

TOWARDS A PAN-EU BUILDING FOOTPRINT MAP BASED ON THE HIERARCHICAL CONFLATION OF OPEN DATASETS: THE DIGITAL BUILDING STOCK MODEL - DBSM

P. Florio^{1*}, C. Giovando², K. Goch³, M. Pesaresi¹, P. Politis⁴, A. Martinez¹

¹European Commission, Joint Research Centre (JRC), Ispra, Italy – (pietro.florio, martino.pesaresi, ana.martinez)@ec.europa.eu

²European Dynamics S.A, Luxembourg – cristiano.giovando@ext.ec.europa.eu

³Institute of Geography and Spatial Organization Polish Academy of Sciences, Warsaw, Poland - kgoch@twarda.pan.pl

⁴European Dynamics S.A., Belgium – panagiotis.politis@ext.ec.europa.eu

KEY WORDS: Building footprints, built-up, satellite imagery, open data, open maps, pan-EU maps, energy transition

ABSTRACT:

This paper presents a hierarchical conflation process applied to open datasets for the creation of a seamless pan-European map of building footprints in vector format, named Digital Building Stock Model – DBSM. The objective is the sequential addition of input components (which currently include OpenStreetMap, Microsoft GlobalML Building Footprints, European Settlement Map), taking into account their limitations, and aiming at the highest level of completeness possible, for planning and evaluating energy transition scenarios at the EU level. The results indicate how DBSM compares robustly against cadastral data from Estonia, used as reference area. The comparison of DBSM with GHS-BUILT-S, a 10 metres resolution grid with worldwide coverage that encodes the built-up surface in each pixel as derived from Sentinel-2 imagery for the year 2018, reveals a relative overestimation of the latter, factored by 0.68 at the EU scale for a sound match.

1. INTRODUCTION

1.1 Context

There are numerous applications that can benefit from the use of reliable harmonized and comprehensive pan-EU maps of the building stock provided in vector format, publicly available, at a level-of-detail LOD0 (according to the CityGML standard), where the buildings' footprints can be identified. These maps are essential for applications in many fields, including architecture, civil and environmental engineering, urban design, energy planning and disaster risk management.

European countries provide vector maps of their building stock through a variety of levels of detail, formats, and tools; data across countries is often heterogeneous in terms of attributes, accuracy and temporal coverage (Biljecki et al., 2021), available through different user interfaces, or hardly accessible due to language barriers. Aggregated data at Member States level are available through the EU Building Stock Observatory¹. Bottom-up solutions from local cadastral data in the framework of the INSPIRE initiative (European Commission. Joint Research Centre., 2016) and top-down standard-setting regulations like the EU Regulation 2023/138 (Commission Implementing Regulation (EU) 2023/138 of 21 December 2022 Laying down a List of Specific High-Value Datasets and the Arrangements for Their Publication and Re-Use., 2022), are increasing and improving the homogeneity in the data availability. Two examples of this are 1) the Copernicus CORDA's multi-country thematic dataset for buildings² and 2) the effort of Eurostat collecting data from authoritative data providers³, which is currently work-in-progress.

Nevertheless, the cost of producing comprehensive vector datasets of building footprints responding to authoritative cartographic standards remains high; even when they are made promptly available, these datasets become soon outdated as the built-up state evolves.

This factor encouraged the development of crowd-sourced providers of building footprint vectors like those included in the OpenStreetMap database, which are gradually covering many countries of the European Union. However, the source of such data is not consistent, with reference, for example, to the date of the imagery used to digitise the building footprints.

Simultaneously, improvements in Earth observation technology resulted in increased resolution of satellite imagery allowing for automatic building footprints segmentation on very high-resolution images based on deep learning algorithms: major organisations in the field of information technology, such as Google and Microsoft, were able to quickly create and publicly disseminate large vector datasets with extensive global coverage. However, such datasets rely on commercial Earth imagery, which affected by some uncorrected distortions. Other research institutions released grid-based maps of built-up areas, covering (a) the whole world at the resolution of 10 metres – e.g., the Built-Up Surface of the Global Human Settlement Layer – GHS-BUILT-S (Pesaresi and Politis, 2022); (b) Europe at the resolution of 2 metres – e.g., the European Settlement Map (Szabo et al., 2023). These grids rely on lower resolution sensors with fixed capture angle, such as Landsat, Sentinel-2 for the former (a) and optical Very High Resolution imagery for the latter (b). Another initiative called EUBUCCO (Milojevic-Dupont et al., 2023) has compiled a vector database of individual building footprints for 200+ million buildings across the 27

