Huerta, Schade, Granell (Eds): Connecting a Digital Europe through Location and Place. Proceedings of the AGILE'2014 International Conference on Geographic Information Science, Castellón, June, 3-6, 2014. ISBN: 978-90-816960-4-3

Towards initiating OpenLandMap founded on citizens' science: The current status of land use features of OpenStreetMap in Europe

Jamal Jokar Arsanjani
GIScience research group
Heidelberg University
Berliner str. 48, 69120
Heidelberg, Germany
jokar.arsanjani@geog.uni-heidelberg.de

Eric Vaz
Department of Geography
Ryerson University
evaz@geography.ryerson.ca

Mohamed Bakillah
GIScience research group
Heidelberg University
Berliner str. 48, 69120
Heidelberg, Germany
mohamed.bakillah@geog.uni-heidelberg.de

Peter Mooney
Department of Computer Science,
National University of Ireland, Maynooth
peter.mooney@nuim.ie

Abstract

Land use inventories are important information sources for scholarly research, policy-makers, practitioners, and developers. A considerable amount of effort and monetary resources have been used to generate global/regional/local land use datasets. While remote sensing images and techniques as well as field surveying have been the main sources of determining land use features, in-field measurements of ground truth data collection for attributing those features has been always a challenging step in terms of time, money, as well as information reliability. In recent years, Web 2.0 technologies and GPS-enabled devices have advanced citizen science (CS) projects and made them user-friendly for volunteered citizens to collect and share their knowledge about geographical objects to these projects. Surprisingly, one of the leading CS projects i.e., OpenStreetMap (OSM) collects and provides land use features. The collaboratively collected land use features from multiple citizens could greatly support the challenging component of land use mapping which is in-field data collection. Hence, the main objective of this study is to calculate the completeness of land use features to OSM across Europe. The empirical findings reveal that the completeness index varies widely ranging from almost 2% for Iceland to 96% for Bosnia and Herzegovina. More precisely, more than 50% of land use features of eight European countries are mapped. This shows that CS can play a role in land use mapping as an alternative data source, which can partially contribute to the existing inventories for updating purposes.

Introduction

Land use/cover maps are essential for environmentalists and land managers for urban and regional planning purposes. These maps identify which features exist on the ground and for which purpose each land parcel is used [26,32]. The process of mapping land related features is called land use/cover mapping e.g., [23,34], which result in land use/cover inventories. Traditionally land surveying and recently remote sensing data and algorithms have been used to map land use/cover patterns e.g., [22,28,30]. Undoubtedly, remote sensing has played a vital role in monitoring and mapping land features. Nevertheless, in-field information is often required to assess the outcomes of remote sensing techniques [3,5]. Additionally, they are used to enrich the land use patterns regarding its attributes and semantic information [13].

Recently, the rise of web 2.0 technologies and CS-based projects has resulted in tremendous amount of geolocated information from citizens [9,16]. As a successful leading CS projects, OSM can be named, which has been increasing receiving new users and contributions. Published investigations on applicability of OSM datasets have shown that OSM provides us a wide variety of datasets for different application including and not limited to routing, Points of Interest (POIs) search, transport mapping, building inventories, etc. OSM also collects the information on land features and shares them with public. So far, little attention to the collected OSM features on land use information has been drawn [4,8], although OSM can provide an alternative source for mapping land use features contributed by citizens. What is remarkable about harnessing OSM for land use mapping is the fact that once OSM users log into OSM, fine resolution image libraries generated from multiple remote sensing imageries are shared in the mapping/editing interface so that the users simply delineate the geometrical tessellation of land use features and additionally insert their personal knowledge of that specific land parcel to it. It is of great importance to note that in this process, the OSM users benefit from user-friendly editing softwares, which display fine-resolution images (even up to 20 cm spatial resolution) in the background, for delineating land parcels and add attributes and metadata about each land parcel to it [21]. In other words, thanks to the fine-resolution images/air-photos as well as users' knowledge of the mapped areas, the process of land use mapping is handled differently so that the in-field information are actively given by the users instead of going to the field for collecting them [20].

