FOSS4G 2018
Dar es Salaam

POLITECNICO
MILANO 1863

CAPACITY BUILDING FOR HIGH-RESOLUTION LAND COVER INTERCOMPARISON AND VALIDATION: WHAT IS AVAILABLE AND WHAT IS NEEDED

M. A. Brovelli, M. Minghini, M. E. Molinari,
H. Wu, X. Zheng, J. Chen

MINISTERO DELL'ISTRUZIONE,
DELL'UNIVERSITÀ E DELLA RICERCA

isprs
information from imagery

NGCC
国家基础地理信息中心
National Geomatics Center of China

Context

- **Land Cover (LC) maps: a key class of global geospatial datasets**
    - LC maps are fundamental for a wide range of users and applications such as **planning**, **nature and biodiversity protection**, **natural resources management**, etc.
    - LC products represent a **key input** to monitor the indicators of the Sustainable Development Goals (**SDGs**)

    - LC data promote evidence-based policy-making on issues like **soil consumption** and **deforestation**

Context

- **High-resolution LC maps** are rapidly increasing due to the continuous advances in Remote Sensing sensors and geospatial technologies
  - Several countries (e.g. EU, USA, Australia) and political organizations have their own high-resolution LC maps
  - Developing countries do not have their own high-resolution datasets but **they can benefits from the free availability of open global high-resolution** LC products (e.g. GlobeLand30)

Context

- It is important that data users and producers in the fields of GIS and RS, especially in developing countries, have:
  - **Awareness** of the existence and importance of global high-resolution LC maps
  - **Capability to perform** high-resolution LC maps **validation** and **inter-comparison** to determine their usability for different applications

“**Capacity Building for High-Resolution Land Cover Inter-comparison and Validation**” project

Outline

- Global high-resolution LC maps
- LC data validation and inter-comparison methods
- Capacity buildings on LC project: tasks and results

