Article

Base Point Split Algorithm for Generating Polygon Skeleton Lines

Lewandowicz Elżbieta [1] and Flisek Paweł [2]

[1] Department of Geoinformation and Cartography/Institute of Geodesy and Civil Engineering/Faculty of Geoengineering/ University Warmia and Mazury in Olsztyn, Poland; leela@uwm.edu.pl

[2] pablo941b@gmail.com *

Correspondence: leela@uwm.edu.pl; Tel.: +48-513-242-658 (E.L.)

Abstract: The article presents the Base Point Split (BPSplit) algorithm for generating a polygon skeleton based on sets of vector data describing lakes and rivers. A key feature of the BPSplit algorithm is that it is dependent on base points representing the source or mouth of a river or a stream. The input values of base points determine the shape of the resulting skeleton. Various skeletons can be generated with the use of different base points. Base points are applied to divide polygon boundaries into segments. Segmentation supports the selection of TIN edges inside polygons. The midpoints of selected TIN edges constitute a basis for generating a skeleton. The algorithm handles polygons with numerous holes, and it accounts for all holes. This article proposes a method for modifying a complex skeleton with numerous holes. In the discussed approach, skeleton edges that do not meet the preset criteria (e.g. that the skeleton is to be located between holes in the center of the polygon), are automatically removed. A simple algorithm for smoothing zigzag lines was proposed.

Keywords: straight skeleton; hydrographic network; network modeling

---

1. Introduction

Automated solutions for processing vector data relating to surface water bodies and rivers have many practical applications, and they can be used to generate a skeleton of a hydrographic network. This is accomplished with the use of generation algorithms. Various generation methods have been described in the literature, including:

- Medial axis transformation [1, 2, 3],
- Chordal axis transform [4],
- Straight skeleton [5,6],
- Delaunay triangulation [6, -9],
  - Splitarea algorithm [10-11],
  - Other [12-14].

In the described methods, the process of plotting skeleton lines begins with the generation of a base skeleton. In successive steps, selected skeleton edges are modified and removed to improve the applicability of the results. Skeletons are modified with the use of various methods, including correction [7, 15], extraction [8], hierarchical feature extraction [16], filtration [17, 18] and generalization [19]. Different types of triangles are identified in polygons that are generated with the use of methods based on triangulated irregular networks (TIN) [10, 7, 20]. Different approaches to generating skeleton edges for various triangle types have been proposed in the literature [7, 10, 20] (Figure 1). In a type 1 triangle, one side overlaps a section of the polygon boundary line, and the two remaining sides intersect the polygon. In a type 2 triangle, two sides are segments of the polygon boundary line. The apices of a type 3 triangle are located on the polygon boundary line, and none of the triangle's sides overlap the polygon boundary. The above also applies to a type 4 triangle.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 September 2020 doi:10.20944/preprints202009.0721.v1

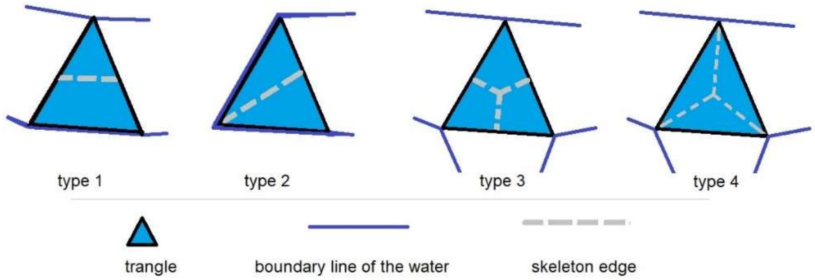


Figure 1. Triangle types: type 1 - link triangle, type 2 - ear triangle, types 3 and 4 - branch triangle

In the literature [8, 10, 14], skeletons were generated based on an analysis of a polygon's surroundings, and neighbors were identified on the right and left side of polygon boundaries. Neighbors are not always easy to identify, and in some cases, the analyzed area has to be expanded [10]. This process is particularly difficult when polygons contain holes (islands on water bodies).

In a previous study [14], the authors applied a simplified method to generate the centerline of an elongated polygon on the example of a river island. Topological data describing the neighborhood of the polygon representing a river channel were used to select TIN edges intersecting the river channel in a direction perpendicular to river flow. The presence of an island on the river complicated the procedure, and the neighborhood relations between the island boundary and the left and right boundary of the river channel had to be considered. The authors concluded that the proposed algorithm should be tested and modified in analyses of other objects with varied shapes, in particular in networks composed of a large number of connected water bodies [14] (p. 17). Such attempts have been made in the current study to eliminate the shortcomings of the previously described algorithm, where the neighborhood relations between polygon boundaries and the neighboring polygons had to be taken into account. This task proved to be particularly challenging in polygons containing numerous holes.

The aim of this study was to develop an algorithm for generating skeletons of polygons that represent elements of a complex hydrographic network, such as a system of lakes connected by rivers or lakes with numerous islands. The developed algorithm was applied in several research tasks. **The first task** relied on the assumption that the designed algorithm can be used to generate an appropriately shaped skeleton already at the beginning of the process, and that it can be applied in various cases. The shape of the resulting skeleton was determined by the location of base points (BP) on the polygon boundary (boundary of a water body). These points should be important for the practical application of the generated skeleton. The identified points will be the hanging nodes in the network. The skeleton will be used to create a geometric model of a hydrographic network for the purpose of hydrographic modeling or navigation. **The second research task** focused on polygons with numerous holes (islands on water bodies), where the skeleton's location between islands was modified automatically based on the preset criteria. **The third task** involved the search for a solution where the skeleton could be generalized by eliminating zigzag effects.

The present study was conducted with the use of programming tools in ArcGIS (ESRI) software [21] and the algorithms developed by the authors in Python.

2. Methods and Materials

2.1. The first research task – characterization of the BPSplit algorithm

The first research task relied on the assumption that the expected shape of a skeleton could be defined already at the stage of data preparation. The analyzed polygon represents a water body, and it is the only element of a hydrographic network composed of lakes connected by a river. The hydrographic network was developed based on vector data. The Base Point Split (BPSplit)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 September 2020 doi:10.20944/preprints202009.0721.v1

algorithm was created. Base points that significantly influence the generated skeleton were defined in the first step. The edges of the Delaunay triangulation (triangulated irregular network, TIN) model inside the polygon were the key elements of the algorithm.

In classical solutions involving TIN models, the skeleton is generated based on all TIN edges inside the polygon with the use of solutions that are applied to various types of TIN triangles (Figure 1). The skeleton is then modified (Meijers et al. 2016). In the authors' previous study [14] and in the current study, selected TIN edges were used to generate a skeleton. In the previous study, TIN edges were selected based on the topological relations between polygon boundaries and the neighboring polygons. In this study, TIN edges were selected based on the adopted set of base points. Base points were localized on the boundary of a polygon, for example a polygon representing the source or mouth of a river intersecting a lake. Base points constitute hanging nodes in the resulting skeleton.

In the developed algorithm, base points divide the polygon's boundary line into segments. Each boundary segment begins and ends in a base point. The segmentation technique is used to select TIN edges. TIN edges that touch segments of the boundary line are selected from the set of TIN edges located inside the polygon. TIN edges that touch base points are removed from this subset. The midpoints of selected TIN edges and base points constitute a basis for generating the skeleton of a hydrographic network. Additional data are required to combine these points into a skeleton. In classical TIN methods, a skeleton is generated with the use of the structures defined for different types of triangles (Fig. 1). Triangles are not analyzed in the proposed solution. Selected TIN edges are used to divide the polygon (for example, a polygon representing a lake) into a set of smaller polygons. Skeleton edges are created based on the midpoints of selected TIN edges and base points. The edges are generated between the points located on the boundary of a single polygon segment.

2.1.1. A comparison of BPSplit and Splitarea algorithms

In 2016, Meijers et al. [10] proposed the Splitarea algorithm for generating skeletons of polygons that represent water bodies. In this study, the Splitarea algorithm was adopted as a classical method, and it was used to modify the previously proposed solution [14]. The Base Point Split (BPSplit) algorithm was described by comparing it with the Splitarea algorithm [10]. Successive stages of the skeleton generation process involving the compared algorithms are presented in Table 2 and Figure 2.

Successive data processing steps are presented in Figure 2. In the solution proposed by Meijets et al. [10], a skeleton is generated based on all TIN edges, and it is modified in successive steps to deliver the required functionality. In the approach involving the BPSplit algorithm, a set of base points is generated. In the presented example, base points are represented by three points on the polygon boundary line. Base points were used to divide the polygon boundary into segments, and the segments were used to select TIN edges that touch different segments of the boundary lines, but not base points. A skeleton was generated based on the midpoints of selected TIN edges and base points.

