

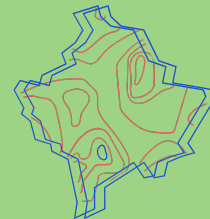
OpenStreetMap as an input source for producing governmental datasets: the case of the Italian Military Geographic Institute

Alessandro Sarretta¹, Maurizio Napolitano², Marco Minghini³

¹ Italian National Research Council – Research Institute for Geo-hydrological Protection, Padova, Italy

² Bruno Kessler Foundation, Trento, Italy

³ European Commission – Joint Research Centre (JRC), Ispra, Italy

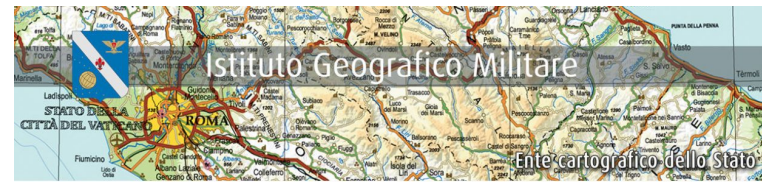


FOSS4G
Prizren, 2023

Introduction

- **Public sector** has historically been the sole responsible for producing & updating geospatial information
- **Citizen-generated data** has challenged this
 - most popular example: **OpenStreetMap** (OSM)
- Extensive research on OSM **quality** and/or **comparison** with authoritative data
- Multiple forms of integration between OSM & authoritative data
 - from the **public sector** – including through OSM **imports**
 - from the **business sector** (e.g. Overture Maps)

Background

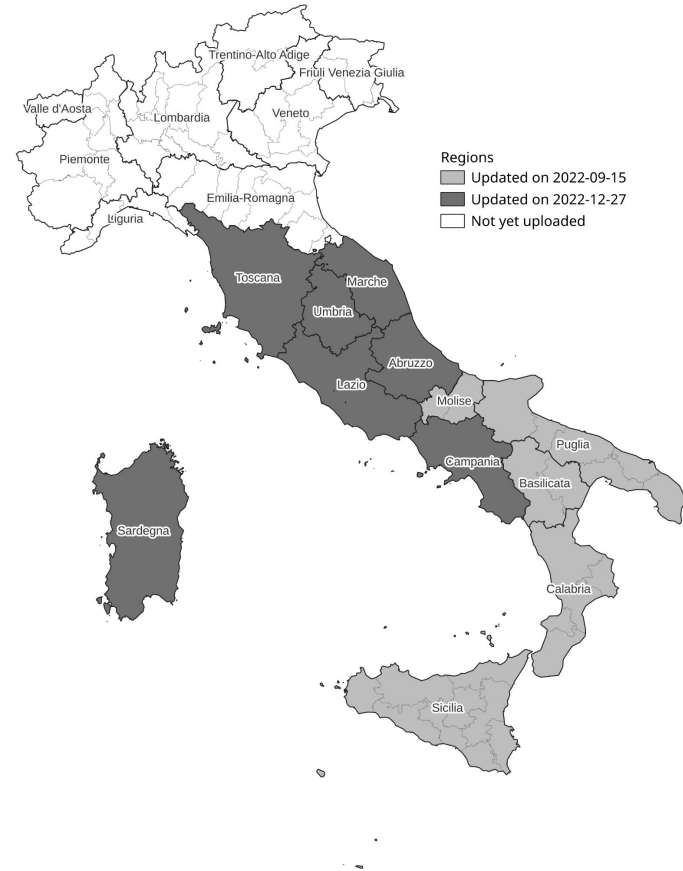


- The **Italian Military Geographic Institute** (IGM), one of the Italian governmental mapping agencies, released the **National Summary Database** (DBSN) in September 2022
 - <https://www.igmi.org/en/dbsn-database-di-sintesi-nazionale>



Background

- DBSN
 - **vector database** of geospatial information relevant for analysis and representation at the national level
 - corresponds to a **medium scale** (1:25,000); it summarises contents of regional databases at larger scales; it is used to derive maps at smaller scales (1:250,000)
 - currently includes data covering **12 out of the 20 Italian regions**



DBSN data sources

- The IGM used several data sources to create the DBSN, including **OSM**
- DBSN database is released free of charge under the Open Database License (**ODbL**), the same license of OSM data

Source code	Source institution	Explanation
01	<i>igm</i>	IGM
02	<i>ortofoto AGEA</i>	AGEA orthophoto
03	<i>osm</i>	OpenStreetMap
04	<i>db_regionale</i>	regional database
...