* Corresponding author

¹ https://energy.ec.europa.eu/topics/energy-efficiency/energy-efficient-buildings/eu-building-stock-observatory_en

² <https://corda.eea.europa.eu>

³ <https://gisco-services.ec.europa.eu/pub/Inspire/ANNEX-1/Buildings/> (to be accessed from: <https://ec.europa.eu/eurostat/web/gisco/geodata/reference-data>).

European Union countries and Switzerland, by merging 50 open government datasets and building footprints from OpenStreetMap, which have been collected, harmonized and partly validated.

1.2 Research objective

The methodology presented here provides a replicable workflow for generating seamless building datasets for each of the EU-27 countries, by combining the most complete available public datasets into a single one, called Digital Building Stock Model (DBSM). The novelty of the presented approach lies in an open and reproducible workflow that allows the database to be recreated at reasonable computation cost using the most up-to-date open data. The development of DBSM was mainly driven by the need of highly precise vector data for continental scale building energy modelling, to foster the implementation of the recent energy transition measures agreed within the European Union.

The final dataset is more comprehensive than the individual input layers and approximates robustly authoritative building footprint data issued from National Mapping and Cadastral Agencies (NMCAs).

2. INPUT DATA AND METHODS

2.1 Input data

After reviewing existing literature and assessing publicly available building footprint data sources, the following core input datasets were identified, besides cadastral maps of buildings from NMCAs (where available):

- OpenStreetMap (OSM)⁴: a free and open-source global dataset of geographic features, including building footprints and attributes;
- Microsoft GlobalML Building Footprints (MSB)⁵: a freely available dataset of building footprints developed by Microsoft using machine learning algorithm on very high-resolution satellite imagery;
- European Settlement Map (ESM) (Szabo et al., 2023): raster dataset of built-up areas at 2-meter spatial resolution, classified using Convolutional Neural Networks from imagery available through the Copernicus⁶ services.

Building footprints are available in OpenStreetMap across all 27 EU countries, but with different levels of completeness and coverage. Human contributors trace data in OSM manually from circa 50 cm resolution imagery, thus the available building footprints are considered of higher geometric quality compared to those extracted by machine learning algorithms of the MSB and ESM datasets. Microsoft provides high resolution building footprints for all 27 countries, but their coverage within the country areas varies considerably. The ESM dataset is derived from a seamless mosaic covering the entire EU-27 area, so it is considered being the most complete in terms of coverage, although the lower resolution (2 metres) and quality does not allow for extracting detailed building footprints as available with OSM and MSB.

Additionally, the Global Built-Up Surface of the Global Human Settlement Layer – GHS-BUILT-S (Pesaresi and Politis, 2022) is used as an independent source of built-up estimates for comparison purposes. GHS-BUILT-S (Pesaresi and Politis, 2022) provides a 10 metres resolution grid that encodes the built-up surface in each pixel for the year 2018. At an aggregated level, the comparison of DBSM with this data source is informative to get an understanding of their level of agreement and potential mutual improvement.

2.2 The conflation process

The combination of the above-listed datasets is carried out with a stepwise hierarchical approach aiming at “filling the gaps”, which starts from the dataset with the highest presumable reliability, downstream until the less reliable one (*Figure 1*).