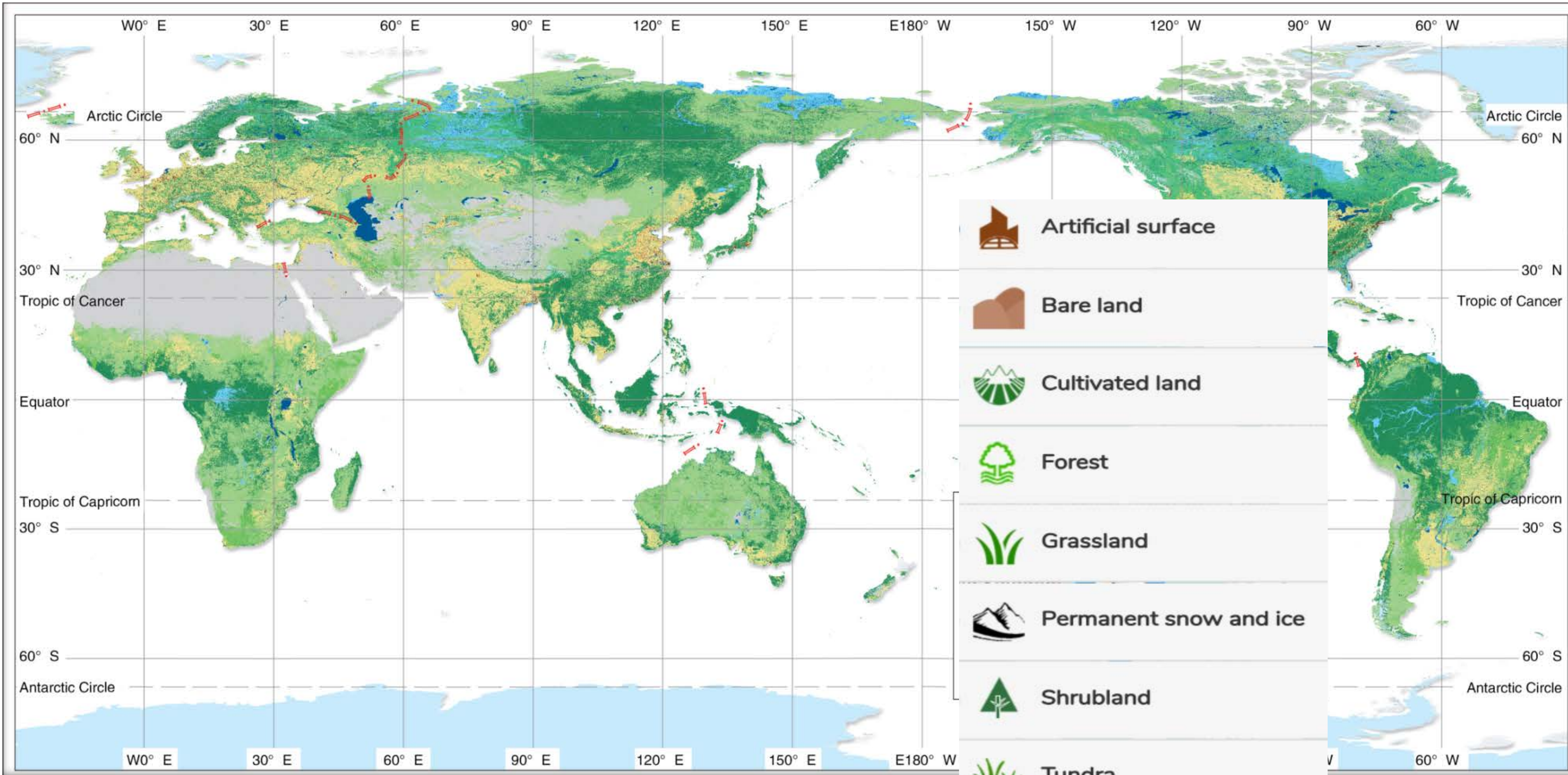
Global and high-resolution LC maps

Name	Resolution (m)	Temporal Coverage	Producer
FROM-GLC	30	2010, 2015	Tsinghua University
GlobeLand30	30	2000, 2010,2015	National Geomatics Center of China (NGCC)

Global and high-resolution LC maps

Name	Resolution (m)	Temporal Coverage	Producer
FROM-GLC	30	2010, 2015	Tsinghua University
GlobeLand30	30	2000, 2010,2015	National Geomatics Center of China (NGCC)
Global Water Surface	30	1984-2015	Joint Research Centre (JRC)
Forest / Non-Forest map	25	2007-2010 2015-2016	Japan Aerospace Exploration Agency (JAXA)
Global Urban Footprint	12	2011	German Aerospace Center (DLR)
Global Human Settlement Layer	38	1975, 1990, 2000, 2014	Joint Research Centre (JRC)
Tree Cover	30	2000	University of Maryland
Global forest cover gain		2000-2012	
Global forest cover loss		2000-2015	

Example: GlobeLand30



www.globallandcover.com/

Validation and inter-comparison methods

- Since the mid-1980s, **error/confusion matrix** is considered *"The standard descriptive reporting tool for accuracy assessment of remotely sensed data"* [1]

	Reference (ground truth) map				
Classified (comparison) map	Class	j=1	j=2	…	j=q
	i=1	n_{11}	n_{12}	…	n_{1q}
	i=2	n_{21}	n_{22}	…	n_{2q}
	…	…	…	…	…
	i=q	n_{q1}	n_{q2}	…	n_{qq}

[1] Lunetta, R.S., Lyon, J.G., 2004. Remote Sensing and GIS Accuracy Assessment, CRC Press, Boca Raton, FL

