Using the OpenStreetMap database to improve an object-based classification process. Application to a landcover product on small islands in the Indian Ocean

Christophe Révillion[1], Vincent Herbreteau[2]

1. Université de La Réunion, UMR 228 Espace-Dev (IRD, UA, UG, UM, UR), 40 Avenue de Soweto 97410 Saint-Pierre, La Réunion, France
2. Institut de Recherche pour le développement, UMR 228 Espace-Dev (IRD, UA, UG, UM, UR), 40 Avenue de Soweto 97410 Saint-Pierre, La Réunion, France

Introduction

The southwest of the Indian Ocean comprises numerous islands of less than 3000 sq. km, including the Comoros Archipelago, the Seychelles Archipelago, the Mascarene Islands (La Réunion, Mauritius and Rodrigues). These small island territories have very fragmented and diversified environments. Land cover and land-use mapping of such environments requires satellite images with high spatial resolution. Unfortunately, the land use products available so far were only at low resolutions (usually from 4 kilometers to 250 meters).That's why we realized a homogeneous high resolution land cover mapping of these small islands by analysing SPOT 5 satellite images.

To improve the results of this classification, we used some data from the free and participative OpenStreetMap (OSM) database as training data. Created in 2004 and often described as the "Wikipedia of Cartography", OSM is a collective project whose multiple collaborative environments aim to feed a shared, freely editable and reusable global cartographic database. The OSM community now consists of more than 4.5 million registered members united around a single focus of interest: "freely mapping the world".

Data mapping in OSM can be done in various ways, including: 1) photo-interpretation of satellite images made available on the OSM website (https://www.openstreetmap.org), 2) massive integration of free vector data (this is the case of cadastral data in some countries), 3) data recording directly on site using smartphones or tablets. In all cases, this requires the creation of an identifier for tracking editions.

OSM data can be easily queried using their keys and values (for example, downloading primary roads on the extent of an island) using the OverPass Turbo tool (http://overpass-turbo.eu/). This tool can be directly accessed in the QGIS software using the QuickOSM plugin, which greatly facilitates its use (Herbreteau et al., 2018).

Our objective was to provide an easily accessible land use / land cover product with a common typology to allow comparative studies on several islands. Using OSM data, this study also aims to propose a methodology that can be reused in other territories
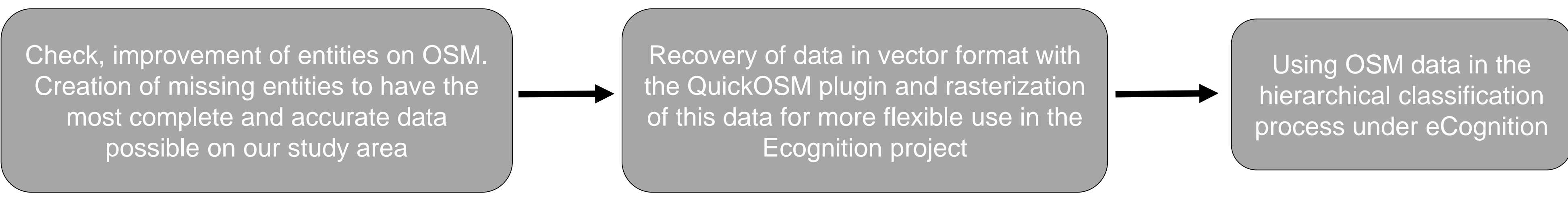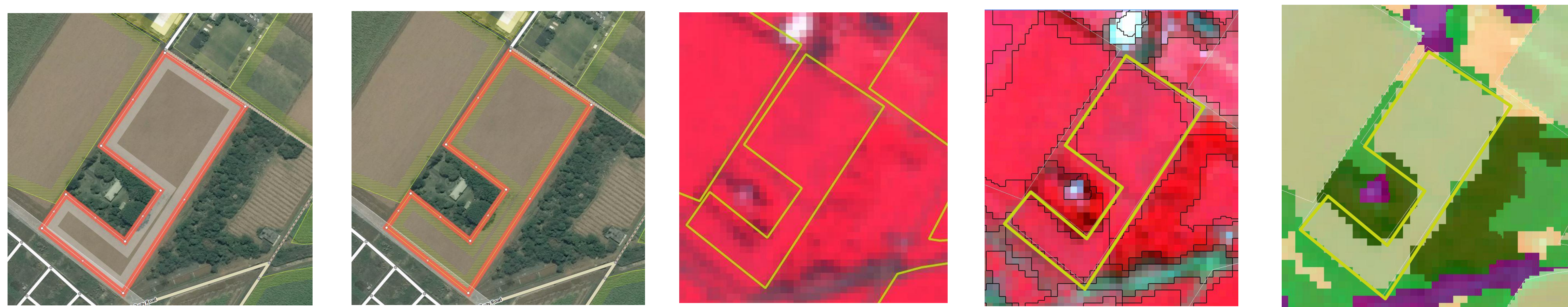
Material and methods

SEAS-OI Station (http://www.seas-oi.org/) provided SPOT 5 satellite images (© CNES, Distribution Airbus Defence and Space) acquired between July 2013 and July 2014. The SPOT5 images used are a pair of images for each island with a panchromatic image at 2.5m resolution and a multispectral image with 4 channels (green, red, PRI and MIR) at 10m resolution.

All additional vector data come from the OSM database. These are the main roads (key=highway; values=primary; secondary; tertiary; residential) and agricultural zoning (landuse=farmland). On these agricultural areas, additional tags are used when it comes to sugar cane (crop=sugarcane) which is a dominant crop on the two largest islands studied, Reunion and Mauritius.

