Front. Earth Sci. 2020, 14(1): 152–170
https://doi.org/10.1007/s11707-019-0765-9

RESEARCH ARTICLE

Geomorphometric relief classification with the *k*-median method in the Silesian Upland, southern Poland

Bartłomiej SZYPUŁA (✉)[1], Małgorzata WIECZOREK[2]

1 Institute of Earth Sciences, Faculty of Natural Sciences, University of Silesia in Katowice, Sosnowiec 41-200, Poland
2 Department of Geoinformatics and Cartography, Faculty of Earth Sciences and Environment Management, University of Wroclaw, Wrocław 50-137, Poland

Abstract The aim of this study is geomorphometric relief classification of a temperate latitude upland area in Central Europe. The Silesian Upland represents diversified structural relief which contains a fan-shaped configuration of long thresholds and wide erosion depressions. A 20 m × 20 m digital elevation model (DEM) provided input data for the analysis. The *k*-median method was applied to examine morphometric variables of the relief. The aim of these activities was to identify clusters with objects of similar mathematical characteristics. These clusters were the basis of landform classification. Smaller numbers of clusters 4 transparently show hypsometric relationships. Key elements of the morphology of the area were clearly visible. The division into 6 clusters gives the best results—a detailed but clear image of the morphological diversity by distinguishing characteristic landform elements. The results for 8 clusters show significant background noise and are ambiguous, which makes them difficult to identify. Our research has confirmed that the *k*-median method is a useful tool for landform classifications. We determined optimal parameters of this method (filtering window size, DEM resolution, number of clusters, aspect influence).

Keywords *k*-median method, relief classifications, digital elevation model (DEM), geomorphometry, Silesian Upland

1 Introduction

The earth's surface is structured into landforms as a result of the cumulative influence of many processes, mainly geomorphic, geological, and hydrological. Landforms are, therefore, widely recognized as natural objects that partition the earth's surface into fundamental spatial entities. Landform entities differ from one another in terms of characteristics such as shape, size, orientation, relief, and contextual position (MacMillan and Shary, 2009). Geomorphometry as a science of quantitative land surface analysis proved to be helpful in recognizing landforms. von Humboldt (1849) was probably the first who defined geomorphometry (or rather morphometry) as the characterization of landforms by quantitative descriptions of the shape of the earth's surface forms and by quantitative measurements of the "physical constitution" of the earth's surface. This kind of analysis has become possible with the development of precise measurement techniques and the availability of high-resolution data like digital elevation models (DEMs). Computer implementation and using DEMs have enabled very large areas (sometimes difficult and not previously available for mapping) to be classified morphometrically (see Guzetti and Reichenbach, 1994). DEMs contain important geomorphological information about quantitative characteristics of the Earth's surface. A grid DEM stores elevation values at regularly distributed points from which characterizations of the form of the land surface can be estimated (Pike, 1988; Moore et al., 1991; Mitášová et al., 1996). Due to the diversity of landforms, the need for their classification arose naturally. The concept of landform classification as spatial taxonomy is not new and dates back to the middle of the 20th century (Hammond, 1954, 1964; Van Lopik and Kolb, 1959; Wood and Snell, 1960), but only the development of computers has enabled its implementation. Computer landform classification can give the geometric signature, "a set of measures that describes topographic form well enough to distinguish geomorphologically disparate landscapes" (Pike, 1988). Landform mapping based on DEMs allows: much more accurate and objective representation of forms, including their boundaries, the analysis of large areas, and

Received November 19, 2018; accepted May 4, 2019

E-mail: bartlomiej.szypula@us.edu.pl

comparisons between different regions. The issues of morphometric studies in geomorphology are still important and research problems related to automatic landform classifications are still valid. Many landform classification methods have been created (e.g., Dikau et al., 1991; Azanon et al. 2004; Drăguț and Blaschke 2006; Iwahashi and Pike, 2007; Deng, 2007; Minár and Evans 2008; Tang and Li, 2008; Mentlik and Novotna, 2010; Drăguț and Eisank, 2011 and 2012; Drăguț et al., 2013; Jasiewicz and Stepinski, 2013). Moreover, many other geomorphometric studies use DEMs: a new approach to crater detection (Luo et al., 2013), automatic detection and identification of watercourses (Broersen et al., 2017), gully boundary extraction based on multidirectional hill-shading from high-resolution DEMs (Yang et al., 2017), object-based classification of global undersea topography and geomorphological features (Dekavalla and Argialas, 2017), LiDAR data analysis detecting subtle landforms of slope failure (Ortuño et al., 2017), new object-based methods for the semi-automated extraction of drumlins footprints (Jorge and Brennand, 2017), a new approach to the openness index for landform characterization (Alonso-Sarría et al., 2018), and the identification and quantitative analysis of the remnants of planation surfaces (Liu et al., 2019). This work follows the trend of geomorphometric research of relief on the basis of DEMs (see also Szypuła, 2017).

The germ of the idea of our research was Tobler's First Law of Geography (Tobler, 1970): "Everything is related to everything else, but near things are more related than distant things". We decided to analyze whether mathematical similarity of morphometric parameters inside given groups reflects morphological similarity. The authors decided to use a method which provides an analysis using multiple variables collectively. Such a method is called clustering, which is defined as a process of putting data into clusters (groups) of similar types using some physical or quantitative measures. One of these clustering methods is the *k*-median, which is a modification of the more common *k*-mean method. The *k*-median method belongs to the group of non-supervised methods; it is a preliminary analysis method allowing similarities occurring in the data themselves to be identified. It is an objective tool that allows grouping results into different numbers of groups (different values of *k*) in statistical terms. This is a completely different approach than the classic landform classifications often used in geomorphometry. Usually, we first need expert knowledge to set parameters. In the *k*-median method, the algorithm exploring the data first works, the result can be evaluated in statistical terms, and only at the end the result is evaluated and interpreted by the expert. Thanks to this, we introduce the element of subjectivism at the end of the analysis, not at the beginning of the analysis. Earlier works with this method for a mountain region with height differences up to 1000 m (Wieczorek, 2008; Wieczorek and Migoń, 2014) and medium-altitude mountainous region with altitudes of 1300–2500 m (Piloyan and Konečný, 2017) demonstrated the method's effectiveness. The obtained results showed principal features of morphology well. We decided to check if the *k*-median method would also work for a temperate latitude upland area.

The main goal of this study was geomorphometric relief classification of the Silesian Upland. The direct result of these analyses was the classification of landforms divided into 4, 6, and 8 classes. The area of the Silesian Upland is an interesting region with diversified topography. Unfortunately, there is no current study of the entire Silesian Upland relief and landform classification based on DEM could fill this gap of knowledge. Existing geomorphological studies of the Silesian Upland have two major limitations. The first is that, although previous maps cover the entire Silesian Upland, the scales are too small, i.e., 1:2000000–1:500000 (Klimaszewski, 1947; Kondracki, 1951; Starkel, 1980; Gilewska and Klimek, 1997). The second limitation is that there are many detailed analyses but they cover only small fragments of the Silesian Upland (Hornig, 1955a and 1955b; Karaś and Starkel, 1958; Klimaszewski, 1959; Karaś-Brzozowska, 1960 and 1963; Gilewska, 1963; Klimek, 1966; Żmuda, 1973; Chmal, 1976; Szczypek, 1977; Jania and Szczypek, 1980; Lewandowski, 1982; Bukowska-Jania, 1983; Szczypek, 1986a and 1986b; Lewandowski, 1987; Szczypek, 1988; Szczypek and Wach, 1991, 1992 and 1993; Dulias, 1994 and 1995; Jania et al., 2014).

This paper focuses on the methodological aspect associated with the use of the *k*-median method. In particular: 1) verification of the use of aspect as an important DEM derivative in the *k*-median calculations, 2) testing optimal parameter settings of the analyses (filtering window size, DEM resolution, the number of clusters, etc.), and 3) comparison of the reliability of the results with a different classification method (Topographical Position Index) by Weiss (2001) and with geomorphological and geological maps and hypsometry.

2 Study area

The first research area was a fragment of the Silesian Upland — known as the "test area", in which initial preliminary computational experiments were carried out. The test area was 20 km × 30 km and would serve as a well-recognized representative area, showing the relief diversity of the Silesian Upland. Secondly, this area was material in checking optimum parameter settings of the *k*-median method to obtain satisfactory results in subsequent steps.

The main study area of this research was the whole Silesian Upland. This region is located in southern Poland within the watershed of the two largest Polish rivers: the Odra in the west and the Vistula in the east (Fig. 1). The

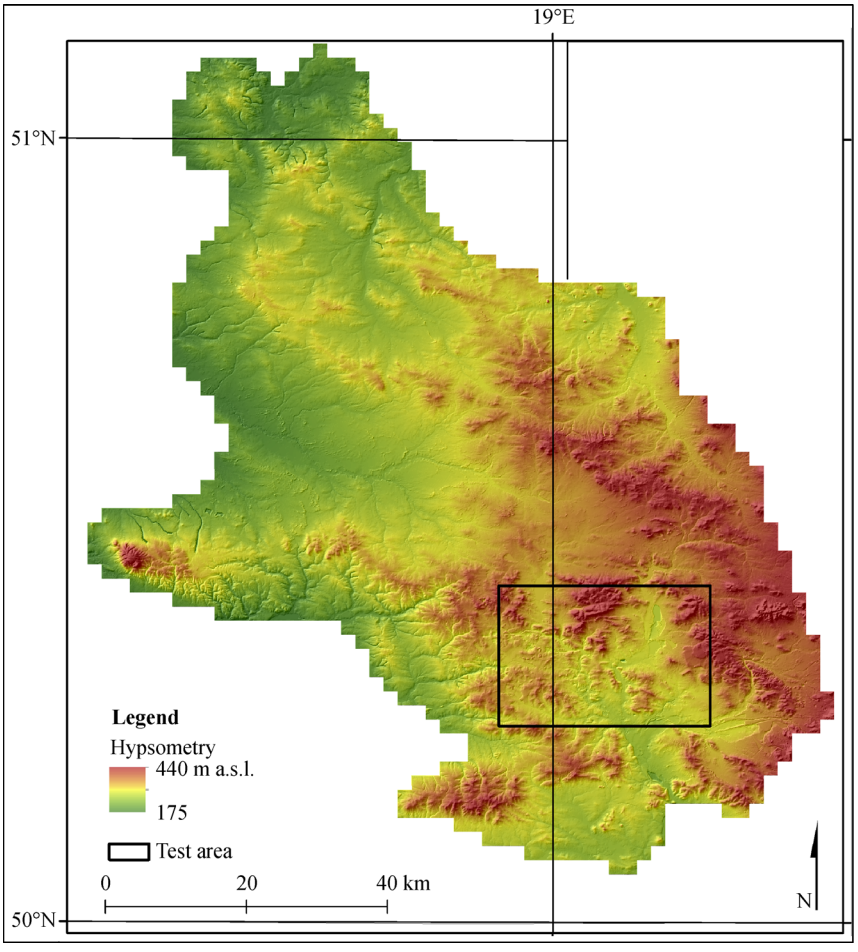


Fig. 1 Location and hypsometry of the Silesian Upland and test area.

current geographical units for the whole of Poland are the result of geomorphological (Galon, 1972; Klimaszewski, 1972; Gilewska, 1986) and regional divisions (Kondracki, 2001). The boundaries and extent of the study area were proposed by Szaflarski (1955) and Gilewska (1986), who identified the Silesian Upland based on morphological, hypsometric, and geological criteria. The eastern border of the Silesian Upland is defined as the Upper Jurassic threshold (part of the Cracow-Częstochowa Upland); the southern border runs along the Krzeszowice Graben and the Vistula River valley; in the southwest the Silesian Upland borders the Racibórz Basin; the western boundary coincides with a 200 m a.s.l. contour line passing through the Silesian Lowland and the northern border is defined by the depression along the upper Warta and Prosna rivers.

