

# Evaluating map specifications for automated generalization of settlements and road networks in small-scale maps

## Abstract

The presented research concerns the methodology for selecting settlements and road networks from 1:250 000 to 1:500 000 and 1:1 000 000 scales. The developed methodology is based on the provisions of the Regulation of the Ministry of Interior from 17 November 2011. The correctness of the generalization principles contained in the Regulation has not yet been verified. Thus this paper aims to fulfil this gap by evaluating map specifications concerning settlement and road network generalizations.

The goal was to automate the selection process by using formalized cartographic knowledge. The selection operators and their parameters were developed and implemented in the form of a generalization model. The input data was the General Geographic Object Database (GGOD), whose detail level corresponds to 1:250 000 scale. The presented research is in line with works on the automation of GGOD generalization performed by the National Mapping Agency (NMA) in Poland (GUGiK). The paper makes the following contributions. First, the selection methodology contained in the Regulation was formalised and presented in the form of a knowledge base. Second, the models for the generalization process were developed. The developed methodology was evaluated by generalizing the settlements and roads in the test area. The results of the settlement and road network generalization for both 1:500 000 and 1:1 000 000 detail levels were compared with the maps designed manually by experienced cartographers.

## Keywords

Automated map generalization • small-scale map specifications • evaluation of generalization principles

© University of Warsaw – Faculty of Geography and Regional Studies

**Karolina Maja Sielicka,  
Izabela Karsznia**

Department of Geoinformatics, Cartography and Remote Sensing, Chair of Geoinformatics and Information Systems, Faculty of Geography and Regional Studies, University of Warsaw, Poland  
e-mail: [k.sielicka@uw.edu.pl](mailto:k.sielicka@uw.edu.pl)  
e-mail: [i.karsznia@uw.edu.pl](mailto:i.karsznia@uw.edu.pl)

Received: 25 December 2018

Accepted: 3 June 2019

## Introduction

Automatic generalization is an essential technological step in the development of modern cartography. It can be a response to the legally imposed obligation to automate the process. An additional motivation is the growing demand of the information society for up-to-date and available data and maps at various detail levels. Currently, despite the strong computerisation and automation in many fields, effective solutions for the automation of selection have not been presented in the scientific community. Until now, scientists have devoted little attention to small-scale maps, which are updated less frequently.

Even partial replacement of the cartographer's work with an automated generalization algorithm would significantly reduce the time required to design a map, the costs and the number of people involved. These challenges have been known to both the scientific community and public administration for a long time. Benefits from automation have been noticed, and the small-scale map specifications have been defined by Polish law, in the form of Geodetic and Cartographic Law (1989.) and Regulation on the Topographic Objects Database and General Geographic Objects Database, as well as standard cartographic products (MSWiA 2011). However, these map specifications have not been verified so far. The motivation of this paper is thus to fulfil this gap by evaluating map specifications concerning small-scale maps

contained in the Regulation. The article presents the formalization and implementation of the rules from the Regulation. This is the first stage in developing a comprehensive solution for automatic small-scale map generalization in Poland.

## Related works

The issue of generalization has been undertaken by research centres, national mapping agencies and commercial companies for years. The subject of this research is very often settlements and road networks, because of the key role they play in the maps.

## Settlements selection

The research conducted on automation of settlements generalization aims at finding a universal answer to the question "how many and which settlements should be shown on the map?" The goal is to correctly present settlements density as well as their most essential elements. The methods that are used to determine the number of settlements to be displayed on the map include, among others: generalization algorithm based on the Voronoi diagram and the Delaunay triangulation models (Ai & Liu 2004; Yan & Weibel 2008), the centroid algorithm (MacQueen 1967) and the Iterative Self-Organizing Data Analysis Technique (ISODATA; Li 2007).

Nevertheless, the answer to the question of which cities should be kept seems to be more complex. In recent research, promising results concerning effective settlements selection methods are reported by Karsznia and Weibel (2018), who used models based on decision trees and genetic algorithms, and also by Bereuter and Weibel who used quadtrees (Bereuter & Weibel 2013; Bereuter & Weibel 2017).

### Roads selection

Researchers have developed a number of methods to simplify road structures that give satisfactory results. Amongst them, two method groups can be distinguished – one where each road segment, usually called a 'stroke,' is analysed individually, and a second group where the road network is analysed as a whole structure.

The first group of methods assumes evaluating the importance of each particular stroke based on its geometrical and attribute properties (Liu et al. 2010). On the other hand, the second group involves methods in which roads are treated as a coherent network, helping maintain their characteristic structure, especially the network density differences as well as their typical elements. For this purpose, graph theory is often used (Mackaness & Beard 1993; Jiang & Claramunt 2004; Benz & Weibel 2014; Weiss & Weibel 2014). However, in the graph-based methods mostly the geometrical road network properties are taken into consideration. This can lead to the deletion of road segments that are important from an administrative or functional point of view. Thus crucial semantic road network properties should also be taken into account in the generalization process.

Most road generalization methods involve making a decision about which road segments (strokes) to leave and which to omit. However, there are also methods of simplification, which reduce the number of vertices to simplify the road network, i.e. Mackaness and Mackechnie's (1999). As a result, several strokes are replaced by one new stroke.

In recent years, the use of artificial intelligence to explore the principles and rules of generalization has been yielding better and better results. Balboa and Lopez (2008) propose the method of artificial neural networks for the automatic classification of road segments. Learning the algorithm takes place in test areas and is carried out by experienced cartographers. The segments can then be easily selected based on the weights from the classification.

### Automated generalization within National Mapping Agencies

Research on the automation of spatial data generalization has been conducted for many years in numerous scientific units, including the French National Geographic Institute (IGN), the British Ordnance Survey, the Kartverket - Kartographic Agency of the Netherlands, the Kadaster - Dutch Cartography Agency in the cooperation with the Dutch Institute of Geoinformation and Earth Observation and also the Universities in Hannover, Zurich and Munich. European centres can cooperate and exchange experiences and knowledge within the framework of activities organized by European Spatial Data Research (EuroSDR) – an organization linking national mapping agencies with research centres and commercial companies. Some of the national mapping agencies have already developed solutions that allow complete automation of the generalization process every time the data is updated. The fully automated mapping process was introduced, among others, by the United Kingdom, Turkey, the United States, Norway, the Netherlands, Denmark, Catalonia, Germany and Switzerland (Stoter et al., 2014; Karsznia & Weibel, 2018). As most agencies have developed their own generalization methods, the resulting materials cannot be used to create supranational maps, because they differ in scale, content range,

cartographic presentation, database structures, as well as the software in which they are developed. The solutions used by national agencies, however are not made public. The algorithms and parameters used are rarely published, thus seldom described in detail.

### The methodology of generalization of settlements and road networks in accordance with the Regulation

Geodetic and Cartographic Law imposes the obligation to prepare and update maps on the Surveyor General of Poland, and thus he is also responsible for automating this process. Regulation, on the other hand, contains a set of map specifications defining the mode and technical standards of map design in a basic scale series. In particular, it includes the principles of selecting objects, resolving graphic conflicts and a collection of cartographic symbols together with their priorities. The order of information provided in annexes to the Regulation suggests the following generalization stages:

- Obtaining data (by selecting from the source object classes),
- Applying appropriate graphic symbols,
- Resolving graphic conflicts among objects.

