Study on spatial distribution of horizontal geodetic control points in rural areas

**Elzbieta Bielecka · Krzysztof Pokonieczny ·
Paweł Kamiński**

Abstract The paper presents the results of research concerning the analysis of spatial distribution of detailed geodetic control points in rural areas in Poland. The analysis takes into consideration the relations between the number of geodetic control points and land cover type (built-up, agricultural, and forest) as well as the effective legal regulations in Poland with regard to the issues. The analysis results were referred to geodetic sections classified by the k-means method with respect to the number of the control points and the built-up, agricultural and forest areas. The research indicates that the number of the geodetic control points in an area depends to the extent of 35 % on the type of land cover (building, agricultural and forest land). Having this in mind, the authors recommend taking the land cover into consideration already in the phase of designing a geodetic control network.

Keywords   Horizontal geodetic control point · Geospatial analysis · K-means clustering ·
Thematic map

1 Introduction

Geodetic controls are of very high priority to various activities such as geodetic, geophysical, surveying and civil engineering, GIS developing and spatial data gathering. Control surveys establish precise horizontal and vertical positions of geodetic monuments. These serve as the basis for originating or checking subordinate surveys for property boundary delineation, route and construction planning, cadastral and topographic mapping. They are also essential

E. Bielecka (✉) · K. Pokonieczny · P. Kamiński
Faculty of Civil Engineering and Geodesy, Institute of Geodesy, Military University of Technology,
2 gen. S. Kaliskiego St., 00-908 Warsaw, Poland
e-mail: ebielecka@wat.edu.pl

K. Pokonieczny
e-mail: kpokonieczny@wat.edu.pl

P. Kamiński
e-mail: pkaminski@wat.edu.pl

as a reference framework for giving locations of data input in Land Information Systems and Geographic Information Systems. Geodetic controls are a fundamental infrastructure for any geo-spatial analysis activities.

The important issues related with designing geodetic controls include defining the number of geodetic control points and their configuration providing required precision, reliability and geometrical strength, and economy. The term *reliability of a geodetic control network*, introduced in 1960s by Baarda (1968), means the ability of the network to detect gross errors in the observations and to be resistant to undetectable errors. The strength criterion, elaborated by Vaniček et al. (2001) deals with strain, and it is a kind of robustness indicator.

The methods of designing and densifying geodetic control networks were the subject of study conducted by Augath (1982), Furaha (1990), Schaffrin and Cothren (2003), Eshagh and Klamerh (2007), Fok et al. (2009), Klein et al. (2012) and Kryński et al. (2013). Some researchers proposed methods for improving the quality of geodetic networks based on entropy (Mrówczyńska 2013), shift-M_{split} estimation (Duchnowski and Wiśniewski 2011), multiple outliers (Klein et al. 2012), or the base station quality index (Han et al. 2012). However, none of the above mentioned researches focused on the impact of topography on the quality of the control networks. Selected studies (Beesley 2003; Xavier and Costa 2007; Zhang et al. 2008; Han and Li 2010; Han et al. 2012) mainly considered the impact of terrain elevation on the control points location measured by means of the GPS techniques. Despite many research works and studies concerning design and densification of geodetic controls, the problem of geodetic networks design is not fully solved, because there is no analytical method that solves the problem of how many stations should be included in a network and where they should be located (Gerasimenko et al. 2000).

The quality of geodetic controls expressed in the studies and researches as reliability and geometrical strength, as well as economy takes the topographic location of the geodetic control points into consideration only to a little extent. In a sense, it is considered in the economical factor which, however, is extremely rarely taken into account during designing the geodetic control (Teunissen 1985). The quality understood in this way, especially with respect to detailed geodetic controls, does not include all information required by surveyor during engineering, parcel boundaries, and topographic measurements, etc. They need not only information concerning the position of the geodetic control points and the accuracy of their coordinates but also information about the number of points available in an area, their topographic and situational location, usability of others geodetic control points with the assumption of the lowest possible cost of all measurement tasks.

The aim of the presented research is an analysis of spatial distribution of horizontal geodetic control points (3rd order) taking into account differentiation of degree of land development. The results are referred to geodetic sections, the lowest unit of local administration which usually comprised of a single village. They are presented in the form of thematic maps, especially choropleth maps. Cartographic presentation of the results will facilitate decisions concerning modernizing and densification of geodetic controls, and it will enable surveyor to assess the time needed for measurement tasks faster and more precisely.

