Vector data model of the Müller's map of Bohemia

J. Cajthaml

Dept. of Mapping and Cartography, Czech Technical University in Prague - Faculty of Civil Engineering, Thákurova 7, 166 29 Prague 6, Czech Republic
jiri.cajthaml@fsv.cvut.cz

Abstract

The Müller's map of Bohemia is one of the most important maps in the Czech history. In last couple of years this map was scanned and analysed a few times. Spatial position of the map symbols is usually compared with current situation. Using this method we can explore changes of the landscape through the ages. The crucial problem of this approach is georeferencing raster image into some well defined coordinate system. The best way how to handle that problem is using ground control points (GCPs) and appropriate type of spatial transformation. The problem of choosing the best set of GCPs is not solved correctly yet. Usually there are used well identifiable points depending on the researcher's meaning. Proposed method is based on creating full vector data model of the map. Having complete vector database of the map, we can test many combinations of GCPs and many types of spatial transformation. Another great advantage of vector database is the fact, that vector data are easier to be analysed (e.g. spatial overlays, proximity analyses, or spatial statistics) in GIS software. As we wanted to make new thorough analysis of the Müller's map of Bohemia, we decided to create its full vector data model. Methods of creating the model are explained in this article. At the end the map was georeferenced and several spatial statistics were done.

Keywords: old maps, Müller's map of Bohemia, georeferencing, vector data model, geodatabase, vectorization

Müller's map of Bohemia

The Müller's map of Bohemia is an extraordinary masterpiece of Czech cartography. J. Ch. Müller, an author of the map, was an important Austrian military engineer and above all an exceptional cartographer (Kretschmer et al., 2004). The map originated in the years 1712 till 1718 as the result of the first systematic topographic mapping of the Czech lands. Firstly it was set out in 1723, unfortunately after Müller's death. The title of the map is „Mappa geographica regni Bohemiae". The scale of the whole map is approx. 1:132,000. The size of one map section is 473 by 557 mm and the entire size is 2,403 by 2,822 mm (5 x 5 map sheets) (Kuchař, 1959).

Classical way of research

The Müller's map of Bohemia was several times an object of cartometric analyses (Krejčí and Cajthaml, 2009). These analyses stand on classical way of research on old maps. Usually on the digital image of the map well identifiable points are collected. These points can serve as ground control points (GCPs) for georeferencing the map into some well defined coordinate system. Using this method we can also determine the scale of the map. The crucial problem of collecting GCPs is in their selection. As we don't have any information about precision of the map we don't know if selected GCPs are identical (with corresponding coordinates in current system and without evident error). Using classical way of georeferencing (any transformation with Least Squares Method of adjustment) these evident errors are hard to be distinguished. Thus we need to have the most complex set of identical points. Within this set we can test many subsets of GCPs, we can also make some statistical tests (e.g. detection of outlying measurements, Least Trimmed Squares method) and choose the best combination of GCPs for georeferencing the map.

Motivation for vector data model

As written above if we want to analyse the map it is necessary to have the most complex set of GCPs. If we take a look at old maps (especially early maps) they contain many point symbols instead of lines or areas. These point symbols are ideal candidates for GCPs. In addition the line crossings or area corners can be used. Therefore it is very interesting to create the whole vector data model of the map and use as many GCPs as possible. The vector data model is created in the local coordinate system of the map.

Second motivation factor for creating vector data model is the fact that the vector data are very easy to handle with. If the model is well designed according to the most reliable standards (usually in spatial database) many kind of spatial analyses can be performed on the data. We can explore the spatial characteristics of the data as it is done with usual GIS data. For example, we can compute the population density, average distance of settlements etc. These results can tell us much about the historical landscape.

Of course, if we have the full vector data model, we can much easier analyse the statistics of the data that hold interesting information. We can count the number of specific map features (villages, churches, castles), we can measure the length of line features (road, rivers) or we can find out the overall area of specific area features (water bodies, forests). In environmental research it is very interesting to monitor the change of landscape through the ages. Once we have vector databases based on a few maps (from different eras) we can compare and measure these changes. Using GIS overlay analysis we are able to distinguish changes of water bodies, changes of forest areas or changes in settlements.

