Using Free and Open Source Software for Visualization and Processing of Big Multidimensional Open Urban Geospatial Data on the Web

Candan Eylül Kilsedar · Maria Antonia Brovelli

Department of Civil and Environmental Engineering, Politecnico di Milano, Milan, Italy

Motivation

The available big urban geospatial data is not fully used while making decisions regarding cities. As a result, city management policies cannot be optimally effective. Moreover, currently, the scientific knowledge driven using this data is not effectively communicated to the public. This results in a lack of engagement of citizens in the decision-making processes.

Proposed Solution

We developed a geoportal to address these issues. For development, we used open standards, free and open source software (FOSS) and open data to maximize interoperability, replicability, reusability, and accessibility. Furthermore, we used Web technologies so that it can run on both mobile and desktop devices regardless of the operating system.

The geoportal encompasses various use cases related to urban areas. The visualization is performed by means of a 3D Web GIS, using mainly NASA WebWorld Wind and CesiumJS APIs. Through effective visualization, the geoportal aims to communicate massive multidimensional data in a clear way. The geoportal also enables to query and process the data.

URBAN GEO BIG DATA Project

The URBAN GEOmatics for Bulk Information Generation, Data Assessment and Technology Awareness (URBAN GEO BIG DATA) project aims to develop innovative GIS tools that use big data from various sources to understand several urban dynamics better and as a result, manage urban resources and infrastructures more effectively. The data sources that enable to monitor the status of the urban areas include the traditional datasets such as topographic databases, LiDAR data and statistical data; earth observation data; the data from Web 2.0, such as volunteered geographic information (VGI) and data from social networking services; passive geo-crowdsourced data, such as SMS, phone calls etc. The GIS tools developed within the project aim to enable the visualization, query, and processing of these data sources. The project focuses on five main cities in Italy, which are Milan, Naples, Rome, Padua, and Turin.

URBAN-GEO BIG DATA

http://www.urbangeobigdata.it/

3D OSM Data Visualization with NASA Web WorldWind

NASA Web WorldWind and 2017 GSoC project 3D OSM Plugin API are used for OpenStreetMap (OSM) data visualization. The building footprints available in the OSM database are used to visualize buildings in three dimensions on top of the virtual globe. The other features in this database are overlayed on the virtual globe in two dimensions.

LiDAR and OSM GeoJSON are used in GRASS GIS to set Milan buildings' height. Urban Atlas Building Height 2012 data of the Copernicus programme and OSM GeoJSON are used in GRASS GIS to set Rome buildings' height. OSM attributes are used to set the buildings' height in Naples, Padua, and Turin.

