

GEOMETRIC ENHANCEMENT OF THE OPENSTREETMAP ROAD NETWORK

Belhouari, Fatima Z. ^{1 *}; Boukerch, Issam¹; Si youcef, Kamel¹

¹Agence Spatiale Algérienne, Centre des Techniques Spatiales, Arzew, Algeria – fbelhouari@cts.asal.dz

Commission IV

KEY WORDS: OpenStreetMap, geometric correction, Delaunay triangulation, road network, geospatial, geometric accuracy.

ABSTRACT:

OpenStreetMap (OSM) is a collaborative project to create a free and editable map of the world. You can think of OSM as the 'Wikipedia' of cartography. An important geospatial component of this database is the road network quality, which is important for applications such as routing and navigation.

The objective of this work is the geometric enhancement of the OSM road network using a standard national map as a reference. We use two transformation methods, the global transformation and the local transformation (Delaunay triangulation).

This study aims to present a new approach to improve the OSM road network geometrically. To this end, we present a three-step approach based on two techniques that leads to the enhancement of the geometric accuracy of the OSM road network. The first step is the global transformation of the OSM road network. The second step consists of applying the local transformation (Delaunay triangulation) on the OSM road network. In the last step, a comparison between the two methods is examined by calculating the mean and the standard deviation of the checkpoints in order to justify which is the best technique for the geometric enhancement of the OSM road network. We will be particularly interested in the application of this approach in the geometric enhancement / correction where each node of the OSM network will have a newly calculated position. Both approaches have been tested in the region of Oran in Algeria as testing example. The reference data is a city map produced by the National Institute of Cartography and Remote Sensing (INCT) in 2006. The proposed techniques show a clear improvement in geometric accuracy.

1. INTRODUCTION

Volunteered geographic information (VGI), which refers to the provision of geospatial content by ordinary people, is now a well-established practice in GIS and geographic domains. The term VGI was introduced in 2007 by Goodchild (Goodchild, 2007) and grown exponentially, not only in the field of Geomatics. VGI is geographic information freely generated by the public and shared on VGI initiatives via the Internet, where users are not required to have special skills to provide the data.

The OpenStreetMap (OSM) project was initiated in 2004 by Steve Coast in the United Kingdom (University College London), particularly encouraged by the lack of accessible or reusable data of roads, railways, rivers, forests, buildings, etc., from all over the world. It is a community and non-profit project that aims to enable the dissemination of geographic data under a free license as a part of the web 2.0 trend. (S. Ruffray, 2012)

For about a decade now there have been several studies on the geometric quality of OpenStreetMap, where most studies are specifically focused on assessing the quality of the road network.

Among the first studies on the quality of OSM road networks, the one proposed by Haklay (2010) and others (2009) measured the completeness and positional accuracy of London and England OSM road datasets against the baseline data provided by the Ordnance Survey of the British NMA. Girres and Touya (2010) extended this work to France and provided more different statistics (geometric accuracy, attribute accuracy, semantic accuracy, completeness, logical consistency, temporal accuracy, lineage, and usage) on OSM roads through a comparison with data provided by the French IGN. Zielstra and Zipf (2010) investigated the completeness of German OSM roads by comparing them with data from TeleAtlas and Helbich et al. (2012) measured the positional accuracy between OSM and

TomTom data using official survey data as the reference dataset. Other comparisons between OSM and reference road network datasets from commercial providers were proposed by Ludwig et al. (2011) and Neis et al. (2011) who used Navteq and Tom Tom data, respectively. In particular, Ludwig et al. (2011) analyzed OSM road network datasets for populated roads in Germany showing a satisfactory completeness of object/names and accuracy of OSM data for map generation purpose. Kounadi (2011) compared the OSM road dataset in Athens with the data provided by the Hellenic Military Geographical Service (HMGS) highlighting an overall good data quality, especially for length completeness, name accuracy, and positional accuracy. Ciepluch et al. (2011) analyzed the completeness of Ireland OSM roads with respect to the Ordnance Survey Ireland dataset. Jokar Arsanjani et al. (2013) compared the OSM roads of Heidelberg (Germany) with an authoritative dataset from the Federal Agency for Cartography and Geodesy. Graser et al. (2014) proposed a novel toolbox for street network comparison. Nasiri, A. et al. (2018) proposed a new approach to improve the positioning accuracy and completeness of OSM road network based on a Voronoi diagram.

