geological Data Analysis*. Blackwell Science, Oxford, 446 pp.

Ślączka, A., Kruglov, S., Golonka, J., Oszczypko, N. & Popadyuk, I., 2006. Geology and hydrocarbon resources of the Outer Carpathians, Poland, Slovakia, and Ukraine. General Geology. In: Golonka, J. & Picha, F. J. (eds), The Carpathians and their foreland: geology and hydrocarbon resources. *AAPG Memoir*, 84: 221–258.

Tyczyńska, M., 1978. The fossil deltas in the Sandomierz Basin. *Folia Geographica. Series Geographica-Physica*, 11: 33–50. [In Polish, with English summary.]

Waskom, M. L., 2021. seaborn: statistical data visualization. *Journal of Open Source Software*, 6: 3021.

Wentworth, C. K., 1922. A scale of grade and class terms for clastic sediments. *The Journal of Geology*, 30: 377–392.

Zingg, T., 1935. Beitrag zur Schotteranalyse. *Schweizerische Mineralogische und Petrographische Mitteilungen*, 15: 39–140.

APPENDIX

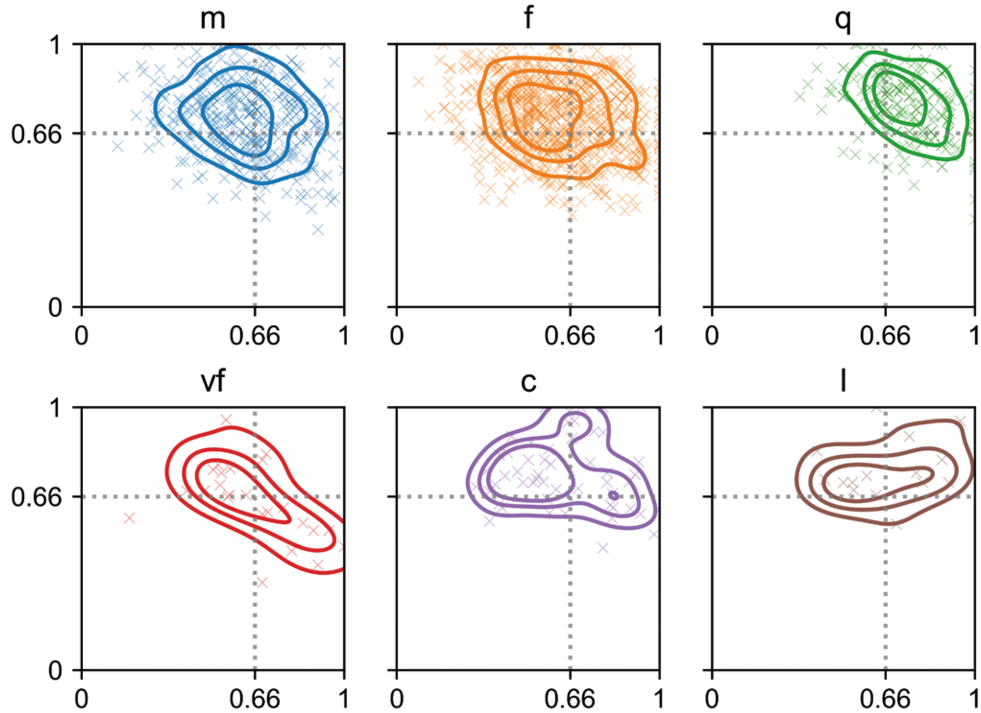

Fig. A1. Zingg (1935) diagrams for grains of certain lithotypes (m – medium-grained sandstone, f – fine-grained sandstone, q – quartz, vf – very fine-grained sandstone, c – coarse-grained sandstone, l – lydite) and kernel distribution estimation (with contours of 25%, 50% and 75% probability mass, i.e., 25% of the probability distribution lies outside the outermost curve, etc.). The plot was prepared, using the kdeplot function from seaborn library (Waskom, 2021) for Python. Chosen parameters were default, except clipping range (0, 1) in both directions and parameters controlling graphics.