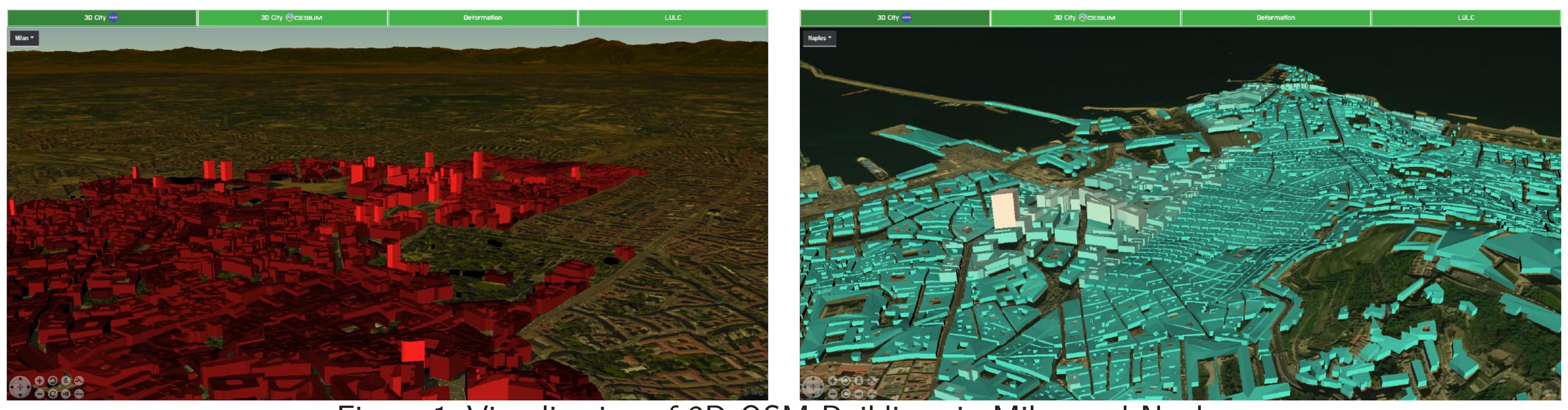


Figure 1. Visualization of 3D OSM Buildings in Milan and Naples

https://github.com/kilsedar/3dosm

Flood Simulation with CityGML and CesiumJS

Floods pose a risk that is likely to worsen in the future due to climate change. Therefore, it is essential that decision makers and domain experts have the tools to evaluate the effects of floods. We developed a tool that visualizes the earth and buildings in three dimensions to simulate floods so that effective strategies can be developed to enhance resilience and mitigate the effects of floods.

As a result of the literature review, we decided to use CityGML and CesiumJS for three-dimensional geospatial data visualization. However, as CityGML data is not available for the cities that our project focuses on, we used a software called shp2city that converts Esri shapefile to CityGML data in LOD0 or LOD1. Currently, CityGML is generated for all the target cities except Rome. Moreover, as CityGML data cannot be immediately used with CesiumJS, we used 3DCityDB to store, represent, and manage the CityGML data; 3DCityDB Importer/Exporter to export the CityGML data in KML/COLLADA/glTF format to be used within the 3DCityDB Web-Map-Client that is based on CesiumJS for visualization. Finally, we simulated floods to aid in the informed decision-making process regarding adaptation measures and mitigation of flooding effects.