A remarkable amount of efforts and money have been inserted into generating global land-use maps, for instance, Global Land Cover (GLC)-2000 [11], Moderate-resolution Imaging Spectroradiometer (MODIS; [10]), and GlobCover [1], among others. At a European level scale, the CORINE 2000 [2] and Global Monitoring for Environment and Security Urban Atlas (GMESUA; [3]) have been prepared. The accuracy of these inventories however, is often questioned by the researchers and further projects on evaluating their accuracies are called [19,25,27,29,33]. To sum up, the process of generating land use inventories actively demands for large amount of budget, while this process in a passive manner diminishes the monetary costs significantly and might result in better results. Furthermore, they need to be updated on a regular basis and therefore, repeating the efforts. As such, the main aim is to evaluate the degree of completeness for OSM land use features in order to see how well OSM can play a role in land use science. Empirical findings reported by [15,20] have addressed the potentials of exploiting OSM for land use mapping. Hence, the main objective of this study is to measure how complete OSM land use features in a European scale are in order to start exploiting them. To be more precise, this research seeks to find out how complete land-use features per each European state are contributed to OSM.

Materials and data processing

3.1 OpenStreetMap dataset

The OSM datasets utilized in this study is the OSM snapshot for February 20, 2014. To retrieve relevant land-use features, A country-wide coverage of forty European countries is sampled in this study. The reason for considering a pan-European wide of datasets is the fact the patterns of contributions are intrinsically heterogeneous as proven by [17,21]. This is also evident through a query to osmatrix.uni-hd.de. Figure 1 displays the extent of this study.

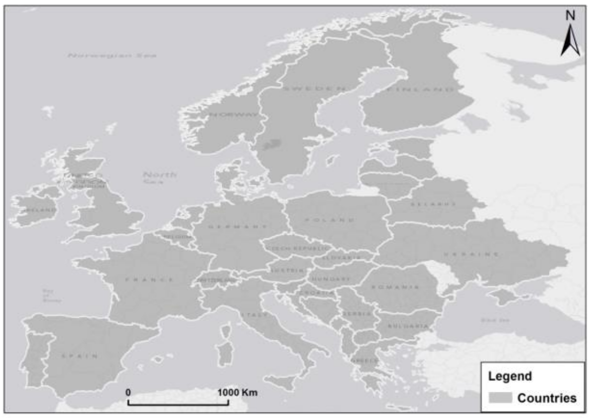

Figure 1: the selected study areas

Methods

Among the purposed criteria by different ISO standards in particular 19157:2013 for assessing the accuracy of geodata internally, completeness plays a vital role as it measures how complete the dataset is [7,14]. Completeness is the major concern for using OSM datasets [18,24] as it is an indicator of how much of the whole has been mapped by volunteers. In contrast to polyline and point features in OSM, the completeness for land-use features is the proportion of mapped areas relate to its overall extent. The completeness index for each country is calculated by calculating the mapped areas by the whole area of extent. This represents a simple indicator to find out how complete a country is mapped i.e., how far we are from having full data coverage.

Results and discussion

Table 1 represents total mapped area and completeness indices for each country. As shown in Table 1, the calculated completeness index values are diverse. While only 1.6% of land use features in Iceland are mapped, 96% of Bosnia and Herzegovina are mapped, which is quite surprising that no study has been already dedicated to further accuracy assessment of the contributed features.

Table 1: the calculated completeness values for each country

Country	Total Area (km2)	Mapped Area (km2)	Completeness (%)	Class
Bosnia & H.	51,209	49,495	96.6	A
Slovakia	49,035	43,698	89.1	A
Netherlands	37,354	30,818	82.5	A
Belgium	30,528	19,221	63.0	A
Romania	238,391	138,737	58.2	A
Luxemburg	2,586	1,426	55.2	A
France	548,500	296,833	54.1	A
Germany	357,114	190,851	53.4	A
Liechtenstein	160	65	41.2	B
Macedonia	25,713	9,432	36.7	B
Czech R.	78,867	28,728	36.4	B
Croatia	56,594	17,591	31.1	B
Andorra	468	144	30.9	B
Poland	312,685	88,489	28.3	B
Austria	83,945	22,764	27.1	B
Denmark	43,094	11,610	26.9	B
Switzerland	41,277	10,803	26.2	B
Cyprus	9,251	2,422	26.2	B
Slovenia	20,273	5,240	25.8	B
Finland	338,419	86,569	25.6	B
Montenegro	13,812	2,916	21.1	B
Spain	505,992	106,131	21.0	B
Greece	131,957	27,181	20.6	B
Great Britain	242,900	46,366	19.1	B
Lithuania	65,300	12,108	18.5	B
Kosovo	10,908	2,004	18.4	B
Norway	386,224	61,706	16.0	B
Moldova	33,846	5,410	16.0	B
Malta	316	48	15.4	B
Hungary	93,028	14,198	15.3	B
Serbia	88,361	11,481	13.0	B
Bulgaria	110,879	14,362	12.9	B
Sweden	441,370	56,657	12.8	B
Italy	301,336	38,024	12.6	B
Ukraine	603,500	68,735	11.4	B
Belarus	207,600	22,968	11.1	B
Ireland	70,273	4,965	7.1	B
Portugal	92,090	3,919	4.3	B
Albania	28,748	897	3.1	B
Iceland	103,000	1,687	1.6	B

