

Large Scaled Topographic Mapping and Issues in Depicting VGI and Open Data

Benedikt Hajek ^{a,*}, Karel Kriz ^b

^a University of Vienna – benedikt.hajek@univie.ac.at

^b University of Vienna – karel.kriz@univie.ac.at

* Corresponding author

Abstract:

Open data and geospatial data collected by volunteers are nowadays easy to obtain and available with worldwide coverage through projects like OpenStreetMap. However, the use of these datasets leads to new challenges in depiction especially in the field of large scale topographic cartography. In addition to quality research, new processing and depiction methods for integrating these data are emerging. In the course of this work, specific problems of maps based on OpenStreetMap and Open Data elevation models are pointed out and possible solutions are introduced. In addition, a method for the preprocessing of contour lines is presented and the process flow is described in more detail. The goal of this work is to give insight into a toolbox of specially adapted and (semi-)automated methods. In this way, the quality standard of the depiction of topographic maps based on Open Data is to be increased, but also limitations are being shown.

Keywords: open data, VGI, OpenStreetMap, Digital Elevation Model, topographic mapping, contour lines

1. Introduction

In the course of time, topographic cartography has undergone changes on the technical level as well as on the basis of its representation methodology. Today's geodata collection methods, which produce increasingly accurate data with the support of aircraft and satellites, are the cornerstone of many modern maps. But nowadays geodata are not only collected by professionals. Thanks to the crowd-sourcing method of Volunteered Geographic Information (VGI) (Goodchild 2007), large amounts of geodata can be collected in a short time. Probably the best-known and most successful example of such data aggregation is the Open Geodata Service OpenStreetMap. With the help of this freely available software, voluntary users can record geodata anywhere in the world, add it to the database, and also retrieve it anytime and anywhere. In the process, a shift in geodata collection from specialists to laymen and hobby cartographers can be observed (Heipke 2010). However, this shift leads to uncertainty in data quality. While official organizations or private specialists follow precise quality-preserving criteria for geodata acquisition, it is difficult to monitor the data in services like OpenStreetMap.

For topographic cartography, elevation models are also of great importance for the creation of contour lines and hill shading. Open Data elevation models are available in a variety of ways. For example, the SRTM elevation model

provided by the USGS is one of the frequently used models, since it can be used very variably with (almost) worldwide coverage and about 30m ground resolution. But also here, due to the data quality, there are often display problems which lead to a reduction of the quality especially for large scale topographic maps.

Despite the problems, the attractiveness of data bases such as OSM and SRTM for use in topographic cartography is high. The reason for this is the worldwide coverage and their free availability. Especially in regions of the world that are not covered by qualitative and up-to-date map material, such as offered by official institutions in Europe, there is a demand for topographic maps for outdoor tourism (hiking, mountain biking, ski touring).

In the course of this article, problem areas for large-scale topographic cartography when using VGI and Open Geodata will be pointed out, approaches to solutions will be presented and limitations due to the data bases will be identified. It is of high priority to capture geodata processing steps, which can be transferred into automated procedures. This work also focuses on an approach to improve contour line representations based on freely available elevation models, which is presented and explained in more detail.

2. Quality of OpenStreetMap

Research on the quality of data in the OSM geodatabase has thus been increasingly conducted (cf. Arsanjani et al 2015). In particular, quality deficiencies due to high data heterogeneity can be identified (see Touya et al 2013). These can be observed not only due to different data density in different regions (urban-rural divergence) (cf. Goodchild et al 2012), but also in the metadata of the objects entered in the database.

The degree of completeness compared among the situation objects in the OSM database also shows significant differences. For example, according to a study from 2017, the transport network is the front-runner and is more than 80% complete worldwide (Barrington-Leigh et al 2017). Elements such as urban areas, on the other hand, are sparse, especially in developing countries, and new methods for collecting this data are needed. (see Grippa et al 2018).

Specifically, the semantic resolution (Ruas and Bianchin 2002) of the data plays an important role in its representation. This denotes the level of detail in the description of the data. The more numerous and precise the attribution of the geodata, the higher the semantic resolution. In the case of OSM data, problems arise in the classification of objects. For example, due to low semantic resolution, river objects cannot be distinguished hierarchically. Also, semantic heterogeneity can occur where objects of one class are labeled with different tags (e.g. "wood" and "forest") (Vandecasteele and Devillers 2015). This leads to problems with automated depiction of the OSM data, as in the case of OpenTopoMap.

