GEODESY AND CARTOGRAPHY
ISSN 2029-6991 print / ISSN 2029-7009 online
2012 Volume 38(3): 98–105
doi:10.3846/20296991.2012.728045

UDK 528.914

CHANGES AFFECTING GENERALIZATION OF LAND COVER FEATURES IN A SMALLER SCALE

Lina Papšienė[1], Kęstutis Papšys[2]

[1]Vilnius Gediminas Technical University, Saulėtekio al. 11, LT-10223 Vilnius, Lithuania
[2]Vilnius University, M. K. Čiurlionio g. 21, LT-03101 Vilnius, Lithuania
E-mails: [1]gkk@vgtu.lt (corresponding author); [2]kestutis.papsys@gf.stud.vu.lt

Received 10 June 2012; accepted 21 September 2012

Abstract. Reference spatial data sets represent the least changing natural and anthropogenic features of terrine. As a rule, such data are stored in different scales and most frequently updated consequently starting with a spatial data set of a larger scale (usually base scale) thus later performing an update of data in smaller scales. The generalization of features in a larger scale is one of the major processes employed in the creation and update of spatial data of a smaller scale. In order to effectively carry out works, it is recommended to use automatic procedures and generalization only in those cases when changes in features are significant, i.e. affect the update of features in a smaller scale. The article discusses the relation between changes in polygon features (identify land cover territories in a base spatial data set) and different generalization processes as well as the evaluation of significance of likely changes.

Keywords: significance of changes in features, update of features, generalization, polygon features, spatial data set, GIS.

1. Introduction

National mapping agencies (NMAs) often maintain reference and thematic spatial data sets to represent various spatial data identifying natural and anthropogenic phenomena of the world (Kazemi 2003). Usually they are stored in several scales, e.g. reference spatial data in Lithuania are collected at 1:10 000 (basic scale), 1:50 000 and 1:250 000 (Papšienė, Papšys 2011), whereas, for example, in Belgium it makes 1:10 000 and 1:50 000 (Bayers 2010). In reference data sets, polygon features serve to store the features representing land cover, such as forests, arable land, built-up territories, hydrographic features, etc. The main task to quickly and effectively update spatial data of a smaller scale is to use spatial data sets of a larger scale more often updated as those of a smaller scale. For example, reference spatial data in Lithuania at a scale of 1:10 000 are updated constantly while those at the scales of 1:50 000 and 1:250 000 - on a 5 year basis or even more.

The automatic generalization of spatial data is one of the most appropriate ways employed in the creation and update of a spatial data set. The use of automatic procedures in the update of spatial data is affected by three principal factors: reduction in work and time resources, the qualification and subjectivity of specialists (Kilpelainen 2000) and data accuracy achieved by manual update (McHaffie 2002). As a rule, along with generalization methods applied in the update of spatial data sets, all features of a larger scale are generalized regardless of whether those features have changed compared to the earlier spatial data set of a larger scale. Therefore, the features that have not really changed are also updated. In this way, each update produces absolutely new features (new data set) having no relation with the earlier feature version. Such update process requires high technological and human resources, as it takes time to generalize all features anew, revise the result later and evaluate whether it meets the set requirements. Accordingly, such generalization is more appropriate for creating rather than for updating a spatial data set based on larger scale data. While updating spatial data sets it is best to generalize only changes reducing costs. The changed features can be identified in the two following ways:

- comparing earlier and later versions of a feature through various queries;
- supporting unique IDs of the features that have to be implemented across all reference spatial data sets at all scales and tracking changes in the life cycle of the feature (Stankevičius 2008; Beconytė *et al.* 2009; Stankevičius *et al.* 2010). Feature IDs must be unique throughout the data set and remain unchanged all through the life cycle of the feature (INSPIRE... 2009).