As a result of presented motivation factors I see many advantages of creating the full vector data model of old maps. If we want to make the thorough analysis of the map, creating the vector database should be the first step. After that we can use the real functionality of GIS software. These findings led me to create the vector database of the largest old map in the Czech history - the Müller's map of Bohemia.

Data preparation

Before designing the vector data model it is necessary to prepare the source data. As the original prints of the Müller's map of Bohemia are stored in several archives we wanted to get the best data. In our previous work we used the data published by the Institute of History of Academy of Sciences of the Czech Republic. These prints of the map sheets had been quartered and a cloth had been used for mounting (Cajthaml and Krejčí, 2007). Despite these problems we analysed the map and made many conclusions. In 2008 we obtained better data from the Central Archives of Surveying, Mapping and Cadastre in Prague. This data has the same image quality as previous, but the map sheets had not been quartered. Both data collections were scanned in true color depth and with 300 dpi resolution.

At the beginning it was necessary to re-project the data into its original size. Due to shrinkage of the paper the data is distorted. We discovered the original size of one map section as 473 by 557 mm (Kuchař, 1959). All 25 map sheets of the Müller's map of Bohemia were transformed into these dimensions using four corner points. As we wanted to discover the degree of shrinkage of the map the affine transformation with Least Squares Method (LSM) adjustment was used. The standard deviation for these 25 transformations varies between 0.002 and 0.509 mm. Three map sheets (number 16, 20, and 24) were transformed only by 3 corner points (fourth corner point was not identified) and thus these map sheets have no standard deviation for affine transformation. The majority of values of the standard deviation are up to 0.25 mm. Only map sheets number 5 and 25 exceeded this value. These sheets are situated in the corners of the whole map and contain only small parts of the map (more than half of these map sheets is covered by decorations).

After transformation the images into correct dimensions it was necessary to design the geodatabase for storing the vector data. First of all the legends of the map were studied thoroughly. The author used many symbols for different types of municipalities, settlements, important institutions, and places of extraction and processing of natural resources. Legends are in Latin and in German also (Krejčí and Cajthaml, 2009). Over 40 point symbols were identified in the legends (see Fig. 1.), 2 more symbols were identified in the map (missing in the legends).

Fig. 1. Part of the legends of the Müller's map of Bohemia

Creating the map sheets geodatabases

On the Müller's map of Bohemia we identified 6 feature categories: settlements and important places (point features), towns (polygon features), river network (line features), water bodies (polygon features), roads (line features), and boundaries (line features). These 6 categories were designed in the geodatabase as feature classes. ArcGIS as the most wide-spread GIS software was used. As we have 25 independent map sheets we started with creating 25 geodatabases for each map sheet separately. For some future analyses it is very important to have separate databases of map sheets for their comparison. At the end of vectorizing we merged (adjacent lines and polygons were united) all these databases together (still in the local paper coordinate system) to make overall statistics and analysis.

Feature class "point symbols" is the most represented in the map. In (Kuchař, 1959) there is mentioned the number 12,495 of point symbols on the map. Vectorizing of this feature class is quite simple. Crucial is to determine the position of the point of reference. The attribute table of this class contains: historic name of the place on the map, current name of the place (in the Czech language) and the feature type (one of more than 40 types).

Some towns were represented by the author as polygons (town within city walls). Due to that fact, a special feature class for these towns was designed. The attribute table copies the previous class. Rivers and streams are depicted very decoratively. It was very difficult to vectorize very sinuous lines of this feature class. Important fact during creating this layer was to keep the topology of the data. In fact the whole layer is represented by a network. Water bodies (especially ponds) were vectorized as polygons. Attribute tables of both layers contain only current names of the features (due to missing labels of hydrologic features in the map). On the Müller's map of Bohemia only a few roads are depicted (with double line symbol). The region boundaries are displayed as dotted lines, the boundary of the Bohemia is displayed with dashed line. Part of the geodatabase is depicted (Fig. 2.).

