Potential and limitation of using OSM for the creation/validation of Land Use Land Cover (LULC) maps

Cidália Costa Fonte

Department of Mathematics – University of Coimbra, Coimbra, Portugal

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

28-30 July 2018 – Milan - Italy

Summary

- Introduction
- Ways to use OSM for LULC mapping
  - Convert OSM data into LULC maps
  - Use OSM data to validate LULC maps
  - Use OSM data to train classifiers to create LULC maps from satellite images
- Conclusions
  - Opportunities/Limitations/Future developments

Introduction

- Main aim of the work under development
  - Add value to available Volunteered Geographic Information (VGI) by either:
    - processing the existing data
    - integrating diverse sources of data

- This work started within COST actions
  - TD1202 (Mapping and the citizen sensor)
  - IC1203 (European Network Exploring Research into Geospatial Information Crowdsourcing: software and methodologies for harnessing geographic information from the crowd - ENERGIC).

(COST EU-funded programme - enables researchers to set up interdisciplinary research networks in Europe and beyond)

Introduction

- Volunteered Geographic Information (VGI)
  - Many types of data
  - Wide variety of projects with very diverse objectives
  - Enormous amounts of data
  - Some enable data download / data accessible by APIs.

Panoramio

YouTube

Introduction

OSM4LULC

How can OSM contribute to the creation / validation of LULC maps?

Introduction

- Production of LULC maps requires
  - The classification of images
  - When supervised classifiers are used training sets are needed
- The created maps need to be validated
  - The quality assessment usually requires reference data

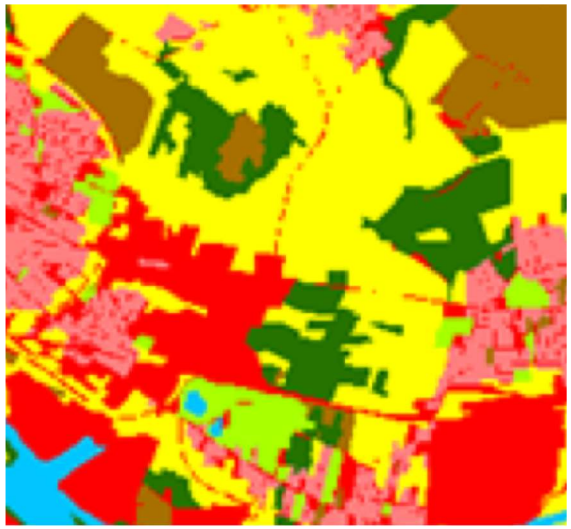
OSM conversion to LULC maps

- OSM may assist the creation and validation of Land Use Land Cover (LULC) maps by
  - Direct creation of LULC maps from OSM
  - Generating reference databases for validation
  - Generating training sets

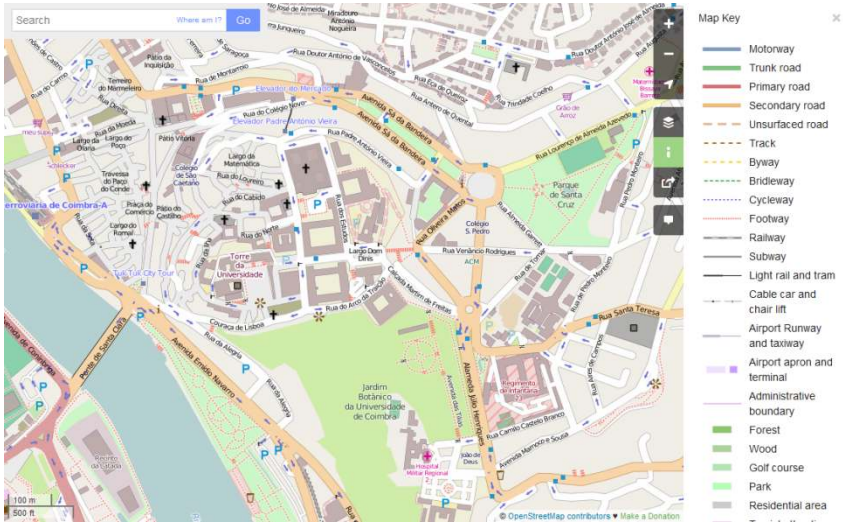
Speed + lower costs

OSM conversion to LULC maps

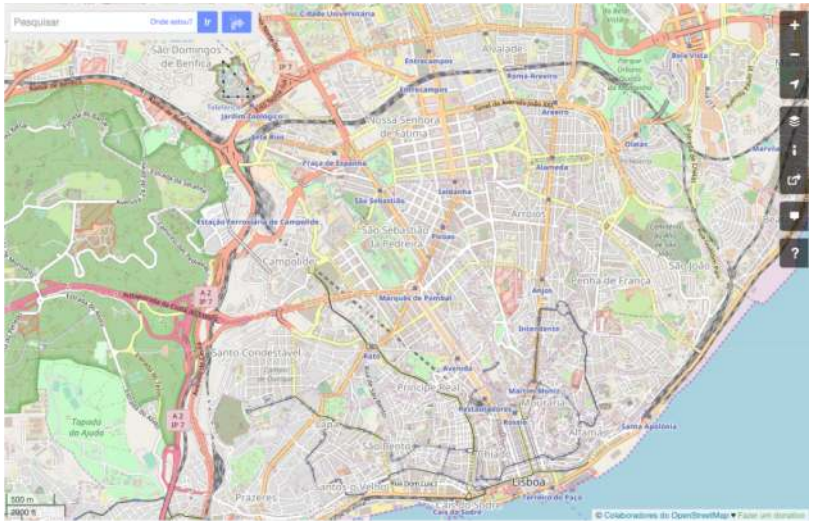
- Web-application to convert automatically the data available in OSM into a LULC map



OSM conversion to LULC maps



OSM – Coimbra (Portugal)



OSM – Lisbon (Portugal)

OSM – Praia (Cape Verde)

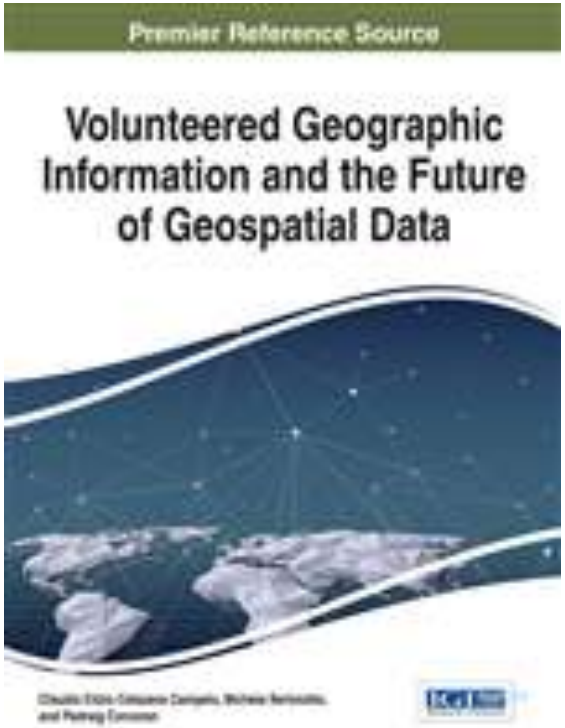
OSM – Milan (Italy)

OSM conversion to LULC maps

- OpenStreetMap (OSM) (http://www.openstreetmap.org/)

- Geospatial entities available in OSM
  - http://wiki.openstreetmap.org/wiki/Map_Features

OSM conversion to LULC maps



Chapter 7

Using OpenStreetMap to Create Land Use and Land Cover Maps: Development of an Application

