GEODESY AND CARTOGRAPHY
ISSN 2029-6991 / eISSN 2029-7009
2016 VOLUME 42(1): 25–30
http://dx.doi.org/10.3846/20296991.2015.1160493

UDK 528.9

TESTING THE POSITIONAL ACCURACY OF OPENSTREETMAP DATA FOR MAPPING APPLICATIONS

Khalid L. A. EL-ASHMAWY

Department of Civil Engineering, Al-Matria Faculty of Engineering, Helwan University
Matria, 11718 Cairo, Egypt
E-mail: khalid85_2002@yahoo.com

Received 06 November 2015; accepted 02 February 2016

Abstract. OpenStreetMap, a web mapping platform, is the most popular web map source for use in location-based services with specific emphasis on pedestrian navigation, tourist guide applications, and other location-based search applications. This paper tests the positional accuracy of OpenStreetMap for the mapping applications using the case study in the campus of UMM El-Qura University, Makah, Saudi Arabia. The proposed testing method consists of statistical comparative approach using OpenStreetMap data and accurate land surveying reference data. The results show that OpenStreetMap data has positional accuracy of 1.57 m which is suitable for generating planimetric maps of scale 1:5000 or smaller. The obtained results open the door for using the OpenStreetMap maps for applications such as general preliminary planning where larger areas are covered but only moderate accuracy is needed. Applications include mapping the general layout of potential construction sites, proposed transportation systems, and existing facilities. The proposed methodology in this paper is of great interest to small engineering firms for the generation of local area maps from OpenStreetMap data.

Keywords: mapping, OpenStreetMap, positional accuracy, two dimensional affine transformation.

Introduction

Mapping surveys are made to determine the locations of natural and cultural features on the Earth's surface and to define the configuration (relief) of that surface. Once located, these features can be represented on maps. Natural features normally shown on maps include vegetation, rivers, lakes, oceans, etc. Cultural (artificial) features are the products of people and include roads, railroads, buildings, bridges, canals, boundary lines, etc. The relief of the Earth includes its hills, valleys, plains, and other surface irregularities. Lines and symbols are used to depict features shown on maps. Names and legends are added to identify the different objects shown.

Two different types of maps, planimetric and topographic, are prepared as a result of mapping surveys. The former depicts natural and cultural features in the plan (X-Y) views only. Objects shown are called planimetric features. Topographic maps also include planimetric features, but in addition they show the configuration of the Earth's surface. Both types of maps have many applications. They are used by engineers, planners, geologists, foresters, architects, agriculturists, archaeologists, geographers, and scientists in numerous fields. Maps are used extensively in geographic information system (GIS) applications. Conducting the surveys necessary for preparing maps and the production of the maps from the survey data are the mainstay of many surveying businesses.

Traditionally, maps were prepared using manual drafting methods. Now however, the majority of maps are produced using computers, computer-aided drafting (CAD) software, and data collectors. Currently, some data collectors include drafting software so that field personnel can display their data in the field to check for mistakes and missing elements.

Mapping surveys are conducted by one of two basic methods: aerial (photogrammetric) or ground (field) techniques, but often a combination of both is employed. However, airborne laser mapping systems may also be used. Satellite positioning systems allow the measurement of features or points anywhere in the world, from space. Ground surveys are still commonly used in preparing large-scale maps of smaller areas. Even when photogrammetry or airborne laser mapping is utilized, ground surveys are necessary to establish control and to field-check mapped features for accuracy.

In recent years, when choosing a free-to-access web-based mapping system for navigation, location-based services, or general information there are three dominant services: Google Maps, Bing Maps, and OpenStreetMap. These web mapping platforms are the most popular web map sources for use in location-based services with specific emphasis on pedestrian navigation, tourist guide applications, and other location-based search applications.