We have tried to use the OSM road network in the study, when we overlaid the OSM road network on a cartographic base, we found some offset of the OSM road network with respect to the map. To solve this problem, we proposed to use the technique of georeferencing the road network by an affine transformation. To test this suggestion, we compared two techniques: Global affine transformation and triangular affine transformation. The question that arises is which is the best technique to improve the road network geometrically compared to the cartographic base.

The first method is the global transformation, i.e. a single parameter of this transformation, which is applied to the entire OSM road network of the geo-referenced road network, and the second technique partitioning the area into triangles and each

triangle had parameters of the affine transformation. This second technique adjusts each segment of the OSM road network with the parameters of the triangle associated with it.

To justify which best technique to apply to adjust the OSM road network, it requires to have the points generated by a topography survey it's better testing, but as we only have a map made by the National Institute of Cartography and Remote Sensing (INCT), this map has been made by an indirect survey (restitution).

2. THE PROPOSED APPROACH

Based on the studies that were done on the geometric accuracy of the OSM data, the most frequent element in the OSM database is the road network.

An approach to enhance the OSM road network has been proposed to carry out this study. A map produced by the National Institute of Cartography and Remote Sensing in Algeria has been used as a reference data.

The first technique is based on georeferencing using the affine transformation (global transformation), and the second technique is the Delaunay triangulation (local transformation), which is used as a technique for partitioning the spatial domain of the network.

After a series of tests, we compared between the results of the two techniques to check which is better to improve the geometry of the OSM road network.

We will be particularly interested in an application to the problem of geometric correction and we will present the work we have done with Matlab.

2.1 Coordinates transformation

The coordinate transformation used for both techniques, global and local (triangular) is an affine transformation.

We have chosen the affine transformation 2D since it corrects six parameters.

The model equation of this transformation is written in the form:

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} C_x \\ C_y \end{pmatrix} = \begin{pmatrix} 1 & \gamma \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \lambda_x & 0 \\ 0 & \lambda_y \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} C_x \\ C_y \end{pmatrix}$$

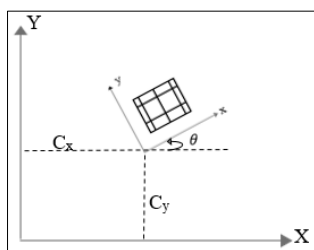


Figure 1. Affine transformation

With: γ , λ_x , λ_y , θ , C_x & C_y are affine transformation parameters,

Where: γ : system orthogonality defect (skewness)
 λ_x : X scale
 λ_y : Y scale.
 θ : rotation angle.

C_x : translation in x.

C_y : translation in y.

2.2 Global transformation

In this application, we mean by global affine transformation a single transformation that transforms the OSM network by a single affine transformation. The parameters of this transformation are calculated using ground control points (GCP's) and the accuracy of these parameters is estimated using independent Check points (ICP's). To do this test we have (X,Y) map (reference) and their counterpart in the OSM system (x,y). To calculate the parameters of the affine transformation we need at minimum three points in coordinates (X,Y) in both systems (reference and OSM).

The least square method is used to solve the system of linear equations derived based on using at least three GCP's.