2. Generalization of Polygon Features

The generalization process is defined as the process of selecting and simplifying representation details appropriate to the scale and/or purpose of the map (ICA… 1973). Digital cartography distinguishes three types of generalization by defining a process from reality to cartographic products (Grünreich 1985; Weibel, Dutton 1999; Cecconi 2003):

- Feature generalization is employed to create the initial abstract image of a phenomenon of the real world (e.g. from satellite images, GPS measurements). Feature generalization produces a primary data model.
- The process of model generalization performs controlled reduction in data. Model generalization is used for creating and updating a data set of a smaller scale from spatial data of a larger scale. Model generalization produces a secondary data model.
- Cartographic generalization is used for developing a cartographic product. This process comprises visualization operations and is employed for the generalization of the spatial features of a primary or secondary model in order to get a cartographic product.

Many years of scientific research had not seen a uniform and systematic classification of generalization operations. McMaster and Shea (1992) identified 12 different generalization functions that were classified according to the transformation type used for generalization:

1. *spatial (graphic) transformation*: simplification, amalgamation, refinement, displacement, smoothing, merging, exaggeration, aggregation, collapse and enhancement;
2. *attribute transformation*: classification and symbolization.

The AGENT project in 1999 specified the former classifications. Thus, generalization functions of spatial transformation were identified depending on the transformation type (*attribute transformation* and *spatial transformation*). In addition, generalization operations under *spatial transformation* have been furthermore divided depending on the features they can be applied for (AGENT... 1999):

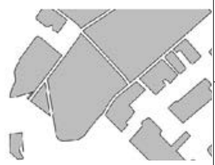
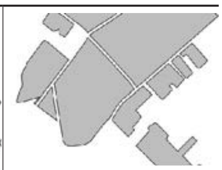
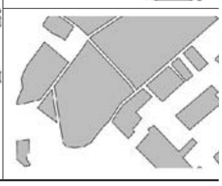
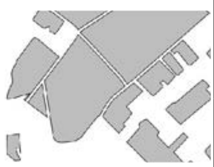
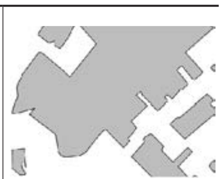
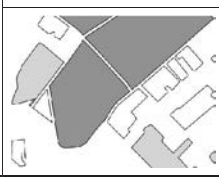
- *individual features*: simplification (weeding, unrestricted simplification) collapse, enhancement (enhancement with regards to geometric constraints, enhancement with regards to semantic constraints);
- *individual features or a set of features*: selection/elimination, displacement;
- *set of features*: aggregation (amalgamation, combine, typification).

Furthermore, in 2006, Li presented a systematic classification of generalization operations depending on what generalization could be applied for various geometric elements representing features. This classification identifies the groups of operations used for individual and a group of point, polyline or polygon features. Individual polygon features may be applied for the operations of collapse (including area-to-point, area-to-line and partial), displacement, exaggeration (including directional thickening, enlargement and widening), elimination, (shape) simplification, split, whereas a group of polygon features is applied for aggregation, agglomeration, amalgamation, dissolving, merging, relocation, (structural) simplification and typification.

Based on Li classification (2006), Table 1 specifies generalization operations that may be employed upon the model generalization of land cover features of a reference data set:

- elimination and simplification of individual polygon features;
- aggregation and dissolving a group of polygon features.

Table 1. Generalization operations for polygon features of land cover

Operation	Larger scale	Smaller scale
	For individual features	
Elimination eliminates small and unimportant features		
Simplification makes the feature shape simpler		
	For a group of features	
Aggregation combines polygon features within a specified distance to each other into a new polygon		
Dissolving aggregates features based on the specified attribute		

In order to properly perform the generalization of spatial features it is, first of all, necessary (Papšienė, Papšys 2011) to:

- determine requirements for features, i.e. the density of features, geometry resolutions, min. area;
- select algorithms and parameters of generalization operations;
- determine the priority of selected algorithms;
- model the generalization process.

The generalization of features must be done in separate object groups represented by the same phenomena of the world. The generalization of land cover features requires, primarily, the selection of proper features according to quality parameters (e.g. selection of deciduous forests). It should be mentioned that first we cannot eliminate at once features according to both their attributes and geometric features (e.g. select only forests with the area over 10 ha). The reason is that in the next step, the aggregation of the selected features according to a minimum distance preliminary defined between the neighbouring features, the features small in the area may, after aggregation, form the conglomerates of a significant size. The size of all features must be evaluated and the features that fail to meet requirements for a minimum feature area must be eliminated only after aggregation. The last step is the simplification of features. The conception process of the generalization of land cover is presented in Fig. 1.