OpenStreetMap, as an example of web mapping, is built by a community of mappers that contribute and maintain data about roads, trails, cafés, railway stations, and much more, all over the world. It emphasizes local knowledge. Contributors use aerial imagery, Global Positioning System (GPS) devices, and field surveying to verify that OpenStreetMap is accurate and up to date. The community of OpenStreetMap is diverse, passionate, and growing every day. Its contributors include enthusiast mappers, GIS professionals, humanitarians mapping disaster-affected areas, and many more (OpenStreetMap 2015).

Web mapping is an open source for geospatial data; therefore it is necessary to identify different elements of spatial data quality (Delavar, Devillers 2010) to determine the appropriate usage for such data. These elements include: positional accuracy, attribute accuracy, temporal accuracy, logical consistency and completeness. Recently, with increasing geospatial data on the Internet, several researches have shown an increased interest in assessing the quality of open source data.

Positional accuracy of data may be far more important from other elements of quality to give users the position correctly. The positional accuracy denotes the coordinate deviation of a spatial object compared to its real location (Haklay 2010). The positional accuracy of the collected data is affected by different influences, e.g. the technological bias like the accuracy of the GPS-receiver used, different data acquisition techniques (e.g., digitizing) or subjective knowledge about the data gathering process. In order to assess the usability of the collected data in varying cases of application the positional accuracy of the used data has thus to be thoroughly evaluated, because missing and imprecise data effect model calibrations and in the worst case leads to false conclusions.

This paper focuses on the statistical analysis of the positional accuracy of OpenStreetMap data for a small area in Makkah, Sudia Arabia. This analysis is based on data of precise surveying techniques. Therefore, the surveying data serves as the reference data, to which OpenStreetMap data is relatively evaluated.

1. Site of the study

In this study the study area was chosen located in the campus of Umm El-Qura University, Makah, Sauda Arabia. Umm El-Qura University Campus lies at about 21° 19' 11" latitude and 39° 56' 43" longitude. The area is approximately 6.50 km by 5.25 km. The longer dimension runs roughly in the north direction. The total geographical extent of the area is about 1650 hectares. The study area contains four GPS ground control points, transportation systems of roads, buildings, trees, lamps post, natural terrain and other details.

2. Preparation of the reference data

The methodology for surveying and mapping the study area for getting the reference data consisted of traversing, detailed surveying, downloading data and mapping.

2.1. Traversing

Traversing was done by using Topcon GTS710 Total Station (Topcon 2015) and the available four GPS ground control points. The total station has the possibility to measure points up to 2400 meters, a large amount of memory to record all the data from the field and software to allow the surveyors to download the recorded data to a computer.

Traversing work consisted of outer and inner traversing. Outer traversing covered the major area. Inner traversing was divided into blocks. The reason for subdividing the traverse into blocks is to minimize the errors in traversing and to start the survey work from any two visible station. After traversing the errors were computed and distributed to all the stations for getting the adjusted ground coordinates of the traverse stations. These coordinates were input to the total station for detailed surveying.

2.2. Detailed surveying

Location sketch of each and every details of the field such as buildings, roads, platforms etc., was drawn. These detailed sketches were helpful during the plotting work. Total Station was used for high-density grid coordinate based survey work. Before commencing a survey it was required to have station and back station coordinates and their exact location. The instrument was kept over the station point and the station and back station coordinates were fed in the total station. The orientation was taken with respect to the station and back station coordinates. The results of orientation were studied carefully and should be within the tolerance (0.01 m). All the remaining points including minor bends in buildings, roads and platforms were identified, coded and the prism kept exactly over the point. The same procedure was continued until the end.

2.3. Downloading and mapping

After completing the surveying work the data was downloaded to the computer using data transfer cable and interface software. Then with the help of mapping software such as LISCAD, the map was generated in the computer. The generated map was exported to the AUTOCAD software for the creation of separate layers for buildings, roads, platforms, trees, departments etc.

