SECOND EUROPEAN CONFERENCE ON EARTHQUAKE ENGINEERING
AND SEISMOLOGY, ISTANBUL AUG. 25-29, 2014

URBAN CHARACTERIZATION ASPECTS AND ASSOCIATED SPATIAL SCALE CONSIDERATIONS IN THE CONTEXT OF GLOBAL-LEVEL EXPOSURE MODELING: COMPARATIVE DATA EVALUATION

Christoph AUBRECHT[1], Rashmin GUNASEKERA[1], Oscar A. ISHIZAWA[1]

Setting the context for global-scale raster-level exposure modeling

In disaster risk assessment the spatially-integrative component exposure is used to identify the elements at risk located in a hazard-prone area that are subject to potential losses or that may suffer damage due to the hazard impact. Exposed elements in that regard can include population, property and infrastructure assets whose location-indicating information can be further combined with the elements' hazard-specific vulnerabilities (social, structural, etc.) in order to estimate the risks in the area of interest. Since effects of a natural hazard can be highly dispersed or very localized, a low level of geographic specificity or spatial resolution in exposure models can lead to errors in risk, damage, and loss assessments. Ongoing efforts aim to make grid-based estimates of exposure at the global scale state-of-the-art in the applied research community. Raster grids provide consistent homogeneous representations as opposed to the spatial heterogeneity inherent in administrative unit-level data (Aubrecht et al., 2013). Most global models such as GED4GEM (Global Exposure Database for the Global Earthquake Model) (Dell'Acqua et al., 2013) and the Global Risk Model prepared for the Global Assessment Report (UN-ISDR, 2013) use population as proxy measure for spatial distribution of physical assets and focus on population and building stock as modeled exposure categories.

The input data that is used for identifying exposed elements at risk (e.g. population and building counts and typology) is available on national or sub-national administrative unit level based on census collection amongst other sources. In order to allow spatial refinement and disaggregation of that information to a regular grid, certain environmental characteristics need to be defined and identified in a spatially explicit manner. Most importantly this refers to the characterization of urban and rural inhabited areas as well as to the type of occupancy (e.g., residential, non-residential, or mixed occupancy). With these characterizations complete in the exposure model setup, building construction types and other features can be proportionally assigned and aggregated asset values can be estimated at the grid level.

The urban definition issue

Distribution patterns and densities of population and associated assets as well as their typologies commonly vary significantly between urban and rural areas. Precisely defining those areas in spatial terms is, however, not a trivial task as there is no single definition of what makes an area 'urban'. In the UN's most recent World Urbanization Prospects (UN-DESA, 2012) definitions used for statistical purposes by a total of 231 countries are analyzed. Administrative criteria are most prominently used for making the distinction between urban and rural, in ˜30% even as sole criteria. Further criteria refer to a variety of population size or density thresholds, social, economic and functional characteristics (e.g., water supply and sewerage systems, electric lighting access), and combinations thereof.

[1] The World Bank, Urban & Disaster Risk Management Unit for Latin America and the Caribbean (LCSDU), Washington DC, USA, caubrecht@worldbank.org

Refraining from mere theoretical administrative definition approaches, the OECD in a joint effort with the European Commission initiated an attempt to define a spatially-explicit harmonized approach for urban identification, however limited to European and a set of other developed countries such as the US, Canada, and Japan (OECD, 2012). On a 1-km grid level contiguous high-population-density cells (1,500 inh./km²) are thereby clustered and a minimum absolute population threshold (50,000 inh.) is applied to identify urban centers. In a final step a city is defined on administrative level if at least half of a municipality's population is located inside the urban center. Using a uniform definition has its benefits, especially in terms of efficiency, cost and comparability across regions. However, as basic socio-economic and urban composition of developing regions vary significantly compared to developed countries, applying fixed thresholds does not seem appropriate at the global scale. Furthermore, when ancillary data need to be integrated that were collected and compiled based on country-specific definitions (e.g., housing census), accounting for the respective national statistical delineation approaches is crucial to ensure retraceable linking, albeit their potential oversimplification.

It becomes clear that most definitions and modeling perspectives in a spatial sense build upon the identification of densely-inhabited areas, mainly built-up areas suitable to 'host' human activities. Therefore, identifying built-up area is considered the first step to define settlements, which can then within its physical boundaries be classified as urban and rural, even if the statistical definition refers to administrative units. For example, the Caribbean island state of Antigua and Barbuda has just one single urban area defined by its national statistical office, namely the capital city St. John's. Refining the spatial level of delineation by accounting for built-up patches within the administrative boundaries significantly enhances the real-world representativeness which is crucial for exposure modeling. Further refinement options include accounting for additional criteria in the urban and rural distinction such as a population count threshold (e.g., 2,500 inhabitants in Mexico) and/or functional characteristics (e.g., electricity, water-supply etc. in Panama).