3. Relation Between the Type of Polygon Changes and Generalization Process

The identification of changes in spatial data includes the analysis of feature versions at different periods (Singh 1989). The primary task of identifying changes in features is to decide which features have changed compared to the earlier version of a spatial data set and what is the type of changes in features that can be evaluated by answering several questions presented in Table 2.

Table 2. Possible types of changes in features

Question	Type of change	
	in a larger scale	in a smaller scale
Is the feature new?	*New feature*	*New feature*
Has an attribute of the feature changed?	*Updated feature attribute*	*New feature* or *Updated feature attribute*
Has the shape of the feature changed?	*Updated feature shape*	*Updated feature shape*
Has the minimum distance between neighbouring features changed?	*Updated feature shape*	*Updated feature shape (aggregated feature)*
Has the feature been deleted?	*Deleted object*	*Deleted object* or *Updated feature shape (no aggregated feature)*

The choice of generalization operations depends on the type of changes in features. Table 2 shows that some types of changes in a larger scale may affect different changes in a spatial data set of a smaller scale. For this reason, it is impossible to make an unambiguous decision as to what generalization is to be applied as long as all changes in features and likely influence

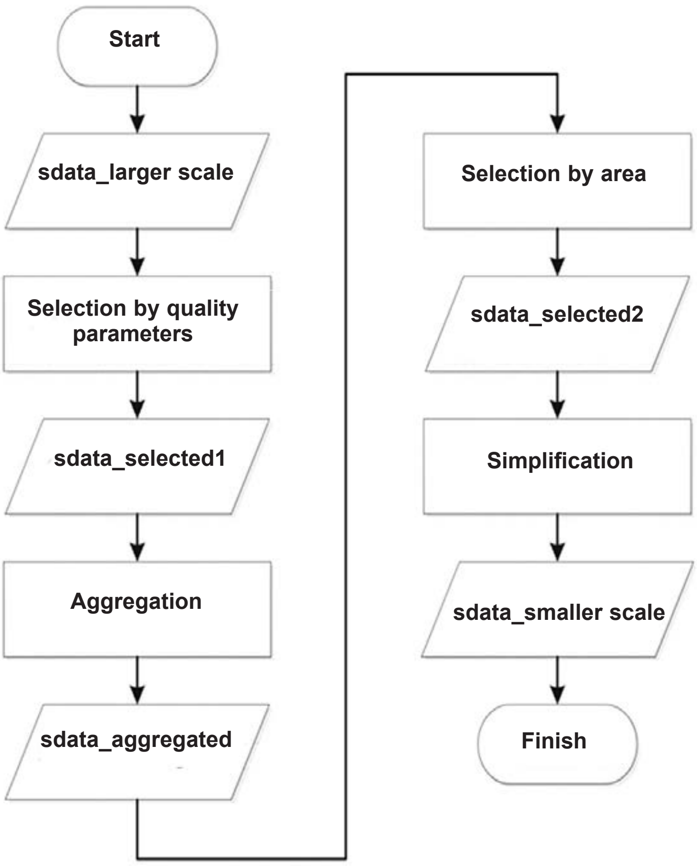


Fig. 1. The conception process of the generalization of land cover

on neighbouring features are not analyzed and evaluated.

A new feature in a smaller scale must be created in two cases (Fig. 2):

– a new feature is identified in a spatial data set of a larger scale with a quality or quantity attribute represented in a smaller scale;
– a feature in a larger scale acquires a new quality or quantity attribute represented in a smaller scale.

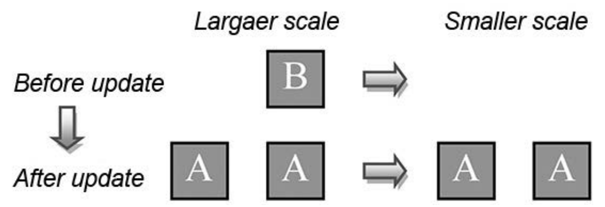


Fig. 2. Cases of creating a new feature

In these two cases, a feature is selected from a spatial data set of a larger scale applying a simplification operation; the achieved result is integrated into a spatial data set of a smaller scale.