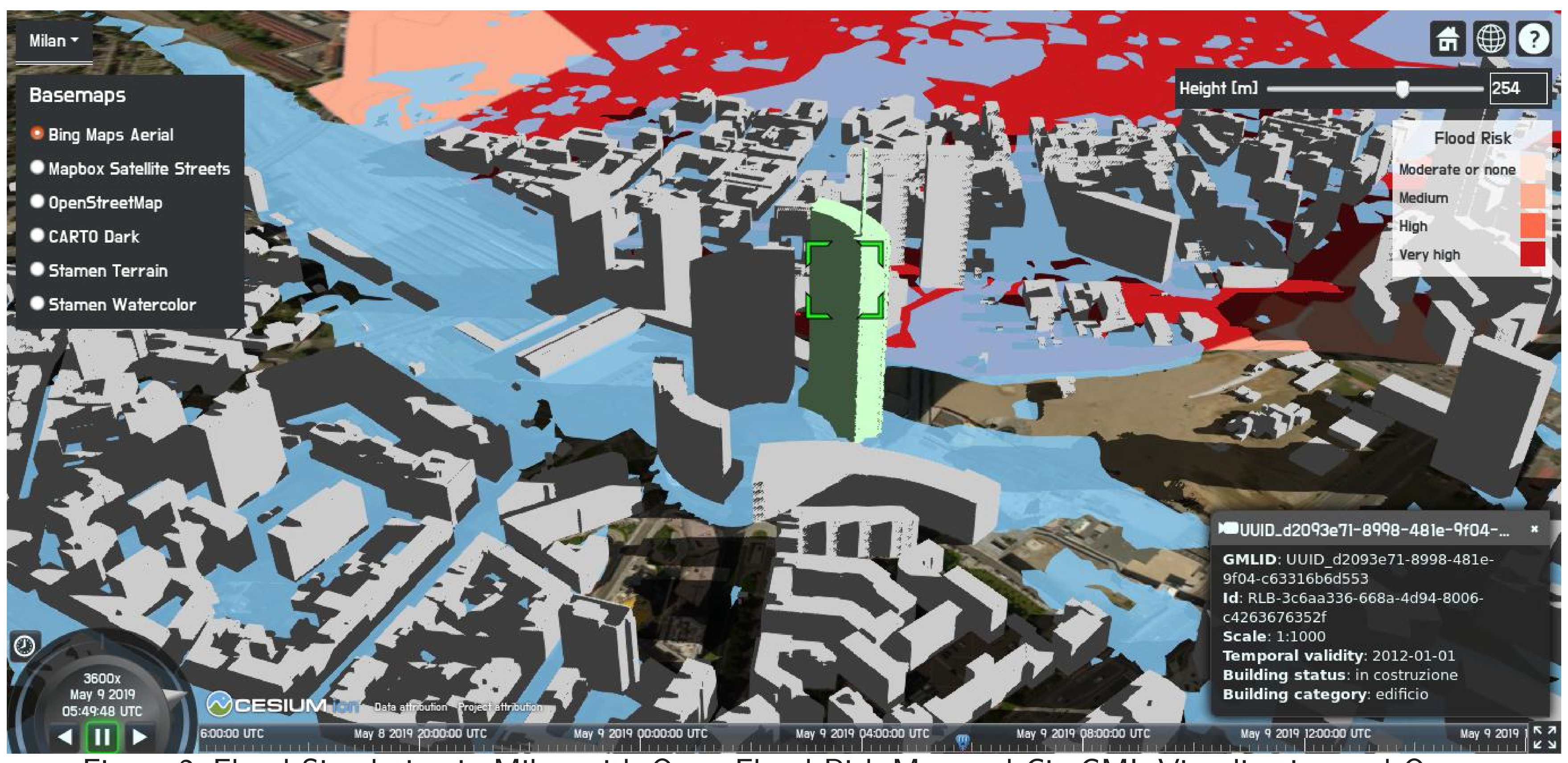


Figure 2. Flood Simulation in Milan with Open Flood Risk Map and CityGML Visualization and Query

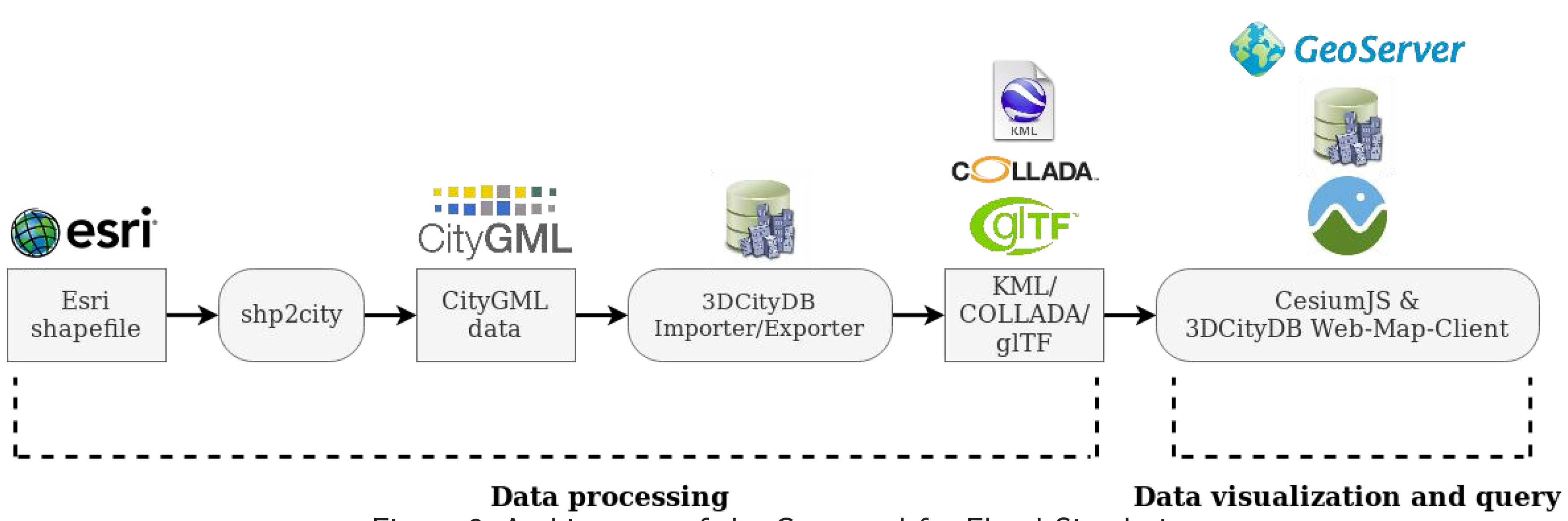


Figure 3. Architecture of the Geoportal for Flood Simulation

https://github.com/kilsedar/urban-geo-big-data-3d

Land Deformation Visualization and Query with CesiumJS

The deformation maps are produced by National Research Council of Italy (Consiglio Nazionale delle Ricerche, CNR)–Institute for Electromagnetic Sensing of the Environment (Istituto per il Rilevamento Elettromagnetico dell'Ambiente, IREA). The technique, which was proposed in 2002, is widely used for the investigation of Earth surface deformation and allows the generation of mean deformation velocity maps, and for each target on the ground deformation time series. The dataset used consists of 164 SAR images collected by the ERS and the Envisat satellites.