Small-scale maps are designed based on the General Geographic Object Database (GGOD), whose detail level corresponds to the 1:250 000 scale. Settlement point layer, in small-scale maps, contains objects selected from GGOD's class of *settlement point features* (OT\_ADMS\_P). The Regulation contains detailed guidelines and object selection principles. At the selection stage at 1:500 000 scale, it is necessary to take into account the population density of the district in which the settlement is located. At the detail levels corresponding to 1:500 000 and 1:1 000 000 scales, six types of settlements are distinguished based on the type of settlement, the administrative significance and the number of inhabitants (population).

Roads contained in the General Geographic Objects Database are obtained from the *road class* (OT\_SKDR\_L), from the thematic group *roads and related objects*. Selection of the most important roads, which should be presented on the small-scale maps, is given in the principles contained in the Regulation. It is based on the following attributes: road class, number of lines, the category of existence and type of geometric representation.

At 1:500 000 scale, nine types of roads are distinguished, while at 1:1 000 000 scale, six types of roads are considered. Each road type is represented with a different symbol.

### Implementation of the generalization methodology

This research aimed to verify and evaluate the correctness of generalization process carried out on the basis of map specifications from the Regulation<sup>1</sup> with particular emphasis on the results of the selection process. The selection was carried out in the ArcGIS program. The process was built in ModelBuilder using the available tools.

Based on the principles contained in the Regulation, generalization models have been designed for each thematic layer and detail level considered. It allows generalization to be carried out nearly fully automatically. The processes include such steps as: data enrichment, selection and symbolization. Enriching the data includes adding additional attributes to the source data, the value of which is important in the selection. The selection stage consists of selecting objects using the given algorithm with specific parameters. Symbolization consists of assigning objects with an appropriate symbol, as defined in the Regulation.

<sup>1</sup> Regulation of the Ministry of Interior on 17 November 2011 on the Topographic Objects Database and General Geographic Objects Database, as well as standard cartographic products, Journal of Laws of 2011, No 279 item 1642.



Figure 1. Research area.  
Source: designed by K. Sielicka

The process was carried out to generalize GGOD for the area of Poland, while the test area was constrained to the border area of the Mazovian, Lublin and Świętokrzyskie Voivodeships (Figures 1 and 2). The Vistula River crosses it from the South to the North. The size of the area is 47 km (West-East) by 55 km (North-South).

The research area includes 2585 km<sup>2</sup>, and within its borders there are 550 settlements and 604,5 km roads in GGOD.

#### Detail level corresponding to 1:500 000 scale

##### Settlements generalization

The first stage in the settlements generalization is the selection. In the designed model, selection is based on settlement attributes and relations to other settlements. Selection is conducted in three steps (Figure 3, Table 1). The first step in the model involves enriching attributes by adding information about population density within particular districts. The information concerning population density within districts can be acquired from the database maintained by the Central Statistical Office of Poland, or can be calculated based on the attributes contained in GGOD. In this study, the population density was calculated based on the number of inhabitants of the settlements and the area of districts from GGOD. Then the selection was conducted based on administrative significance, number of inhabitants and the district's population density.

Settlement selection rules defined in the Regulation assume retaining on the target map the settlements that fulfil the following conditions:

- Settlements of type 'city' ("rodzaj" = 'Ms'),
- Settlements of type 'village' with municipal offices ("rodzaj" = 'Ws' AND "siedzibaUr" = '1')

- Settlements of type 'village' with population > 100 inhabitants, within a district with population density < 50 inhabitants per km<sup>2</sup>, ("rodzaj" = 'Ws' AND "liczbaMies" > 100 AND "gestZalPow" < 50)
- Settlements of type 'village' with population > 700 inhabitants in the districts with population density between 50 and 100 inhabitants per km<sup>2</sup>, ("rodzaj" = 'Ws' AND "liczbaMies" > 700 AND "gestZalPow" > 50 AND "gestZalPow" < 100)
- Settlements of type 'village' with population > 1000 inhabitants, within districts with population density > 100 inhabitants per km<sup>2</sup> ("rodzaj" = 'Ws' AND "liczbaMies" > 1000 AND "gestZalPow" > 100)

The selection is done by dividing objects with the same symbolization into groups (Table 2).

The second step of settlement generalization is assigning appropriate symbols to a particular settlement. The final step is evaluation and solving topological conflicts. As underlined by numerous authors, the evaluation of the results plays an important role in the generalization process (Bard 2004; Stoter et al. 2009; Stoter et al. 2014).

##### Roads generalization

Road network generalization is a challenging process. Numerous characteristics and relationships must be considered. According to the provisions of the Regulation, road selection should be determined based on road's attributes, the role that the road plays in the network and its relation to settlements and other thematic layers. In the Regulations the following features are indicated as the most important: road class (klasaDrogi), management category (katZarzadzania), number of lines (liczbaJezdniDrogi), category of existence (katIstnienia) and

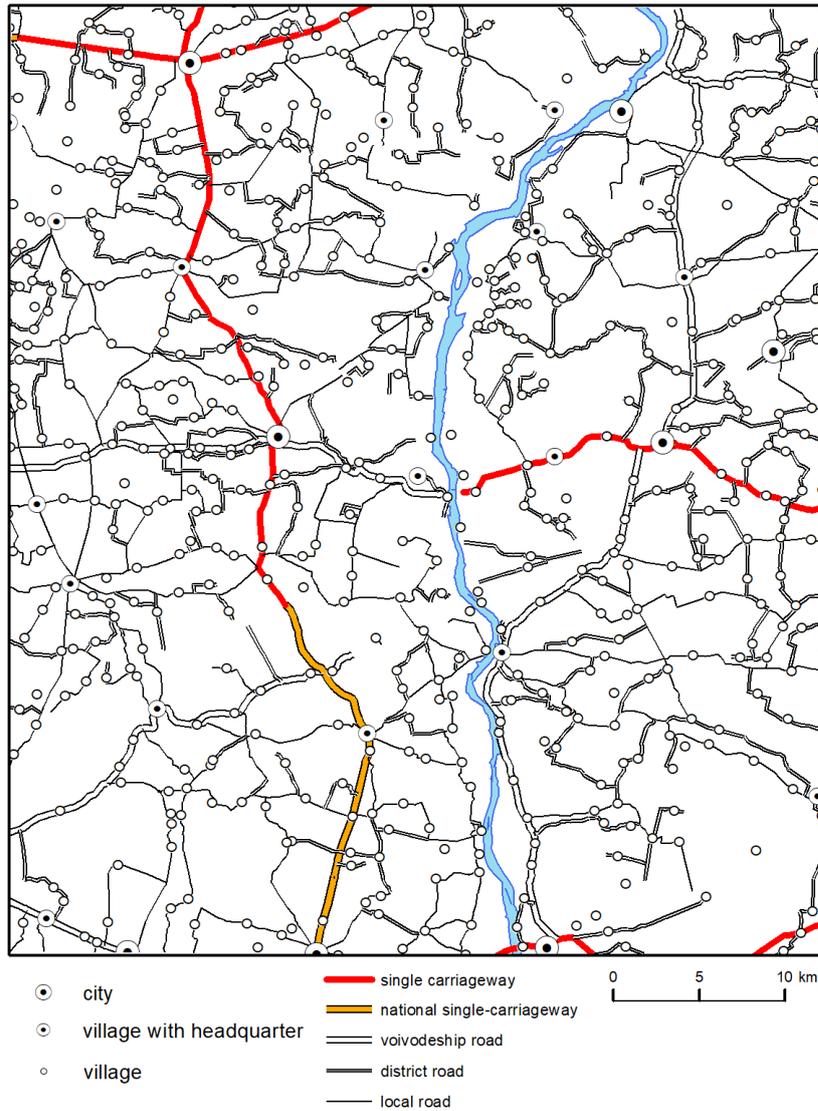


Figure 2. Layers of roads and settlements from GGOD in the research area.  
Source: designed by K. Sielicka