Table 2. Process of generating a skeleton with the Splitarea algorithm and the Base Point Split algorithm

	Classical algorithm: Splitarea [10]	Proposed algorithm: BPSplit
(1)	Triangulation	Segmentation of polygon boundaries with the use of base points
(2)	Selection of internal triangles	Triangulation
(3)	**Skeleton generation**	Selection of internal triangles

(4)	Generation of connectors	Selection of TIN edges based on segments of the polygon boundary
(5)	Edge labeling and skeleton pruning	**Generation of the final skeleton**
(6)	Generation of the final skeleton	Skeleton smoothing

Successive steps of the Splitarea algorithm proposed by Meijers et al. (2016)		Successive steps of the BPSplit algorithm proposed in this study

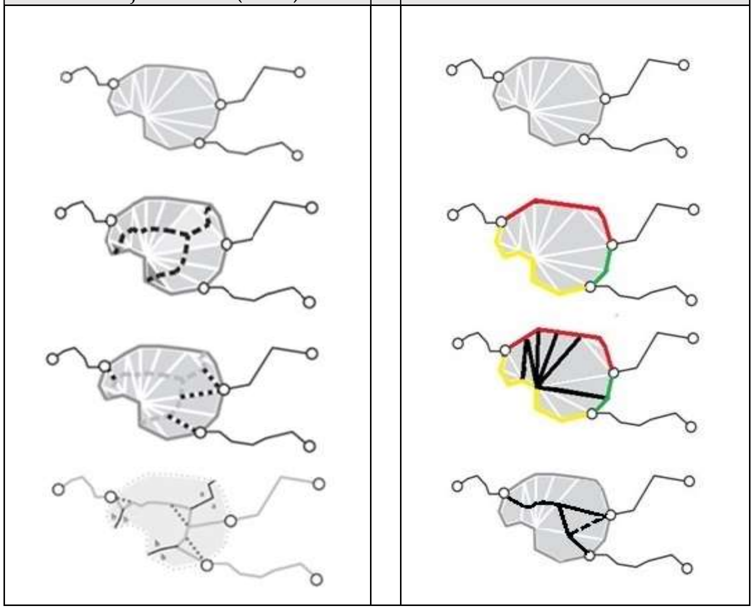


Figure 2. Successive steps of the algorithm generation process with the use of Splitarea [10] and BPSplit algorithms.

2.1.3. Skeleton adjustment with the BPSplit algorithm

The skeleton generated with the BPSplit algorithm can contain loops (Figures 2 and 3). The resulting skeleton can be used to model navigation networks. The modeled hydrographic network should not contain loops. There are two approaches to removing loops. In the first method, a loop is replaced with a loop centroid (Figure 3b and 3d). This solution has been used in the literature for type 3 triangles (Figure 1). The second approach involves a tool for generating the shortest path between the source and the mouth of a river (the spanning tree tool can also be used). The results produced by the shortest path tool are combined to generate a skeleton without loops (Figure 3 c).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 September 2020 doi:10.20944/preprints202009.0721.v1

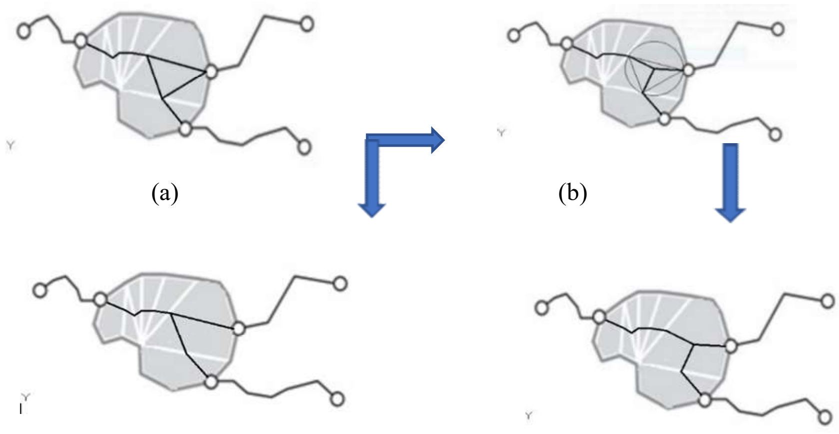


Figure 3. Elimination of loops from a skeleton: (a) skeleton with a loop; (b) the loop is replaced with the loop centroid (type 3 triangle); (c) the skeleton is modified with the use of network analysis tools; (d) the resulting skeleton where the loop was replaced with the loop centroid.

The skeletons generated with the use of four different methods are presented in Figure 4. The skeletons generated with the use of the solutions proposed by [10] are presented in Figures 4a and 4b. The skeletons generated with the use of the solutions proposed in this study are presented in Figures 4c and 4d.

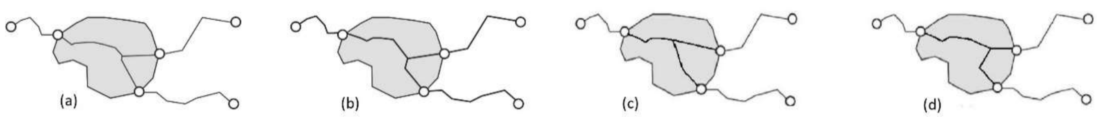
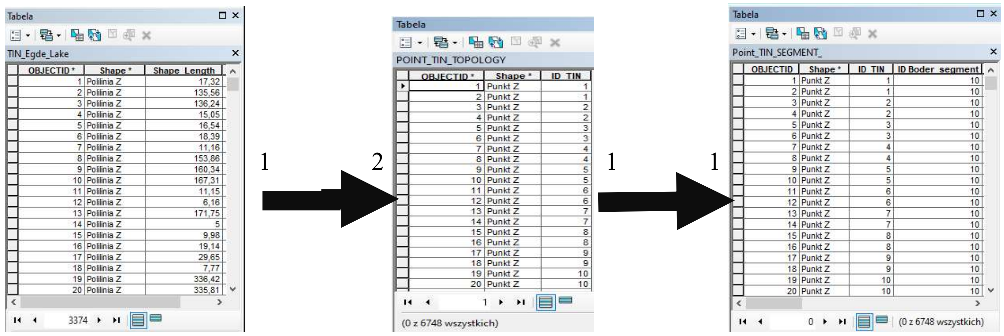


Figure 4. Skeletons of the same polygon generated with the: (a) straight skeleton method (Gold 2017); (b) Splitarea algorithm [10]; (c) and (d) BPSplit algorithm and skeleton adjustment.

2.1.2. Method of selecting TIN edges in the BPSplit algorithm

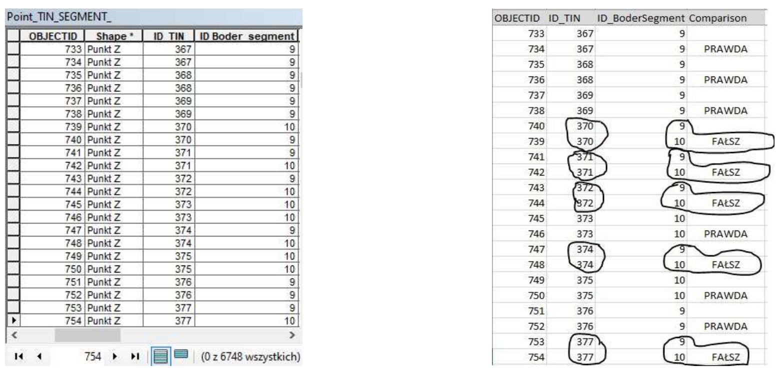
The process of selecting TIN edges plays an important role in the BPSplit algorithm. TIN edges are selected based on segments of polygon boundaries. TIN edges that touch different boundary segments are selected. The relations between TIN edges and polygon boundary segments have to be determined for this purpose. This is accomplished by analyzing topological relations during data processing. A set of points that constitute the apices of TIN edges is generated. These points are assigned the identifiers of TIN edges (ID_TIN) in attribute tables (Figure 5 b). The resulting data set contains twice as many apex points as TIN edges. As a result, the attribute table contains two points with identical ID_TIN (Figure 5 b). In the next step, the algorithm connects boundary segments with TIN apices that touch them. The ID of the boundary segment (IDBoder_segment) that touches TIN apices is included in the attribute table (Figure 5c). These data are used to select the ID of TIN edges touching different boundary segments. The algorithm compares the values of ID_TIN and IDBoder_segment (Figure 6) and selects IDBoder_segment values that differ for the same values of ID_TIN. In the discussed example, 1029 TIN edges were selected from the overall set of 3374 TIN edges. In the next step, TIN edges that touch base points are removed from the set of selected TIN edges that touch different segments of the polygon boundary.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 September 2020 doi:10.20944/preprints202009.0721.v1



(a) (b) (c)

Figure 5. Topological relations during data processing: (a) attribute table of TIN edges; (b) attribute table of TIN edge apices with ID_TIN values; (c) attribute table of TIN edge apices that touch boundary segments (containing the values of ID_Boder_segment).