The classification algorithm is a hierarchical object-based image analysis method, implemented in eCognition® software. The hierarchical classification allows to group the objects by the formulation of rules that can be combined, resulting from the observation and the knowledge of the operator. Each object is described by attributes related to its reflectance, texture and geometry.

Our hierarchical classification process is organized into two levels. A first level of segmentation with the objective of delineating agricultural and natural objects and a second level of segmentation with a smaller scale factor to extract urban areas, isolated buildings and main roads.
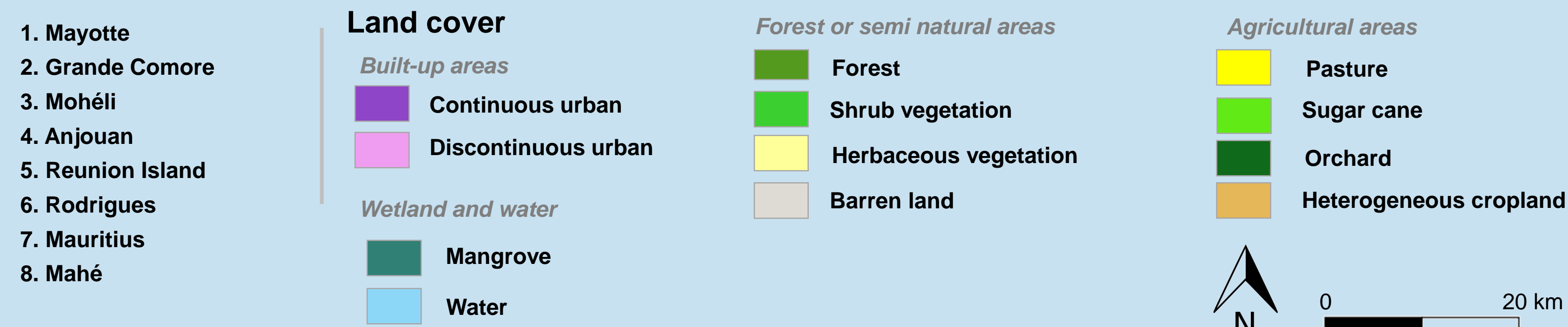
More in detail, the methodology used for the integration of OSM data in the classification process is 1) correction, improvement and creation of data on OSM 2) recovery of this data in shp format using the QuickOSM plugin under the QGIS software (rasterization) 3) integration of the data in the eCognition project and in the classification process. In the classification process OSM data is used as training data and an object that is more than 90% covered by an entity extracted from OSM retrieves OSM information (agricultural area, road).

Check, improvement of entities on OSM. Creation of missing entities to have the most complete and accurate data possible on our study area

→

Recovery of data in vector format with the QuickOSM plugin and rasterization of this data for more flexible use in the Ecognition project

→

Using OSM data in the hierarchical classification process under eCognition

Land cover classification results

We classified the land cover / land use of these tropical islands into 12 major classes. Helped by a good knowledge of the field and the use of OSM data, the classification reached a good general accuracy with a kappa of 0.85.

This high resolution land cover is available on a web portal (**http://homisland.seas-oi.org/**) to ensure a wider distribution to researchers and thematicians using land cover information.

Land cover

1. Mayotte
2. Grande Comore
3. Mohéli
4. Anjouan
5. Reunion Island
6. Rodrigues
7. Mauritius
8. Mahé

Built-up areas
- Continuous urban
- Discontinuous urban

Wetland and water
- Mangrove
- Water

Forest or semi natural areas
- Forest
- Shrub vegetation
- Herbaceous vegetation
- Barren land

Agricultural areas
- Pasture
- Sugar cane
- Orchard
- Heterogeneous cropland

0   20 km

Discussion and prospects

Discussion

The integration of participatory approach data into satellite image classification processes is an opportunity to improve the quality of remote sensing products that often lack in situ data. It also raises many methodological questions and involves taking a critical look at this data to make the best use of it. In our study, these reflexes are at a preliminary stage.

For the use of OSM data some elements have to be taken into account :

- It is necessary to check the completeness of the data used, which sometimes involves spending time editing on the online database

- Make sure that the satellite background used for OSM editing is a data close to the date of the image you want to classify

By using OSM data for this work, we helped to assess the quality of this data and improve its accuracy and completeness. This mapping work (preliminary to the classification) will therefore be free and available to any other user via the OSM database.

More globally in a period where new field of study for automated remote sensing image classification opens with mathematical approaches such as deep learning or data harvesting, the OSM participatory database can have a great contribution to these fields of study.

Prospects :

- Broaden the scope of users by increasing the access possibilities (for instance, by WMS) ;

- Update this land cover by using SPOT6 / 7 images. It will allow diachronic studies.

- A next step will be to integrate land cover classes obtained by remote sensing image analysis into OSM. This could be difficult when these classes are already partially existing in OSM since it is necessary to compare with the old data and avoid deleting any relevant information. In this sense, we have already integrated the mangroves on the Comoros. By integrating remote sensing data into OSM, it is an effective way to share and release such data.

References

Herbreteau V., Révillion C., Trimaille E.. GeoHealth and QuickOSM, Two QGIS Plugins for Health Applications. In : Baghdadi N. (ed.), Mallet C. (ed.) Zribi M. QGIS and Generic Tools Volume 1 – QGIS in Remote Sensing. London: ISTE, 2018, (1), p. 257-286. (Earth Systems – Environmental Sciences: QGIS in Remote Sensing SET; 1). ISBN 978-1-78630-187-1

GEOBIA 2018
GEOBIA IN A CHANGING WORLD