Polish Cartographical Review
Vol. 54, 2022, pp. 75–91
DOI: 10.2478/pcr-2022-0006
Received: 25.05.2022
Accepted: 22.08.2022

IGA AJDACKA
University of Warsaw, Faculty of Geography and Regional Studies,
Department of Geoinformatics, Cartography and Remote Sensing
Warsaw, Poland
orcid.org/0000-0001-5543-404X; i.ajdacka@student.uw.edu.pl

IZABELA KARSZNIA
University of Warsaw, Faculty of Geography and Regional Studies,
Department of Geoinformatics, Cartography and Remote Sensing
Warsaw, Poland
orcid.org/0000-0001-5510-8770; i.karsznia@uw.edu.pl

Verifying and improving map specifications of river network selection for automatic generalization of small-scale maps

Abstract. Automated generalization is highly desired for effective map production. This research focuses on the initial stage of generalization, namely object selection. The study aims to conduct river network automatic selection based on map specifications contained in the Minister of Internal Affairs and Administration regulation. The research covers river network selection from the General Geographic Objects Database from 1:250,000 to 1:500,000 detail level. Within the research scope, three selection variants were designed. The first was a basic variant that only included the implementation of the specifications contained in the regulation. The other two were experimental variants: an extended variant and an extended-modified variant with the parameters and data enrichment proposed by the authors. The extended variant has been supplemented with the Id_MPHP index usage, derived from the Map of Hydrographic Division of Poland (MPHP), which defines the hierarchy of watercourses in the river network. The extended-modified variant was implemented according to the guidelines of the regulation, with the use of the Id_MPHP index and additionally with the help of the parameter denoting "priority" watercourses, which was assigned by the authors. The results of the work constitute the generalization models designed in ArcMap 10.8. with the use of Model Builder functionality as well as the maps presenting the selection variants output visualizations. The results were compared visually as well as verified with the reference atlas map generalized by an experienced cartographer. As a result, the map specifications concerning the selection process presented in the regulation proved to be insufficient to generalize river networks properly. The variants proposed in this research made it possible to improve the selection results and enabled the automation of the river selection process. Additional specifications and parameters proposed in this work may constitute an essential supplement to the guidelines contained in the regulation.

Keywords: cartographic generalization, river network, automation, small-scale maps

1. Introduction

Every map is a model of reality because not everything situated in space can be shown on a map. Maps are characterized by a lower level of detail than reality therefore, the information they contain must be selected and simplified. The degree of object generalization on a given map depends on its purpose and map scale. The map scale determines the level of reduction of a fragment of reality. The process of excluding unnecessary details while maintaining or even highlighting the important ones is called cartographic generalization.

There are many definitions of cartographic generalization. The International Cartographic Association (ICA) defines generalization as the selection and simplification of geographic information according to the scale and purpose of the map (1973). According to Robinson et al. (1978), generalization includes several activities which can be classified into four primary generalization operators:

– selection and simplification assumes defining relevant data features and making them more visible,

– classification consists of the selection of an appropriate scale and grouping of data,

– symbolization concerns graphic coding of the object groups based on their common values or position,

– induction assumes the usage of the logical inference process in cartography.

These four elements are strongly associated with each other. Selection and simplification are important from the point of view of maintaining the clarity and fidelity of the presentation. One of the most complex challenges for cartographers is the issue of which elements on the map should be omitted and which ones should be kept. It is possible to implement these operations in automatic selection mode by assigning them various thematic and spatial attributes. Classification is the second element of generalization that divides objects into classes in order to organize them. According to Robinson et al. (1978), the most frequently used classification processes include grouping similar objects into classes of phenomena, selecting the location of an element and its possible modification based on the placement of adjacent objects, and other forms of classification that may consist of aggregating or creating new data to obtain specific classes. With symbolization, it is possible to visualize the previously classified objects and the outputs of the simplification. Symbolization leads to emphasizing the most important phenomena on the map and generalizing the objects to a small extent. Induction tends to be used if there is a need to obtain more information than the data used to design the map. An example of the data characterizing a continuous distribution would be isolines presenting air temperature. Although the temperature is readable from any place on the map, the data used for isolines design includes a limited number of sites with measured values. Occasionally cartographers may use induction subconsciously when using their knowledge to extend the information presented on a given map.

In this paper we focused on one of the initial stages of generalization, which is selection. It is a primary operator of generalization, significantly affecting its final result. The study aims to conduct river network automatic selection based on map specifications contained in the Minister of Internal Affairs and Administration regulation (Ministerstwo Rozwoju, Pracy i Technologii [MRPiT], 2021; Ministerstwo Spraw Wewnętrznych i Administracji [MSWiA], 2011). The research covers river network selection from the General Geographic Objects Database (GGOD) from 1:250,000 to 1:500,000 detail level within three test areas. The contributions of this research are as follows. First we verify the river selection specifications contained in the regulation by implementing them in the form of the basic variant (the implementation of the specifications contained in the regulation). Second, we propose to enrich the basic variant by implementing two experimental variants: an extended variant and an extended-modified variant. The extended variant has been supplemented with the Id_MPHP[1] index usage, derived from the Map of Hydrographic Division of Poland (MPHP),[2] which defines the hierarchy of watercourses in the river network. The extended-modified variant was implemented according to the guidelines of the regulation using the Id_MPHP index and additionally with the help of the parameter denoting "priority" watercourses. Third, we verify all variants by comparing them to the reference atlas map, designed by an experienced cartographer.

2. Related work

Cartographic generalization, can be conducted in the manual or digital mode. Because of the development of technology and the specificity of digital and interactive maps, at present the precedent goal is to create methods that allow designing maps fully automatically. The evolution of computer science and digital cartography brought new possibilities for constructing generalization methods and tools. Thus the algorithms and methods leading to automation have changed and become more advanced.

The first attempts to automate cartographic generalization were made in the 1960s. These attempts mainly referred to the simplification of

[1] Id_MPHP is an attribute that determines the order of importance.

[2] MPHP presents the hydrographic network of Poland at the detail level corresponding to the 1:50,000 or 1:10,000 scale (Otwarte Dane; n.d.; https://dane.gov.pl/pl/dataset/2167,mapa-podzialu-hydrograficznego-polski-w-skali-110).

the object geometry, while at the same time, the geographic nature of the objects was ignored. Tobler (1964) conducted one of the first studies in this field. In his research, various ways of simplifying the coastline by eliminating some points that make up its shape were described. The research on linear objects generalization was continued in subsequent studies. Douglas and Peucker (1973) developed a line simplification method as the main element of automatic generalization. The Douglas-Peucker iterative algorithm is based on the concepts of a baseline and tolerance zone. The baseline connects the start point to the endpoint, and a user-defined tolerance zone is provided to model the end character of the line. Generalization involves reducing the number of points representing a given object while maintaining its original shape.

Many scientists have attempted cartographic generalization of river networks. Because hydrography is the map element that is most sensitive to generalization (Brewer et al., 2011), it is quite challenging to develop generalization methods that would be universal and useful. Christensen (1999) presented a proposal of line generalization based on the simplification of a coastline and transformation of its medial axis. It has been shown that the chosen method maintains the general shape of the line even during large-scale changes and enables the aggregation and elimination of objects where needed. Perkal's research was used in this study (Perkal, 1966). Perkal's theory assumed an objective method of generalization based on the geometric transformation of each area into a generalized area.