If a feature is to be eliminated, no generalization operation is performed.

A feature in a smaller scale will have to be deleted in cases opposite to those of creating a new feature, i.e. when (Fig. 3):

– the deleted feature is identified in a spatial data set of a larger scale having a quality or quantity attribute usually represented in a smaller scale;
– the feature in a larger scale acquires a new quality or quantity attribute not represented in a smaller scale.

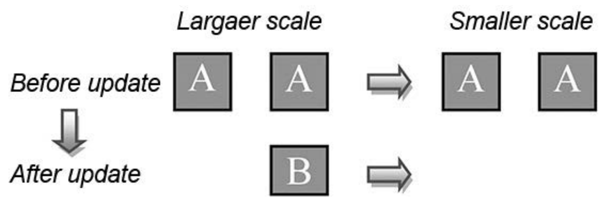


Fig. 3. Cases of deleting a feature

A feature in a smaller scale is updated when its quality or quantity attributes or shape in a spatial data set of a larger scale are changed (Fig. 4). In the first case, feature attributes and in the second, the feature shape is updated.

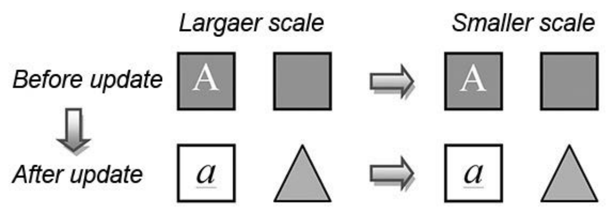


Fig. 4. Cases of updating features

If only the attributes of a feature have changed, no generalization is needed (only the attribute is updated) while in case of changes in the shape, feature simplification is to be carried out.

Additionally, the evaluation of the above cases shows it is necessary to evaluate the distance to the neighbouring features with the same attribute, i.e. whether it is above or below the minimum distance allowed:

1. a new feature or the feature that "moved towards" the neighbouring feature will be aggregated with it, i.e. the feature in a smaller scale will enlarge (e.g. when a new residential area emerged close to a former built-up territory) (Fig. 5);
2. upon elimination or "receding" the feature, that was earlier aggregated with the neighbouring one, will have to be eliminated from the aggregated polygon feature in a smaller scale, i.e. the feature will be reduced (e.g. gardening was started in one of the adjacent fields of the arable land) (Fig. 6);
3. a feature of the changed shape in a larger scale will have effect on the shape of the feature produced by aggregating neighbouring features (e.g. a part of one of adjacent forests was cut down) (Fig. 7).

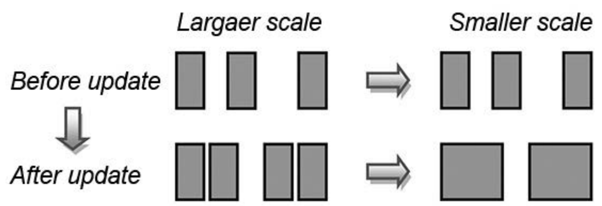


Fig. 5. Cases of enlarging an aggregated feature

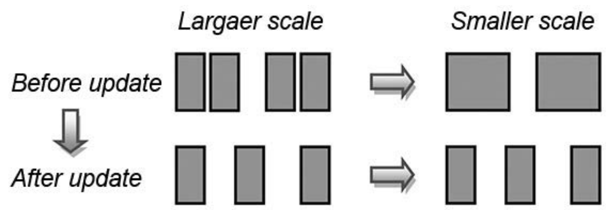


Fig. 6. Cases of reducing an aggregated feature

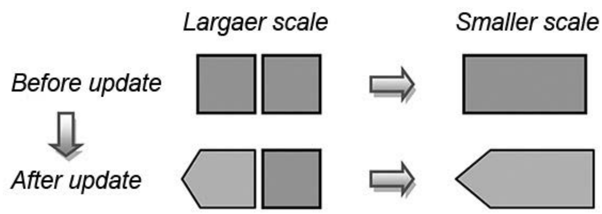


Fig. 7. Cases of changing the shape of an aggregated feature

The first case demands a feature simplification operation as well as aggregation with a neighbouring feature from a spatial data set of a smaller scale. In the second case, the "unsuitable" feature needs to be eliminated from the earlier aggregated polygon feature in a smaller scale. However, instead of the elimination function, it is enough to newly simplify and integrate the changed features (additionally, the aggregation function is used for merging the rest of the features). In the third case, similarly to the first one, the simplification of changed features -aggregation with adjacent features - is to be performed.