3. Getting the corrected OpenStreetMap data

The proposed method for getting the corrected OpenStreetMap data comprises the following:

- Downloading data from OpenStreetMap.
- Two Dimensional affine transformation.
- Developing software for correcting the OpenStreetMap data.

3.1. Downloading data from OpenStreetMap

It is possible to obtain access to vector data for OpenStreetMaps. This facility is not available for all open source for geospatial data such as Google Maps or Bing Maps. Consequently a quantitative coordinate-based comparison between OpenStreetMap and the surveying reference data is possible.

The following approach was taken for getting the OpenStreetMap in vector format. OpenStreetMap data for the case study was downloaded in OpenStreetMap XML format. Using FME Quick Translator tool (Safe Software 2015), this OpenStreetMap file was converted to DWG file for use in AutoCAD. The projected coordinate system was defined as Universal Transverse Mercator (UTM) same as for the reference data. Using the obtained DWG file, the creation of separate layers for buildings, roads and platforms can be easily done.

3.2. Two dimensional affine transformation

The reduction of OpenStreetMap coordinates to reference data system involves a coordinate transformation in a plane. A more common approach is to use affine transformation which has the characteristic that true shape is retained after transformation. This transformation is a three-step process that involves (Ghosh 2005):

- Scaling to create equal dimensions in the two coordinate systems.
- Rotation to make the reference axes of the two systems parallel.
- Translations to create a common origin for the two coordinate systems.

The transformation function has one rotation angle (α), two scale factors in x and y direction (S_x and S_y), and two origin shifts (a_1, a_2). The equation is:

$$\left.\begin{aligned} x' &= a_1 + S_x \cos\alpha \cdot x + S_y \sin\alpha \cdot y \\ y' &= a_2 + S_x \sin\alpha \cdot x + S_y \cos\alpha \cdot y \end{aligned}\right\} . \tag{1}$$

Equation (1) can be simplified as:

$$\left.\begin{aligned} x' &= a_1 + b_1 x + c_1 y \\ y' &= a_2 + b_2 x + c_2 y \end{aligned}\right\} , \tag{2}$$

where x', y' are the points coordinates in the reference data system;

x, y are the points coordinates in OpenStreetMap data system; and

$a_1, \ldots, c_2$ are certain transformation coefficients, to be determined.

In order to obtain the transformation coefficients, the coordinates in reference and OpenStreetMap data systems for at least three points should be available. More than three points would enable a least squares solution (El-Ashmawy 1999).

After obtaining the coefficient of transformation, values of α, S_x and S_y can be derived as:

$$\alpha = \tan^{-1} \cdot \frac{b_2}{b_1}, \; S_x = \frac{b_1}{\cos\alpha} \text{ and } S_y = \frac{c_1}{\sin\alpha} . \tag{3}$$

It is very important to understand the effect of these coefficients (Friedman, Kohler 2003). A positive value of a_1 indicates a west-to-east shift and a negative value indicates an east-to-west shift. Similarly, positive values of a_2 are in accordance with a south-to-north shift and vice versa.

The two scale factors in x and y direction (S_x and S_y) cause the magnitude of contraction or expansion. When the scale factor value is >1 indicates expansion and when the scale factor value <1 means contraction relative to the reference data.

The positive value for the rotation angle (α) means a counter clockwise rotation and negative α a clockwise one.

3.3. Developing software for correcting the OpenStreetMap data

Based on the above mentioned method of coordinates transformation, a software was developed to facilitate the production of the corrected OpenStreetMap map. The functions of the developed software are:

- Reading the x and y coordinates of the well distributed points in the reference data and the corresponding points in OpenStreetMap map.
- Computing the six coefficients of affine coordinates transformation.
- Reading the OpenStreetMap in DXF-ASCII file format and computing the transformed, corrected, *x* and *y* coordinates of each point on the map.
- Generating a new DXF file with the same layers for the corrected OpenStreetMap map.