☐ **Accetto la licenza d'uso.**

La base di dati denominata DBSN (DataBase di Sintesi Nazionale) e' resa disponibile con Licenza Open Data Commons Open Database License (ODbL) ver. 1.0 <https://opendatacommons.org/licenses/odbl/1-0/>.

IGM non e' responsabile per qualunque tipo di danno diretto, indiretto o accidentale derivante dall'utilizzo delle informazioni contenute nella base di dati.

Objective and methodology

- Identifying the actual **role played by OSM** in the production process of the DBSN (and the reasons for that)
 - **analysis** of the DBSN and OSM data models
 - **mapping** between the DBSN and OSM data models
 - assessment of **OSM contribution to DBSN** in all Italian regions & provinces
 - **comparison** between OSM and DBSN for specific objects
 - buildings
 - roads and railways

DBSN data model

- Specifications are defined in an Annex to the **Ministerial Decree** of 10 November 2011 “*Technical rules for the definition of content specifications of geotopographical databases*”
- **Hierarchical structure** composed of
 - 10 layers
 - 30 themes
 - 93 classes (with their attributes)

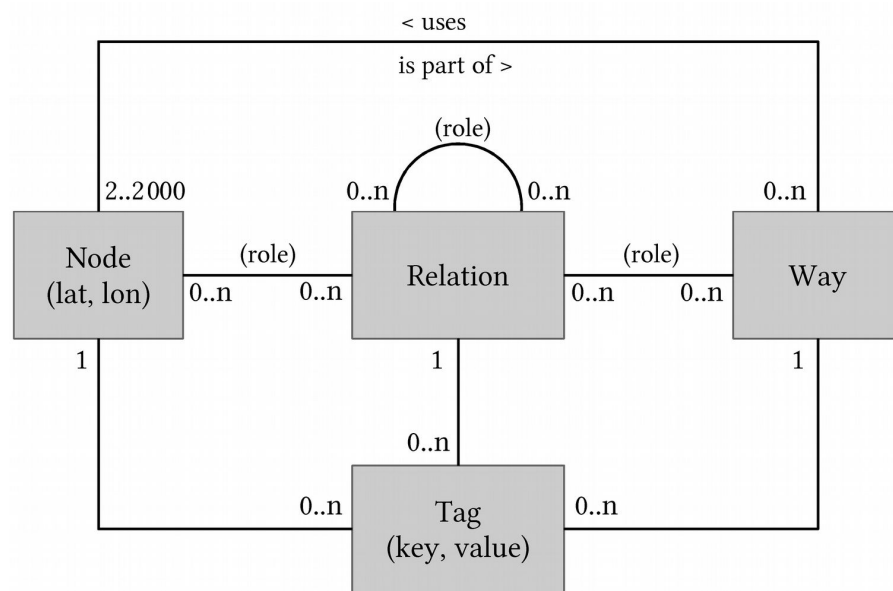
DBSN layer	DBSN theme
Geodetic and photogrammetric information	Geodetic information
	Cartographic and meta-information
Roads, mobility and transport	Roads
	Railways
	Other transport
Buildings and human settlements	Buildings
	Man-made constructions
	Transport infrastructure works
	Soil support and defence works
	Hydraulic, defence and hydraulic regulation works
Hydrography	Inland and transitional waters
	Marine waters
	Glaciers and perennial snow-fields
	Hydrographic network
Orography	Altimetry
	Bathymetry
	Terrain forms
	Digital terrain models (tin, dem/dtm)
Vegetation	Agro-forestry areas
	Urban green
Underground utility networks	Water supply network
	Electricity network
	Gas distribution network
	Oil pipelines
Significant places and cartographic markings	Significant places
	Cartographic markings
Administrative areas	Local authority administrative areas
Appurtenant areas	Transport services
	Appurtenances
	Quarries-dumps



OpenStreetMap data model



- 3 geometric primitives
 - nodes
 - ways
 - relations
- attributes (**tags**)
 - key-value pairs
 - minimum one tag per object



Source: Ramm F., Topf J. & Chilton S. (2010). *OpenStreetMap: Using and Enhancing the Free Map of the World*. Cambridge: UIT.