4. Possibilities of Evaluating the Significance of Changes in the Polygon Feature

When identifying changes in spatial features, it is essential to determine significant and to reject insignificant changes (Richard *et al.* 2005). A change will be significant when:

– the acquired new attribute is represented in a spatial data set of a smaller scale;
– a change in the feature shape will be seen in a spatial data set of a smaller scale.

Upon evaluating the significance of changes in features, a single changed feature needs to be analyzed establishing how its change will influence surrounding features according to the types of changes specified in the above section.

The evaluation of changes in features is to be conducted in groups according to the type of changes in the following procedure.

Group: update of the feature attribute. When evaluating the significance of changes in the feature attribute of a larger scale, it suffices to know what attributes are significant for (i.e. attributable to) the feature in a smaller scale and comparing different versions of the feature (feature version before and after the update in a larger scale) to single out the features that have acquired new "significant" attributes, e.g. the name of the lake has been specified.

Evaluation in this group of changes requires:

– to link feature attributes before and after the update through unique feature IDs (if there are any) or spatial join that creates a table join in which the field attributes of features before and after the update are presented based on the relative locations of the features;
– to find changed appropriate attributes through queries (*"AttributeValue_BeforeUpdate"* <> *"AttributeValue_AfterUpdate"*).

Group: creating a feature. The appearance of a new feature will be significant in cases similar to those of updating the feature attribute, i.e. if a new feature has proper quality and quantity attributes.

Evaluation requires:

– to eliminate, through queries, features (before and after the update) lacking set quality and quantity attributes, i.e. to select the proper ones;
– to find new features using information on the life cycle of the feature (if there is any) or selecting features "after the update" not intersecting the feature "before the update".

Group: eliminating a feature. Deleting a feature will be significant in a larger scale if the feature was earlier represented in a smaller scale.

Evaluation requires finding the eliminated features by using information on the life cycle of the feature (if there is any) or selecting features "before the update" not intersecting the feature "after the update".

Group: a feature under aggregation. The changed feature will be significant in respect of aggregation operation in case it "has approached" closer than the minimum distance allowed between neighbouring features considering the attributes of the same quality and quantity. Searching for such changes may be easily implemented through spatial analysis query looking for intersections between the buffer around the feature of a changed shape and/or new features and a neighbouring feature. The width of the buffer under creation must be equal to the defined minimum distance between neighbouring features.

Group: changing the shape of features. The evaluation of the significance of changes in the feature shape is more complex than that found in the cases mentioned above.

The significance of changes in the feature shape is determined by comparing the size of changes referring to the fixed minimum change allowed, which, first of all, should depend on the scale (this scale affects the resolution of the map) and specificity of a spatial data set. Additionally, the expected changes in the size of the object have to be evaluated after generalization.

Before the analysis of changes in the feature shape is started, the features with changed shape are immediately rejected, if they are not represented in smaller scale maps according to parameters of quality and quantity.

A model has been developed to evaluate the significance of changes in the feature shape using spatial analysis queries (by *ESRI ArcGIS* software). The purpose of the model is to compare the features of a smaller scale before and after the update, to find changes in the feature shape, to evaluate them and to select the significant ones that have to be simplified. The model has been developed on the presumption that changes in the feature shape will be significant if the vertex/vertexes of changes (expressed in polygon) will be moved from the feature boundary (before the update) at a distance higher than the set (*s*) is placed, which must, as mentioned above, depend on the scale (usually it makes 0.05 cm of the map scale).