First, building footprints cadastral data from authoritative sources should be added to the map: at the current stage, though, this step is not yet implemented. As a first step, at the moment, footprints are extracted from OSM ways and relations through the OGR2OGR utility, via a query based on the “buildings” tag. Subsequently, the MSB dataset is compared to OSM: MSB features are checked for topological validity, discarding the invalid ones (e.g. self-intersecting) and dissolved to flatten overlapping or adjacent footprints; after conversion to single-part features, MSB buildings are selected whenever they overlap or intersect OSM for less than 20% of the surface. This threshold is deemed sufficient to avoid most duplicates, but cannot withstand uncorrected parallax distortions in the source imagery of MSB at the basis of many footprints misplacements that persist. Before conflation, MSB buildings below 40 m² of surface are filtered out as outliers. This cut-off is based on the assumption that, below this living standard, buildings in Europe are probably not inhabited all over the year, and consequently often unheated. Besides, the Energy Performance of Buildings Directive (Directive 2010/31/EU of the European Parliament and of the Council on the Energy Performance of Buildings (Recast), 2010), exempts buildings below 50 m² of floor area from setting minimum energy performance requirements: sometimes they are garages and other storage spaces; frequently they are related with wrong detections by the machine-learning classifier. Thus, they do not serve the scope of DBSM, which specifically focuses on the estimation of buildings energy demand. In the following step, the ESM data is clipped to match country boundaries at the highest resolution openly available online⁷, then compared to the combined OSM and MSB buildings and vectorised, to fill in any gap that is not covered by the latter. ESM-derived building footprints are dissolved, converted to single-part features and buffered with a negative offset of 4 metres, to reduce the area overestimation and improve the building footprint delineation. Holes smaller than 500 m² are removed to create coherent shapes and limit concavity, even if this may result in including unroofed areas in the proximity, like courtyards and streets. The resulting polygons are filtered to retain only features above 100 m² of surface. Such filter is arbitrarily chosen to exclude probable incorrect detections, intrinsic in the ESM data given the resolution of its source imagery (2 metres). The output of this process is incorporated in the DBSM dataset, whenever their overlap with the intermediate conflation OSM+MSB does not exceed 30% of the surface, to avoid duplicates.

⁴ www.openstreetmap.org

⁵ Microsoft Global ML Building Footprints version 2022-07-08. <https://github.com/microsoft/GlobalMLBuildingFootprints>

⁶ https://spacedata.copernicus.eu/web/guest/optical-very-high-resolution-coverage-over-europe-vhr_image_2018-plus-vhr_image_2018_enhanced-and-dem_vhr_2018-

⁷ EUROSTAT GISCO Countries 2020, scale 1 : 1 million <https://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/administrative-units-statistical-units/countries>

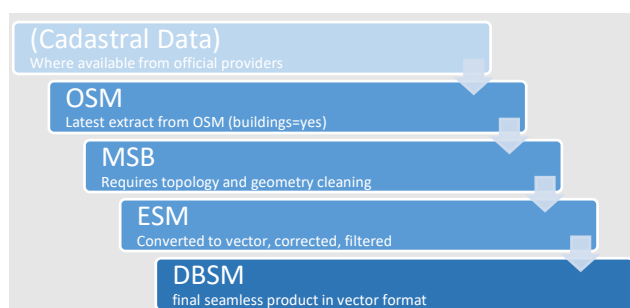


Figure 1. Overview of the conflation process.

To implement and automate the described logical workflow, an interactive model is developed to work in the popular QGIS desktop software and attached to this paper's online resources. The QGIS model builder allows for building logical processing workflows by linking input data forms, variables and all the analysis functions available in the software. The conflation process is conducted at the country level, since OSM and MSB sources are already conveniently provided in country extent packages. Depending on the geographic size of each country and the amount of data included, some countries are further split into tiles for processing. The resulting building footprints from each input dataset are kept in separate files for easier handling, but can be combined visually in GIS software or physically merged in a single file. All datasets are re-projected to EPSG:3035⁸ standard geographic projection for Europe and saved in the FlatGeobuf⁹ efficient binary format.

2.3 Comparison with GHS-BUILT-S

The DBSM buildings dataset is compared with the European Commission's GHSL Built-up surface layer – GHS-BUILT-S (Pesaresi and Politis, 2022) to get an understanding of the respective coverage at pan European level. GHS-BUILT-S is a spatial raster dataset depicting the distribution of the built-up surfaces estimates between 1975 and 2030 in 5 years intervals, in 100 m resolution. The dataset is created through the spatio-temporal interpolation of five observed collections of satellite imageries: Landsat (MSS, TM, ETM sensor) supports the 1975, 1990, 2000, and 2014 epochs; Sentinel-2 composite supports the 2018 epoch (Corbane et al., 2020). For the temporal anchor point of 2018 the data is available at finer 10 m resolution (GHS), as observed from the S2 image data (European Commission. Joint Research Centre., 2022). The 10 m resolution layer for reference year 2018 (referred to as GHS for brevity) is used for comparison with DBSM building dataset. In fact, GHS is an independent resource to compare against, as it relies on Sentinel-2 imagery, fully captured in year 2018. This eliminates all discrepancies between country authorities, typical of cadastral building data, and temporal misalignments, typical of community-based data like OSM.