3. OpenTopoMap – OSM and SRTM Data

The OpenTopoMap project serves as a showcase for the use of Open Geodata in the creation of topographic maps (see OpenTopoMap 2021). OpenTopoMap is a free, topographic map with global coverage that is generated automatically from data in the OSM database and SRTM elevation data. While the automated generation of topographic maps of the entire surface of the earth based on Open Data is a remarkable achievement, significant shortcomings in quality, particularly with regard to high mountainous areas, can be monitored.

3.1 Classification of water lines

Due to the previously mentioned low semantic resolution of the OSM data, there may be problems with the generalization for different scales when displaying the water lines. Depending on the accuracy of the attribution of the river objects by the users, the classification of the water lines is often not possible. The result is the overload of the map image especially in smaller scales/zoom levels since lower-ranking water lines cannot be distinguished from higher-ranking ones. Figure 1 shows a section of the OpenTopoMap from Austria east

of Linz at the zoom level 11. In the west of the section a dense network of water bodies can be seen, while in the east only a few more prominent streams are shown. The western area, however, does not have more water bodies in reality, but non-perennial water bodies were tagged with the same tag "stream" as perennial water bodies. A generalization based on the OSM classification is therefore not possible.

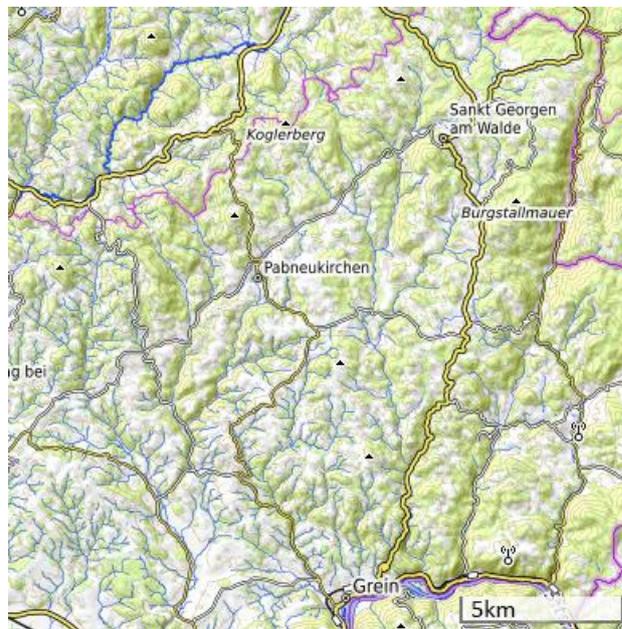


Figure 1. Poor stream classification, OpenTopoMap (2021)
Data: © OpenStreetMap contributors, SRTM
Cartography: © OpenTopoMap (CC-BY-SA)

An approach for automated reclassification of OSM water lines was shown using a channel network method (Hajek et al 2019). Here, the existing OSM water lines are assigned a hierarchy ID that is obtained from a computed channel network based on an elevation model. Due to the similarity of the location of the OSM data and the possible calculated water course, the lines can be matched.

3.2 Water lines and contour lines

The generation of contour lines from a wide variety of elevation models can lead to offset errors when merged with water networks such as the data from OSM. At smaller scales, these display errors are not immediately apparent, but web maps such as OpenTopoMap allow for high zoom levels and thus viewing at large scales. It shows that the terrain described by the contour lines does not match the water courses. In the severest cases, situations arise where it appears that the water is flowing uphill (see Fig. 2).



Figure 2. Stream and contour line misalignment, OpenTopoMap (2021) Data: © OpenStreetMap contributors, SRTM Cartography: © OpenTopoMap (CC-BY-SA)

A method for automated adjustment of the elevation model to the water courses was presented by Samsonov (2020). A major advantage of this novel method is that existing terrain shapes are reshaped and adapted to the watercourse situation instead of creating new terrain shapes.

3.3 Land cover

Land cover plays an important role in topographic cartography, since the map reader can establish a rough but quick relation to reality by its indication in the map. Especially the representation of forest areas and urban areas therefore serve as a means of orientation. The further differentiation of, for example, mountain pines and forest areas can serve to improve the orientation basis, especially in mountainous regions.