The Silesian Upland is the most westerly protruding segment of the middle-Polish uplands (Gilewska, 1972) which occurs in the form of asymmetric tectonic uplift. Geologically, it is monoclinal in structure. The Silesian Upland consists of Paleozoic and Mesozoic deposits with variable erosional resistance, dipping toward the north-east, which created convenient conditions for the development of structural relief (Gilewska, 1999a). Morphogenetically, this area represents past glaciated areas with transitional characteristics at the limits of the South-Polish glaciation (Sanian = Mindel) and the Middle-Polish glaciation (Odranian = Riss) (Gilewska, 1999b). The study area was invaded twice by the Scandinavian inland-ice (Mindel, Riss).

The Silesian Upland occupies an area of approximately 6500 km^2. It is a region with diversified relief (elevation range: 180 m to 440 m, mean: 288 m a.s.l.). Among its characteristic relief features are: asymmetry in the NW-SE cross-section; a radial arrangement of younger-age structural thresholds (Middle-Triassic, Upper-Triassic, Middle-Jurassic, Upper-Jurassic) with the highest elevations in the east and the lowest ones in the north and west (Fig. 1). The analyzed area consists of a fan-shaped configuration of long cuestas and wide erosion depressions. These often divided cuestas are present in the central and northern parts of the Upland. They are large-scale forms (50–90 km long), with relative heights of about 100–150 m. They are characteristic of undulating plateaus as well as broad ridges and hills. These forms are mainly composed of resistant carbonate rocks: Triassic and Jurassic limestones, dolomites, marls, and sandstones. Wide depressions and basins

between these cuestas are drained by the major rivers of the region. They are widespread flat or slightly undulating areas filled with clay sediments, clays, and Pleistocene glacial tills, sands, and gravels that reach up to tens of meters in thickness.

The Silesian Upland is an example of structural relief — the character of the landforms reflects their internal geological structure (mainly resistance) (Szypuła, 2009). This area is a typical mid-latitude upland and, thus, it can be compared with similar regions in Europe (Parisian Basin, London Basin, Swabian-Franconian Basin, Weser river basin) (Klimaszewski, 1991).

3 Data and methods

3.1 Source data

The quality of digital data — as the core material for later analysis — is the key item which provides reliable results. Even the most sophisticated geomorphometric algorithm will not fix artifacts and errors in the input DEMs. The best results are obtained from homogeneous material at good resolution. A DEM of Poland (DTED-2, 2001) with an accuracy approaching that of topographical military 1:50,000 scale maps (Czajka, 2009) was the basic data used for this study. This DEM has a spatial resolution of 25 m × 25 m. The assumed absolute horizontal accuracy is < 23 m and vertical accuracy is < 12 – 18 m (DMA, 2000), but our tests have shown much higher accuracy (vertical accuracy approx < 3.5 m).

Due to the fact that the DTED-2 model contains only meter integer values, the height varies in a discrete manner. To ensure surface continuity, its transformation to real values was performed as proposed by Urbański (2012). The procedure was as follows: 1) a value of 0.5 was added to each DTED-2 cell; 2) from the processed DEM above, 1 m interval contours were generated; 3) information about the river system and water bodies was derived from the Digital Map of Hydrographical Division of Poland (MPHP, 2010); and 4) an interpolation using Topo-to-Raster in ArcGIS was performed (which is based on an ANUDEM model — the best method for interpolating terrain elevations (Hutchinson, 1989 and 2011) using 1-m contours and vector data of rivers and water bodies. This resulted in a continuous hydrologically corrected new DEM with 20 m × 20 m resolution. This resultant DEM became the basis for final calculations.

3.2 Morphometric parameters

As Dikau (1989) claimed, an accurate morphometric classification may be performed using four basic parameters obtained from a DEM (slope, aspect, plan, and profile curvatures). These four primary local attributes (Wilson and Gallant, 2000) were derived from the DEM using proper transformations available in ArcGIS software (ESRI, 2017). True height as one additional data set was calculated from the DEM. True height (elevation range) was evaluated in 5 × 5 raster cells (e.g., 100 m × 100 m), as this best captures the natural amplitude of topography. The median filter was used for all variables (excluding aspect) to remove random isolated extreme values of morphometric variables. For aspect, regular median function could not be applied due to its directional nature. Instead, an average vector was calculated. This filtration did not have any impact on the statistical distribution. The result of these transformations was a set of layers, subsequently subject to final clustering.

3.3 Statistics of morphometric variables

Table 1 presents descriptive statistics of values for the six data sets for the whole Silesian Upland and the test area. The statistics for the Silesian Upland show an asymmetry of data values frequency distribution. Even if the mean is equal to the median value (as for plan and profile curvature), it is caused by a huge number of raster pixels with a value close to the median value, which is proved by a very high value of kurtosis. Skewness is extremely low or high for both curvature parameters, which suggests asymmetric distribution (although the asymmetric factor is one of the lowest in this set of data). The most symmetric variable in this case is the initial altitude. Wieczorek (2011) showed that the size of input data influences the results of *k*-median clustering. Therefore, descriptive statistics were calculated also for the test area. The conclusions are similar to the previous ones for the Silesian Upland. The mean and median values are equal for plan and profile curvature which have extreme values of skewness and kurtosis. Moreover, DEM statistics of the test area are similar to the entire Silesian Upland area, as well as for slope and aspect — i.e., the test area is a good representative area for the whole of the Silesian Upland.

3.4 *k*-median method

The authors searched for a method which provides an analysis of variables collectively, not separately. Therefore, a method which can be used in a multidimensional analysis was needed. Data mining is the process of analyzing huge sets of data from different perspectives. Clustering is defined as a process of putting data into clusters (groups) of similar types using some physical or quantitative measures (Larose, 2005). One of these clustering methods is the *k*-median, which is a modification of the more common *k*-mean method. The goal of *k*-median clustering is to separate data into distinct clusters based on data differences using medians as measures of each subset center. The *k*-median method is robust to outliers and results in compact clusters, while the *k*-mean is easily affected by outliers. In sets of data describing relief and its morpho-

Table 1 Descriptive statistics of Silesian Upland and test area

Variable	Statistics of area	Silesian Upland	Test area
DEM	Min	176.02	217.03
	Max	436.26	400.31
	Mean	267.94	288.68
	Median	266.01	282.29
	Skewness	0.27	1.19
	Kurtosis	−0.26	1.34
	Asymmetric Factor	0.04	0.24
True height	Min	0.00	0.00
	Max	36.47	36.47
	Mean	1.20	1.67
	Median	0.85	1.25
	Skewness	3.60	2.89
	Kurtosis	27.43	20.73
	Asymmetric Factor	0.26	0.22
Slope	Min	0.00	0.00
	Max	45.05	45.05
	Mean	1.36	1.91
	Median	0.97	1.44
	Skewness	3.54	2.96
	Kurtosis	26.67	23.31
	Asymmetric Factor	0.26	0.21
Aspect	Min	0.00	0.00
	Max	360.00	360.00
	Mean	183.76	177.63
	Median	195.21	183.69
	Skewness	−0.14	−0.06
	Kurtosis	−1.14	−1.03
	Asymmetric Factor	−0.17	−0.12
Plan Curvature	Min	−10.69	−10.69
	Max	7.62	7.62
	Mean	0.00	0.003
	Median	0.00	0.001
	Skewness	−8.93	−8.65
	Curtosis	3593.43	2210.70
	Asymmetric Factor	0.09	0.12
Profile Curvature	Min	−10.48	−10.48
	Max	8.89	8.89
	Mean	0.00	0.003
	Median	0.00	0.001
	Skewness	8.25	6.44
	Curtosis	1990.34	817.89
	Asymmetric Factor	0.03	0.08

metric parameters, the data are very seldom symmetric (see Evans, 1972; Evans and Cox, 1999; Wieczorek and Migoń, 2014). The method was expected to handle directional variables, such as aspect, in an appropriate way. Therefore, the *k*-median method with the Manhattan metric system was chosen for multidimensional clustering of morphometric parameters. The clustering algorithm was described in Wieczorek and Migoń (2014). All variables have the same importance in evaluating the distance to the cluster center, so the values are standardized by using standard deviation. An observation is, thus, a six-dimensional vector of six filtered morphometric variables for each pixel. Cluster analysis was performed using software specially prepared for this operation.

4 Experiments

As demonstrated by Wieczorek (2011), the results of operations on digital data (terrain models) are "sensitive" to changes in the input parameters settings. Therefore, prior to the calculations for the entire Silesian Upland, initial measurements were conducted in different variants in small test areas. First, 5 test fields of 2 km × 2 km were selected in a well-recognized topography. Each test field represented different kind of relief (Fig. 2):

I – an area of residual hills due to rock hardness (Lower-Triassic dolomites and marls) on river terraces, with local relief of ⩽100 m.

