THE ANALYSIS OF DATA QUALITY ON TOPOGRAPHIC MAPS
ANALIZA KAKOVOSTI PODATKOV NA TOPOGRAFSKIH KARTAH

Miro Govedarica, Mirko Borisov

UDK: 528.93/.94
ABSTRACT

The data quality of topographic maps involves a degree of deviations of geometric and thematic content from their actual position and status in nature, respectively. For determining the geometric accuracy, comparative way of testing is usually applied, namely the comparison of measured values of topographic maps with "true" or "conditionally true" values (numeric data of positional and elevation network or data of a map in a larger scale, where the content errors of individual elements could be neglected in comparison due to differences in the scale of mapping). This method gives a direct accuracy assessment of printed maps and reproduction originals, as opposed to the experimental method which considers and determines errors in certain phases in the process of making topographic maps. Geometric, and to a large extent thematic accuracy of topographic maps are not always reliable enough and depend on the way of surveying and the type of contents, i.e. from geodetic control points, which are more accurate, to vegetation whose positional accuracy is the worst.

KEY WORDS

cartography, topographic map, classic map, digital map, data quality

IZVLEČEK

Podatki na topografski karti se geometrično in tematsko deloma razlikujejo od dejanskega stanja v naravi. Za določitev geometrične natančnosti se uporablja primerjava podatkov na karti s »pravimi« in »pogojno pravimi« vrednostmi na natančnejšem viru. Pri tem se običajno uporabljajo numerični podatki položajnih in višinskih mrež ali podatki kart večjega merila, na katerih je mogoče primerjalno zanemariti napake posameznih objektov glede na merilo. Takšna metoda omogoča neposredno oceno natančnosti tiskanih kart in reprodukcijskih originalov v primerjavi z eksperimentalno metodo, pri kateri se obravnavajo in določajo napake posameznih faz v procesu izdelave topografskih kart. Geometrična in v veliki meri tematska natančnost topografskih kart nista vedno zadovoljiva, kar je odvisno od načina izmere in tipa vsebine, na primer od geodetskih točk, ki so natančnejše, do podatkov vegetacije, katere položajna natančnost je najslabša.

KLJUČNE BESDE

kartografija, topografska karta, klasična karta, digitalna karta, kakovost podatkov

1 INTRODUCTION

Military Geographical Institute (MGI) created a system of topographic maps whose main content is based on topographic and photogrammetric land survey conducted from 1947 to 1967, and the periodic cycles of revision and amendments after that time period. On this basis, MGI obtained classic maps at the following scales: 1:25.000, 1:50.000, 1:100.000, 1:200.000,
1:300,000, 1:500,000 and 1:1,000,000. All the earlier editions of maps which were made before and after the Second World War and which were periodically updated and used for a long time are being replaced with aforementioned topographic maps. Basic features of the new maps are based on the same geographical and mathematical basis. They were created using the system of derived maps and represented in the conformal (Gauss-Kruger or Lambert) cartographic projection, with the Greenwich meridian as default. The same or similar cartographic-editorial decisions are applied. In other words, all maps are created in a single geodetic and cartographic system, which considerably facilitates their application.

This article focuses on topographic maps at the scales 1:25,000 (TM25), 1:50,000 (TM50) and 1:100,000 (TM100). The main goal was to represent their quality in the classic (older) and digital (newer) form. One of the main goals of MGI is ongoing activity on the conversion from analog form of topographic maps to digital form considering various aspects of digitization of cartographic sources. Significant effort is made to come up with the entire range of scales of maps published by MGI and to create an integrated topographic and cartographic information system.

2 ANALYSIS OF DATA QUALITY ON CLASSIC MAPS

The data quality of classic topographic maps includes both geometric accuracy and thematic fidelity of cartographic content. For the analysis of data quality, classic topographic maps at the scales TM 25, TM 50 and TM 100 have been considered. Thematic integrity and data fidelity were accomplished in the examination and evaluation of other content, apropos the previous assessment of accuracy of maps (Mihajlović, 2003). Considering that, it can be stated that the input data at the time were very reliable. However, the general quality of information has declined over time, so it is necessary to periodically update the content.