2 Object and method of research

2.1 Horizontal geodetic control in Poland

In Poland, the geodetic control has been categorized as primary, secondary, or detailed since 2012. The course of action and technical standards concerning establishing and maintaining

geodetic controls, as well as the rules of collecting information about the geodetic controls databases are governed by the Regulation of the Ministry of Administration and Digitization of Poland of February 14, 2012 as regards geodetic control, gravimetric and magnetic networks (MAiC 2012). The detailed horizontal geodetic control consists of a set of points that develop the basic horizontal geodetic control; the detailed geodetic control points are used as reference for measurement controls and for conducting detailed geodetic measurements. The density of the control points is diverse and it depends on the terrain development status: in urbanized areas it is one or more points in 20 ha, while in arable and forest areas—it is at least one point in 120 ha. The regulation recommends to locate the control points in a way that provides the greatest number of cases of mutual capability to observe the neighboring control points of the same or higher order.

The detailed control may be established by static GNSS measurements method, measurements conducted in the ASG-EUPOS system or traditional measurements of angles and distances. The mean error of control point's location is in the range from 0.05 to 0.10 m, and for new established control points it should not exceed 0.07 m. The horizontal geodetic control networks up to the 3rd order are established and maintained on a national basis and they are a part of the Polish Spatial Information Infrastructure. According to the Head Office of Geodesy and Cartography, the number of the detailed geodetic control points is 1,378,383 which yields the average density of 1 in 22.5 ha. However, it is worth mentioning that many of the points require reestablishing because their stabilizing marks in terrain have been damaged.

The Regulation of the Ministry (MAiC 2012) provides also a common methodology for creating datasets of horizontal coordinate values for geodetic control points represented by survey monuments. Information concerning geodetic control points is collected in detailed geodetic controls database and managed at the level of the administrative districts (second level of administrative division of Poland). The database includes information about the point number, its coordinates in the national 2D orthogonal coordinate system, mean location error, stabilization type and topographic description (enabling to recover or reconstruct the point in terrain).

The analysis of the spatial distribution of geodetic control points was conducted for 3rd order horizontal geodetic control points in the district of Losice. The presented research enables not only to show the control with respect to topographic information but also to make a quick assessment whether in the area of interest of the surveyor there is a sufficient number of control points or a densification by means of a measurement control is required. Consequently, availability of such information will facilitate and speed up planning and assessment of measurement tasks cost. The assumption is that in future this information will be available in spatial data network services.

2.2 Method of research

It was assumed that the analysis of the spatial distribution of the detailed geodetic control points (further referred to as GCP) would be conducted in the basic fields for which geodetic sections (further referred to as GS) were selected. The sections are the smallest units of the administrative and geodetic division. The aim of the analysis is to classify the geodetic sections with respect to the number of control points and their distribution. K-means clustering was used for the analysis which enables to group geodetic sections with respect to the similarity of the following features: (1) the number of geodetic control points, (2) percentage of built-up areas, (3) percentage of the agricultural land and (4) percentage of forest. K-means

algorithm belongs to the Exploratory Data Analysis (EDA) and is used to group sets of objects in such a way that objects in the same group (called a cluster) are more similar to each other than to those in other groups (clusters). Particularly, k-means clustering aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean, serving as a prototype of the cluster. The number of classes was determined initially using the V-fold cross-validation test and then it was verified after the initial analyses.

The analyses were conducted according to the following scheme:

(1) Initial data processing including:

    a. selecting object classes from VmapL2 and creating (by means of agregarion) a polygon thematic layer concerning land cover;

    b. calculating the percentage of built-up areas, agricultural and forest areas in the GSs, which in further statistical analyses would be treated as explanatory variables;

    c. calculating the number of GCPs located in each GS, treated as the explained variable;

    d. analysis of correlation between the explanatory variables and the explained variable and variables standardizing;

    e. tests verifying the target number of classes.

(2) Cluster analysis—k-means algorithm (with the option of Euclidean distance).
(3) Analysis and geovisualization of the analysis results.

Choropleth map and cartographic compilation, a method named by Medyńska-Gulij (2010), which enables to present information from miscellaneous thematic data files on one map, were used for visualizing the results.