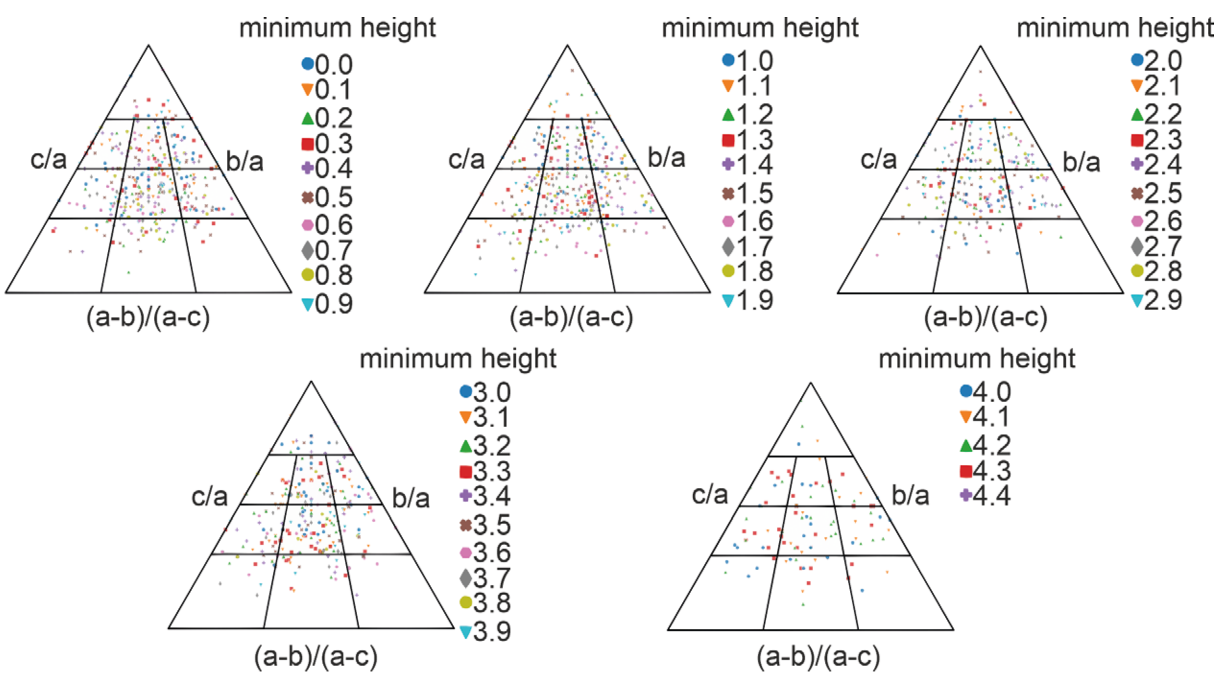

Fig. A2. Long, intermediate and short orthogonal axes of each measured grain, plotted on particle shape diagrams after Sneed and Folk (1958) separately for 5 log intervals (0.0–0.9, 1.0–1.9, 2.0–2.9, 3.0–3.9, 4.0–4.4 m). Different colours represent intervals of 10 cm (minimum height of each interval indicated).