Cidália Costa Fonte
University of Coimbra, Portugal & INESC Coimbra, Portugal

Joaquim António Patriarca
INESC Coimbra, Portugal

Marco Minghini
Politecnico di Milano, Italy

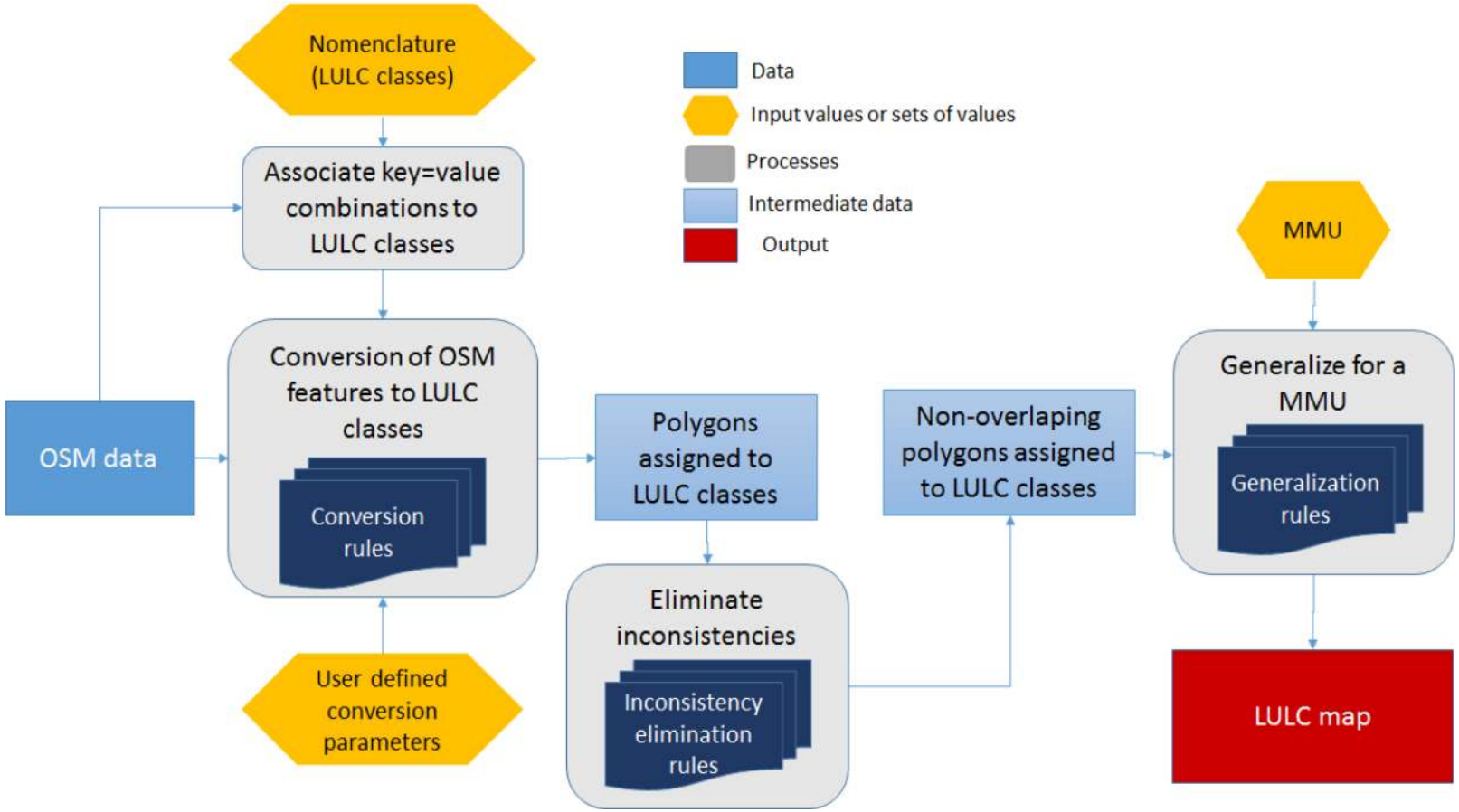
Vyron Antoniou
Hellenic Military Geographical Service, Greece

Linda See
International Institute for Applied Systems Analysis, Austria

Maria Antonia Brovelli
Politecnico di Milano, Italy

OSM conversion to LULC maps

Methodology of the tool created for the conversion of OSM into LULCM



OSM conversion to LULC maps

- Nomenclatures
  - Urban Atlas (UA) – European product
    - Global Monitoring for Environment and Security Urban Atlas
      - Detailed classification of LULC of the European cities with more than 100 K inhabitants + some cities with more than 50 k inhabitants since 2012
      - 12 thematic classes
  - Corine Land Cover (CLC) – European product
    - LULC classification of Europe
      - Minimum mapping Unit of 25 ha
      - 44 thematic classes
  - GlobeLand 30 (GL30)
    - Map produced by the “National Geomatic Center of China” from Landsat imagery
      - Global coverage
      - Raster format – spatial resolution of 30m
      - 10 thematic classes

OSM conversion to LULC maps

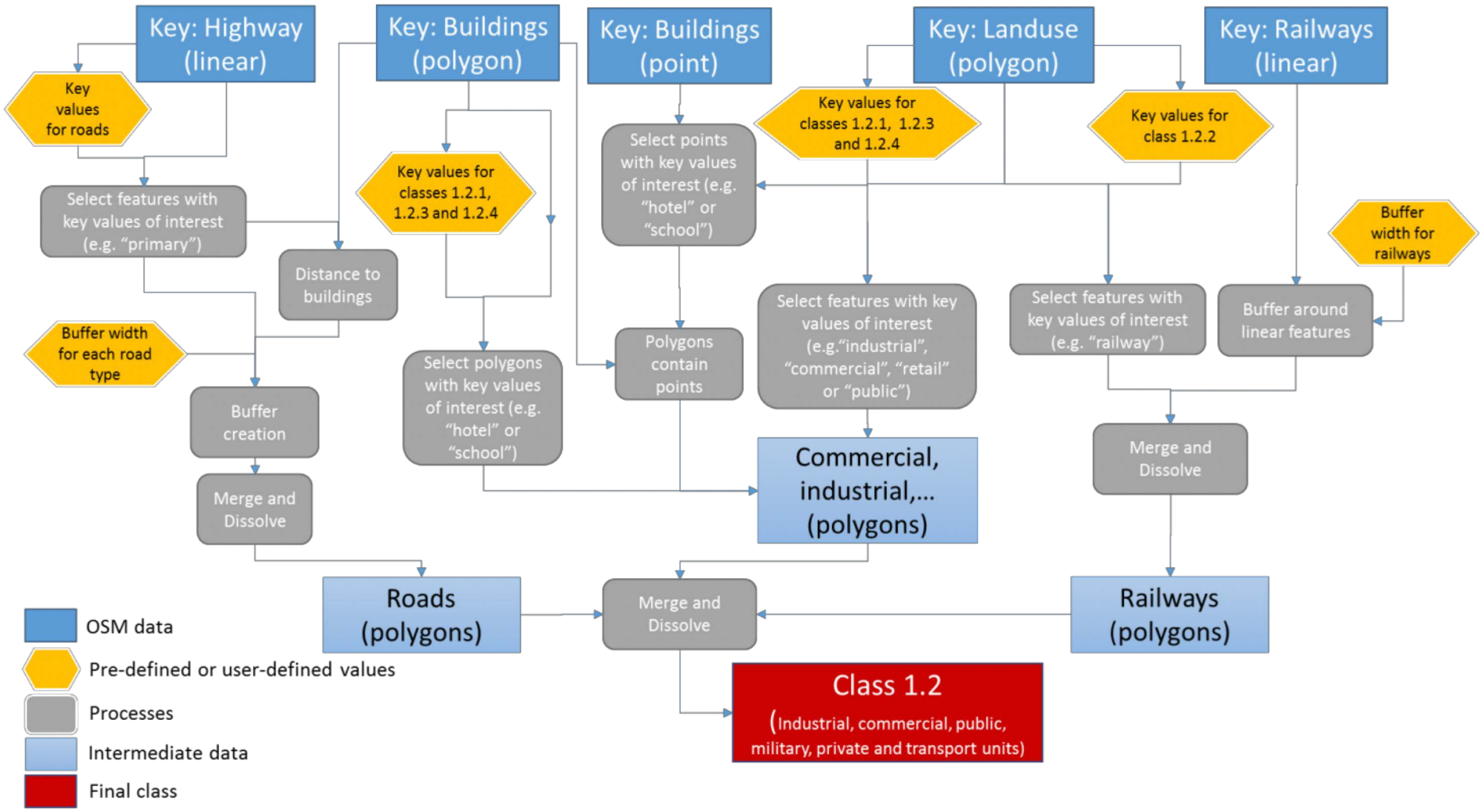
UA and CLC nomenclatures

Urban Altas (UA)			CORINE Land Cover (CLC)		
Nível1	Nível 2	Nível 3	Nível 1	Nível 2	Nível 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.Artificial Surfaces	1.1 Urban Fabric	1.1.1 Continuous urban fabric 1.1.2 Discontinuous urban fabric
	1.2 Industrial, commercial, public, military, private and transport units	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		1.2 Industrial, commercial, public, military, private and transport units	1.2.1 Industrial or commercial 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 constructio sites	1.3.1 Mineral extraction and dump sites 1.3.3 Construction sites 1.3.4 Land without current use		1.3 Mine, dump and construction sites	1.3.1 Mineral extraction 1.3.2 Dump sites 1.3.3 Construction sites
	1.4 Artificial non-agricultural vegetated areas	1.4.1 Green urban areas 1.4.2 Sports and leisure facilities		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			2.Agricultural areas	2.1 Arable land	2.1.1 Non-irrigated arable land 2.1.2 Permanently irrigated land 2.1.3 Rice fields
				2.2 Permanent crops	2.2.1 Vineyards 2.2.2 Fruit trees and berry plantations 2.2.3 Olive groves
				2.3 Pastures	2.3.1 Pastures
				2.4 Heterogeneous agricultural areas	2.4.1 Annual crops associated with permanent crops 2.4.2 Complex cultivation patterns 2.4.3 Land principally occupied by agriculture, with significant areas of natural vegetation 2.4.4 Agro-forestry areas
3. Forests			3. Forest and semi natural areas	3.1 Forests	3.1.1 Broad-leaved forest 3.1.2 Coniferous forest 3.1.3 Mixed forest
				3.2 Scrub and/or herbaceous vegetation associations	3.2.1 Natural grasslands 3.2.2 Moors and heathland 3.2.3 Sclerophyllous vegetation 3.2.4 Transitional woodland-shrub
				3.3 Open spaces with little or no vegetation	3.3.1 Beaches, dunes, sands 3.3.2 Bare rocks 3.3.3 Sparsely vegetated areas 3.3.4 Burnt areas 3.3.5 Glaciers and perpetual snow
-------			4. Wetlands	4.1 Inland wetlands	4.1.1 Inland marshes 4.1.2 Peat bogs
				4.2 Maritime wetlands	4.2.1 Salt marshes 4.2.2 Salines 4.2.3 Intertidal flats
5. Water			5. Water	5.1 Inland waters	5.1.1 Water courses 5.1.2 Water bodies
				5.2 Marine waters	5.2.1 Coastal lagoons 5.2.2 Estuaries 5.2.3 Sea and ocean