(a) (b)

Figure 6. Selection of TIN edges (ID_TIN) based on: (a) attributes of TIN apices based on the topological relations during data processing; (b) ID of TIN edges selected based on FALSE values (370, 371, 372, 374, …).

2.1.3 Description of the BPSplit algorithm

The BPSplit algorithm generates a skeleton of a polygon with the use of base points and the midpoints of selected TIN edges. The algorithm relies on various types of polygon segments (Figure 7a) and different connections with base points (Figure 7b) in the process of generating a skeleton.

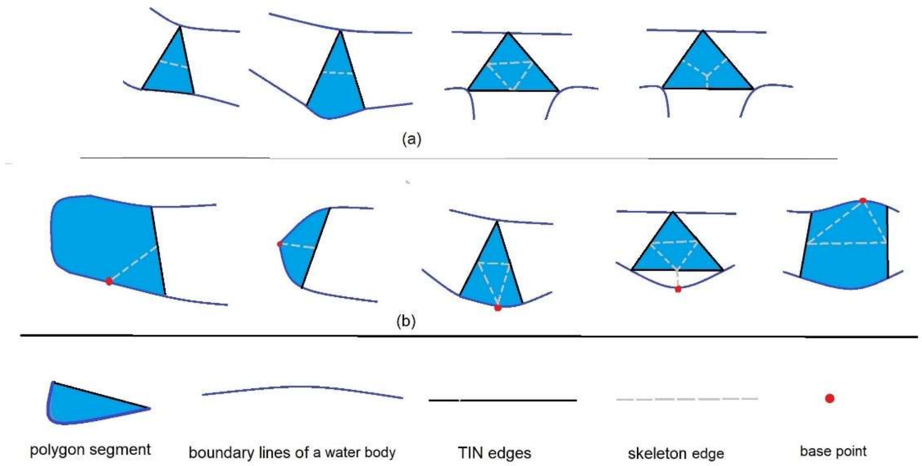


Figure 7. Editing of skeleton edges based on various types of polygon segments: (a) different types of polygon segments with skeleton edges; (b) different locations of base points in segments with a visualization of skeleton edges.

The proposed BPSplit algorithm relies on classical skeleton generation methods that are based on TIN edges. The relations between polygon boundaries and the neighboring polygons on the right and left side are not taken into account. A skeleton is generated based on the rules applicable to the identified types of polygon segments, as well as additional rules, with different locations of base points (Figure 7).

2.1. Methodology associated with the second research task – polygons with holes

The second research task relied on the assumption that the location of a skeleton representing a polygon with numerous holes (islands on water bodies) can be automatically modified between islands according to the preset criteria. In selected software solutions, including ESRI, holes are eliminated during skeleton generation. This is the simplest solution. In the approach described by [10], holes were taken into account, and their location inside the polygon was considered (Figure 8). A special case where the holes inside a polygon come into contact is presented in Figure 8b.

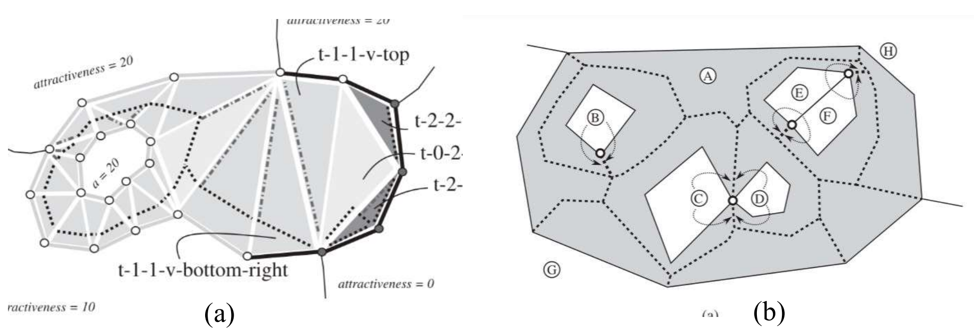


Figure 8. Skeletons of polygons with holes: (a) a polygon with one hole; (b) a polygon with numerous holes [10].

Skeletons were generated with the use of the BPSplit algorithm based on the polygons presented by [10]. In the first step, base points were identified as common points located on both the polygon boundary and the lines outside the polygon contour. Three base points were identified for the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 September 2020 doi:10.20944/preprints202009.0721.v1

polygon in Figure 8a. Three base points were also identified for the polygon in Figure 8b, but two points are located on the polygon's external boundary, whereas the third point is located on the boundary of two contacting holes (the third point was adopted to generate skeleton edges between islands). The segments of polygon boundaries are marked with different colors in Figure 9. The boundaries of holes inside the polygon constitute separate segments. Polygon skeletons were generated with the BPSplit algorithm. TIN edges were generated based on polygon apices as well as the midpoints of polygon edges. These points had to be accumulated on boundary lines because in preliminary tests where points were not accumulated, individual edges of the skeletons generated with the BPSplit method intersected islands. This result points to an insufficient number of points on the polygon boundary for generating TIN edges with the BPSplit algorithm. In ESRI software [21], points are accumulated by default when the centerline of a polygon is generated with the use of the *create centerline from polygon features* tool.

The skeletons generated by the BPSplit algorithm are presented in Figures 9a and 9b. The skeletons surround the holes inside the polygon, and they contain loops. Loops can be replaced with loop centroids based on the rule applicable to type 3 triangles (Figure 1, Figure 3d, Figure 7a 4).

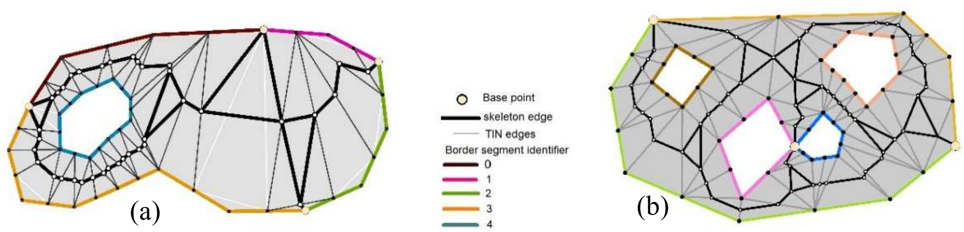
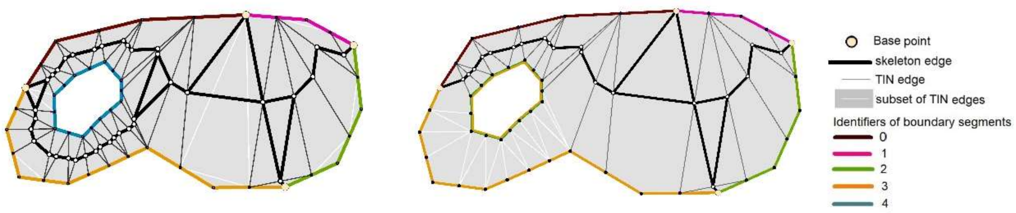


Figure 9. Skeletons generated by the BPSplit algorithm based on the spatial data presented in Figure 8 [10]: (a) skeleton of a polygon with a single hole; (b) skeleton of a polygon with four holes.

In the second research task, the authors assumed that skeletons which account for all holes inside polygons can be modified according to need. The process of modifying a skeleton will be presented on the example of the skeletons shown in Figure 9.

