Instituto de Engenharia de Sistemas e Computadores de Coimbra

Institute of Systems Engineering and Computers

INESC - Coimbra

Nuno Martinho, Cidália C. Fonte

Assessing the applicability of OpenStreetMap data to validate Land Use/Land Cover Maps

No. 5 2015

ISSN: 1645- 2631

Instituto de Engenharia de Sistemas e Computadores de Coimbra

INESC - Coimbra

Rua Antero de Quental, 199; 3000-033 Coimbra; Portugal

www.inescc.pt

The work of Cidália C. Fonte was partially supported by the Fundação para a Ciência e a Tecnologia (FCT) under project grant UID/MULTI/00308/2013.

This page was intentionally left blank

Assessing the applicability of OpenStreetMap data to validate Land Use/Land Cover Maps

Nuno Martinho[(1)], Cidália C. Fonte[(1,2)]

[(1)] Department of Mathematics, University of Coimbra, Coimbra, Portugal

[(2)] Institute for Systems Engineering and Computers at Coimbra (INESC Coimbra), Coimbra, Portugal

Corresponding author: cfonte@mat.uc.pt

September 2015

Abstract

The validation of Land Use/Land Cover Maps (LULCM) is usually performed using a reference database consisting of a sample of points or regions to which the "real" land cover/use class is assigned. This assignment is usually performed by specialists using photo-interpretation of high resolution imagery or field visits, which are time consuming and expensive processes. The aim of the article is to assess if the data extracted from the collaborative project OpenStreetMap (OSM) may be used as reference data to validate LULCM. For this aim two case-studies were analysed where the validation of the GMESUA LULCM is performed with two reference databases, produced for the same sample of points, but in one case using photo-interpretation and the other data extracted from OSM. The results show that for GMESUA level 1 classes' the results obtained from OSM are even better than the ones obtained using photo-interpretation without local knowledge, while for level 2 classes the results are not satisfactory.

Keywords: land use/land cover, accuracy assessment, OpenStreetMap, reference data

1. Introduction

Land Use Cover Maps (LULCM), usually created through the classification of satellite imagery, are fundamental for many areas of application. However, one key point is their accuracy, which is usually assessed using reference data collected either by interpretation of higher resolution satellite or aerial images or by field visits. The process of collecting reference data is a time-consuming and expensive process, which may prevent the LULCM validation from being performed. Therefore, it is necessary to develop methodologies that enable its execution with reliable results but limiting the costs of the process and time needed. This last aspect gains particular importance due to the increasing temporal resolution of satellite imagery and the need to create LULCM with increasing frequency.

The development and maturing of Web 2.0 technologies, as well as the democratization of internet use, enabled the creation of a broad number of initiatives to collect, store and share geographic information provided voluntarily by the general public. This information is usually referred to as Volunteered Geographic Information (VGI) (Goodchild 2007), but other terms may also be used, such as contributed geographical information, or collaborative mapping (Coleman 2013; Haklay 2013; MacGillavry 2003). Initiatives like Flickr, Degree Confluence Project, Wikimapia, GeoWiki or OpenStreetMap (OSM), among others, store rich datasets, with potential application to multiple problems in several fields, such as risk assessment and emergency response, navigation, environmental studies, decision support and LULCM creation and validation (McCormick 2012; Bastin et al. 2013; Estima and Painho 2013; Jokar Arsanjani et al. 2013; Fonte et al. 2015).

Created by the OpenStreetMap Foundation, OSM is a collaborative mapping project to create an editable map at global scale. It is one of the most studied and eventually the most well-known VGI initiative. The spatial data contributed to OSM may be field data acquired with portable GPS devices, features identified over aerial and satellite imagery, local knowledge, or upload of free spatial data (Mooney and Corcoran 2012). Additional information, such as descriptions, names and other tags can be saved with the geographic features. The flexibility of use, data availability, free access to the latest information on a daily basis, large number of user's and the existence of data that is not traditionally available makes OSM a valued source of information for several types of applications.

Jokar Arsanjani et al. (2013) showed that the data available in OSM in urban areas may generate LULCM with high quality. This study was performed for an urban region, in Vienna city, where the information available in OSM is rich (Jokar Arsanjani et al. 2013).

As the work already developed shows that OSM is a source of information with potential to be used as data to assist the creation of LULCM, the aim of this study is to assess if OSM data can also be used to validate LULCM created, for example, through the automatic classification of satellite images, using pointwise sampling units (Stehman and Czaplewski 1998). The analysis focus on determining if the results obtained from OSM data are comparable to the ones obtained with the photo-interpretation of high resolution imagery without any additional local knowledge of the study area or field visits. The key points that need to be assessed are: 1) assess the influence in the validation results of the fact that OSM data are in vector format and therefore gaps may exist in the coverage, which may result in the lack of data for point features used as reference; 2) determine if data extracted from OSM are equally appropriate for all classes; 3) assess if the spatial accuracy of OSM is good enough to generate point reference data with enough quality to provide reliable validation results.

To analyze the three aspects mentioned above regions that have a good coverage of OSM data were selected. So that some prior indicative knowledge of the expected LULCM accuracy was known, LULCM that were already validated with a highly reliable approach were used as case studies. Therefore, Global Monitoring for Environment and Security Urban Atlas (GMESUA) data were used, considering level 1 and level 2 classes, since these maps available metadata provides complete information about the validation results.

Two regions were analyzed, one in the vicinity of London and the other in the vicinity of Paris. To analyze the applicability of OSM for validation, two reference databases were created for the same set of points obtained using a stratified random sampling approach, considering the GMESUA classes as strata (Stehman 2009). For one of the reference databases the thematic data was obtained extracting the classes from OSM and for the other photo-interpretation of very-high resolution imagery was used.