Data is stored on GeoServer and used through Web Map Service (WMS) on a virtual globe built with CesiumJS. Each target on the ground can be queried to display the deformation time series plot. Time series are plotted using Plotly. Deformation is also visualized in cm/year for 16 years as animation using Web Map Tile Service (WMTS) and ImageMosaic through GeoServer and timeline and animation widgets of CesiumJS to demonstrate the movement of the land for Naples, Milan, and Turin. The same will be implemented also for Rome and Padua.

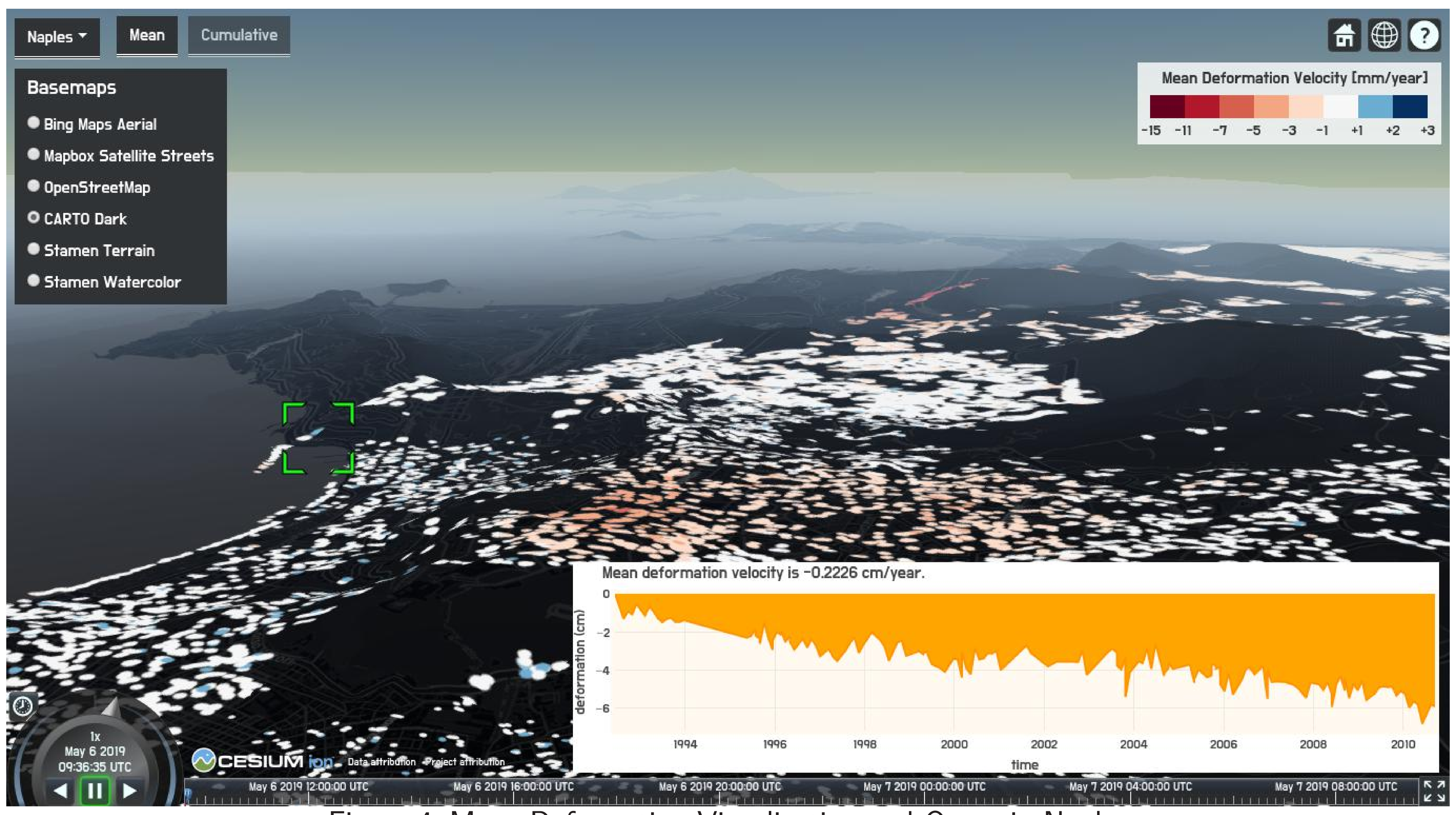


Figure 4. Mean Deformation Visualization and Query in Naples

https://github.com/kilsedar/urban-geo-big-data-3d

Big Multidimensional Raster Data Visualization and Processing with CesiumJS

EO-derived datasets GlobeLand30 of 2000 and 2010; Global Human Settlement Layer of 1975, 1990, 2000 and 2014; built-up area map from ISPRA (Italian Institute for Environmental Protection and Research) of 2012, 2015, 2016 and 2017; land cover map from ISPRA of 2012 are visualized and processed on the Web. The spatial extent of all the datasets is restricted to Italy because of the limited hardware. The spatial resolutions of the datasets in order are 30 m, 40 m, 10 m, and 10 m.

The datasets of multiple years are visualized using animation to enable detecting changes in land cover or soil consumption visually using the same method for deformation animation. The datasets are processed using Web Coverage Processing Service (WCPS) through rasdaman (raster data manager), an Array DBMS. At the moment, the processing involves returning the change of land cover classes or soil consumption through years for the clicked coordinates. In the future, we will calculate the amount of change of a land cover class or soil consumption for an area drawn by the user for two selected years.

Volunteered geographic information (VGI) on land cover classification, collected with the application Land Cover Collector, using the nomenclature of GlobeLand30 is overlayed on top of the raster maps. The VGI data can be queried.