The assessment of previous GHSL built-up surface estimates showed a tendency for the built-up surface overestimation (Uhl and Leyk, 2022). Therefore, it is necessary to compute an adjustment factor, corresponding to the ratio of the built-up area estimated using the reference DBSM data, to the built-up area derived from the GHS layer. The adjustment factor can facilitate the evaluation of built-up area derived from GHS-BUILT-S layers, given the temporal accuracy and the coverage of the reference data used. Here, the adjustment factor is estimated at

pan-European level using DBSM data as reference. First, it is necessary to compute the ratio of the built-up area derived from DBSM to the built-up area derived from GHS on a 10 x 10 km aggregated grid level, excluding the grids with no built-up area in any of these layers. The same estimation is performed at the country level. Secondly, the grid samples are subset by range of one standard deviation of the area ratio, from the mean value among all grids. The GHS adjustment factor is computed as the median value of area ratios from the subset samples (Figure 2, top).

For the spatially explicit pan-European level analysis, the GHS layer is multiplied by the computed adjustment factor to obtain the adjusted layer GHS', mitigated for the built-up surface overestimation. The DBSM dataset is compared to the adjusted GHS' layer at pan-European level, by 10 km-side grid cells. The completeness check of DBSM data against the adjusted GHS' layer requires computing the ratio of the difference between areas derived from both layers to their sum (a completeness check). The completeness check results in an indicator with values ranging from -1 to +1, where -1 refers to a situation where only GHS' data is available for a given grid cell, and +1 refers to the situation where only DBSM data is available for a given grid cell. Value 0 means the convergence of the two datasets in terms of built-up area per grid cell.

2.4 Local comparison with cadastral building data

A more focused look into the comparison with available cadastral data for a particular area of interest provides a preliminary understanding of the accuracy of the DBSM layer along with its limitations. The selected country of Estonia for the local comparison, given the completeness and soundness of the cadastral building data¹⁰, withstands the authors' careful visual inspection against updated very-high-resolution imagery.

The comparison includes built-up area surface derived from the adjusted GHS' layer and the DBSM layer, as well as with the DBSM input layers, before conflation: ESM, OSM and MBF (Table 1).

For each data source (GHS adjusted, DBSM and its input layers) a completeness check is performed against the cadastral building data, in 25 x 25 km² grid. Vector layers are rasterised to 1 metre resolution for comparison with raster layers. The cadastral building data is considered as the "ground truth" observation and the completeness check is calculated as the ratio of the difference between areas derived from the evaluated dataset and cadastral data to their sum: $\frac{x-cadastral}{x+cadastral}$. The completeness check results in an indicator with values ranging from -1 to +1, where -1 refers to a situation where only cadastral building data is available for a given grid cell, and +1 refers to the situation where only evaluated data are available for a given grid cell. Value 0 means the full agreement of the two datasets in terms of built-up area per grid cell.

The comparison features the total area of built-up surface in Estonia for each data source under consideration.

⁸ <https://epsg.io/3035>

⁹ <http://flatgeobuf.org/>

¹⁰ Estonian Land Board 3.10.2022:

<https://geoportaal.maaamet.ee/eng/Maps-and-Data/Estonian-Topographic-Database/Download-Topographic-Data-p618.html>

Data	Format	Spatial resolution	Temporal resolution
OSM	Vector	Rasterized to 1 m	Acquisition date: September 2022
ESM	Raster	2 m	2018
MSB	Vector	Rasterized to 1 m	2014-2022
GHS	Raster	10 m	2018
Cadastral building data for Estonia	Vector	Rasterized to 1 m	2022

Table 1. Data used for local comparison in Estonia

3. RESULTS

The conflated DBSM dataset is published under an open license together with the conference proceedings. A large share of DBSM is composed of OSM data (78%), followed by MSB footprints (14%) and ESM vectorisation (7%). Overall, there is sufficient match between DBSM and GHS-BUILT-S, the former estimating 0.66 times less total built-up surface in EU-27 compared to the latter (Figure 2, middle). The adjustment factor based on the comparison of GHS data against the DBSM data returns a correction value of 0.68 (Figure 2, top). The relative underestimation, represented by the DBSM / adjusted GHS' ratio, is not uniform throughout EU countries (Figure 2, below), but there is a fair homogeneity among continental Member States.