2. Materials and Methods

2.1 Study Area and Datasets

Two study areas were selected, each with 100 km2, in regions with different characteristics (Figure 1). One of the areas is located in the northern region of London (area A), where urban, agricultural and forest areas are present as well as surface water bodies. The other region (area B) is located in France, approximately 60 km northwest of Paris and presents different characteristics, namely smaller and lower density urban areas and predominance of agriculture and forest regions. The corresponding OSM datasets were downloaded using Geofabrik at April 4, 2014. The polygon features “buildings”, “landuse” and “natural areas” were used, as well as the “railways”, “roads” and “waterways” linear features. The feature “points of interest” (POIs) was not considered in the present work.

a)

b)

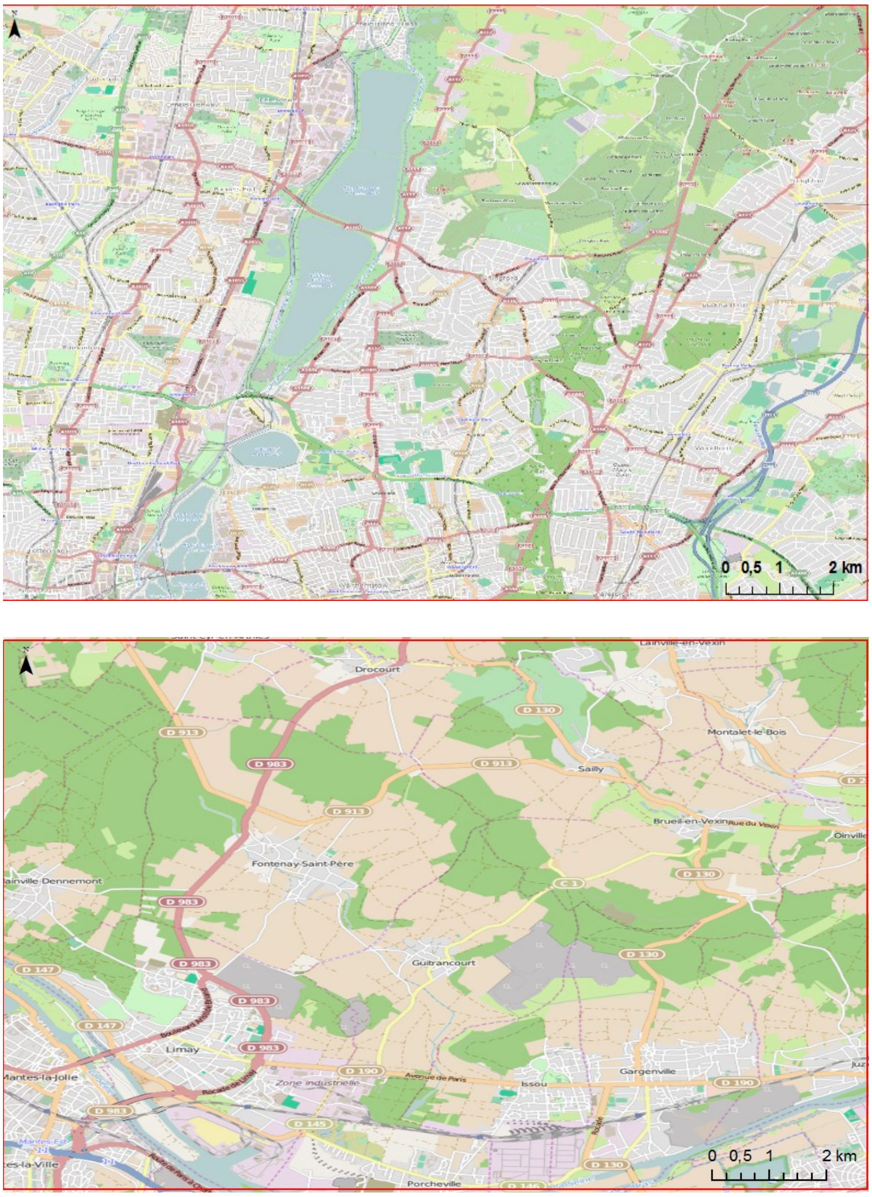

Figure 1: a): OSM data of study area A located in the northern region of London. b): OSM data of the study area B located northwest of Paris.

GMESUA initiative aims to provide high resolution LULCM for Pan-European regions with more than 100 000 habitants. The data is freely available for download through the European Environmental Agency website (http://www.eea.europa.eu/data-and-maps/data/urban-atlas) in vector format. The minimum mapping units is 0,25ha for area features and 100m for linear features, having a geometric scale of 1:10 000. GMESUA LULCM thematic classes are presented in Table 1 (European Union, 2011). GMESUA LULCM for each study area were used, considering in this study only level 1 and level 2 classes.

Table 1. GMESUA classification schema with 3 levels.

Level 1	Level 2	Level 3
1.Artificial Surfaces	1.1 Urban Fabric	1.1.1 Continuous urban fabric 1.1.2 Discontinuous urban fabric 1.1.3 Isolated Structures
	1.2 Industrial, commercial, public, military and private units or transport units are predominant	1.2.1 Industrial, commercial, public, military and private units 1.2.2 Road and rail network and associated land 1.2.3 Port areas 1.2.4 Airports
	1.3 Mine, dump and construction sites	1.3.1 Mineral extraction and dump sites 1.3.3 Construction sites 1.3.4 Land without current use
	1.4 Artificial non-agricultural vegetated areas	1.4.1 Green urban areas 1.4.2 Sports and leisure facilities
2. Agricultural, semi-natural areas, wetlands		
3. Forests		
5. Water		

Table 2 shows the confusion matrices available as metadata for level 1 GMESUA LULCM that include each of the study areas. Even though these results were obtained for the entire map region, and the study regions are subsets of the maps, these results are indicative of the expected accuracy. The global thematic accuracy of the GMESUA LULCM for level 1 classes that includes Area A is 93% and for the map that includes Area B is 94%. The class "Forests" have the higher omission and commission errors in both cases, resulting mainly from a confusion with class 1 ("Artificial Surfaces"). Table 3 shows the accuracy indices available for level 2 classes.