https://github.com/kilsedar/land-cover-collector

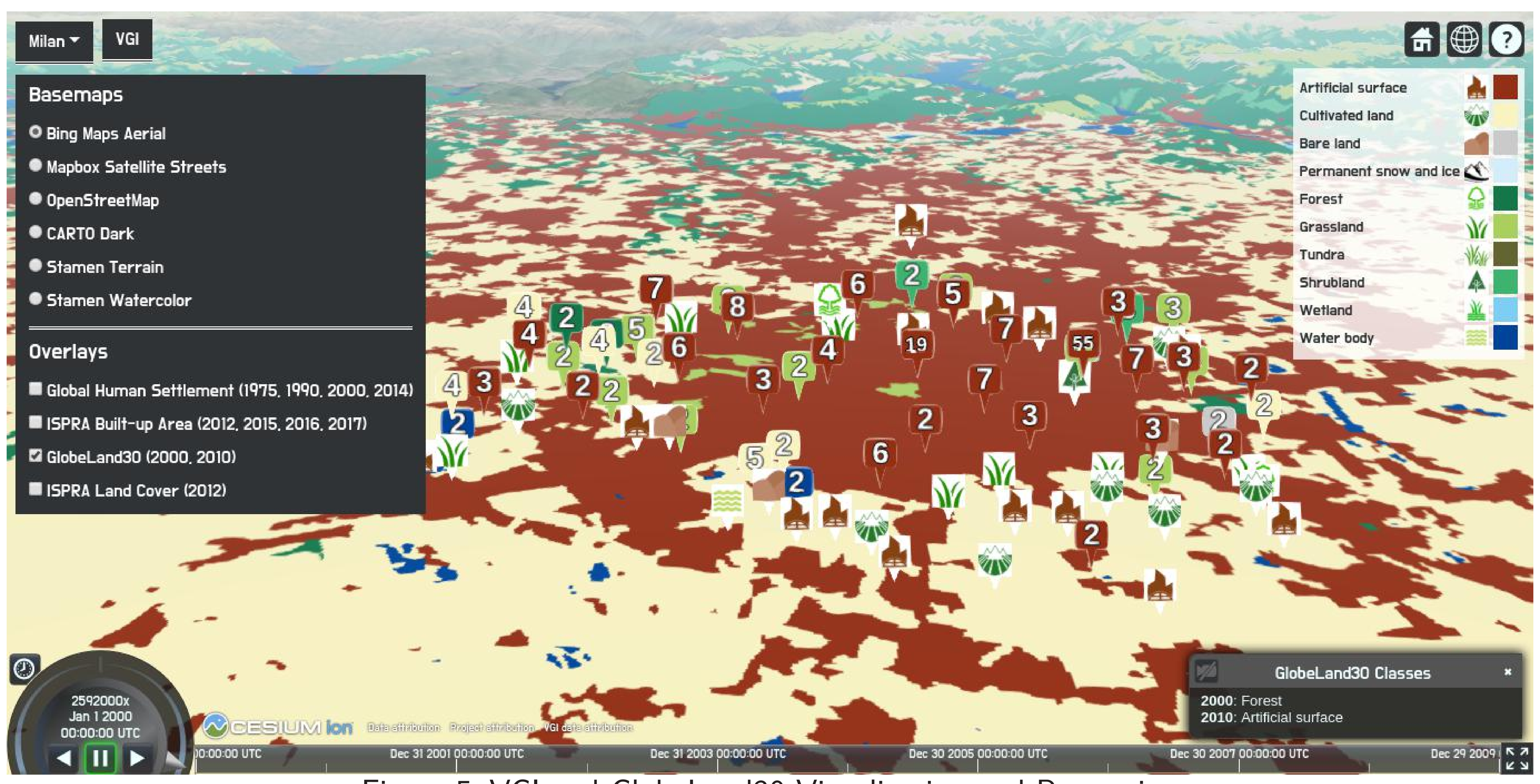


Figure 5. VGI and GlobeLand30 Visualization and Processing

https://github.com/kilsedar/urban-geo-big-data-3d

Acknowledgements

This work is supported by URBAN GEOmatics for Bulk Information Generation, Data Assessment and Technology Awareness (URBAN GEO BIG DATA), a project of national interest (PRIN) funded by the Italian Ministry of Education, University and Research (MIUR)—id. 20159CNLW8.

Prof. Francesco Pirotti and Dr. Francesca Fissore from the University of Padova developed the software called shp2city and provided us the CityGML data. Prof. Antonio Pepe, Prof. Gloria Bordogna, and Eng. Luca Frigerio from CNR provided us the deformation maps.

Empowering Accuracy Assessment Procedures for Global Land Cover Maps with Spatial Association Analysis

Daniele Oxoli, Gorica Bratic, Maria Antonia Brovelli
Department of Civil and Environmental Engineering, Politecnico di Milano, Milan, Italy

Background & Objectives

The availability of **Global Land Cover** (GLC) products has increased over the last few years driven by the modern Earth Observation (EO) platforms capabilities (frequent pass | **high-resolution** | global coverage). However, the accuracy of GLC maps not always meets the users' requirements making the use of regional land cover maps often preferred. The **accuracy assessment** of GLC maps still represents a pivotal task in order to promote the use of **GLC map for local applications**.

- The study focuses on the validation of the GlobeLand30 (GL30) map at a regional scale by empowering traditional accuracy assessment procedures with spatial association statistics and error patterns mapping