II – flat and wide low depression after sand mining

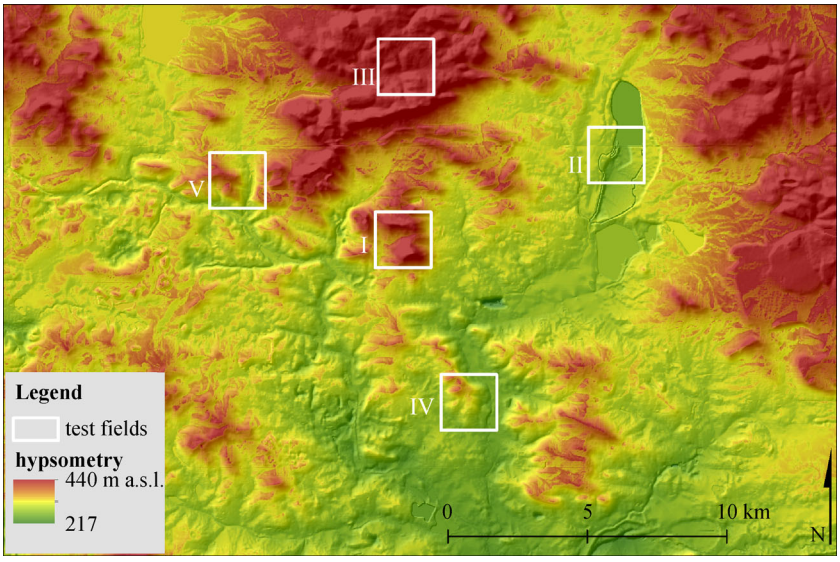


Fig. 2 Test area with the test fields.

(Pleistocene and Holocene sands and gravels), nowadays filled with water, consisting in part of an artificial river channel.

III – a plateau cut by floodplain terraces and a small fragment of a planation surface in watershed setting, hills with numerous pits and quarries (local relief⩽50 m); low- and mid-Triassic sandstones, marls, and dolomites.

IV – fragments of floodplain and river terraces within a plateau which form a high river bank near an artificial river channel (local relief⩽54 m); low- and mid-Triassic marls and dolomites and covered by Pleistocene glaciofluvial sands and gravels.

V – fragments of small plateaus (mid-Triassic dolomites) cut by artificial river channels (filled with Holocene and Pleistocene glaciofluvial sands and gravels) with floodplain terraces and small fragments of planation surfaces (local relief⩽63 m).

The next step was selecting the optimal computational settings which give the best results. Based on previous research (Wieczorek, 2008), the authors decided to use clustering into 4, 6, and 8 clusters (Table 2). In all the experiments we used the same morphometric variables, but they were defined by various settings of input parameters. Because of a high level of mosaic (background noise) in the output maps, different types of filtering (median, focal statistics, and block statistics) were tested. Therefore, for each of the 7 named experiments, a dozen or so alternative computational variants were made. Table 3 shows exemplary settings of selected variants of these experiments.

Thirdly, the results were assessed. All the results were thoroughly examined at a large scale using shaded relief raster (from DEM), the topographical maps, and topography knowledge of the test fields. The experiments

Table 2 Sum of square between clusters for different number of clusters

Number of clusters (*k*)	4	5	6	7	8
*BSS**	20.50	26.28	31.65	37.44	44.40
BSS/*k***	5.13	5.26	5.27	5.35	5.55

Notes: **BSS* = blind source separation value; ***k* = number of classified groups

assessment proves that the results were frequently highly generalized (ex 1A, 1B), with a large share of background noise (ex 2), block statistics resulting in artificial image divisions (ex 3), and too coarse final resolution (ex 4). The best results were obtained from the settings of experiment 5. The determined clusters properly describe the relief of the test fields for 4 and 6 clusters. Table 1 contains statistics of the morphometric variables for the entire area (Silesian Upland) and the test area according to the settings of experiment 5. (Fig. 3). The division into 8 clusters is dubious — in some parts it looks better than the division into 6 clusters, and in others much worse. However, at this stage all the three divisions into 4, 6 and 8 clusters were kept. Moreover, it can be observed that the application of median filtering successfully removed background noise. Better results were achieved when the median filtering window was the same size as the basic analysis window (in experiment 5 we have the same size 7×7 cells). This is probably a mathematical consequence of the applied analysis window size.

5 Results

5.1 Landform classification

The results of landform classification of the Silesian

Table 3 Input parameters of the selected variants of experiments

Experiment number	Input DEM resolution	Size of filter window (grids)	Number of clusters	Morphometric variables*	Additional function	Final raster resolution
1A	50 m×50 m	5×5	4, 6, 8	h, s, a, p, pc	median filtering (3 × 3)	50 m×50 m
1B	50 m×50 m	7×7	4, 6, 8	h, s, a, p, pc	median filtering (3 × 3)	50 m×50 m
2	20 m×20 m	5×5	4, 6, 8	h, s, a, p, pc	median filtering (3 × 3)	20 m×20 m
3A	20 m×20 m	3×3	4, 6, 8	h, s, a, p, pc	focal statistics-median (3 × 3)	20 m×20 m
3B	20 m×20 m	3×3	4, 6, 8	h, s, a, p, pc	block statistics-median (3 × 3)	60 m×60 m
4	20 m×20 m	5×5	4, 6, 8	h, s, a, p, pc	focal statistics (5 × 5)	100 m×100 m
5	20 m×20 m	7×7	4, 6, 8	h, s, a, p, pc	median filtering (7 × 7)	20 m×20 m

Notes: * h—true height, s—slopes, a—aspect, p—profile curvature, pc—plan curvature

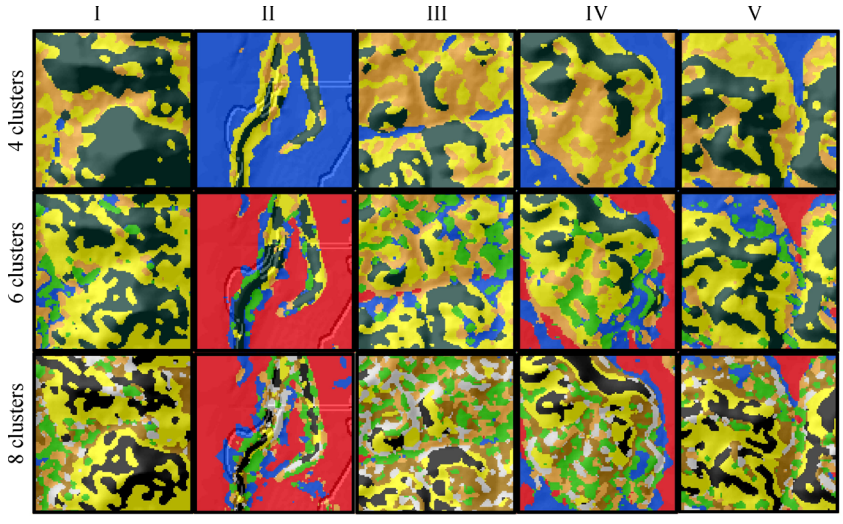


Fig. 3 Results of experiment no. 5 (I-V numbers of test fields).

Upland are described for each cluster and a list of all distinguished landforms is presented in Table 4. To define the distinguished landforms in the clusters we applied a proposal of morphological type (topographic position) classes of Speight (1990) and, additionally, the latest geomorphological map (Jania et al., 2014) which included approximately 20% of the study area.

At the beginning, one has to state that clusters with the same name (A, B, C, etc.), but in different divisions (4, 6, 8 clusters) do not correspond to each other exactly (e.g., cluster B within 4 clusters does not match cluster B within 6 or 8 clusters). These clusters may be more or less similar, but will never be exactly the same due to the differences in the method of calculating the similarity (i.e., similarity is defined differently for a smaller number of clusters than for a bigger one).

Regardless of the number of clusters, the application of the *k*-median method showed proper differentiation of the Silesian Upland morphology. Landform classification reflected the main features of the study area and the conducted calculations allowed to draw up their quantitative characteristics. Although the clusters consist of different landforms (see Table 4), each cluster is a set of fragments of the study area with similar mathematical characteristics, which resulted from the calculated variables.

The whole set of data contains 6 rasters of 16.5 million pixels each and is divided into 4, 6, and 8 clusters respectively. Table 5 contains general statistics of each variable in a particular cluster. Regarding aspect, we can see that each cluster has the same range of data from 0 to 360 degrees. Differences can be observed in the mean

Table 4 Landform classification in divisions into 4, 6, and 8 clusters

Cluster name	Distinguished landforms and occupied area					
	4 clusters	(%)	6 clusters	(%)	8 clusters	(%)
A	- plains - wide depressions - big valleys - valley bottoms	43.6	- plains	28.6	- lower-slopes - plains	23.6
B	- ridges and peaks - steep slopes - deep valleys escarpments	11.2	- upper-slopes - steep slopes - ridge-lines escarpments	5.5	- steep slopes - fragments of ridges	4.1
C	- incisions - flow-lines - thin valleys - lower-slopes	22.9	- flow-lines - small valley bottoms - lower-slopes	10.7	- lower-slopes - valley bottoms	11.8
D	- lower-slopes	22.3	- slopes	27.6	- plains	25.1
E	-	-	- mid-slopes	12.8	- mid-slopes	7.1
F	-	-	- low ridges - peaks	14.9	- hill-tops - peaks	4.8
G	-	-	-	-	- lower-slopes	13.0
H	-	-	-	-	- small hills - little ridges	10.5
Total:		100.0		100.0		100.0

value and standard deviation according to the rule that the more classes are used, the lower the standard deviations can be observed while the differences between mean values are more significant.

The minimum values of the DEM in all clusters are very similar and the maximum values exceed 400 m, excluding cluster A in the divisions of 4 and 6 clusters. Clusters A, B, C, and D have a similar range, mean value, and standard deviation in all divisions. This tendency does not exist for true height and slope.

Profile and plan curvature values that concentrate around zero are independent of cluster and division, although the largest dispersion of values is visible in clusters B and E, which is indicated by the standard value in these clusters. Considering all divisions and variables, cluster B is the least homogeneous one.

The above relief classification (due to its specificity, i.e., mathematical homogeneity) indicates the elements of landforms (i.e., lower-slope, fragments of upper ridges, etc.) rather than entire landforms in the classic approach. In this perspective, classification of relief can be used as a quick and objective tool to create maps presenting spatial relations between relief elements and quantitative characteristics of individual clusters (see Table 3).

5.1.1 Results for 4 clusters

Observations from cluster A (Fig. 4(b)) have the smallest mean height (256 m) and true height (24 m), very small slopes (mean 0.5°), curvature nearly 0, and a north-western aspect. This cluster includes flat terrains, — plains and/or very wide depressions, which are located in large valleys and valley bottoms (Mała Panew Valley, Liswarta Basin, Górna Warta Basin). Cluster A appears with the highest frequency and accounts for almost half of the area (43%).

Cluster B has the highest mean height (292 m) and true height (35 m), very large slope values (max 44°, mean 3.7°), curvatures slightly below and above zero, and a north-eastern aspect. Cluster B represents the highest areas of the Silesian Upland — mainly ridges and peaks (Mid-Triassic Escarpment, Tarnowskie Góry Ridge, Jaworzno Hills, Mikołów Edge). There are also steep slopes in deep valleys and escarpments. This cluster occupies 11% of the area.