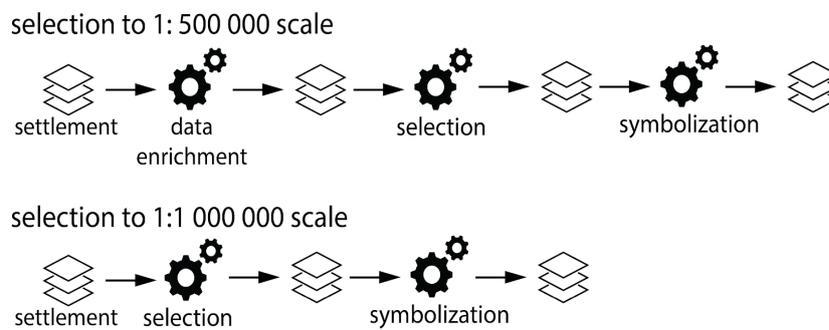


Figure 3. Settlement generalization models from 1:250 000 to 1:500 000 and 1:1 000 000 detail levels.  
Source: designed by K. Sielicka

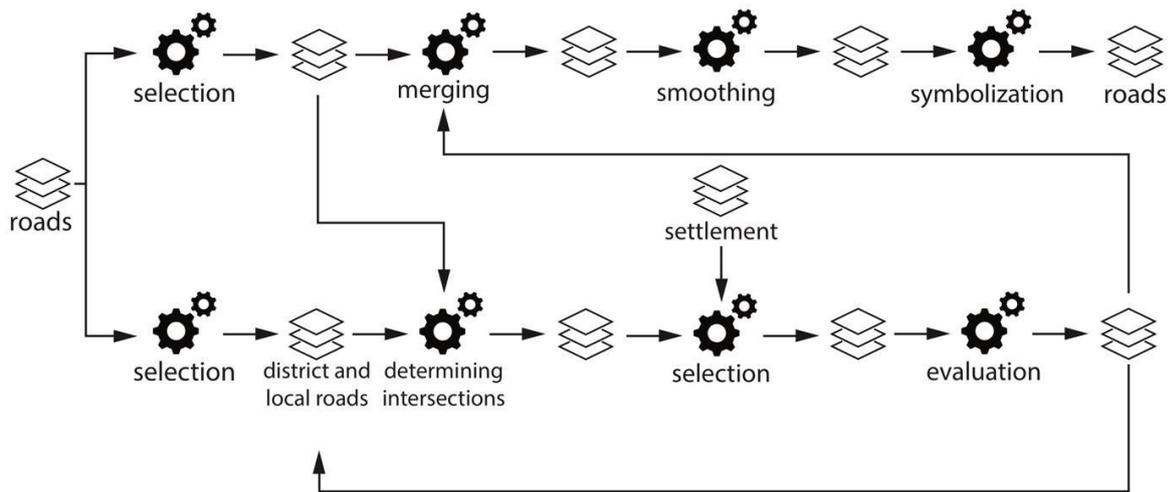


Figure 4. Model of roads generalization to the level of detail corresponding to the 1:500 000 scale  
Source: designed by K.Sielicka

Table 1. Knowledge base (settlement in 1:1 000 000 scale)

| Generalization step  | Implementation  |
|----------------------|---|
| 1<br>Data enrichment | <i>Spatial join</i> - sum of „liczbaMies” from settlement was added to district’s boundaries layer;<br><i>Calculate field</i> - calculate field “density” in district’s layer: sum of „liczba_Mies”/ shape_area ;<br><i>Join attributes by location</i> - add district’s density attribute to settlement by location. |
| 2<br>Selection       | <i>Select by attributes</i> using sum of selection algorithms from Table 1.   |
| 3<br>Symbolization   | <i>Apply symbology</i> - assigning symbols (specified in the Regulation) to objects, divided into groups from Table 1.  |

Source: own elaboration

surface material (materialNawierzchni). The algorithms and parameters of selection are divided into groups of objects with the same symbolization (Table 4).

It is essential to preserve the characteristic features as well as the connectivity of the entire road network. The most complicated problem is selecting sections of the lower road classes – the local ones and the ones managed by municipalities. Not all of them should be maintained, and the selection criteria should prevent leaving segments not connected to the rest of the network or cul-de-sacs. The authors propose carrying out the process simultaneously and separately for the segments of higher classes and lower classes of roads (Figure 4). Local roads (district surfaced roads) should be preserved only if they are a crucial connection, complementing the main network. This is checked by indicating intersections with higher category roads (ArcGIS tool *Production Create Points At Intersections*, Table 4). If a local road (meeting the criteria of district surfaced road

Table 2. Settlement symbolization classes in accordance with the provisions of the Regulation - settlements in 1:500 000 scale

| Settlement class              | Selection algorithm  |
|-------------------------------|--|
| More than 500 000 inhabitants | „rodzaj” = ‚Ms’ AND „liczbaMies” >=500000  |
| 100 000 - 500 000 inhabitants | „rodzaj” = ‚Ms’ AND „liczbaMies” >=100000 AND „liczbaMies” <500000   |
| 25 000 - 100 000 inhabitants  | „rodzaj” = ‚Ms’ AND „liczbaMies” >=25000 AND „liczbaMies” <100000  |
| 5 000 - 25 000 inhabitants    | „liczbaMies” >=5000 AND „liczbaMies” <25000  |
| 1 000 - 5 000 inhabitants     | „liczbaMies” >=1000 AND „liczbaMies” <5000   |
| Less than 1 000 inhabitants   | „rodzaj” = ‚Ws’ AND „liczbaMies” < 1000 AND ((„liczbaMies” > 100 AND „gestosc_za” < 50) OR („liczbaMies” > 400 AND („gestosc_za” < 50 OR („wojewodztw” = ‚łódzkie’ AND „jpt_nazwa_” NOT IN ( ‚powiat wieluński’ , ‚powiat łódzki wschodni’ )) OR „jpt_nazwa_” = ‚powiat miechowski’ )) OR („liczbaMies” > 700 AND „gestosc_za” <= 100 AND „gestosc_za” >= 50)) |

Source: own elaboration

from Table 3) is connected on both sides with roads of a higher category, or leads from such a road to a settlement, it should be shown on the map. Higher class roads should undeniably be left, and should be shown on the target maps both at 1:500 000 and 1:1 000 000 detail levels. In order to correctly select lower class roads, the key features of the whole road network should be maintained, including sections that are connected to the network on both sides, or connecting roads leading to the settlement shown on the map. The maintained road sections should preserve, in a simplified manner, the density and character of