Mapping between DBSN & OSM data models

- DBSN layers, themes & classes are **semantically matched** with corresponding OSM objects through tags.
- Some of them have no match as they are beyond the scope of OSM.

DBSN theme	OSM tags
Roads	highway=*
Railways	railway=*
Other transport	aerialway=*, route=ferry, ...
Buildings	building=*
Man-made constructions	man_made=*
Transport infrastructure works	bridge=*, highway=* + tunnel=*, ...
Soil support and defence works	man_made=embankment
Hydraulic, defence and hydraulic regulation works	waterway=dam, man_made=dyke, embankment=dyke



Analysis

1. Assessment of OSM contribution to DBSN

Data download

Python scripts available at:
<https://github.com/napo/dbsnosmcompare>

DBSN: <https://www.igmi.org/en/dbsn-database-di-sintesi-nazionale>

OSM: <https://osmit-estratti.wmcloud.org>

OSM as source

Data enrichment

Aggregation by
province/region

Buildings/roads & railways: area/
length calculation for both datasets

Buildings: spatial intersection for
calculating % of OSM/DBSN

2. Comparison of DBSN & OSM building, road and railway datasets

OSM contribution to DBSN – Regions

- Percentage of DBSN objects derived from OSM, for each layer & theme

	Layers								
	Geodetic and photogrammetric information	Roads, mobility and transport	Buildings and human settlements	Hydrography	Orography	Vegetation	Underground utility networks	Significant places and cartographic markings	Administrative areas
Abruzzo		0.0	0.0				0.3		1.8
Basilicata		0.5	0.0				1.3		0.0
Calabria		0.0	0.3				0.1		5.4
Campania		24.2	0.1	0.0	0.0		0.8		4.2
Lazio		0.1	0.0				0.0		0.7
Marche		2.2	0.2						2.0
Molise		20.8	1.8			1.9	94.0		51.8
Puglia		0.3	1.1				0.1		3.9
Sardegna		0.3	0.0						0.1
Sicilia		3.0	0.1				0.2		11.2
Toscana		3.4	0.2				99.6		13.0
Umbria		95.9	0.8				0.2		16.2

Abruzzo	0.0	0.0	0.0	0.0	4.7					0.3			0.6	2.4	0.9
Basilicata	0.5			0.0	0.1					1.3				0.0	
Calabria	0.0	2.4		0.0	1.0					0.1			1.6	6.2	1.9
Campania	23.6	88.5	20.0	0.0	1.2	0.1	0.8	0.0		0.8				5.3	0.2
Lazio	0.0	1.6		0.0						0.0				0.5	5.3
Marche	2.2	5.2		0.1		33.4								2.3	
Molise	21.2	5.5	100	0.0	77.3				1.1	100	94.0		8.3	60.0	22.2
Puglia	0.3	0.0		0.0	8.5					0.1			0.1	5.1	3.0
Sardegna	0.3	1.3		0.0										0.2	
Sicilia	2.8	1.3	6.3	0.1	0.1	0.4				0.2			0.9	11.4	34.5
Toscana	3.4	0.5		0.2	1.6					99.6			0.3	16.2	0.3
Umbria	96.1	89.0		0.6	1.4					0.2			0.6	18.9	7.0

Roads															
Railways															
Other transport															
Buildings															
Man-made constructions															
Hydraulic, defence and hydraulic regulation works															
Inland and transitional waters															
Terrain forms															
Agro-forestry areas															
Urban green															
Electricity network															
Transport services															
Appurtenances															
Quarries-dumps															