To compare the validation results obtained with OSM data to results obtained through photo-interpretation of a high resolution satellite images, a reference database was built performing photointerpretation of the base maps available in ArcGIS software, with a 30cm spatial resolution for both study areas. The images available for the London region (area A) are from 2011 and the ones available for Paris regions (area B) are from 2010.

Table 2. Confusion matrixes available for GMESUA LULCM containing study areas A and B.

Area A (London)

	1	**2**	**3**	**5**	Total	User's Accuracy
1	**1556**	72	38	2	1668	93%
2	26	**522**	4	0	552	95%
3	12	10	**107**	1	130	82%
5	2	2	0	**28**	32	88%
Total	1596	606	149	31	2382	Global accuracy
Producer's Accuracy	98%	86%	72%	90%		93%

Area B (Paris)

	1	**2**	**3**	**5**	Total	User's Accuracy
1	**1273**	18	33	2	1326	96%
2	34	**571**	7	0	612	93%
3	22	9	**250**	1	282	89%
5	4	0	0	**55**	59	93%
Total	1333	598	290	58	2279	Global accuracy
Producer's Accuracy	96%	96%	86%	95%		94%

Table 3. Accuracy indices available for GMESUA level 2 classes for study areas A and B.

	Area A (London)			Area B (Paris)		
Class	Producer's accuracy	User's accuracy	Global accuracy	Producer's accuracy	User's accuracy	Global accuracy
11	89%	98%	88%	84%	98%	88%
12	91%	68%		89%	79%	
13	82%	86%		78%	67%	
14	92%	71%		82%	65%	
20	86%	95%		95%	93%	
30	72%	82%		86%	89%	
50	90%	88%		95%	93%	

2.2 Methodology

The use of OSM data to create a reference database requires the harmonization between OSM datasets and the LULCM nomenclature, in this case the GMESUA classes. This was established based on the work published by (Estima and Painho 2013) and (Jokar Arsanjani et al. 2013) (see Table 4). For both study areas the classification using the OSM data was done considering the attribute "type" of each feature and the established correspondence between nomenclatures, using an automated procedure done through an ArcPy script. Since the "buildings" dataset has many features that do not have any value for the attribute "type", these cases were considered to belong to the "Artificial Surfaces" of level 1 (class 1). The information about construction sites (class 13) were obtained from type "construction" available for the "landuse" features.

For the linear features such as "roads", "railway" and "waterlines", since there is no information available on their width, a pre-assessment was necessary to determine the typical width for each feature type. To this aim, a sample of features of each type was selected, an estimation of their width was done and a buffer was generated for each line. Subway networks as well as minor and residential roads were not considered.

Table 4. Correspondence between OSM "buildings" and "natural" datasets and the GMESUA classes.

OSM		**GMESUA**	
dataset	**values of attribute "type"**	**Level 1**	**Level 2**
Buildings	*church*	1	12
	clinic	1	12
	commercial	1	12
	industrial	1	12
	library	1	12
	museum	1	12
	public	1	12
	retail	1	12
	school	1	12
	townhall	1	12
	apartments	1	11
	garages	1	11
	residential	1	11
	Other	1	11
Natural	*forest*	3	30
	park	1	14
	riverbank	5	50
	water	5	50

The classified datasets are then combined into a single dataset using a union operation and the GMESUA nomenclature is applied. The existence of areas classified with differing classes in different datasets causes classification difficulties, which were resolved choosing the class occupying the highest level in a proposed class hierarchy (e.g. roadways [level 3 in GMESUA nomenclature] overlaps waterways [level 1 in GMESUA nomenclature]). This enabled the creation of a LULCM using the GMESUA nomenclature for level 1 and level 2 classes, based on OSM.

To generate the reference datasets a stratified random sample of 100 points per class was used, considering level 2 classes of the GMESUA LULCM as strata (e.g. Stehman 2009). One reference database is generated assigning the class to each point using the OSM data and the other is created assigning a class to the points using photo-interpretation of the high resolution satellite imagery. Contingency tables were generated to compare both reference databases as well as the GMESUA LULCM with the two reference datasets. Accuracy measures were computed in all cases, namely the overall accuracy and the user's and producer's accuracy for each class.

3. Results

3.1 Preliminary analysis

Since the methodology used to extract the classes from OSM data produces a LULCM, it is possible to compare these maps to the GMESUA LULCM. Figure 2 shows the results obtained for area A regarding level 1 and level 2. The main aspects that can be identified with a visual analysis are: 1) there are regions of the study area that lack OSM data, resulting in empty spaces in the maps generated from OSM. Based on GMESUA LULCM, omissions occur mainly over agricultural, semi-natural areas or wetlands (class 2) and artificialized areas (class 1). 2) There is some divergence in the classification of some forested regions, which are classified as “Forests” in OSM and as “Artificial non-agricultural vegetated areas” (class 14) in GMESUA, corresponding to green urban spaces (141), and therefore assigned to class 1 (“Artificial Surfaces”) in level 1.