Table 3. Road symbolization classes in accordance with the provisions of the Regulation - roads in 1:500 000 scale.

| Road group                                      | Selection algorithm  |
|---|--|
| Highway   | KlasaDrogi = ,A' AND liczbaJezdniDrogi > 1 AND x_katlstnienia = ,Eks'  |
| Highway under construction                      | KlasaDrogi = ,A' AND liczbaJezdniDrogi > 1 AND x_katlstnienia = ,Bud'  |
| Dual-carriageway expressway or main fast road   | KlasaDrogi = IN (,S', ,GP') AND liczbaJezdniDrogi > 1 AND x_katlstnienia = ,Eks'   |
| Single-carriageway expressway or main fast road | KlasaDrogi = IN (,S', ,GP') AND liczbaJezdniDrogi < 2 AND x_katlstnienia = ,Eks'   |
| Expressway or main fast road under construction | KlasaDrogi = IN (,S', ,GP') AND x_katlstnienia = ,Bud'   |
| Dual-carriageway national road                  | KatZarzadzania = ,K' AND klasaDrogi = ,G' AND liczbaJezdniDrogi > 1 AND x_katlstnienia = ,Eks'                             |
| Single-carriageway national road                | KatZarzadzania = ,K' AND klasaDrogi = ,G' AND liczbaJezdniDrogi < 2 AND x_katlstnienia = ,Eks'                             |
| Voivodeship road                                | katZarzadzania = ,W' AND klasaDrogi = ,G' AND x_katlstnienia = ,Eks'   |
| District surfaced road                          | katZarzadzania IN (,P', ,G') AND klasaDrogi IN (,L', ,I') AND materialNawierzchni IN (,T', ,U') AND x_katlstnienia = ,Eks' |

Source: own elaboration

Table 4. Knowledge base (roads in 1:500 000 scale)

|    | Generalization step       | Implementation  |
|----|---------------------------|---|
| 1a | Selection                 | <i>Select by attributes</i> using selection algorithm from table 2 - "District surfaced roads".   |
| 1b | Selection                 | <i>Select by attributes</i> using sum of selection algorithms from table 2, except "District surfaced roads".   |
| 2  | Determining intersections | <i>Production Create Points At Intersections</i> create point on every intersection of "District surfaced roads" with road of higher class;<br><i>Spatial join</i> counts number of points intersecting the road, sum is stored in "intersections" attribute.   |
| 3  | Selection                 | <i>Select by location</i> select settlements that are more than 1 km away from intersections;<br><i>Select by location</i> select objects from "District surfaced roads" if are closer than add value of attribute "toSettlement" 1 if any settlement object is closer than 2 km;<br><i>Select by attributes</i> select District surfaced roads" with algorithm: "intersections" = 2 OR ( "intersections" = 1 AND "toSettlement" = 1) |
| 4  | Evaluation                | Manual verification of the correctness of selection - whether there are cul-de-sacs or discontinuities in the road network  |
| 5  | Merging                   | <i>Merge</i> "District surfaced roads" with roads of higher classes.  |
| 6  | Smoothing                 | <i>Smooth Line</i> algorithm Polynomial Approximation with Exponential Kernel (PAEK), tolerance 5 km.   |
| 7  | Symbolization             | <i>Apply symbology</i> - assigning symbols (specified in the Regulation) to objects, divided into groups from table 2.  |

Source: own elaboration

Table 5. Knowledge base (settlement in 1:1 000 000 scale)

|   | Generalization step | Implementation  |
|---|---------------------|---|
| 1 | Selection           | Select by attributes using the algorithm: „rodzaj” = ‚Ms’ OR („rodzaj” = ‚Ws’ AND “siedzibaUr” = ‘1’)           |
| 2 | Symbolization       | Apply symbology - assigning symbols (specified in the Regulation) to objects, divided into groups from table 1. |

Source: own elaboration

Table 6. Knowledge base (roads in 1:1 000 000 scale)

|   | Generalization step | Implementation   |
|---|---------------------|--|
| 1 | Selection           | Select by attributes using the algorithm: KlasaDrogi IN (‚K’, ‚S’, ‚GP’) OR KatZaradzania IN (‚K’, ‚W’),   |
| 2 | Smoothing           | Smooth Line algorithm Polynomial Approximation with Exponential Kernel (PAEK), tolerance 5 km.   |
| 3 | Symbolization       | Apply symbology - assigning symbols (specified in the Regulation) to objects, divided into groups: highways, highways under construction, expressways or main fast roads, expressways or main fast roads under construction, national roads, voivodeship roads . |

Source: own elaboration

the road network. Further, the selected road network should be simplified and smoothed. Then the appropriate symbolization should be assigned.

The final step in the generalization of both thematic layers is the evaluation of the result. This step consists of controlling minimum distances, verifying readability and density of the generalized objects. ArcGIS tools such as *Buffer* (0.1 mm on the map, which gives a minimum distance between the signatures of 0.2 mm), *Calculate Density* (which should have a similar distribution as the input data) or *Select by Location* (selecting objects less than 0.2 mm away) are used for this. In case of errors, the procedures to solve them, including moving and deleting objects, should be implemented. The verification tasks should be undertaken iteratively until an optimal result is achieved.

**Detail level corresponding to 1:1 000 000 scale**

**Settlements generalization**

Due to the limited space, small-scale maps contain only the most important objects. In the case of settlements, only the main ones should be kept. The settlements selection is carried out based on their administrative significance. According to the Regulation provisions, the following objects should be selected and presented on the target map:

- Settlements of ‘city’ type,
- Settlements of ‘village’ type with municipal offices.

The steps are analogous to those performed in the case of generalization to the level of detail corresponding to the scale 1:500 000 (Figure 3, Table 5). After validating the minimum distances, the settlement symbols are depicted on the map. The last stage of generalization is controlling the quality of results in the context of legibility and spatial conflicts. If errors occur, deletion or displacement of objects can be considered. The most important thing is to identify key relationships between objects and show them on the target map. At the same time, one should be aware that small-scale maps are mainly meant to signal the presence of certain phenomena and objects, not to show their precise location.

**Roads generalization**

Road network generalization to 1:1 000 000 detail level consists of three main steps (Figure 5, Table 6). The first step is the selection process. The road selection criteria are very restrictive and, to a large extent, the selection is carried out by objects attribute values. In the case of roads, only the sections belonging to high classes and categories are shown on the target map at the considered detail level. According to the provisions of the Regulation, the following roads must be selected:

- Highways (KlasaDrogi = ‘A’),
- Highways under construction (KlasaDrogi = ‘A’ AND x\_katlstnienia = ‘Bud’),
- Expressways or main fast roads (KlasaDrogi IN (‘S’, ‘GP’)),
- Expressways or main fast roads under construction (KlasaDrogi IN (‘S’, ‘GP’) AND x\_katlstnienia = ‘Bud’),
- National roads (KatZaradzania = ‘K’),
- Voivodeship roads (KatZaradzania = ‘W’).