In similar research, Touya tried to answer the question of which watercourses are essential to be shown on a small-scale map (Touya, 2007). The author noticed that the difficulty of hydrography generalization mainly relates to the selection of appropriate objects. The study was carried out on two IGN French National Cartographic Agency (IGN) databases. These were the BD TOPO® database with a detail level of approximately 1:15,000 as a source database and the BD CARTO® database with a detail level of approximately 1:100,000 as a target. The databases were created independently of each other. The author mentions an important issue concerning Gestalt psychology (Wertheimer, 1923/1938), as used in graphics and cartography to give the recipient an impression of perceptual clustering. The human eye notices some patterns formed by objects with different geometry and then perceives them as a whole (Spielman et al., 2020). The strokes which were listed by the author are different line segments and sets of arcs. Those strokes make an impression of continuity (Thomson & Richardson, 1999). In Gestalt psychology, such regularity is called the principle of good continuation. This principle was used in the study to develop a method of generalizing rivers on a map. The source layer contained information only concerning river sections. Thus the input data was enriched with additional attributes to make the selection possible. "Source" and "Sink" classes have been added to define the start and end of watercourses. The input data was also supplemented with two additional classes to manage the selection of individual islands located on a river. Moreover, a class has been added to store information concerning irrigation zones. Then the Strahler method (Strahler, 1952) and Horton method (Horton, 1945) were used to select watercourses. As a result, a generalized river network for the region of France of 7,600 km^2 was obtained. The last step was comparing the generalization results to the BD CARTO® database.

River hierarchy and data enrichment were also used in Stanislawski's work (Stanislawski et al., 2014). The author presents an algorithm of simplifying a river network using density analysis and data enrichment. The algorithm selects both natural and artificial watercourses at various scales. Four stages of the workflow were suggested: enriching the source data, eliminating the number of watercourses, simplifying sections of the braided rivers, and adjusting the density of the watercourses. The elimination of watercourses using the presented algorithm takes place within two stages. Initially, all rivers shorter than the defined minimum length threshold are removed. Then, the number of watercourses is limited in those network segments where the watercourse density is too high. Watercourses are removed only if all three previously calculated values (S – the river's hierarchy according to Strahler's classification, L – the distance to the farthest source, B – the number of branches of the watercourse counted from the source) are below the user-defined thresholds. The algorithm changes if there are braided rivers. Then its operation is concentrated not on rivers but on islands. If the size of an island is below a user-defined threshold, it is either merged with nearby islands or enlarged. The final process is related to adjusting the watercourse density. Density is calculated using a grid that divides the data set into individual cells. The algorithm computes the average cell density of the input and generalized data, and then the difference between these values is calculated. At the last stage, the algorithm equalizes the variability in the network density by eliminating or supplementing the watercourses in individual cells until the threshold of the set value is reached.

An attempt to simplify rivers based on the Map of Hydrographic Division of Poland (MPHP) database was conducted in the study by Szombara and Lupa (2018). The authors proposed a new solution for generalizing a river network on large-scale and small-scale maps. In the research General Geographic Objects Database (GGOD) and the MPHP were used. An attempt was made to carry out the cartographic generalization of the GGOD from the detail level corresponding to the 1:250,000 scale, to 1:500,000, and 1:1,000,000 scale. Only the layer named "SWRS_L" from the GOOD was used. This layer refers to linear hydrographic objects and includes all rivers and streams in Poland. The data of the rivers in both GGOD and MPHP databases were consistent with each other. In the research by Szombara and Lupa (2018) river selection and simplification were conducted. Before the simplification processes, all watercourses up to 750 meters length were eliminated from the database. The research methodology was based on two stages. The first step was to apply the hierarchy in the database. The objects were sorted according to the order of watercourses (Id_MPHP), starting from the lowest value. Then, each object was analysed (both a segment and the entire line), and if more than one linear feature had the same Id_MPHP number, the sections were joined. The watercourses were selected according to one of two criteria:

– The object recognition standard that is described on the basis of an elementary triangle. It is the concept of a model, a standard triangle with the most diminutive dimensions, ensuring line recognition. Its dimensions depend on the scale of the map and the width of the line representing the object (Chrobak et al., 2016).

– The criteria of the total length less than 750 m, specified in the regulation of the Minister of the Interior and Administration of 17 November 2011 on the database of topographic objects and general geographic objects, as well as standard cartographic studies.

Based on the criterion of the object recognition standard, it is assessed whether the watercourse meets the recognition standard. The standards used in the tests consider the shorter arm and the base of the elementary triangle. The standard also considers the line width specified by the user as one of the input parameters. The identification of the watercourse is checked by adding up the lengths of the triangles and comparing them with the value from the regulation. If the watercourse is recognizable, the algorithm moves to the following ID_MPHP number. Otherwise, the watercourse is removed, and the two following sections of the higher hierarchy watercourse are joined. All watercourses in GGOD meet the object recognition standard within generalized maps in this research. While in the case of 1:1,000,000 map, individual watercourses do not meet the criteria of recognition. In the case of MPHP generalization, approximately 20% of watercourses remain at 1:500,000 scale, while on the scale of 1:1,000,000, only approximately 5% of watercourses meet the criterion. The authors conclude that while developing maps based on GGOD, it is challenging to generalize objects to the level of detail corresponding to 1:500,000 and 1:1,000,000 scale. The reason for this difficulty is that the attributes of objects associated with data visualization in these scales are incorrect or lacking. The article states that in order to be able to generalize the objects from the GGOD, it would be necessary to enrich it with further thematic attributes. As a result, the number of visible objects on the map at the detail level of 1:500,000 and 1:1,000,000 scale would be larger. The study also showed that the MPHP database could be used for more effective small-scale topographic maps design, especially to enrich the GGOD database. According to the authors, the visual analysis of the generalization results shows that the presented standard of object recognition is suitable for simplifying the river network derived from the GGOD database.

Van Altena and Stoter also dealt with the automatic generalization of hydrography (Van Altena & Stoter, 2016). The scope of the study covered the simplification of the Danish water supply network at the detail level corresponding to the 1:10,000 scale. The research methodology consisted of four stages: data preparation, application of three existing algorithms for object selection, evaluation of the results for individual areas, and comparison of the results of each algorithm. The authors noticed that the algorithms commonly used to reduce the number of objects on the map, resulted in the loss of continuity of some watercourses. This result was undesirable and needed improvements. It was decided to use information about the landscape in which a given river network is located. The authors divided the current approaches of river network selection into three categories:

– semantic selection that uses the features of the river network (i.e., the type of watercourse, width, order of importance, which does not guarantee the correct topology and geometry of the objects (Liu et al., 2010),

– selection based on perceptual grouping, so called “stroke-based approach” that selects single linear objects based on their significance for maintaining essential network features (Thomson & Richardson, 1999),

– mesh-based approach that reduces the density of the river network in a given cell of the previously created mesh (Chen et al., 2009).

The study was carried out using appropriate algorithm implementations based on the stroke and mesh-based approaches. The Esri “Thin Road Network” tool was also used for the selection. The first stage of the work was data preparation, which consisted of the automatic identification of landscape types. The identification consisted of calculating the humidity statistics of the selected area by comparing the ratio of wet to dry areas and then determining the dominant shape of the waterline. The next step of data preparation was a repair of the geometry of the river network. Relevant line features have been added to the existing river network to improve its continuity.

The authors identified three measures leading to an objective comparison of the results:

– significant reduction of the details on the map,

– maintaining the similarity of the results to the structures in the input data,

– maintenance of the continuity of the network.

Based on these measures, it was decided that the most appropriate approach was the use of the “Thin Road Network” tool. The mesh-based algorithm turned out to be the least useful approach. It has also been shown that using various algorithms does not cause differences in the obtained results for particular landscapes types. Therefore, using different types of selection for different landscapes, does not improve the overall selection results for the investigated area.