2.3 Triangular affine transformation

A Delaunay triangulation for a given set P of discrete points in a plane is a triangulation DT (P) such that no point in P is inside the circumcircle of any triangle in DT (P). Delaunay triangulations maximize the minimum angle of all the angles of the triangles; they tend to avoid sliver triangles. (Delaunay, Boris, 1934)

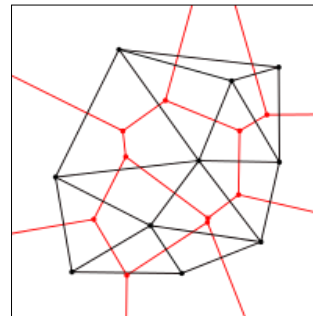


Figure 2. Superposition of a Voronoi diagram (in red) and its Delaunay triangulation (in black).

In this second technique we propose to calculate a triangulation using the GCP's, then we use the three points forming the triangles to calculate an affine transformation, this calculated transformation will be valid to correct points located inside the considered triangle from the target dataset.

3. DATA AND METHODOLOGY

3.1 Data

3.1.1 Reference Data: To study the geometric enhancement of OSM, we use a map that was produced from a photogrammetric survey based on aerial photographs taken at 1:20 000.

Realized in 2006 by the National Institute of Cartography and Remote Sensing (INCT) on a scale of 1/7,500 and covering an area of 72 KM², Spatial reference: WGS 1984_UTM zone 30N.



Figure 3. The city map of Oran (Reference Data)

3.1.2 OSM Data: The OpenStreetMap database has been downloaded from the site (<http://download.geofabrik.de/>), on 12/09/2019.

The OSM database was treated in ArcGis, just keeping the highway and clipping out the road network OSM in the study area.

3.2 Methodology Diagram

The following diagram illustrates the different functions applied to calculate the parameters of the triangular affine transformation (TAT) for the geometric improvement of the OSM road network.

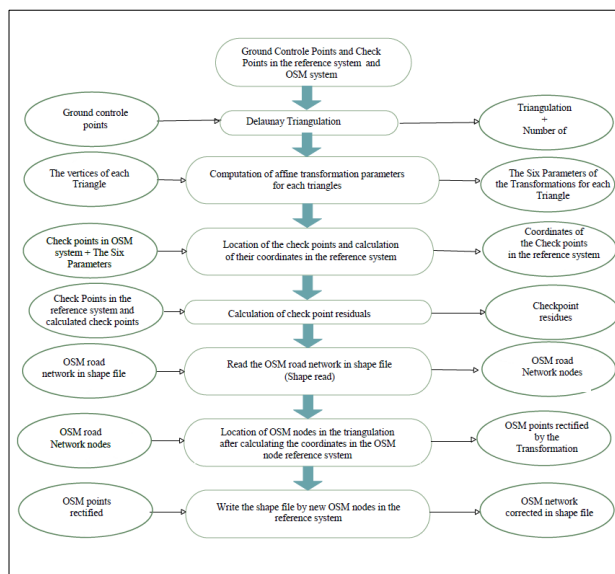


Figure 4. Framework of Triangular Affine Transformation (TAT)

4. RESULTS AND DISCUSSIONS

We programmed the two techniques, the Triangular Affine Transformation (TAT) (Delaunay triangulation) and the Global Transformation (GT) for the geometric enhancement of the OSM road network using a reference map. For both techniques, we used three tests. In each test, a different number and distribution of ground control points were used.

From the reference data, we selected a set of 118 ground control points. Four of these points represent the convex hull to ensure that the other points are distributed within the convex hull. Those four points are always considered among the GCPs which is the minimum set of ground control points.



Figure 5. Distribution of the 118 reference points (green dot) in the region

The number of GCPs varies from 4 to 90 points, each time the four points of the convex hull are included, i.e. the distribution of the other points wherever the number of points are interior of the convex hull, we have chosen only a number of points to use for each test, and we add complementary points to calculate the residuals of each transformation, this set of complementary points is chosen at random from the set.