Cluster C has a mean height of 277 m and true height of 36 m. Slope values are the largest (max 45°, mean 2.0°), curvatures are around zero, and the cluster has a north-eastern aspect. It follows incisions, flow-lines, and narrow valleys (or valley axes). There are also lower-slopes in places. This cluster constitutes 23% of the area.

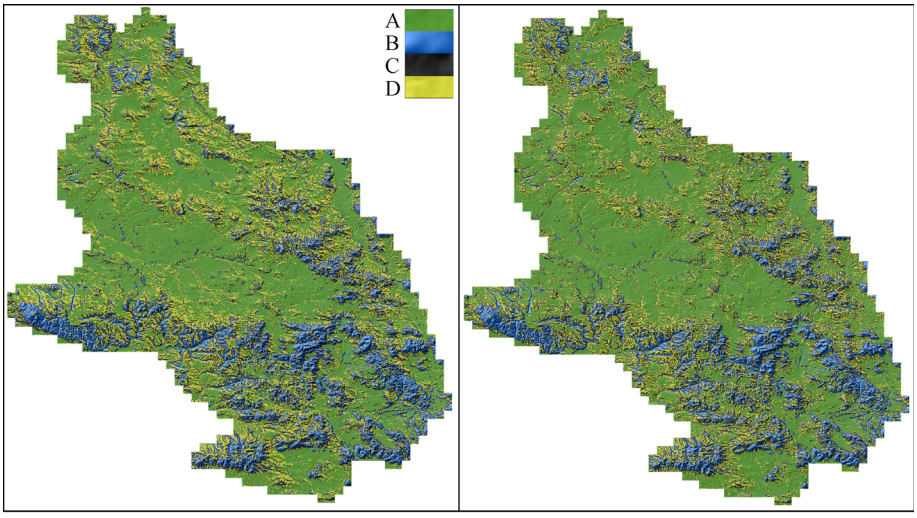
Cluster D includes observations with the mean height of 269 m and true height of 36 m. Mean slope values are 1.2°, curvatures around zero, and it has a north-eastern aspect. Cluster D represents lower-slopes, adjacent to high ridges and hills. This cluster occupies 22% of the area.

5.1.2 Results for 6 clusters

In the classification into six clusters (Fig. 5(b)), observations from Cluster A (just like those divided into 4 clusters)

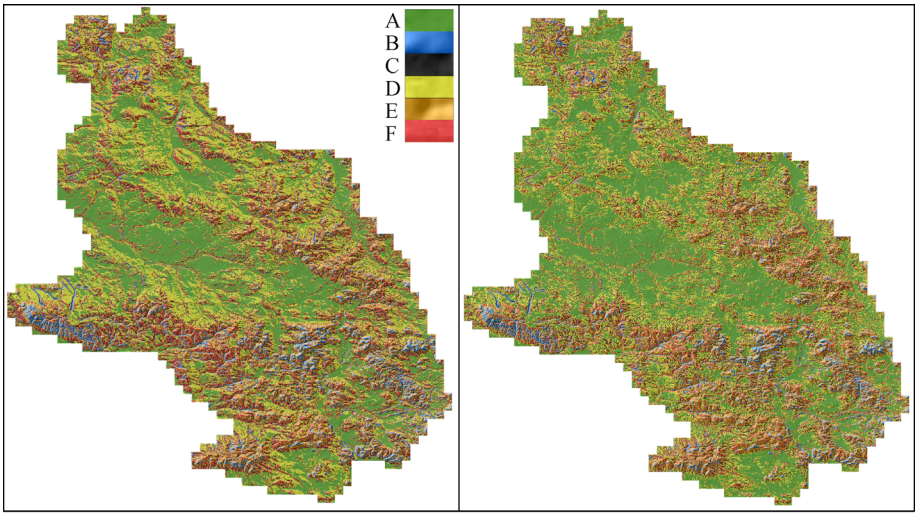
Table 5 Descriptive statistics for division into 4, 6, and 8 clusters

Variable	4 clusters				6 clusters						8 clusters							
	A	B	C	D	A	B	C	D	E	F	A	B	C	D	E	F	G	H
Aspect																		
Min	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Max	360	360	360	360	360	360	360	360	360	360	360	360	360	360	360	360	360	360
Mean_vector	324	40	36	56	17	48	32	55	43	51	286	43	48	56	40	77	54	1
Std	101.9	101.7	96.5	113.5	91	102	85	120	114	83	63.3	108.9	84.6	130.0	102.9	88.4	121.3	77.3
DEM																		
Min	176.0	177.1	177.8	177.0	176.0	178.3	177.5	176.8	177.1	177.8	176.0	180.1	177.5	176.5	177.2	178.6	177.8	178.5
Max	378.8	436.3	415.6	404.9	378.5	436.3	403.3	403.6	425.2	425.0	401.9	436.3	414.6	402.7	422.8	426.3	404.9	397.3
Mean	256.1	292.0	277.4	269.2	254.5	290.4	272.3	263.0	288.7	273.5	255.4	289.8	273.3	258.6	293.4	287.8	276.7	266.5
Std	32.3	42.3	37.7	34.3	32.3	45.6	36.1	32.5	38.8	37.1	33.5	46.3	36.9	31.1	40.4	38.6	35.8	34.2
True height																		
Min	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Max	23.7	35.5	36.5	36.5	23.7	35.7	14.5	17.1	35.8	36.5	35.7	29.0	16.0	35.7	35.8	15.7	36.5	16.8
Mean	0.4	3.5	1.5	1.2	0.3	4.2	1.0	0.8	2.4	1.5	0.4	4.5	1.2	0.6	2.9	2.3	1.6	1.0
Std	0.3	1.9	0.9	0.5	0.3	2.4	0.5	0.4	1.0	0.6	0.4	2.4	0.7	0.4	1.4	0.6	0.6	0.3
Slope																		
Min	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Max	33.2	43.7	45.1	36.5	33.2	41.8	7.3	41.4	45.1	29.3	41.8	32.9	12.5	41.8	43.7	11.4	45.1	6.0
Mean	0.5	3.7	2.0	1.2	0.4	4.4	1.4	1.0	3.0	1.4	0.5	4.4	1.6	0.7	3.8	2.1	1.8	1.0
Std	0.4	2.2	1.1	0.5	0.3	2.6	0.5	0.5	1.3	0.7	0.4	2.53	0.8	0.4	1.9	0.7	0.6	0.4
Plan																		
Min	−9.80	−10.27	−10.69	−7.93	−4.10	−10.27	−1.85	−7.93	−10.69	−8.03	−8.03	−10.27	−1.37	−7.93	−10.69	−1.64	−9.43	−0.85
Max	7.58	7.62	6.16	4.97	6.15	6.07	0.82	4.97	7.62	4.02	4.86	5.36	0.89	6.15	7.62	0.97	6.16	0.46
Mean	0.00	0.02	−0.01	0.01	0.00	0.03	−0.02	0.00	0.00	0.01	0.00	0.03	−0.01	0.00	−0.01	0.03	0.00	0.01
Std	0.03	0.10	0.05	0.04	0.02	0.12	0.04	0.03	0.08	0.04	0.02	0.12	0.05	0.03	0.10	0.04	0.05	0.03
Profile																		
Min	−5.83	−10.48	−9.11	−8.80	−5.83	−8.69	−1.26	−4.08	−10.48	−5.65	−8.69	−5.47	−1.64	−8.69	−10.48	−1.48	−8.80	−1.09
Max	8.03	8.81	8.89	8.06	7.37	8.81	1.16	8.81	8.89	7.49	8.78	7.70	1.88	8.81	8.89	1.57	8.35	0.69
Mean	0.00	−0.02	0.02	−0.01	0.00	−0.03	0.02	0.00	0.02	−0.01	0.00	−0.05	0.02	0.00	0.04	−0.03	0.00	−0.01
Std	0.05	0.15	0.07	0.05	0.04	0.18	0.04	0.04	0.12	0.06	0.05	0.18	0.05	0.06	0.15	0.05	0.07	0.03



(a) (b)

Fig. 4 Results for 4 clusters without aspect (a) and with aspect (b).



(a) (b)

Fig. 5 Results for 6 clusters without aspect (a) and with aspect (b).

refer to the lowest places of the study area (max 378 m, mean 254 m) and true height (24 m). This cluster includes flat areas and plains, as confirmed by the lowest mean slopes (0.4°) and both curvatures with zero values. A NNE aspect (17°) dominates in these areas. This class again occupies the largest area of all clusters (29%).

Cluster B comprises the highest and the steepest parts of the study area, with absolute heights of 436 m, a mean of 290 m and local relief of 36 m. It represents mainly higher-slopes, ridge-lines, and escarpments with a mean aspect vector toward the NE (48°). These observations have the highest slope values of all clusters (maximum 48° and mean 4.4°). This cluster occupies merely 5% of the area.

Cluster C is formed by flow-lines, small valleys, and sometimes lower parts of slopes with maximum heights of 403 m, the mean of 272 m, and true heights of only 14 m. The dominant aspect vector is to the NE (32°), and maximum slope values are merely 7°. This cluster occupies around 11% of the area.

Gentle slopes (mean 1°) of a NE aspect (55°), which highlight the greatest elevations of the area about the direction of NW-SE (Fig. 5), occur in Cluster D. Both the mean height (263 m) and true heights (17 m) are small. The plan and profile curvatures are zero, and the whole cluster occupies over one quarter of the area (27%).

The last two clusters (E and F) often appear side by side: Cluster E mostly consists of mid-slopes, and Cluster F is formed by low ridges and peaks. Both have the same NE mean aspect. Both clusters are present at similar heights (mean 273–288 m) and true heights (36 m). Mean slopes in

Cluster E (3°) are greater than in Cluster F (1.4°). Cluster F has a curvature of zero, while Cluster E — with slightly positive profile curvatures, probably indicates planar hillslopes. Both clusters cover slightly more than one quarter of the total area (28%).

5.1.3 Results for 8 clusters

As a final option the area was divided into eight clusters (Fig. 6). As in the divisions above — Cluster A consists of flat areas and lower-slopes, mainly west-oriented (mean 286°). Again, these are the lowest-lying areas (mean 255 m), although they have a significant local relief (36 m). The mean slope is 0.5° with a curvature of zero, which agrees with the flatter lower-slopes. This cluster covers 23% of the area.

Cluster B forms the smallest area of all clusters (only 4%), because it consists of steep slopes occurring in narrow valleys and fragments occur from the west to eastern large ridges (Fig. 6**)**. This cluster reaches the maximum height (436 m) for the study area (the average height amounts to 289 m with a predominant NE (43°) aspects). This cluster has the highest average true heights (4.5 m) and the highest mean slopes (above 4°). It shows slight positive plan curvatures and negative profile curvatures, which confirms the nature of narrow ridges and rock ribs.