3. Research methods

To achieve this study’s research goals, it was necessary to complete the stages shown in Figure 1. The first step was to select the appropriate test areas. After that, hydrography data was obtained from the GGOD. The reference scale for GGOD is 1:250,000. It is a database that contains data concerning geographic objects located in Poland. This study uses GGOD layers containing hydrography objects, namely “OT_SWRS_L” (rivers), “OT_SWKN_L” (channels), and “OT_PTWP_A” (water reservoirs). Each river from the database contains the information concerning width, length, type, course, intermittency, location (below or above the ground), navigability, and an identifier of the watercourse from the Map of Hydrographic Division of Poland (MPHP).

The second step was the selection of the previously obtained data using the variables included in the regulation of the Ministry of Interior and Administration (MSWiA, 2011). This stage also includes improving the selection carried out according to the guidelines of the regulation to obtain better results of river network selection. Three selection variants based on variables from the regulation were implemented:

– Basic variant, based only on the map specifications.

– Extended variant, based on the map specification and with the use of Id_MPHP index.

– Extended-modified variant, based on the map specification with the use of the division into priority watercourses (fig. 1).

Results were assessed and compared to a reference atlas map at 1:500,000 scale contained in the Atlas of the Republic of Poland (Polish Academy of Sciences. Institute of Geography and Spatial Organization, 1993).

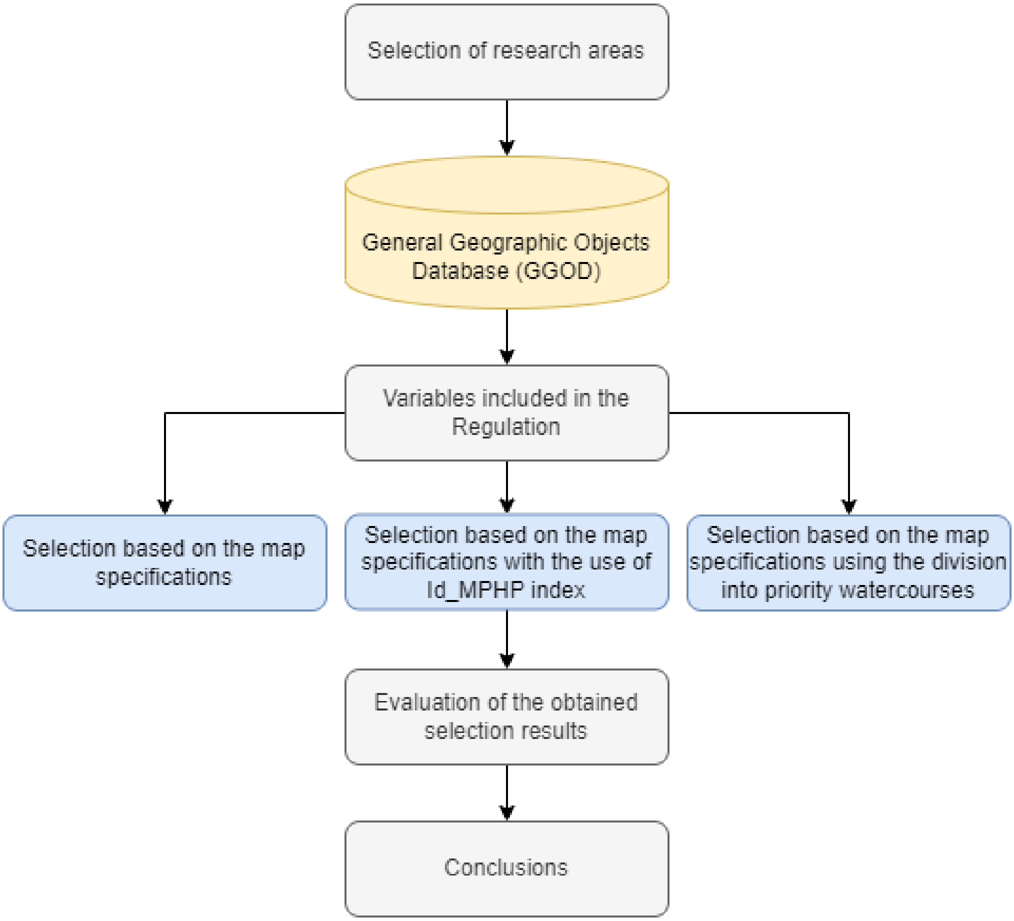

Fig. 1. Research schema, own elaboration. The yellow colour symbolizes the database used, the grey colour represents the methodology stages, and the blue colour represents particular variants of selection

The last element of the research methodology was to draw conclusions based on the comparison of the implemented variants.

The scope of the research covered three river basins. While selecting the river basin areas, the primary considerations were the variety of terrain where basins are situated as well as the different nature of the watercourses. These factors play a significant role since they help to verify the correctness of the algorithm implementation for rivers located in the diverse areas. The river basins selected for the study are located in three different geographical regions:

– Nysa Kłodzka basin,
– Biebrza basin,
– Radomka basin.

The selected river basins are presented in figure 2.

The basin of the Nysa Kłodzka River is located in the southern part of the Dolnośląskie Voivodeship and the western part of the Opolskie Voivodeship. The basin area covers 4,874.1 km², and the river is 182 km long. The Nysa Kłodzka is situated within twenty-three topographically diverse mesoregions (Regionalny Zarząd Gospodarki Wodnej we Wrocławiu, n.d.). The river starts in the Śnieżnik Massif, and flows through the Kłodzka Valley, the Bardzkie Mountains,

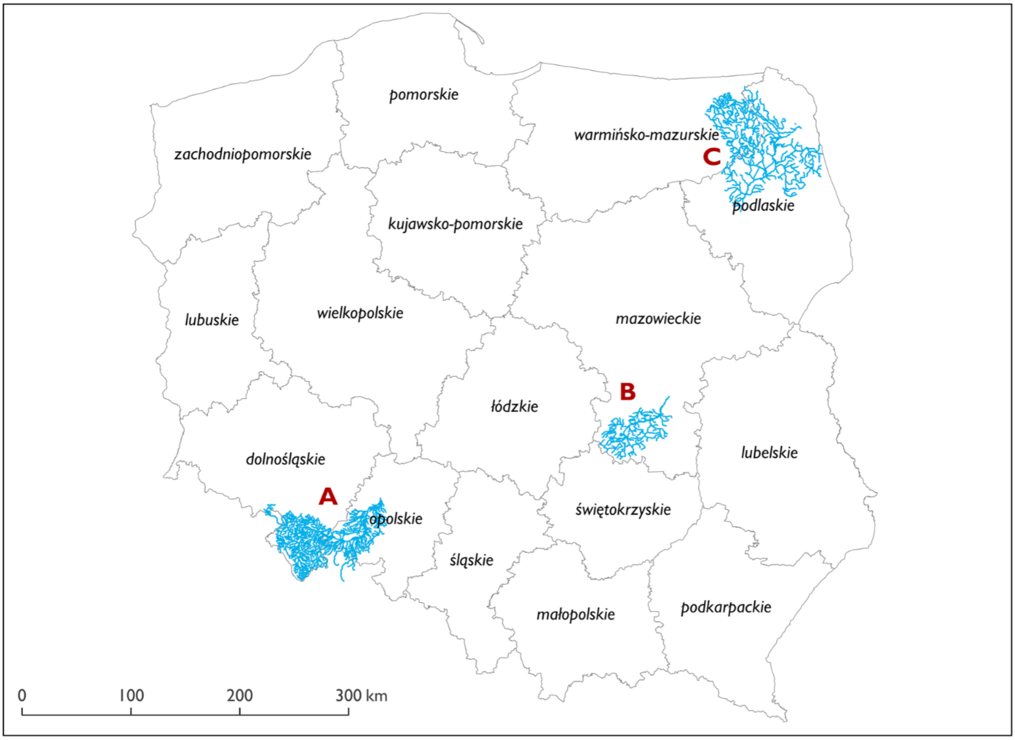

Fig. 2. Selected research areas in Poland, own elaboration. A – the Nysa Kłodzka basin, B – the Radomka basin, C – the Biebrza basin

the Otmuchowskie Depression, and the Nysa Kłodzka Valley. The river flows into the Wrocław Urstromtal where Wrocław Urstromtal joins the Odra River (Wydawnictwo Naukowe PWN, n.d. a).