The primary goal of the road selection process is to ensure the continuity and cohesion of the overall network. In the second step, the evaluation and control of road network unity are carried out. As a last task, appropriate symbols are assigned to the selected object.

**Discussion of the generalization results**

The generalization process was carried out in accordance with the principles contained in the Regulation. In order to assess the quality and correctness of the generalization results, they were compared with reliable, manually designed small-scale maps and databases elaborated at the corresponding scales. The verification process consisted of visual inspection of objects presented in the results and also shown in the reference materials, namely atlas maps. Based on the visual results inspection, quantitative analyses have been carried out and summed up in tables 3-8.

**Generalization results evaluation at 1:500 000 detail level**

The results obtained were verified by comparing them with maps created by experienced cartographers; these maps in the study are called reference materials. In the case of the 1:500 000 scale, the following reference maps were considered:

- General map of Lublin Voivodeship (PPWK 1999);
- Atlas of the Republic of Poland (PPWK 1993-1997).

The digitized settlements and road network from the reference maps are shown in Figure 6.

The result obtained was assessed visually by the authors, and then compared with well-designed paper maps. Uneven distribution of the settlements on the resulting map is visible. During the generalization, the uniformity of the arrangement of the objects and the density of the settlements is not taken into account, while it is very important in the case of the manual map design done by the cartographer.

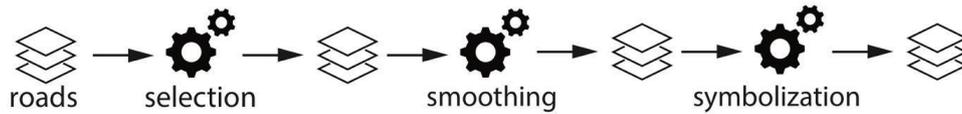


Figure 5. Model of roads generalization to the level of detail corresponding to 1:1 000 000 scale.  
Source: designed by K. Sielicka

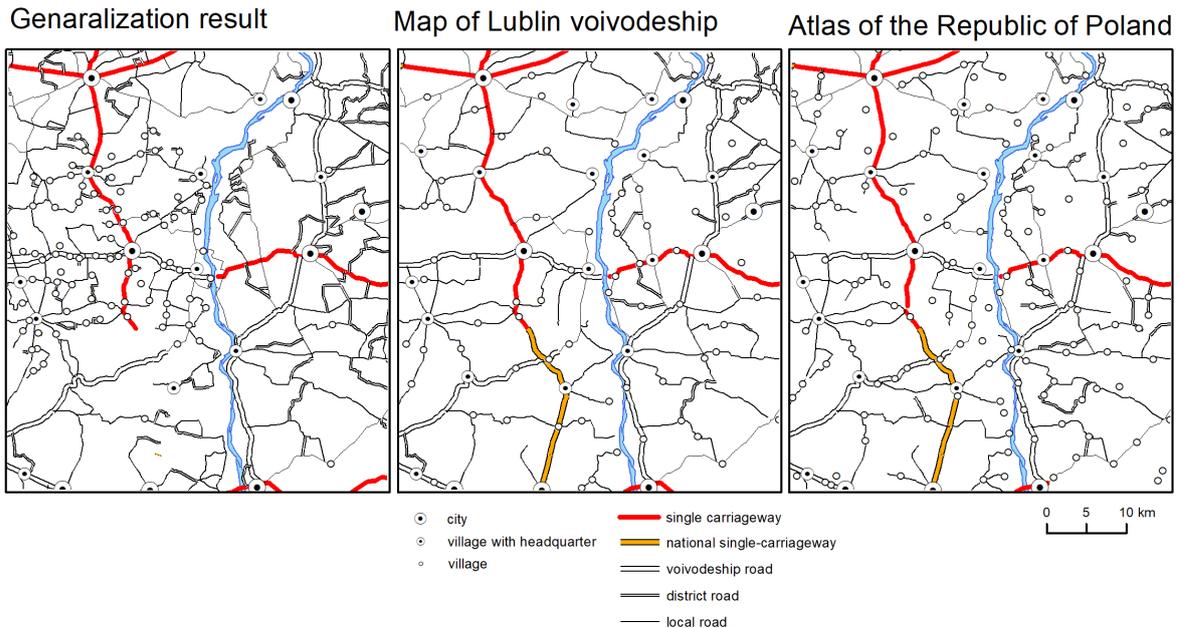


Figure 6. Visualization of settlements and road network from different sources: generalized to 1:500 000.  
Source: designed by K. Sielicka

In the resulting map, the boundaries of districts are clearly visible because of the different settlement densities. For example, in Lipski District, there is a higher concentration of settlements (Figure 7). The reason is the low population density in this district (47 people / km<sup>2</sup>), which is taken into account as a key factor in the settlement selection process according to the map specifications contained in the Regulation. While in neighbouring districts, with higher population density, the selection process is stricter. This situation results in exceedingly significant differences among neighbouring districts.

In order to evaluate the generalization results, the authors have also conducted a quantitative analysis (Figure 8). In Tables 7-9, the result of the comparison of the number of settlements presented in the reference maps and databases is presented. Table 5 shows all the combinations of the presentation of the settlements. The first row shows the settlements shown on all maps, the second row is the settlements omitted in EuroGlobalMap and presented on the remaining ones, etc. The highest value was assigned to the settlements that were shown on both reference maps and omitted in the generalization result. For these settlements, regularities were sought according to which of them, despite the small number of inhabitants, were selected and displayed on the map. Their primary functions, as well as the distance to the nearest neighbour, were also calculated. As a result, no regularities were found by examining the attributes of objects. However, their presence on the target map seems to

be necessary for correctness and aesthetic reasons. It is difficult to indicate the objective factors determining the choice of these settlements. It was likely decided by the subjective assessment of the cartographer and supported by his experience.

In the case of roads, the network in the resulting map is unevenly dense, and there are cul-de-sacs. While comparing it to the manually designed reference maps, it is visible that some connections are missing and there are places where the road density is too high according to this detail level. In the case of a map from the Atlas of the Republic of Poland, it can be seen that the terrain situation has changed and it is out of date, as there are no roads that were built after 1997.

In summary, GGOD generalized to the detail level corresponding to the scale 1:500 000 has 112 settlements, while according to the well known Topfer and Pillewizer (1966) Radical Law, this layer should contain 391 objects<sup>2</sup>. The Radical Law can be used in this case, in the small-scale map the settlements are the only objects represented as point objects.

#### Generalization results evaluation at 1:1 000 000 detail level.

In the case of the 1:1 000 000 scale, the following reference map and database were considered:

- Geographical Atlas (Demart 2006);
- EuroGlobalMap, developed by EuroGeographics<sup>3</sup>.