Observations from Cluster C include lower-slopes and narrow valley bottoms covering about 12% of the area. The mean aspect has a direction towards the NE (48°). These forms range from the lowest to the highest places in the study area, and the local relief is only 16 m. These represent flatter areas (with maximum slopes of only 12.5°) and concave/convex areas (as indicated by slightly negative plan curvatures and positive profile curvatures).

Cluster D is the most predominant cluster in the area (25%). This cluster represents plains (mean slope 0.7°) with a NE aspect (56°). This truly emphasizes the NW-SE course of the most significant elevations within the study area. Average heights (258 m) within Cluster A are the lowest in this division and zero curvatures are further supported by the flat and balanced character of these forms.

Mid-slopes form Cluster E most often accompany the ridges from Cluster B. These forms are of a similar height to Cluster B (mean 293 m) and have the same aspect NE (40°). The mean slopes (4°) are also similar. Clearly, positive profile curvatures and a little negative plan curvatures suggest slightly concave slopes and their foothills. This cluster occupies a very small area (7%).

Forms of Cluster F occupy the smallest area (less than 5%) consisting of hill-tops and peaks of a mainly eastern aspect (77°), arranged from the lowest to highest altitudes within the whole research area. The maximum slopes total 11°, which suggests flat summits and plateaus. Negative profile curvatures and positive plan curvatures clearly indicate hill tops.

Cluster G reaches the height of 277 m and is NE exposed (54°). It contains lower-slopes that gently transform into plains (mean slope 1.7°, and curvatures of zero). This cluster occupies an area of 13%.

The final cluster (H) is created by small hills and little ridges that usually accompany the valleys of Cluster C. This is the only cluster of landforms that has a clear northern aspect (1°). The shape of curvatures is similar to Cluster F, but these landforms are 20m below (266 m) the average. Maximum slopes are only 6° with the average of 1°, which proves, beyond doubt, the flat nature of the summits. This cluster represents 10% of whole area.

5.1.4 Reduction to 4 and 6 clusters

As presented in the above figures, the division into 4

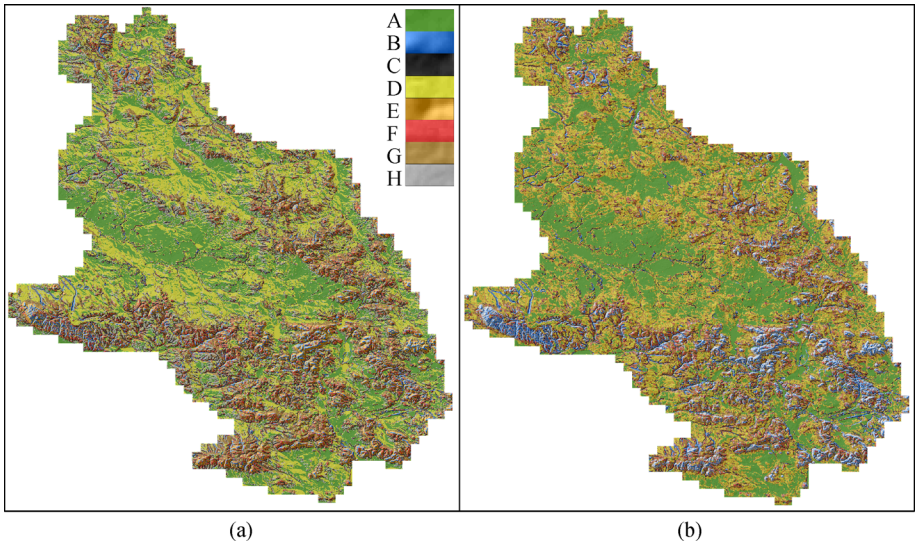


Fig. 6 Results for 8 clusters without aspect (a) and with aspect (b).

clusters provides the clearest picture of morphology. The most fundamental landforms are conspicuous here: wide land depressions and river valleys (green), peaked ridges protruding over relief (blue), and accompanying lower slopes transforming into slightly inclined wide depressions (yellow). The relief is also well represented in the division into 6 clusters, as apart from the above-mentioned elements, it includes slopes and lower-slopes of a NE aspect, which effectively emphasize the outline of ridges.

It was therefore decided to verify whether the reduction of 6- and 8-cluster divisions into a 4-cluster one (Fig. 7), and an 8-cluster division into a 6-cluster one (Fig. 8) would produce similar results. The reduction of 6 clusters to 4 ones (Fig. 7(b)) increased the share of ridges and culminations (blue) almost threefold, while the number of lower-slopes and slopes rose by nearly 1/4 (yellow), at the expense of wide depressions (green). Moreover, the share of flow-lines and incisions (black) decreased by half. For the reduction of 8 clusters to 4 (Fig. 7(c)), the layout and surface of ridges (blue), as well as flow-lines and incisions (black) remained almost the same. Similarly, the share of lower-slopes (yellow) increased at the cost of wide depressions (green). Cutting down 8 clusters to 6 ones (Fig. 8) gave almost identical results with an increased number of steep slopes and fragments of ridges (blue) in the central and north-western part of the area.

To sum up, qualitative generalization (decreasing the number of clusters) did not alter the spatial morphological image of the examined area. The range of occupied areas changed but spatial relations involving general character of

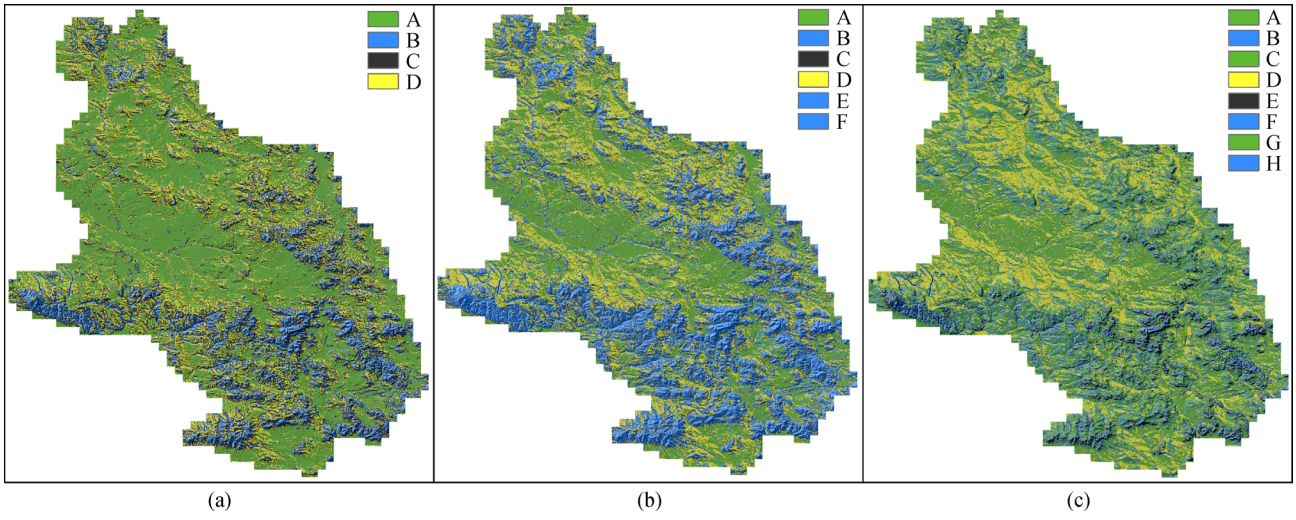


Fig. 7 Results for 4 clusters (a), reduction from 6 to 4 clusters (b) and reduction from 8 to 4 clusters (c).

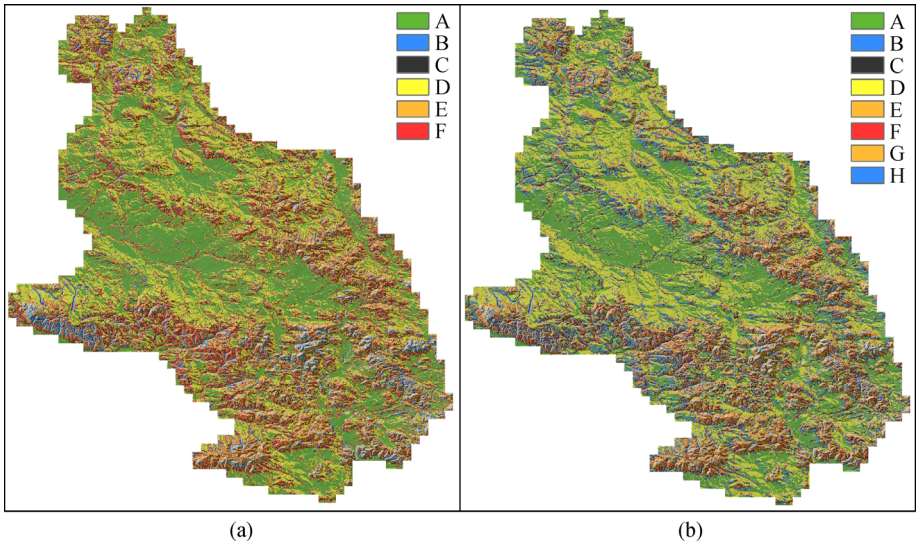


Fig. 8 Results for 6 clusters (a) and reduction from 8 to 6 clusters (b).

relief remained the same. It proves that all 3 divisions (into 4, 6, and 8 clusters) classified landforms correctly. An increased number of clusters provides a more detailed image (distinct lower-slopes and slopes), however, it requires an analysis in a larger scale, which is inconvenient for such a vast area (over 6000 km^2).

5.2 Influence of aspect on the results

In previous studies, Wieczorek (2008) noticed that the influence of aspect on the final results should be taken into account. Clearly, aspect is one of the key elements that have an effect on results. Table 6 shows how aspect affects the area occupied by each cluster. There are no significant differences in the division into 4 clusters, with or without the use of aspect. There are small changes in the percentage share of individual clusters, like their spatial location (Fig. 4).

When divided into 6 clusters, we can notice a distinct change in spatial distribution (Fig. 5). The results without aspect show a strong dispersion of Cluster D and with the use of aspect — the pattern and density of Cluster D is concentrated in the areas with a north, north-eastern, and north-western aspect, as well as flat areas (no aspect). In addition, the surface area of Cluster F increased by half, which clearly highlights the course of low ridges.

The largest quantitative changes occur in the areas divided into 8 clusters (Table 6). In the analysis with the use of aspect, Cluster D increased from 14% to 25%, while Cluster E decreased its surface area almost threefold (from 20% to 7%). Again, Cluster D is concentrated very strongly in the N, NW, and NE directions (Fig. 6).