3 Case study

3.1 Area and input data

The analysis of the spatial distribution of the control points was done for the district of Losice located in the eastern part of Poland. The district consists of 7 communes including one town of Losice. Its total area is 77,030 ha and its number of inhabitants is 32,925 (GUS 2012). The district is of mainly agricultural character, the land usage structure is dominated by agricultural land which is 72.9 % (including arable land—56.2 %, meadows and pastures—15.4 % and orchards—1.3 %), forests cover slightly over 20 %, and water and built-up areas—6.7 %.

The detailed controls data are collected by the Geodetic and Cartographic Center in Losice. The analyzed control consists of 1926 points, of the RMS equals to 0.08 m. Location of GCPs is presented in Fig. 1.

The information concerning the land cover is derived from VmapL2. It is a geospatial, vector database, equivalent (in terms of details) to the topographic map of the 1:50,000 scale, developed and maintained by the Polish Military Geography Directorate. Object coding is realized in accordance with the DIGEST (Digital Geographic Information Exchange Standard) standard. VmapL2 covers the entire area of Poland. The data in the database are grouped in 11 thematic categories, such as boundaries, elevation, hydrography, industry, physiography, aero, transportation, cultural, vegetation. They are then subdivided into subcategories and classes. Owing to the agricultural character of the district, for the purposes of the research it was decided to distinguish four dominating land cover classes: built-up areas, arable land,
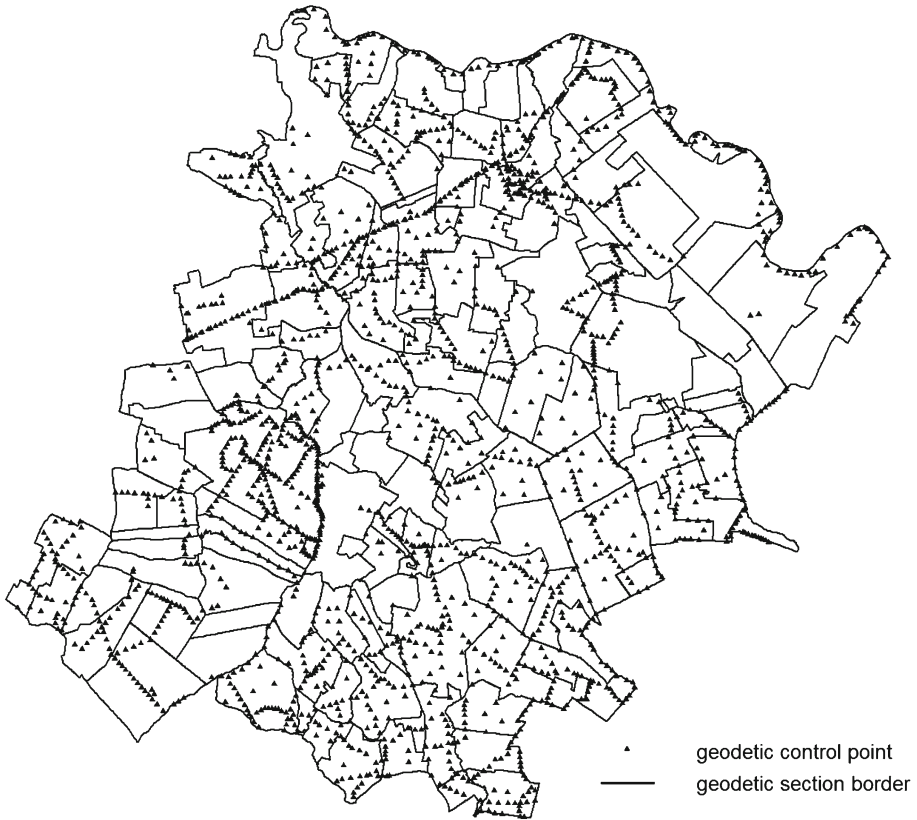
Fig. 1 Location of first order GCPs

forests, and water. Water areas were neglected in further analyses because they cover only 0.4 % of the district's area and by definition they do not include any geodetic control points.

3.2 Research results and discussion

According to Section 6 point 8 of the Regulation of the Ministry of Administration and Digitization of Poland of February 14, 2012 concerning geodetic, gravimetric and magnetic controls (MAiC 2012), the density of detailed control points together with the basic control points should not be lower than 1 point in 20 ha in urbanized areas, and it should be related with the needs of the local land-use plans in arable and forest areas, however, it should not be lower than 1 point in 120 ha. The area of the built-up area in the analyzed territory equals 1,666 ha, while the area of the arable lands and forest equals 75,321 ha. According to the Regulation, in the built-up areas there should be at least 84 GCPs, while in the arable lands and forest there should be at least 628 GCPs.

