UNCERTAINTY OF OPENSTREETMAP DATA FOR THE ROAD NETWORK IN CYPRUS

Demetris Demetriou*[a, b]

[a]Public Works Department, 165 Strovolos Avenue, 2048 Nicosia, Cyprus
[b]School of Geography, University of Leeds, Leeds LS2 9JT, UK

ABSTRACT

Volunteered geographic information (VGI) refers to the geographic data compiled and created by individuals which are rendered on the Internet through specific web-based tools for diverse areas of interest. One of the most well-known VGI projects is the OpenStreetMap (OSM) that provides worldwide free geospatial data representing a variety of features. A critical issue for all VGI initiatives is the quality of the information offered. Thus, this report looks into the uncertainty of the OSM dataset for the main road network in Cyprus. The evaluation is based on three basic quality standards, namely positional accuracy, completeness and attribute accuracy. The work has been carried out by employing the Model Builder of ArcGIS which facilitated the comparison between the OSM data and the authoritative data provided by the Public Works Department (PWD). Findings showed that the positional accuracy increases with the hierarchical level of a road, it varies per administrative District and around 70% of the roads have a positional accuracy within 6m compared to the reference dataset. Completeness in terms of road length difference is around 25% for three out of four road categories examined and road name completeness is 100% and around 40% for higher and lower level roads, respectively. Attribute accuracy focusing on road name is very high for all levels of roads. These outputs indicate that OSM data are good enough if they fit for the purpose of use. Furthermore, the study revealed some weaknesses of the methods used for calculating the positional accuracy, suggesting the need for methodological improvements.

Keywords: Volunteered geographic information; OpenStreetMap; data quality; road network; GIS, Cyprus

1. INTRODUCTION

Volunteered Geographic Information (VGI)[1] also called crowdsourcing[2] geospatial data or neogeograpgy[3] involves the geographic data collected, created and disseminated voluntarily by individuals on the Internet through specific web-based tools for various fields of interest. VGI is a considerable issue in the geospatial information community because it extends the perspectives of authoritative information provided by the National Mapping Agencies (NMAs)[4,5]. However, a critical issue arisen regarding the quality [6,7] of geographic data in general and VGI[8] in particular, is uncertainty[9]. Uncertainty is the discrepancy between geographic reality and its representation, either as a result of observations or propagation through data management and analyses in a GIS environment. Ignoring uncertainty may lead to incorrect analysis, interpretations, predictions and eventually to inappropriate related decision making, policies and actions. Therefore, modeling uncertainty [10,11,12], quantifying and managing[13] it are very important tasks. Several quality standards for geographic information (GI) have been produced by several organisations (e.g. The US Geological Survey, the International Cartographic Association). Though, the most well-known are those developed by the International Standardization Organization (ISO) which have been adopted by the European Committee for Standardization[14]. These standards, which have been updated in 2013 and denoted as ISO19157: Geographic Information-Data Quality[15] define six data quality measures: completeness, positional accuracy, thematic accuracy, logical consistency, temporal quality, and usability element among which the first three are employed in this research.

The most well-known VGI project is OpenStreetMap (OSM)[16] which has been created in 2004 and provides worldwide free geospatial data representing a wide range of features (e.g. roads, streets, tracks and buildings). Several studies have investigated the quality of OSM for various countries, e.g. UK[17], France[18], Germany[19], Ireland[20] and Iran[21]. Nevertheless, there is a lack of such a study for Cyprus road network. Consequently, this paper aims to quantify the uncertainty of OSM data for the main road network in Cyprus. The case study involves the four main road categories (A, B, E and F) with a total length 1275 Km (20% of the entire road network and 43% of the network under the authority of the Public Works Department-PWD). This task is timely, very useful because recently (on February, 2016) it has been published the Geospatial Information Portal of Cyprus (http://inspire.cyprus.moi.gov.cy/geocatalog/srv/eng/catalog.search#/search

<http://inspire.cyprus.moi.gov.cy/geocatalog/srv/eng/catalog.search>), as a commitment of EU members, based on INSPIRE Directive[22], that provides free of charge the reference road network utilized in this study. Results showed that OSM data are sound enough if they conform to the purpose of use and they may be used complementary to the authoritative data. Further analysis revealed a spatial variation of data quality and some weaknesses of the methods employed.

The sequent Section 2 briefly outlines the standards and methods utilized for evaluating the quality of GI and apposes some related work. Afterwards, Section 3 presents the case study area, the road network under consideration and the datasets compared. Then, the outputs are presented and discussed in Section 4 and eventually some conclusions are summarized in Section 5.