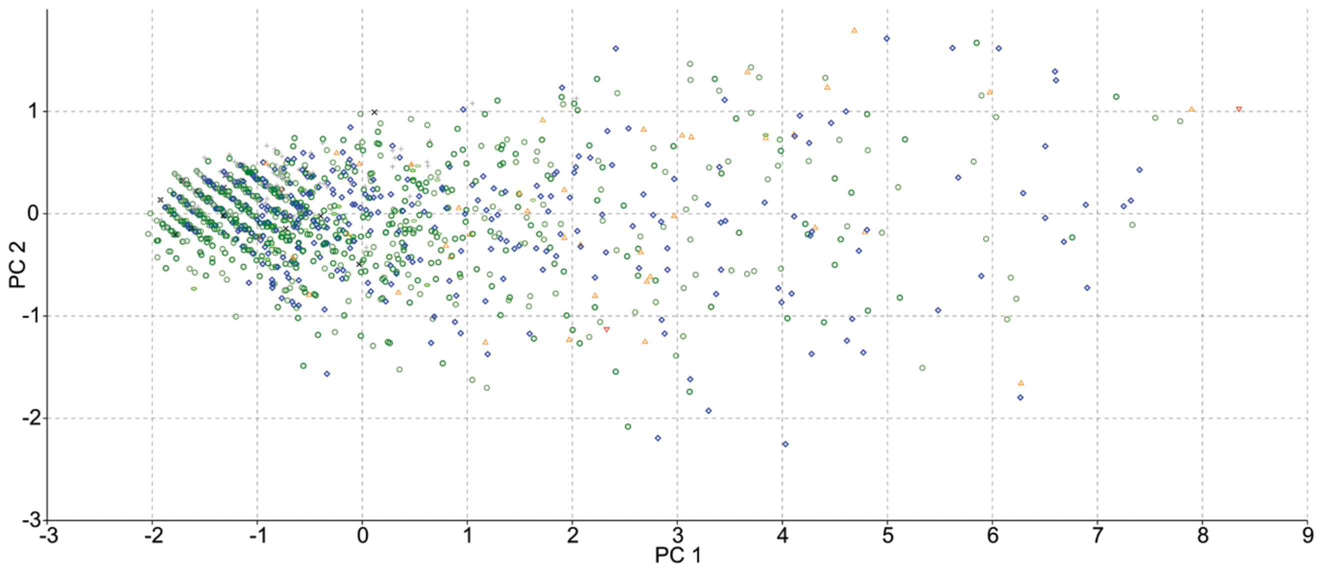

Fig. A3. Scatterplot of data point clouds grouped by lithotype on PC1 and PC2 axes.

Table A1

Percentage of variance explained by particular principal components.

PC	Eigenvalue	% variance
1	3.523720	88.0930
2	0.244642	6.1160
3	0.140266	3.5066
4	0.091368	2.2842

Table A2

Loadings of principal components per measured variables.

Loadings	PC 1	PC 2	PC 3	PC 4
hgt	-0.088720	-0.005660	-0.237090	-0.084760
a	0.503560	-0.497970	0.075820	0.701930
b	0.506670	-0.338630	0.450580	-0.652380
c	0.484850	0.795000	0.315840	0.182050
wt	0.504610	0.073067	-0.831550	-0.220350

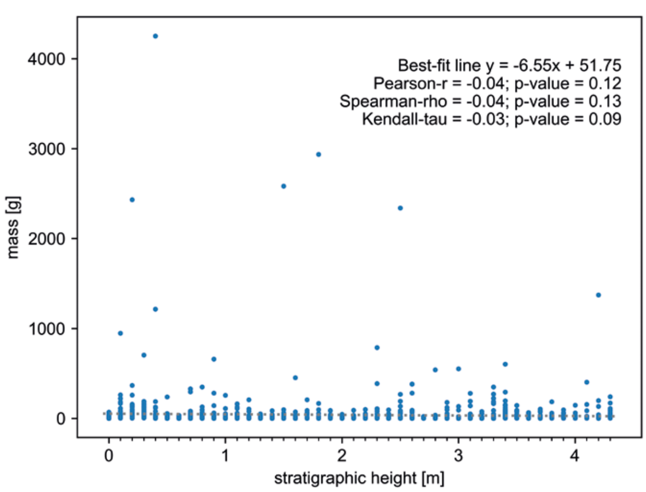

Fig. A4. Scatterplot of mass in relation to stratigraphic height.

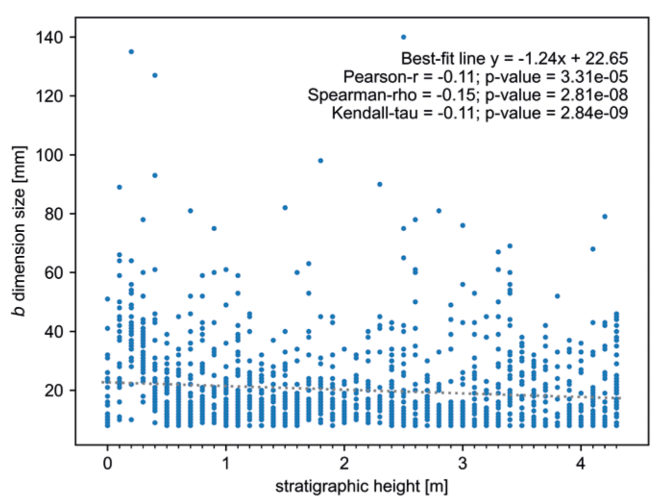

Fig. A5. Scatterplot of *b* dimension size in relation to stratigraphic height.