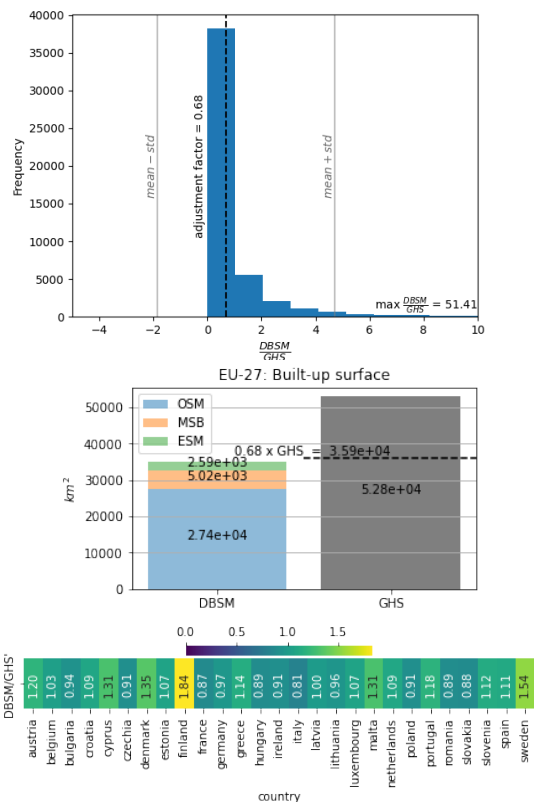


Figure 2. GHS adjustment factor based on the median grid-level ratio between DBSM and GHS built-up surface (top); Composition of the built-up surface of DBSM vs GHS (middle); Ratio between DBSM and GHS' (adjusted) built-up surface, by country (bottom).

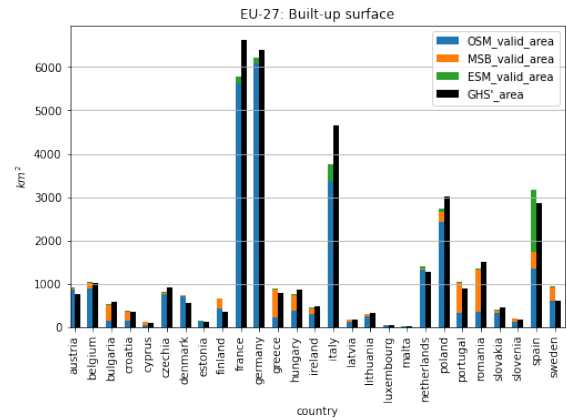


Figure 3. Composition of the built-up surface of DBSM vs GHS' (adjusted), by country in EU-27.

Larger countries (France, Germany, Italy, Spain, Poland) are affected by higher discrepancies, given the accumulation of differences in a larger extension (Figure 3). ESM becomes essential to cover sparse mountainous settlements in Spain, where there is no coverage from MSB and OSM. However, open cadastral data may complement the current datasets incorporated in DBSM. In many countries (Portugal, Greece, Romania, Bulgaria, Sweden), MSB makes a difference.

Southern and Eastern Europe feature wide zones of lower DBSM estimates, with some less pronounced spots of higher DBSM estimates, compared to adjusted GHS' (Figure 4). The former seems to match some specific administrative boundaries, remaining isolated in specific sub-regions. This may be linked with the incompleteness of OSM data derived from local authorities, or with the poor availability of imagery in such zones, affecting both the community-based and the remotely sensed digitalisation of footprints. On the contrary, Scandinavia marks a widespread under-detection of footprints in GHS data compared to DBSM.