2. EVALUATING THE QUALITY OF GEOGRAPHIC INFORMATION

2.1 Standards for spatial data

ISO19157 document[15] encompasses six quality elements, that is, *completeness, positional accuracy, thematic accuracy, logical consistency, temporal quality and usability*. In particular, *positional accuracy* is defined as the accuracy of the location of an object in the database related to the reality on the ground based on a reference system. Usually, relative or internal accuracy is employed to represent positional accuracy i.e. it measures the closeness of the relative positions of features in a dataset to their respective relative positions accepted as or being true. *Completeness* denotes the presence and absence of features, their attributes and relationships reflecting two data quality elements, namely, omission and commission. In other words, the question addressed is: how many objects are expected to be found in the dataset but are missing (omission) and how many are excess data (commission)?

Attribute (or thematic) accuracy, refers to the accuracy of quantitative and non-quantitative attribute information of the objects concerned, their classification and relationships. In the former case, it is investigated how close is the value of a quantitative attribute to a value accepted as or known to be true, while the latter measures whether a non-quantitative attribute is correct or incorrect. *Logical consistency* represents the degree of conformance to logical rules of data structure, attribution and relationships. *Temporal quality* reflects the quality of the temporal attributes and relationships of features and *usability* evaluation is based on user requirements which may encompass all the aforementioned quality elements and additional specific user requirements about a dataset's suitability for a particular application. This research employs the first three quality elements i.e. positional accuracy, completeness and thematic accuracy, because they are the most important for the scope of this study and they cover the main quality aspects for the certain dataset. The methods utilized to represent and quantify these three quality standards are discussed below.

2.2 Methods utilized in this work

For the assessment of positional accuracy there are four main line-based methods noted by Ariza-Lopez et al.,[23] : the Hausdorff Distance; the single Buffer Overlay; the Mean Distance and the Double Buffer Overlay. This research employs the second and third methods denoted in this paper as: *buffering method* (BM) and the *average distance method* (ADM) which are both appropriate for the evaluation of road networks. The BM developed by Goodchild and Hunter[24] and has been used in several studies with different scope [25,26,27,28]. The method provides that, for a given pair of linear elements, it measures the percentage of length of a tested line that falls within the buffer zone (or error band) of a changeable width x (in meters) which is defined by the user along the true line e.g. the real road centerline (Figure 1). The definition of a suitable buffer distance may be based on the desired or the expected positional accuracy of data. The noted method refers to a single buffer while a modified version proposed latter by Tveite[28] uses buffers in both line elements.

The ADM developed by White[29] and McMaster[30] and it measures for a pair of linear elements, the average distance (in meters) between the tested and a reference-real line (the dashed and the continuous line, respectively, shown in Figure 2) by dividing the enclosed area with the length of the tested line. This method has been also employed by several studies in different fields [31,32].

Completeness consists of two quality elements: the *road length* and *road name*. The former measures the difference in length between the matched roads in the reference and dataset. This difference can be plus or minus and is checked per road category and per road, for a sample of roads. The latter is quantified by measuring the percentage of road names found in the test data. These names can be correct or wrong. Thus, *attribute accuracy* identifies the percentage of these road names which are correct compared to the reference dataset.

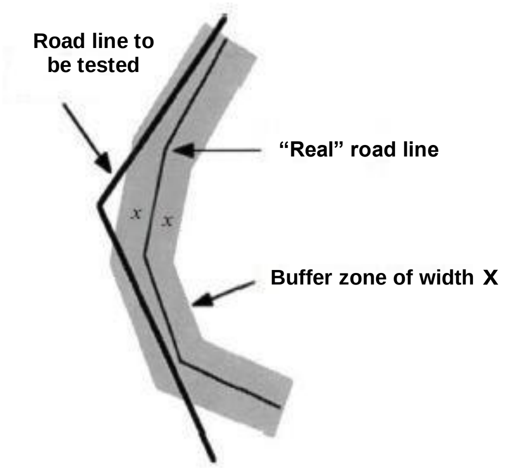

Figure 1. The buffering method

$$dA(A,B): \frac{Area_{(line\ A, line\ B)}}{length(line\ being\ evaluated)}$$

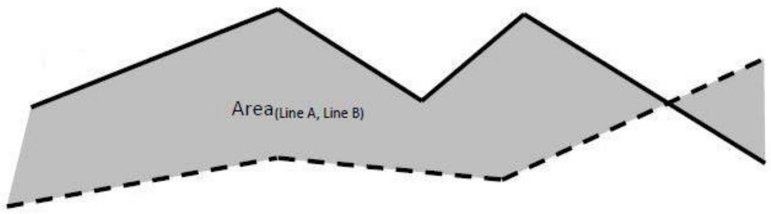

Figure 2. The average distance method

2.3 Related work

The investigation of the quality of data provided by OSM has been the subject of several studies for various countries (mainly European) which have been carried out the last six years. In particular, Haklay[17] compared OSM data with Ordnance Survey data for the road network in England and London. Analysis showed that OSM are fairly accurate, that is, on average within about 6 m of the position of the authoritative data and around 80% overlap of motorways objects between the compared data. Similarly, Zielstra and Zipf[33] and Neis et al.[34] compared the OSM network of Germany with that provided by TeleAtlas and TomTom Multinet, respectively. They both showed that OSM data are often more complete in urban areas than commercial datasets, while OSM is poor in rural areas. In addition, Girres and Touya[18] evaluated the quality of OSM for France involving not only line based data (i.e. roads) but also points and polygons. They found that positional accuracy was, on average, better than 6m, while the main attribute information was very good for all feature types, in contrast to the secondary information which was poor. Further studies noted in chronological order are those for: Ireland[20], Iran[21], Greece[35] and Sweden[14]. Other studies assessed both contributors and data of OSM[36].