The survey of forest, mountain pines, urban areas and also rocky areas is of great importance for large-scale topographic mapping. Land cover data from the OSM database also suffer from heterogeneous quality states like its other data. Therefore, its use is not always sufficient for large scale depictions. Other Open Data projects such as CORINE Landcover, a project of the European Environment Agency (EEA) in the framework of EU Copernicus, also has Europe-wide coverage but insufficient ground resolution (100m/250m) to be used in maps with scales of 1:75.000 and larger.

Techniques for completion therefore include land cover classification using high-resolution satellite imagery, preferably with bands in different spectral ranges to automatically delineate forest and mountain pines, for example.

Grippa et al (2018) take a combination of OSM traffic data, remote sensing data, and spatial metrics with their

approach to mapping road blocks. In test areas, this method was able to delineate urban areas with up to 84% accuracy.

A current and promising option for depicting land cover is the project *ESRI 2020 Land Cover* where areas are derived from Sentinel-2 satellite data. A deep-learning model generates the different classifications from 6 bands of Sentinel-2 surface reflectance data. The model processes multiple images throughout the year and the results are then assembled into a final map for 2020. The dataset includes 10 classes, including Trees, Grass, and Built Areas with a Cell Size of 10m (ESRI 2021). However, the previously mentioned distinction between forest and mountain pines cannot be made here either.

3.4 Contour lines in flat terrain

In modern cartography, contour lines are calculated from elevation models. While in the meantime standard tools in every GIS can output contour lines in every interval distance, there are still depiction issues in certain types of terrain. Due to inaccuracies of the elevation models in flat regions, there are often unsteady line progressions, which do not describe the terrain situation correctly. An example of this specific problem can be seen in Figure 3.

To correct the erroneous lines in flat terrain, smoothing the elevation model is a simple but effective approach. However, this introduces the problem that steep terrain shapes are also smoothed and thus important information about the typical nature of the relief is lost (compare Fig 3 with Fig 4).

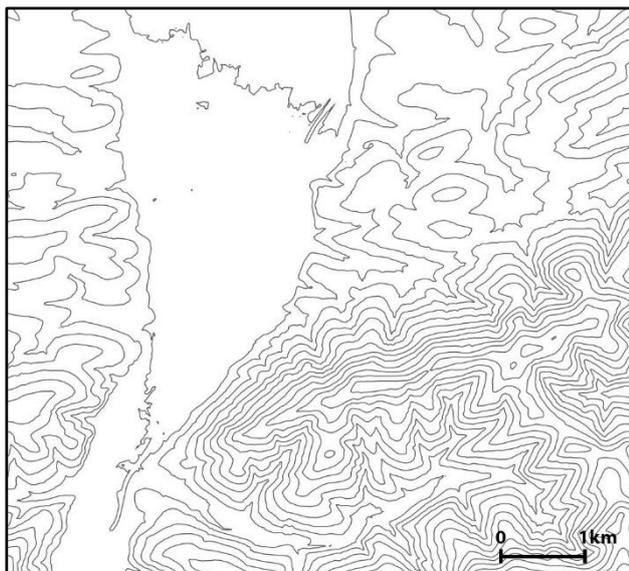


Figure 3. Contour lines from raw DEM, depiction issues in flat area. (Data: DGMÖ 2021)

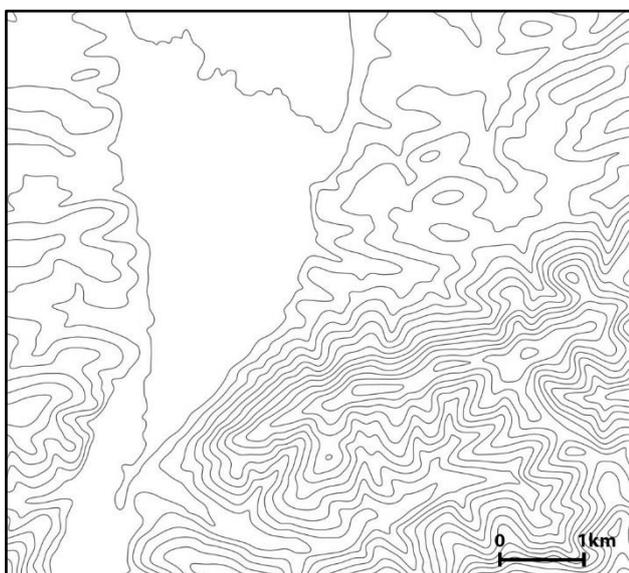


Figure 4. Contour lines from smoothed DEM, relief information is lost in steep areas. (Data: DGMÖ 2021)

To avoid this problem, different smoothing parameters have to be applied to the elevation model depending on the relief shape. In the course of this work an automated method was developed which leads to the desired results. In the following, the procedure of this method will be described in more detail.