2.1 The positional accuracy

The positional accuracy of graphical data in classic topographic maps of larger scale is given in Table 1. The table shows that the ratio of geometric accuracy of classic topographic maps is almost proportional to the reduction of sheet area, apropos the increasing of the surface of signatures and width of the lines, which represents that the reduction in geometric accuracy of classic topographic maps is, on a certain level, the consequence of cartographic - reproduction processes. Considering the successive methodology and technology in creating of drafting-publishing originals, it can be assumed that the mean square error of point objects is in the range from ± 3,7 m to ± 43 m, at the map scale.

Therefore, the mean square position error of geodetic control points for sheets of scales TM25, TM50 and TM100 ranges from ± 3,7 m to ± 29 m. This confirms that examined sheets of topographic maps have the correct absolute position within the used projection and there is no significant increase of, i.e. reduction, rotation, translation or shearing of the mapping surface. The reference value which would prove this is, for example, graphic accuracy of 0,2 mm at TM25 is 5 m on the ground, where 3,7 m is below 5 m. Also, the graphic accuracy of 0.3 mm at TM50 and 0.35 mm at TM100 (Buder, 1978). Positional accuracy of other points of topographic content is
worse and is given specifically for the detailed points which are printed in black colour, as well as the detailed points which are printed in green colour. Only clearly defined detailed points were taken into account during testing. The mean square error of detailed points which are printed in black ranges from ± 10,1 m to ± 40 m. Also, the mean square error of detailed points which are printed in green colour ranges from ± 14,6 m to ± 43 m, in the scale of mapping (Barut, 1974).

Since hydrography is represented in a more detailed and more accurate way compared to the other content in topographic maps, and other elements of geographic content with transition to smaller scale moves in accordance with hydrography, it could be said that the mean square error of hydrography in the mentioned classic topographic maps ranges in within ± 10,1 m to ± 43 m, at the appropriate map scale. In other words, the first value is an indication of the accuracy of detailed points printed in black and the other value of the positional accuracy of the detailed points printed in green colour. The accuracy of hydrography is between these two values (Barut, 1974). The lowest accuracy of the detailed points, which were printed in green is expected since it is most dependent content of topographic maps because it accords with all other data.

### 2.2 The height accuracy

The height accuracy of classic topographic maps refers to precision which represents different points of land forms in relation to the selected zero height reference surface. There should be the difference distinguished between direct certain height accuracy and precision of height contour line points. Heights of the points are determined during the survey, and depend on surveying method, i.e. photogrammetry where heights of the points depend on the accuracy of control points and altitude of aircraft. According to (Buder, 1978) it is:

\[
\sigma_h = \pm 0.2 \% \times H_0 \text{ (m)}
\]  

(2.1)

where \( H_0 \) is the relative altitude of aircraft.

The height accuracy of classic topographic maps also depends on the precision of the representation of relief, i.e. of the accuracy of contour lines as the main way of the relief's
representation and of the accuracy of numerically marked elevation points (levels). The mean square errors of positions of contour lines, which determine the height accuracy, are commonly used as an elevation indicator of the accuracy of classic topographic maps, as well as the mean square errors of height points in relation to the nearest geodetic control points (Table 2).

<table>
<thead>
<tr>
<th>The accuracy criterion</th>
<th>TM25</th>
<th>TM50</th>
<th>TM100</th>
</tr>
</thead>
<tbody>
<tr>
<td>The height accuracy of detailed points</td>
<td>$\sigma_h = \pm 2.0$</td>
<td>$\sigma_h = \pm 2.6$</td>
<td>$\sigma_h = \pm 3.2$</td>
</tr>
<tr>
<td>The height accuracy of detailed points by Koppe</td>
<td>$\sigma_h = \pm (1.4 + 4.1 \cdot \tan \alpha)$</td>
<td>$\sigma_h = \pm (1.3 + 14 \cdot \tan \alpha)$</td>
<td>$\sigma_h = \pm (1.1 + 3 \cdot \tan \alpha)$</td>
</tr>
</tbody>
</table>

Table 2: The comparison of height accuracy of TM25, TM50, and TM100 (Barut, 1974)

Formulas and many expressions used as indicators of the quality of the relief represented by contour lines, do not give a precisely and easily understandable picture of deviation of contour lines of their actual position. The accuracy of elevation points of contour lines is determined by many complex procedures. One of the commonly used equations is defined by Koppe:

$$\sigma_h = \pm (A + B \cdot \alpha)$$

where

- $\alpha$ is the slope of terrain and,
- A and B are coefficients which are determined depending on type of the cartographic material.