3. CASE STUDY

3.1 Study area and road classification

The study area refers to the free part of Cyprus which is under the control of the Republic of Cyprus and it does not include data from the northern part, which is occupied by Turkey since 1974. Cyprus is administratively divided into six Districts, that is, Nicosia, Limassol, Larnaka, Paphos, Famagusta and Kerynia. The road network in Cyprus is split up into four main categories denoted by: A (motorways); B (main roads), E (secondary roads) and F (local roads). Both Districts and the whole road network (per road category) of the study area are shown in Figure 3. Especially, 'A' stands for roads with divided carriageway and central reserve, two directions, with at least two lanes each and a hard shoulder, with interchanges. Similarly, 'B' represents main roads with or without dividing carriageway, connecting important locations between them. In urban areas, these roads they form the main road network together with the motorways, and they are connected directly to the national road network. 'E' roads connect locations of secondary importance with the motorways or the main roads. In urban regions, such roads connect the main roads with the various population centres and the core of commercial activities. Lastly, local roads ('F') connect remote communities or centres between them or with the secondary roads. In the urban fields, such roads are the roads that lead to households.

The whole road network of the study area has in total around 6.5 thousand Km among which approximately 2.8 thousand Km is under the authority of the Public Works Department which is responsible for the transport planning/engineering in Cyprus. The test network, which is depicted in Figure 4 involves in total 1275 Km (i.e. 43% of the road network, which is under the dominance of the PWD and 20% of the entire road network) and in particular per category as follows A: 251 Km; B: 378 Km; E: 470Km; and F: 177 Km.

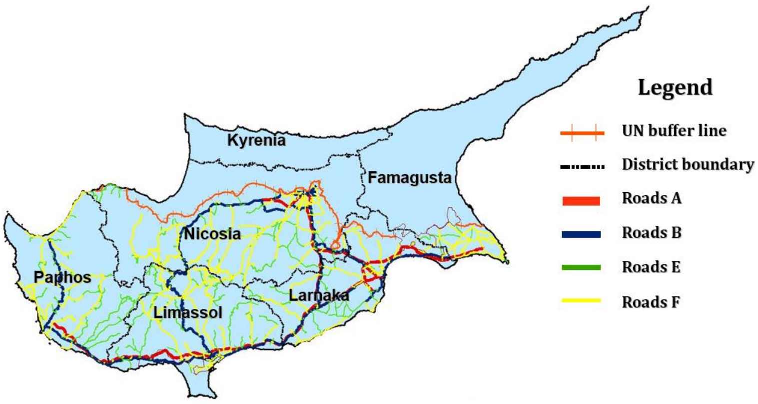

Figure 3. Districts in Cyprus and the whole road network of the study area per road category of Cyprus

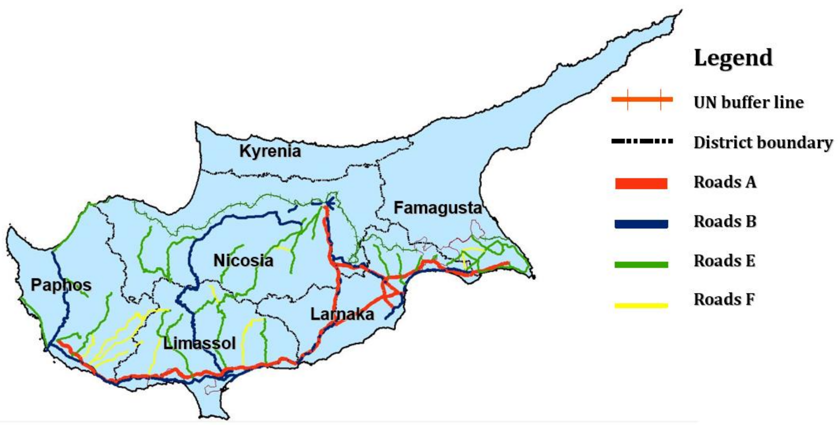

Figure 4. The road network under consideration

3.2 The datasets compared

This study employs external direct quality evaluation approach, that is, it determines the data quality through the comparison of the tested dataset with the external reference dataset. Namely, the former dataset is the road network provided by the OSM which had been downloaded as a shapefile by Geofabrik website on March 2015 whilst the latter is the official road network stored in the GIS of the PWD. The second dataset is a result of digitization of aerial photos with a resolution of 0.5 m. OSM network has more line features because it includes tracks, cycle routes and paths whilst the reference dataset not.