Fig. 2. Geodatabase of the map

Merging the geodatabases

After having all 25 vectorized databases the point was to merge all databases together. Because every geodatabase was vectorized in the map sheet coordinate system (the origin in the lower left corner) it was necessary to transform the data to the right position in the merged map. Of course identity transformation was used. After transforming the data into right position (in fact only spatial shift) the edges of map sheets were studied. Many map features were divided by the map sheet edges and we had to re-establish their topology. But there appeared following problems.

We supposed that the edges of map sheet are perfectly straight. Unfortunately maps are during their archiving distorted and they are usually somehow bended. Our approach is based only on linear transformations. If we want to make the edges fit together more precisely we should use another method of merging map sheets, for example constrained polynomial fit (Molnár, 2010).

Of course, there are some other problems such as evident errors made by the author of the map. One of these errors had been published earlier (Krejčí et al., 2009). The town Kladno is on the Müller's map of Bohemia depicted twice (on two map sheets). This fact was proven and other edges were inspected to find similar errors – none was found. The second type of the author's incorrectness was the topological inconsistency. For example, the boundary or road is depicted on the other side of the river on different map sheets (see Fig. 3). The third group of errors consists of missing map features. There are some evident missing roads (not connected network) and missing boundary in the northwest Bohemia (see Fig. 4).

Fig. 3. Topological error on the map

Fig. 4. Missing boundary on the map

Statistical analysis of the map

As we have the complete geodatabase of the Müller's map of Bohemia it is very easy to perform statistical analysis of the map. These numbers are very interesting when confronted with early published data. On the map we found: 15,215 point symbols (11,997 villages, 677 small towns, 44 towns, 233 places of extraction and processing of natural resources, etc.). Compare this number with early published 12,495. In the polygon class we found 133 towns (45 royal towns within this class).

On the map there are 2154 water bodies with area approx. 28,300 mm^2 (490 km^2 in reality). The river network consists of approx. 185,400 mm of rivers and streams (24,400 km in reality). These data will be analysed during following research. Other two feature classes are not so interesting in the statistics field. The road network is not complete and contains only selected roads (25,950 mm on the map). Nevertheless it would be interesting to compare these main roads with today motorway and highway network. The boundaries class is very interesting. It shows the shape of Bohemia (western part of the Czech Republic) almost the same as today. That points out the fact that the boundaries has not changed through almost 300 years.

Georeferencing the map

The main goal of our research is to analyse spatial information on the map. The best way how to make such research is to georeference the map. Then we can compare it with some other maps. Having a few maps in the same coordinate system the full functionality of GIS software can be used (e.g. spatial overlay). We can also perform hydrologic analysis like (Székely, 2009). Fluvial pattern of the rivers could be evaluated, changes of their channel could be revealed.

In our previous work (Krejčí and Cajthaml, 2009) we used 83 GCPs for transformation the map into S-JTSK (Czech national coordinate system). As written above the crucial problem is in the choice of these points. In the current research we used as many points as possible. On the map we found 4,409 points with corresponding points on the current map (database of the Czech statistical office). Current coordinates were added to the attribute table of these points. As these coordinates were added automatically (by joining tables based on the name of the settlement) there occurred many errors. In the Czech Republic there are many villages with the same name and thus the name is not unique (it can't serve as primary key, ID of the table). The errors were eliminated by original method. The map coordinates were transformed approximately only using a few (confident) control points. Then the azimuth and the distance between pairs of GCPs were calculated. After statistical testing, the points with outlying values were excluded. After elimination the set of GCPs consisted of 3906 points.

There are many transformation methods to georeference old maps. Usually affine transformation is used because the shrinkage of the paper could be eliminated and the map image is not badly distorted. The better results can be achieved with polynomial transformations. Unfortunately, the image is distorted and the lines are transformed into the curves. As we used the transformation (to the right dimensions of the map) at the beginning of our research, we eliminated the shrinkage of the paper. Then if we don't want to damage the map image similarity transformation should be the best.