We need to emphasize that aspect is one of the elements of landform morphometry that should be taken into account. It can be observed that the use of aspect in relief analysis with the *k*-median method results in a more orderly and spatially coherent picture of landforms (especially when divided into 6 clusters). This is why an analysis with aspect provides a more complete picture from a geomorphological point of view.

6 Discussion

6.1 Results comparing *k*-median vs TPI method

To calculate the errors of a method, one needs to compare the model values with the observed patterns, which are not available here. We do not have a pattern here, so we cannot compare it. Therefore, an expert assessment of the results obtained is needed. Another way to quantify the obtained results is to compare them with the results of the classification made by another method. We decided to use the TPI method — well-known and simple, in which expert knowledge is indispensable. The TPI forms the basis of the classification system and is simply the difference between the cell elevation value and the average elevation of adjacent cells. Positive values mean the cell is higher than the surrounding ones, while negative values mean that it is lower (Weiss, 2001). The Topographic Position Index is used for classifying the landscape into 3, 4, or 6 broad topographic positions. We used the extension for ArcGIS (Jenness et al., 2013) to calculate options for 4 and 6 classes.

Figure 9 represents a test area divided into 4 and 6 clusters classified by the two methods. An overall picture of the relief looks similar. The key elements of the morphology of this area have been well highlighted: a wide river valley with a water reservoir in the east and a cluster of hills to the NE, N, and NW. Unfortunately, the entire south-western section of the area looks much worse when analyzed with the TPI method. It is not an easy terrain to analyze statistically, but the image resulting from the *k*-median method looks clearer. The situation is similar in the division into 6 clusters. The TPI image leaves only the most obvious relief elements and omits all the rest.

The analysis performed using the TPI method shows a simplified picture of relief elements classification. It is inherent in using a very simple algorithm associated with height differences. Many nuances or details of morphology are missed. The *k*-median method is based on a multivariate analysis, i.e., a greater number of factors affect the

Table 6 Aspect influence on cluster divisions

Cluster	4 clusters (% of area)		6 clusters (% of area)		8 clusters (% of area)	
	without aspect	with aspect	without aspect	with aspect	without aspect	with aspect
A	45.3	43.6	33.7	28.6	29.5	23.6
B	10.6	11.2	4.5	5.5	4.7	4.1
C	23.9	22.9	11.5	10.7	7.9	11.8
D	20.2	22.3	27.8	27.6	14.4	25.1
E	–	–	12.9	12.8	20.2	7.1
F	–	–	9.6	14.9	5.3	4.8
G	–	–	–	–	11.6	13.0
H	–	–	–	–	6.5	10.5

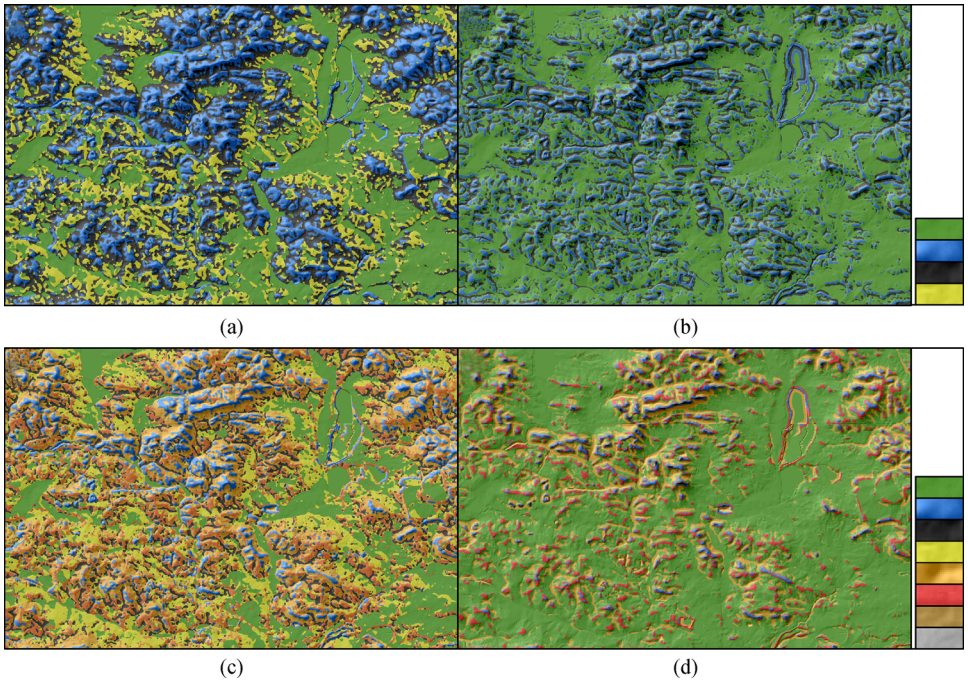


Fig. 9 Results of *k*-median for 4 (a) and 6 (c) clusters and TPI for 4 (b) and 6 (d) classes in the test area.

results, with the final effects being more volatile. In both examples (4 and 6 clusters) using the *k*-median method, Cluster D (yellow) shows lower and gentle slopes. This cluster is absent using the TPI method, which significantly simplifies the picture of relief.

Table 7 shows a quantitative comparison of distinguished clusters (classes) with both methods. As can be seen from Table 7 and Fig. 9, the TPI method passes over gentle slopes (yellow) in 4 classes and heavily underestimates the areas of ridges (blue) and incisions and lower-slopes (black) in 6 classes. Both methods have different methodological assumptions and it seems, that for a thorough geomorphologist who is provided with a high-resolution DEM, the *k*-median method proves to be better for relief research. Landform classification is clearly better, especially for less evident forms (gentle-slopes and lower-slopes) or those sometimes hidden in larger forms (small valley bottoms and flow-lines and incisions between ridges). The TPI method loses these kinds of forms, classifying (combining) them together with wide depressions and plains. Probably in this situation of extracting less distinct forms, the only comparison of elevation differences (which is the basis of the TPI method) is not enough. The *k*-median method deals much better with this issue.

Table 7 Comparison of areas occupied by *k*-median clusters and TPI classes

Clusters/classes no.	*k*-median method	TPI method
	4 clusters (%)	4 classes (%)
1 (black)	32.8	22.0
2 (green)	28.0	57.6
3 (blue)	20.9	20.4
4 (yellow)	18.4	0.03
	6 clusters (%)	6 classes (%)
1 (green)	18.9	72.2
2 (blue)	10.8	3.0
3 (black)	12.8	0.4
4 (yellow)	19.0	0.01
5 (orange)	23.2	14.0
6 (red)	15.3	10.4

6.2 *k*-median method in relation to geology and geomorphology

Comparing cluster results using the *k*-median method with geological 1:200000 maps (Haisig and Wilanowski, 1979; Kotlicka and Kotlicki, 1979; Biernat et al., 1980; Kaziuk and Lewandowski, 1980) did not provide satisfactory results. Generally, in many places there is some concurrence. Cluster A is found mainly in Pleistocene and Holocene alluviums (sands, gravels, silts) but, unfortunately, in other types of rocks and sediments as well. Similarly, Clusters B and C are usually located in places of resistant rocks (Mid-Triassic dolomites, limestones, and marls) and Carboniferous rocks (clays, mudstones, sandstones), but also within less resistant Pleistocene loess covers. Cluster D is concentrated in Pleistocene sediments (clays and fluvioglacial sands and gravels) but also occurs

in Mid-Triassic rocks (limestones and marls). This inconsistency is understandable if we believe that the Silesian Upland relief is polygenetic (double glaciation: Mindel and Riss I and intensive human impact on the relief over the past 200 years) (Gilewska, 1999a and 1999b) and, thus, the shape of forms may be similar, despite their different origins.

The results achieved with the *k*-medians for 4 clusters are consistent with the hypsometric image and fairly well describe the morphological differentiation of the Silesian Upland area. Cluster A refers to aligned areas (plains), Clusters B and C are associated with the most important elevations in the relief (ridges, hills, and summits), while Cluster D describes the lower parts of slopes. The comparison of this data with classic geomorphological 1:50000 maps (Klimaszewski, 1959; Jania et al., 2014) and with detailed subdivision into geomorphological regions at a scale of 1:100000 (Karaś-Brzozowska, 1960) shows surprising compliance. Although the *k*-median method refers to the shape and measurements of landforms (defined by variables) — an outline of the clusters corresponds to the above referenced geomorphological maps. A comparison of the *k*-median results with a map of a smaller scale, i.e., 1:500000 (Starkel, 1980) gives poor results, because a map at this scale contains much less detail than the results for the four clusters.

In general, we can say that the *k*-median analysis divided into four clusters is the easiest to read. The distinguished clusters leave no doubt as to the assignment of specific landforms (see Table 4) and closely refer to the hypsometric image of the Silesian Upland (see Figs. 1 and 2). It seems that the most valuable result from a geomorphological point of view is the division into six clusters. The final cartographic picture is evidently less diversified than in geomorphological maps but shows the principal features of morphology fairly well (see Wieczorek and Migoń, 2014). In this case, the expansion of the relief image distinguished low ridges, peaks, mid-slopes, and small valley bottoms that are not present in the 4 clusters division. The division into 8 clusters definitely distorts the picture of the relief elements over such a large area and makes the image unreadable. Certainly, everything depends on the scale at which we consider a given area but the Silesian Upland, with the surface area of about 6500 km^2, should be considered at a scale of 1:250000 – 1:100000.

6.3 Benefits and faults of *k*-median method

The *k*-median method, like any computer calculation method, has its own strengths and weaknesses. The undeniable advantages include its quantitative character, which eliminates qualitative (subjective) judgments on research to a great extent. Another important feature is its repeatability and, thus, the opportunity to work on data sets describing different research areas, which allows comparative studies between them. One has to note, however, that it is a statistical method, i.e., all objects will not always be allocated to precisely the same clusters. Assignment to a cluster depends on intercluster variability in the data set and the primary division into clusters. In the case of a morphometric data set containing hundreds of thousands of observations (pixels) receiving exactly the same division twice is highly unlikely. However, the differences between the results obtained with the same method parameters are negligible when we get a similar value of intercluster variability. A positive feature of the *k*-median method is also the application of a clearly defined mathematical algorithm (Wieczorek and Migoń, 2014) and minimum requirements for input data. The only needed tool is a hydrologically correct DEM (obviously, vertical and horizontal spatial resolution determines the accuracy of any further analysis).

In the Drăguț and Blaschke (2006) and Piloyan and Konečný (2017) articles the areas were divided in 9 classes according to Dikau (1989) delimitation. The *k*-median method provides an opportunity to obtain different numbers of classes and the statistics indicate the best clustering result and allows us to find the proper number of classes in the Silesia Upland for a different application. In both above mentioned works, the classification was made without information about aspect. Although Piloyan and Konečný (2017) agreed that aspect is an important derivative of elevation, aspect was not included in basic information sources because of its nonlinear measurement method. The study presented in this article shows how to solve the problem with aspect and use this basic data in clustering process.