The research indicates unambiguously that in the Losice district the requirements are met (Table 1). In the built-up area there are 184 points, while in the arable and forest terrains there are in total 1,742 points, i.e. 90.5 % of all the points. It is worth to note that these numbers exceed more than twice the numbers of points in the specific terrain types, required by the Regulation.

Table 1 Land cover structure and the number of GCPs

	Area		Geodetic control points	
	(ha)	%	Number	%
Built-up area	1,666.49	2.16	184	9.55
Arable land	58,143.15	75.24	1,490	77.36
Forest	17,177.56	22.23	252	13.08
Water	292.87	0.38	0	0.00
Total	77,280.07	100.00	1,926	100.00

Table 2 Descriptive statistics of the analyzed set of the detailed geodetic control points

	Value
Number of geodetic sections	132
Minimum number of GCPs	0
Maximum number of GCPs	162
Average number of GCPs in a geodetic section	14
Median	11
Standard deviation	16.18
Total number of GCPs	1926

The number of geodetic control points in individual geodetic sections is significantly differentiated and equals from zero (for 6 typically arable geodetic sections) up to 162 in Losice which is a town of slightly over 7 thousand inhabitants. The average number of GCPs in the geodetic section is 14 (Table 2). The density of the GCPs (the number of GCPs in 1 km^2) is also significantly differentiated. It has the highest values (over 5 GCP/km^2) for built-up areas and the smallest for the forests (less than 1 GCP/ km^2). The differentiation is shown in Fig. 2.

The V-fold cross-validation test suggested dividing the geodetic sections into 4 classes with respect to the differentiation of the area covered by built-up areas, forests and agriculture land and the number of geodetic control points, treated as the diagnostic variables. This division resulted in distinguishing geodetic sections of built-up and agricultural, agricultural, forest, and mixed cover character. Having done numerous analyses the authors decided to add one more class (division into five clusters). This enabled to distinguish a separate built-up area class. The class characteristics are presented in Fig. 3, where each class is represented by different line types. Analysis of Fig. 3 indicates that areas covered by built-up and agriculture lands are the most diversifying factors designated geodetic sections classes (clusters). The area occupied by forests is indistinguishable in classes 2, 3 , and 4. It varies only in classes 1 and 5. The density of GCPs differentiates from other classes only class 4, where it is three time.

Geodetic sections including all types of land cover classes without a significantly dominating one ("mixed" cover sections) were classified in class 1. In GSs belonging to class 2, similarly to class 3, a significant area is covered by agricultural lands. However, in class 2, a relatively big part of built-up areas was distinguished, hence the name: "agricultural and built-up". Class 3 includes geodetic sections with dominating arable lands, while the area of built-up areas and forests is the smallest ("agricultural" sections). Class 4 cumulates two sections in which there are two biggest towns of the district, it is the "built-up" class. In class
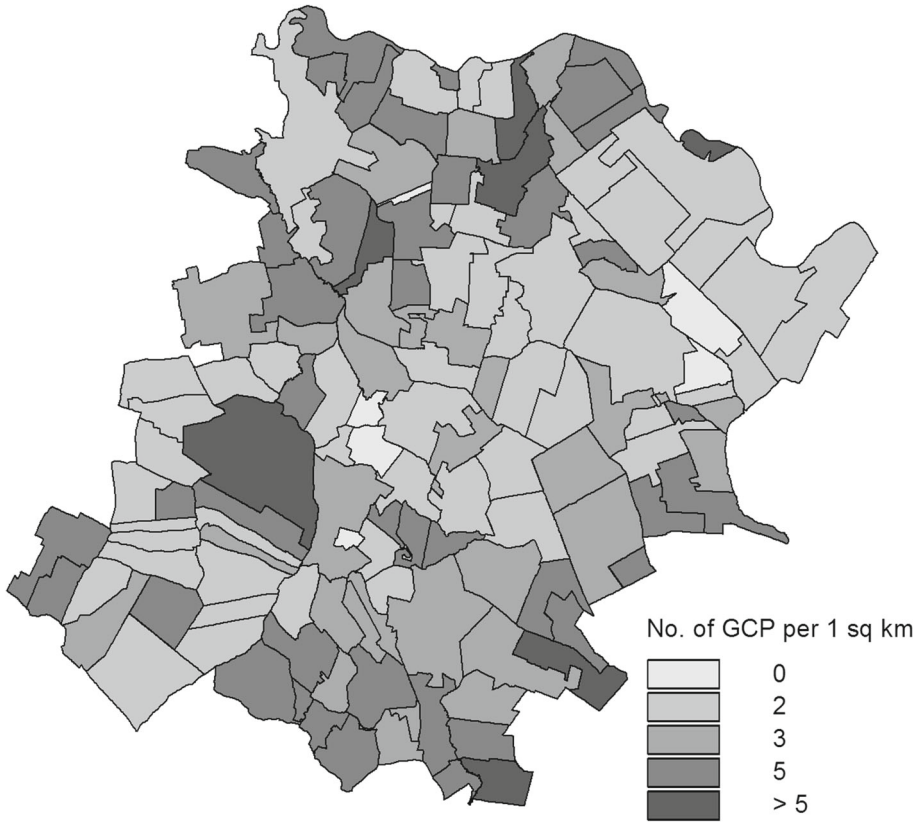
Fig. 2 Density of GCP in the Losice district