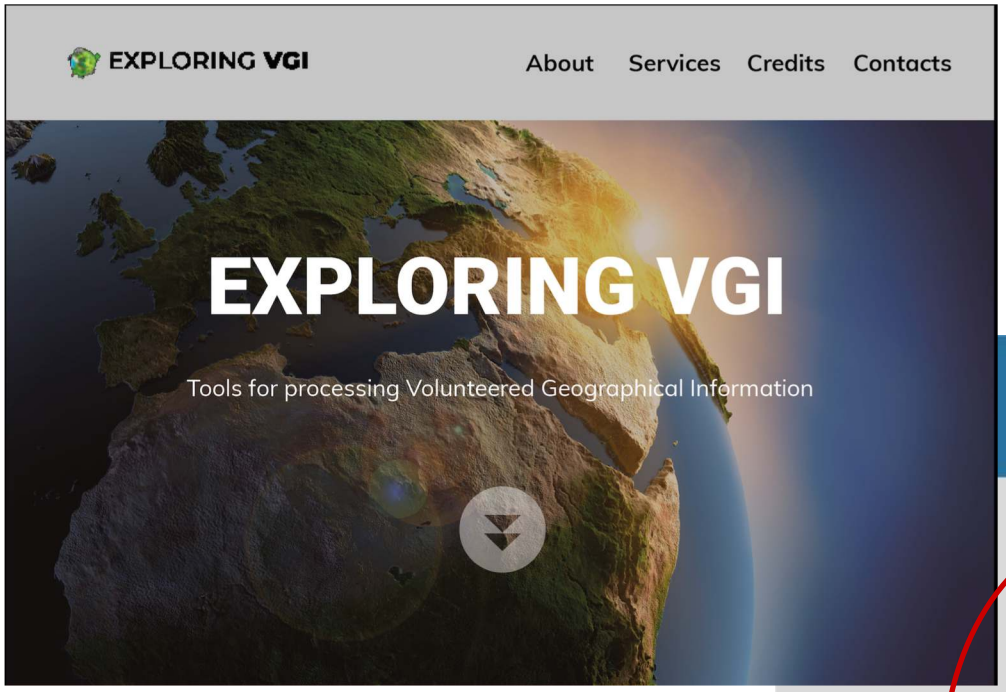
OSM conversion to LULC maps
GlobeLand30 (GL30) nomenclature

Class Code	Class Name	Class Description	Minimum Mapping Unit (km^2)
10	Cultivated land (CL)	Arable land (cropland): dry land, paddy field, land for greenhouses, vegetable fields, artificial tame pastures, economic cropland in which shrub crops or herbaceous crops are planted, and land abandoned with the reclamation of arable land	0.0324
20	Forest (F)	Broadleaved deciduous forest, evergreen broad-leaf forest, deciduous coniferous forest, evergreen coniferous forest, mixed broadleaf-conifer forest	0.0576
30	Grassland (GL)	Typical grassland, meadow grassland, alpine grassland, desert grassland, grass	0.09
40	Shrubland (SL)	Desert scrub, mountain scrub, deciduous and evergreen shrubs	0.09
50	Wetland (WL)	Lake swamp, river flooding wetlands, seamarsh, shrub/forest wetlands, mangrove forest, tidal flats/salt marshes	0.0729
60	Water bodies (WB)	Open water: lakes, reservoirs/fishponds, rivers	0.0009 (Rivers) 0.0081 (Lakes)
70	Tundra (T)	Brush tundra, poaceae tundra, wet tundra, bare tundra, mixed tundra	Not provided
80	Artificial surfaces (AS)	Settlement place, industrial and mining area, traffic facilities	0.0144
90	Bareland (BL)	Saline-alkali land, sand, gravel, rock, microbiotic crust	0.0324
100	Permanent snow/ice (SI)	Permanent snow, ice sheet and glacier	0.0081

OSM conversion to LULC maps

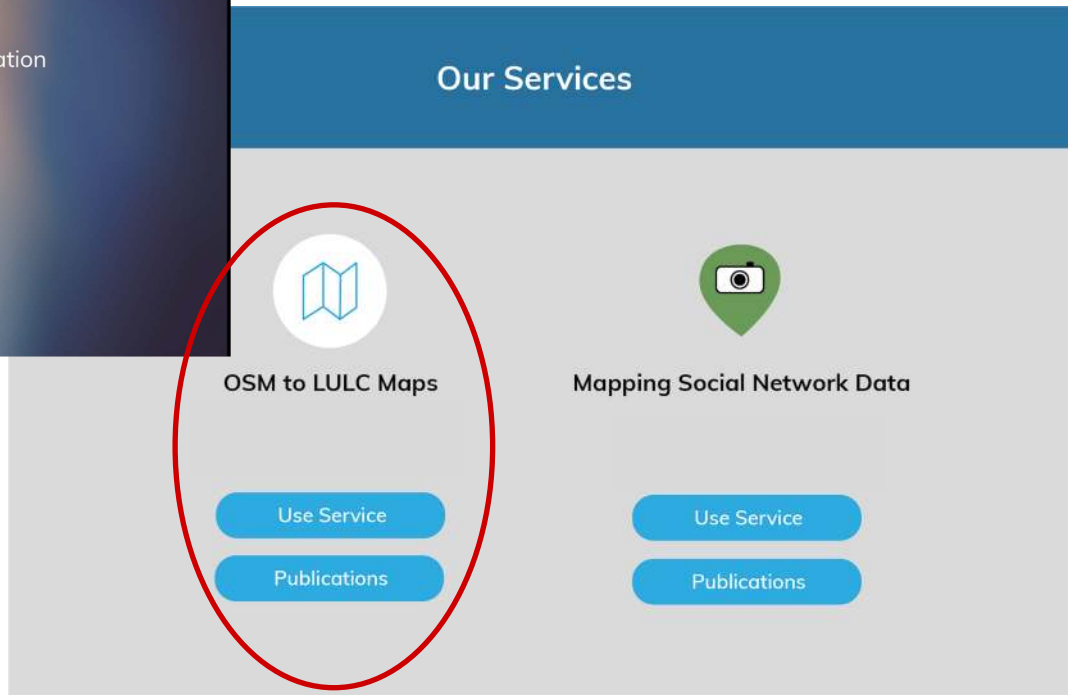


OSM conversion to LULC maps



Open source software:

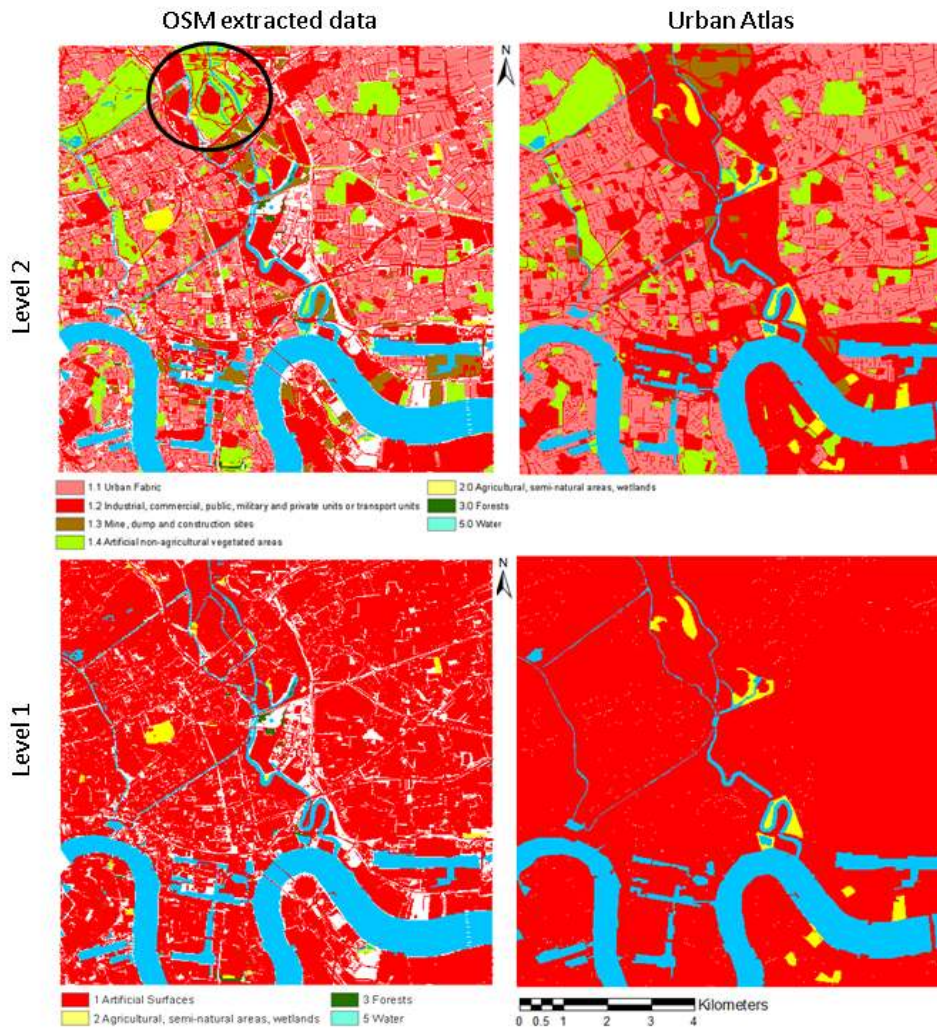
(Django, Apache, Tweepy, GDAL/OGR, Grass GIS. PostgreSQL, Angular JS, Leaflet, ...)