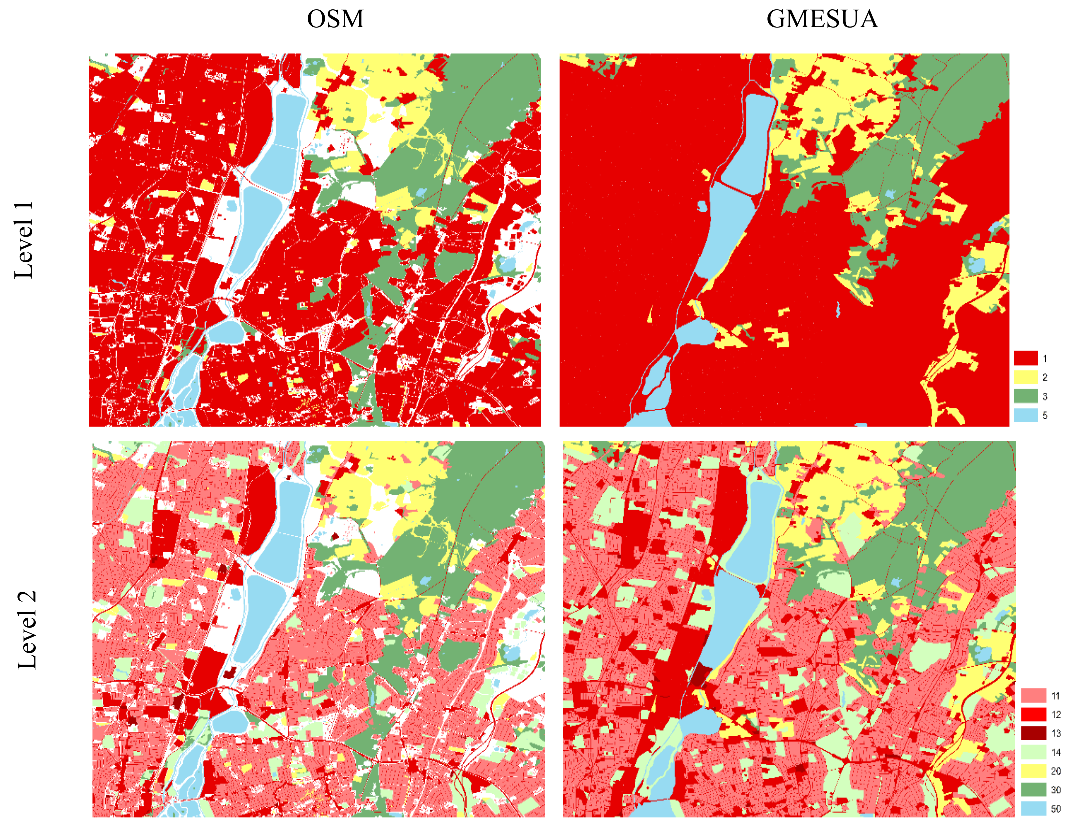

Figure 2 – Land cover map obtained from OSM data and correspondent GMESUA LULCM considering level 1 and level 2 GMESUA nomenclature for study area A.

Figure 3 shows the results obtained for study area B. As obtained for study area A, there are also regions that lack OSM data resulting in empty spaces in the maps generated from OSM. Both in levels 1 and 2 the map obtained from OSM presents a higher degree of generalization, mainly for artificialized areas. It can also be noted that the number of roadways is smaller in OSM derived LULCM, probably due to the consideration of only major and classified roads.

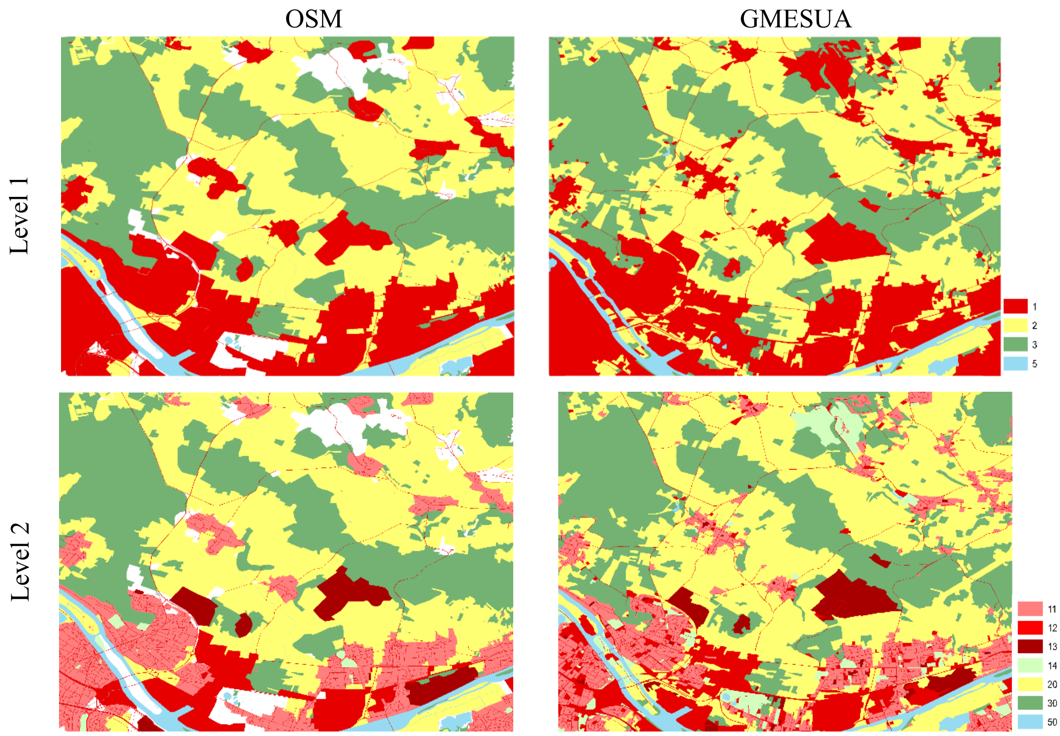

Figure 3 - Land cover map obtained from OSM data and correspondent GMESUA LULCM considering level 1 and level 2 GMESUA nomenclature for study area B.

3.2 Comparison of the reference databases obtained from OSM and by photo-interpretation

The reference database was created considering 100 points per GMESUA LULCM level 2 classes (7 classes, as indicated in Table 1) and therefore 700 points were sampled for each study area. For study area A 11% of the points were located in the empty areas obtained in the OSM generated LULCM (77 points), and for study area B 2.4% (17 points). These points were excluded from the rest of the analysis.

The comparison, for the sample points, between the class derived from OSM data and the class assigned by photo-interpretation of high resolution satellite images was done building a confusion matrix for each study area, considering results of the image photo-interpretation as reference. The results can be seen for area A in Table 5 and Table 6, respectively for level 1 and level 2, and for study area B in Tables 7 and 8 considering the same division.