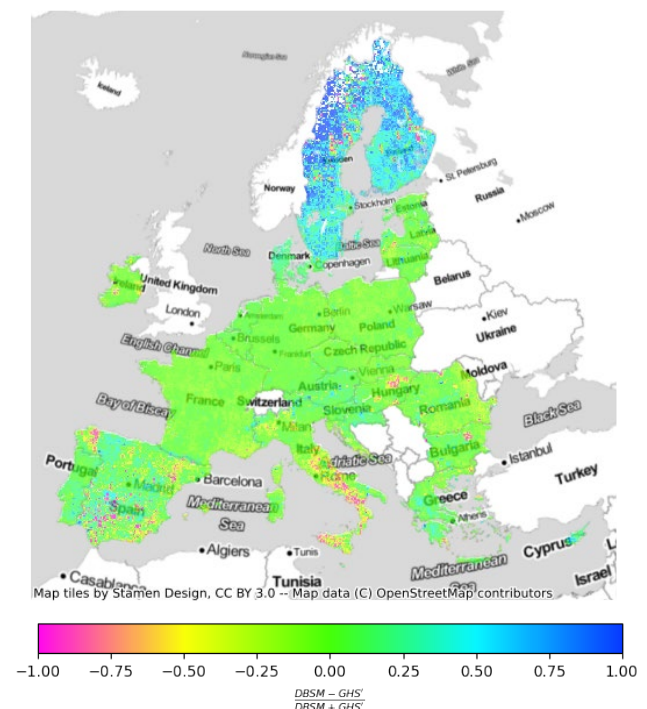


Figure 4. Map of the completeness of DBSM dataset against GHS' (adjusted), on a grid with 10 km resolution.

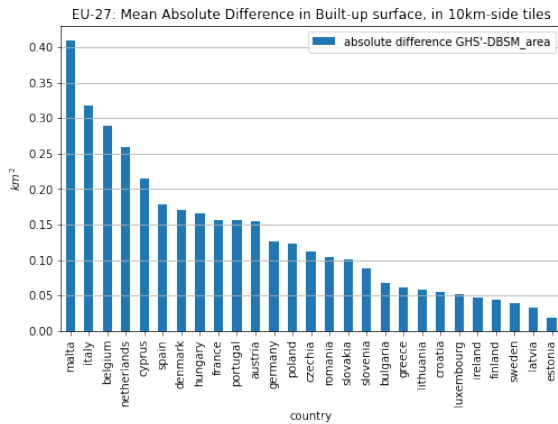


Figure 5. Mean Absolute Difference (MAD) between DBSM and GHS' (adjusted), on a grid with 10 km resolution, by country.

This may be linked, among other factors, with the temporal misalignment of datasets, with the recent urbanisation of some areas, as well as with the large presence of forestry, which decreases the performance of the GHS classification.

However, looking at the Mean Absolute Difference (MAD) by country (*Figure 5*) reveals that in Scandinavia, where settlements are sparse, with low cumulative built-up surface, the mismatch between DBSM and GHS is not significant in absolute terms.

It is possible to compare the performance of DBSM against its source components before conflation (OSM, MSB, ESM) and with cadastral data of spot countries. The case of Estonia (*Figure 6*, *Figure 7*) shows that, when assuming cadastral data as 100% (reference), the adjusted version of GHS-BUILT-S misses a small share of built-up surface, and it is best approximated by DBSM, while its individual components underperform.

Looking at the distribution of Built-up surface on the Estonian map (*Figure 7*) shows a significant discrepancy in the MSB dataset, which clearly misses some areas in the North-West of the country, along coastlines, including the capital city - Tallin.