3.4.1 Existing Research

Jaara et al (2011) distinguish mountainous and flat regions using an algorithm based on terrain height differences. A new elevation model is then computed using an elastic grid.

Kettunen et al (2017) arrived at similar results as in the method of this paper using their method "adaptive DEM smoothing". The Topographic Position Index (TPI) is used to distinguish between flat and mountainous terrain. The

TPI characterizes locations in the DEM in terms of variation from surrounding elevations (Weiss A. D. 2001).

The difference to the method presented here is the use of the slope factor to delineate steep and flat relief. In addition, to ensure smooth transitions between the differently smoothed areas, a transition raster consisting of gradient values between the two areas is used as a connector.

4. Slope Dependent DEM Smoothing

The method of Slope Dependent DEM smoothing aims exclusively at preparing the elevation model in a pre-processing phase for the generation of optimized contour lines. The process has been automated using the GDAL library. This allows the quick modification of some crucial parameters which have a great impact on the result. The method can be applied to elevation models of different ground resolutions without any complications. In the course of the work, the Open Data Digital Elevation Models AustriaDGM (10m) (DGMÖ 2021) and the SRTM (approx. 30+m) (SRTM 2021) were used. In the following, the individual steps are presented and explained in an overview (see also Fig. 5 and 6).

- 1) The original DEM is smoothed to create the smoothed DEM.
- 2) A slope raster is calculated from the original DEM and used to create masks.
- 3) The masks for the smoothed DEM ($0^\circ - 2^\circ$ slope), the original DEM ($2^\circ - 90^\circ$) and the transition DEM ($1^\circ - 3^\circ$) are created from the slope raster.
- 4) The slope factor raster is calculated from the slope raster with the following formula.

$$\frac{\text{slope value} - \text{min. transition value (1)}}{\text{max. transition value (3)} - \text{min. transition value (1)}}$$

- 5) The transition DEM is calculated from the original DEM and the smoothed DEM based on the slope factor raster with the following formula.

$$\text{smoothed DEM} * (1 - \text{slope factor}) + \text{original DEM} * \text{slope factor}$$

- 6) The 3 masks are applied to the 3 associated DEMs and the new rasters are added together. The new elevation model is created from which optimized contour lines can be generated (Fig. 7).

