EGU General Assembly 2024

EGU24-14888, updated on 06 Dec 2024
https://doi.org/10.5194/egusphere-egu24-14888
EGU General Assembly 2024

From Shelters to Skyscrapers: A Worldwide Exploration of Buildings and Building Types Using Volunteered Geographic Information and Earth Observation Datasets

Laurens J.N. Oostwegel, Tara Evaz Zadeh, and Danijel Schorlemmer
GFZ German Research Centre for Geosciences, Potsdam, Germany (laurens@gfz-potsdam.de)

The location and type of buildings are incredibly useful data in any phase of the disaster management cycle. During the prevention and preparedness phases, the exposure and vulnerability of the population to natural hazards can be identified, using building inventories. To get to know the extent of damage in the recovery phase, one needs to know the state of the buildings before the disaster. In the response phase, knowledge about the exact population distribution can prove crucial. This can be derived from the knowledge of building locations and how and when they are populated. The common spots for shelter, such as hospitals and schools, should be identified immediately after a disaster has struck.
Humanitarian mapping has been a key support for disaster relief. Usually, the mapping is channeled through OpenStreetMap (OSM). By MapSwipe and automated completeness assessments it can quickly become clear where data is lacking. Volunteers often map areas through Mapathons organized by the Humanitarian OpenStreetMap Team (HOT) or Missing Maps. It is partly per these endeavors that OSM contains almost 600 million buildings as of 1 January 2024.
Recently, datasets using AI methods, largely based on Earth Observation (EO) data have been created to identify the world's buildings. The Google Open Buildings dataset, the Microsoft Global ML Building Footprints contain semi-automatically generated building footprints. Both are of near-global extent and they contain respectively 1.8 and 1.3 billion buildings, but neither is fully complete.
Unfortunately, unlike OSM, these datasets lack building attributes. The Microsoft dataset does include height in some limited areas, such as the USA and parts of Europe. There are datasets that contain more information, but they span a much smaller area. For example the USA Structures dataset defines occupancy types of buildings based on land use. However, in an ideal situation the rich information structure found in OSM is combined with the extensiveness of building footprints from the EO-derived datasets.
We investigated the wide range of features and attributes available from OSM, such as land use, amenities and points of interest and used these to classify all building footprints found in OSM itself and the EO-derived Google and Microsoft datasets. This resulted in three datasets of together 3.7 billion buildings. However, many of these buildings are overlapping. Therefore, a grid has been established on a resolution of roughly 100x100 meter. Each tile in the grid contains

buildings exclusively from one of the three datasets, with the priorities from high to low: OSM, Google, Microsoft.

If we use the quantity of the EO-derived datasets with the elaborate OSM tagging scheme, we can make better data-informed decisions during all phases of the disaster management cycle: (1) A detailed high-resolution global building inventory leads to better risk forecasting models. (2) Knowing both location and type of buildings results in a broad understanding of the common shelter spots and better estimates of the population distribution at the time of the disaster. (3) post-disaster situations can be better analyzed in scenario-based damage assessments.