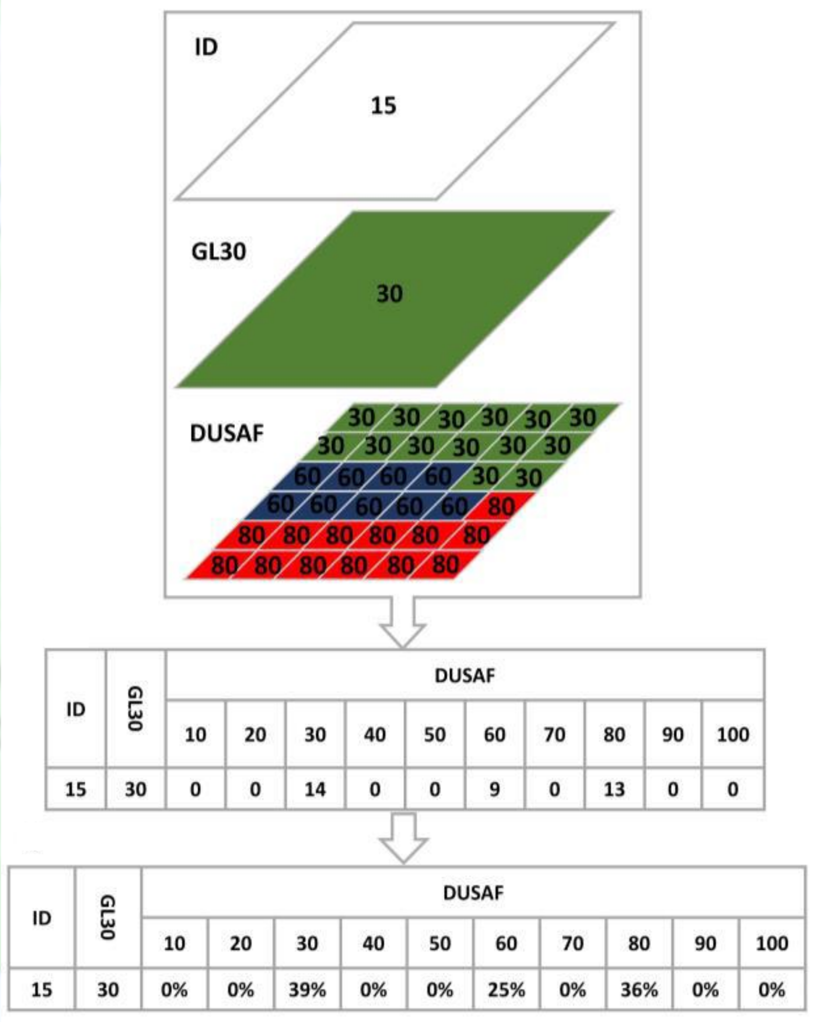
Case Study & Data Collection

The **Lombardy Region** (Northern Italy | ~ 23870 km²) is chosen as study area. The study makes use of the following datasets:

- **GL30**: the most frequently updated (2000, 2010, and 2015 announced) high-resolution (30m) GLC multiple-class map currently available, employed as target map for the accuracy assessment
- **DUSAF**: the official land cover (vector) maps of Lombardy Region at a scale 1:10000, employed as reference map for the classification accuracy

Location of Lombardy Region in Northern Italy

Data Processing



Schematic of the data overlay procedure to compute sub-pixel error on a sample target map pixel

The DUSAF vector map is rasterized using a 5m pixel to take into account the minimum mapping unit declared by the producer. The dataset is then harmonized with the GL30 in terms of classification legend, coordinate reference system and projection.

To enable the investigation of error spatial patterns, an overlay procedure for sub-pixel errors detection between the target and the reference map was designed and implemented by means of GRASS GIS. This allows preserving both the original resolution of the reference map and the spatial reference (ID) of the target map pixels, by producing a single table including pixel-wise disagreement counts (i.e. errors) for each map class.

The table is processed by means of the DASK Python library that provides support for multithreading computation allowing manipulation and analysis of larger-than-memory datasets - such as the errors table of the case study (> 10 GB) - on a standard desktop machine .

GL30 class code	Class
10	Cultivated Land
20	Forest
30	Grassland
40	Shrubland
50	Water bodies
60	Wetland
70	Tundra
80	Artificial surfaces
90	Bareland
100	Permanent snow and ice