Both datasets stored in ArcGIS and calculations of the quality metrics noted above either facilitated and/or automated by developing groups of GIS functions through Model Builder tool[37]. An example of these models for the BM and the ADM are shown in Figure 5a and 5b, respectively. The comparison based on the road centerlines hence in the case of motorways, which are represented as dual lines, the ArcGIS tool called 'Collapse Dual Lines To Centerline' employed to convert motorways to a single centerline. For the BM the buffer zone has been chosen to be 3m and 6m. The former has been selected because based on the Accuracy Standards for Digital Orthophotos (with a resolution around 0.50m) provided by the American Society for Photogrammetry and Remote Sensing[38], is expected to be 2.97m with a confidence interval level 95%. On the other hand, the expected accuracy of data collected through a GPS receiver, which are usually fed to OSM, is on average 6m. The investigation of positional accuracy for two different buffer distances provides a better overall picture of quality which can be linked with the purpose of use of data.

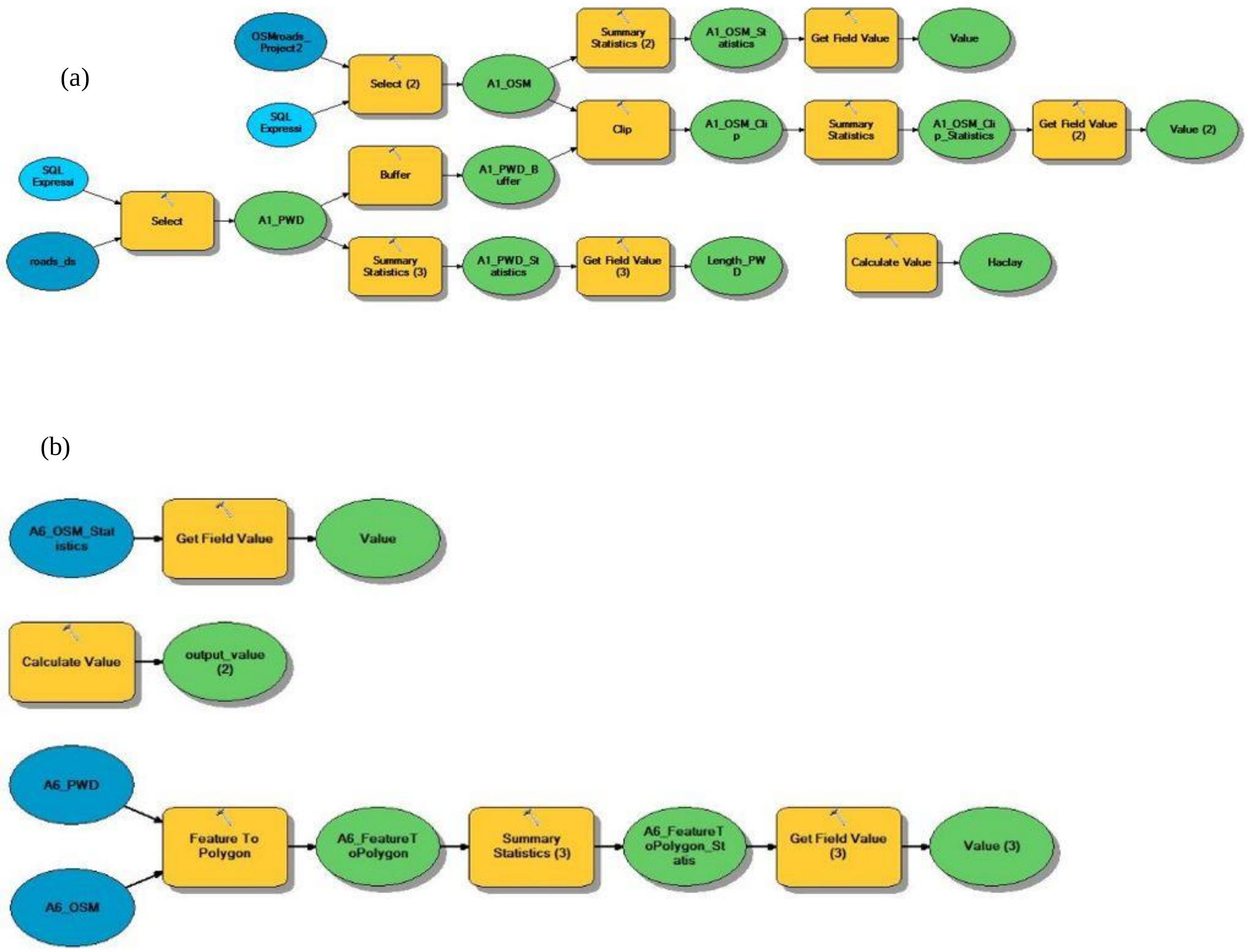

Figure 5. Models for automating: (a) Buffering method and (b) Average Distance method

Further to the use of the noted methods, it is remarkable, noting that recently several studies have focused on developing automated methods to assess data quality of OSM [14,39] which have broader application perspectives. Other authors[40] tried to improve the quality metrics through an unsupervised manner i.e. without employing a reference dataset.