Comparative evaluation of global-level spatial data for urban area characterization

Addressing the urban area characterization issue on global scale in spatial terms, there is a limited list of processed spatial datasets available that provide a consistent basis to derive relevant information. On the one hand these include overall global land cover datasets such as GLC-2000 (Global Land Cover 2000) and GlobCover. On the other hand various datasets have been developed that already put a dedicated focus on certain characteristics relevant for urban identification, specifically addressing the delineation of built-up area. Potere and Schneider (2010) found huge area variations (at an order of magnitude) between the differently derived urban extents. MODIS-500, derived from 500-m spatial resolution optical satellite imagery, was identified as most detailed and since then has been considered sort of the global standard. Besides MODIS-500, the 1-km resolution Global ISA (Impervious Surface Area) data, derived from DMSP-OLS nighttime satellite imagery, has frequently been adopted for urban extent identification. Nighttime lights data give an indication on the distribution of human activities which facilitates its implementation in urban characterization as well as global population distribution models such as GRUMP (Global Rural-Urban Mapping Project) and LandScan (both 30 arc sec. ~1-km grids). GRUMP was first to have an urban mask applied for weighted reallocation in the population disaggregation process. LandScan went even further in implementing complex disaggregation weights and modeling ambient population as opposed to a pure residential population representation. Recently another population dataset has become increasingly prominent at the global level. The WorldPop initiative provides an order of magnitude higher spatial resolution (~100-m) for residential patterns. However, currently it is not yet available for all countries.

There are several recent developments and initiatives focusing on improved and spatially refined built-up area identification (further details are provided by Gunasekera et al., 2014). The European Commission's Joint Research Centre (JRC), within the Global Human Settlement Layer (GHSL) initiative, has produced a so-called Built-Up Reference Layer (BUREF). This dataset is modeled by using MODIS-500-derived built-up areas as 'sampling' regions in order to extract further inherent built-up areas from LandScan data. Another initiative is the German Aerospace Center's Global Urban Footprint (GUF) project that relies exclusively on Synthetic Aperture Radar (SAR) satellite data and provides built-up information at a resolution of 50-70 m, an order of magnitude higher than the before-mentioned datasets. Current efforts also enter the very high resolution domain in terms of the level of detail analyzed, previously unavailable at the global scale. The U.S. Oak Ridge

National Laboratory (ORNL) is developing ‘LandScan HD’ based on a settlement layer derived from meter-level satellite imagery. Similarly, JRC is working on a high resolution population and built-up classification scheme in the course of their GHSL development efforts.

Considering the heterogeneity in nature, scale and sophisticaton of the described built-up area datasets, comparative spatial evaluation was carried out for sample areas in Central America and the Caribbean region. Our objective is the integration of urban and rural built-up areas in the makeup of a 1-km exposure grid for the implementation of asset value disaggregation. Therefore, the focus lies on determining the best option for built-up and urban delineation at that specific scale level which implies certain resampling and thresholding considerations of higher resolution products. Potential strengths and weaknesses of the different datasets were analyzed, including inherent characteristics of optical versus radar data derived products. The radar-based GUF has the advantage that built-up areas can be limited to actual building clusters excluding road networks by accounting for the inherent height information. On the other hand, terrain and slope aspects can cause misidentification problems with radar data particularly in coastal areas. Analyzing the higher-resolution optical imagery-based ORNL settlement layer, more granular built-up features are detected at the very local level. However, at the same time certain misclassification issues appear related to features with similar reflectance characteristics (e.g., sand, bare soil) as well as cloud-covered regions. Furthermore, roads are commonly included in optical-derived built-up datasets due to the difficult distinction lacking height information. At the 1-km grid level, building on the 500-m resolution MODIS data enables the BUREF product to show continuous built-up ratios. The significance of those cell ratios was analyzed for several sample regions with regard to thresholding for binary built-up masking. Furthermore, a similar resampling approach was tested to derive GUF-based built-up ratios and subsequent binary masks at the same scale.

Given the availability of data at various resolutions and with different setup characteristics as outlined above, exposure model input data must be selected cautiously. For example, operating at a final output resolution of 1-km grid cells, built-up input data at the 1-m level would be over-detailed and efforts required for resampling feasibility outplay the actual benefits of increased input data accuracy. Furthermore, returning to the urban characterization issue, integration of population data is essential. Here, the use of residential as well as ambient representations can be beneficial for identifying relevant socio-economic activity patterns. Proper selection and integration of the available datasets is crucial for ongoing exposure modeling efforts on global scale.

The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent. The following not-publicly-available datasets were provided for use in the above-cited project: BUREF (by the European Commission’s Joint Research Centre), GUF (by the German Remote Sensing Data Center), ORNL level-1 settlement layer (by the U.S. Oak Ridge National Laboratory).

REFERENCES

Aubrecht C, Özceylan D, Steinnocher K, Freire S (2013) “Multi-level geospatial modeling of human exposure patterns and vulnerability indicators,” *Natural Hazards*, 68(1):147-163

Dell’Acqua F, Gamba P, Jaiswal K (2013) “Spatial aspects of building and population exposure data and their implications for global earthquake exposure modeling,” *Natural Hazards*, 68(3):1291-1309

Gunasekera R, Aubrecht C, Ishizawa O, Freire S (2014) Global Exposure Modeling: an End User Perspective OR ‘5 things you should know about exposure modeling!’, Contributing paper to the Global Assessment Report on Disaster Risk Reduction 2015, United Nations Office for Disaster Risk Reduction, Geneva, in press

OECD (2012) Redefining “Urban”, Organisation for Economic Co-operation and Development, Paris

Potere D, Schneider A (2010) “Comparison of Global Urban Maps,” In Gamba P, Herold M, eds, Global Mapping of Human Settlement, Remote Sensing Applications Series, CRC Press, 269-308

UN-DESA (2012) World Urbanization Prospects, the 2011 Revision, United Nations Department of Economic and Social Affairs - Population Division, New York

UN-ISDR (2013) Global Assessment Report on Disaster Risk Reduction 2013, United Nations Office for Disaster Risk Reduction, Geneva, 246 pp