The selected GCPs are then used to calculate the parameters of the global transformation and the set of parameters of the triangular affine transformation. To evaluate these transformations, 20 independent and random check points are selected among the rest of the points which are not included in the calculation of the parameters of the both transformations.

The differences between the reference and the transformed coordinates of the check points are calculated using the parameters obtained, the mean and the standard deviation are estimated and stored for this set.

As the selection of the additional control points and the 20 control points is done randomly, we perform this process 100 times then the results are averaged to obtain the final results (see figure 6).

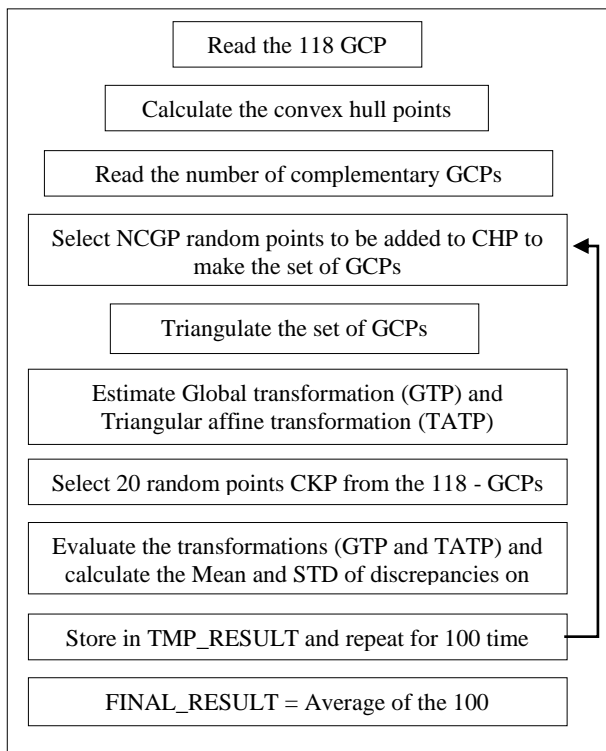


Figure 6. Organigram of the evaluation procedure

4.1 Statistic evaluation using the two transformations

Table 1 illustrates the advantage of using the triangular transformation since it divides the area into triangles and each triangle calculates the parameters of the affine transformation from the vertex. The triangular transformation gives better results when the area is densified by points. The global transformation is commonly used for improving the geometric quality of geodata in general, it gives a better result by using more points until it reaches about 6.5m and remains constant; if more points are added, its standard deviation also stabilizes around 6.3m.

Number of GCP	Mean for TAT	STD for TAT	Mean for GT	STD for GT
4	11.3413	6.4705	9.5748	6.3761
5	8.422	6.4834	8.653	6.4627
14	7.6754	6.9591	7.4319	6.5623
34	6.9488	5.948	6.8283	6.3487
54	6.5199	5.8187	6.681	6.403
74	6.371	5.7253	6.5371	6.2835
94	5.9828	5.4526	6.7104	6.5464

Table 1. The mean and standard deviation for TAT and GT after 100 runs of the check point.

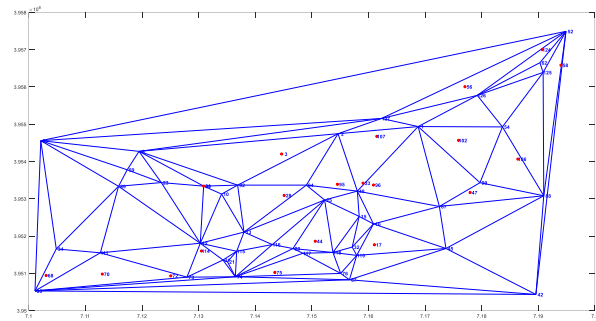


Figure 7. Example of the four convex hull and 50 complementary points triangulation (triangles vertices) with 20 check points (in red).

The proposed approach is dynamic, because the more points we add the triangles became smaller and the associated transformation is regionalized, this is why the mean and the standard deviation of the error is getting smaller when more points are added without a constant stabilization.