Using all 3906 GCPs we acquired following parameters for similarity transformation: standard error of position 1.74 km and the mean scale of the map 1:131,580. Compare these numbers with previous research (Kuchař, 1959), (Krejčí and Cajthaml, 2007). If we transformed separated map sheets the mean standard error of position for map sheet is 1.48 km. Other transformation methods will be tested.

Conclusions

The aim of this article was to present the full vector data model of the Müller's map of Bohemia. Creating this model was very work-intensive, but led to the data basis prepared for future research. First outputs of this work are statistics of the map. The numbers differ from till now published data. Using this new research we can correct 50 years old results still cited in the literature.

The first step in the map analysis is the georeferencing. For the first time more than 3900 control points were used. The results of similarity transformation of the data led to the standard error of position 1.74 km. There is much space for testing many types of transformation or different groups of GCPs (e.g. only churches, only towns, only towns within the walls) in the future. Such testing can show us if some map features are depicted more precisely than others.

Having the full georeferenced vector data model there is a space for many GIS analyses. The settlements can be analysed, disbanded or flooded villages can be identified. The river network can be compared with some other maps and changes can be detected. The area of water bodies has changed through the ages very dramatically. The area of ponds before 300 years can be compared with the current situation. The road network on the map can be compared with today's motorways and highways. These and many other analyses can be performed on the basis of the vector data model.

Georeferenced data should be also distributed on the internet. Till today old maps are usually distributed only in web applications (Brůna and Křováková, 2006). Georeferenced raster data of the Müller's map of Bohemia will be distributed by WMS service, the vector data model will be available through WFS service. Web map services allow users free and comfortable access to the map.

Acknowledgements: This research has been supported by the GA CR grant No. 205/09/P102.

References:

Brůna V., Křováková K. (2006): Staré mapy v prostředí GIS a Internetu [Old maps in the GIS and Internet environment (in Czech with English summary)]. GEOS 2006 - 1st International Fair of Geodesy, Cartography, Navigation and Geoinformatics - Conference Proceedings. Praha.

Cajthaml J., Krejčí J. (2007): Müllerovy mapy Českých zemí, jejich digitalizace a zpracování. [Müller's maps of the Czech lands, their digitizing and processing. (in Czech with English summary)]. Kartografické listy vol. 15: pages 51-59, Kartografická spoločnosť SR a Geografický ústav SAV, Bratislava.

Krejčí J., Cajthaml J. (2009): Müller's maps of the Czech lands and their analysis. Acta Geodaetica et Geophysica Hungarica, vol. 44: pages 27-38, Akadémiai Kiadó, Budapest.

Krejčí J., Mikšovský M., Zimová R. (2009): Possible relation between the Müller's map of Bohemia and the First Military Survey — A case study of the Kladno region. Acta Geodaetica et Geophysica Hungarica, vol. 44: pages 39-48, Akadémiai Kiadó, Budapest.

Kretschmer I., Dörflinger J., Wawrik F. (2004): Österreichische Kartographie. Wiener Schriften zur Geographie und Kartographie – Band 15. Institute für Geographie und Regionalforschung der Universität Wien, 318 p.

Kuchař K. (1959): Vývoj mapového zobrazení Československé republiky I.- Mapy českých zemí do poloviny 18. století [Evolution of mapping of the Czechoslovak Republic I. – Maps of Bohemia till the middle of the 18th century (in Czech)]. Praha : Ústřední správa geodézie a kartografie.

Molnár G. (2010): Making a georeferenced mosaic of historical map series using constrained polynomial fit. Acta Geodaetica et Geophysica Hungarica, vol. 45: current issue, Akadémiai Kiadó, Budapest.

Székely, B. (2009): Rediscovering the old treasures of cartography — What an almost 500-year-old map can tell to a geoscientist. Acta Geodaetica et Geophysica Hungarica, vol. 44: pages 3-16, Akadémiai Kiadó, Budapest.

Figure captions:

Fig. 1. Part of the legends of the Müller's map of Bohemia
Fig. 2. Geodatabase of the map
Fig. 3. Topological error on the map
Fig. 4. Missing boundary on the map