The best result of clustering in the Silesia Upland was obtained for 6 clusters, which corresponds with the number of clusters in a mountain area presented in the Wieczorek and Migoń (2014) article. Both studies were made using DEM with similar resolution (25 m in the Sudetes and 20 m in the Silesia Upland). The spatial size was similar (125 m × 125 m in the Sudetes and 140 m × 140 m in the Silesia Upland) although a different number of cells in a filter window were used (5 × 5 in the Sudetes and 7 × 7 in the Silesia Upland). It shows that the method could be useful not only in a mountainous area. However, morphometric classifications are prepared for the whole counties or mountain areas, which has been proved by many articles (Dikau et al., 1991; Drăguț and Blaschke, 2006; Wieczorek and Migoń, 2014; Piloyan and Konečný, 2017)

Undoubtedly, a distinguishing feature of the *k*-median method should be its multidimensional character, i.e., a simultaneous analysis of multiple variables and, consequently, the ability to determine the similarity of the elements making up the cluster on the basis of the achieved results. In addition, the *k*-median method is a modification of the *k*-means method — a known, used (see Niemann and Howes, 1991; Arrell, 2001; Burrough et al., 2001; Arrell

et al., 2007; Khan, 2012), and tested calculation method available in many GIS packages. As it turns out, relief (height adjustment) or morphometric variables as a set of mathematical data do not have a symmetric distribution[1)], so the value that best describes them is the median rather than average (see Evans, 1972; Evans and Cox, 1999).

The key concern is whether to give weight to each morphometric variable or to pay attention to the characteristic (most important) features of distinguished landforms.

However, the *k*-median method also has its drawbacks and limitations. For example, the selection of well-defined parameters — both morphometric and mathematical (e.g., filtering window size) will affect the results. This situation makes us carry out a series of tests in order to determine the optimal settings that will significantly affect the achieved results. Thus, a researcher's judgement of the settings is extremely important. The experiments conducted in the test fields have proved very helpful. According to the authors, this is a necessary process from which one should start such kind of research.

7 Conclusions

Our research led to the following conclusions:

The results of relief classification using the *k*-median method properly show general topography diversification of the Silesian Upland. Due to spatial DEM resolution (20 m × 20 m) this classification corresponds to 1:50000–1:25000 map scales. A smaller number of clusters 4 clearly shows the key elements of the morphology of the area. The division into 6 clusters retains a clear image and provides more details to the image of the morphological diversity by distinguishing characteristic elements of landforms (e.g., steep slopes of a NE aspect, small valley axes, incisions and flow-lines, etc.). The results for the 8 clusters bring on significant data noise and are ambiguous (i.e., similar landforms correspond to different clusters), which makes them difficult to identify. Unfortunately, the *k*-median results failed to compare so well with the geological data, because the method used quantitative, not qualitative, data.

Using aspect as one of the variables was reasonable. The influence of aspect on quantitative results for 4 and 6 clusters is low, but it leads to a much better spatial image. Using aspect well highlights the general slope of the Silesian Upland area to the NE and the course of the most important elevations (long ridges) in a NW-SE direction.

The *k*-median method is a multidimensional analysis. The principle of such an analysis consists in a simultaneous examination of the entire data set. This information and observations provide evidence that, for the same settings and using the *k*-median method, the final results depend heavily on the extent and amount of analyzed data. It is important that the studies take into account the entire geomorphological unit (herein, the Silesian Upland), not just a selected part because the *k*-median method is sensitive to input parameters and the results may not be satisfactory and meaningful. To maximize the correctness of the results, one should choose the size of the study area to make it representative for the given landforms that are to be distinguished. Such a procedure will allow us to capture geomorphological homogeneity of the area.

Using the *k*-median method along with reducing the size of the filtering window increases the degree of the classification accuracy and vice versa. One has to bear in mind that a too-small filtering window cannot recognize major landforms. Following different attempts and experiments, it seems that for such a large area (6500 km^2) the most optimal filtering window size is no less than 7 × 7 cells for a DEM of a resolution 20 m × 20 m.

Although the *k*-median method has its advantages (quantitative character, repeatability, objectivity, and minimal requirements for input data), it may also have disadvantages (quality of input data, area extent, input settings). On the basis of the conducted research it can be argued that the *k*-median method is a fast mathematical tool for allocating landforms (landform complexes) into clusters (groups) of similar morphographic (shape) and morphometric (measurements) properties, which describe the relief of an area correctly and can be used in geomorphological examinations.

References

Alonso-Sarría F, Gomariz-Castillo F, Cánovas-García F (2018). A new approach to the openness index for landform characterization. Comput Geosci, 119: 68–79

Arrell K E (2001). A fuzzy k-means classification of elevation derivatives to extract the natural landforms in Snowdonia, Wales. In: Proceedings of 9th National Conference on GIS Research UK (GISRUK 2001)

Arrell K E, Fisher P F, Tate N J, Bastin L (2007). A fuzzy c-means classification of elevation derivatives to extract the morphometric classification of landforms in Snowdonia, Wales. Comput Geosci, 33 (10): 1366–1381

Azanon J M, Delgado J, Gómez A (2004). Morphological terrain classification and analysis using geostatistical techniques. In: Proceedings of ISPRS Congress. Istanbul, 12–23

Biernat S, Haisig J, Lewandowski J, Wilanowski S (1980). Geologic Map of Poland 1:200000, sheet Częstochowa. Warszawa: Instytut Geologiczny

Broersen T, Peters R, Ledoux H (2017). Automatic identification of watercourses in flat and engineered landscapes by computing the skeleton of a LiDAR point cloud. Comput Geosci, 106: 171–180

Bukowska-Jania E (1983). Contemporary fluvial processes in the eastern part of Silesian Upland. Dissertation for the Doctoral Degree, Wrocławski: Uniwersytet Wrocławski (in Polish)

1) The exception is curvature which is characterized by a strongly peaked distribution and symmetry.

Burrough P A, Wilson J P, van Gaans P F M, Hansen A J (2001). Fuzzy *k*-means classification of topo-climatic data as an aid to forest mapping in the Greater Yellowstone Area, USA. Landsc Ecol, 16(6): 523–546

Chmal H (1976). Processes of the erosion forms development on the dumps of the coal mining in the Upper Silesian Basin. Dissertation for the Doctoral Degree, Wrocławski: Uniwersytet Wrocławski (in Polish)

Czajka W (2009). Database of the terrain elevations in DTED format. Kwartalnik BELLONA-90 lat geografii wojskowej (wydanie specjalne). MON, Warszawa, 26–30 (in Polish)

Dekavalla M, Argialas D (2017). Object-based classification of global undersea topography and geomorphological features from the SRTM30_PLUS data. Geomorphology, 288: 66–82

Deng Y (2007). New trends in digital terrain analysis: landform definition, representation and classification. Prog Phys Geogr, 31(4): 405–419

Dikau R (1989). The application of a digital relief model to landform analysis. In: Raper J F ed. Three dimensional applications in Geographical Information Systems. London: Taylor and Francis, 51–77

Dikau R, Brabb E E, Mark R M (1991). Landform classification of New Mexico by computer. Open File Report 91–634. U.S Geological Survey. 15

DMA (Defense Mapping Agency) (2000) Performance specification digital terrain elevation data (DTED)

Drăguț L, Blaschke T (2006). Automated classification of landform elements using object-based image analysis. Geomorphology, 81(3–4): 330–344

Drăguț L, Csillik O, Minár J, Evans I S (2013). Land-surface segmentation to delineate elementary forms from Digital Elevation Models. Geomorphometry, 16–20

Drăguț L, Eisank C (2011). Object representations at multiple scales from digital elevation models. Geomorphology (Amst), 129(3–4): 183–189

Drăguț L, Eisank C (2012). Automated object-based classification of topography from SRTM data. Geomorphology (Amst), 141–142(4): 21–33

DTED-2 (2001). Digital Elevation Model of Poland level 2. Warszawa

Dulias R (1994). Documentation of the group of the aeolian landforms between Woszczyce and Kleszczówka. Sosnowiec (in Polish)

Dulias R (1995). Dunes of the southern part of the Silesian Upland. In: Proceedings of III Zjazd Geomorfologów Polskich: procesy geomorfologiczn,. Sosnowiec. 1, 19–20 (in Polish)

ESRI (Environmental Systems Research Institute) 2017. ArcGIS Desktop: Release 10.5. Redlands, CA

Evans I S (1972). General geomorphometry, derivatives of altitude and descriptive statistics. In: Chorley R ed. Spatial Analysis in Geomorphology. London: Methuen and Co., 17–91.

Evans I S, Cox N J (1999). Relation between Land Surface Properties: Altitude, Slope and Curvature. In: Hergarten S, Neugebauer H J eds. Process Modelling and Landform Evolution. Heidelberg: Springer, 13–45

Galon R, ed. (1972). Geomorphology of Poland vol. 2. Warszawa: PWN (in Polish)

Gilewska S (1963). Relief of the Mid-Triassic escarpment in the vicinity of Będzin. IG PAN, Prace Geograficzne nr 44, Warszawa: Wydawnictwa Geologiczne, 119 (in Polish)

Gillewska S (1972). Silesian-Małopolskie Uplands. In: Klimaszewski M ed. Geomorphology of Poland vol. 1. Warszawa: PWN, 232–339 (in Polish)

Gilewska S (1986). Geomorphological subdivision of Poland. Przegląd Geograficzny, 58(1–2): 15–40 (in Polish)

Gilewska S (1999a). Relief. In: Starkel L ed. Geografia Polski. Geography of Poland. Natural environment. Warszawa: PWN, 243–287 (in Polish)

Gilewska S (1999b). Development of the environment of Poland in Tertiary. In: Starkel L ed. Geography of Poland. Natural environment. Warszawa: PWN, 38–66 (in Polish)

Gilewska S, Klimek M (1997). Relief origin and age. 1: 1500000. IGiPZ PAN, Atlas Rzeczypospolitej, Warszawa: PPWK (in Polish)

Guzzetti F, Reichenbach P (1994). Toward the definition of topographic divisions for Italy. Geomorphology 11:57–75

Haisig J, Wilanowski S (1979). Geologic Map of Poland 1:200000 sheet Kluczbork. Warszawa: Instytut Geologiczny

Hammond E H (1954). Small-scale continental landform maps. Ann Assoc Am Geogr, 44(1): 33–42

Hammond E H (1964). Analysis of properties in land form geography: an application to broad-scale land form mapping. Ann Assoc Am Geogr, 54(1): 11–19