For decades there has been a specific problem of testing the accuracy of contour lines, and a general solution has not yet been found (Buder, 1978). The essence of the problem is in a desire to mutually compare two different surfaces that have very complicated structure and are undefined in the continual numerical sense. The first one is the part of the physical surface of the Earth, it has many details and is designed to scale down the observed map (approximately), and the other surface has been given as conditional representation of the third dimension using contour lines. In other words, contour lines may be incorrect in different ways in relation to the actual "natural" curve. More specifically, this case is about the height discrepancy on a map of displayed area which is framed by given contour line of the actual terrain surface height given by the contour lines form in nature.

2.3 Thematic accuracy

Thematic accuracy of classic maps includes content and time determinations, namely data fidelity of topographic maps. It is impossible to find mathematical expression for thematic accuracy. Only the number of errors in relation to the total number of data can be determined, such as names or signs in a single sheet of a topographic map. Obsolete, outdated, and thus the defective or incorrect content decreases the quality and usefulness of topographic map, since
they do not show exact reality as it is, but the condition as it was. Therefore, there are not set tasks related to mapping only, but scientific and practical researches in order to modernize and improve the content of topographic maps, namely their renewal.

When it comes to showing the land forms, by their nature that element of content of topographic maps should be the least changeable. However, in the relief area there is a variety of changes due to human activities (reclamation, mining, etc.) and there is also the influence of natural forces (landslides, erosion, etc.). In addition, substantial changes are visible in a view of the relief by applying specific methods of surveying and cartography (topographic and photogrammetric surveys, base height, generalization of field types, etc.). Also, construction of power plants, accumulation of lakes and a network of channels cause modifications of confluences and river flows.

There are very noticeable changes in land cover and there are perhaps the biggest changes close to populated places. Populated places and infrastructure have been recently subjected to rapid changes. Entirely new settlements arise, and the old ones change their old look and all those changes influence the fidelity of the topographic content. Major changes were made in the transport network, settlements and geographical names, too. Thematic accuracy of classic topographic maps, especially in the content, is the highest on the maps made by direct topographic and photogrammetric surveying, because that is connection between surveyor-cartographer compiler from its original situation and real state in the field. Moreover, in derived classic topographic maps (in this particular case TM50 and TM100) there is presented degradation of thematic accuracy and that is the direct consequence of generalization and omission errors.

When it comes to other types of maps obtained from the cartographic sources, quality is even worse because there are both, a question of accuracy of thematic content and the update of cartographic sources.

3 ANALYSIS OF DATA QUALITY ON DIGITAL MAPS

When discussing the quality of data obtained by digitizing the existing topographic maps, one should, above all, bear in mind that all data have already been processed. Specifically, these data were already reviewed which includes removal of irregular details on the digitized content as a result of surface deformation and random errors. On the other hand, the processing that has already been made using these data carries certain risks. Since these data have already been processed, it is practically impossible to remove all the errors from data which were made during that modelling, because there is usually no information which could be used for reconstructing the original data. Errors which are contained in data obtained by digitalization of the existing maps contain both errors of the original data collection and processing (both surveying and mapping errors) and also transformation errors (scanning, georeferencing, vectorization). The limiting factor is certainly the original data collection errors.

Data quality of digital topographic maps depends on the graphics (cartographics) sources, numerical sources (catalog of geodetic points), the process of analog-digital conversion, and the methods of data modelling and processing (system of visualization, technical possibilities and
human factor). However, the general quality depends on the positional and height accuracy of topographic content, and of the updating of data.

After scanning and vectorizing TK25, TK50 and TK100 (paper maps and reproduction originals) in digital form, an additional activity was a control and modelling of the data. That control was used for checking reference data (geodetic points) and implementing modifications and updates, according to available reference data sources. Correction and processing was performed also for geometric and thematic data.