Digitized settlements and road network from the maps are

<sup>3</sup> <http://www.eurogeographics.org/>, access date: 15.07.2017

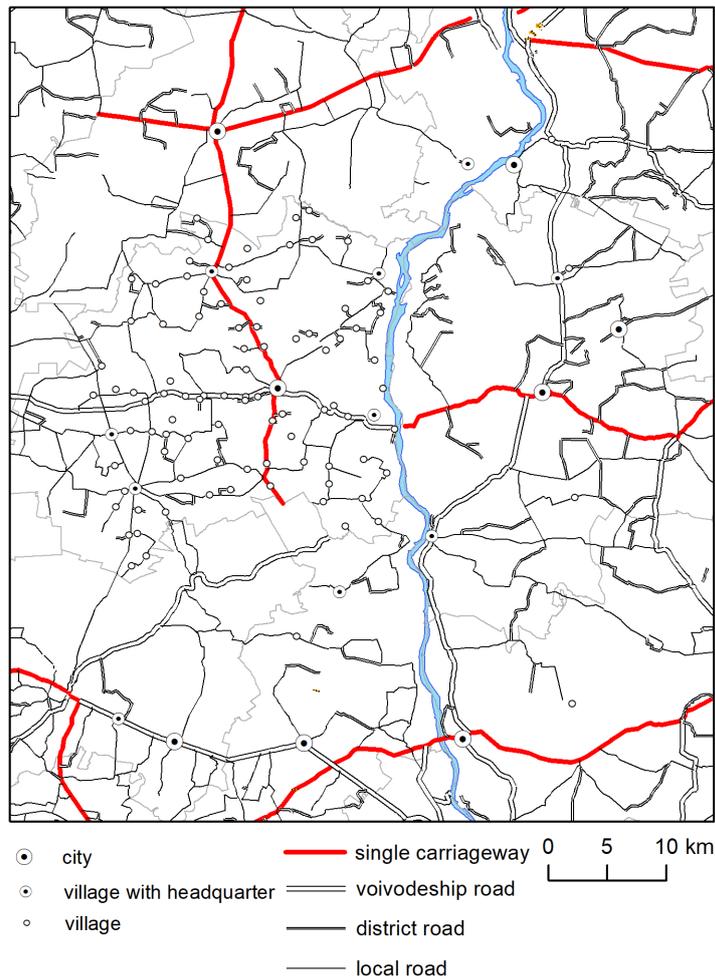


Figure 7. Settlements from GGOD after generalization to 1:500 000 scale. Source: designed by K.Sielicka

shown in Figures 9 and 10.

The small number of settlements and the strict selection criteria made the result of the generalization seem more correct than in the 1:500 000 scale. Particularly interesting is the comparison with EuroGlobalMap. The database corresponding to the small-scale detail level was derived from the databases, whose detail level corresponds to the 1:250 000 scale. This process was conducted in a fully automated mode. There are very few settlements presented in this database. According to the map from the geographical atlas, not enough settlements were presented. Therefore, this database should also be improved regarding settlements selection.

The number of objects as a result of the selection is comparable to the number of settlements on the Demart map developed by the cartographer (Table 10 and Figure 10). This indicates a relatively correct density of settlements. In Table 11 it can be seen that only a few settlements have been presented on one of the maps. Only one settlement was presented on the EuroGlobalMap and omitted in the atlas map and generalization result (marked with an asterisk in Table 12). What is more, six settlements were shown only in the Demart atlas (marked with an asterisk in table 12). An interesting case is Bochoznica, presented on both the atlas map and EuroGlobalMap, while it

was omitted in generalization result. This town is essential from the communication point of view because it is located at the intersection of provincial roads.

In conclusion, in GGOD generalized to the detail level corresponding to 1:1 000 000 scale, 18 settlements were presented, while according to Topfer and Pillewizer (1966) Radical Law, this layer should contain 276 objects<sup>4</sup>. Such a significant difference may partly result from the fact that the source database is very rich, much richer than the level of detail on which it was developed.

**Conclusions and future work**

The result of automatic generalization differs to some extent from the content developed by cartographers and shown on the manually designed maps. However, these differences are not very significant. At the 1:500 000 scale, settlements selected on the basis of map specifications contained in the Regulation are not uniformly distributed. This is due to the conditions of district density set during the selection process. It suggests that the generalization principles contained in the Regulation should be modified and supplemented.

In the case of roads, the biggest problem is the selection of local roads, which should complement the network and maintain

Table 7. A number of settlements depicted on the maps

|                      | Generalization result | Atlas of the RP | PPWK map | Radical Law |
|----------------------|-----------------------|-----------------|----------|-------------|
| Number of settlement | 112                   | 161             | 94       | 391         |

Source: own elaboration

Table 8. A number of settlements - divided into the number of shows

| Shown on              | One map | Two maps | All three maps |
|-----------------------|---------|----------|----------------|
| Number of settlements | 147     | 62       | 32             |

Source: own elaboration

Table 9. The number of settlements that were presented in the generalization results and on the reference maps.

| Generalization result | Atlas of the RP | PPWK map | Number of settlement | % of total settlement |
|-----------------------|-----------------|----------|----------------------|-----------------------|
| shown                 | shown           | shown    | 32                   | 5,8                   |
| shown                 | shown           | omitted  | 19                   | 3,4                   |
| omitted               | shown           | shown    | <b>38</b>            | <b>6,9</b>            |
| shown                 | omitted         | shown    | 5                    | 0,9                   |
| shown                 | omitted         | omitted  | 56                   | 10,1                  |
| omitted               | shown           | omitted  | 72                   | 13,1                  |
| omitted               | omitted         | shown    | 19                   | 3,4                   |
| omitted               | omitted         | omitted  | 312                  | 56,4                  |
|                       |                 | Sum      | 553                  |                       |

Source: own elaboration

Table 10. Number of settlements depicted on the maps

|                       | Generalization result | Demart atlas | EuroGlobalMap | Radical Law |
|-----------------------|-----------------------|--------------|---------------|-------------|
| Number of settlements | 18                    | 23           | 13            | 276         |

Source: own elaboration

Table 11. Number of settlements - divided into the number of shows

| Shown on              | One map | Two maps | All three maps |
|-----------------------|---------|----------|----------------|
| Number of settlements | 7       | 10       | 9              |

Source: own elaboration

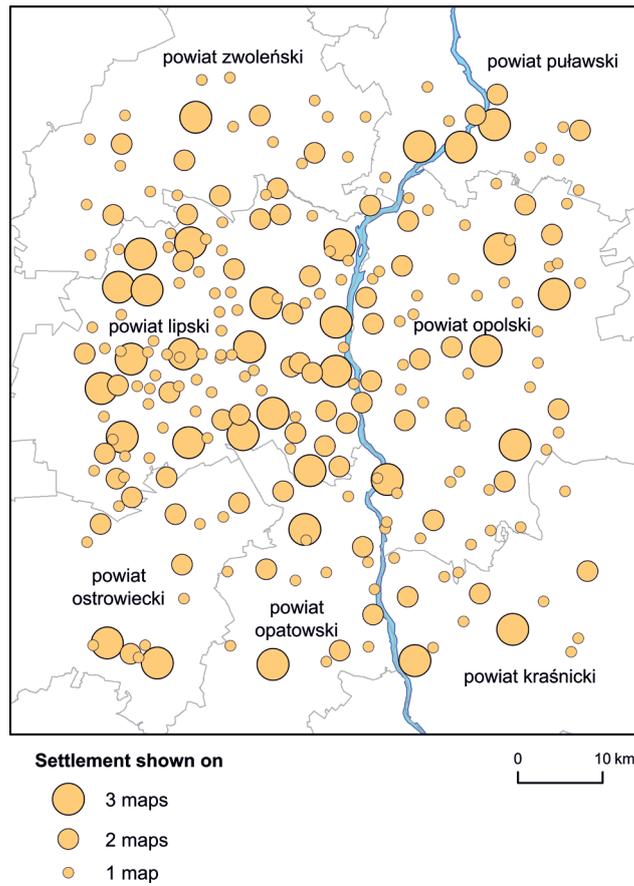


Figure 8. Visualization of settlements on 1:500 000 scale from different sources.  
Source: designed by K.Sielicka