This conclusion is made on this dataset for the city of Oran, which is a flat area and seems coherent, but in areas where the terrain is more rugged, the proposed approach may be more effective in enhancing the OSM road network, to test this hypothesis unfortunately we do not have data that covers a more rugged area.

4.2 Statistic evaluation after perturbation

To evaluate the previous hypothesis for testing on a more rugged area, the data set we have on the Oran area is regionally disrupted. The reference dataset (map) is broken down into four regions as follows : North-East, South-East, North-West, and South-West, each of these regions being perturbed by a constant error. The range of this error has been set at $\pm 10m$, $\pm 20m$, and $\pm 30m$ as shown in Table 2. This leads to the creation of the three simulated data sets which are used to perform the tests as in the first case.

The range of the error (m)						
	± 10		± 20		± 30	
	X	Y	X	Y	X	Y
N-E	-3.706	-1.654	16.518	2.803	-17.433	-25.500
S-E	-3.983	-6.964	-9.668	-5.399	25.340	-6.121
N-W	6.958	3.630	0.404	-1.244	10.648	21.754
S-W	1.205	-4.475	6.801	19.541	-16.155	-29.997

Table 2. The perturbation error added to each region for the three simulated datasets.

The results of the mean and standard deviation of the check point for both transformations after perturbation the GCP's by different ranges of error is illustrated in the following figures:

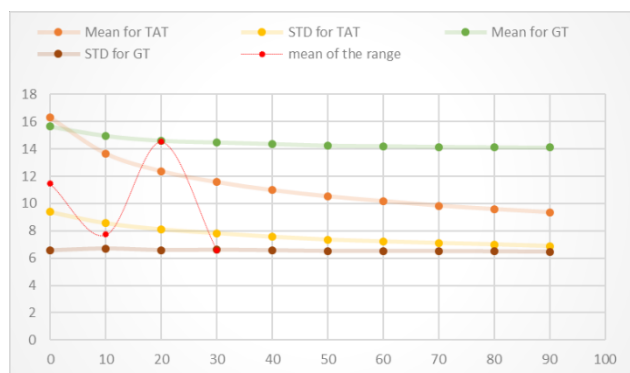


Figure 8. Plot of the resulting mean and std of the check point for both transformations after perturbation of GCPs by ± 10 m.

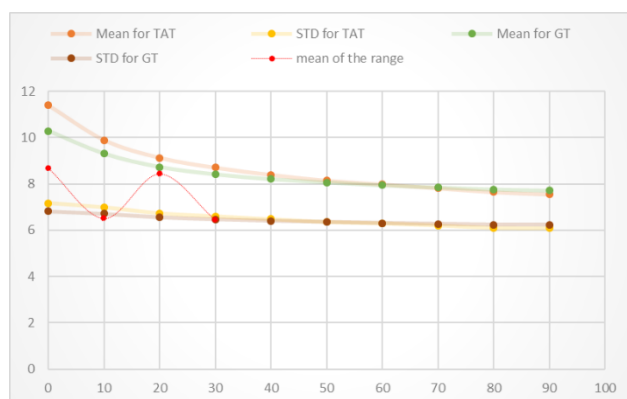


Figure 9. Plot of the resulting mean and std of the check point for both transformations after perturbation of GCPs by ± 20 m.

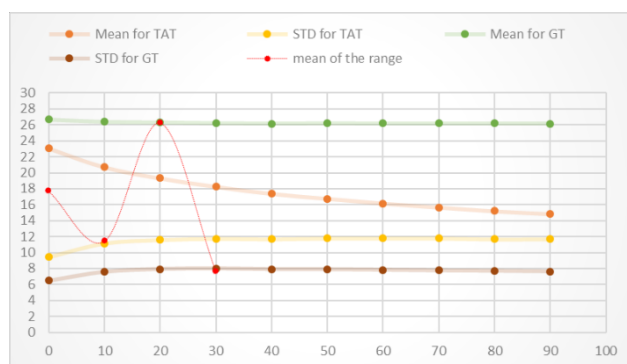


Figure 10. Plot of the resulting mean and std of the check point for both transformations after perturbation of GCPs by ± 30 m.