OSM contribution to DBSN – Regions

- Highly **variable contribution of OSM** as a source of information for the DBSN among the 12 available regions
- For DBSN layers, in half of the cases OSM contribution is lower than 1%
- 16 out of the 30 DBSN themes do not include objects derived from OSM
 - *Geodetic information; Cartographic and meta-information; Transport infrastructure works; Soil support and defence works; Marine waters; Glaciers and perennial snowfields; Hydrographic network; Altimetry; Bathymetry; Digital terrain models (tin, dem/dtm); Water supply network; Gas distribution network; Oil pipelines; Significant places; Cartographic markings and Local authority administrative areas*
- Percentages depend on the total number of objects
 - a high percentage does not necessarily correspond to several OSM objects being used

OSM contribution to DBSN – Provinces

- Percentage of DBSN objects derived from OSM, for each theme

Abruzzo	Chieti	0.0	0.0	0.0	4.5					0.9	2.3
	L'Aquila	0.0	0.1	0.1	0.0				0.4	1.8	0.3
	Pescara	0.0			2.0			0.5	4.2	1.1	2.0
	Teramo	0.0		0.0	8.0			0.7		5.0	
Basilicata	Matera	0.0		0.0	0.0						
	Potenza	0.9			0.1			1.8		0.1	
Calabria	Catanzaro	3.3		0.0	0.4			0.1		6.2	2.3
	Cosenza	0.0	2.3	0.0	2.5				1.0	3.8	2.0
	Crotone	1.3			0.6			0.1	6.7	10.7	4.7
	Reggio di Calabria	0.0	2.5	0.0	0.0			0.1	1.4	8.6	0.7
	Vibo Valentia	1.8		0.0	0.1			0.2	6.3	5.9	
Campania	Avellino	18.0	79.3	0.0	4.2	1.1		1.1		3.5	
	Benevento	16.9	76.7	0.0	5.0	0.7		1.0		2.9	
	Caserta	26.0	90.2	0.0	0.7	0.6	0.3	10.5		9.6	1.1
	Napoli	47.5	91.4	33.3	0.0	0.1	3.4	6.4		5.8	
	Salerno	16.4	88.7	0.0	0.5	0.1	0.8	0.2		4.3	
Lazio	Frosinone	0.0									
	Latina	0.1	0.1								
	Rieti										
	Roma	0.0	2.8	0.0						0.2	
Marche	Viterbo	0.0		0.0				0.0		3.6	21.7
	Ancona	3.6	0.1	0.2		43.5				4.2	
	Ascoli Piceno	3.0	18.4	0.1						0.4	
	Fermo	1.8	12.7	0.1		48.5				1.1	
	Macerata	1.4	8.9	0.2		26.2				1.3	
Molise	Pesaro e Urbino	1.6		0.1		35.0				3.4	
	Campobasso	19.9	3.6	0.0	82.0		1.4	100	93.5	13.6	63.0
	Isernia	24.3	8.3	100	0.0	68.0	0.6	100	94.9	54.1	22.2

Roads
Railways
Buildings
Other transport
Man-made constructions
Hydraulic, defence and hydraulic regulation works
Inland and transitional waters
Agro/forestry areas
Urban green
Electricity network
Transport services
Appurtenances
Quarries-dumps