polygon features labelled with “Land-use” and “Natural” tags are filtered. While the features with “Natural” tag describe a wide variety of physical features, features with “Land-use” tag identify the land use features. These features are then merged together to create a uniform dataset.

3.2 Study area

The completeness indices are then arbitrarily categorized into two classes ranging between zero to hundred percent with 50 percent interval. To be more precise, while class “A” represents countries that completeness index exceeds 50 percent, class “B” identifies countries that less than half of them are mapped. According to this categorization, 8 countries place within the class “A” and 32 countries are classified as “B”. Belgium, Bosnia & Herzegovina, Germany, France, Luxemburg, the Netherlands, Romania, and Slovakia are those which are well-mapped. Spatial distribution of the mapped features within Europe is displayed in Figure 2. Green cells represent the contributed features regardless their attributes. It should be mentioned that the European countries have different populations and population densities, and physical characteristics and the completeness values should not be used for refereeing the topology of citizen participations in collaborative mapping practices [17]. For instance, Iceland with an area of 103,000 km^2 and nearly 300 thousand inhabitants is the least mapped country. This is not comparable with the Netherlands, holding an area of 41,500 km^2 and nearly 17 million inhabitants, corresponding to the best mapped country (82%). This inequality of public participation should be further investigated.

Huerta, Schade, Granell (Eds): Connecting a Digital Europe through Location and Place. Proceedings of the AGILE'2014 International Conference on Geographic Information Science, Castellón, June, 3-6, 2014. ISBN: 978-90-816960-4-3

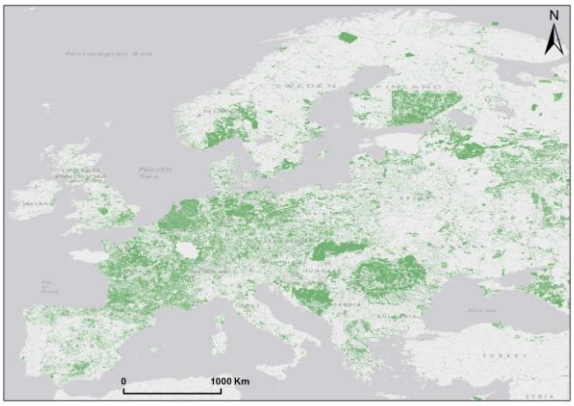

Figure 2: spatial distribution of contributed land use features in Europe

Conclusion

The contemporary emergence of citizen science projects, namely OSM, has drawn the attention of large number of citizens to share their information, as well as records of their GPS-enabled devices, with the public. This collaboratively collected information have been implemented in several applications such as navigation, context-aware routing, indoor mapping, and tourism recommendations. Exceptionally, OSM collects the land use features from contributors and therefore its potential for land use science has to be assessed.

This study aimed at assessing the completeness of land use features across European countries to find out how completely these features have been mapped. The calculated indices reveal that the degree of completeness is heterogeneous and ranges between 1 to 96 percent. More than half of 8 countries as listed in Table 1 are mapped in terms of land use features by OSM mappers. Apart from barely mapped countries, this means that volunteered mappers express their interest in mapping landscape related information as well and this opens avenues for further research towards harnessing CS for land use science. Future research directions should be conducted towards accuracy assessment of the land use attributes versus ground truth or proprietary datasets, e.g., the pan-European urban atlas and CORINE datasets.

As a final conclusion, the contributed OSM land use information suggest a promising alternative data source for land use mapping independent from applying computational image processing techniques. Whereas the degree of completeness in OSM increases over time, further contributions from volunteers should be expected within a short period of time. Further to this, the findings attempt to draw the attentions of volunteers to map the landscape-related objects as well so that citizen science could greatly contribute to collecting up-to-date information of our land resources. The following recommendations are suggested to environmentalists and land-use scientists that contributed features enable us to either consider the OSM features as an alternative data source or take advantage of the partially mapped areas for updating the existing and outdated inventories as outlined by [12]. It should be mentioned that applying data mining and data fusion techniques with other available features in OSM help to complete the incomplete areas.