The results of the second test, after perturbation of the GCP coordinates, show that the proposed approach improves the OSM road network geometrically in rugged areas efficiently.

We notice that when the GCPs coordinates is perturbed by ± 10 m, the results of the mean and the standard deviation for both transformations have almost the same results. But when the threshold of the GCPs coordinates perturbation is augmented to

± 20 m and ± 30 m, the results of the global transformation are stable at higher value for the mean residuals (14 and 26, respectively) with the standard deviation of about 7m. On the contrary, the results of the mean of the error using triangular affine transformation decreases to achieve 9 and 15m, respectively (see figures 9 and 10).

According to the results of this test, we conclude that it is not suitable to use the global affine transformation for geometric correction in rugged terrain or in some regions where the error is not stable, but the triangular affine transformation (local transformation) generates better results in such cases.

4.3 Visual comparison between both transformations

We will confirm these results by a visual comparison of the downloaded OSM road network and after correction using both transformations (see figures 12 and 13).

Reading the OSM road network is done using the function provided by Matlab is the first step.

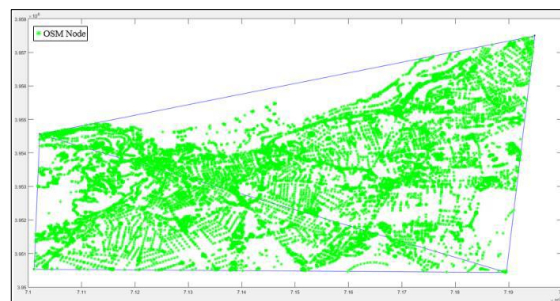


Figure 11. OSM node after reading the road network (green dot)

After reading the OSM data, we proceed to the finding of the triangle that each node is belonging to. The inclusion of the convex hull of the control points data will ensure that each point in the dataset will be included in one triangle. When the triangle is found we extract the correspondent transformation parameters, then these parameters are used to calculate the correct position of the node.

In the case of the global transformation, the same parameters are applied to the whole set of OSM nodes. At the end the program exports the OSM data with the new corrected node as shape file readable in GIS software.

The next figures illustrate the results of the correction using the maximum set of GCPs for both methods (GT and TAT) along with the original OSM data.

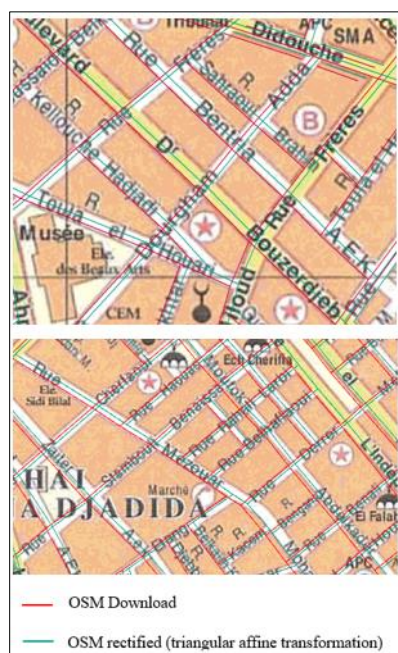


Figure 12. OSM network before and after geometric correction using triangular affine transformation