Table 5. Thematic accuracy for OSM data (level 1) considering as reference the data extracted from the High Resolution Satellite Images (HRSI) for area A.

		Reference: HRSI (*ESRI Basemaps*)					
		1	**2**	**3**	**5**	**Total**	**User's accuracy**
OSM	**1**	**313**	15	13	0	341	92%
	2	8	**57**	3	0	68	84%
	3	4	1	**116**	0	121	96%
	5	0	0	0	**93**	93	100%
	Total	325	73	132	93	**623**	**Global accuracy**
Producer's accuracy		96%	78%	88%	100%		93%

Table 6. Thematic accuracy for OSM data (level 2) considering as reference the data extracted from the High Resolution Satellite Images (HRSI) for area A.

		Reference: HRSI (*ESRI Basemaps*)								
		11	**12**	**13**	**14**	**20**	**30**	**50**	Total	**User's accuracy**
OSM	**11**	**128**	8	4	5	1	0	0	146	88%
	12	5	**63**	6	1	2	3	0	80	79%
	13	0	1	**28**	0	0	0	0	29	97%
	14	0	4	6	**54**	12	10	0	86	63%
	20	0	2	0	6	**57**	3	0	68	84%
	30	0	3	1	0	1	**116**	0	121	96%
	50	0	0	0	0	0	0	**93**	93	100%
	Total	133	81	45	66	73	132	93	**623**	**Global accuracy**
Producer's accuracy		96%	78%	62%	82%	78%	88%	100%		87%

For area A, level 1, a global accuracy of 93% was achieved, representing a good agreement between the OSM derived LULCM and the data obtained from photo-interpretation. The majority of land cover classes present also relatively high producer's and user's accuracy (Table 5). The higher omission and commission errors occur for class 2 ("Agricultural, semi-natural areas, wetlands"), due mainly to some confusion between this class and class 1 (Artificial Surfaces). The results obtained for level 2 (Table 6) show a decrease in the overall accuracy (87%). The analysis of the matrix enables to access that the confusion in level 1 between OSM extracted class 1 and photo-interpreted class 2 is mainly due to OSM "Artificial non-agricultural vegetated areas" (class 14) being classified as class 2 ("Agricultural, semi-natural areas or wetlands"). Moreover, in Table 6 a similar situation can be found between OSM class 14 (Artificial non-agricultural vegetated areas) and class 3 ("Forests"), resulting in a user's accuracy for class14 of only 63%. This situation illustrates how VGI may provide land use information (in this case vegetated areas that are green urban areas and sports facilities) that can be hardly obtained just by photo-interpretation of aerial or satellite images, without additional local knowledge.

A closer analysis of the situation where mismatches are found shows that the confusion between subclasses of class 1 are essentially due to: the width of the buffer applied to linear features such as roadways does not correspond to the real width of the roads (see Figure 4); lack of local knowledge during the photointerpretation process, which makes the distinction between, for example, urban fabric (class 11) and commercial areas (class 12) difficult; omission in OSM of smaller areas due to generalization into larger and more representative areas (i.e. urban fabric [11] may contain regions such as sports and leisure facilities [14]).

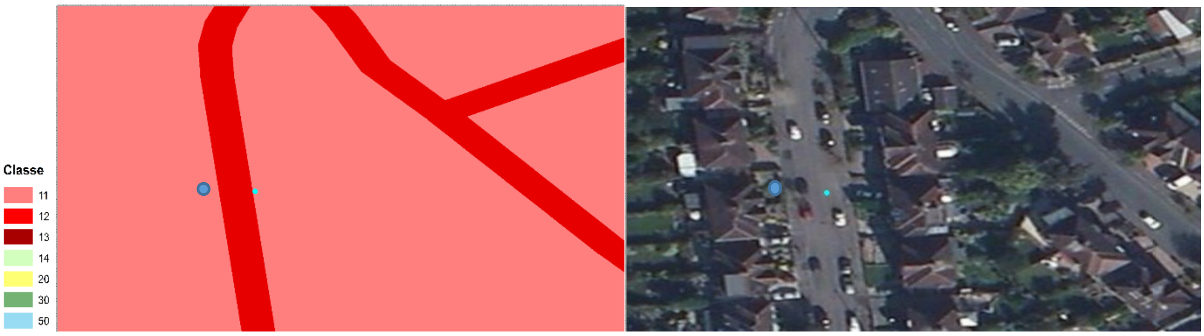

Figure 4 – Divergence in classification of roadway feature motivated by the adoption of an average buffer value.

For Area B an overall accuracy of 87% was obtained for level 1 (Table 7). For most classes the mismatches present similar characteristics when compared to the ones observed for study Area A, except for class 5 (“Water”). In this region there is considerable confusion between the class “Water” in OSM and classes 1 and 2, due mainly to the fact that many points are located in transition zones between the river and the surrounding area, resulting in difficulties in class assignment.

Table 7. Thematic accuracy for OSM data (level 1) considering as reference the data extracted from the High Resolution Satellite Images (HRSI) for area B.

		Reference: HRSI (*ESRI Basemaps*)					
		1	**2**	**3**	**5**	**Total**	**User’s accuracy**
OSM	**1**	**295**	32	12	0	339	87%
	2	16	**110**	4	1	131	84%
	3	6	5	**105**	0	116	91%
	5	4	11	0	**82**	97	85%
	Total	321	158	121	83	**683**	**Global accuracy**
Producer’s accuracy		92%	70%	87%	99%		87%

For level 2 information (Table 8) an overall agreement of 77% was achieved between OSM data and photo-interpreted data. The analysis of level 2 results show that the mismatches between class 1 and class 2 identified in level 1 are mainly due to the assignment of regions located during the photo-interpretation process in class 2 that in OSM are assigned to classes 11 ("Urban fabric") and 13 ("Mine, dump and construction sites"). A closer look to these cases reveals, for the first situation, that OSM feature limits are irregular and not always coincident with the parcel limits and since some control points are located near the boundary of the features (Figure 5) a small positional difference between the image and OSM data influences the results.