3.1 The positional accuracy

Applied methods of data conversion from analog to digital format (scanning, interactive digitizing, automatic and semi-automatic vectorization) caused various errors. Their values do not exceed the maximum error values of source materials in analog form. Therefore, the accuracy of digital data which were obtained after processing and conversion from analog form can be described by the following types of errors:

\[
m_{iz} = \sigma ;
\]

\[
m_{proc} \leq m_{iz} ;
\]

\[
m_{max}^2 = m_{iz}^2 + m_{proc}^2 = 2m_{iz}^2 ;
\]

\[
m_{max} = \sqrt{2m_{iz}^2} ;
\]

where:

- \( m_{iz} \) - maximum error of source material;
- \( m_{proc} \) - maximum error of processing (conversion and data processing);
- \( m_{max} \) - maximum error of digital data and
- \( \sigma \) - standard (Gaussian mean square) error of data source.

The accuracy of the source maps (reproduction originals) that were used as input data is defined by the Gaussian mean square error and maximum permissible error (double value of the mean square error). The expression for the Gaussian mean square error is:

\[
\sigma = \sqrt{\frac{1}{n} \sum \Delta_i^2} ;
\]

where:

- \( \sigma \) - standard (Gaussian mean square) error of data source;
- \( \Delta \) - deviation (difference) of real values;
- \( n \) - number of discrepancies in the sum and
- \( i \) - positive integer variable.
Based on the expressions 3.1 and 3.2, the mean square errors were calculated and the quality assessment of digital data was shown. The accuracy of the mathematical base (cartographic projection, referent-marked points and geodetic points) of digital topographic maps is given by the mean square error in Table 3.

<table>
<thead>
<tr>
<th>Elements of the mathematical base</th>
<th>The mean square error at map scale (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cartographic projection (lines)</td>
<td>± 0.10</td>
</tr>
<tr>
<td>Reference (marked) points</td>
<td>± 0.10</td>
</tr>
<tr>
<td>Geodetic points</td>
<td>± 0.18</td>
</tr>
</tbody>
</table>

*Table 3: The mean square errors of the mathematical bases of TM25, TM50 and TM100*

The accuracy of position of other objects on digital topographic maps is determined compared to the nearest control (geodetic) point. The mean square errors of well-identified objects (cross-road, railways, religious objects, powerlines, telecommunication and other facilities) are given in Table 4.

<table>
<thead>
<tr>
<th>Objects represent in digital topographic maps</th>
<th>The mean square error at map scale (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Permanent, well-identified objects (facilities, utilities, religious objects)</td>
<td>± 0.35</td>
</tr>
<tr>
<td>Poor identified objects (features and characteristics of settlements and soil)</td>
<td>± 0.55</td>
</tr>
</tbody>
</table>

*Table 4: The mean square errors of important objects of TM25, TM50 and TM100*

In areas with forests and/or mountains and sandy hills, objects can be 1.5 times of less accuracy, with maximum of 0.75 mm. Where the symbols of topographic maps were not relocated, the absolute value of horizontal (positional) errors of strictly defined lines that refer to the nearest control point or the network does not exceed 0.5 mm. When it comes to vegetation and contours which are marked, the maximum positional error reaches up to 1 mm at the scale of map.

Positional accuracy of the other elements of digital topographic maps is not controlled. Since the number of surveyed geodetic points and important objects was of high density, and the content source accorded and carefully positioned in relation to surveying and mapping network, there is a probability for good positional accuracy and the remaining contents of the digital topographic maps.
Also, the process of scanning and georeferencing of reproduction originals was carried out, as were also the settlements, roads and contour lines at the scale of TM25. After the overlay of independently scanned and georeferenced sources in the working environment of software package DigiScan (and some other software packages) very good agreement of a content was noticeable (Figure 1). This confirms the quality of error elimination using collocation based on all points of the map grid and agrees with the values shown in Table 5.