References

[1] Bontemps, S., Defourny, P., Van Bogaert, E., Arino, O., Kalogirou, V., Ramos, P., Jose, J., (2011). GLOBCOVER 2009 Products description and validation report. Université catholique de Louvain (UCL) & European Space Agency (ESA), Vers. 2.2, pp 53.

[2] Buettner, G., Feranec, J., and Jaffrain, G., (2002). Corine land-cover update 2000. EEA. Technical Report, vol. 89, 17 December 2002. Copenhagen.

[3] Cihlar, J. and Jansen, L.J.M., (2001). From land-cover to land-use: a methodology for efficient land-use mapping over large areas. The Professional Geographer, 53 (2), 275–289.

[4] Comber, A., Brunsdon, C., See, L., Fritz, S., and McCallum, I. (2013). Comparing Expert and Non-expert Conceptualisations of the Land: An Analysis of Crowdsourced Land Cover Data. In T. Tenbrink, J. Stell, A. Galton, & Z. Wood (Eds.), Spatial Information Theory SE - 14 (Vol. 8116, pp. 243–260). Springer International Publishing.

[5] De Leeuw, J., Said, M., Ortegah, L., Nagda, S., Georgiadou, Y., & DeBlois, M. (2011). An assessment of the accuracy of volunteered road map production in Western Kenya. Remote Sensing, 3(2), 247-256.

[6] Devillers, R., and Jeansoulin, R., (2006). Fundamentals of Spatial Data Quality, ISTE, London.

[7] Devillers, R., Bédard, Y., Jeansoulin, R., and Moulin, B. (2007). Towards spatial data quality information analysis tools for experts assessing the fitness for use of spatial data, International Journal of Geographical Information Sciences, 21(3), 261–282.

[8] Estima, J., & Painho, M. (2013). Exploratory analysis of OpenStreetMap for land use classification. In Proceedings of the Second ACM SIGSPATIAL International Workshop on Crowdsourced and Volunteered Geographic Information (pp. 39-46).

[9] Foody, G.M., See, L., Fritz, S., Van der Velde, M., Perger, C., Schill, C., and Boyd, D. S., (2013). Assessing the Accuracy of Volunteered Geographic Information arising from Multiple Contributors to an Internet Based Collaborative Project. Transactions in GIS, 17(6), pp. 847–860.

[10] Friedl, M.A., McIver, D.K., Hodges, J.C., Zhang, X.Y., Muchoney, D., Strahler, A.H., and Schaaf, C., (2002). Global land cover mapping from MODIS: algorithms and early results. Remote Sensing of Environment, 83(1), 287-302.

[11] Fritz, S., Bartholomé, E., Belward, A., Hartley, A., Stibig, H. J., Eva, H., Mayaux, P., ... and Defourny, P. (2003). Harmonisation, mosaicing and production of the Global Land Cover 2000 database (Beta Version) (p. 41). Luxembourg: Office for Official Publications of the European Communities.

[12] Fritz, S., McCallum, I., Schill, C., Perger, C., Grillmayer, R., Achard, F., Kraxner, F., Obersteiner, M., (2009). Geo-Wiki.Org: The Use of Crowdsourcing to Improve Global Land Cover. Remote Sensing, 1, 345-354.

[13] Fritz, S., Mccallum, I., Schill, C., Perger, C., See, L., Schepaschenko, D., … Obersteiner, M. (2012). Environmental Modelling & Software Geo-Wiki: An online platform for improving global land cover. Environmental Modelling and Software, 31, 110–123.

[14] Gervais, M., Bédard, Y., Levesque, M., Bernier, E., and Devillers, R., (2009). Data Quality Issues and Geographic Knowledge Discovery, 99–116.

[15] Hagenauer, J., and Helbich, M., (2012). Mining urban land-use patterns from volunteered geographic information by means of genetic algorithms and artificial neural networks. International Journal of Geographic Information Science, 26 (6), 963–982.

[16] Haklay, M. (2010). How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. Environment and Planning B: Planning and Design, 37(4), 682–703.