Table 8. Thematic accuracy for OSM data (level 2) considering as reference the data extracted from the High Resolution Satellite Images (HRSI) for area B.

		Reference: HRSI (*ESRI Basemaps*)								
		11	**12**	**13**	**14**	**20**	**30**	**50**	**Total**	**User's accuracy**
OSM	**11**	**90**	13	2	37	11	5	0	158	57%
	12	0	**68**	1	1	3	2	0	75	91%
	13	0	10	**50**	0	17	5	0	82	61%
	14	0	0	0	**23**	1	0	0	24	96%
	20	2	4	0	10	**110**	4	0	131	84%
	30	0	5	0	1	5	**105**	0	116	91%
	50	0	2	0	2	11		**82**	97	85%
	Total	92	102	53	74	158	121	83	**683**	**Global accuracy**
Producer's accuracy		98%	67%	94%	31%	70%	87%	99%		77%

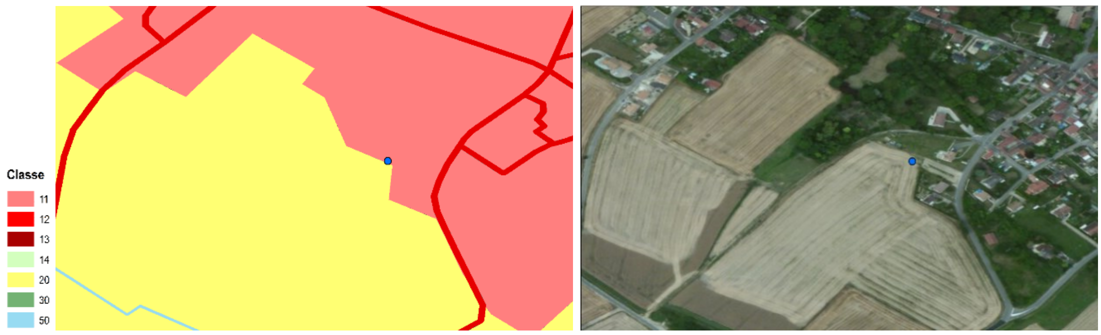

Figure 5 – Classification divergence due to location of control point near the boundary of the feature.

The confusion between OSM class 13 and class 2 results from areas identified by OSM users as quarries and photo-interpreted as semi-natural areas, due to vegetation seen in the high resolution satellite images. Commission errors associated with class 5 are motivated by confusion with class 2, due to control points located near the boundary of the features and in transition zones between the two classes, enhancing the influence of positional discrepancies.

Poor values of user's accuracy are obtained for classes 11 and 13 (respectively 57% and 61%). The first case is due mainly to the inclusion of "sports and leisure facilities" and "green urban areas" (class 14) into class 11 ("urban fabric") where users did not differentiate this type of structures when inserting data into OSM, and also to the identification of "roadways" (class 12) in the high resolution satellite images that were not considered in OSM data or due to the width of the considered buffer, as observed to area A. The assignment of urban green areas to classes 14 or 11 depends on the public or private nature of the area (European Union, 2011) which requires additional knowledge, and therefore is sometimes difficult to differentiate using only photo-interpretation.

Regarding producer's accuracy, low values are obtained for classes 14 ("Artificial non-agricultural vegetated areas") and 12 ("Industrial, commercial, public, military and private units or transport units are predominant") (respectively 31% and 67%). The first case is also due to the confusion between classes 11 and 14, the second case to confusion between 12, 11 and 13.

An analysis of the results obtained for both study areas shows that higher degree of discordance is found between the two sources of data for area B, which has lower urban density and more fragmented urban agglomerates with smaller dimension that alternate with green areas.

3.3 GMESUA LULCM accuracy assessment

In order to access the feasibility of using OSM data as reference information when compared to photo-interpretation, the accuracy assessment of the GMESUA LULCM of both study areas were assessed using the reference databases extracted from OSM and the ones obtained by photo-interpretation. The obtained results were then compared to the accuracy metadata available for GMESUA LULCM. Table 9 shows the accuracy measures obtained for area A considering the two reference databases for level 1 classes and Table 10 for level 2 classes. Tables 11 and 12 show the corresponding results for area B.

For both study areas the overall accuracy of the analysis performed for level 1 with OSM reference database is slightly larger than the one obtained with photo-interpretation (87% versus 85% for area A and 88% vs 84% for area B).

For area A the ranking of classes according to the user's accuracy is the same, while for the producer's accuracy there is a change in the ranking position between classes 2 and 3. Higher users' accuracy values were obtained with the OSM when compared to the results obtained with photo-interpretation for classes 1 and 5 while for classes 2 and 3 larger user's accuracy are obtained when using the photo-interpreted reference data. However, the differences are always smaller than 8%. All obtained accuracy values are larger than 70%, except the producer's accuracy obtained for class 3 with the photo-interpreted reference data, which is 67%, and the producer's accuracy obtained for class 2 with the OSM reference data, which is 69%.

For study area B, level 1 (Table 11), the ranking of classes according the producer's accuracy is the same, while for the user's accuracy there is a change in the ranking position between classes 2 and 5. As for study area A, higher users' accuracy values were obtained with the OSM when compared to the results obtained with photo-interpretation for classes 1 and 5 while for classes 2 and 3 larger user's accuracy are obtained when using the photo-interpreted reference data. In all cases the differences are relatively small with the two larger differences corresponding to 10% for the producer's accuracy of class 2 and user's accuracy of class 5. All obtained accuracy values are larger than 75%, except the producer's accuracy of class 2 which is 54% when the photo-interpreted reference data was used and 64% when the OSM generated references data was used.

Table 1 – Accuracy measures obtained for study area A GMESUA LULCM for level 1 classes considering as reference databases the LULCM obtained with photo-interpretation and from OSM data.