The BPSplit algorithm generates a skeleton based on the midpoints of selected TIN edges. The selected TIN edges have to come into contact with different polygon boundary segments. A polygon with a single hole features segments of the external boundary line and the hole boundary line which is regarded as the successive segment. If the ID of the hole boundary is replaced with the ID of a single segment of the external boundary, a skeleton will not be generated between the hole and that segment of the external boundary. This solution was applied to a polygon with a single hole (Figure 9a). It was assumed that a skeleton should not be located south of the hole. Therefore, the hole ID was replaced with the ID of the segment of the external boundary located south of the hole. As a result, the hole and the segment of the external boundary were assigned identical IDs. TIN edges that touch the hole and the analyzed segment of the external boundary were not considered in the process of selecting TIN edges. A subset of TIN edges can be created by comparing the selection of TIN edges with the BPSplit algorithm based on source data (Figure 10a) and the selection of TIN edges based on the modified ID of the hole boundary line (Figure 10b). The subset of TIN edges is the difference between TIN sets. To modify the skeleton, the edges that touch TIN edges are eliminated from the above subset. The skeleton generated based on source data is presented in Figure 10a, and the modified skeleton is shown in Figure 10b. Changes in the color of the hole boundary and the subset of TIN edges for modifying the skeleton are presented in Figure 10b.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 September 2020 doi:10.20944/preprints202009.0721.v1

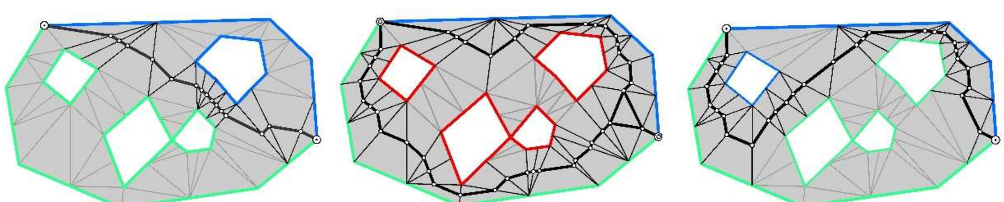


(a) (b)

Figure 10. Skeletons generated for a polygon with a hole: (a) skeleton generated based on source data; (b) skeleton modified on the assumption that the skeleton should not be located south of the hole (the island ID was modified by changing the color of the island boundary and the subset of TIN edges – the edges touching the subset of TIN edges were eliminated).

Other solutions can be applied to polygons containing a large number of holes (Figure 9b):

- To automate the process, each hole boundary (island boundary) can be assigned an ID corresponding to the nearest segment of the external polygon boundary (based on hole centroids). TIN edges between holes and the external boundary lines will be disregarded in the process of selecting TIN edges that touch different segments of polygon boundaries. The edges of the modified polygon will be located between holes, on the polygon's centerline (Figure 11a).
- If a single ID is assigned to all holes, TIN edges between holes will not be considered during the automatic selection of TIN edges (Figure 11b). The edges of the skeleton will be located near two external polygon boundaries (Figure 11b).
- The ID of a selected segment of the external polygon boundary can be assigned to the boundary lines of selected holes (islands). TIN edges between selected holes and TIN edges between selected holes and selected segments of the external boundary are not taken into account in this solution (Figure 11c).



(a) (b) (c)

Figure 11. Skeleton modification based on three criteria: (a) the skeleton should be located in the center of the polygon between holes (islands); (b) the skeleton should omit holes (islands) and should be located along the polygon's external boundary; (c) the skeleton can also be modified by changing the ID of hole boundaries.

The third research task is presented in the next section, based on the discussed example.

3. Results

The BPSplit algorithm was applied to generate a skeleton of a hydrographic network based on vector data from the Database of Topographic Objects in 1:10000 scale (BDOT_10). The database was created by public administration services based on aerial images. Data concerning lakes and rivers were selected from the database, and they were used to present the operation of the BPSplit algorithm.

 doi:10.20944/preprints202009.0721.v1

3.1. The application of the BPSplit algorithm to different sets of base points in a selected object

In the first example, the algorithm was applied to generate a skeleton of a hydrographic network comprising two large lakes and two small lakes connected by a river, as well as seasonal inflows to the largest lakes (Figure 12). All lakes are represented by polygons. The river flows through all four lakes. The river and seasonal streams are represented by lines.

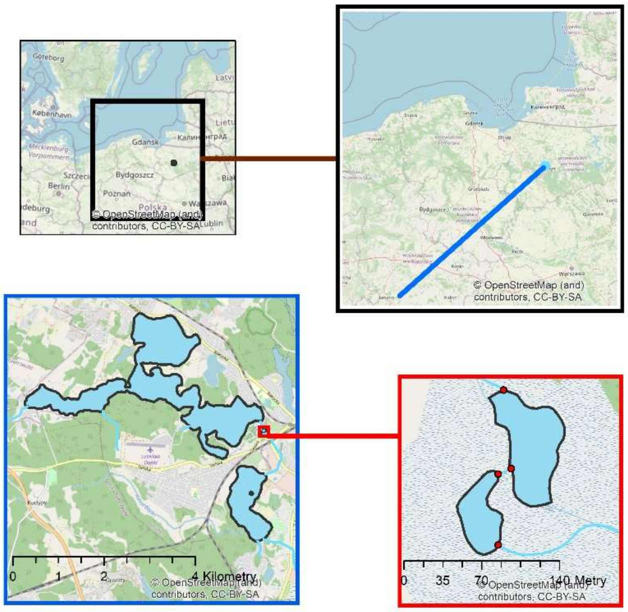


Figure 12. The analyzed object – four lakes connected by a river with seasonal streams flowing into the lakes.

The skeleton of the hydrographic network was generated based on source data in the following steps:

1. A set of base points was created.
2. Polygon boundaries were divided into segments based on base points.
3. TIN edges were generated inside polygons based on polygon apices and base points.
4. TIN edges that touch different segments of polygon boundaries, but do not touch base points, were selected.
5. The midpoints of selected TIN edges were generated.
6. Polygons were divided into segments.
7. Skeleton edges were generated between the midpoints of selected TIN edges, and between the midpoints of TIN edges and base points.

The modeled hydrographic network consists of polygon skeletons and river lines.

Three solutions were applied to generate a skeleton of the hydrographic network. In the first solution, the skeleton was created based on the examined lakes and the intersecting river. Eight (8) base points were generated to indicate the locations where the river flows in and out of the analyzed lakes. Seasonal streams entering the examined lakes were considered in the second solution, and, eleven (11) base points were generated. In the third solution, the skeleton of the largest lake was modeled for the purpose of navigation. Attractive points on the lake shore were selected as the base points. The results produced by each solution are presented in Figures 13-19.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 September 2020 doi:10.20944/preprints202009.0721.v1

3.1.1 Generation of a skeleton of a hydrographic network based on lakes and one river

In the first solution, the Kortówka River that flows through the four analyzed lakes was represented with a line. Eight base points marking the locations were the river flows in and out of the studied lakes were selected. Base points were used to divide lake boundaries into nine segments, including one segment representing the island boundary (Figure 13).

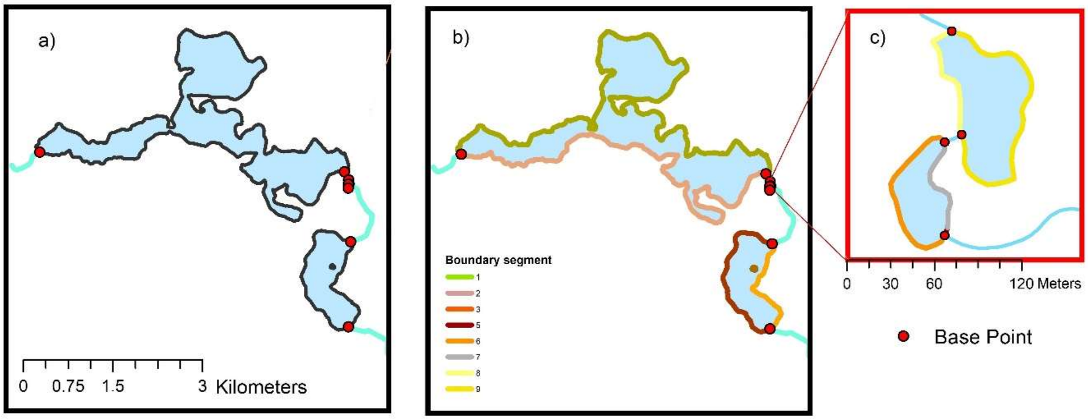


Figure 13. Source data for generating the first skeleton based on 8 base points and 9 segments of polygon (lake) boundaries; a) lakes with the location of base points; b) segmentation of the boundary lines of lakes; c) two small lakes with segments of their boundary lines.

Boundary segments were used to select TIN edges which play an important role in the process of developing skeletons. The skeletons of lake polygons were generated based on the midpoints of selected TIN edges and the set of base points (Figure 14).