The software is menu based software. It has been implemented using Visual C++ Compiler V6.0 (Gregory 1998; Malik 2010) and designed to be flexible and portable to 32-bit Windows platforms (Windows XP, Windows 8).

4. Testing the accuracy of the corrected OpenStreetMap map

The DWG files for the reference and corrected OpenStreetMap data can be easily overlaid. In order to assess the positional accuracy, a comparison for the location of well-defined points such as road intersection, roads centre lines, building corners, etc. in OpenStreetMap layer and reference data layer were selected. An AutoLISP (Autodesk 2012) program was developed for automatic storing, in a TXT file, the x and y coordinates of the corresponding points in the two layers. The number of the selected points is 2057 points.

The positional difference in *x* and *y* coordinates for each point was computed and the Root Mean Square Error (RMSE) was computed using the well known formula (Ghosh 2005):

$$RMSE_X = \sqrt{\sum_{i=1}^{n} (X_t - X_r)_i^2 / n} \; ;$$

$$RMSE_Y = \sqrt{\sum_{i=1}^{n} (Y_t - Y_r)_i^2 / n} \; , \qquad (4)$$

where: $RMSE_X$, $RMSE_Y$ are RMSE for X and Y coordinates;

X_r, Y_r are the coordinates of the reference data point,

X_t, Y_t are the coordinates of the tested point, *n* is the number of points.

5. Results

The nature of the geometric distortion of the OpenStreetMap data is studied using the transformation coefficients as shown in Table 1. The computed coefficients indicate how the OpenStreetMap data must be transformed to get the reference data.

Table 1. Estimated coefficients of affine coordinate transformation

Coefficient	a_1	a_2	α	S_x	S_y
Value	–65.35	2.73	–0.016	1.0001	1.0002

OpenStreetMap map is shifted east-to-west as well as south-to-north. Furthermore, there are expansion in *x* and *y* directions. α refers to a clockwise rotation of the OpenStreetMap.

To check the validation of the developed methodology, values of RMSE, Mean Absolute Error (MAE) and Mean Error at the tested points were computed before and after applying the affine coordinates transformation. The obtained values are tabulated in Table 2.

Table 2. Values of RMSE, MAE and Mean Error before and after applying the developed method

Method	Before applying affine transformation (values in meters)	After applying affine transformation (values in meters)
Minimum Absolute Error	0.92	0.53
Maximum Absolute Error	14.36	8.65
Mean Error	6.89	3.35
RMSE	4.82	1.57

From Table 2, the following conclusion can be obtained:

- Using affine transformation decreases the value of mean error by about 3.2 meter.
- The results show an 67% improvement in positional accuracy (RMSE).

Moreover, the obtained RMSE values in Table 2 were compared with the permissible limits according to the specifications of ASPRS (American Society for Photogrammetry and Remote Sensing) (ASPRS 1993) as tabulated in Table 3. Taking into account only the specifications for the highest accuracy (Class I Maps), it can be seen that the planimetric accuracy is suitable for generating maps of scale 1:5000 or smaller rather than 1:20000 or smaller without using the developed method.

The obtained results open the door for using the OpenStreetMap maps for applications such as general preliminary planning where larger areas are covered but only moderate accuracy is needed. Applications include mapping the general layout of potential construction sites, proposed transportation systems, and existing facilities.

Conclusion

This research emphasizes the determination and improvement of the positional accuracy of OpenStreetMap data using the case study in the campus of UMM El-Qura University, Makah, Suidia Arabia.

The positional accuracy of OpenStreetMap for the mapping applications was tested. The proposed testing method consists of statistical comparative approach using OpenStreetMap data and accurate land surveying reference data. Additionally, the comparison is based on the RMSE values determined from the differences between the OpenStreetMap data and the reference data.