	GMESUA vs Photo-interpretation			GMESUA vs OSM		
Class	Producer's accuracy	User's accuracy	Global accuracy	Producer's accuracy	User's accuracy	Global accuracy
1	95%	83%	85%	95%	86%	87%
2	74%	78%		69%	70%	
3	67%	93%		70%	90%	
5	90%	96%		89%	100%	

Table 2 – Accuracy measures obtained for study area A GMESUA LULCM for level 2 classes considering as reference databases the LULCM obtained with photo-interpretation and from OSM data.

	GMESUA vs Photo-interpretation			GMESUA vs OSM		
Class	Producer's accuracy	User's accuracy	Global accuracy	Producer's accuracy	User's accuracy	Global accuracy
11	58%	80%	72%	57%	88%	68%
12	71%	64%		65%	58%	
13	96%	43%		93%	27%	
14	62%	47%		57%	52%	
20	74%	78%		69%	70%	
30	67%	93%		70%	90%	
50	90%	96%		89%	100%	

Table 3 – Accuracy measures obtained for study area B GMESUA LULCM for level 1 classes considering as reference databases the LULCM obtained with photo-interpretation and from OSM data.

	GMESUA vs Photo-interpretation			GMESUA vs OSM		
Class	Producer's accuracy	User's accuracy	Global accuracy	Producer's accuracy	User's accuracy	Global accuracy
1	96%	79%	84%	97%	85%	88%
2	54%	91%		64%	87%	
3	76%	96%		80%	95%	
5	94%	83%		95%	93%	

Table 4 – Accuracy measures obtained for study area B GMESUA LULCM for level 2 classes considering as reference databases the LULCM obtained with photo-interpretation and from OSM data.

Class	GMESUA vs Photo-interpretation			GMESUA vs OSM		
	Producer's accuracy	User's accuracy	Global accuracy	Producer's accuracy	User's accuracy	Global accuracy
11	77%	71%	76%	55%	89%	74%
12	67%	72%		77%	59%	
13	91%	49%		85%	70%	
14	92%	73%		96%	25%	
20	55%	91%		64%	87%	
30	78%	96%		80%	95%	
50	100%	83%		95%	93%	

The obtained results for level 1 classes' are very similar for both reference databases and in general not very different from the metadata provided by the GMESUA LULCM (see Table 2), however since those values apply to the whole map and not only to the region occupied by the two study areas, differences in the results are expected and natural. Moreover, since both reference data used in this study are less reliable than the one used by GMESUA, because the quality of VGI is not controlled and the reference database created with the photo-interpretation did not include any local knowledge of the area nor field visits, less reliable results were to be expected *à priori*.

The accuracy assessment results obtained for level 2 classes shows significant however differences. The global accuracy of the results obtained with the photo-interpreted reference data are larger for both study areas, when compared to the results obtained with the OSM extracted data (4% for area A and 2% for area B). In both cases particularly low values of accuracy are obtained for some classes. For area A when OSM data are used as reference the user's accuracy of class 13 (mine, dump and construction sites) is only 27%, which corresponds to large commission errors. In this case the great majority of sites classified in GMESUA LULCM as 13 were split by OSM almost equally for classes 11, 12 and 14. Low user's accuracy was also obtained for classes 13 and 14 with the photo-interpreted reference database (respectively 43% and 47%). For class 13 a difference of 16% in the user's accuracy was obtained with the two reference databases. For area B, even though a slightly larger value of global accuracy was obtained, still larger differences were found between the results generated with both reference data for some classes. The most problematic class with the OSM data was also class 14, for which a user's accuracy of only 25% was obtained, due mainly to the fact that in OSM most of these location were assigned to class 11 (Urban Fabric), while with the photo-interpreted reference data a user's accuracy of 73% was reached, corresponding to a difference of 53%. For the assessment made with the photo-interpreted reference database the worst result was 49% user's accuracy for class 13, while with the OSM reference data a value of 70% was reached, corresponding to a difference of 21%. For class 11 a difference of 22% was also obtained for the producer's accuracy with the two reference data. The accuracy results provided in GMESUA metadata for level 2 (Table 3) show all class accuracy values larger than 65% in both study areas, with overall accuracy of 88% in both cases.

4. Conclusions

To assess the possibility of using OSM as reference data for the accuracy assessment of a LULCM, when compared to the results obtained with photointerpretation of high resolution satellite images, the accuracy assessment of GMESUA LULCM of two regions, one close to London and the other near Paris, was made. The use of OSM data requires the harmonization between OSM features and the GMESUA LULCM nomenclature, which was made as proposed by Estima and Painho (2013) and Jokar Arsanjani et al.(2013).

A stratified random sample of points using the GMESUS LULCM as strata was used to build two reference databases, one using data from the LULCM extracted from OSM and the other using photo-interpretation of high resolution satellite images. The samples generated for both study areas included points which were located in areas where no OSM data was available, (11% for area A and 2.4% for area B). The percentage of points in this situation did not raise problems since the minimum number of points per class usually considered as enough to obtain statistically meaningful results (50 points) was available for all classes (Congalton and Green 1998). Therefore, the points where no OSM data was available were excluded from the subsequent analysis. An initial comparison of the two databases was done, showing better correspondence for area A than for area B, corresponding to a global accuracy of respectively 93% and 87% for areas A and B. User's and producer's accuracy of all classes was always larger than 78% for area A and 70% for area B for level 1 classes'. The correspondence between classes of level 2 was much worse, corresponding to a global accuracy of 87% for area A and 77% for area B. A particularly low value of producer's accuracy of 31% was obtained for class 14 ("Artificial non-agricultural vegetated areas") for area B. The obtained differences for level 2 classes result mainly from positional discrepancies between the OSM data and the satellite images; difficulties in interpretation of land use classes in the satellite image without additional local knowledge; differences in the degree of thematic generalization in OSM data, corresponding to lack of detailed information provided by the volunteers; and processing of OSM linear features to generate areas for features such as “roads” and “waterways”.