Deformation errors of map sheet and scanning to a large extent can be minimized, but only under certain conditions. In fact, to detect surface deformation which occurs during the time, it is necessary that map points be plotted with known reference coordinates. This is especially important in cases of strains of a local character, where it is necessary that there is a greater amount of points with relatively uniform and dense distribution over the useful area of a map. On the basis of known theoretical coordinates of these points and their positions on a map, it is possible to determine the deformation of the map sources. In the most favorable cases, those
points represent the vertices of the map grid or graticule. Scanning errors can be avoided by using high-quality scanners and software, in the same way as in the case of identification and elimination of map media deformation. The optimum case is the usage of high quality flatbed scanner of large format. An additional problem with digitization of the existing hardcopy (cartographic sources) is the questionable updatedness with real situation, since these map sources represent the situation several years old, or even several decades old (Borisov, 2001).

3.2 The height accuracy

In digital cartography the height system is usually represented in the form of digital elevation model (DEM). It could be said that the accuracy of digital elevation representation depends directly on the sources which are used for measuring the plane coordinates, on the density of elevation points (i.e. grid size) and their geometric positions on the slope of terrain.

The quality of DEM which is obtained by digitizing contour lines from maps includes all sources of errors both in the process of its conversion into digital form and generating of DEM by interpolation. Specifically, the process of the DEM production consists of digitalization (scanning and vectorization) of contour lines, adding heights to contour lines, transforming vector lines into cloud of points and generating of DEM model. The overall mean elevation error, as an indicator of accuracy of the elevation model, is calculated using the formula:

$$\sigma_{HP} = \sqrt{\sum \sigma_i^2}$$  \hspace{1cm} (3-3)

where:

- $\sigma_{HP}$ - error level obtained by digitizing the contours from the existing reproduction original of contour lines,
- $\sigma_i$ - error as a result of data processing at a certain stage of works.

The most significant sources of errors that can affect the quality of DEM, and that should be considered in the calculation of accuracy are:

a. primary data collection of elevation model (photogrammetric restitution):
   - orientation errors of photogrammetric model;
   - mapping errors (lowering of marks of photogrammetric instrument and drawing contour lines);
   - errors of cartographic processing;

b. errors after mapping and during the process of secondary data collection, namely data conversion in digital form:
   - map media deformation errors;
   - scanning errors;
   - georeferencing errors of a scanned map;
   - digitization errors (vectorization of contour lines and other lines and of individual points).
The accuracy of DEM is defined by the positional accuracy of individual points that are directly measured and their quantity with respect to quality of the sample that represents the terrain surface. From topologic aspect, there are two basic forms which can represent the digital elevation model - the regular grid and the triangular irregular network (TIN). The precision of the grid model is determined by the dimensions of the cell. High density of the grid with points of high accuracy produce accurately approximated surface. But TIN model is in the form of triangulated irregularly scattered points, dependent on the degree of a surface slope and curvature. For the increased accuracy of surface approximation, the TIN model requires dense points and the inclusion of characteristic break lines.

For digital models generated from the classic cartographic material, it is usually assumed that the elevation accuracy is from 1/4 to 1/5 of equidistance of represented contour lines (Frančula, 2000). Furthermore, based on experience and empirical values achieved for the value of the graphical scribing accuracy is usually set to 0.2 mm. In Table 6, the empirical accuracies of the DEM created by digitization of cartographic material are presented.

<table>
<thead>
<tr>
<th>Scale</th>
<th>Equidistance (m)</th>
<th>Height accuracy (m)</th>
<th>Graphical scribing accuracy (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: 25,000</td>
<td>10</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>1: 50,000</td>
<td>20</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>1: 100,000</td>
<td>20</td>
<td>5</td>
<td>15</td>
</tr>
</tbody>
</table>

Table 6: The quality of DEM obtained by digitizing TM25, TM50 and TM100

To control the quality of the obtained models (the previously explained method for both procedures) coordinates of the points from the Catalog of state points of trigonometric network (Peterca, 1987), which are located in that area, were used. So, there were visual and exact methods for assessing the data quality of DEM. For evaluating the accuracy the equation for standard error is used:

$$\sigma_H = \sqrt{\frac{\sum (H^*_i - H_i)}{n}}; \quad (3.3)$$

where:

- $H^*_i$ - values of interpolation elevations;
- $H_i$ - values of measured (control) elevations and
- $n$ - number of control points.

