Input Paper prepared for the Global Assessment Report on Disaster Risk Reduction 2015

A GLOBAL EXPOSURE MODEL FOR GAR 2015

ANDREA DE BONO AND BRUNO CHATENOUX

UNEP/GRID-GENEVA

JULY 2014

Table of Contents

1. Introduction

This report describes how a global exposure database was generated for the quantification of the exposure, to support the earthquakes and cyclones probabilistic risk modeling in the Global Assessment Report 2015 (GAR2015) on Disaster Risk Reduction process. The Global Exposure database for GAR 2015 (GEG-2015) is an open exposure global dataset which integrates population and country-specific building typology, use and value.

The main objective of GEG-2015 was to generate a global evaluation of exposed assets in urban and rural areas in order to provide a specific exposure input data to be used in the CAPRA platform

An exposure database that includes a global inventory of critical facilities based on a pure bottom up approach would require considerable human and economic efforts, and is beyond the scope of this project. As the bottom-up approach is not available a spatial disaggregation was employed. It consists on a top-down or "downscaling" approach, where information including socio-economic, building type and capital stock at a national sub-national scale (statistical data) are transposed onto a regular grid, using Gis data such as geographic population and Gross Domestic Product (GDP) distribution models as proxies (Figure 1).

This approach offers the considerable advantage of ensuring good uniformity and ability to compare results across the different countries, together with the relative availability of the base datasets necessary in order to build the model. On the other hand, however, the results remain highly correlated with population, because strictly dependent on it.

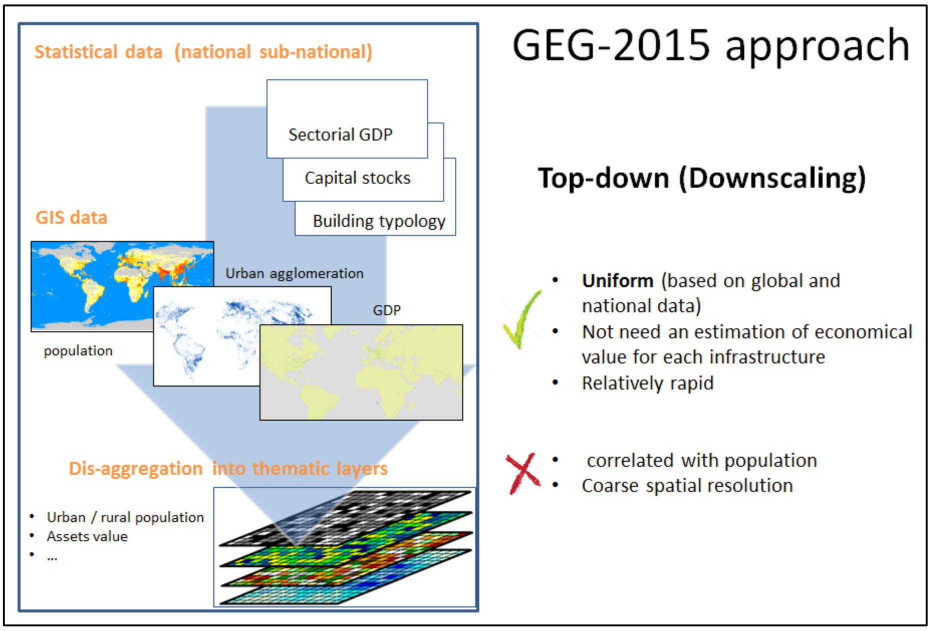

Figure 1: GEG 2015: a top-down approach

Exposure are the elements at risk to potential losses (UNISDR, 2009) or that may suffer damage due to a hazard impact. The exposed elements include people, resources, infrastructures, production, goods, services or ecosystems and coupled social-ecological systems. In the Global Exposure database for GEG-2015 the physical exposures are represented through the inventory of buildings in urban and rural areas, called here "the building stock". Structural types in each country were provided by the World Agency of Planetary Monitoring & Earthquake Risk Reduction (WAPMERR), conforming to the PAGER construction type classification scheme (Jaiswal et al., 2010)..

The urban areas definition in the GEG-2015 context is based on physical attributes; that are areas that are dominated by the "built environment". The built environment includes all non-vegetative, human-constructed elements, such as buildings, roads, runways, etc. (Schneider et al., 2009). Therefore, in this context when we refer to urban areas or urban agglomerations we implicitly refer to the above definition of built-up environment.