5, the clearly dominating element of the land cover is forests ("forest" sections). The smallest part of the arable lands is also characteristic of the class. Figure 4 shows that in the defined ranges, the number of the GCPs is similar, only in the "built-up" class the number of the geodetic control points is greater.

The number of geodetic control points and their density in the individual classes of geodetic sections is presented in Table 3. It shows that in the Losice district, the most numerous are the "agricultural" and "mixed" geodetic sections which correspond to the rural character of the district. These two classes cumulate the most of the geodetic control points (almost 65 % of all points). It results from the area included in the two classes which is altogether about 58 % of the entire district area. The Pearson correlation coefficient between the area and the number of GCPs is 0.63. The density of points in class 1 and 3 is similar and equals about 2.5 points in 1 km^2. The smallest number of points was observed in geodetic sections classified as "forest". There are there about 11.5 % of all geodetic points, while the area of the sections is almost 20 % of the Losice district. According to the above mentioned Regulation (MAiC 2012), the greatest numbers of geodetic control points in 1 km^2 are observed in sections classified in classes 2 and 3. There are almost 28 % of all control points there, while the area constitutes 22 % of the entire district area. The results of the cluster analysis are visualized in the form of a choropleth map in Fig. 4.
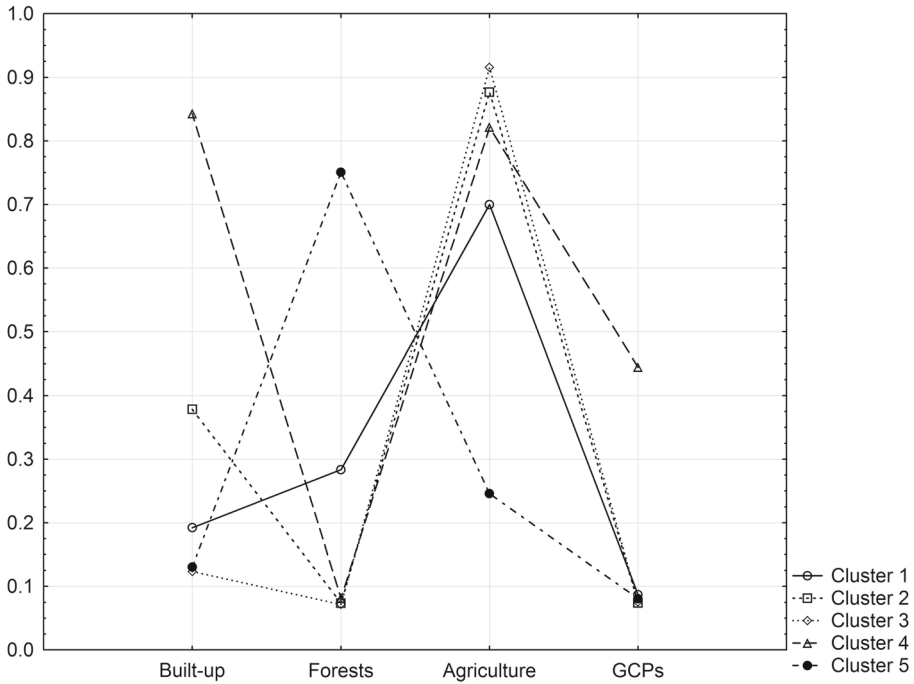
Fig. 3 Differentiation of class attributes distinguished by means of the k-means method