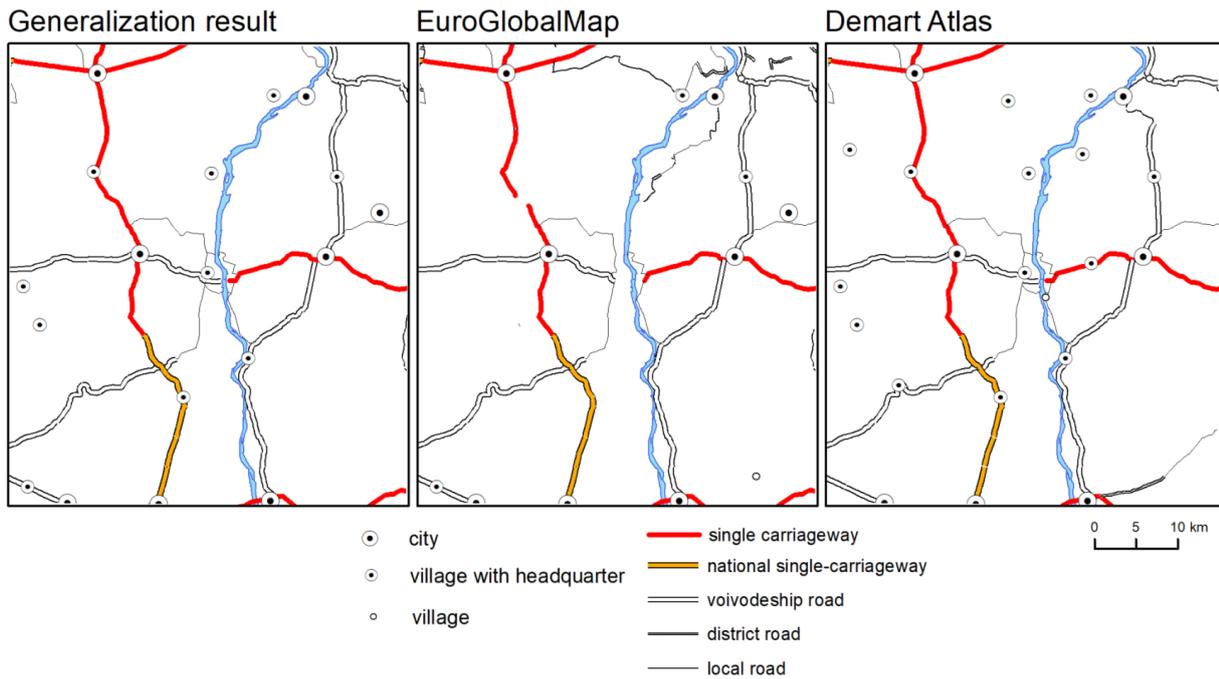


Figure 9. Visualization of settlements and road network from different sources.  
Source: designed by K.Sielicka

Table 12. The number of settlements presented as a result of selection and on reference maps

| Generalization result | Demart atlas | EuroGlobalMap | Number of settlements | % of total settlements |
|-----------------------|--------------|---------------|-----------------------|------------------------|
| shown                 | shown        | shown         | 9                     | 1,63                   |
| shown                 | shown        | omitted       | 7                     | 1,27                   |
| omitted               | shown        | shown         | 1                     | <b>0,18</b>            |
| shown                 | omitted      | shown         | 2                     | 0,36                   |
| shown                 | omitted      | omitted       | 0                     | 0                      |
| omitted               | shown        | omitted       | 6*                    | 1,08                   |
| omitted               | omitted      | shown         | 1*                    | 0,18                   |
| omitted               | omitted      | omitted       | 527                   | 95,3                   |
|                       |              | Sum           | 553                   |                        |

Source: own elaboration

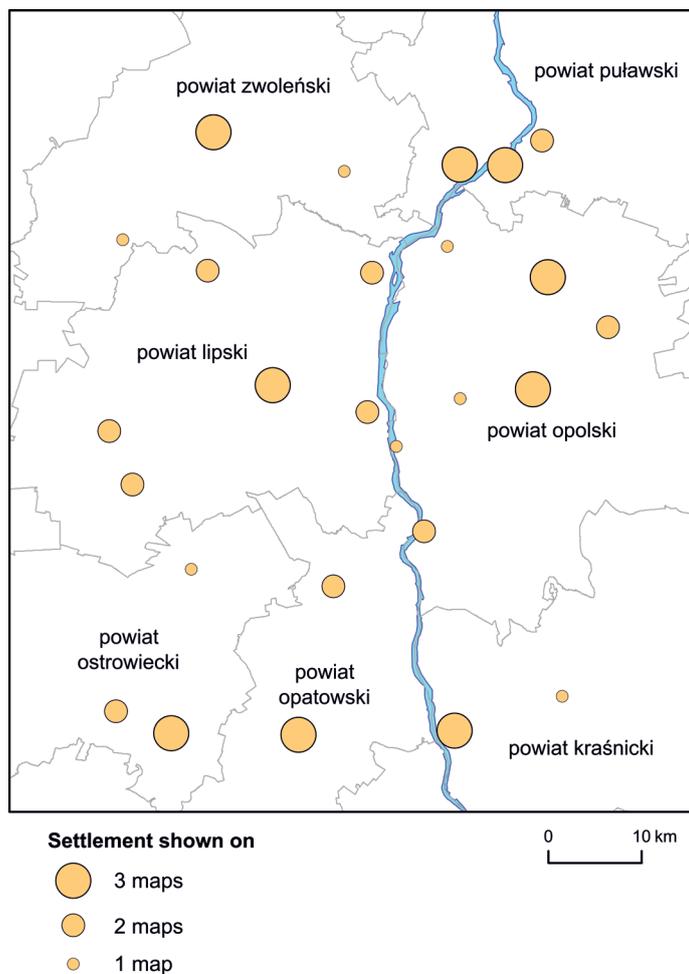


Figure 10. Visualization of settlements on 1:1 000 000 scale from different sources

Source: designed by K. Sielicka

Table 13. Length of the roads (in kilometres) by management category and number of occurrences in the reference materials (maps and databases)

| Management category | 1   | 2  | 3 occurrences |
|---------------------|-----|----|---------------|
| Municipality        | 18  | 57 | 73            |
| District            | 87  | 20 | 26            |
| Voivodeship         | 0.5 | 9  | 219           |
| National            | 0   | 5  | 90            |

Source: own elaboration

a connection with the settlements shown on the map.

Automatic road selection at 1:1 000 000 scale, due to simple and formalised selection criteria, can be considered as relatively correct. There are few differences between the generalization result and the reference maps designed manually by a cartographer.