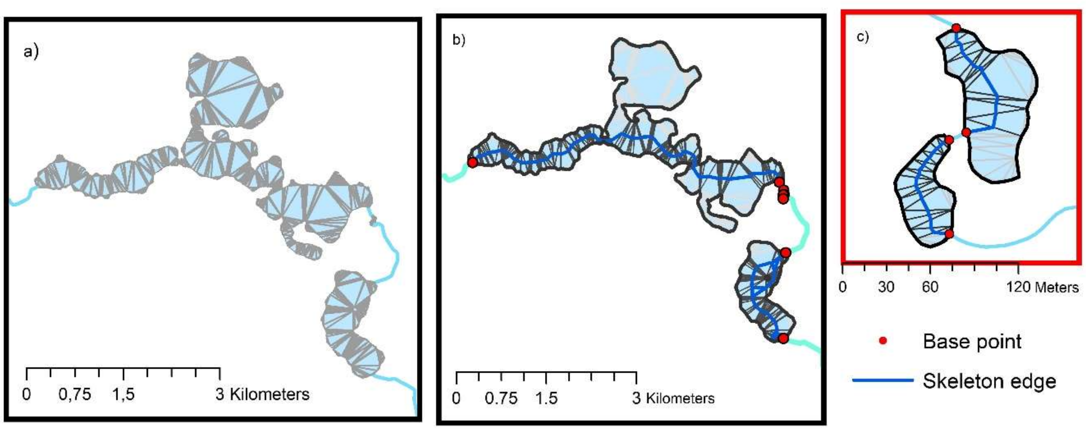


Figure 14. Skeleton of a hydrographic network generated with the BPSplit algorithm based on selected base points; a) TIN edges in polygons (lakes) ; b) selected TIN edges; c) two small lakes with selected TIN edges.

The network was generated based on polygon segments which were obtained by dividing polygons with selected TIN edges (Figure 15). Only several segments had the shape of type 1 triangles. Type 1 triangles are encountered in narrow water bodies with a monotonous shore line,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 September 2020 doi:10.20944/preprints202009.0721.v1

whereas type 3 triangles are found in the vicinity of islands. Skeleton lines were plotted in these segments (triangles) according to the rule presented in Figure 7.

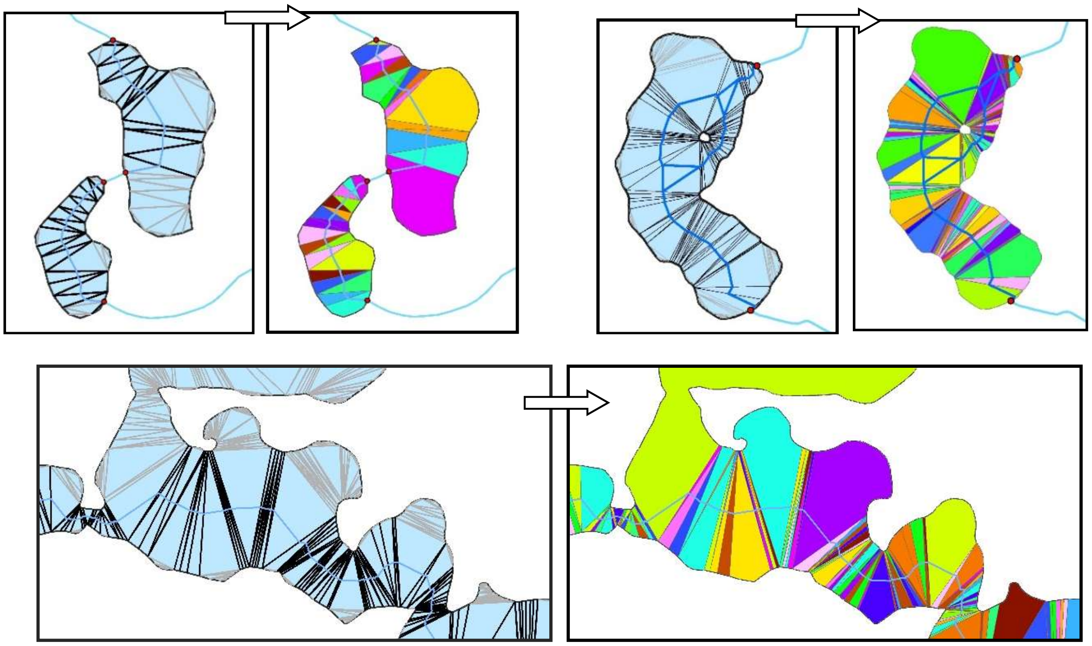


Figure 15. Examples of different polygon segments for generating a skeleton based on the midpoints of selected TIN edges. A few segments have the shape of type 1 and 3 triangles from Figure 7.

3.1.2 Generation of a skeleton of a hydrographic network based on lakes, one river and three streams

In the second solution, the main river intersecting all of the analyzed lakes and three seasonal streams flowing into these lakes were taken into account in the process of generating a skeleton. The number of base points increased by three, relative to the previous solution (3.1.1). These points were added to the boundaries of polygons representing the two largest lakes. The location of base points, selected TIN edges, skeletons generated on lakes, and seasonal streams are presented in Figure 16. Geometric figures in boundary segments touching base points are shown in Figure 17.

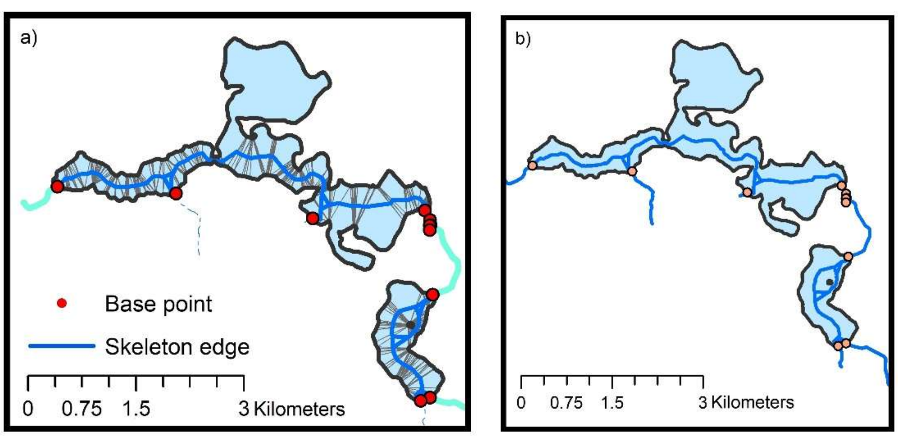


Figure 16. Skeleton of the hydrographic network generated based on 11 base points in four lakes, one river intersecting all lakes, and three seasonal streams flowing into the analyzed lakes; a) generation process; b) results.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 September 2020 doi:10.20944/preprints202009.0721.v1

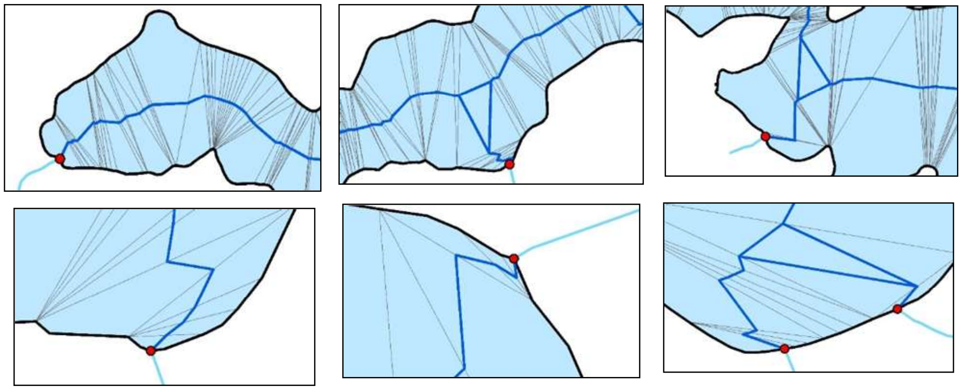


Figure 17. Geometric figures in boundary segments touching base points.

3.1.3 Generation of a skeleton for navigation in the largest lake with a diverse shore line

The third solution involves only the largest lake as well as base points marking the most attractive sites on the lake shore. The generated skeleton can be used for navigation (Figure 18).

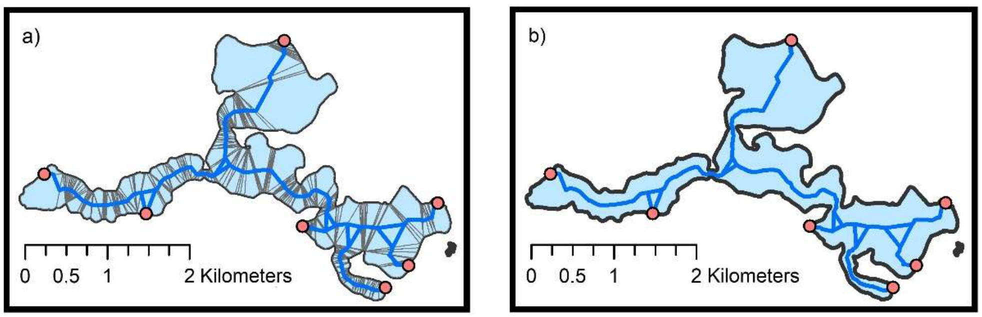


Figure 18 Skeleton of a polygon generated based on 7 base points marking attractive sites on the lake shore. The skeleton can be used for navigation; a) generation process; b) results.