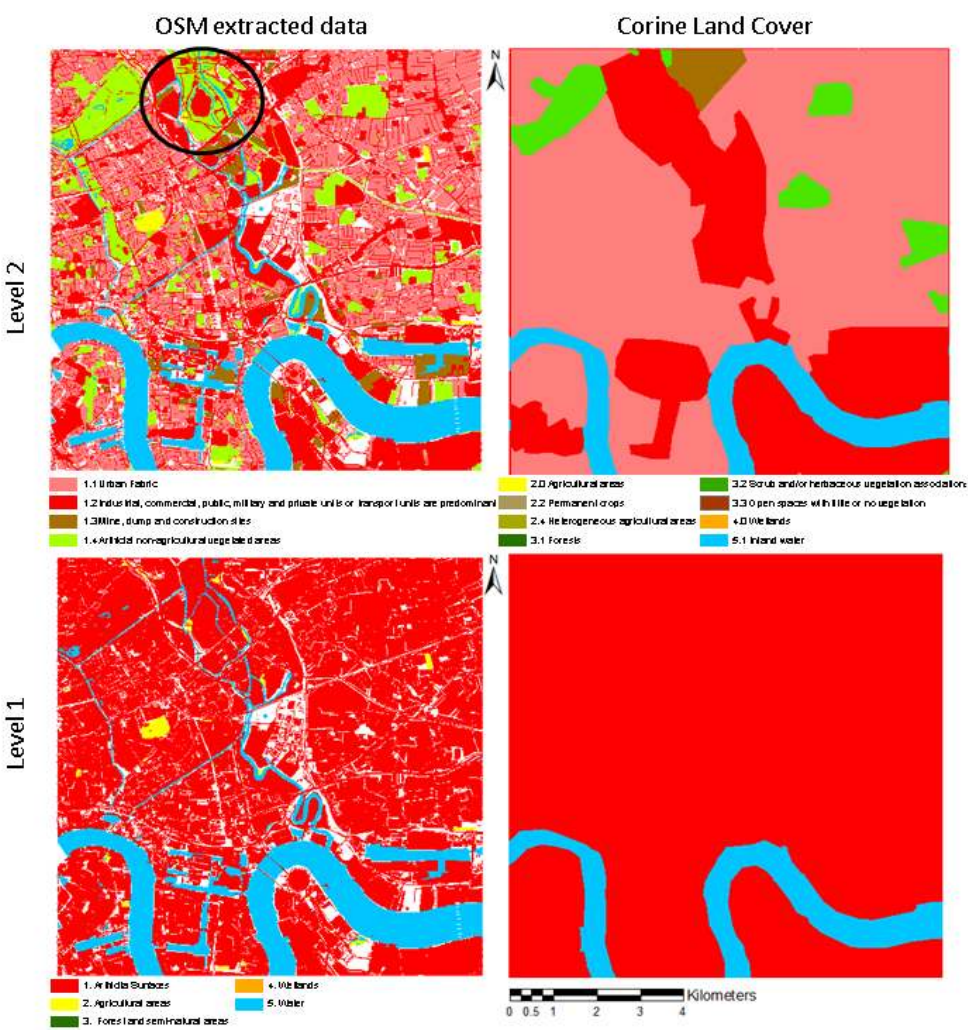


vgi.uc.pt

OSM to LULC maps

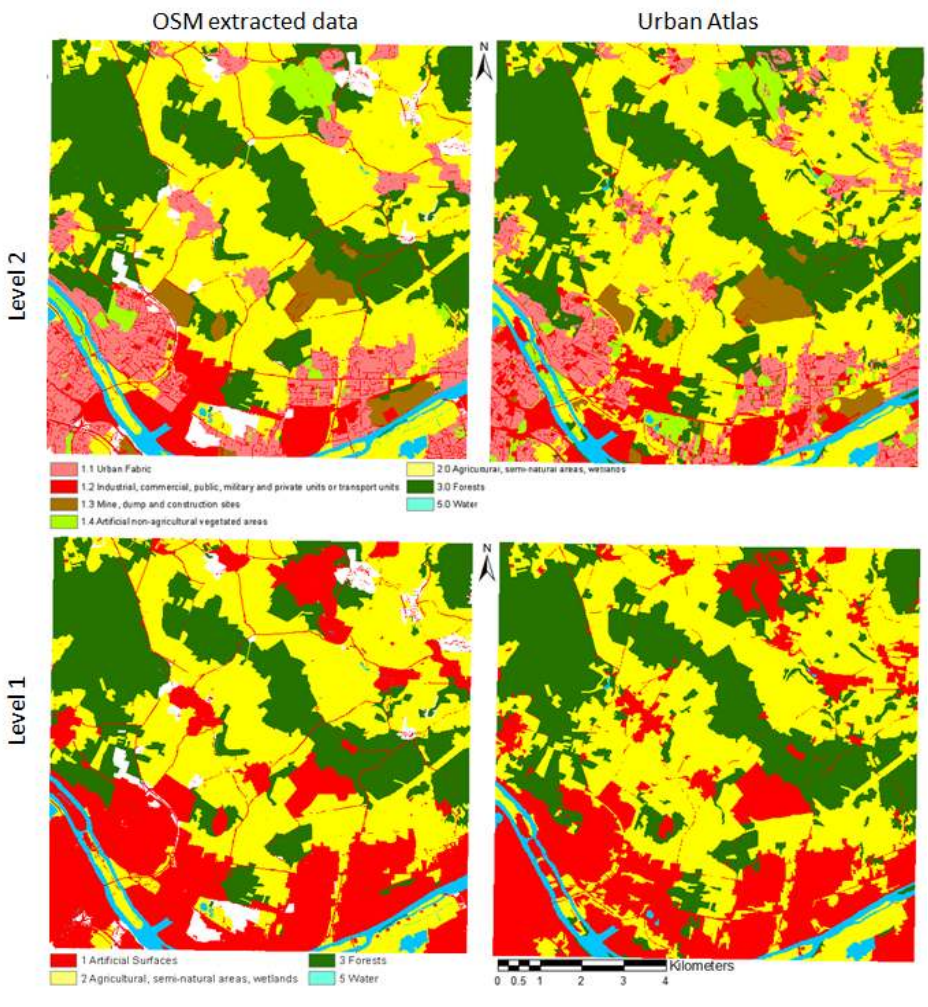
London

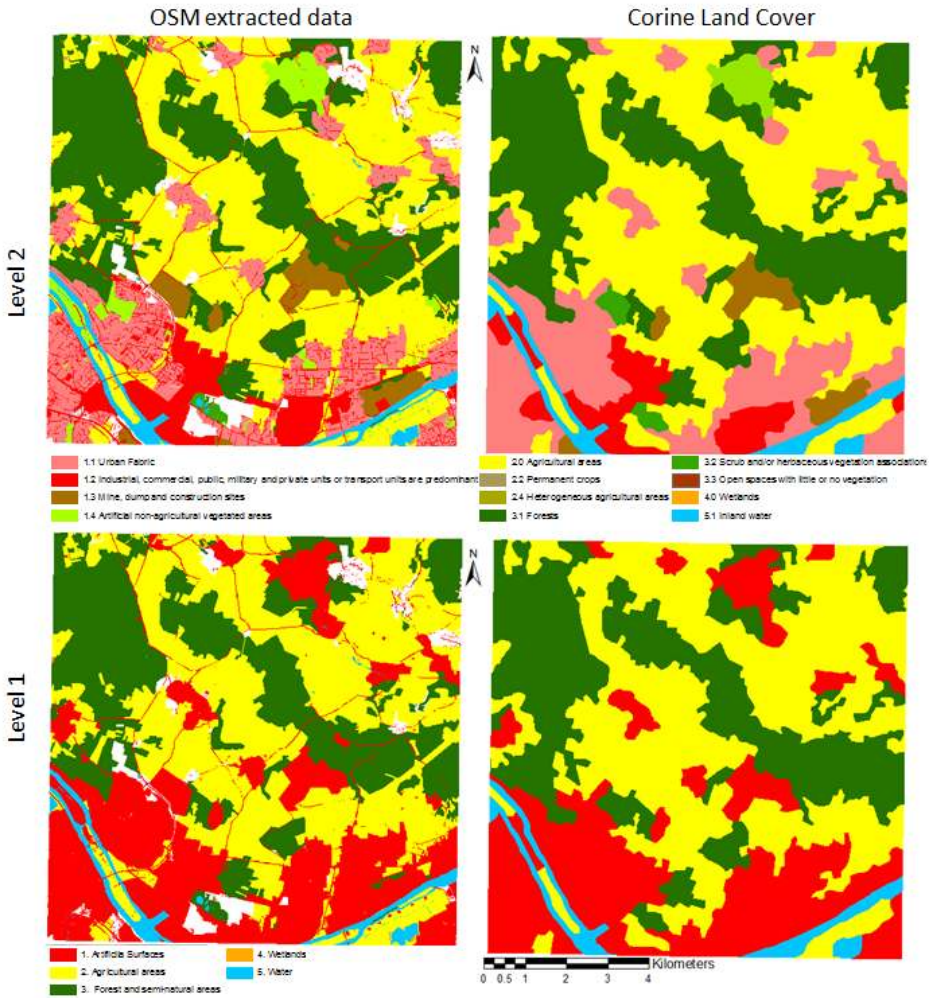




OSM to LULC maps

Paris





OSM conversion to LULC maps

isprs International Journal of *Geo-Information*

MDPI

Article

Generating Up-to-Date and Detailed Land Use and Land Cover Maps Using OpenStreetMap and GlobeLand30

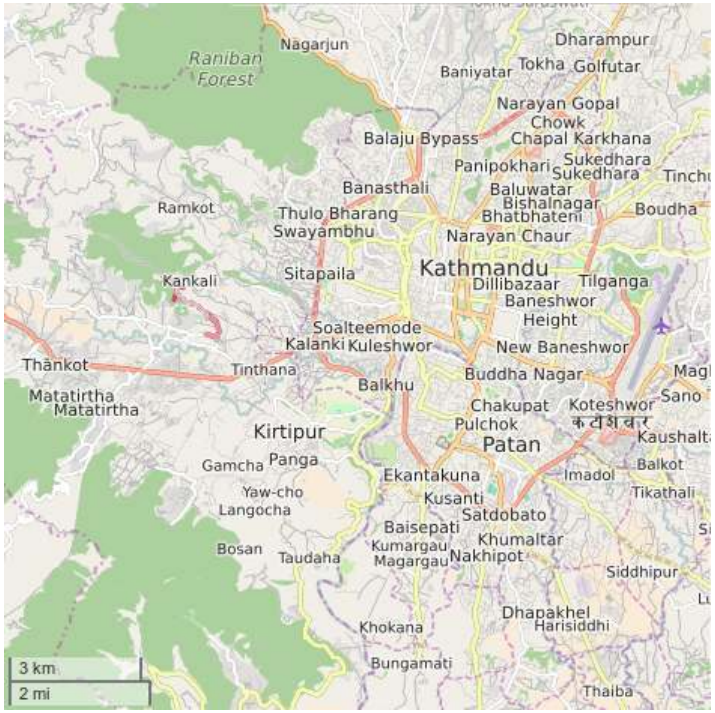
Cidália Costa Fonte [1,2,*], Marco Minghini [3], Joaquim Patriarca [2], Vyron Antoniou [4,5], Linda See [6] and Andriani Skopeliti [7]

1 Department of Mathematics, University of Coimbra, Largo D. Dinis, 3001-501 Coimbra, Portugal
2 INESC Coimbra, Rua Sílvio Lima, Pólo II, 3030-290 Coimbra, Portugal; joaquimapatriarca@gmail.com