4. RESULTS AND DISCUSSION

4.1 Positional accuracy

The results for the BM (3 and 6 meter buffer zones) for each road category are listed in Table 1. Results are poor for the 3m buffer and quite good for 6m buffer. This finding reveals that the majority of the OSM data fits more to the provided handheld GPS accuracy and small scale road maps (i.e. 1:50,000), rather than to more accurate resources i.e. high resolution orthophotos. In particular, the weighted average overlay of the two lines is only 43% of the former case while it is significantly increased to almost 70% in the latter case. It is also clear that, higher level roads i.e. A and B present better accuracy than road categories E and F. This happens because higher level roads are more important, hence they get more attention in terms of contribution and availability of resources. It was also found that there is no correlation between the accuracy and the length of roads, confirming that accuracy is more dependent on the road category and its importance. In both cases, the range of accuracy variability is large, rather as a result of the variety of skills and reliability of contributors.

Table 1. Positional accuracy results based on the buffering method

Road category	Buffer 3m (% overlay)	Buffer 6m (% overlay)
A	45.95	73.28
B	46.66	72.30
E	37.76	62.46
F	42.60	69.83
Variability	4.45 - 69.68	8.98 - 95.30
W. Aver.	42.68	68.52

A better picture regarding the spatial distribution of positional accuracy (for a 3m buffer) and the relevant classification of roads is provided by map in Figure 6.

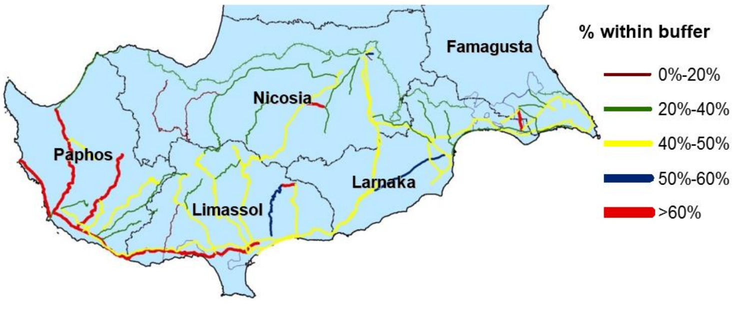

Figure 6. The spatial distribution of positional accuracy (for a 3m buffer) and the relevant classification of roads

It is obvious that positional accuracy is better for Paphos District and the worst of Nicosia District. In addition, the majority of roads classified in the yellow group, i.e. they have an overlay between 40-50%. Furthermore, even visually it is clear that closer roads present similar results as it will also be confirmed later by spatial autocorrelation analysis. Results for 6m buffer are similar apart from the fact that the majority of roads has an overlay between 60-80% since accuracy is increased.

Further to BM the outputs of the ADM are summarized in Table 2. The results present the same trends as those with the BM. Specifically, accuracy gradually increased with the level of road category and there is no evidence based on the calculated correlation coefficient R for a relation between the accuracy and the length of roads. Variability of accuracy per road also presents a wide range leading to the aforementioned conclusion. It is also remarkable that no correlation exists between the outcomes (per road) of the two noted methods, suggesting that they indicate different things and they may lead to different conclusions. In particular, BM reflects a direct absolute measure i.e. the percentage of the length of a line that falls within the buffer that involves a clear interpretation while; ADM represents an indirect metric through a ratio i.e. the mean distance difference of the two polylines which is sensitive to outliers and relevant errors.

Table 2. Positional accuracy results based on the average distance method

Road category	Area/Length
A	2.73
B	3.38
E	4.78
F	6.11
Variability	0.6 - 63.79
W. Aver.	5.82

Likewise with BM, Paphos presents the best outcome and Nicosia the worst among Districts. Furthermore, the majority of roads has an accuracy between 0-6 m which is the most common accuracy can be achieved by a regular handheld GPS. It was also visually clear that closer roads present similar accuracy. In addition, spatial autocorrelation statistics have been calculated by employing the relevant tool within ArcGIS and it was found that both positional quality measures (BM, ADM) and length completeness present a positive SA. In all cases z-score and p-value indicate statistical significance, hence a positive Moran's index value, providing evidence for a tendency towards clustering (Figure 7). In other words, uncertainty at proximal roads correlated, meaning that the same contributor(s) work out these adjacent roads.