3.2. The results generated by the BPSplit algorithm on a lake with numerous islands

In the next solution, the BPSplit algorithm was tested on a polygon containing many holes. The studied object was Lake Wydmińskie which is characterized by a diverse shore line and numerous islands (Figure 19). A skeleton was generated for the purpose of navigation. Base points representing the launch point and mooring points were selected on the lake shore.

Figure 19. Lake Wydmińskie with numerous islands

Data were processed with the BPSplit algorithm. The results of successive transformations are presented in Figure 20. The generated skeleton is shown in Figure 21.

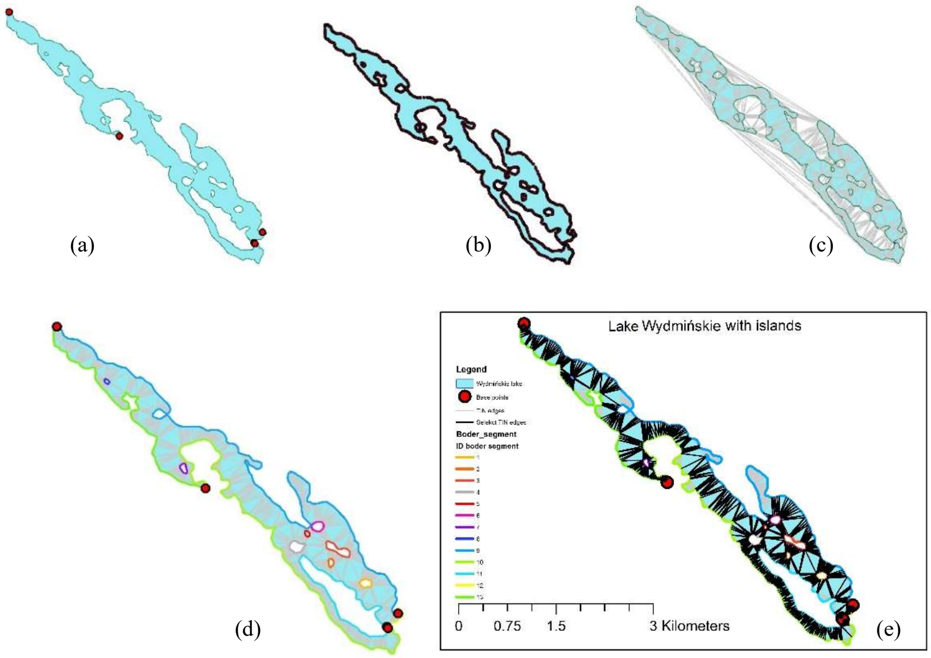


Figure 20. Data transformation in the skeleton generation process: (a) location of base points, (b) apices of a polygon representing the analyzed lake; (c) generated TIN edges; (d) TIN edges inside the polygon representing the analyzed lake, with a visualization of boundary segments; (e) process visualization, with an indication of TIN edges selected for skeleton generation.

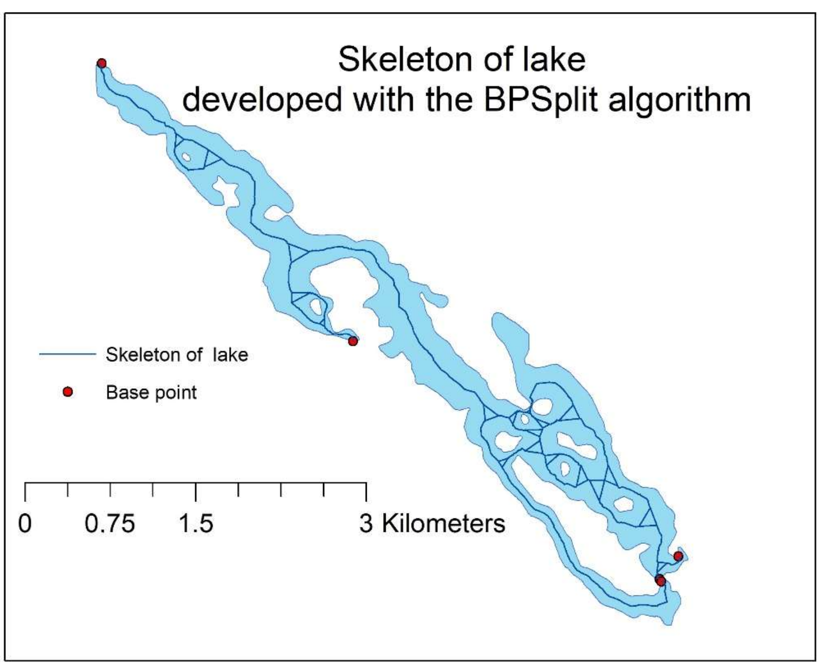


Figure 21. Skeleton of a lake with numerous islands developed with the BPSplit algorithm based on the adopted set of base points.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 September 2020 doi:10.20944/preprints202009.0721.v1

The presented skeleton was generated based on the location of base points, and it does not cover the entire lake (polygon). The skeleton accounts for significant areas around selected base points. The skeleton covers all islands. Skeleton edges were generated on both sides of the islands.

3.3. Modification of a skeleton representing a polygon with many holes – validation of the solution proposed in the second research task

The assumption that a skeleton between islands can be modified automatically was validated on the example of Lake Wydmińskie (Figure 21) as the base skeleton. Selected TIN edges for building the skeleton were used in the modification process. These edges formed the base set of TIN edges. Two modifications involving automatic changes in the ID of island boundaries are presented below.

3.3.1 Modification of the base skeleton to generate a skeleton between islands in the center of the lake by assigning the ID of the nearest segment of the polygon boundary to island boundaries

The base skeleton was modified by replacing the ID of island boundaries with the ID of the nearest segment of the external lake boundary. When the ID of island boundaries was modified, a new set of TIN edges touching different segments of the polygon boundary was selected. The base set of TIN edges was compared with the new set of TIN edges, and a subset of TIN edges was created based on the difference between the compared sets. The base skeleton was modified by removing the edges touching the subset of TIN edges. The modification process is presented in Figure 22. The new IDs of island boundaries are marked in color in Figure 22a. The subset of TIN edges for modifying the skeleton is presented in Figure 22b. The process of eliminating skeleton edges that touch TIN edges is shown in Figure 22c. The resulting skeleton is presented in Figure 22d.

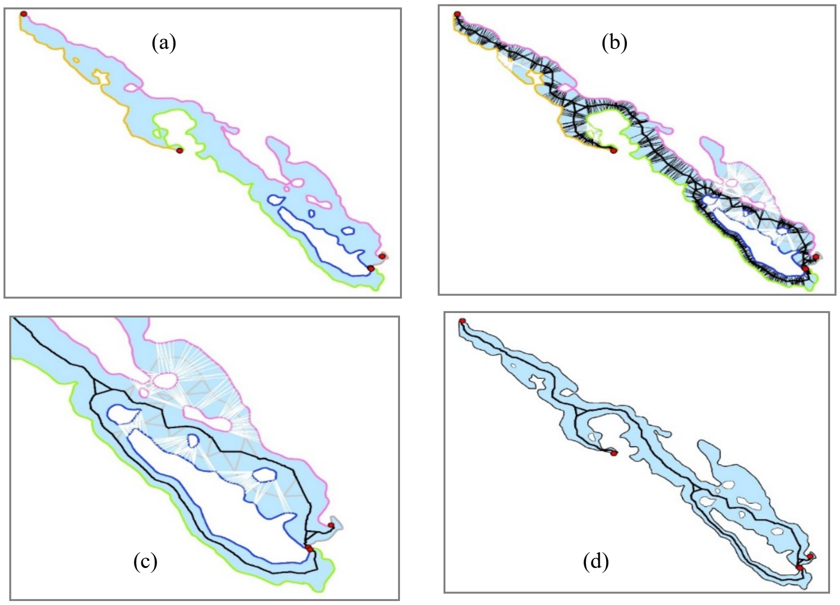


Figure 22. Modification of the base skeleton to obtain a skeleton in the center of the polygon between islands: (a) island boundaries are assigned new identifiers, and the resulting changes are marked in color; (b) a subset of TIN edges for modifying the skeleton is marked in white; (c) elimination of skeleton edges that touch TIN edges; (d) the resulting skeleton.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 September 2020 doi:10.20944/preprints202009.0721.v1

3.3.2. Modification of the lake skeleton to obtain a skeleton that is not located between islands when island boundaries are assigned an identical ID

If all island boundaries are assigned an identical ID, the modified lake skeleton will be located along the lake's external boundaries, and it will not be located between islands. The successive steps of the skeleton modification procedure, during which the edges of the base skeleton touching the subset of TIN edges were removed, are presented in Figure 23.