[17] Haklay, M., (2013), Citizen Science and Volunteered Geographic Information – overview and typology of participation in Sui, D.Z., Elwood, S. and M.F. Goodchild (eds.), Crowdsourcing Geographic Knowledge. Berlin: Springer. pp. 105-122

[18] Hecht, R., Kunze, C., and Hahmann, S., (2013). Measuring Completeness of Building Footprints in OpenStreetMap over Space and Time. ISPRS International Journal of Geo-Information, 2(4), 1066–1091

[19] Herold, M., Mayaux, P., Woodcock, C. E., Baccini, A., & Schmullius, C., (2008). Some challenges in global land-cover mapping: An assessment of agreement and accuracy in existing 1 km datasets, Remote Sensing of Environment 112(5), 2538–2556.

[20] Jokar Arsanjani, J., Helbich, M., Bakillah, M., Hagenauer, J., and Zipf, A., (2013a). Toward mapping land-use patterns from volunteered geographic information, International Journal of GIS, 27(12), 2264-2278.

[21] Jokar Arsanjani, J., Helbich, M., Loos, L., Bakillah, M., (2014:in-press). The emergence and evolution of OpenStreetMap: A cellular automata approach, International Journal of Digital Earth.

[22] Kandrika, S. and Roy, P.S., (2008). Land-use land-cover classification of Orissa using multi-temporal IRS-P6 AWIFS data: a decision tree approach. International Journal of Applied Earth Observation and Geoinformation, 10 (2), 186–193.

[23] Kasetkasem, T., Arora, M. K., & Varshney, P. K. (2005). Super-resolution land cover mapping using a Markov random field based approach. Remote Sensing of Environment, 96(3–4), 302–314.

[24] Koukoletsos, T., Haklay, M., and Ellul, C., (2012). Assessing data completeness of VGI through an automated matching procedure for linear data. Transactions in GIS, 16 (4), 477–498.

[25] Mayaux, P., Eva, H., Gallego, J., Strahler, A. H., Herold, M., Member, S., Agrawal, S., (2006). Validation of the Global Land-cover 2000 Map, IEEE Transactions on Geoscience and Remote Sensing, 44(7), 1728–1739.

[26] Paneque-Gálvez, J., Mas, J.-F., Moré, G., Cristóbal, J., Orta-Martínez, M., Luz, A. C., Reyes-García, V., (2013). Enhanced land use/cover classification of heterogeneous tropical landscapes using support vector machines and textural homogeneity. International Journal of Applied Earth Observation and Geoinformation, 23, 372–383.

[27] Pontius, R.G. Jr., and Petrova, S.H., (2010). Assessing a predictive model of land change using uncertain data. Environmental Modelling & Software, 25 (3), 299–309.

[28] Qi, Z., Yeh, A. G. O., Li, X., and Lin, Z., (2012). A novel algorithm for land use and land cover classification using RADARSAT-2 polarimetric SAR data. Remote Sensing of Environment, 118, 21-39.

[29] Robinson, D.T. and Brown, D.G., (2009). Evaluating the effects of land-use development policies on ex-urban forest cover: an integrated agent-based GIS approach. International Journal of Geographical Information Science, 23 (9), 1211–1232.

[30] Saadat, H., Adamowski, J., Bonnell, R., Sharifi, F., Namdar, M., and Ale-Ebrahim, S. (2011). Land use and land cover classification over a large area in Iran based on single date analysis of satellite imagery. ISPRS Journal of Photogrammetry and Remote Sensing, 66(5), 608–619.

[31] Seifert, F., (2009). Improving Urban Monitoring toward a European Urban Atlas. In Global Mapping of Human Settlement. CRC Press.

[32] Sexton, J. O., Urban, D. L., Donohue, M. J., and Song, C. (2013). Long-term land cover dynamics by multi-temporal classification across the Landsat-5 record. Remote Sensing of Environment, 128, 246–258.

[33] Strahler, A. H., Boschetti, L., Foody, G. M., Friedl, M. A., Hansen, M. C., Herold, M., Mayaux, P., (2006). Global Land-cover Validation: Recommendations for Evaluation and Accuracy Assessment of Global Land-cover Maps. Luxemburg : Office for Official Publications of the European Communities, (25).

[34] Thenkabail, P.S., Schull, M., and Turral, H., (2005). Ganges and Indus river basin land use/land cover (LULC) and irrigated area mapping using continuous streams of MODIS data. Remote Sensing of Environment, 95(3), 317–341.