3 Department of Civil and Environmental Engineering, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milan, Italy; marco.minghini@polimi.it
4 Hellenic Military Academy, Leof. Varis—Koropiou, 16673 Vari, Greece; v.antoniou@ucl.ac.uk
5 Hellenic Military Geographical Service, 4, Evelpidon Str., 11362 Athens, Greece
6 International Institute for Applied Systems Analysis (IIASA), Schlossplatz 1, A2361 Laxenburg, Austria; see@iiasa.ac.at
7 School of Rural and Surveying Engineering, National Technical University of Athens, 9 H. Polytechniou, 15780 Zografou, Greece; askop@survey.ntua.gr
* Correspondence: cfonte@mat.uc.pt; Tel.: +351-239-791-150

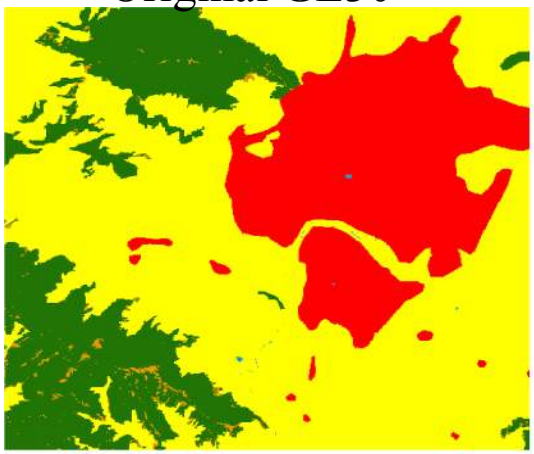
Academic Editor: Wolfgang Kainz
Received: 4 March 2017; Accepted: 17 April 2017; Published: 22 April 2017