The accuracy assessment of the GMESUA LULCM was then made with the two references databases and the results compared. The study shows that for level 1 classes’ the results obtained with OSM are very similar to the ones obtained by photo-interpretation of high resolution imagery with no prior local knowledge of the area, generating for both case-studies even higher global accuracy than the results obtained with photo-interpretation. For level 2 classes larger discrepancies are found when using both reference datasets, but these are more problematic for OSM data. When using photo-interpretation alone, most of the difficulties are due to some of the classes correspond to land use and that type of information is harder to obtain without local knowledge. When using OSM as reference, the reasons for the discrepancies are the ones stated above when comparing the OSM with the photo-interpreted data, namely positional discrepancies, generalization and processing of linear features. In the two case-studies considered numerous mismatches were generated for area A due to the assignment of points classified as “Mine, dump and construction sites” in GMESUA LULCM to all other level 2 urban classes (11, 12 and 14) in OSM, which might eventually correspond to time discrepancies between the data used. For area B most problems were raised by the assignment of several points classified as GMESUA LULCM “green urban areas” to simply “urban fabric” (class 11) in OSM, which corresponds to a more general classification.

The results show that OSM has potential to be used for the accuracy assessment of LULCM for regions that have high levels of available data in OSM, such as the regions used in this study. The creation of a LULCM from OSM data resulted in relatively low percentages of empty spaces, and therefore the difficulty of not having enough data was not raised in the considered case-studies. However, for regions that have lower levels of data in OSM this aspect may invalidate the use of OSM for this purpose. The results obtained for level 1 classes’ are quite promising, showing that OSM can be a possible alternative to photo-interpretation of satellite imagery when no additional local knowledge is available, providing in some cases even better results. However, for level 2 classes’ large discrepancies start to emerge, generating anomalous results. This might however be minimized if additional data are used, such as OSM “Points of interest”, which may provide useful land use information, or additional VGI, such as Wikimapia (http://wikimapia.org), Wikiloc (http://www.wikiloc.com) or photographs available in collaborative projects like Panoramio (http://www.panoramio.com) or Flickr (https://www.flickr.com).

One advantage of using data extracted from the LULCM generated from the OSM data is that, since the classes can be automatically assigned to the points of the reference database, samples with more points than the ones usually needed can be used, which overcomes the problem of not having data in some locations. This can also present additional advantages, such as generate outcomes statistically more representative; enable the use other sampling approaches which may present disadvantages when the number of points that can be used is limited, such as systematic samples with a large density of points; and even enable the possibility of not using sampling at all, overlapping both LULCM and perform a direct comparison of both.

References

Bastin, L., G. Buchanan, A. Beresford, J. -F. Pekel, and G. Dubois. 2013. "Open-Source Mapping and Services for Web-Based Land-Cover Validation." Ecological Informatics 14 (March): 9–16. doi:10.1016/j.ecoinf.2012.11.013.

Coleman, David J. 2013. "Potential Contributions and Challenges of VGI for Conventional Topographic Base-Mapping Programs." In Crowdsourcing Geographic Knowledge, edited by Daniel Sui, Sarah Elwood, and Michael Goodchild, 245–63. Springer Netherlands. http://link.springer.com/chapter/10.1007/978-94-007-4587-2_14.

Congalton, Russell G., and Kass Green. 1998. Assessing the Accuracy of Remotely Sensed Data: Principles and Practices. CRC Press.

Estima, Jacinto, and Marco Painho. 2013. "Exploratory Analysis of OpenStreetMap for Land Use Classification." In Proceedings of the Second ACM SIGSPATIAL International Workshop on Crowdsourced and Volunteered Geographic Information, 39–46. GEOCROWD '13. New York, NY, USA: ACM. doi:10.1145/2534732.2534734.

European Union, 2011. Mapping guide for a European Urban Atlas. Available from: http://www.eea.europa.eu/data-and-maps/data/urban-atlas [Accessed 15 April 2014].

Fonte, C.C., Bastin, L., See, L., Foody, G., Lupia, F., 2015. Usability of VGI for validation of land cover maps. International Journal of Geographical Information Science. Accepted. doi:10.1080/13658816.2015.1018266

Goodchild, Michael F. 2007. "Citizens as Sensors: The World of Volunteered Geography." GeoJournal 69 (4): 211–21. doi:10.1007/s10708-007-9111-y.

Haklay, Mordechai. 2013. "Neogeography and the Delusion of Democratisation." Environment and Planning A 45 (1): 55–69. doi:10.1068/a45184.

Jokar Arsanjani, Jamal, Marco Helbich, Mohamed Bakillah, Julian Hagenauer, and Alexander Zipf. 2013. "Toward Mapping Land-Use Patterns from Volunteered Geographic

Information." International Journal of Geographical Information Science 27 (12): 2264–78. doi:10.1080/13658816.2013.800871.

MacGillavry, Edward. 2003. "Collaborative Mapping." Webmapper. http://www.webmapper.net/carto2003/.

McCormick, Sabrina. 2012. "After the Cap: Risk Assessment, Citizen Science and Disaster Recovery." Ecology and Society 17 (4). doi:10.5751/ES-05263-170431.

Mooney, Peter, and Padraig Corcoran. 2012. "The Annotation Process in OpenStreetMap." Transactions in GIS 16 (4): 561–79. doi:10.1111/j.1467-9671.2012.01306.x.

Stehman, Stephen V. 2009. "Sampling Designs for Accuracy Assessment of Land Cover." International Journal of Remote Sensing 30 (20): 5243–72. doi:10.1080/01431160903131000.

Stehman, Stephen V., and Raymond L. Czaplewski. 1998. "Design and Analysis for Thematic Map Accuracy Assessment: Fundamental Principles." Remote Sensing of Environment 64 (3): 331–44. doi:10.1016/S0034-4257(98)00010-8.