The values collected in Table 13 are the confirmation that regional and national roads in all analysed sources were chosen fairly. Among municipal and district roads, many more roads were selected and presented at EuroGlobalMap.

The verification of map specifications leads to the conclusion that in the future additional selection criteria should be added. After comparing the results of automatic generalization with the reference materials, it is possible to conclude that considering more geometrical and semantic settlements and road characteristics may contribute to improving the quality of the objects selection.

In the case of settlements, an important additional criterion for selection can be the relation to neighbouring objects; for instance, the proximity of the river, road network nodes or a railway line (in particular the stations). Features such as the function of the settlement and its distance to the nearest neighbour, as well as the settlements density, may play an equally important role.

In the case of road network generalization, the essential aspect that should be considered is road network connectivity, so it is necessary to take into account the neighbourhood of the settlements, bridges, harbours and railway stations. It is

also important whether other roads are performing the same connectivity function, as well as if the network density is correctly represented on the generalized map.

The research conducted has proven that it is impossible to determine generalization parameters that would allow for obtaining correct results throughout Poland, therefore smaller units should be considered for this purpose. However, they should not necessarily be based on administrative boundaries. What is more, the algorithms for automatic generalization of the road network and settlements should take into account the relationships between objects from different layers. At the same time, manually designed maps are and should be material for checking the quality and correctness of automatic generalization results.

While conducting the research, the incompleteness and need to enrich the GGOD structure was found. Based on the analysis of the content of the EuroGlobalMap settlement layer, it can also be concluded that the generalization algorithms used to obtain this layer should also be improved. The lessons learned from this and future research could be used to enhance the quality of both GGOD and EuroGlobalMap databases. Last but not the least, it can be stated that the map specifications contained in the Regulation need to be complemented and improved to be able to generalize maps automatically.

## References

- Ai, T & Liu, Y 2004, 'Analysis and simplification of point cluster based on Delaunay triangulation model' in *Advances in spatial analysis and decision making*, eds Z Li, Q Zhou & W Kainz, Taylor & Francis, London, pp. 9–19.
- Balboa, GJL & López, AFJ 2008, 'Generalization-oriented road line classification by means of an artificial neural network', *Geoinformatica*, 2008, vol. 12, no. 3, pp. 289-312. DOI: 10.1007/s10707-007-0026-z.
- Bard, S 2004, 'Quality assessment of cartographic generalization', *Transactions in GIS*, vol. 8(1), pp. 63-81. Available from: <https://doi.org/10.1111/j.1467-9671.2004.00168.x>. [10 May 2018].
- Benz, S & Weibel, R 2014, 'Road network selection for medium scales using an extended stroke-mesh combination algorithm', *Cartography and Geographic Information Science*, vol. 41(4), pp. 323-339.
- Bereuter, P & Weibel, R 2013, 'Real-time generalization of point data in mobile and web mapping using quadrees', *Cartography and Geographic Information Science*, vol. 40, no. 4, pp. 271-281. DOI: 10.1080/15230406.2013.779779.
- Bereuter, P & Weibel, R 2017, 'Variable-scale maps in real-time generalization using a quadtree data structure and space deforming algorithms', *International Journal of Cartography*, vol. 3(1), pp. 1-14. DOI: 10.1080/23729333.2017.1304189.
- Demart 2006, *Geografia, Świat, Polska - atlas geograficzny, liceum*, [Geography, World, Poland - geographic atlas, highschool] Demart, Warszawa.
- EuroGeographics, *EuroGlobalMap*. Available from: <http://www.eurogeographics.org/>. [15 July 2018].
- Geodetic and cartographic law*, The Act of May 17, 1989, Journal of Laws of 1989 No. 30 item 163.
- Jiang, B & Claramunt, C 2004, 'A structural approach to the model generalization of an urban street network', *Geoinformatica*, vol. 8, no. 2, pp. 157-171.
- Karsznia, I & Weibel, R 2018, 'Improving settlement selection for small-scale maps using data enrichment and machine learning', *Cartography and Geographic Information Science*, vol. 45(2), pp. 111-127. DOI: 10.1080/15230406.2016.1274237.
- Li, Z2007, 'Digitalmapgeneralizationattheageoftheenlightenment: A review of the first forty years', *The Cartographic Journal*, vol. 44(1). DOI: 10.1179/000870407X173913.
- Liu, X, Zhan, FB & Ai, T 2010, 'Road selection based on Voronoi diagrams and "strokes" in map generalization', *International*

- Journal of Applied Earth Observation and Geoinformation*, vol. 12, Supplement 2, pp. 194–202.
- Mackaness, WA & Beard, MK 1993, 'Use of graph theory to support map generalization', *Cartography and Geographic Information Systems*, vol. 20, no. 4, pp. 210-211.
- Mackaness, WA & Mackechnie, G 1999, 'Automating the detection and simplification of junctions in road networks', *GeoInformatica*, vol. 3, no. 2, pp. 185-200.
- Polskie Przedsiębiorstwo Wydawnictw Kartograficznych im. Eugeniusza Romera S.A. 1999, *Mapa województwa lubelskiego [Map of the Lublin voivodeship]*, Warszawa-Wrocław.
- Polskie Przedsiębiorstwo Wydawnictw Kartograficznych im. Eugeniusza Romera S.A. 1993-1997, *Atlas Rzeczypospolitej Polskiej [Atlas of Republic of Poland]*, Główny Geodeta Kraju, Warszawa.
- MacQueen, J 1967, 'Some methods for classification and analysis of multivariate observations', *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics*, University of California Press, Berkeley, California, pp. 281-297. Available from: <<https://projecteuclid.org/euclid.bsm/1200512992>>. [15 July 2018]
- MSWiA 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* [Regulation of the Ministry of Interior from 17 November 2011 On the Topographic Objects Database and General Geographic Objects Database, as well as standard cartographic products] (Dz.U., nr 279, poz. 1642). Available from: <<http://bip.mswia.gov.pl/bip/rozporzadzenia-ministr/2011/20378,DzU-2011-nr-279-poz-1642.html>>. [10 August 2018].
- Stoter, J, Burghardt, D, Duchêne, C, Baella, B, Bakker, N, Blok, C, Pla, M, Regnauld, N, Touya, G & Schmid, S 2009, 'Methodology for evaluating automated map generalization in commercial software', *Computers, Environment and Urban Systems*, vol. 33(5), pp. 311-324.
- Stoter, J, Zhang, X, Stigmar, H & Harrie, L 2014, 'Evaluation in generalization' in *Abstracting Geographic Information in a Data Rich World. Lecture Notes in Geoinformation and Cartography*, eds D Burghardt, C Duchêne & W Mackaness, Springer, Cham.
- Töpfer, F & Pillewizer, W 1966, 'The principles of selection', *The Cartographic Journal*, vol. 3, no. 1, pp. 10-16.
- Weiss, R & Weibel, R 2014, 'Road network selection for small-scale maps using an improved centrality-based algorithm', *Journal of Spatial Information Science*, vol. 9, pp. 71-99.
- Yan, H & Weibel, R 2008, 'An algorithm for Point Cluster Generalization based on the Voronoi Diagram', *Computers & Geosciences*, vol. 34, no. 8, pp. 939-954.