The Biebrza River is located in the area of the North Podlasie Lowland. It has its source at the foot of the Sokólskie Hills and flows through the Biebrza Basin. The basin covers an area of 7,057.4 km², and the river is 162.8 km long. A remarkable feature of the Biebrza River basin is its significant asymmetry. The right bank part of the river constitutes 75.5% of the whole river basin, while the left bank part only 24.5% (Zawadzka, 2020).

The basin area of the Radomka River is located within the borders of the Masovian Voivodeship. Its area covers approximately 2,100 km². The Radomka is a left tributary of the Vistula River with a total length of 106 km. It flows from the northern part of Garb Gielnowski and flows through the Radom Plain and Kozienice Plain area (Wydawnictwo Naukowe PWN, n.d. b).

The topography of selected areas of the basin area is diverse, which influences the diversity of the river basins located in the area. The difference is visible on the example of the course of the river system and the change of the tortuosity or the width of individual watercourses. Therefore, we believe the three different basins selected for the analysis constitute a good research sample.

After downloading the data covering the research areas from the website dane.gov.pl (Otwarte Dane, n.d.),[3] the next step was to prepare the hydrography data. Hydrographic objects were selected from the GGOD. The purpose of this selection is to generalize the data from the GGOD to a detail level corre-

[3] https://dane.gov.pl/pl/dataset/782,baza-danych-obiektow-ogolnogeograficznych

sponding to the 1:500,000 scale. In the generalization process, the regulation of the Minister of Internal Affairs regulation of 17 November 2011 on the database of topographic objects and the database of general geographic objects and standard cartographic studies (MSWiA, 2011) was used. The regulation contains guidelines for the generalization of large-scale maps, elaborated on the basis of Topographic Objects Database (TOD) and GGOD. However, this act is no longer in force. In the Polish law, there is also the regulation of the Minister of Development, Labor and Technology of 27 July 2021 on the database of topographic objects and the database of general geographic objects and standard cartographic studies (MRPiT, 2021). The current regulation differs from the previous one only concerning the information about the class of objects and the method of obtaining the vector layers from the GGOD needed for the selection. As the regulation from 2011 is more detailed and not contradictory to the new one, for this analysis, the guidelines from both 2011 and the current regulation were used to facilitate the selection process of GGOD.

ArcMap version 10.8.1 and its "ModelBuilder" functionality were used to automate the selection according to the specifications contained in the regulation. The program allows for the creation, editing, and management of models that enable the systematization of individual geoprocessing tools into appropriate sequences, thus automating the selection process. According to the regulation, three variants of generalization models have been developed: the basic, extended, and extended-modified model.

3.1. Model designed using the Ministry of Interior and Administration guidelines (basic variant)

The first model was designed based only on the guidelines of the Ministry of Interior and Administration. The following variables were taken into account:

– watercourse length of over 10 km (2 cm in the map scale) in lowlands and in highlands over 7.5 km long (1.5 cm in the map scale) in mountain areas,

– watercourse width, less than or equal to 250 m (0.5 mm in the map scale),

– operational status (navigable and non-navigable watercourse),

– reservoir type (stagnant and flowing water).

In the initial stage, the watercourses in a linear form were considered. Although there is no information in the regulation concerning the necessity to connect sections of watercourses into rivers, they were joined in order to maintain topological continuity. As it is logical from the cartographic practice point of view, the watercourses were first joined using the "unsplit line" tool in ArcMap. The next step constituted the value of the length attribute calculation, as it was missing in the source database. Then, based on the calculated geometry, only the watercourses longer than 10 km were selected. In the next step, watercourses less than or equal to 250 m wide were selected. The next stage was to divide the rivers into navigable and non-navigable according to the "Z" and "NZ" attribute values. The next part of the model was the selection of channels with the operating status "Z", i.e., navigable channels. The remaining watercourses were then saved as a shapefile layer. The last stage of the model was the selection of stagnant and flowing waters. In order to obtain the surface area of rivers, a water layer was also selected from the database with the "Pp" attribute (flowing waters).

3.2. Model designed using the guidelines of the Ministry of Interior and Administration and additionally using the Id_MPHP index (extended variant)

This model was designed with the use of the Ministry of Interior and Administration regulation, but also with the "Id_MPHP" index enrichment to the source database. The selection rules contained in the regulation did not take into account the Id_MPHP index. However, expecting that it may be an essential and helpful selection criterion, it was included in this model variant alongside the other selection rules. The Id_MPHP index contains the data from the MPHP database, which level of detail corresponds to the 1:50,000 scale. In the MPHP the hierarchy of watercourses in the overall river network has been defined. As a next step in the developed model the "dissolve" tool was used based on the idMPHP attribute, allowing the aggregation of individual river sections.

The length attribute was then added to allow the selection of watercourses over 10 km long. Subsequently, the division of the layer with the “intersect” tool was used to supplement the attributes lost during the aggregation process. Thanks to this, it was possible to perform the selection again based on width smaller or equal to 250 m. The navigable and non-navigable rivers were then distinguished according to the “Z” and “NZ” attribute value. The subsequent selection steps were carried out in the same order as the selection under the regulation guidelines. The next step was to select the channels with the “Z” operating status, i.e., navigable channels, and then save the layer in shapefile format. The last step in the extended model was the selection of stagnant and flowing waters. They were selected based on the attribute specifying the nature of a lake or pond by the type “Ps” and a surface area greater than or equal to 500,000. From the water reservoirs layer, those with the attribute equal to “Pp” were also selected to obtain the polygonal layer of rivers.

3.3. Model designed with the use of the guidelines of the Ministry of Interior and Administration and the division into priority watercourses (extended-modified variant)

This is another experimental variant of the model designed with the use of the guidelines of the Ministry of Interior and Administration as well as with the additional attributes. The Id_MPHP index from the GGOD was also used in this variant. Additionally the source database was enriched with the attribute denoting “priority” of the watercourses. In the case of this selection, the first step was to use the “dissolve” tool based on the Id_MPHP attribute, which allowed the aggregation of individual river sections. The second step consisted of selecting “priority” watercourses, the Id_MPHP index of which was less than or equal to the specified n-digit number.[4] The value of the number “n” was selected experimentally, visually assessing its usefulness. Then the length attribute was added to allow further the selection of watercourses over 10 km long. Subsequently, the watercourse layer was split using the “intersect” tool to supplement the attributes that were lost during the aggregation process. As a result, it was possible to re-use the selection based on a width less than or equal to 250 m.

In the next stage, the navigable and non-navigable watercourses were distinguished according to the “Z”, and “NZ” attributes. The next stages of selection did not differ from the basic variant’s. The next step was to select channels with the “Z” operating status, i.e., navigable channels, and then save the layer to a shapefile. The last element of the model was to select the stagnant and flowing waters. They were selected after the character attribute of a lake or pond of the type “Ps” (stagnant waters) and a surface area greater than or equal to 500,000. From the water layer, those with the attribute “Pp” (flowing waters) were also selected to obtain the polygonal layer of the rivers.