Validation and inter-comparison methods

- Many **accuracy indexes** can be derived from confusion matrix:
  - Most commonly used:
    - ✓ Overall accuracy (P0)
    - ✓ Producer’s accuracy (PA)
    - ✓ User’s Accuracy (UA)
  - Derived from P0, PA, UA
    - ✓ Average of user’s accuracy or of producers accuracy
    - ✓ Combined user’s or producer’s accuracy
    - ✓ Hellden’s mean accuracy
    - ✓ Short’s mean accuracy
    - ✓ Classification success index and its variations: Group Success Index and Individual Classification Success Index
  - Margfit

Validation and inter-comparison methods

- Derived from information theory
  - ✓ Average mutual information and different ways of normalizing it (arithmetic mean, geometric mean)
- Kappa and kappa-like indexes
  - ✓ Standard kappa index
  - ✓ Conditional kappa
  - ✓ Weighted kappa
  - ✓ Tau
  - ✓ Aickin's alpha
  - ✓ Ground truth index
- Indexes of disagreement
  - ✓ Quantity disagreement
  - ✓ Allocation disagreement

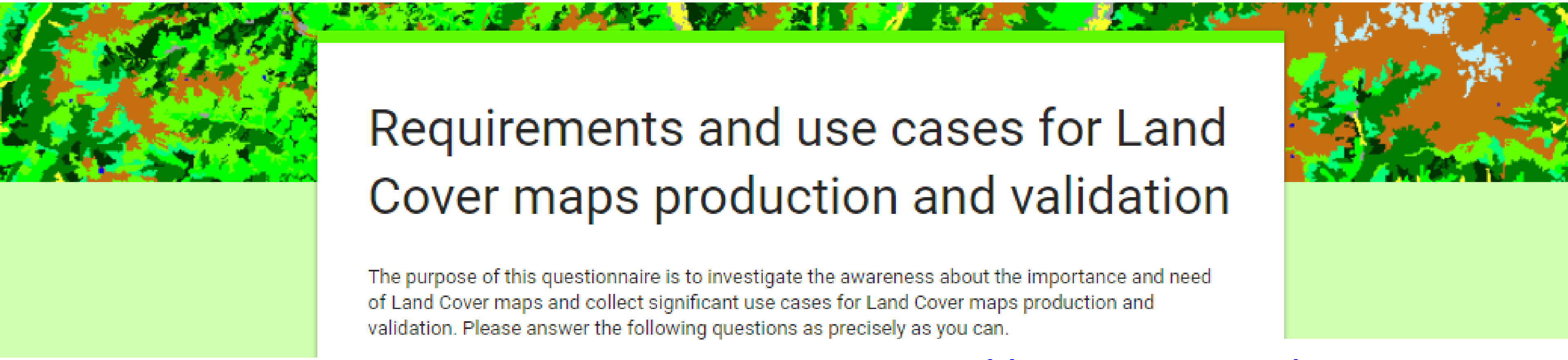
Capacity buildings on LC mapping project

- It is funded by the **International Society for Photogrammetry and Remote Sensing** (Educational and Capacity Building Initiatives 2018: http://www.isprs.org/society/ecbi/default.aspx)
- The aim is to **create new knowledge and tools to educate and raise awareness on the inter-comparison and validation** of global high-resolution LC maps, mainly in developing countries
- The initiative started in February 2018 and will end in January 2019. It is composed of four tasks

Capacity buildings on LC mapping project

- **TASK 1**: Analysis of the **needs, requirements and limiting factors** in using and validating LC maps from a user's perspective, with special focus on developing countries