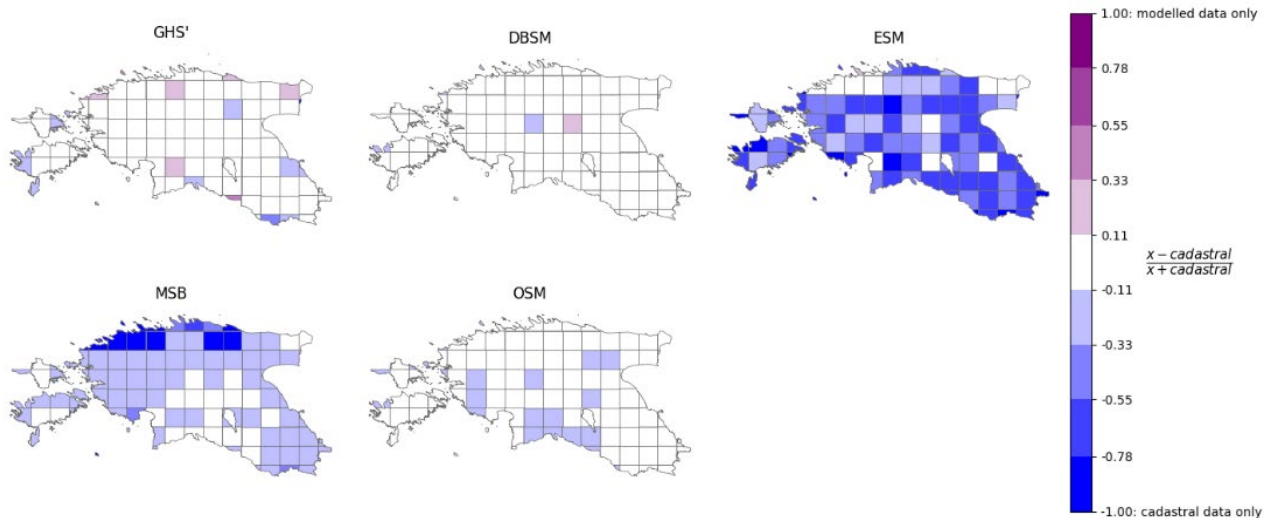


Figure 6. Map of completeness in Estonia derived from different data sources before conflation, in 25 x 25 km² tiles. Values for GHS layer are adjusted by factor 0.68 (GHS').

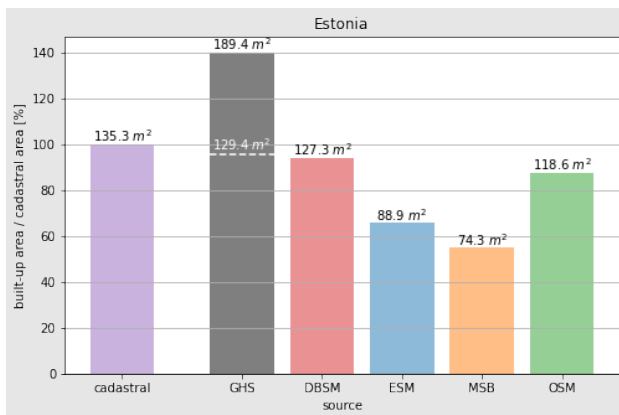


Figure 7. Built-up area in Estonia derived from different data sources, in relation to the built-up area derived from cadastral data. From the left: cadastral data, GHS and DBSM, and DBSM components before conflation: ESM, MSB and OSM data. Labels on top of the bars show the absolute value of built-up surface area. White dashed line and white label show the values for GHS data adjusted by 0.68 (GHS').

4. DISCUSSION AND CONCLUSION

This paper presents a conflation method to integrate heterogeneous data sources covering Europe, with the objective of minimising gaps and maximising completeness over EU-27. The conflation process is hierarchical to prioritise most reliable data sources and incorporates filters to minimise false positives. However, as datasets are merged in a progressive addition, false positives might propagate through the different steps. There are several known limitations to the data and the processing workflow:

- Many MSB building footprints present irregular geometries that are caused by faulty image interpretation or by image distortion. These can be filtered by calculating the vertex angle values of each polygon and removing specific outlier values. A method was tested at small scale, but it was not possible to implement it at country scale yet.
- The ESM geometries do not accurately describe the actual building footprints but only the rough block outline. While ESM has seamless coverage, its best application would be for guiding additional feature

extraction from VHR imagery in areas where OSM and MSB have poor coverage.

- The default overlap thresholds (i.e. 20%, 30%) could be tweaked and dynamically adjusted, based on the built-up pattern (e.g., lower in urban areas, higher in rural areas).
- Filters of minimum feature size of 40 m² for MSB and 100 m² for ESM can be optimised to find the most robust balance between including non-building features and actual smaller buildings.

Despite the limitations discussed above, DBSM approximates soundly cadastral data in Estonia, improving the coverage of its individual components taken separately (OSM, MSB and ESM). Such an assessment can be extended to other countries, where up-to-date cadastral data is available.