4. Results

The results constitute two elements. The first covers the three selection variants implemented in the form of generalization models. The second is the generalized river network at the detail level corresponding to the 1:500,000 scale, obtained from three variants of selection implemented with the use of the guidelines of the regulation and enriched with additional river attributes. In Figures 3–5 the three variants of the generalization models designed for the selected basins are presented.

Figures 6–8 show the results of the automatic selection performed with the use of the “ModelBuilder” application. The results include selection according to the regulation (basic variant), selection according to the regulation with the enrichment of the Id_MPHP index (extended variant), and selection according to the regulation using both the Id_MPHP and the division into priority watercourses (extended-modified variant). The resulting maps also include water reservoirs and channels, generalized in the basic variant and added to all maps to maintain the river network continuity.

5. Discussion and summary

Based on the visual analysis it can be noticed that none of the three selection variants entirely

[4] Id_MPHP has its own numeric order. The higher the value of “n”, the higher the order of the watercourse which should be shown on the map. The maximum value for “n” is 9.

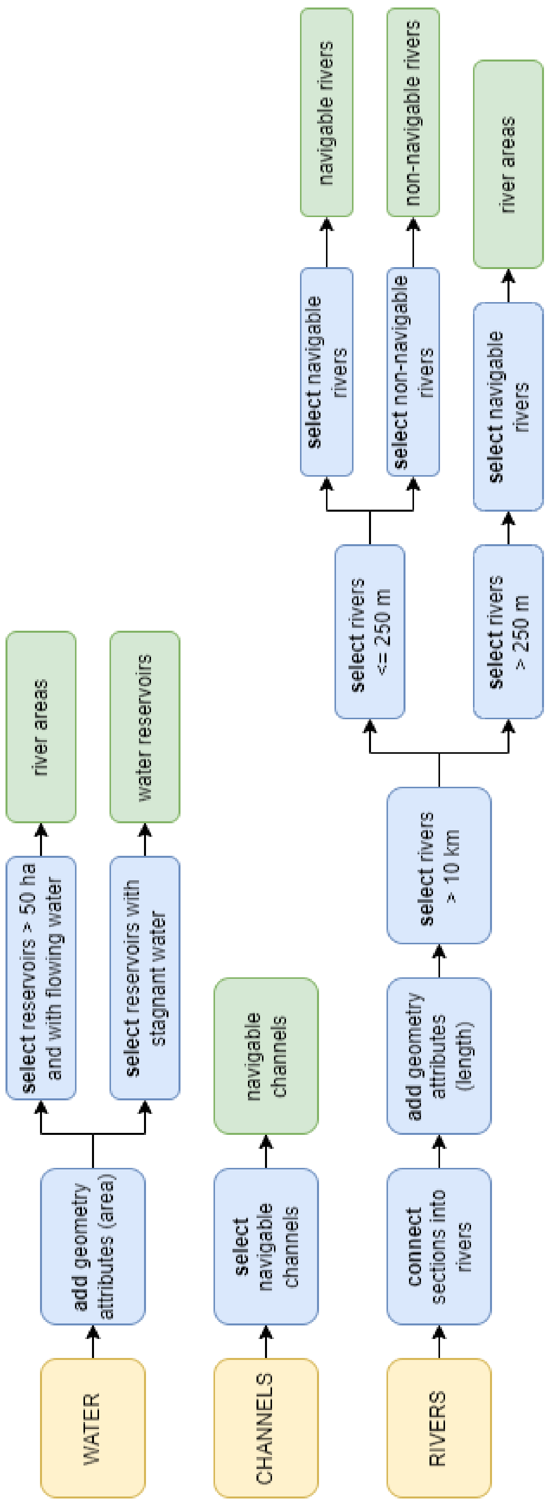

Fig. 3. Generalization model according to the guidelines of the regulation (basic variant). The yellow colour symbolizes the input layers, the blue colour represents individual stages of selection, and the green colour represents the output layers. Own elaboration

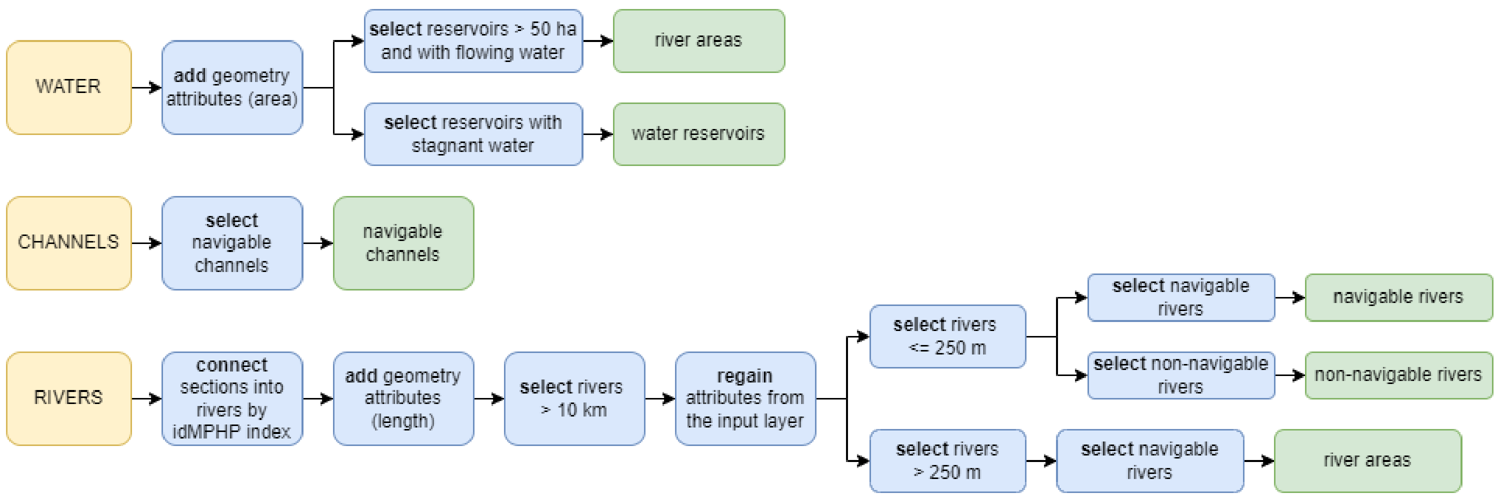

Fig. 4. Generalization model according to the guidelines of the regulation with the use of Id_MPHP index (extended variant). The yellow colour symbolizes the input layers, the blue colour represents individual stages of selection, and the green colour represents the output layers. Own elaboration