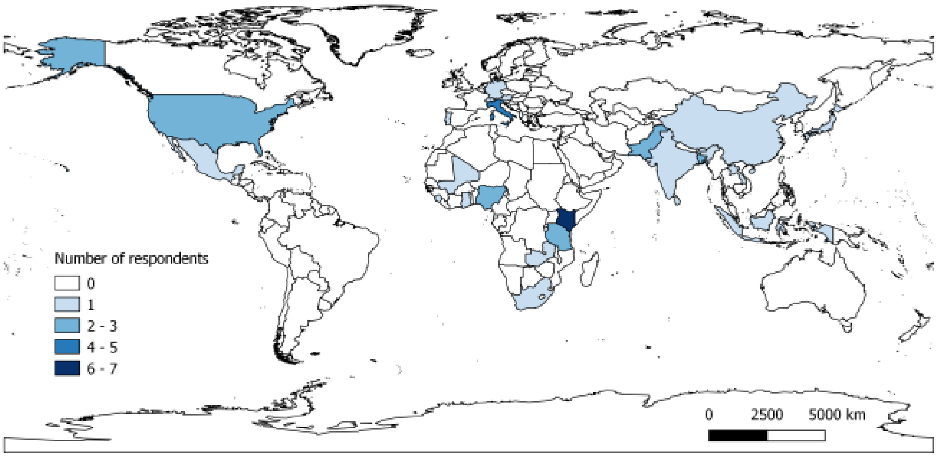
  ➔ Ad hoc questionnaire aimed at assessing the general awareness about the existence and importance of LC maps as well as the need for their inter-comparison and validation

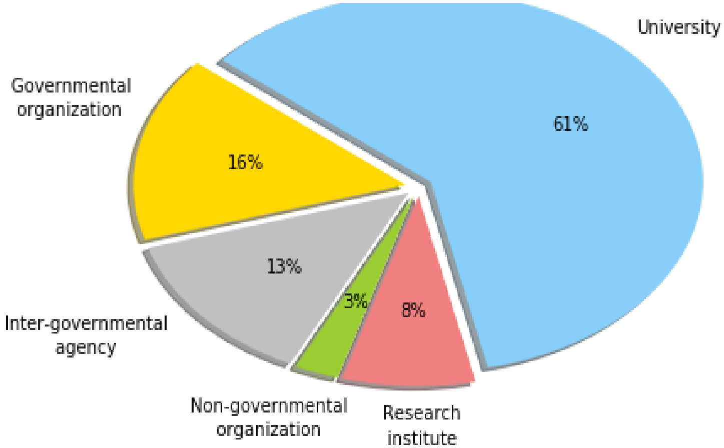


https://tinyurl.com/ydgg59ua

TASK 1 results: ad hoc questionnaire

- **39 respondents from 24 different countries (most of the respondents were from developing countries)**





- **92%** of respondents are **familiar with LC** maps and **85%** of them **have used LC** maps in their research or professional work

- **1/3** of the respondents have used LC maps derived from Landsat imagery; 15% have used **CORINE LC** and **13%** have used **GL30**. Other LC maps have been used by about 5% of the respondents (maps from MODIS and Sentinel, Urban Atlas, Global Urban Footprint, etc.)

TASK 1 results: ad hoc questionnaire

- Respondents highlighted the use of LC maps for a huge variety of applications:
  - **Environmental applications**: agricultural monitoring, disaster management, inventories of forests and greenhouse gases, urban planning and urban growth analysis, studies on climate change and food security, etc.
  - **Research applications**: comparison and validation of diverse LC maps, also with gamification approaches; creation of hybrid LC maps, including OSM-derived maps
  - **RS and spatial analysis teaching**

- Main reasons for LC validation and inter-comparison:
  - need to count on accurate products to quantify and report environmental indicators
  - make users aware of the large disagreements existing between available LC maps
  - **identify the best datasets to be used at specific scales and for specific applications**
  - guide error-free scientific research

- The best resolution of LC maps depends on the specific application: for instance <10m for urban applications, between 30 and 500m for agricultural applications. High-resolution maps have a wider applicability than medium and low-resolution ones