1. The actions in the query of spatial analysis follow the order below (Fig. 8).
2. Union of spatial features before (*sdata.v1*) and after the update (*sdata.v2*). Selecting the changed part of the feature (*ChangesMinus_all, ChangesPlus_all*).
3. Determining the significance of buffer width. The size of the buffer depends on specified resolution that relates to the resolution of spatial data.
4. Creating buffer zones (according to the specified significance of buffer width) around the source feature before the update (*Buffer*).
5. Selecting changes in feature geometry outside the buffer zone (*ChangesMinus, ChangesPlus*).
6. Simplifying changes in the geometry of the selected feature (*ChangesMinus_simplify, ChangesPlus_simplify*).

7. Creating buffer zones (according to the specified significance of buffer width) around updated objects (*Buffer*).
8. Selecting changes in the geometry of the objects outside the buffer (*ChangesMinus_Significant, ChangesPlus_Significant*). The resulted changes will be significant.

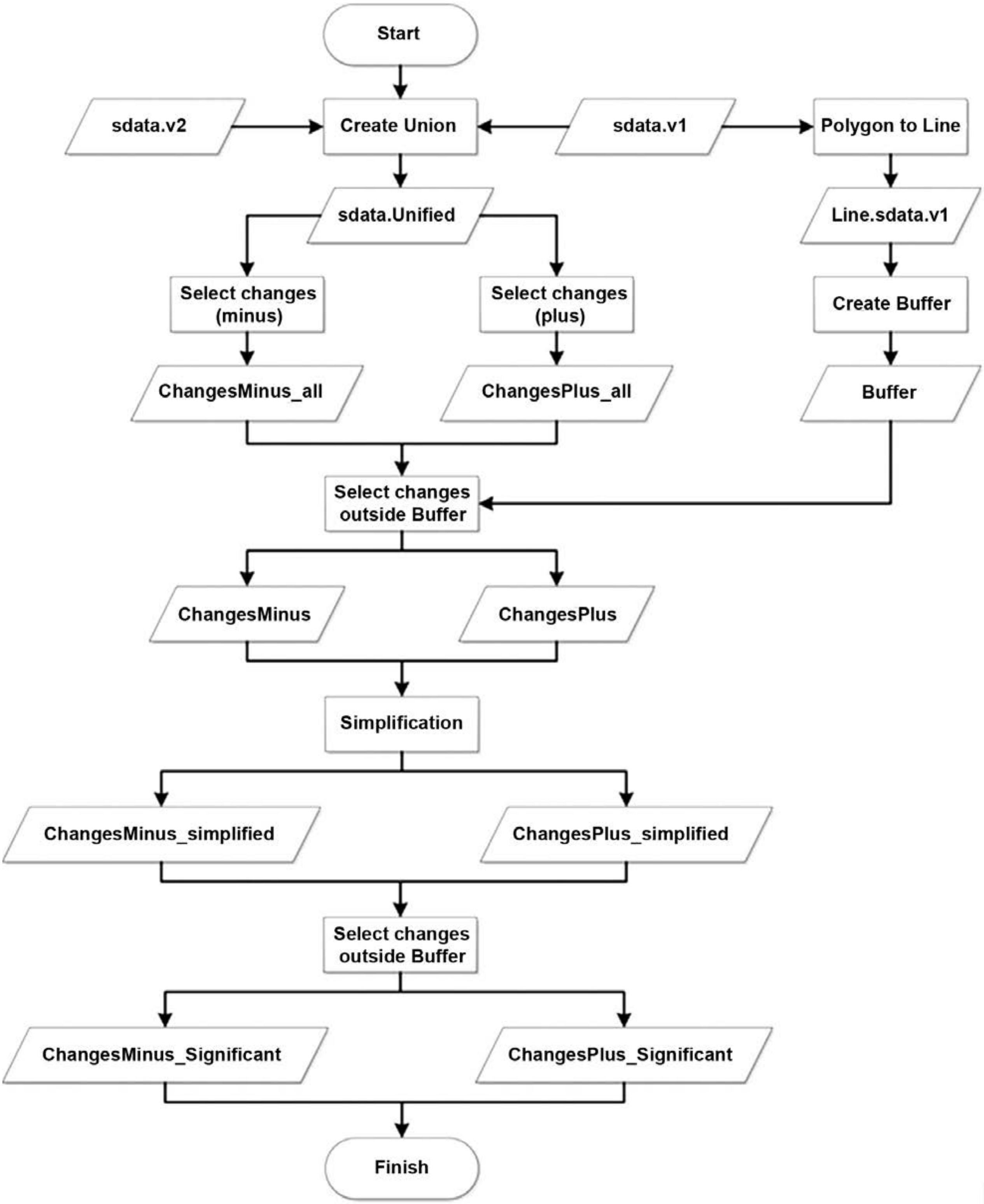


Fig. 8. Processes for determining significant changes in the features of land cover with a changed shape

5. Evaluation Test on Changes in Polygon Features

The period from 2009 to 2010 faced the development of Lithuanian digital raster orthophotographic map ORT10LT at a scale of 1:10 000, which served as a base for updating the features of Lithuanian reference spatial data set at a scale 1:10 000 in 2011. Furthermore, the period from 2011 to 2012 has been witnessing the update of Lithuanian reference spatial data set at a scale of 1:50 000 using the already updated spatial data at a scale of 1:10 000.

Research task: check the correctness of the above described evaluation methods for feature changes. Research object: features representing built-up territories in Lithuanian reference spatial data sets at a scale of 1:10 000. The target territory: the municipality of Molėtai Region.

Accordingly, the following features have been analyzed and compared with:

Lithuanian reference spatial data set at a scale of 1:10 000 before the update (2010 version);

Lithuanian reference spatial data set at a scale of 1:10 000 after the update (2011 version).

The research results are presented in Table 3.

Table 3. Test results

Group of changes	Build-up territories	
	Changes in 1:10 000	Changes significant for 1:50 000
Update of the feature attribute	0	0
Creation of the feature	1184	1038*
Elimination of the feature	198	199
Feature under aggregation	Not applicable	2034** 199***