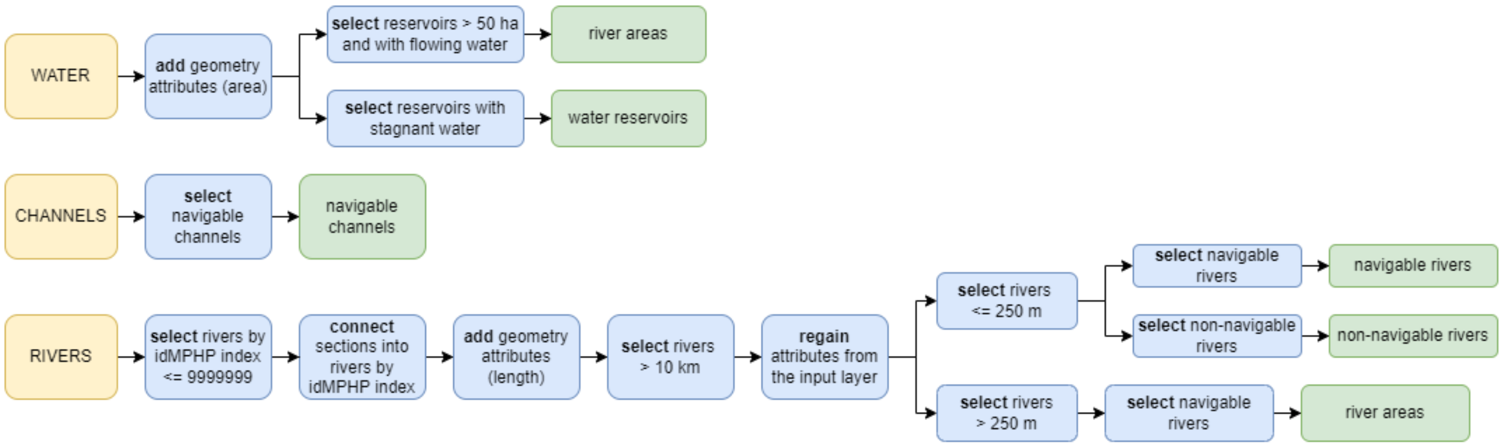

Fig. 5. Generalization model according to the guidelines of the regulation using the division into priority watercourses (extended-modified variant).
The yellow colour symbolizes the input layers, the blue colour represents individual stages of selection, and the green colour represents the output layers.
Own elaboration

coincides with the river network acquired from the reference atlas map at 1:500,000 scale. The worst performing variant is the basic one constituting the selection according to the guidelines contained in the regulation. In the case of this variant, the selected watercourses do not constitute an integral network. It is caused by the fact that the generalization rules specified in the regulation are not adjusted to the structure of the GGOD. The watercourses from the GGOD are divided into sections. Each change in the watercourse width divides it into

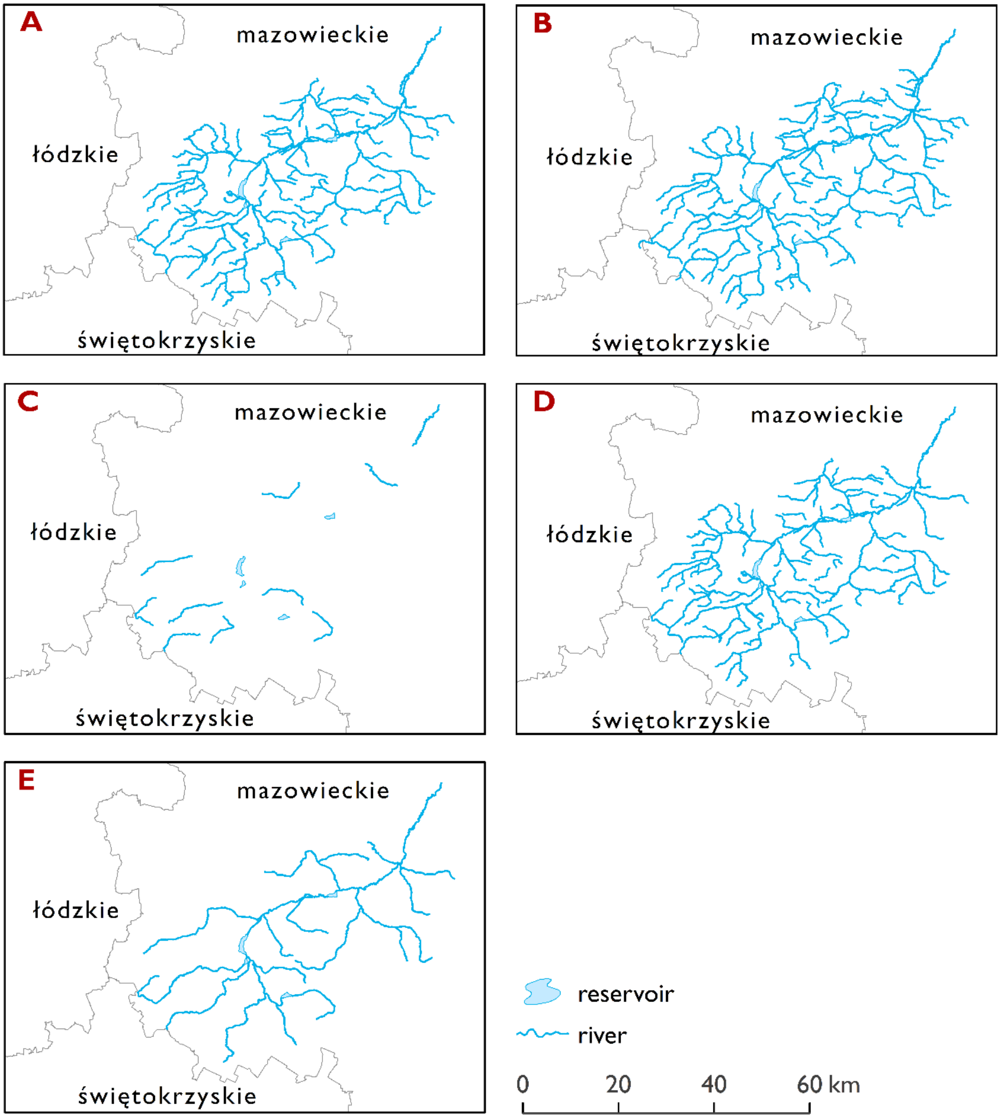

Fig. 6. The results of the generalization of the Radomka basin. A – source data, B – vectorized data from the reference atlas map, C – selection according to the regulation, D – selection according to the regulation using the idMPHP index, E – selection according to the regulation with the use of priority watercourses. Own elaboration

smaller segments. Therefore, each time the attribute related to the watercourse length is used in the selection process, individual segments of the river are selected instead of the longer river fragments.

In order to enable more efficient selection according to the regulation guidelines, the segments of watercourses belonging to one watercourse should be joined. This was applied in the extended variant. As a result, in the extended variant the best selection outcomes were obtained, namely the closest in comparison to the reference atlas map (Figures 6D–8D). Due to the connection of rivers with the use of