Figure 4 shows that GSs belonging to the individual classes do not form spatial clusters. It is also confirmed by the value of the Global Moran's Index close to zero (Moran's Index $= -0.004484$) which indicates that it is not possible to find spatial clusters of GSs with similar number of GCPs and land cover structure.

Analysis of Fig. 5 confirms non-uniform distribution of geodetic control points and its dependence on land cover type. In 30 % of geodetic sections (highlighted in light grey) the number of geodetic control points is lower than the one required by the Regulation. 21 of them belong to agricultural and built-up class, 8—to mixed, 5—to agriculture, and 3—to forest. All together these sections occupied 1.4 km^2, which constitute the 18.3 of the area of Losice district. For 97 remaining geodetic sections (presented in dark gray in Fig. 5) the number of geodetic control points is higher than the one required by the Regulation. Many more control points (5 times more) are in geodetic section located near border of analysed district and in the area were cadastre has been modernized.

To determine the influence of the land cover type on the number of control points in the geodetic sections, the multiple correlation coefficient R was determined by means of the Pavlowski's method. Its value is $R = 0.59$ which means that the determination coefficient $R2$ is 0.35, and the number of GCPs in the area depends as much as in 35 % on the land cover type (built-up, agricultural and forest areas).

4 Conclusions and recommendations

The proposed method of analysis of spatial distribution of geodetic control points will facilitate solving the optimization problems for densification of geodetic control networks accord-
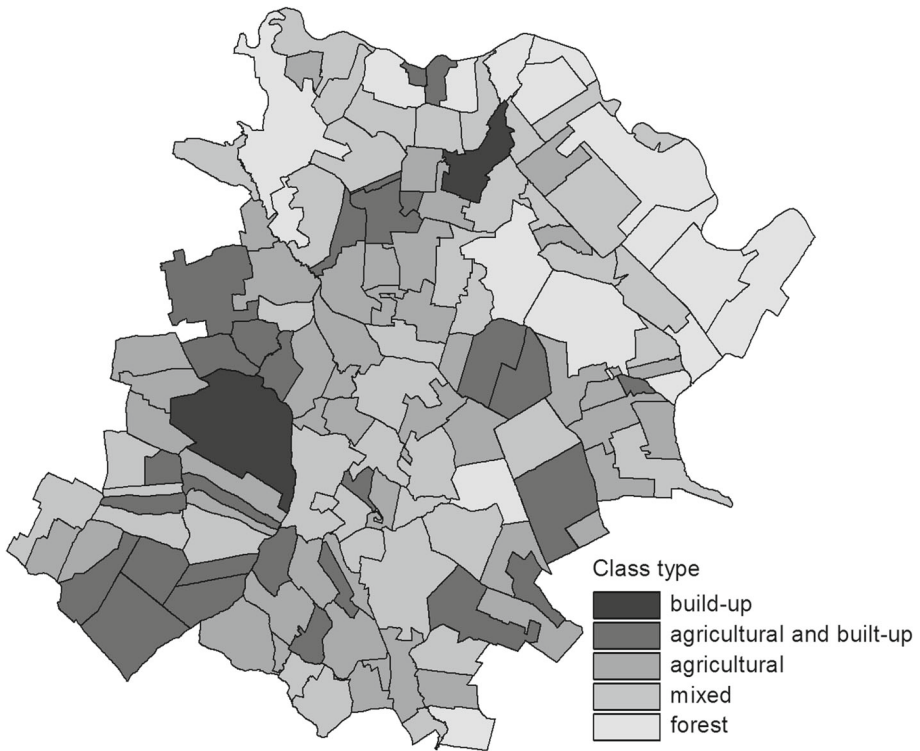
Fig. 4 Choropleth map of GSs classification by means of the k-means method

Table 3 The number of geodetic control points and their density in the individual classes of geodetic sections

Class no.	Class type	Number of GSs	% of GSs	Number of GCPs	Density of GCPs (per 1 km²)	Area (km²)
1	Mixed	42	30.9	590	2,45	241.19
2	Agricultural and built-up	28	20.6	325	2.33	139.74
3	Agricultural	47	34.6	575	2.76	208.42
4	Built-up	2	1.5	214	6.88	31.12
5	Forest	17	12.5	222	1.46	152.01

ing to the third-order design (THOD)—increasing the number of observations to improve an existing network. The research indicates that information concerning the land cover structure should not be neglected in the phase of designing the location of detailed geodetic control points because it has a significant influence on the number of GCPs.