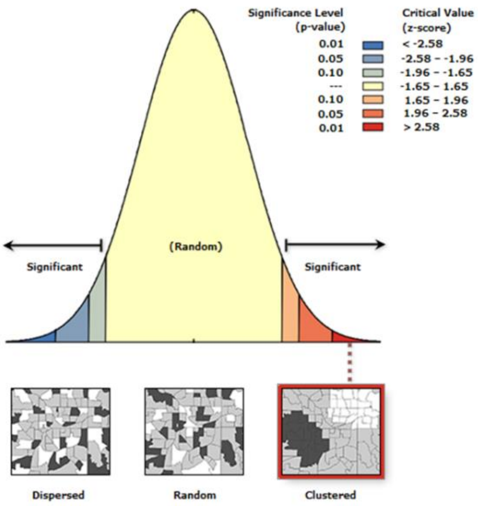

Figure 7. The spatial autocorrelation statistics

Further to the positional accuracy findings, the application of both methods has revealed some weaknesses. Specifically, the BM can give the same result for two tested lines which may be closely located to the extremes i.e. on the external boundary of a buffer and very close to the center line. Thus, the method cannot catch accuracy below the buffer length.

In this example shown in Figure 8a, while the dashed line is very close to the true line except a small segment in the middle, it presents the same overlay percentage with the bold continuous line which is quite farther from the true line. Likewise, the ADM cannot take into account outlier segments i.e. isolated parts of a tested line that may located extremely far from the “normal range” of distance difference with the true line. An example is provided in Figure 8b with an average distance difference between the two roads lines 64m, while this is limited only to that outlier segment whilst the rest of the road has an average difference around 8 meters. Hence the method does not treat well this outlier distance which is a result of a random or a technical mistake. These findings indicate the demand for improving both methods.

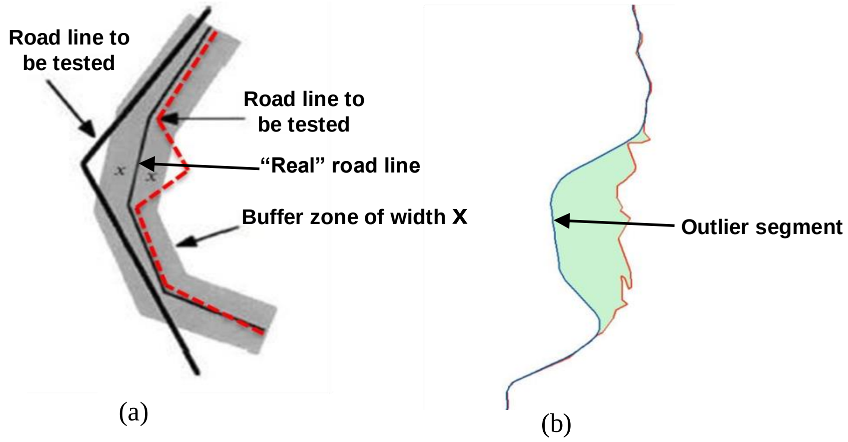

Figure 8. Weaknesses of (a) buffer method and (b) average distance method

4.2 Completeness

Road length completeness has been carried out through SQL query analysis within ArcGIS based on manual road matching and visual inspection. The percentage road length difference per road category between the two datasets, by summing up both less and excess length cases, is expressed in Table 3. Results show an identical length matching for B-road category and a difference between 21-27% of the other three classes. More specifically, it was found that OSM network has 4.7% less and 7.37% excess length in total.

Table 3. Road length completeness results

Road category	% road length difference
A	-24.43
B	0.61
E	-26.80
F	-21.31
Variability	-63.26 to + 113.85

The map in Figure 9 shows graphically the completeness per road with less or excess length. Excess length is a result of a wrong name, in other words, that road or a segment of a road should have been assigned to a different category. All Districts have both problems (less or excess length) and the majority of roads have ± 0-25% length difference. Variability has also a large bound.

The second quality metric of completeness i.e. road name completeness is perfect for higher level roads i.e. categories A and B with 100% matching and poor for lower level roads i.e. categories E and F with a matching slightly more than 40%, as indicated in Table 4.

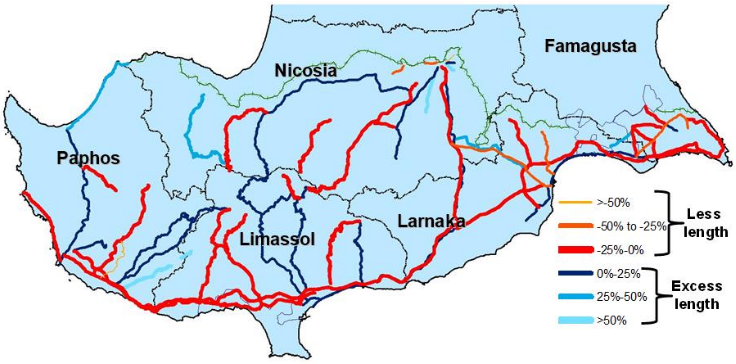

Figure 9. Length completeness per road with less or excess length.

Table 4. Percentage of road names included in OSM data

Road category	% road name completeness
A	100
B	100
E	42.08
F	44.40

4.2 Attribute accuracy

The last measure is attribute accuracy that refers to the road name which is expressed as the percentage of correct road names included in OSM data. Results presented in Table 5 are perfect for motorways and also very good, exceeding 86% of matching, for the other three road categories.