Capacity buildings on LC mapping project

- **TASK 2**: assessment and classification of the available training material on inter-comparison and validation of global high-resolution LC maps
  - **Manual review of the state of the art** (FOSS4G packages and training material)
  - Ad hoc **questionnaire** that asks respondents to indicate their knowledge of training or educational material

https://tinyurl.com/yde9ykqg

TASK 2 results: review of the state of the art

- The most relevant **FOSS4G packages** for LC mapping (including specific modules or extensions) and related **learning material**

TASK 2 results: review of the state of the art

The most significant tools provided by QGIS to handle LC maps are available into 2 dedicated plugins:

1. **Semi-Automatic Classification Plugin (SCP)**
   - ✓ Download, preprocessing and post-processing of imagery by ASTER, Landsat, MODIS and Sentinel-2
   - ✓ Algorithms for the supervised and unsupervised classification of satellite imagery
   - ✓ Step-by-step tutorial on how to perform land cover classification

TASK 2 results: review of the state of the art

The most significant tools provided by QGIS to handle LC maps are available into 2 dedicated plugins:

1. Semi-Automatic Classification Plugin (SCP)
   - ✓ Computation of confusion matrix and **overall accuracy, producer's accuracy, user's accuracy, kappa coefficient and its variance**
   - ✓ Availability of learning material about validation (set of tutorial, videos and screenshots)
   - ✓ Plugin available for QGIS 2 and recently rewritten for QGIS 3

TASK 2 results: review of the state of the art

QGIS

The most significant tools provided by QGIS to handle LC maps are available into 2 dedicated plugins:

2. **Accuracy Assessment plugin**

   - ✓ Comparison of two LC raster maps available in QGIS and computation of a confusion matrix as a Comma Separated Values (CSV) file
   - ✓ Short plugin explanation included in the GitHub repository README file. No tutorial or other learning material was found
   - ✓ Plugin only available for QGIS 2. Its development apparently stopped in 2014

TASK 2 results: review of the state of the art

The relevant GRASS modules for managing LC maps belong to the modules for image data and raster data:

- ✓ i.maxlik, i.cluster, i.gensig, g.gui.iclass: supervised and unsupervised image classification
- ✓ r.kappa: validation of LC products. It computes the confusion matrix and returns accuracy measures such as overall **kappa, conditional kappa, overall accuracy, user's accuracy and producer's accuracy**
- ✓ The GRASS GIS reference manual contains a section related to image processing, which includes both image classification and validation. A short tutorial on image classification is available

TASK 2 results: review of the state of the art

ORFEO ToolBox

It is specifically developed for Remote Sensing purposes:

- ✓ Algorithms for unsupervised and supervised classification provided by TrainVectorClassifier and TrainImagesClassifier applications
- ✓ Validation capabilities thanks to the ComputeConfusionMatrix application, which returns the error matrix in CSV format together with the **overall accuracy and the user's and producer's accuracy** for each LC class
- ✓ A guideline on classification and validation and a tutorial on how to perform supervised classification are available

TASK 2 results: review of the state of the art

- ✓ Tools for performing classification of imagery (both supervised and unsupervised)
- ✓ Tools for validating the classification results through the creation of a confusion matrix and the derived coefficients (**kappa and overall accuracy**)
- ✓ Tutorials on both supervised and unsupervised imagery classification and a tutorial on LC change mapping (including classification and inter-comparison steps) are available

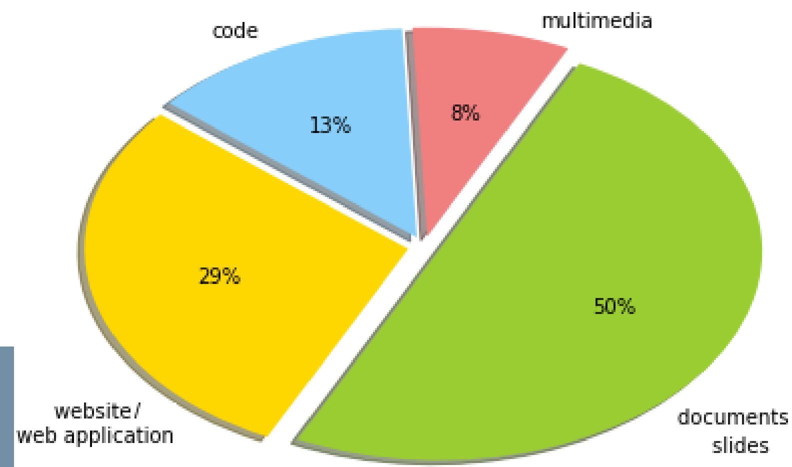
TASK 2 results: review of the state of the art