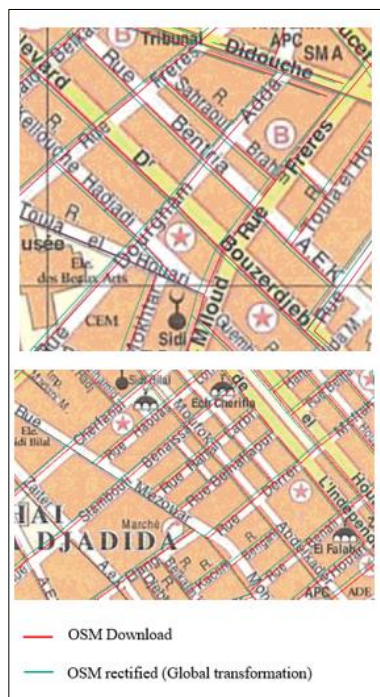


Figure 13. OSM network before and after geometric correction using global transformation

5. CONCLUSION

The triangular affine transformation method proposed for the geometric correction of the OpenStreetMap road network using reference data generates a successful correction according to the results. The analysis was based on a statistical evaluation of the residuals calculated using random sets of check points and a different number of control points. The residue statistics indicate that the affine triangular transformation gave the best geometric

correction of the OSM road network and we confirmed the results in the case of rough terrain on an original data set and a data set perturbed by a random error in a different range from 10m to 30m. The visual inspection demonstrates the improved quality of the position.

In a future perspective, we aim to develop on one hand an automated method for collecting the control point, and on the other hand the possibility to use polyline vector data to check the accuracy of the transformation by measuring the similarity between vector maps.

REFERENCES

- Baley, M, Touya, G. ; Intégration et correction automatique de données OpenStreetMap laboratoire COGIT, IGN, Janvier-Février 2014.
- Brovelli, m. A., Minghini, M, Molinari, M. E; an automated grass-based procedure to assess the geometrical accuracy of the OpenStreetMap Paris road network.
- Ciepluch, B. , Mooney, P. , Jacob, R. , Zheng, J. (2011), Winstanely, A. Assessing the quality of open spatial data for mobile location-based services research and applications.
- Delaunay, Boris (1934). 'Sur la sphère vide'. Bulletin de l'Académie des sciences de l'URSS, classe des sciences Mathématiques et naturelles. 6 :793-800.
- Girres, J. F., Toua, G., 2010. Quality assessment of the French OpenStreetMap dataset. Transactions in GIS.
- Goodchild, M. F., 2007. Citizens as sensors: the world of volunteered geography.
- Graser, A., Straub, M., Dragaschnig, M., (2013). Towards an Open Source Analysis Toolbox for Street Network Comparison: Indicators, Tools and Results of a Comparison of OSM and the Official Austrian Reference Graph.
- Haklay, M., 2010. A comparative study of OpenStreetMap and Ordnance Survey datasets.
- Helbich, M., Amelunxen, C., Neis, P., Zipf, A., 2012. Comparative Spatial Analysis of Positional Accuracy of OpenStreetMap and Proprietary Geodata.
- Jokar Arsanjani, J., Barron, C., Bakillah, M., Helbich, M., 2013. Assessing the Quality of OpenStreetMap Contributors.
- kounadi, O (2009) Assessing the quality of OpenStreetMap data.
- Ludwig, I., Voss, A., Krause-Traudes, M., 2011. A comparison of the street networks of Navteq and OSM in Germany.
- Nasiri, A, Abbaspour, R. A, Chehregghan, A and Jokar, Arsanjani J (2018), Improving the Quality of Citizen Contributed Geodata through Their Historical Contributions: The Case of the Road Network in OpenStreetMap.

Petit O., Billon P., Follin JM. (2012). Évaluation de la qualité des données OpenStreetMap : sur la Sarthe et réflexion sur le processus de contribution.

Ruffray. S, Hamez.G; 5. Généralités liées à OpenStreetMap et la complétude des données.

Siebritz. L.A., Assessing the Accuracy of OpenStreetMap Data in South Africa for the Purpose of Integrating it with Authoritative Data; February 2014.

Zielstra. D, Zipf. A (2010). Quantitative Studies on the Data Quality of OpenStreetMap in Germany.