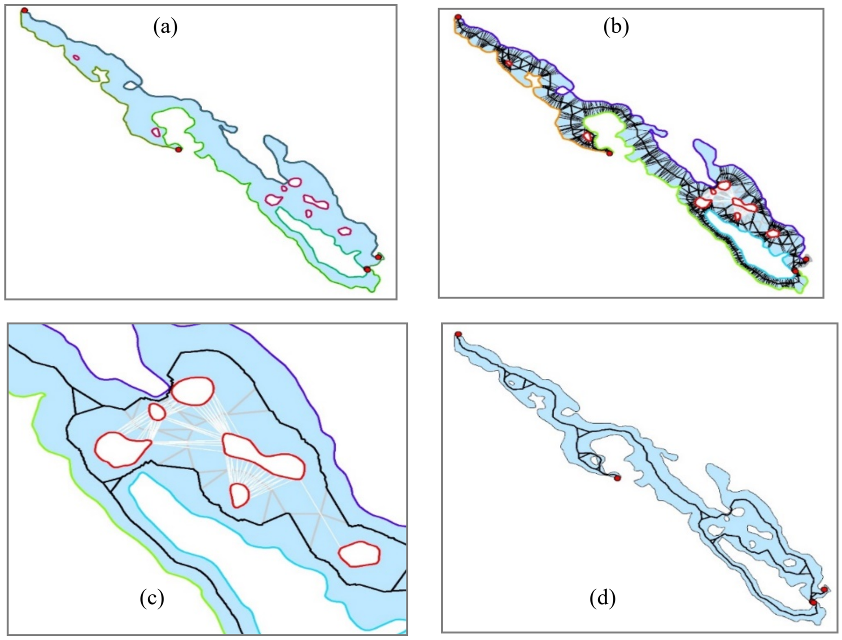


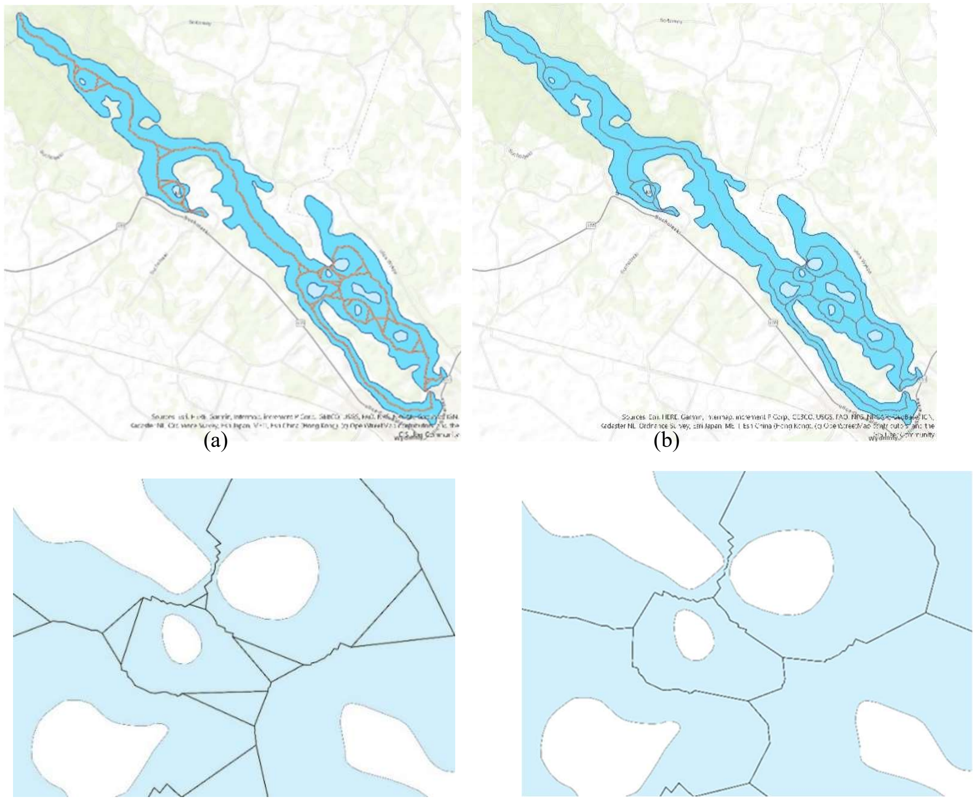
Figure 23. Modification of the base skeleton to obtain a skeleton located along the lake's external boundaries when the boundaries of all islands are assigned an identical ID: (a) the segments of island boundaries are assigned new identifiers; island boundary lines are marked with the same color; (b) the subset of TIN edges for modifying the skeleton is marked in white; (c) elimination of skeleton edges that touch the subset of TIN edges; (d) the resulting skeleton.

A base skeleton with holes can also be modified by changing the ID of island boundaries based on island attributes. The described procedures can be applied to predict the outcome of such modifications.

4. Skeleton generalization (smoothing) – validation of the third research task

In most solutions where skeletons are generated based on TIN edges, concave polygons and polygons with holes (islands) are characterized by the presence of loops and zigzag effects. The third research task was validated by smoothing the skeleton of Lake Wydmińskie (Figure 21). The loop was replaced with loop centroids (Figure 24) with the use of an algorithm developed by the authors in Python. A similar procedure was applied in the solution for type 3 triangles. The elimination of the loop reduces the number of nodes in the skeleton.

 doi:10.20944/preprints202009.0721.v1

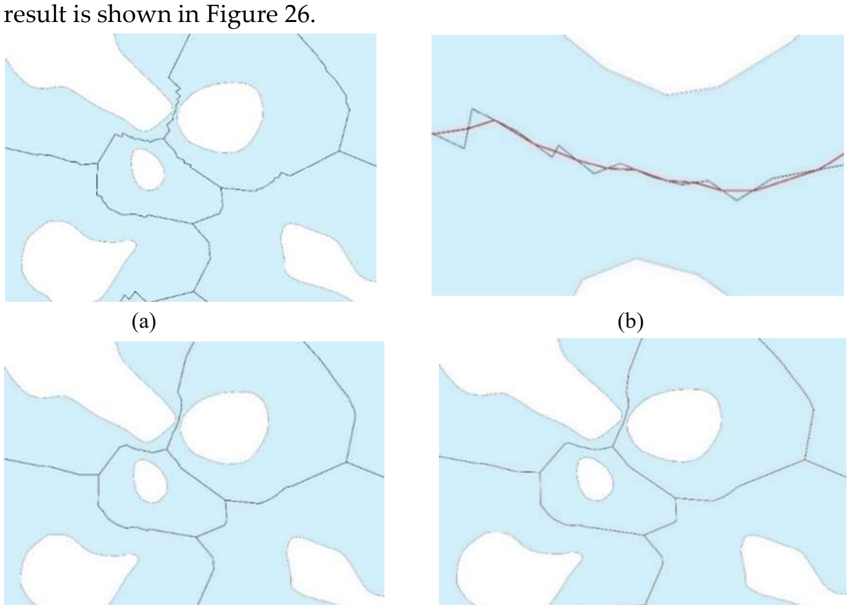


(a) (b)

(c) (d)

Figure 24. Skeleton model: (a) model with loops; (b) model where loops were replaced with loop centroids; (c, d) fragments of the skeleton before and after loop removal.

Most polygons generated based on TIN edges have zigzag effects. Zigzag effects are minimal when the polygon has uniform width, and they are more extensive in polygons with varied width (Figures 24c; Figure 25a). Various line simplification methods are described in the literature [23-27]. An alternative smoothing method was proposed in this study. In this approach, skeleton segments are replaced by new segments based on the midpoints of successive skeleton segments (Figure 25b). To prevent the formation of new loops during smoothing, the points representing network nodes were adopted as fixed points. The set of base points was combined with the set of fixed points. It was assumed that the location of fixed points would not change during smoothing. Successive steps (iterations) of the smoothing procedure are presented in Figures 25c and 25d, and the result is shown in Figure 26.



(a) (b)

(c) (d)

Figure 25. Elimination of zigzag effects in the network skeleton by establishing fixed points at network nodes: (a) zigzag effects in the network; (b) visualization of the smoothing method; (c) results of smoothing after the first iteration; (d) results of smoothing after the second iteration.

In successive iterations of the smoothing process, the location of the resulting object should be controlled inside the polygon.