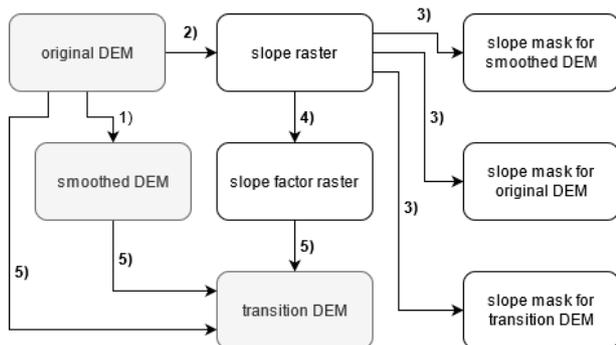


Figure 5. Slope Dependent DEM Smoothing process diagram: steps 1 - 5

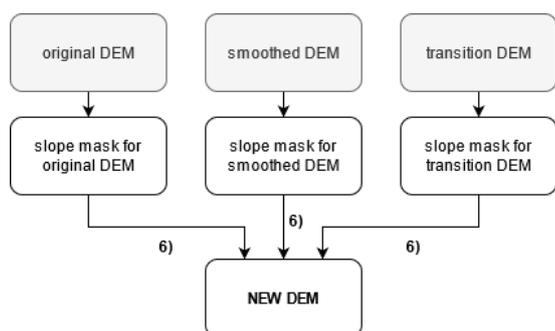


Figure 6. Slope Dependent DEM Smoothing process diagram: step 6

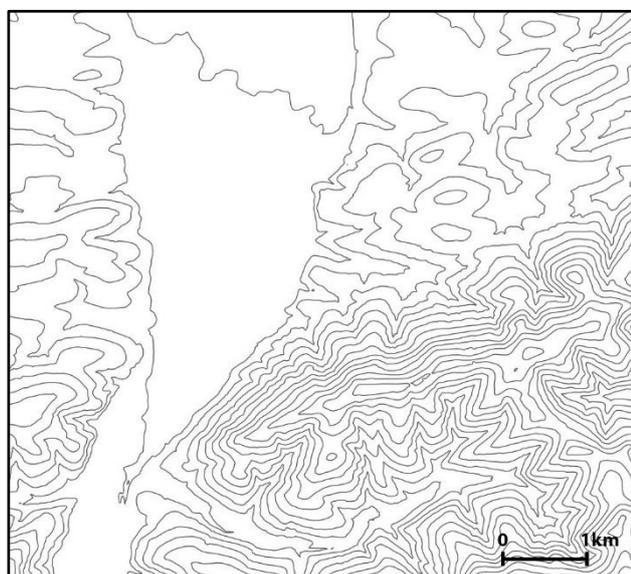


Figure 7. Contour lines from slope dependent smoothed DEM, flat areas are smoothed, relief information is kept for steep areas. (Data: DGMÖ 2021)

4.1 Transitional DEM Merging

Merging elevation models of different quality can lead to display problems at the "seam" of the DEMs. In some cases, this results in the formation of edges in the subsequent contour line representation (Fig. 8).

Part of the applied method of Slope Dependent DEM Smoothing can also be used for this case. Here, a transition factor raster is generated, which is multiplied by the DEMs to be intersected. At the point of contact of the DEMs, a smooth transition is thus generated, which leaves no visible trace in the calculated contour lines (Fig. 9).

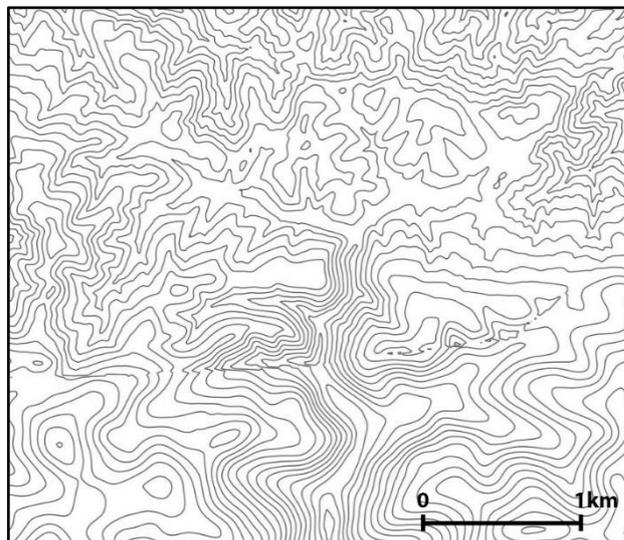


Figure 8. Merge of Austrian DEM (DGMÖ) and SRTM, visible edge error at the seam of both DEMs.

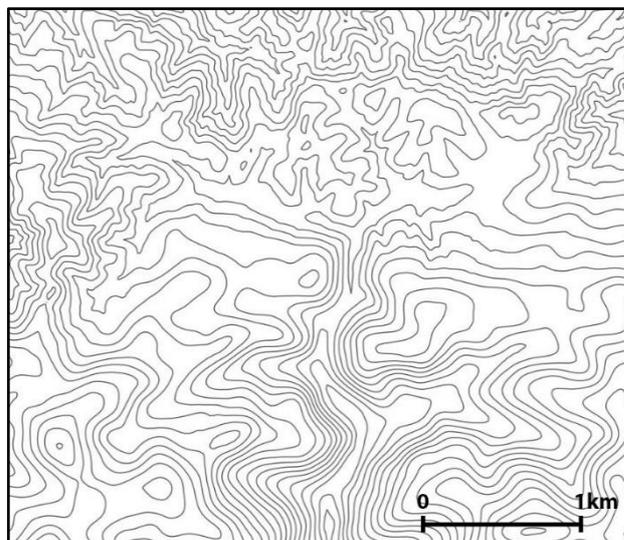


Figure 9. Transitional merge of Austrian DEM (DGMÖ) and SRTM along Austrian border, seamless transition.

5. Conclusion

The project of this work is to increase the display quality of topographic maps based on Open and VGI data. The special quality state of these data requires increased cartographic attentiveness in map production.