The number of persons subdivided by socio-economic class living in a specific construction type in a determined geographical location is used as a base for distributing the exposed economic value of the building stock. In other words, population data will be used to obtain information on the constructed area. For example, a certain percentage of people, employed in the industrial sector per a determinate country, will be related to the proportion of industrial buildings in that country (De Bono and Mora, 2014).

In order to establish a monetary evaluation of the exposed assets, like it was done in GAR 2013, the concept of stock is employed. In the context of natural disasters, stocks represent the usual choice of unit for measuring exposure. This is especially true considering that a natural disaster could cause asset damage greater than the annual flow.

The GEG-2015 uses World Bank data (World Bank, 2011) on produced capital stock per country and focuses directly only on building stocks without taking into account other typologies of infrastructures such as roads or bridges.

2. Data and methodology

2.1 Main components

This section describes the main datasets (base data) and underlying databases that have been used to produce the GEG-2015.

2.1.1 Reference grid(s)

During the disaggregation process, socio-economic, building type and capital stock information will be transposed onto uniform geographical units that are the reference grid at 2'30" (or approx. 5x5 km at equator). The choice of the resolution is justified by two significant reasons:

- a 5x5 resolution is considered as satisfactory in order to capture effects from large-scale hazards such as earthquakes and cyclones
- to guarantee consistency in the results: socio-economic indicators are provided at a national scale, for certain large and non-uniform countries a disaggregation to smaller cell sizes would stretch the downscaling process too far.

Moreover this resolution is appropriate in order to optimize the time of analyses needed for obtaining the results from the probabilistic risk calculation.

It is important to emphasize that the size of the cell is only one side of the "resolution" of a dataset, which also includes the thematic aspect, or in other words the amount and the quality of the information captured within it.

The so-called 5x5 reference grid for GEG 2015 includes the whole earth land surface, comprising uninhabited land areas. In this way the reference grid will be able to handle eventually future data on crops pastures and forest areas. The total number of cells of the grid is 9'008'829, while inhabited cells correspond to 4'574'010.

GEG 2015 also includes a second grid at 30" resolution (around 1x1 km at equator) was set in order to hold exposure data related to coastal areas. The grid was only built for a sector including the first 10 km of coast worldwide. A finer resolution is required in order to capture the magnitude of effects from small-scale hazards such as storm surges and tsunamis. Data and methodologies used to generate the 1x1 grid correspond to those used for the 5x5 reference grid.

2.1.2 Population distribution

The primary source of global exposure information is a model that represents the distribution of people on the earth surface. A gridded population dataset is the result of this model, where each cell indicates the number of people living on it.

Presently only three (complete) global gridded population data sets are available: LandScan developed by The Oak Ridge National Laboratories (ORNL), (Bright, E.A, 2002, Bhaduri et al., 2002), GPW and GRUMP from SEDAC (Socioeconomic Data and Applications Center), (2012). LandScan refers to the concept of "ambient population", a time-weighted average (over 24 hours) of the number of people in each cell, while other grids usually locate each person in his/her dwelling.

The GEG-2015 integrates the new LandScan layer published in June 2012 by the ORNL, which refers to the population as of July 2011 with a resolution of 30" (approx. 1 km at equator).

LandScan is a reasonable option for GEG needs, indeed, we already used it since the 2007 edition of GAR, but we have to keep in mind that, it has two main, not negligible limitations:

- The in-deep methodology and complete algorithms are not published.
- Its license does not allow it to be distributed, such as all other GAR datasets, to the public domain.

Following the same methodology used in the past edition of GAR, the population figures in LandScan were adjusted to match with UN official data using World Population Prospects[1].

2.1.3 Modeled urban settlements (BUREF)

For GAR 2013 the building stock was delineated by using the built-up areas class extracted from the MODIS 500-m map of global urban extent [2]. Even if this dataset was for far the best choice for built-up areas estimation, it displays some inconsistencies for high density countries in Southeast Asia (De Bono, 2013).

[1] http://esa.un.org/wpp/

[2] http://sage.wisc.edu/people/schneider/research/data_readme.html

For the GEG 2015 we employed a new dataset developed by Joint Research Center or JRC (Pesaresi et al., 2012) that is called BUREF (Built-Up REFerence).

BUREF layer includes the best estimation of presence of built-up areas at 30" cell resolution at global scale. Preliminary tests confirm BUREF more accurate than MODIS 500 especially in small-scattered settlement areas.