On the comparison side, a generalised over or underestimation factor weighting GHS or DBSM uniformly in the whole EU depends on the geographically inhomogeneous performance of such datasets. Moreover, in continental Europe, DBSM compares rather stably with GHS. As these two sources are independent of each other, the completeness check facilitates quantitative and visual analysis of both layers in terms of their completeness and accuracy. Further comparisons with cadastral data in Scandinavian and southern-European areas, will provide a better understanding of the differences encountered between DBSM and GHS-BUILT-S, including those generated by the temporal misalignment between the component datasets of DBSM and with GHS. From visual inspection, it emerges that areas where remotely sensed data are not well-performing include sparse settlements in forest areas in Scandinavia, dry highlands in Spain and Italy, rural areas in Eastern countries (Hungary, Bulgaria, Romania) and in Ireland, jagged coastlines, riverbeds that can be confused with roads, harbour areas that can be confused with buildings (e.g. in Malta).

The incorporation of cadastral-based datasets for Europe, like the ones consolidated in EUBUCCO or EUROSTAT GISCO, could increase the completeness of DBSM in the short term. In the medium-longer term, the provision of authoritative data by Member States in the framework of the High-Value Dataset Regulation will complete the building tessellation for EU, expected in the coming years. However, at its current stage, DBSM is deemed to start constituting already a valuable source for planning and evaluating energy transition scenarios at the EU level with sufficient effectiveness.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the technical advice of Carmen Maduta (European Commission, Joint Research Centre – JRC, Ispra, Italy, unit C.2) and of Michele Melchiorri (European Commission, Joint Research Centre – JRC, Ispra, Italy, unit E.1).

REFERENCES

Biljecki, F., Chew, L. Z. X., Milojevic-Dupont, N., and Creutzig, F., 2021. *Open government geospatial data on buildings for planning sustainable and resilient cities*. doi.org/10.48550/ARXIV.2107.04023

Corbane, C., Syrris, V., Sabo, F., Politis, P., Melchiorri, M., Pesaresi, M., Soille, P., and Kemper, T., 2020. Convolutional neural networks for global human settlements mapping from

Sentinel-2 satellite imagery. *Neural Computing and Applications*. doi.org/10.1007/s00521-020-05449-7

Commission Implementing Regulation (EU) 2023/138 of 21 December 2022 laying down a list of specific high-value datasets and the arrangements for their publication and re-use., no. 2023/138, European Commission, 2022. eur-lex.europa.eu/eli/reg_impl/2023/138

European Commission. Joint Research Centre, 2016. *Buildings related datasets accessible through the INSPIRE geoportals: European Union Location Framework (EULF) project energy pilot*. Publications Office. data.europa.eu/doi/10.2791/46341

European Commission. Joint Research Centre, 2022. *GHS data package 2022: Public release GHS P2022*. Publications Office. data.europa.eu/doi/10.2760/19817

Directive 2010/31/EU of the European Parliament and of the Council on the energy performance of buildings (recast), Official Journal of the European Union, 18 06 2010, 2010. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2010:153:013:0035:en:PDF

Milojevic-Dupont, N., Wagner, F., Nachtigall, F., Hu, J., Brüser, G. B., Zumwald, M., Biljecki, F., Heeren, N., Kaack, L. H., Pichler, P.-P., and Creutzig, F., 2023. EUBUCCO v0.1: European building stock characteristics in a common and open database for 200+ million individual buildings. *Scientific Data*, 10(1), 147. doi.org/10.1038/s41597-023-02040-2

Pesaresi, M., and Politis, P., 2022. *GHS-BUILT-S R2022A - GHS built-up surface grid, derived from Sentinel-2 composite and Landsat, multitemporal (1975-2030)*. European Commission Joint Research Centre. http://data.europa.eu/89h/d07d81b4-7680-4d28-b896-583745c27085

Szabo, F., Corbane, C., Politis, P., Goch, K., and Kemper, T., 2023. *The European Settlement Map 2018* (Publications Office of the European Union) [Map].

Uhl, J. H., and Leyk, S., 2022. A scale-sensitive framework for the spatially explicit accuracy assessment of binary built-up surface layers. *Remote Sensing of Environment*, 279, 113117. doi.org/10.1016/j.rse.2022.113117