It turns out that due to their heterogeneity, VGI data must be processed with special methods to optimize them for automated representations such as OpenTopoMap. Automated generalization of VGI data often proves

problematic. Non-uniform meta-information of the geodata, which is added by the users themselves, often does not allow standardized generalization methods. Since the quality of these freely available geodata is usually not sufficient to produce large-scale topographic maps with high accuracy, new methods for improved representation must be devised, and limitations must also be discussed. Also the use of elevation models for the calculation of contour lines leads in the example of OpenTopoMap to representation problems. Methods like the adjustment of the contour lines and the water network or the completion and correction of the land cover play an important role to increase the quality of large scale topographic maps. The methods for Slope Dependent DEM smoothing and Transitional DEM merging presented in this work are intended to serve as part of the toolbox for improving the data basis and depiction quality for maps based on Open Data. A consolidation and collection of all (semi-) automated tools for increasing cartographic quality can bring Open Data cartography closer to the high standards of institutional map production.

6. References

- Barrington-Leigh C., Millard-Ball A. (2017). The world's user-generated road map is more than 80% complete. *PLoS ONE* 2017, 12, e0180698.
- DGMÖ (2021). Digitales Geländemodell Österreich. License: Creative Commons Namensnennung 3.0 Österreich, <https://www.data.gv.at/katalog/dataset/d88a1246-9684-480b-a480-ff63286b35b7> (retrieved: 20.04.2021)
- ESRI (2021). ESRI 2020 Land Cover. <https://www.arcgis.com/home/item.html?id=d6642f8a4f6d4685a24ae2dc0c73d4ac> (retrieved: 10.08.2021)
- Goodchild M. F. (2007). Citizens as Sensors, The World of Volunteered Geography. In: *GeoJournal*, 69, 4, pp. 211-221.
- Goodchild M. F., Li L. (2012). Assuring the quality of volunteered geographic information. In: *Spatial Statistics*, 2012, 1, pp. 110-120.
- Grippa T., Georganos S., Zarougui S, Bognounou P. Diboulo E., Forget Y., Lennert M., Vanhuyse S. Mboga N. Wolff E. (2018). Mapping Urban Land Use at Street Block Level Using OpenStreetMap, Remote Sensing Data, and Spatial Metrics. In: *SPRS International Journal of Geo-Information*, 7(7):246. <https://doi.org/10.3390/ijgi7070246>
- Hajek B., Kriz K. (2019). Topographic Maps Based on OpenStreetMap Data: Optimising Depiction of Water Streams. In: *Proceedings of the 11th ICA Mountain Cartography Workshop*, pp. 29-36.
- Heipke C. (2010). Crowdsourcing geospatial data. In: *ISPRS Journal of Photogrammetry and Remote Sensing*, 2010, 65, pp. 550-557.
- Jaara, K., & Lecordix, F. (2011). Extraction of cartographic contour lines using digital terrain model (DTM). *The Cartographic Journal*, 48(2), pp. 131–137.
- Jokar Arsanjani J., Mooney P., Zipf A., Schauss A. (2015). Quality Assessment of Contributed Land Use Information from OpenStreetMap Versus Authoritative Datasets. In: *OpenStreetMap in GIScience - Experiences, Research and Applications*. Springer International Publishing Switzerland, pp. 37-58.
- Kettunen P., Koski C., Oksanen J. (2017). A design of contour generation for topographic maps with adaptive DEM smoothing. In: *International Journal of Cartography*, 3(1), pp. 19-30
- OpenTopoMap. (2021). Topographische Karten aus OpenStreetMap. <https://opentopomap.org/>
- Samsonov T. E. (2020). Automated Conflation of Digital Elevation Model with Reference Hydrographic Lines. In: *ISPRS International Journal of Geo-Information*. 2020; 9(5):334. <https://doi.org/10.3390/ijgi9050334>
- SRTM (2021). U.S. Geological Survey, Shuttle Radar Topography Mission. Digital Elevation Model. License: Public Domain, <https://earthexplorer.usgs.gov/> (retrieved: 20.04.2021)
- Touya G., Brando-Escobar C. (2013): Detecting Level-of-Detail Inconsistencies in Volunteered Geographic Information Data Sets. In: *Cartographica, The International Journal for Geographic Information and Geovisualization*, 48, 2, pp. 134-143.
- Weiss, A. D. (2001). Topographic position and landforms analysis. Poster presented at the Esri User Conference, San Diego, CA, USA. http://www.jennessent.com/downloads/tpi-poster-tnc_18x22.pdf