Although in the whole area of the district, the number of geodetic control points exceeds the minimum number required by the Regulation (MAiC 2012), the differentiation of their density results in a situation that in some geodetic sections classified as "forest" and "agricultural and built-up" areas the number is much smaller. The "forest" sections are located inside a vast forest, so reduction of the number of GCPs there is justified from the economical point of view. In contrast, "agricultural and built-up" are
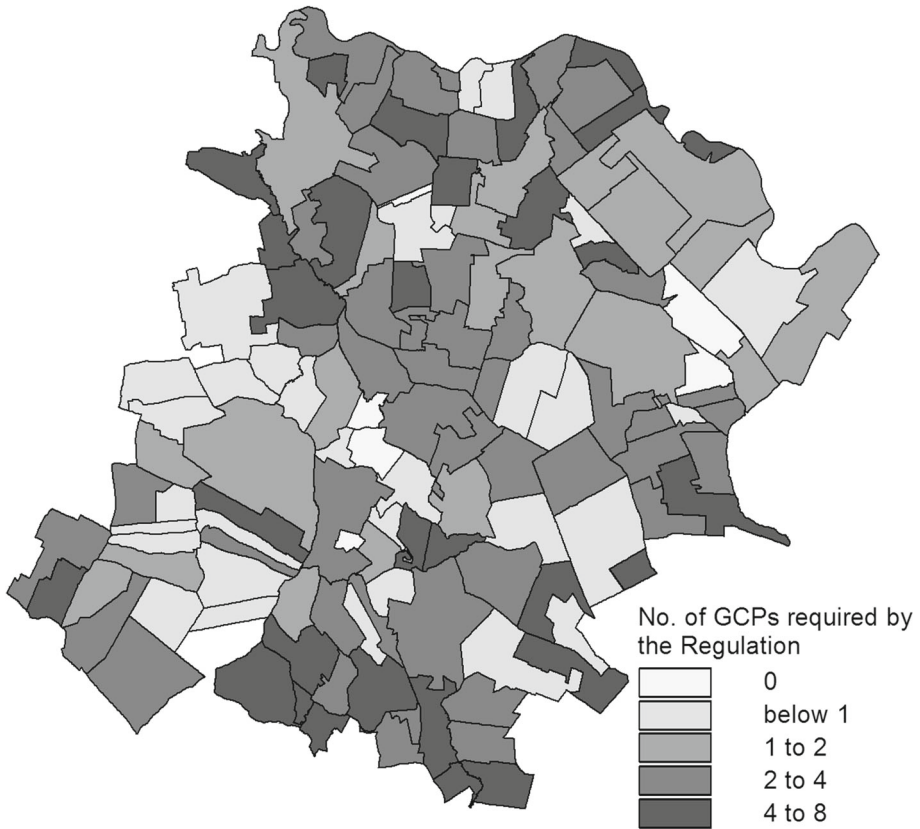
Fig. 5 Number of GCPs in: 20 ha of urban areas and 120 ha for the remaining geodetic sections

the areas with dispersed settlements. It is related with the peculiarity of built-up areas in rural terrains which is characterized by significant dispersion of the buildings and level of detail of the VmalL2 database corresponding to the map of 1:50,000 scale in which single buildings are not taken into consideration. It is worth to note the fact that "urbanized areas" are not defined in the Regulation and that they are of different character in urban regions, urban and rural, or typically rural which is the case of the analyzed area.

The planned densification of the geodetic control should at first include the indicated in Fig. 5 areas with the number of geodetic control points smaller than the requirement. The geodetic sections classification with respect to the number of GCPs and the land cover structure will make it easier for the surveyor to assess geodetic control points availability as well as to assess time and cost of measurement tasks realization. For the geodetic and cartographic administration it will facilitate planning work related with modernization of geodetic controls.