OSM conversion to LULC maps

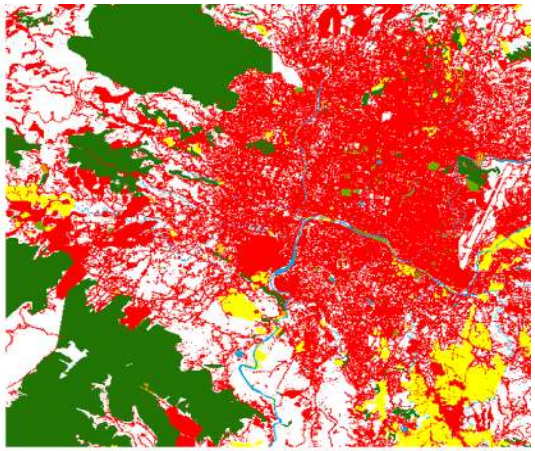
Kathmandu

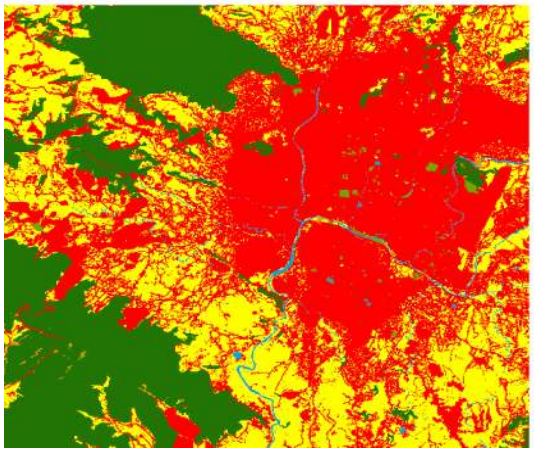


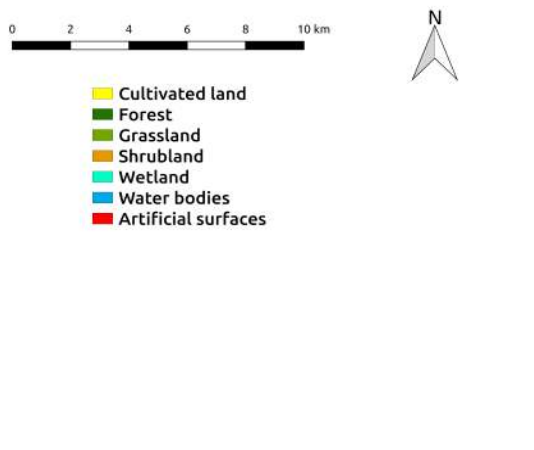
Original GL30



LULCM derived from OSM





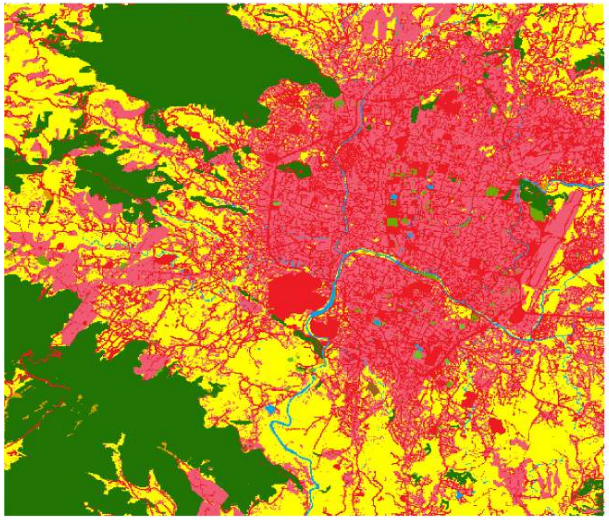


GL30 updated

OSM conversion to LULC maps

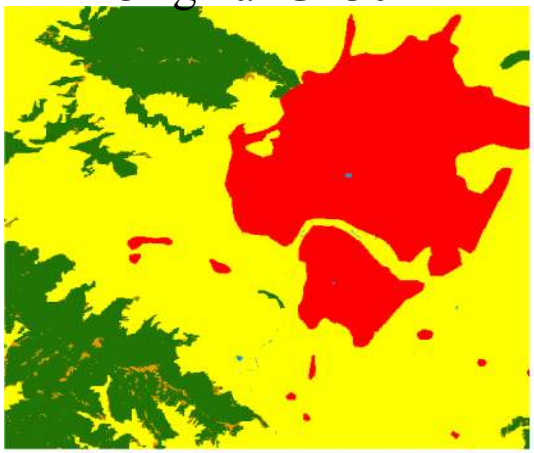
Kathmandu

GL30 detailed

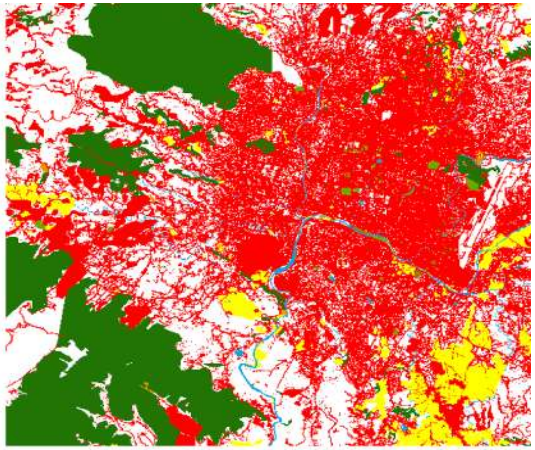


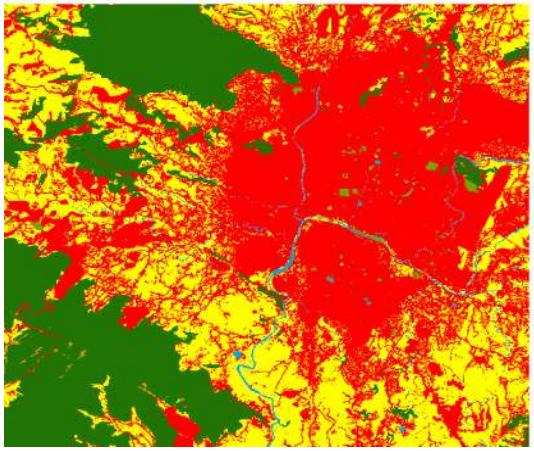
- Urban Fabric
- Industrial, commercial, public, military, private and transport units
- Mine, dump and construction sites
- Artificial non-agricultural vegetated areas
- Cultivated land
- Forest
- Water bodies
- Grassland
- Shrubland
- Wetland

Original GL30

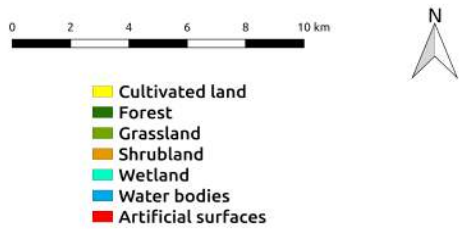


LULCM derived from OSM





GL30 updated

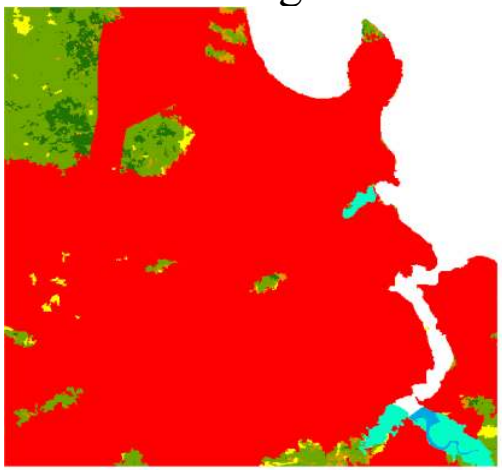


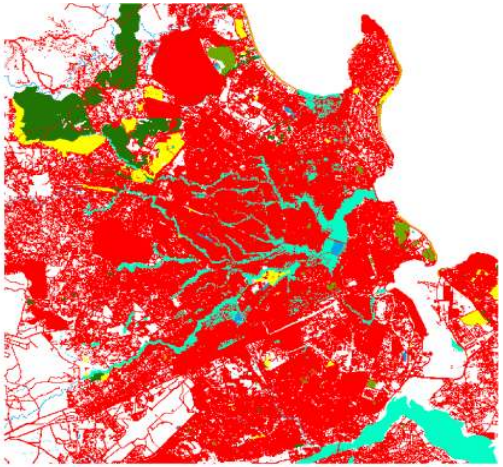
OSM conversion to LULC maps

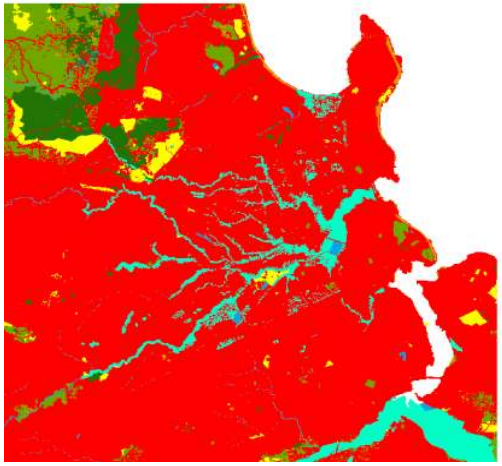
Dar es Salaam

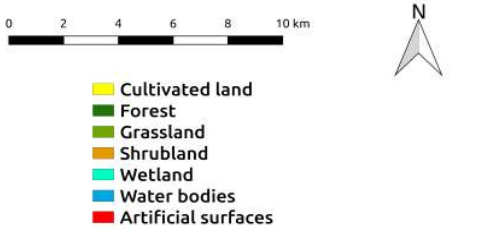
GL30 Original

LULCM derived from OSM







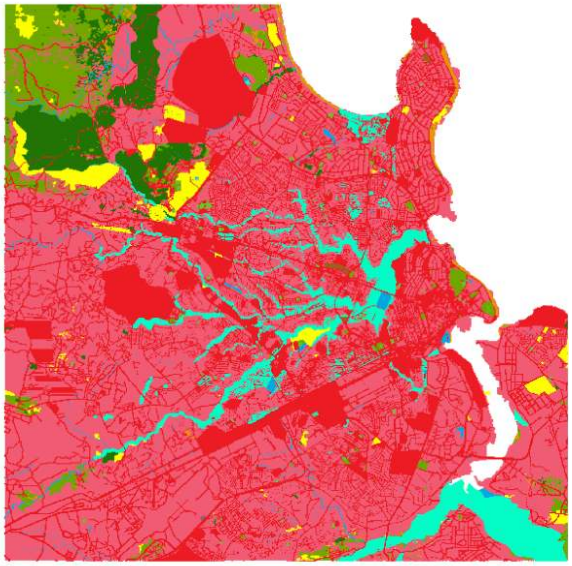


GL30 updated

OSM conversion to LULC maps

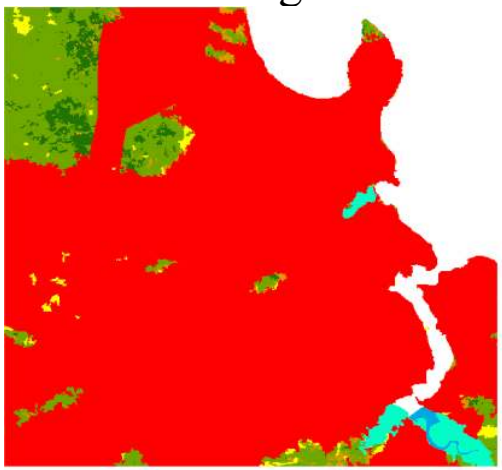
Dar es Salaam

GL30 detailed

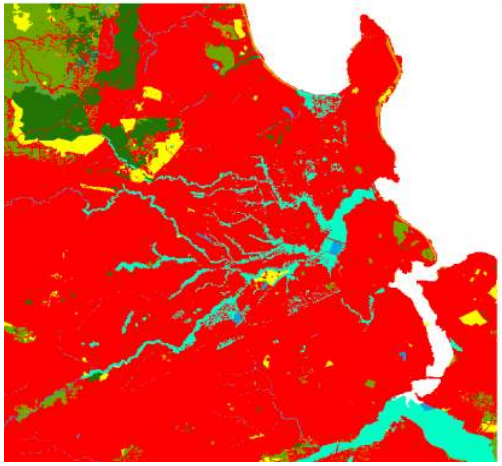


Urban Fabric
Industrial, commercial, public, military, private and transport units
Mine, dump and construction sites
Artificial non-agricultural vegetated areas
Cultivated land
Forest
Water bodies
Grassland
Shrubland
Wetland