Table 3. ASPRS planimetric coordinate accuracy requirement for well-defined points (Class I maps)

Planimetric (X or Y) Accuracy (limiting RMSE in Meters)	Typical Map Scale
0.0125	1:50
0.025	1:100
0.050	1:200
0.125	1:500
0.25	1: 1,000
0.50	1:2,000
1.00	1:4,000
1.25	1:5,000
2.50	1:10,000
5.00	1:20,000

Two dimensional affine coordinates transformation was used to reduce the OpenStreetMap coordinates to reference data system. RMSE values before and after applying the affine transformation were determined and analyzed.

The results show that using affine coordinates transformation has significant effects on decreasing the value of mean error by about 3.2 meter and improving the positional accuracy (RMSE) by about 67%.

The obtained accuracies, for the case study, allow the generation of planimetric maps at scale of 1:5000 or smaller using OpenStreetMap data. However, the results obtained in this study are encouraging and show the possibility of using OpenStreetMap maps for applications such as general preliminary planning, and mapping the general layout of potential construction sites, proposed transportation systems, and existing facilities.

The main disadvantages of OpenStreetMap map are:

- The incompleteness of the planimetric features such as buildings, parking areas, etc. comparing with the reference data.
- Lack of information about how the data was captured. As is known, OpenStreetMap data is obtained from a variety of different sources, and the accuracy investigation in one place, may not provide knowledge about other places accuracy. To overcome this disadvantage, the user of OpenStreetMap maps has to test the accuracies of the obtained maps using the proposed methodology in this study.

The results of this study are of great interest to small engineering firms for the generation and testing the accuracies of local area maps.

References

American Society for Photogrammetry and Remote Sensing (ASPRS). 1993. *ASPRS Accuracy Standards for Large-Scale Maps*, USA, 1993.

Autodesk. 2012. *AutoCAD 2013: AutoLISP Developer's Guide*. USA: Autodesk, Inc.

Delavar, M. R.; Devillers, R. 2010. Spatial data quality: from process to decisions, *Transactions in GIS* 14(4): 379–386. http://dx.doi.org/10.1111/j.1467-9671.2010.01224.x

El-Ashmawy, K. 1999. A cost-effective photogrammetric system for engineering applications: Ph.D. Thesis. Department of Civil Engineering, University of Roorkee, Roorkee, India.

Friendman, A.; Kohler, B. 2003. Bidimensional regression: A method for assessing the configural similarity of cognitive maps and other two-dimensional data, *Psychological Methods* 8(4): 468–491.

Ghosh, S. K. 2005. *Fundamentals of computational photogrammetry*. New Delhi: Concept Publishing Company.

Gregory, K. 1998. Special edition using visual C++ 6. USA: Que.

Haklay, M. 2010. How good is volunteered geographical information? A comparative study of OpenStreetMap and ordnance survey datasets, *Environment and Planning B: Planning and Design* 37(4): 682–703.

Malik, D. S. 2010. Data structures using C++. 2nd ed. USA: Cengage Learing, Inc..

OpenStreetMap. 2015. [online], [cited 12 August 2015]. Available from Internet: http://www.openstreetmap.org/

Safe Software. 2015. [online], [cited 7 May 2015]. Available from Internet: https://www.safe.com/fme/fme-desktop/

Topcon. 2015. [online], [cited 23 September 2015]. Available from Internet: http://www.topcon.co.jp/en/index.html

Khalid L. A. EL-ASHMAWY (Dr) is Associate Professor of Surveying & Digital Mapping at the Department of Civil Engineering, Al-Matria Faculty of Engineering, Helwan University, Egypt. Dr Khalid El-Ashmawy developed many surveying softwares which were published in national and international journals and conferences, such as SoftPhotoMap, PHOTOMAP, MATHP, ScanCal, LandSurMap, Calendar, RSA (Road Safety Assessment), and MathDTM softwares.

His area of expertise includes photogrammetry, land surveying, space photography, digital mapping, GIS and development of surveying softwares.