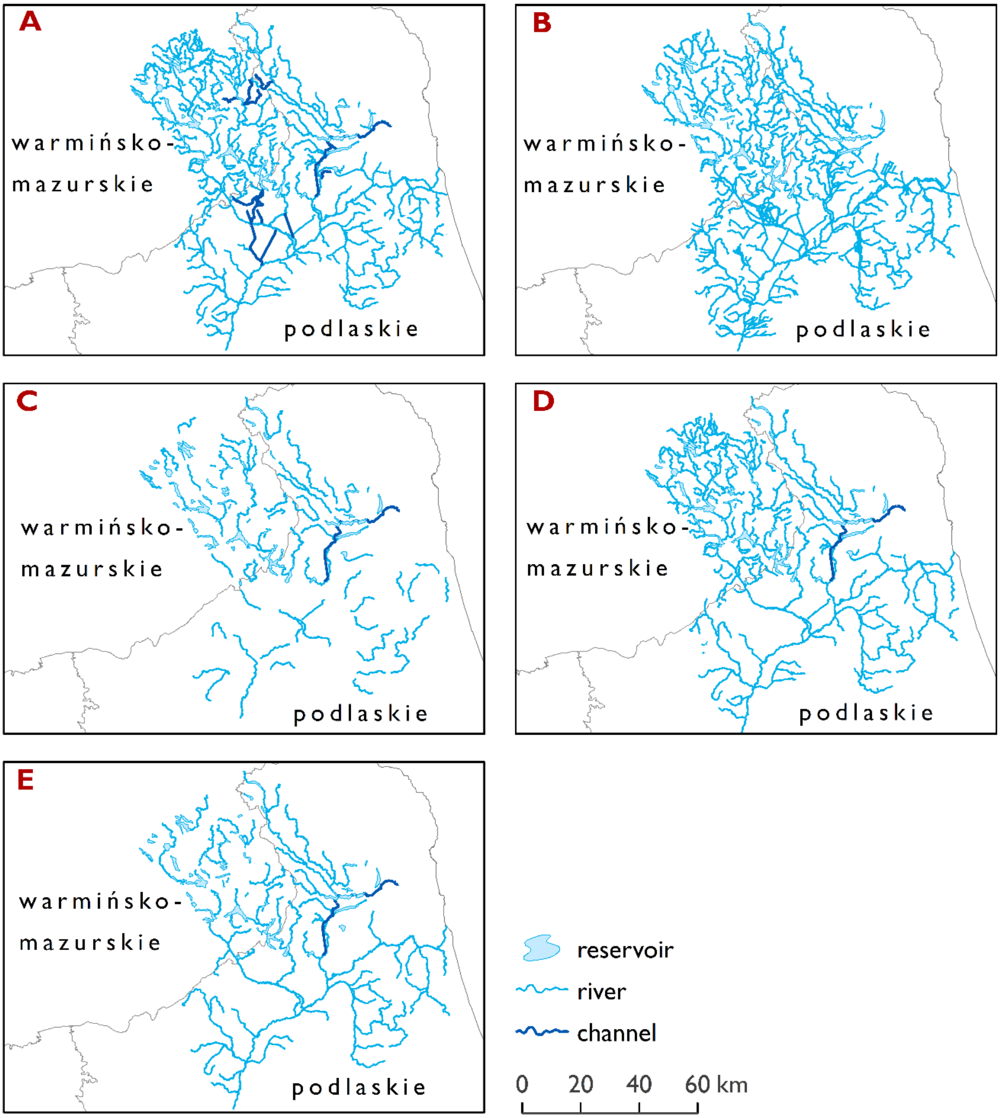

Fig. 7. The results of the generalization of the Biebrza basin. A – source data, B – vectorized data from the reference atlas map, C – selection according to the regulation, D – selection according to the regulation using the idMPHP index, E – selection according to the regulation with the use of priority watercourses. Own elaboration

Id_MPHP index, the number of individual segments of the watercourses is much smaller in comparison to the basic variant. What is more, this solution allowed for maintaining the general character of the river network.

The results obtained from the selection with the use of the division into priority watercourses show a homogeneous river network not containing any single river segments (Figure 6E–8E). On the other hand, a significant disadvantage of the extended-modified variant is the insufficient number of watercourses compared to the number of rivers present on the reference atlas map. The deficit is caused by some watercourses

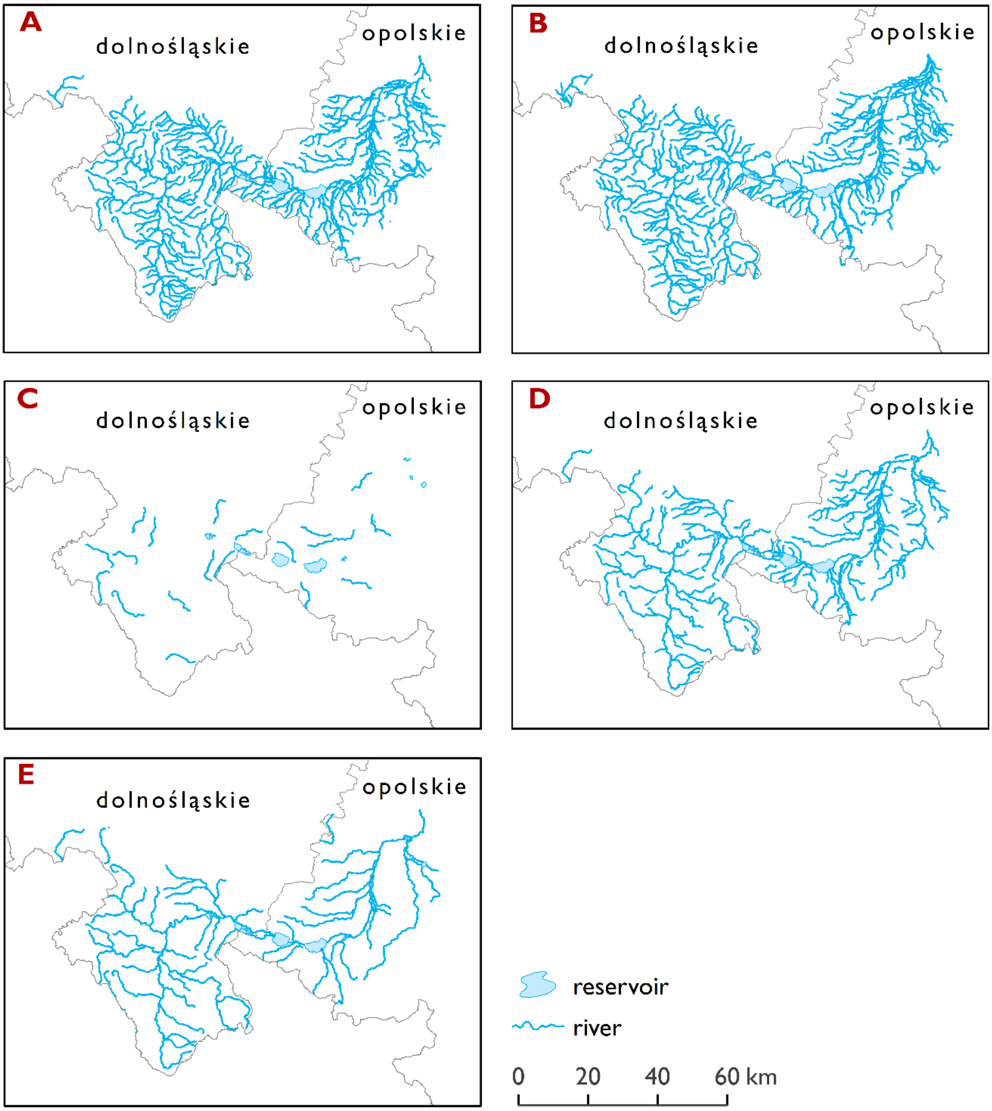

Fig. 8. The results of the generalization of the Nysa Kłodzka basin. A – source data, B – vectorized data from the reference atlas map, C – selection according to the regulation, D – selection according to the regulation using the idMPHP index, E – selection according to the regulation with the use of priority watercourses. Own elaboration

with an Id_MPHP identifier equal to 0, making it impossible to connect them to the overall river network, even if the Id_MPHP number is specified as a nine-digit number (maximum value). However, the possibility of experimental determination of the n-digit value and automatic elimination of less important watercourses suggest that the extended-modified variant can be used as a basis for selection for scales smaller than 1:500,000.

Based on the conducted research, it can be concluded that although the regulation contains the information concerning the selection guidelines, it is insufficient to correctly generalize the hydrography to the detail level corresponding to the 1:500,000 scale. With this research, it has been shown that enriching the source database with the Id_MPHP index within the extended selection variant makes the selection more efficient and helps to obtain better results, which are more similar to the river network contained in the reference atlas map. On the other hand, enriching the source database with a prioritization index allows for obtaining a uniform river network and may be the basis for generalizing rivers to scale smaller than 1:500,000. Moreover, considering additional attributes may constitute an essential supplement to the regulation guidelines.