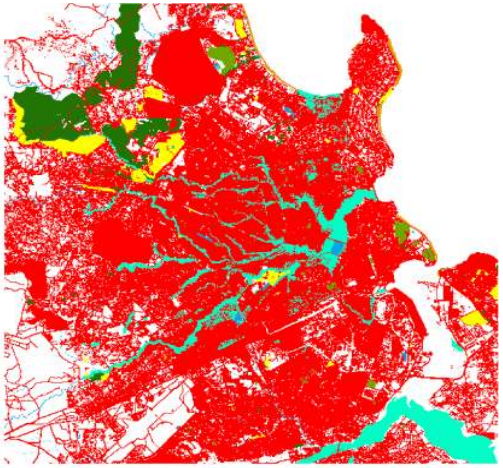
GL30 Original



GL30 updated

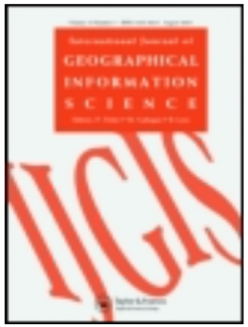


LULCM derived from OSM



Cultivated land
Forest
Grassland
Shrubland
Wetland
Water bodies
Artificial surfaces

LULC map validation with OSM



Taylor & Francis
Taylor & Francis Group

International Journal of Geographical Information Science

ISSN: 1365-8816 (Print) 1362-3087 (Online) Journal homepage: http://www.tandfonline.com/loi/tgis20

Assessing the applicability of OpenStreetMap data to assist the validation of land use/land cover maps

Cidália C. Fonte & Nuno Martinho

To cite this article: Cidália C. Fonte & Nuno Martinho (2017): Assessing the applicability of OpenStreetMap data to assist the validation of land use/land cover maps, International Journal of Geographical Information Science, DOI: 10.1080/13658816.2017.1358814

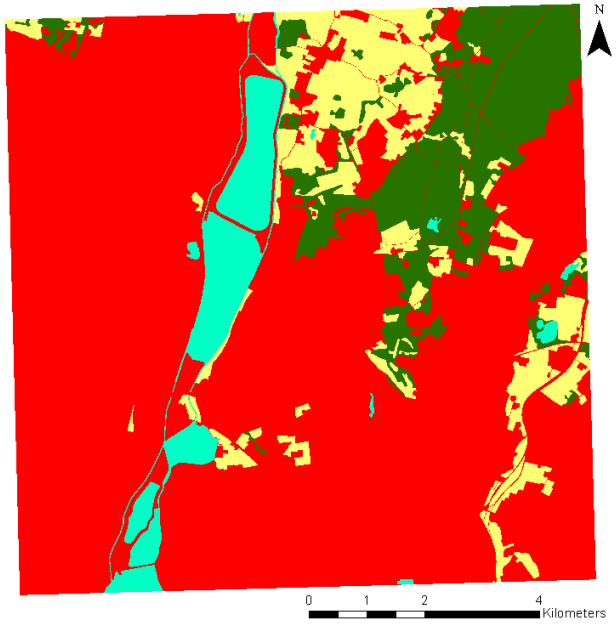
To link to this article: http://dx.doi.org/10.1080/13658816.2017.1358814

LULC map validation with OSM

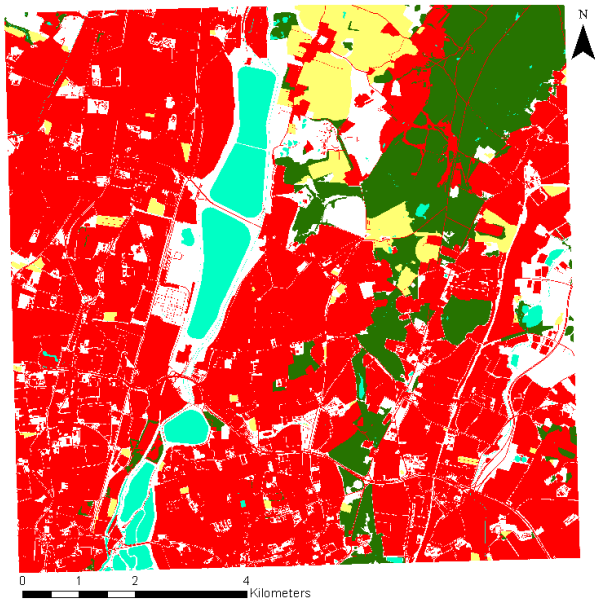
- Possible approaches
  - Direct comparison with a reference map
  - Comparison with a reference sample
    - Creation of a random sample of points stratified per class
    - Creation of the reference data with:
      - Photointerpretation (PI)
      - Data extracted automatically from OSM + PI
  - Accuracy assessment of UA with both reference databases

LULC map validation with OSM

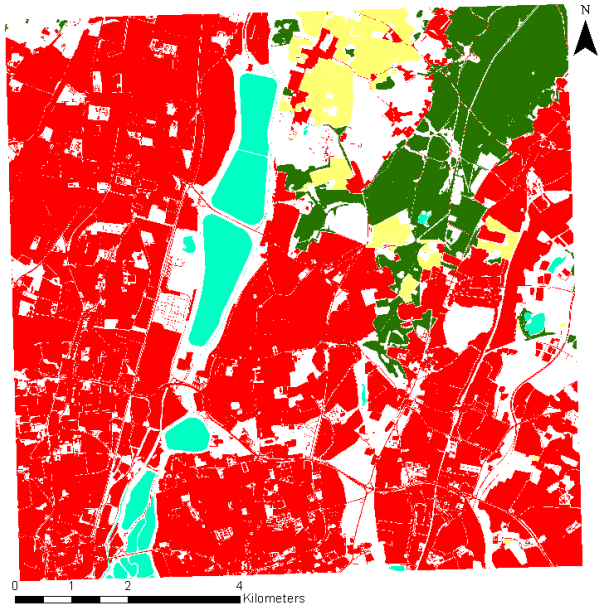
Urban Atlas



OSM extracted LULC



Agreement



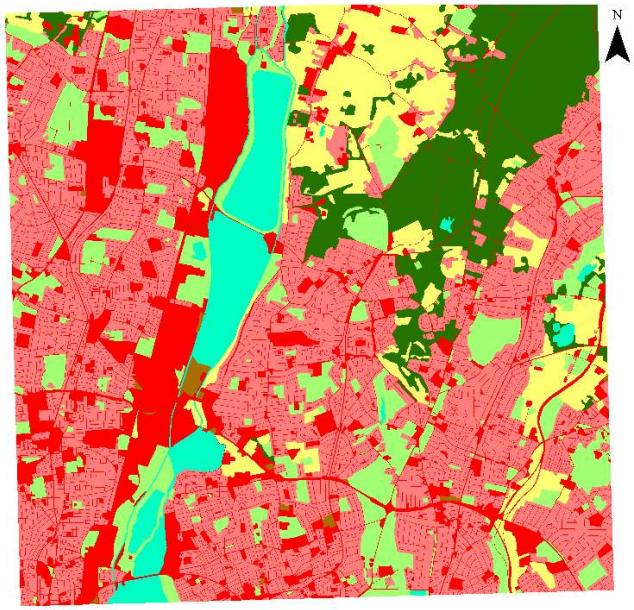
London

Urban Atlas nomenclature – Level 1

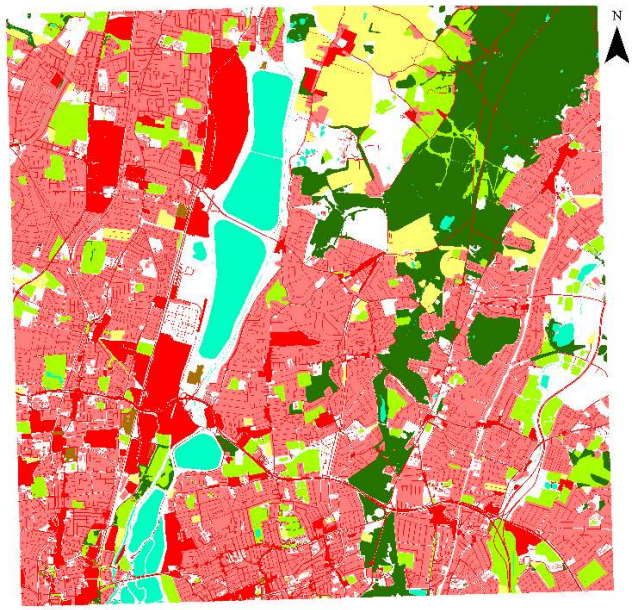
- 1 Artificial Surfaces
- 2 Agricultural, semi-natural areas, wetlands
- 3 Forests
- 5 Water