Sentinel-2 toolbox offers a rich set of tools for the visualisation, analysis and processing of high-resolution optical imagery

- ✓ Algorithms for both supervised and unsupervised classification are available
- ✓ **Validation or inter-comparison tools are not available**
- ✓ Specific tutorials or learning materials are not available but ESA is regularly organizing dedicated training events on land Remote Sensing topics

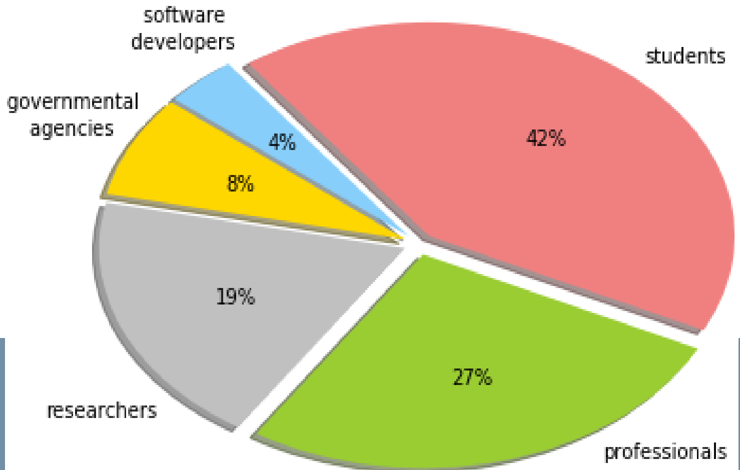
TASK 2 results: ad hoc questionnaire

- 29 respondents from 21 different countries; most of them were the same as in the Task 1 survey. Their geographic distribution is thus similar to the Task 1 survey, with the exception of one new respondent for each of the following countries: Brazil, Canada, Nepal, The Netherlands and Uganda
- The proportions of the respondents' affiliation types are pretty much the same as those obtained from Task 1 survey
- Regarding the type of material the emerging panorama is very heterogeneous:



TASK 2 results: ad hoc questionnaire

- Educational material is based on ESRI ArcGIS (40%), QGIS (28%), ERDAS IMAGINE (17%), OTHERS (ENVI, GRASS GIS, TerrSet, eCognition, IMPACT Toolbox, Google Earth, LULC Mapper, GEOVAL, LACO-Wiki , GL30 Platform)
- Answers about **license** were discarded (many respondents indicated license of the software instead of the material; others provided generic answers, etc.). **It seems that researchers and professionals are unfamiliar with licenses and/or they do not give them so much importance**
- Training material objectives and target users:



Capacity buildings on LC mapping project

- **TASK 3**: development of new computer aided educational material on the intercomparison and validation of global high-resolution LC maps

➡ Creation of ad hoc teaching material released under open access licenses, and software-based material released under open source licenses to maximize the exploitation and impact within the community

TASK 3 results: teaching material and software

- Desktop procedure for LC map validation, implemented in QGIS and using GlobeLand30 as sample dataset. The teaching material license is CC BY 4.0
- Validation is performed taking advantage of custom scripts for PyQGIS (https://github.com/GoricaB/Land-cover-validation) written by M. Molinari and G. Bratic under the guidance of M.A. Brovelli

Use case 1

Objective

Validation of GlobeLand30 by means of a comparison with a reference points dataset obtained from LUCAS, a land use and land cover survey programme promoted by Eurostat.