Acknowledgements. This research was funded by the National Science Centre, Poland, grant number UMO-2020/37/B/HS4/02605, "Improving Settlement and Road Network Design for Maps of Small Scales Using Artificial Intelligence and Graph Theory".

Literature

Brewer, C. A., Buttenfield, B. P., & Stanislawski, L. V. (2011). Choosing between Geometry Change and Display Change for Multiscale Mapping—The Role of Elimination in Design. *Proceedings of the 25th International Cartographic Conference, Paris, France, 3–8 July*. https://icaci.org/files/documents/ICC_proceedings/ICC2011/Oral%20Presentations%20PDF/D1-Feature%20selection%20and%20typification%20in%20generalisation/CO-310.pdf

Chen, J., Hu, Y., Li, Z., Zhao, R., & Meng, L. (2009). Selective omission of road features based on mesh density for automatic map generalization. *International Journal of Geographical Information Science*, *23*(8), 1013–1032. https://doi.org/10.1080/13658810802070730

Christensen, A. H. J. (1999). Cartographic line generalization with waterlines and medial-axes. *Cartography and Geographic Information Science*, *26*(1), 19–32. https://doi.org/10.1559/152304099782424893

Chrobak, T., Szombara, S., Kozioł, K., & Lupa, M. (2016). A method for assessing generalized data accuracy with linear object resolution verification. *Geocarto International*, *32*(3), 238–256. https://doi.org/10.1080/10106049.2015.1133721

Douglas, D. H., & Peucker, T. K. (1973). Algorithms for the Reduction of the Number of Points Required To Represent a Digitized Line or Its Caricature. *Cartographica: The International Journal for Geographic Information and Geovisualization*, *10*(2), 112–122. https://doi.org/10.3138/fm57-6770-u75u-7727

Horton, R. E. (1945). Erosional development of streams and their drainage basins; hydrophysical approach to quantitative morphology. *Geological Society of America Bulletin*, *56*(3), 275–370. https://doi.org/10.1130/0016-7606(1945)56[275:edosat]2.0.co;2

International Cartographic Association. Commission II: Definition, Classification and Standardization of Technical Terms in Cartography. (Ed.). (1973). Generalization. In *Multilingual dictionary of technical terms in cartography* (p. 137). Franz Steiner.

Liu, X., Zhan, F. B., & Ai, T. (2010). Road selection based on Voronoi diagrams and "strokes" in map generalization. *International Journal of Applied Earth Observation and Geoinformation*, *12*(Supplement 2), 194–202. https://doi.org/10.1016/j.jag.2009.10.009

Ministerstwo Rozwoju, Pracy i Technologii. (2021). *Rozporządzenie Ministra Rozwoju, Pracy i Technologii z dnia 27 lipca 2021 r. w sprawie bazy danych obiektów topograficznych oraz bazy danych obiektów ogólnogeograficznych, a także standardowych opracowań kartograficznych* [The regulation of the Minister of Economic Development and Technology of 27 July 2021 on topographic objects database, general geographic objects database and standard maps]. (Dz.U. 2021 poz. 1412).

Ministerstwo Spraw Wewnętrznych i Administracji. (2011). *Rozporządzenie Ministra Spraw Wewnętrznych i Administracji z dnia 17 listopada 2011 r. w sprawie bazy danych obiektów topograficznych oraz bazy danych obiektów ogólnogeograficznych, a także standardowych opracowań kartograficznych* [The regulation of the Minister of Internal Affairs and Administration of 17 November 2011 on topographic objects database, general geographic objects database and standard maps]. (Dz.U. 2011 nr 279 poz. 1642).

Otwarte Dane. (n.d.). https://dane.gov.pl/

Perkal, J. (1966). An attempt at objective generalization. *Michigan Inter-University Community of Mathematical Geographers*, (10), 58–76.

Polish Academy of Sciences. Institute of Geography and Spatial Organization. (1993). *Atlas of the Republic of Poland*. Surveyor General of Poland.

Regionalny Zarząd Gospodarki Wodnej we Wrocławiu. (n.d.). *Opracowanie charakterystyki zlewni bilansowej rzeki Nysy Kłodzkiej*. https://wroclaw.rzgw.gov.pl/files_mce/Planowanie%20w%20gospodarowaniu%20wodami/charakterystyka_nysa_klodzka.pdf

Robinson, A., Sale, R., & Morrison, J. (1978). *Elements of cartography* (4th ed.). John Wiley and Sons.

Spielman, R. M., Jenkins, W. J., Lovett, M. D., & Czarnota-Bojarska, J. (2020). *Psychologia*. OpenStax. https://openstax.org/books/psychologia-polska/pages/5-6-zasady-postrzegania-w-psychologii-gestalt

Stanislawski, L. V., Buttenfield, B. P., Bereuter, P., Savino, S., & Brewer, C. A. (2014). Generalisation operators. In D. Burghardt, C. Duchêne & W. Mackaness (Eds.), *Abstracting geographic information in a data rich world. Methodologies and Applications of Map Generalisation* (pp. 157–195). Springer. https://doi.org/10.1007/978-3-319-00203-3_6

Strahler, A. N. (1952). Dynamic basis of geomorphology. *Geological Society of America Bulletin*, *63*(9), 923–938. https://doi.org/10.1130/0016-7606(1952)63[923:dbog]2.0.co;2

Szombara, S., & Lupa, M. (2018). Cartographic Generalization of the River Network Including Standards of the Objects Recognition. *Proceedings of the 2018 Baltic Geodetic Congress, BGC–Geomatics 2018*, 117–120. https://doi.org/10.1109/BGC-Geomatics.2018.00028

Thomson, R. C., & Richardson, D. E. (1999). The ‘good continuation’ principle of perceptual organization applied to the generalisation of road networks. *Proceedings of the 19th International Cartographic Conference*, *Ottawa, Canada*, 1215–1223 http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=4DA1A15D31A459AF418AD8381C2062F5?doi=10.1.1.202.4737&rep=rep1&type=pdf

Tobler, W. R. (1964). *Technical report No. 1, „An experiment in the computer generalization of maps”*. Department of Geography University of Michigan. https://deepblue.lib.umich.edu/bitstream/handle/2027.42/7972/aja5779.0001.001.pdf?sequence=5&isAllowed=y

Touya, G. (2007). River Network Selection based on Structure and Pattern Recognition. *23rd International Cartographic Conference, Moscow, Russia, August 2007*, 4–9. https://doi.org/10.13140/RG.2.1.1594.4167

van Altena, V., & Stoter, J. (2016). Context-aware thinning of artificial water networks for map generalization. *Journal for Geographic Information Science*. *GI_Forum*, *4*(1), 12–29. https://doi.org/10.1553/giscience2016_01_s12

Wertheimer, M. (1938). Laws of organization in perceptual forms. In W. D. Ellis (Ed.), *A source book of Gestalt psychology* (pp. 71–88). Kegan Paul, Trench, Trubner & Company. (Original work published 1923). https://psychclassics.yorku.ca/Wertheimer/Forms/forms.htm

Wydawnictwo Naukowe PWN. (n.d.). Nysa Kłodzka. In *Encyklopedia PWN: źródło wiarygodnej i rzetelnej wiedzy*. https://encyklopedia.pwn.pl/haslo/Nysa-Klodzka;3949141.html

Wydawnictwo Naukowe PWN. (n.d.). Radomka. In *Encyklopedia PWN: źródło wiarygodnej i rzetelnej wiedzy*. https://encyklopedia.pwn.pl/haslo/Radomka;3965526

Zawadzka J. (2020, April 7). *Wody. Biebrzański Park Narodowy.* https://archiwum.biebrza.org.pl/46,wody