Table 5. Percentage of correct road names included in OSM data

Road category	% correct road names
A	100
B	96.72
E	86.66
F	93.72

5. CONCLUSIONS

This paper explores the quality of data provided by the OSM for the Cyprus main road network. Positional accuracy estimated to be within 6m and 3m for around 70% and 45% of the examined roads respectively, compared to the position recorded by the authoritative data. Higher level roads (A and B) present a better accuracy compared to lower level roads (E and F). Road length completeness present a difference around 25% compared to the official dataset for all road categories apart from B that yields an identical cumulative length matching. Road name completeness found to be perfect for higher level roads and poor for lower level roads whilst road name correctness (attribute accuracy) is excellent for all road classes. These data quality findings indicate that information delivered by OSM for the Cyprus road network is fairly accurate, hence it is good enough if they fit to purpose. In particular, the use of these data is appropriate for a broad scope of subjects, e.g. for planning, visualization and navigation and can be used as complementary to the official data recently provided free of charge by the Geospatial Information Portal of Cyprus in the context of INSPIRE.

Further analysis showed that, data quality regionally varies perhaps due to different skill-levels of contributors. In contrast, spatial autocorrelation has been confirmed for the outputs of positional accuracy methods and completeness, practically meaning that proximal roads present similar quality, possibly because they have been worked out by the same contributors. Moreover, it was found that there is not a direct correlation between the length of a road and accuracy, suggesting an independence of the workload and quality of data. Furthermore, in that respect is no correlation between the results of the two methods used to quantify positional accuracy showing that they bear a different setting and reliability. Especially, the average distance method is somehow flawed because is prone to the outliers and the weaknesses of the average whilst buffering method is more reliable because it measures a direct metric.

The study has also shown a weakness for each of the methods employed for estimating positional accuracy signifying the need for improving the current methods in terms of reliability and accuracy. These methodological improvements, the integration of OSM data (and VGI data in general) with authoritative data and the automation of quality evaluation process, constitute interesting future research directions.

ACKNOWLEDGEMENTS

This research has been carried out in the context of my participation in the COST Action TD1202-"Mapping and the Citizen Sensor", hence I thank very much the people involved in that Action who made useful comments about it. I also thank very much the Public Works Department, which kindly provided the reference dataset.

REFERENCES

[1] Goodchild, M.F., "Citizens as sensors: the world of volunteered geography" GeoJournal, 69, 211-21(2007).

[2] Heipke, C., "Crowdsourcing geospatial data" ISPRS Journal of Photogrammetry and Remote Sensing 65(6), 550-557 (2010).

[3] Turner, A.F., [Introduction to Neogeography], Sebastopol, CA, O'Reilly, (2006)

[4] Al-Bakri, M., Fairbairn, D., "Assessing the accuracy of 'crowdsourced' data and its integration with official spatial data sets" In Proceedings of the Ninth International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, Leicester, UK, 20–23 July, 317–320 (2010).

[5] Olteanu-Raimond, A-M., Hart, G., Foody, G., Touya, G., Kellenberger, T., Demetriou, D., "The Scale of VGI in Map Production: A Perspective on European National Mapping Agencies" Transactions in GIS doi: 10.1111/tgis.12189 (2016)

[6] Bernhardsen, T., [GIS: An introduction], Wiley, New York, (2002).

[7] Heywood, I., Cornelius, S., Carver, S., [An Introduction to Geographical Information Systems]. Pearson Education Limited, Harlow, (2002).

[8] Goodchild, M.F., Li, L., "Assuring the quality of volunteered geographic information" Spatial Statistics 1, 110-20 (2012).

[9] Longley, P., Goodchild, M., Maguire, D., Rhind, D., [Geographic Information Systems and Science], Wiley, Chichester, 123-143 (2011).

[10] Shahin, F., "Modeling positional uncertainty if linear features in geographic information systems" An-Najah Univ. J. Res., 11 23-38, (1997).

[11] Fisher, P.F., "Models of Uncertainty in Spatial Data" In: [Geographical Information Systems: Principles, Techniques, Management and Applications] Longley P., Goodchild, M. Maguire D. and Rhind D. (eds), New York, John Wiley & Sons, 191–205 (1999).

[12] WenZhong, S., "Modeling uncertainty in geographic information and analysis" Science in China Series E: Technological Sciences, 51, 38-47 (2008).

[13] Hunter, G.J., “Managing uncertainty in GIS” In: [Geographical Information Systems: Principles, Techniques, Management and Applications] Longley, P.A., Goodchild, M.F., Maguire, D.J., Rhind, D.W., (eds.) (2005).

[14] Will, J., [Development of an automated matching algorithm to assess the quality of the OpenStreetMap road network: A case study in Goteborg, Sweden], Student thesis series nr 317, Lund University (2014)

[15] International Organization for Standardization (ISO)., “ISO/FDIS 19157 Geographic Information–Data Quality”, ISO/TC 211, Geographic Information/Geomatics Secretariat: SN (Norway) (2013).