Changes in the shape of the feature	27333	11932

* new features bigger than 0.01 ha.

** features have to be aggregated in a smaller scale (distance less than 5 m)

*** features have to be eliminated from aggregated features in a smaller scale (distance more than 5 m)

6. Conclusions

The update of polygon features identifying land cover territories has to be carried out using larger scale generalization, which would include only significant changes rather than all features from a reference spatial data set. Changes in specific features must be evaluated according to their influence on the update of the data of a smaller scale. The accepted significant changes are those "seen" in a spatial data set of a smaller scale, i.e. having "appropriate" quality and quantity attributes for a smaller scale and meeting minimum requirements for a geometric attribute. Depending on the type of feature changes, feature update may vary in a spatial data set of a smaller scale. Besides, the type of a change affects different analysis of its significance. Therefore, the evaluation of changes in features must be performed consequently according to the type of changes in the feature and follow the procedure: updated attributes, new features, deleted features, features that have to be aggregated, features having an updated shape.

References

AGENT Consortium 1999. Selection of Basic Algorithms, version 2.3. Technical report. Department of Geography, University of Zurich, Zurich. Available from Internet: http://agent.ign.fr/deliverable/DD2.pdf

Bayers, E. 2010. Updating topographic reference data (processes, methods, needs, research): progress report in the Belgian NGI-IGN, in *ISPRS Archives,* vol. 38, part 4-5-2/W9: 185-188. Available from Internet: http://www.isprs.org/proceedings/XXXVIII/4_8_2-W9/papers/final_190_NGI_Belgium_Paper_Eric_BAYERS_EN.pdf

Beconytė, G.; Kryžanauskas, A.; Papšienė, L.; Papšys, K.; Stankevičius, Ž. 2009. Lietuvos geografinės informacijos infrastruktūra – kelias į bendrą geografijos metodologiją, *Geografija* [Geography] 45(1): 1–10. ISSN 1392-1096.

Cecconi, A. 2003. *Integration of Cartographic Generalization and Multi-Scale Database for Enhanced Web Mapping*: Ph.D. Dissertation. Zurich University, Switzerland. 155 p. Available from Internet: http://www.geo.uzh.ch/fileadmin/files/content/abteilungen/gis/research/phd_theses/thesis_AlessandroCecconi_2003.pdf

Grünreich, D. 1985. Computer assisted generalisation, in *Papers CERCO-Cartography Course*. Frankfurt am Main, Institut für Angewandte Geodäsie.