Area of interest

Lombardy Region (Northern Italy)

Datasets

- **GlobeLand30 2010** raster maps covering the Lombardy Region area: N32_40_2010LC030, N32_45_2010LC030 (available in DATA\GL30_Italy folder). The data are provided in WGS84/UTM32N coordinate system (EPSG: 32632)
- **LUCAS 2009** dataset related to Italy (available here or in DATA\LUCAS folder). The data are provided in WGS84 reference system (EPSG: 4326)

isprs

18

Use case 2

Objective

Validation of GlobeLand30 by means of a comparison with a reference raster dataset obtained from DUSAF, a land use and land cover database of Lombardy Region, Italy.

Area of interest

Como Province, Lombardy Region (Northern Italy)

Datasets

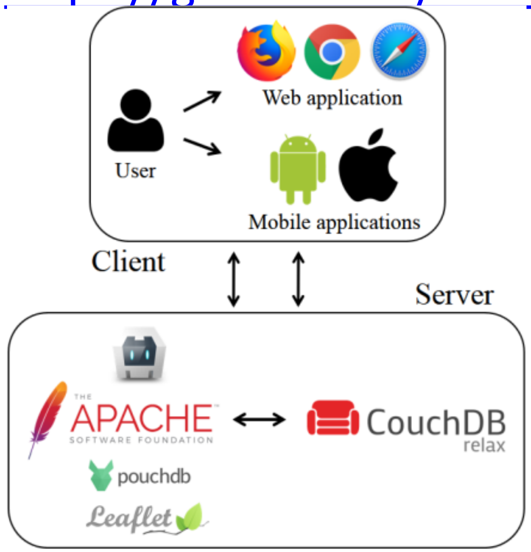
- **GlobeLand30 2010** raster map covering the Como Province area: N32_45_2010LC030 (available in DATA\GL30_Italy folder). The map is provided in the WGS84/UTM 32N coordinate system (EPSG: 32632)
- **DUSAF 4.0 – Use of soil 2012** database consists of vector maps for every province in Lombardy Region, as well as for the whole Lombardy Region (available here). The map is in WGS84 reference system, UTM 32N projection (EPSG:32632).

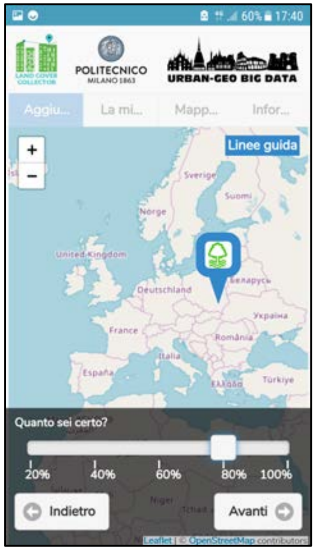
isprs

65

TASK 3 results: teaching material and software

- Development of a web app, named **Land Cover Collector**, to allow users to collect field data according to the same LC nomenclature of GlobeLand30
- The web app is released under the GPL 3.0. Collected data are released under the Open Database License (ODbL) and developed by C. E. Kilsedar under the guidance of M. A. Brovelli
- Code is available at https://github.com/kilsedar/land-cover-collector





Capacity buildings on LC mapping project

- **TASK 4**: organization of three workshops on the intercomparison and validation of global high-resolution LC maps

Workshop 2: Capacity Building for High-Resolution Land Cover Intercompariso...

by ISPRS Working Group IV/4

Free

Sales Ended

DETAILS

DESCRIPTION

High-resolution Land Cover maps are fundamental for many applications such as natural resources management, ecological and hydrological modelling and study of phenomena like soil consumption and deforestation and climate change assessment.

The availability of these global datasets as open data is a great richness everywhere, and especially where these products are not available, but sometimes there are limitations in their usage due to the lack of capacity in managing and processing them. Moreover, there is a need to assess them at the local level to evaluate their accuracy and, more in general, their fitness-for-use.