GlobeLand30 classification legend

Preliminary Results

1. Traditional Accuracy Assessment

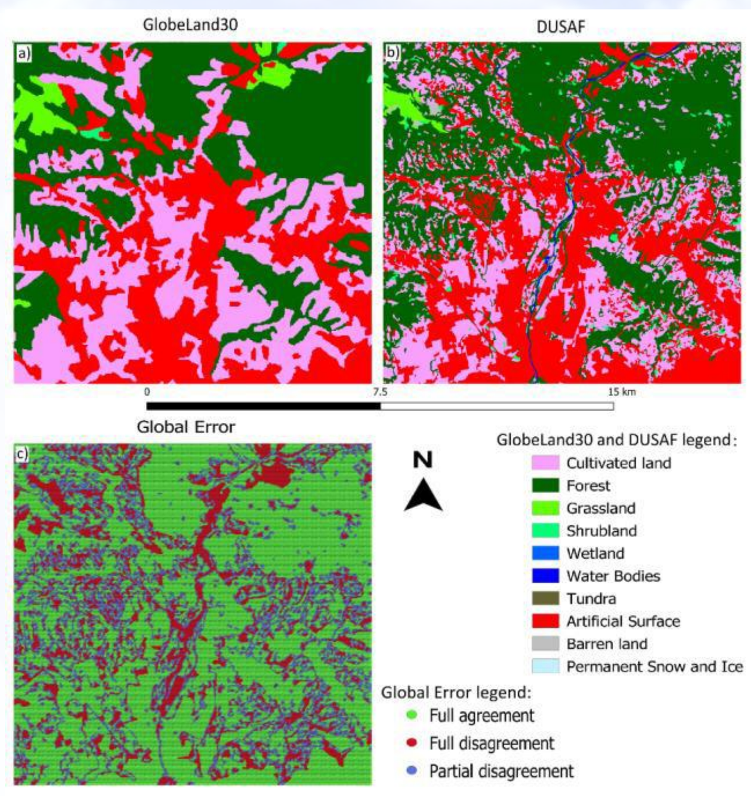
From the errors table, the confusion matrix is extracted. The computed Overall Accuracy of the GL30 map is 79% for the Lombardy Region. The agreement (diagonal values) of class 40 is the lowest, and that the highest confusion (extra-diagonal values) is between class 40 and class 20.

Class		GlobeLand30								
		10	20	30	40	50	60	80	90	100
DUSAF	10	90	11	1	20	35	9	30	1	0
	20	5	79	13	42	16	3	2	7	0
	30	1	3	51	13	4	0	1	7	0
	40	0	3	10	14	0	0	0	5	0
	50	0	0	0	0	35	1	0	0	0
	60	0	0	0	0	8	83	0	0	0
	80	3	2	0	2	1	2	66	0	0
	90	0	1	24	8	1	1	0	79	19
	100	0	0	0	0	0	0	0	0	81

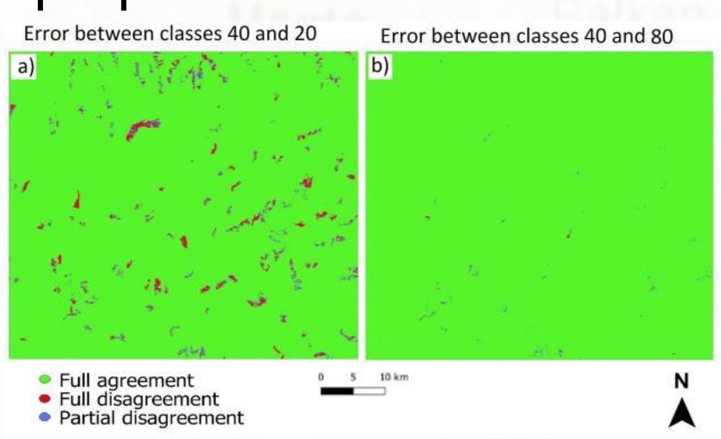
Normalized confusion matrix [%]

2. Error Spatial Patterns Investigation

Maps are derived from the errors table to provide visual insight into the spatial patterns of global, inter-class, or intra-class errors. To quantify these patterns in terms of spatial association, the global Moran's I index is computed. The index computed for the global error map is 0.80 confirms the marked positive spatial association that can be also visually detected on the global error map. Additional tests on the error spatial patterns can be performed starting from the proposed errors table.



Example of global error pattern maps



Example of inter-class error pattern maps

E.g.: Focusing on the highest and the lowest confusions for class 40, i.e. class 20 and class 80, the resulting Moran's I for these inter-class errors is respectively 0.82 and 0.62 thus providing evidence of an underlying connection between errors from the confusion matrix and the spatial association characterizing their patterns.

Conclusions & Acknowledgements

The traditional accuracy assessment provides robust indicators to describe the global accuracy of land cover maps but no insights into the errors spatial distribution. The proposed errors table provides a comprehensive input dataset to detailed accuracy assessments facilitating both visual and statistical analysis of error spatial patterns, thus - in turn – potentially improving the GLC maps accuracy evaluation at a local or regional scale.

This work was supported by URBAN-GEO BIG DATA, a PRIN project funded by the Italian Ministry of Education, University and Research (MIUR) - id. 20159CNLW8.