An overview of accuracy assessment of DEM obtained based on information gathered by photogrammetric method, and by creating various types of DEM and different interpolation is given in Table 7.
Table 7: The quality of DEM obtained on the basis of photogrammetric mapping of aerophotos

The accuracy of DEM obtained based on digitized contour lines from the cartographic substrates (reproduction originals) are given in Table 8. In this case, beside of grid and TIN, the grid model is created by two methods of interpolation and visualizing elevation surface. One of them is inverse distance weighted (IDW) and the other method is spline. Also, they are presented in different resolution.

Table 8: The quality of DEM obtained based on digitized contour lines of reproduction originals

Figure 2: Example of matching of digitized content of the elevation model (red) and raster map TM25
The limits of contour lines errors are given in Table 9, by positional mean square error in meters. Deformations can be 1.5 times higher in the covered field.

<table>
<thead>
<tr>
<th>Flat terrain</th>
<th>Hilly terrain</th>
<th>Mountainous terrain</th>
</tr>
</thead>
<tbody>
<tr>
<td>$0^9 - 2^9$</td>
<td>$2^9 - 6^9$</td>
<td>$&gt; 6^9$</td>
</tr>
<tr>
<td>$\pm 4 \text{ m}$</td>
<td>$\pm 5 \text{ m}$</td>
<td>$&gt; \pm 5 \text{ m}$, (depends on the slope)</td>
</tr>
</tbody>
</table>

Table 9: Limits of contour lines errors, types of terrain and the slope

In the mountainous regions the number of contour lines is adjusted to the height difference between the top and the foot of the mountain. Types of terrain and slope are given also in the same table.

4 CONCLUSION

Emergence of new technologies for collecting, displaying and analyzing geospatial data is also affected in our case to review and upgrade so far achieved results in the field of digital cartography and GIS. Specifically, the purpose of that is to create topographic-cartographic database. From many aspects, choosing how to collect data of space in database may well be important decision.

This paper analyzed the data quality on topographic maps, especially in digital form. Analyzing of data quality on maps includes geometric accuracy and thematic fidelity of the content being mapped. The geometric accuracy of digital data is according to the data derived from classic maps, and they have to meet requirements generally formulated by the national standards. Positional errors of well-identified objects do not exceed the values on the classic map. But positional errors of features forming the graphic skeleton of the digital map in some cases are over than on the classic map. It can be explained so that the data are digitized from different source materials, and they are burdened with the additional errors of georeferencing and digitizing. Also, the height accuracy of the content depends on the applying methods of interpolation and visualizing elevation surface. Generally, values of the positional accuracy and the DEM of topographic maps are according to the indicators and they do not exceed the maximum error values of source material in classic form.

Specifically, the data quality on topographic maps primarily depends on the chosen scale and method of production (direct topographic and photogrammetric surveys or using generalization of the content). Moreover, in derived classic topographic maps (in this case TM50 and TM100) there is presented degradation of geometric and thematic accuracy, and that is the direct consequence of generalization and updatedness. The state topographic map at the scale 1:25.000 is a good base for creating a national infrastructure of topographical data. In that case, the primary sources of topographic-cartographic database are reproduction originals of TK25. Supplementary sources are the copies of the other topographic maps and the statistical data served as textual and numerical source material.

At the end it is very important to say that the concept of digitalization topographic maps is
more useful and practical especially in the case when topographic data is available. Also, it is a cheaper solution than new surveying and gathering data. The existing topographical data can be updated by, e.g., GPS, and finally cartographic presentation can be carried out.

**Literature and reference:**


DMA (2000). The Digital Geographic Information Exchange Standard, USA.


**Received for publication: 14 June 2011**

**Accepted: 16 October 2011**

**Associate Professor Miro Govedarica, Ph.D.**

University of Novi Sad, Faculty of Technical Sciences

Trg Dositeja Obradovića 6, 21 000 Novi Sad, SERBIA

E-mail: miro@uns.ac.rs

**Assistant Professor Mirko Borisov, Ph.D.**

Military Geographical Institute

Mije Kovačevića 5, 11 000 Beograd, SERBIA

E-mail: mborisov@EUnet.rs