The analyses results may be published in the form of thematic maps on geo-portals or other web-mapping services. Examining geodetic control point in the web map together with topographic data allows to get a clear idea of the set of geodetic control points that establish

a link between real world locations and their correct corresponding locations via maps and mapping applications.

References

Augath W (1982) Accuracy and reliability measures concerning design and qualification of densification networks. In: Borre K, Welsch W (eds) International federation of surveyors—FIG—proceedings survey control networks meeting of study group 5B, 7th–9th July, 1982, Aalborg University Centre, Denmark Schriftenreihe des Wissenschaftlichen Studiengangs Vermessungswesen der Hochschule der Bundeswehr Mün-chen, Heft 7, pp 51–63

Baarda W (1968) A testing procedure for use in geodetic networks. Publ Geodesy, New Series 2, No. 5 Nederland Geodetic Commission, Delft

Beesley BJ (2003) Sky viewshed modeling for GPS use in the urban environment. Proceedings of the twenty-third annual ESRI user conference, San Diego, California

Duchnowski R, Wiśniewski Z (2011) Shift-M_{split} estimation. Geod Cartogr 60(2):79–97. doi:10.2478/v10277-012-0009-6

Eshagh M, Klamerh R (2007) A strategy for optimum designing of the geodetic networks from the cost, reliability and precision views. Acta Geod Geoph Hung 42(3):297–308. doi:10.1556/AGeod.42.2007.3.4

Fok HS, Baki Iz H, Schaffrin B (2009) Comparison of four geodetic network densification solutions. Surv Rev 41(311):44–56. doi:10.1179/003962608X389997

Furaha NL (1990) Rigorous mathematical models for the densification and integration of geodetic networks. Bull Géod 64(3):219–229

Gerasimenko MD, Shestakov NV, Kato T (2000) On optimal geodetic network design for fault-mechanics studies. Earth Planets Space 52:985–987

GUS (2012) Demographic data for Łosice as of December 31 2013. www.stat.gov.pl/vademecum/.../losicki/Losice.pdf? Accessed 20 Dec 2013

Han JY, Li PH (2010) Utilizing 3-D topographical information for the quality assessment of a satellite surveying. Appl Geomat 2(1):21-32

Han JY, Wu Y, Liu RY (2012) Determining the optimal site location of GNSS base stations. Bol Ciênc Geod 18(1):154–169. doi:10.1590/S1982-21702012000100009

Klein I, Matsuoka MT, Souzas F, Collischonn C (2012) Design of geodetic network reliable against multiple outliers. Bol Cienc Geod 8(3):480–507

Kryński J, Olszak T, Barlik M, Dykowski P (2013) New gravity control in Poland—needs, the concept and the design. Geod Cartogr 62(1):3–21. doi:10.2478/geocart-2013-0001

MAiC (2012) Regulation of Ministry of Administration and Digitization of Poland of 14 February 2012 as regards geodetic control, gravimetric and magnetic networks. Dz. U. 2012 No. 0, entry. 352 (in Polish)

Medyńska-Gulij B (2010) Map compiling, map reading, and cartographic design in "Pragmatic pyramid of thematic mapping". Quaest Geogr 29(1):57–63. doi:10.2478/v10117-010-0006-5

Mrówczyńska M (2013) Analysis of the horizontal structure of a measurement and control geodetic network based on entropy. Geod Cartogr 62(1):23–31. doi:10.2478/geocart-2013-0002

Schaffrin B, Cothren J (2003) Hierarchical data fusion with photogrammetric applications. Braz J Cartogr 55(2):25–34

Teunissen PJG (1985) Optimization and design of geodetic networks. In: Grafarend EW, Sanso F (eds) Quality control in geodetic networks. Springer, Berlin, pp 526–547

Vaniček P, Craymer MR, Krakiwsky EJ (2001) Robustness analysis of geodetic horizontal networks. J Geod 75:199–209

Xavier P, Costa E (2007) Simulation of the effects of different urban environments on land mobile satellite systems using digital elevation models and building databases. IEEE Trans Vehicular Technol 56(5–2):2850–2858

Zhang K, Liu GJ, Wu F, Densley L, Retscher G (2008) An investigation of the signal performance of the current and future GNSS in typical urban canyons in Australia using a high fidelity 3D urban model. Location based services and telecartography II: from sensor fusion to ubiquitous LBS. Springer, Heidelberg, pp 407–420