[16] OpenStreetMap, “About OpenStreetMap”, <http://wiki.openstreetmap.org/wiki/About_OpenStreetMap>, (2016).

[17] Haklay, M., “How good is volunteered geographical information? A comparative study of OpenStreetMap and ordnance survey datasets” Environment and Planning B 37, 682-703 (2010).

[18] Girres, J-F., Touya, G., “Quality assessment of the French OpenStreetMap dataset” Transactions in GIS 14, 435-459 (2010).

[19] Ludwig I., Voss A. & Krause-Traudes M., “How Good is OSM? - Method and Results for Germany”, Proc. Sixth International Conference on Geographic Information Science, Zurich, Switzerland 14-17 Sep. (2010).

[20] Ciepluch, B., Mooney, P., Jacob, R., Zheng, J., Winstanely, A.C., “Assessing the quality of open spatial data for mobile location-based services research and applications” Archives of Photogrammetry, Cartography and Remote Sensing, 22, 105-116 (2011).

[21] Forghani, M.; Delavar, M.R. A quality study of the OpenStreetMap dataset for Tehran. ISPRS Int. J. Geo-Inf. 2014, 3, 750–763 (2014).

[22] European Commission, [DIRECTIVE 2007/2/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL: Establishing an Infrastructure for Spatial Information in the European Community (INSPIRE)], Official Journal of the European Union, (2007).

[23] Ariza-Lopez, F.J., Mozas-Calvache, A.T., Urena-Camara, M.A., Alba-Fernandez, V., Garcia-Balboa, J.L., Rodriguez-Avi, J., Ruiz-Lendinez, J.J., “Influence the sample size on line-based positional assessment methods for road data” ISPRS Journal of Photogrammetry and Remote Sensing, 66, 708-719 (2011).

[24] Goodchild, M.F., Hunter, G., “A simple positional accuracy measure for linear features” International Journal of Geographical Information Science 11(3), 299-306 (1997)

[25] Van Niel, T., McVicar, T.R., “Experimental evaluation of positional accuracy estimates from linear network using point and line based testing methods” International Journal of Geographical Information Science, 16 (5), 455–473 (2002).

[26] Wiedemann, C., “External evaluation of road networks” International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 34, 93–98 (Part 3/W8) (2003).

[27] Chen, C.C., Knoblock, C.A., Shahabi, C., “Automatically and accurately conflating raster maps with orthoimagery” Geoinformatica 12 (3), 377–410 (2008).

[28] Tveite, H., “An accuracy assessment method for geographic line data sets based on buffering” International Journal of Geographical Information Science, 13(1), 27-47, (1999).

[29] White, E.R., “Assessment of line-generalization algorithm using characteristic points” The American Cartographer, 12, 17-27 (1985).

[30] McMaster, R., “A statistical analysis of mathematical measures for linear simplification” The American Cartographer 23, 103–17 (1986).

[31] Skidmore, A., Turner, B., “Map accuracy assessment using line intersect sampling” Photogrammetric Engineering and Remote Sensing 58(10), 1453–1457 (1992).

[32] Veregin, H., “Quantifying positional error induced by line simplification” International Journal of Geographical Information Science 14(2), 113–130 (2000).

[33] Zielstra, D., Zipf, A., “Comparative Study of Proprietary Geodata and Volunteered Geographic Information for Germany” Proc. AGILE 13th AGILE International Conference on Geographic Information Science, ed. Painho. M., M. Y. Santos, and H. Pundt. Springer Verlag. Guimaraes, Portugal, (2010).

[34] Neis, P., Zipf, A., “ Analyzing the contributor activity of a volunteered geographic information project: The case of OpenStreetMap” ISPRS International Journal of Geoinformation 1, 146-65 (2012).

[35] Trigonis, I., “Assessing the quality of OpenStreetMap data: A case study from Greece”, MSC thesis, Harokopio University Department of Geography, Greece, (2014).

[36] Arsanjani, A.J., Helbich, M., Bakillah, M., Loos, L., “The emergence and evolution of OpenStreetMap: A cellular automata approach” International Journal of Digital Earth 8, 76-90 (2013).

[37] ESRI, [ModelBuilder – Creating Tools Tutorial], ESRI Press, (2010).
[38] ASPRS Accuracy Standards for Digital Geospatial Data, American Society for Photogrammetry and Remote Sensing, (2013).
[39] Koukoletsos, T., Haklay, M., Ellul, C., "Assessing data completeness of VGI through an automated matching procedure for linear data" Transactions in GIS 16, 477-98 (2012).
[40] Mooney, P., Corcoran, P., Winstanley, A., "Towards quality metrics for OpenStreetMap" Proc. 18th International Conference on Advances in Geographic Information Systems, 514-517, ACM, New York, (2010).