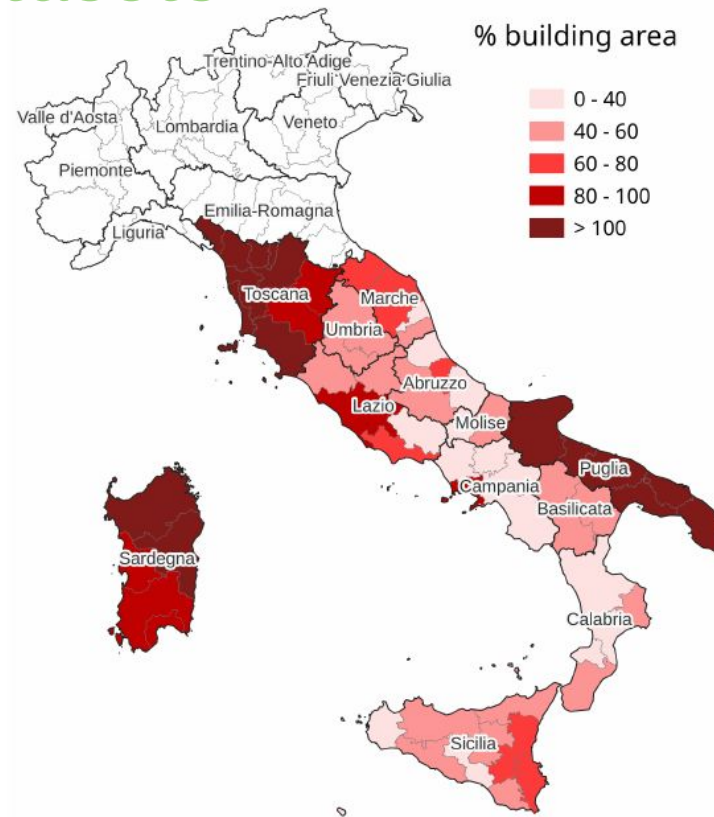
Puglia	Bari	0.2		0.0	6.2					6.1	15.0
	Barietta-Andria-Trani	0.1		0.2	8.2					13.3	4.0
	Brindisi	0.3			10.4						
	Foggia	0.6		0.0	13.5			0.1	0.3	13.1	1.0
	Lecce	0.3			4.8					0.3	2.2
Sardegna	Taranto	0.2		0.0	12.7					0.3	0.3
	Cagliari	0.1	11.5							0.2	
	Nuoro	0.1									
	Oristano	0.0		0.0							
	Sassari	0.8		0.0							1.4
Sicilia	Sud Sardegna	0.1									
	Caltanissetta	2.9		0.0	0.1					14.6	51.9
	Catania	3.1	0.4	13.3	0.4	0.0				1.7	33.3
	Enna	2.5	0.3	0.0	0.1					9.1	60.0
	Messina	5.1	3.5	0.3	0.0					0.2	
Toscana	Palermo	2.3	1.4	0.0	0.0						37.5
	Ragusa	2.6	0.3	0.0	0.0					8.3	19.0
	Siracusa	3.2	1.7	0.1	0.0					1.6	20.6
	Trapani	2.4		0.0	0.4	0.8				6.1	22.7
	Arezzo	0.4		0.0	1.1					99.5	12.7
Umbria	Firenze	0.4	2.1	0.1	3.0					99.9	13.3
	Grosseto	8.5		0.1	1.0					99.7	18.0
	Livorno	0.6		0.1	0.8					100	1.4
	Lucca	0.9		0.1	0.3					100	21.3
	Massa Carrara	4.3		0.7	7.9					98.5	14.0
	Pisa	0.4		0.4	1.1					99.9	12.6
	Pistoia	1.1		0.1	1.8					100	28.3
	Prato	0.2		0.0	3.1					100	29.6
	Siena	7.9		0.4	0.6					98.9	17.3
	Perugia	95.9	88.5	0.5	0.1					0.1	1.2
	Terni	96.8	89.9	0.9	12.5					0.3	11.8

Roads
Railways
Buildings
Other transport
Man-made constructions
Hydraulic, defence and hydraulic regulation works
Inland and transitional waters
Agro/forestry areas
Urban green
Electricity network
Transport services
Appurtenances
Quarries-dumps



Comparison of building datasets

- Ratio between the **total area of buildings** in OSM & in the DBSN
 - indirect measure of **completeness**
 - only in 3 regions (Sardegna, Toscana, Puglia) the areas of OSM buildings is almost equal to that of DBSN buildings
 - high **variability** across regions & provinces - possible reasons
 - demographic density
 - attractiveness
 - presence or absence of OSM local communities
 - imports



Comparison of building datasets

- Ratio between the **total area of buildings** in OSM & in the DBSN for each region
 - high variability, from 35% to 100+%
 - high variability of the standard deviation across provinces in the same region