ICA – International Cartographic Association. 1973. *Multilingual Dictionary of Technical Terms in Cartography.* Wiesbaden, F. Steiner. 573 p.

INSPIRE – Infrastructure for Spatial Information in Europe. 2009. INSPIRE Generic Conceptual Model, version 3.2. Available from Internet: http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/D2.5_v3.2.pdf

Kazemi, F. 2003. A generalization framework to derive multiscale GEODATA, in *Proc. of the Spatial Science Conference*. September, 2003. Canberra, Australia, 1–12.

Kilpelainen, T. 2000. Maintenance of multiple representation databases for topographic data, *The Cartographic Journal* 37(2): 101–107.

Li, Z. L. 2006. *Algorithmic Foundation of Multi-scale Spatial Representation*. CRC Press (Taylor & Francis Group), Bacon Raton. 310 p. ISBN 9781420008432.

McHaffie, P. H. 2002. Towards the automated map factory: Early automation at the U.S. Geological Survey, *Cartography and Geographic Information Science* 29(3): 193–206. Available from Internet: http://www.geography.wisc.edu/histcart/v6initiative/07mchaffie.pdf

McMaster, R.; Shea, S. 1992. *Generalization in Digital Cartography*. Association of American Geographers, Washington, USA. 134 p.

Papšienė, L.; Papšys, K. 2011. Possibilities of updating small-scale basic spatial data in Lithuania using generalization methods, *Geodesy and Cartography* 37(4): 143–148. ISSN 2029-6991.

Richard, J. R.; Srinivas, A.; Omar, A. K.; Badrinath, R. 2005. Image change detection algorithms: a systematic survey, *IEEE Transactions on Image Processing* 14(3): 294–307. http://dx.doi.org/10.1109/TIP.2004.838698

Singh, A. 1989. Digital change detection techniques using remotely-sensed data, *International Journal of Remote Sensing* 10(6): 989–1003.
http://dx.doi.org/10.1080/01431168908903939

Stankevičius, Ž. 2008. Feasibility study of united national cartographical model, in *7th International Conference "Environmental Engineering": Selected papers*, vol. 3. Vilnius: Technika, 1483–1487.

Stankevičius, Ž.; Beconyte, G.; Kalantaitė, A. 2010. Automation of update of digital national geo-reference databases, *Technological and Economic Development of Economy* 16(2): 254–265.

Weibel, R.; Dutton, G. 1999. Generalising Spatial Data and Dealing with Multiple Representations, in *Geographic Information Systems: Principles and Technical Issues*, vol. 1. Ed. by Longley, P. A.; Goodchild, M. F.; Maguire, D. J.; Rhind, D. W. New York: Wiley, 125–155.

Lina PAPŠIENĖ. PhD student at the Department of Geodesy and Cadastre, Vilnius Gediminas Technical University, Saulėtekio al. 11, LT-10223 Vilnius, Lithuania; Ph +370 5 274 4703, Fax +370 5 274 4705, e-mail: *gkk@vgtu.lt*. A graduate from Vilnius Gediminas Technical University (MSc in measurement engineering / geodesy and cartography, 2000), actively participates in the processes of developing spatial information infrastructure and producing basic (reference) spatial data at the scales of 1:10 000, 1:50 000, 1:250 000 in Lithuania. Research interests: generalization and harmonization of spatial data, SDI.

Kęstutis PAPŠYS. PhD student at the Centre of Cartography, Vilnius University, M. K. Čiurlionio g. 21, LT-03101 Vilnius, Lithuania; Ph +370 5 272 4741, Fax +370 5 373 7723, e-mail: *kestutis.papsys@gf.stud.vu.lt*. A graduate from Vilnius University (MSc in geography, 1999). Actively participates in the process of developing spatial information infrastructure in Lithuania. Research interests: GIS methods for predicting natural, social and ecological hazards, modelling with GIS, SDI.