DATE AND TIME

Mon, 1 October 2018
1:30 pm – 3:00 pm Central European Summer Time Netherlands Time
Add to Calendar

LOCATION

TU Delft Aula Conference Centre
5 Mekelweg
2628 CC Delft
Netherlands

REGIONAL CENTRE FOR MAPPING OF RESOURCES FOR DEVELOPMENT

...ver Inter-comparison ...lation

Nairobi, 3 September 2018, 9:00-13:00 am
Regional Centre for Mapping of Resources for Development (RCMRD)
Kasarani Road, Off Thika Road
P.O Box 632-00618
Nairobi, Kenya
Room: GIS Training Lab
Speakers: Prof. Chen Jun, Prof. Maria Antonia Brovelli, Mr. Peng Shu, Dr. Marco Minghini
Local Organizers: Mrs. Phoebe Oduor and Kenneth Kasera

High-resolution Land Cover maps are fundamental for many applications such as natural resources management, ecological and hydrological modelling and study of phenomena like soil consumption and deforestation and climate change assessments.
The availability of these global datasets as open data is a great resource, and especially where these products are not available, but sometimes there are limitations in their usage due to the lack of capacity in managing and processing them. Moreover, there is a need to assess them at the local level to evaluate their accuracy and their fitness-for-use.

High-Resolution Land Cover Intercomparison and Validation

Dar Es Salaam, 1 September 2018, 9:00-13:00 am
World Bank
Room: Room No. 110 First floor, Address (50 Mirambo Street)
Speakers: Prof. Chen Jun, Prof. Maria Antonia Brovelli, Mr. Peng Shu, Dr. Marco Minghini
Local Organizers: Mr. Msilikale Msilanga, Miss. Devotha Laurent

High-resolution Land Cover maps are fundamental for many applications such as natural resources management, ecological and hydrological modelling and study of phenomena like soil consumption and deforestation and climate change assessment.
The availability of these global datasets as open data is a great richness everywhere, and especially where these products are not available, but sometimes there are limitations in their usage due to the lack of capacity in managing and processing them. Moreover, there is a need to

Training material

GIS Team

POLITECNICO MILANO 1863

GEOLaB Geomatics and Earth Observation lab

Home Topics Projects Applications Team Publications Affiliations Software

GEO4D (2017-2020)

GEO4D is a Capacity Building in Higher Education project, funded by the Erasmus+ Programme of the European Commission with the aim to modernize higher education in geodesy in order to support sustainable development in Jordan. The specific project's objective is to establish 3 new geodesy/GIS laboratories at 3 Jordanian partner universities during 2018, develop and start 3 new master programmes in autumn 2019 and introduce e-learning, Problem-Based Learning (PBL) and quality assurance in geodesy education during 2020.

isprs information from imagery

Capacity Building for High-Resolution Land Cover Intercomparison and Validation (2018)

Funded as one of the ISPRS Education and Capacity Building Initiatives 2018, the project aims to create computer-aided teaching and learning material about the intercomparison/validation of global land cover maps and to organize three workshops, two of which are held in developing countries (Tanzania and Kenya). The training material for validation with QGIS can be downloaded here, and the training material on how to use the Land Cover Collector application can be downloaded here. Principal investigators are Politecnico di Milano and the National Geomatics Center of China. The development of the Land Cover Collector application was supported by Italian Ministry of Education, University and Research (MIUR) thanks to the URBAN GEO BIG DATA project.

http://geomobile.como.polimi.it/website/

POLITECNICO
MILANO 1863

Thank you for the attention

Maria Antonia Brovelli

Politecnico di Milano
Department of Civil and Environmental Engineering
Piazza Leonardo da Vinci, 32 - 20133 Milano (Italy)

maria.brovelli@polimi.it