Hornig A (1955a). Formy powierzchni ziemi stworzone przez człowieka na obszarze Wyżyny Śląskiej (Landforms made by human in the Silesian Upland area). In: Wrzosek A ed Górny Śląsk. Wydawnictwo Literackie, Kraków: 127–149 (in Polish)

Hornig A (1955b). On some monuments of inanimate nature of the Silesian Upland. Chrońmy przyrodę ojczystą 6: 8–18 (in Polish)

von Humboldt A (1849). Ansichten der Natur: mit wissenschaftlichen Erläuterungen. Stuttgart: J.G. Cotta'scher Verlag. 407

Hutchinson M F (1989). A new procedure for gridding elevation and stream line data with automatic removal of spurious pits. J Hydrol (Amst), 106(3–4): 211–232

Hutchinson M F (2011). ANUDEM Version 5.3. User Guide. Fenner School of Environment and Society, Australian National University

Iwahashi J, Pike R J (2007). Automated classification of topography from DEMs by an unsupervised nested-mean algorithm and a three-part geometric signature. Geomorphology, 86(3–4): 409–440

Jania J, Dulias R, Szypuła B, Tyc A (2014). Digital Geomorphological Map of Poland 1:100000, sheet Katowice. Poznań: GUGiK, Gepol

Jania J, Szczypek T (1980). An attempt to distinguish of the eolian sediments and landforms in the selected areas of the Silesian Highland by means of the photointerpretation. Fotointerpretacja w geografii 4: 25–40 (in Polish)

Jasiewicz J, Stepinski T F (2013). Geomorphons—a pattern recognition approach to classification and mapping of landforms. Geomorphology, 182: 147–156

Jenness J, Brost B, Beier P (2013). Land Facet Corridor Designer: Extension for ArcGIS. Flagstaff: Jenness Enterprises

Jorge M G, Brennand T A (2017). Semi-automated extraction of longitudinal subglacial bedforms from digital terrain models—two new methods. Geomorphology, 288: 148–163

Karaś C, Starkel L (1958). Extent of the Middle Polish glaciation in the southern part of the Silesian Upland) Przegląd Geograficzny, 30:

263–270 (in Polish)

Karaś-Brzozowska C (1960). Geomorphological characteristics of the Upper Silesian Industrial District. Warszawa: Biuletyn PAN (in Polish)

Karaś-Brzozowska C (1963). Extent of the Middle Polish glaciation in the Racibórz Basin. Przegląd Geograficzny, 35: 431–442 (in Polish)

Kaziuk H, Lewandowski J (1980). Geologic Map of Poland 1:200000 sheet Kraków. Warszawa: Instytut Geologiczny

Khan F (2012). An initial seed selection algorithm for *k*-means clustering of georeferenced data to improve replicability of cluster assignments for mapping application. Appl Soft Comput, 12(11): 3698–3700

Klimaszewski M (1947). Geomorphic map of the Southern Poland 1:1800000. Czas Geogr, 17: 133–182 (in Polish)

Klimaszewski M (1991). A geomorphological comparison of structural thresholds. Wrocław-Warszawa-Kraków: Dokumentacja Geograficzna (in Polish)

Klimaszewski M (1959) Geomorphological Map of the Uppersilesian Industrial Region, 1:50000. Warszawa: Komitet itd. GOP PAN (in Polish)

Klimaszewski M, ed. (1972) Geomorphology of Poland vol. 1. Warszawa: PWN (in Polish)

Klimek K (1966) Deglaciation of northern part of Silesia-Cracow Upland during the Middle-Polish glaciation. Warszawa: Prace Geograficzne IG PAN 53, 136 (in Polish)

Kondracki J (1951). Geomorphological map of Poland, 1:2000000. Przegląd Geograficzny 23 (in Polish)

Kondracki J (2001). Regional geography of Poland. Warszawa: PWN, 441 (in Polish)

Kotlicka G N, Kotlicki S (1979). Geologic Map of Poland 1:200000 sheet Gliwice. Warszawa: Instytut Geologiczny

Larose D T (2005). Discovering Knowledge in Data: an Introduction to Data Mining. New York: John Wiley & Sons, 240

Lewandowski J (1982). Extent of ice sheet of Middle-Polish glaciation in the Silesian Upland. Biuletyn Instytutu Geologicznego, 337(26): 115–136 (in Polish)

Lewandowski J (1987). Odra glaciation in the Silesian Upland. Biuletyn Geologiczny, 31: 247–301 (in Polish)

Liu F, Gao H, Pan B, Li Z, Su H (2019). Quantitative analysis of planation surfaces of the upper Yangtze River in the Sichuan-Yunnan Region, Southwest China. Front Earth Sci, 13(1): 55–74

Luo L, Mu L, Wang X, Li C, Ji W, Zhao J, Cai H (2013). Global detection of large lunar craters based on the CE-1 digital elevation model. Front Earth Sci, 7(4): 456–464

MacMillan R A, Shary P A (2009). Landforms and landform elements in geomorphometry. In: Hengl T, Reuter H I, eds. Geomorphometry. Concepts, Software, Applications. Amsterdam: Elsevier, 227–254

Mentlik P, Novotna M (2010). Elementary forms and 'scientific reliability' as an innovative approach to geomorphological mapping. Journal of Maps 6(1): 564–583

Minár J, Evans I S (2008). Elementary forms for land surface segmentation: the theoretical basis of terrain analysis and geomorphological mapping. Geomorphology, 95(3-4): 236–259

Mitášová H, Hofierka J, Zlocha M, Iverson R L (1996). Modelling topographic potential for erosion and deposition using GIS. Int J Geogr Inf Syst, 10(5): 629–641

Moore I D, Grayson R B, Ladson A R (1991). Digital terrain modelling: a review of hydrological, geomorphological, and biological applications. Hydrol Processes, 5(1): 3–30

MPHP (Digital Map of Hydrographical Division of Poland) (2010). IMiGW, Warszawa

Niemann K O, Howes D E (1991). Applicability of digital terrain models for slope stability assessment. ITC J, 3: 127–137

Ortuño M, Guinau M, Calvet J, Furdada G, Bordonau J, Ruiz A, Camafort M (2017). Potential of airborne LiDAR data analysis to detect subtle landforms of slope failure: Portainé, Central Pyrenees. Geomorphology, 295: 364–382

Pike R J (1988). The geometric signature: quantifying landslide-terrain types from digital elevation models. Math Geol, 20(5): 491–511

Piloyan A, Konečný M (2017). Semi-automated classification of landform elements in Armenia based on SRTM DEM using *k*-means unsupervised classification. Quaest Geogr, 36(1): 93–103

Speight J G (1990). Landform. In: McDonald R C, Isbell R F, Speight I G, Walker J, Hop M S eds. Australian Soil and Land Survey Field Handbook. Melbourne: Inkata Press, 9–57

Starkel L (1980). Geomorphological Outline Map of Poland, 1:500000. Warszawa: IGiZP PAN

Szaflarski J (1955). Overview of the relief development of the Silesian Upland. In: Wrzosek A ed. Górny Śląsk. Kraków: Wydawnictwo Literackie, 65–121 (in Polish)

Szczypek T (1977). Eolic activities and deposits in the southern part of the Silesian Upland. Katowice: Prace Naukowe (in Polish)

Szczypek T (1986a). Aeolian cover sands in the northern part of the Silesian Upland. Geographia. Studis and Dissertationes, 9: 45–56 (in Polish)

Szczypek T (1986b). Dune forming processes in the middle part of the Cracow-Wieluń Upland against a background of the neighbouring area. Katowice: Prace Naukowe UŚ 823. 183 (in Polish)

Szczypek T (1988). Aeolian activity in the eastern part of the Silesian Upland on the example of the Bukowno vicinity. Geographia. Studia et Dissertationes, 11: 7–22 (in Polish)

Szczypek T, Wach J (1991). Development of the modern dune in the strong human impact conditions. Katowice: Prace Naukowe (in Polish)

Szczypek T, Wach J (1992). Human impact and course of natural morphogenetic processes on the example of Silesian Upland. Kształtowanie środowiska geograficznego i ochrony przyrody na obszarach uprzemysłowionych i zurbanizowanych, 4: 5–12 (in Polish)

Szczypek T, Wach J (1993). Anthropogenic scarp dune at Bukowno on the Silesian Upland in the period 1989–1993. Katowice: Uniwersytet Śląski, 50 (in Polish)

Szypuła B (2009). Research on the rock strength of the Silesian Upland using Schmidt hammer. Geographia. Studia et Dissertationes, 31: 65–80 (in Polish)

Szypuła B (2017). Quantitative studies of the morphology of the south Poland using Relief Index (RI). Open Geosci, 9(1): 509–524

Tang G, Li F (2008). Landform classification of the loess plateau based on slope spectrum from grid DEMs. In: Advances in Digital Terrain Analysis (Lecture Notes in Geoinformation and Cartography), 107–124

Tobler W (1970). A computer movie simulating urban growth in the

Detroit region. Econ Geogr, 46(2): 234–240

Urbański J (2012). GIS in the environmental research. Gdańsk: Uniwersytetu Gdańskiego, 252 (in Polish)

Van Lopik J R, Kolb C R (1959). A technique for preparing desert terrain analogs. U.S. Army Engineer Waterways Experiment Station. Vicksburg, MS, Tech. Rept. 3–506

Weiss A (2001). Topographic Position and Landform Analysis. Poster presentation, In: ESRI User Conference. San Diego

Wieczorek M (2008). The classification of landforms based on Digital Elevation Model. Dissertation for the Doctoral Degree. Wrocław: Uniwersytet Wrocławski, 104 (in Polish)

Wieczorek M (2011). An influence of spatial range of input data set on terrain relief form classification homogeneity for glacial area. In: Ruas A ed. Advances in Cartography and GIScience, Vol. 2 Selection from ICC 2011. Paris: Springer, 357–369

Wieczorek M, Migoń P (2014). Automatic relief classification versus expert and field based landform classification for the medium-altitude mountain range, the Sudetes, SW Poland. Geomorphology, 206: 133–146

Wilson J P, Gallant J (2000). Terrain Analysis. Principles and Applications. London: John Wiley & Sons Inc., 479

Wood W F, Snell J B (1960). A quantitative system for classifying landforms. Technical Report EP-124. U.S. Army Quartermaster Research and Engineering Center, 20

Yang X, Li M, Na J, Liu K (2017). Gully boundary extraction based on multidirectional hill-shading from high-resolution DEMs. Trans GIS, 21(6): 1204–1216

Żmuda S (1973). Anthropogenic changes in the natural environment of the Upper Silesian conurbation. Warszawa-Kraków: PWN, 207 (in Polish)