Region	% OSM/DBSN building area	St. dev. of % OSM/DBSN building area for provinces	% OSM building area disjoint from DBSN
Abruzzo	44.2	9.9	2.9
Basilicata	46.8	7.4	3.1
Calabria	35.4	11.6	2.2
Campania	50.6	30.8	3.1
Lazio	71.9	22.2	3.9
Marche	57.8	16.7	6.7
Molise	47.6	7.8	3.1
Puglia	105.2	1.7	6.4
Sardegna	99.1	4.1	5.2
Sicilia	50.0	11.9	2.3
Toscana	103.7	5.2	7.1
Umbria	51.9	7.2	2.0

Comparison of building datasets

- Fraction of the area of OSM buildings **not intersecting** any DBSN building
 - values usually low, but this proves that OSM includes buildings not available in DBSN

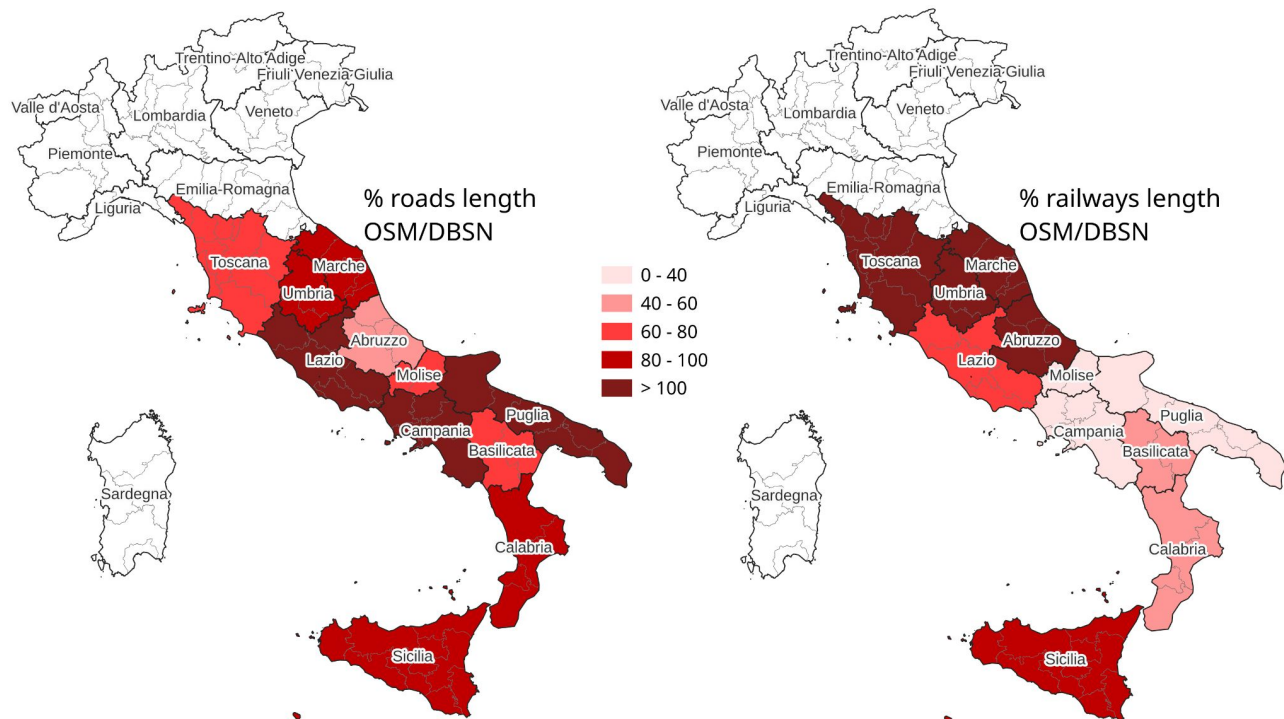
Region	% OSM/DBSN building area	St. dev. of % OSM/DBSN building area for provinces	% OSM building area disjoint from DBSN
Abruzzo	44.2	9.9	2.9
Basilicata	46.8	7.4	3.1
Calabria	35.4	11.6	2.2
Campania	50.6	30.8	3.1
Lazio	71.9	22.2	3.9
Marche	57.8	16.7	6.7
Molise	47.6	7.8	3.1
Puglia	105.2	1.7	6.4
Sardegna	99.1	4.1	5.2
Sicilia	50.0	11.9	2.3
Toscana	103.7	5.2	7.1
Umbria	51.9	7.2	2.0