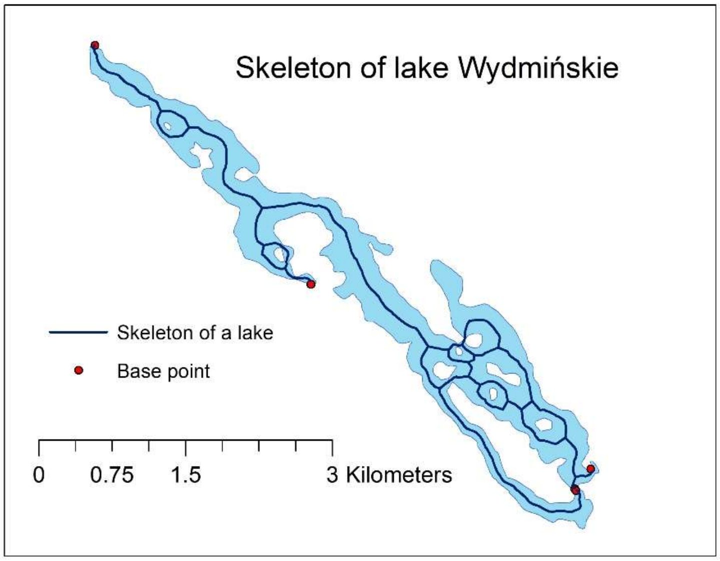


Figure 26. A model of the network skeleton generated in a polygon representing Lake Wydmińskie based on the adopted set of base points, after loop elimination and two iterations of the smoothing procedure to remove zigzag effects.

5. Discussion

The proposed BPSplit algorithm is an alternative method of generating polygon skeletons. The algorithm can be applied to various sets of spatial data, and to specific cases. The BPSplit algorithm for generating a polygon skeleton requires a set of base points on the polygon boundary line. Base points determine the shape of the generated skeleton, and they play an important role during the selection of TIN edges. TIN edges for skeleton generation are selected based on topological relations during data processing. The set of base points supports the determination of initial conditions in the process of generating a skeleton.

The proposed algorithm was validated on data sets with varied geometry, based on different sets of base points. The algorithm validated the first research task postulating that the anticipated shape of the generated skeleton can be defined.

In lakes with numerous islands (polygons with numerous holes), the BPSplit algorithm generates skeleton edges between islands, and between islands and external lake boundaries. The location of the skeleton between islands can be modified by the user. The presented examples validate the second research task.

The third research task was also successfully validated by removing loops and smoothing the skeleton based on the defined nodes and base points as fixed points.

Author Contributions: Conceptualization, Elżbieta Lewandowicz; Methodology, Elżbieta Lewandowicz; Software, Elżbieta Lewandowicz; Software by Python, Paweł Flisek; Validation, Elżbieta Lewandowicz; Formal Analysis, Elżbieta Lewandowicz; Writing-Original Draft Preparation, Elżbieta Lewandowicz; Writing-Review & Editing, Elżbieta Lewandowicz; Visualization, Elżbieta Lewandowicz. All authors have read and agreed to the published version of the manuscript.

 doi:10.20944/preprints202009.0721.v1

Funding: This research was financed as part of a statutory research project of the Faculty of Goengineering of the University of Warmia and Mazury in Olsztyn, Poland, entitled "Geoinformation from the theoretical, analytical and practical perspective", (No. 29.610.008, timeline: 2017–2020).

Acknowledgments: The authors are grateful to the Regional Center for Geodetic and Cartographic Documentation in Olsztyn for providing access to the Database of Topographical Objects in 1:10 000 scale. The acquired data supported analyses in the tested objects.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lee, D.T. Medial Axis Transformation of a Planar Shape. *IEEE Trans. Pattern Anal. Mach. Intell* **1982,** 4, 363–369.
2. Chin, F., Snoeyink, J., Wang, C. A. Finding the Medial Axis of a Simple Polygon in Linear Time. – Proc. 6th Ann. Int. Symp. *Algorithms and Computation* **1995**
3. Karimipour F., Ghandehari M., Ledoux H. Watershed delineation from the medial axis of river networks, *Computers & Geosciences* **2013**; Volume 59, Pages 132-147.
4. Prasad, L., 1997: Morphological Analysis of Shapes. *CNLS Newsletter 139:1.* **1997;** http://www.lanl.gov/Highlights/-07
5. Eppstein, D., Erickson, J., Raising Roofs, Crashing Cycles, and Playing Pool: Applications of a Data Structure for Finding Pairwise Interactions. – Discrete Computer Geometry **1999**; 22, pp. 569-592.
6. Gold C., Spatial modeling: Voronoi diagrams. In: Richardson D, Castree N, Goodchild MF, Kobayashi A, Liu W, Marston RA (eds) *International encyclopedia of geography: people, the earth, environment and technology*. Wiley, Hoboken, **2017;**
7. Li Ch., Yin Y., Wu P., Wu W.. Pattern recognition and correction method for skeleton lines at small patch boundaries", *Transactions in GIS*, **2020**; https://doi.org/10.1111/tgis.12643
8. Li Ch., Yin Y., Wu P., Wu W. A method for the extraction of partition lines from long and narrow patches that account for structural features , *Transactions in GIS*, **2019**; 23(5), DOI:10.1111/tgis.12523
9. Ai, T., Ke, S., Yang, M., & Li, J. Envelope generation and simplification of polylines using Delaunay triangulation. *International Journal of Geographical Information Science* 2017; 31(2), 297–319.
10. Meijers M., Savino S., Van Oosterom P.. SPLITAREA: an algorithm for weighted splitting of faces in the context of a planar partition. *Journal International Journal of Geographical Information Science* **2016**; Volume 30, 2016 - Issue 8
11. Meijers, M.,(2011); Variable-scale Geo-information. Published by: NCG, *Nederlandse Commissie voor Geodesie, Netherlands Geodetic Commission, Delft, the Netherland* **2011**; pp. 236
12. Aigner W., Aurenhammer F., Jüttler B. 2015; *On triangulation axes of polygons. Mathematics, Computer Science, Information Processing Letters* 2015; Volume 115, Issue 1, pp. 45-51.
13. Wang Z., Yan H. An algorithm for extracting main skeleton lines of polygons based on main extension directions. International Conference on Electronics, Communications and Control (ICECC), Ningbo, **2011,** pp. 125-127. DOI: 10.1109/ICECC.2011.6067859
14. Lewandowicz E., Flisek P. A Method for Generating the Centerline of an Elongated Polygon on the Example of a Watercourse. *ISPRS International Journal of Geo-Information* 2020; *9*(5), 304.
15. Li, C.; Yin, Y.; Wu, P.; Liu, X.; Guo, P. Improved Jitter Elimination and Topology Correction Method for the Split Line of Narrow and Long Patches. *ISPRS Int. J. Geo-Inf.* **2018**, 7, 402.
16. Wang T.. Extraction of optimal skeleton of polygon based on hierarchical analysis. Computer Science 2010
17. Giesen, J., Miklos, B., Pauly, M. Medial axis approximation of planar shapes from union of balls: a simpler and more robust algorithm. In: Canadian Conference. *Computational Geometry* **2007;** pp. 105–108.
18. Giesen, J., Miklos, B., Pauly, M., Wormser, C. The scale axis transform. In: Proceedings of the 25th Annual Symposium on *Computational Geometry* **2009**; pp. 106–115. doi:10.1145/1542362.1542388.
19. Haunert, J.H.; Sester, M. Using the straight skeleton for generalization in a multiple representation environment. *In Proceedings of the ICA Workshop on Generalization and Multiple Representation, Leicester, UK, 20–21 August* **2004.**

20. Szombara, S. Comparison of methods used in cartography for the skeletonisation of areal objects. Geoinform. Pol. **201;** 14.
21. ESRI 2020, ArcGIS https://www.esri.pl/,
22. ESRI 2020, Tool: Creates centerlines from polygon features. https://pro.arcgis.com/en/pro-app/tool-reference/topographic-production/polygon-to-centerline.htm
23. Douglas, D.H.; Peucker, T.K. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature.Can. *Cartography* 1973; 10, 112–122.
24. Li, Z.L.; Openshaw, S. 1992. Algorithms for automated line generalization based on a natural principle of objective generalization. *Int. J. Geogr. Inf. Syst.* **1992**, 6, 373–38.
25. Qian H., Zhang M., Wu F. A New Simplification Approach Based on the Oblique-Dividing-Curve Method for Contour Lines. *ISPRS Int. J. Geo-Inf.* **2016**; 5(9), 153; https://doi.org/10.3390/ijgi5090153
26. Shi W., Cheung Ch., K. () Performance Evaluation of Line Simplification Algorithms for Vector Generalization. The Cartographic Journal **2006**; 43,(1),27-44, Taylor & Francis, https://doi.org/10.1179/000870406X93490
27. Kozioł K., Znaczenie punktów stałych obiektu w procesie upraszczania (The importance of fixed points in the simplification proces), Archiwum Fotogrametrii, Kartografii i Teledetekcji **2012**; Vol. 23, 2012, s. 169–177.