LULC map validation with OSM

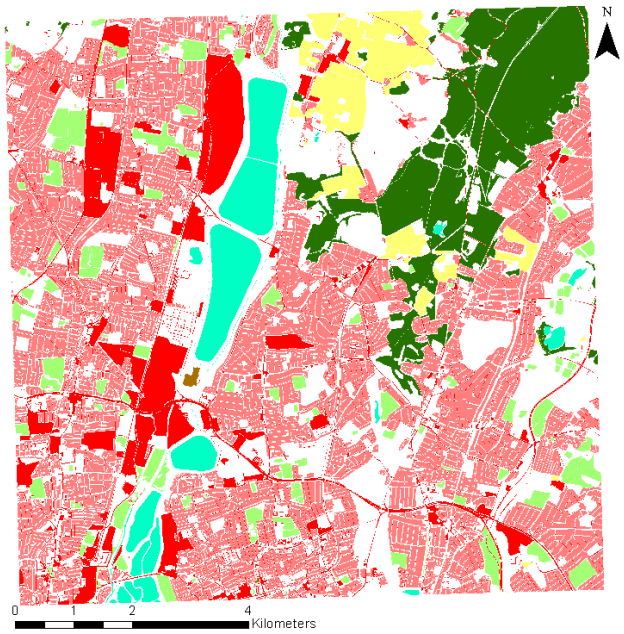
Urban Atlas



OSM extracted LULC



Agreement



London

Urban Atlas nomenclature – Level 2

- 1.1 Urban Fabric
- 1.2 Industrial, commercial, public, military and private units or transport units
- 1.3 Mine, dump and construction sites
- 1.4 Artificial non-agricultural vegetated areas
- 2.0 Agricultural, semi-natural areas, wetlands
- 3.0 Forests
- 5.0 Water

LULC map validation with OSM

	Class	UA	OSM	No data in OSM	NoOSM /UA	NoOSM /Area$_{NoOSM}$
		(Area in km^2)			(%)	(%)
Level 1	1	72.84	60.24	11.52	16	67
	2	9.99	4.83	3.62	36	21
	3	11.46	12.91	0.87	8	5
	5	5.66	4.87	1.12	20	7
Level 2	1.1	39.51	36.15	2.71	7	16
	1.2	20.28	15.82	4.79	24	28
	1.3	0.38	0.26	0.19	50	1
	1.4	12.68	8.01	3.83	30	22
	2.0	9.99	4.83	3.62	36	21
	3.0	11.46	12.91	0.87	8	5
	5.0	5.66	4.87	1.12	20	7

NoOSM/UA - Percentage of area per class with no data in OSM relative to class area in UA

NoOSM/Area$_{NoOSM}$ - Percentage of area per class with no data in OSM relative relative to the total area of the regions with no data in OSM

LULC map validation with OSM

		UA versus OSM		
	Class	MPA_{OvOSM} (%)	MPA_{OvUA} (%)	TPA_{Ov} (%)
Level 1	1	95	93	90
	2	71	54	
	3	74	90	
	5	92	98	
Level 2	1.1	86	84	76
	1.2	63	64	
	1.3	24	33	
	1.4	59	53	
	2.0	71	54	
	3.0	74	90	
	5.0	92	98	

$$MPA_{OvOSM}(i) = \frac{Area\ of\ class\ i\ in\ OSM\ derived\ data\ that\ is\ in\ agreement\ with\ UA}{Total\ area\ of\ class\ i\ in\ OSM\ derived\ data}$$

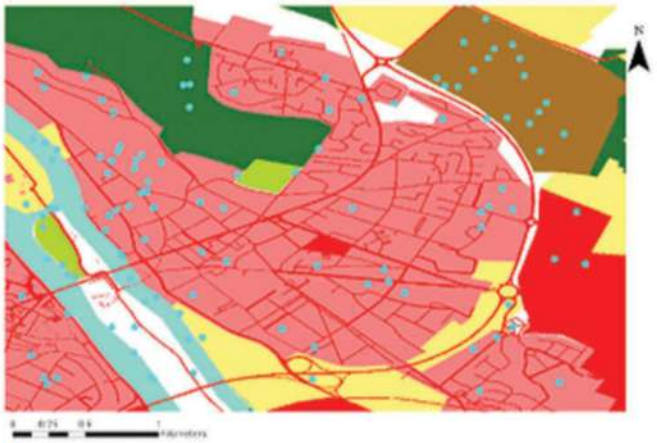
$$MPA_{OvUA}(i) = \frac{Area\ of\ class\ i\ in\ UA\ that\ is\ in\ agreement\ with\ the\ OSM\ derived\ data}{Total\ area\ of\ class\ i\ in\ UA\ where\ OSM\ data\ exists}$$

$$TPA_{Ov} = \frac{Area\ of\ agreement\ between\ UA\ and\ OSM\ derived\ data}{Total\ area\ with\ data\ in\ the\ OSM\ derived\ map}$$

LULC map validation with OSM

Urban Atlas

Data extracted from OSM



Sample points + satellite images

LULC map validation with OSM

		Producer's accuracy (%)		User's accuracy (%)		Overall accuracy (%)	
	Class	UA vs. PI	UA vs. OSM/PI	UA vs. PI	UA vs. OSM/PI	UA vs. PI	UA vs. OSM/PI
Study area A							
Level 1	1	97 [96,98]	93 [92,94]	83 [80,86]	86 [82,89]	84 [81,87]	84 [81,87]
	2	65 [47,82]	69 [48,90]	78 [75,81]	59 [56,62]		
	3	58 [44,71]	58 [44,72]	93 [91,95]	88 [86,90]		
	5	76 [54,97]	75 [53,96]	96 [95,97]	96 [95,97]		
Level 2	1.1	83 [79,88]	85 [81,89]	80 [75,85]	84 [79,89]	75 [71,79]	74 [71,78]
	1.2	70 [60,81]	72 [61,82]	64 [60,68]	65 [61,69]		
	1.3	36 [1,71]	21 [0,46]	43 [42,44]	27 [26,28]		
	1.4	64 [50,78]	52 [42,62]	47 [43,51]	51 [47,55]		
	2.0	67 [57,77]	71 [59,83]	78 [75,81]	59 [56,62]		
	3.0	67 [61,73]	68 [61,74]	93 [91,95]	88 [86,90]		
	5.0	86 [77,95]	84 [75,94]	96 [95,96]	96 [95,97]		
Study area B							
Level 1	1	88 [84,92]	88 [85,92]	79 [77,81]	85 [83,87]	89 [86,92]	89 [86,93]
	2	88 [83,92]	89 [84,94]	91 [87,95]	88 [84,92]		
	3	91 [84,96]	92 [87,98]	96 [94,98]	95 [93,97]		
	5	100 [100,100]	66 [40,92]	83 [82,84]	92 [91,93]		
Level 2	1.1	86 [80,92]	68 [62,74]	71 [68,74]	84 [81,87]	86 [83,89]	83 [81,87]
	1.2	64 [52,77]	70 [56,84]	72 [69,75]	60 [57,63]		
	1.3	67 [41,93]	53 [34,72]	49 [47,51]	70 [68,72]		
	1.4	74 [57,91]	53 [0,100]	73 [71,75]	9 [8,10]		
	2.0	89 [86,92]	90 [87,94]	91 [87,95]	88 [84,92]		
	3.0	91 [86,96]	93 [88,98]	96 [94,98]	95 [93,97]		
	5.0	100 [100,100]	68 [43,93]	83 [82,84]	92 [91,93]		

Training sets creation with OSM

- Possible approaches
  - Training classifiers using all data extracted from OSM
  - Select regions from OSM to train classifiers
    - Requires quality / reliability assessment

Training sets creation with OSM

- Approaches tested:
  - Use NDVI (Normalized Difference Vegetation Index) to identify regions with and without vegetation from the training sets
  - Identify regions where there are more than one possible LULC class in OSM and exclude them from the training set
  - Select a few regions for training instead of using all obtained data for each class

Training sets creation with OSM

- Results:
  - In some cases the accuracy of the LULC maps obtained is similar to the ones obtained with the manual identification of training sets
  - The results depend a lot on the nomenclature (LULC classes) used
  - …

OSM contribution to LULC mapping
Opportunities

- Automate the creation and validation of LULC maps - less expert intervention
  - Faster
  - Cheaper
- Have more current LULC maps
- Facilitate the creation of high resolution LULC for regions of the world where they are not available

OSM contribution to LULC mapping
Opportunities

- OSM
  - Large quantities of available data
  - Low costs associated to data collection
  - The dynamic characteristics of this type of data enables the collection of updated data
  - Citizens may have local knowledge
    - May provide more reliable data

OSM contribution to LULC mapping Limitations

- Main problems:
  - Lack of data in many places where this would be more useful
  - Data Quality
    - Traditional aspects of geospatial data quality
      - Positional accuracy, Thematic accuracy
      - Completeness
      - Logical consistency
      - Temporal consistency
      - …

OSM contribution to LULC mapping Limitations

- Data Heterogeneity
  - The available data may have different levels of quality (detail, accuracy, completeness,…) associated to:
    - Different regions
    - Data provided by different volunteers
    - At different time stamps
    - For different classes
    - ….
- Data inconsistencies, such as:
  - Overlapping polygons with different meanings

OSM contribution to LULC mapping
Future work

- Improve the tools created so far
  - Conversion process for some types of features / classes
  - Keep working on the extraction of reliable data for classifiers training
- Integration with other sources of data

GET INVOLVED IN CONTRIBUTING WITH DATA TO OSM!!

Youthmappers?

A group for HOT?

Mapping parties?

Conclusions

Many opportunities!!

OSM4LULC

Many challenges!!!!

Thank you !

Cidália Costa Fonte

Department of Mathematics – University of Coimbra, Coimbra, Portugal

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

28-30 July 2018 – Milan - Italy