OSM buildings not available in DBSN

- Possible reasons
 - “in favour” of OSM
 - current workflow adopted by the IGM to filter the OSM database (through tags) **fails to capture** potentially relevant objects
 - the continuous evolution of the OSM database happens at a **pace** that the IGM can hardly cope with
 - “against” OSM
 - elements not included in DBSN as buildings, e.g. greenhouses or roofs
 - demolished buildings still in OSM
 - ‘imperfect’ overlap between the DBSN and OSM building datasets

Comparison of road and railway datasets

- Ratio between the **total length of roads & railways** in OSM & in the DBSN
 - roads are more complete than railways
 - variability is higher for railways



Conclusions & discussions points

- Official cartographic production needs (also) **crowdsourced data** to close some information gaps
 - especially in baseline data layers
- OSM can be a reference source also for governmental bodies *but* it is **not yet ready** (at least as far as building coverage is concerned)
- Interesting case in terms of **data reuse/licensing**



Future work

- Extend the analysis to the remaining 8 Italian regions once the DBSN is released
 - spot possible **geographical trends**
 - identify other OSM objects having a high(er) potential for integration in the DBSN
- Perform analyses on additional **DBSN layers and themes**
- Evaluate **correlation** between OSM quality & OSM use in the DBSN
- Assess the potential to use the DBSN as a source for **imports or data update** in OSM



Paper



OPENSTREETMAP AS AN INPUT SOURCE FOR PRODUCING GOVERNMENTAL DATASETS: THE CASE OF THE ITALIAN MILITARY GEOGRAPHIC INSTITUTE

A. Sarretta ¹, M. Napolitano ², M. Minghini ^{3*}

¹ Research Institute for Geo-Hydrological Protection, National Research Council (CNR-IRPI), Padova, Italy –
alessandro.sarretta@irpi.cnr.it

² Fondazione Bruno Kessler, Trento, Italy – napo@fbk.eu

³ European Commission, Joint Research Centre (JRC), Ispra, Italy – marco.minghini@ec.europa.eu

KEY WORDS: Data integration, Data quality, Geospatial, Governmental dataset, Italian Military Geographic Institute, OpenStreetMap, Open data, Volunteered Geographic Information.

ABSTRACT:

The role of Volunteered Geographic Information (VGI) to integrate, update or complement authoritative datasets released by governments has become increasingly important. This work analyses the contribution of OpenStreetMap (OSM), the most popular VGI project, as one of the input sources that the Military Geographic Institute (IGM), one of the Italian governmental mapping agencies, has used for producing the National Summary Database (DBSN). This database, which was recently released for 12 out of the 20 Italian regions, has a schema organised into a hierarchical structure composed of 10 layers, 30 themes and 93 classes, where each geospatial object carries information on the specific data source it was derived from. For each DBSN layer and theme, we first calculated the fraction of objects derived from OSM in all the Italian regions and related provinces. We found a heterogeneous picture with OSM contribution generally being limited, with the exception of few regions and layers/themes where the DBSN was almost exclusively derived from OSM. An in-depth comparison between the DBSN and OSM building datasets showed that OSM building completeness is varying across Italian regions and provinces, but in all regions there are buildings in OSM that are not included in the DBSN. The work shed light on the opportunities and obstacles for OSM to become a primary input source for the production of governmental datasets.

<https://isprs-archives.copernicus.org/articles/XLVIII-4-W7-2023/193/2023/>

Thank you for the attention!

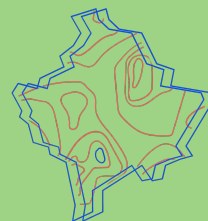


marco.minghini@ec.europa.eu



@MarcoMinghini

slides: https://bit.ly/OSM-DBSN_analysis



FOSS4G

Prizren, 2023