Differently than MODIS 500 that is a dichotomy, BUREF describes the built-up areas as a continuous value, as an estimation of the percentage of built-up area with respect to the total surface of the cell. Values expressed in the range [0..1] (**Error! Reference source not found.**). After going through many checks, only values from 0.1 to 1 were considered in the urban mask for GEG-2015.

The BUREF layer was produced by merging LandScan and MODIS 500 as input data. This combination is obtained by reclassification of LandScan using MODIS as training set, with a spatial learning kernel of 500 kilometers (Pesaresi et al., 2012).

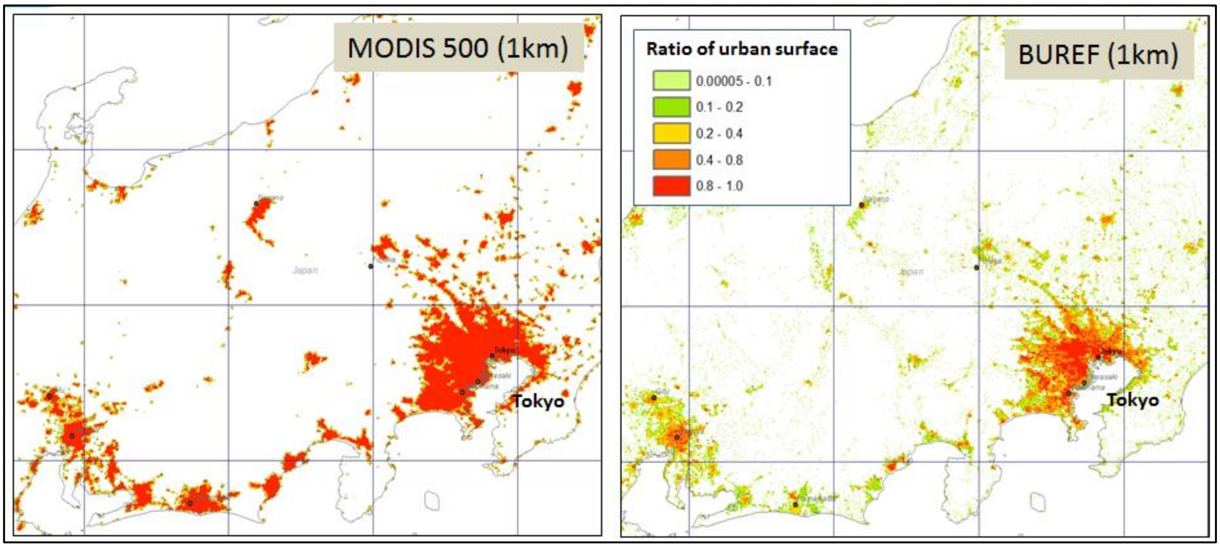

Figure 2: Comparison between MODIS 500 (resampled to 1km) and BUREF, region of Tokyo (Japan) Legend refers to BUREF map only; it indicates the amount of built up per cell. A value of 1 means a100% constructed surface.

2.1.4 Socio-economic indicators (SEI): Building stock characterization

Socio-economic indicators (SEI) will be used as proxies to estimate the sectorial use of the building stock in accordance with the city size of each settlement area. They consist of tabular data at country level Table 1: The list of socio economic indicators (SEI) and their sourcesTable 1).

A first group of indicators includes resident population subdivided into four classes, based on the income per capita, according with the World Bank (WB) classification.

A second group of indicators including non-residents is subdivided by economic activities (industry, services and government), health coverage (number of hospital beds in public/private) and education for private and public sectors (number of pupils).

Sector		Socio-economic indicator	Data sources
Population		country population in 2011	UN WPP 2013[3]
Capital stock		Produced capital (machinery and urban)	WB 2011
Residents	Income	1) % of people living below 1005 $	WDI 2013[4]
		2) % of people living between 1005 $ and 3975 $	
		3) % of people living between 3975 $ and 12275 $	
		4) % of people living above 12276 $	
Non residents	Employment	5) % of people employed in industry	WDI 2012, ILO 2012[5], UNSD 2012[6]
		6) % of people employed in services	
		7) % of people employed in government	
	Health	8) number of beds per 1000 people private	WDI 2012
		9) number of beds per 1000 people public	
	Education	10) number of pupils (private)/total population	UNESCO 2012[7]
		11) number of pupils (public)/total population	

Table 1: The list of socio economic indicators (SEI) and their sources

SEI are commonly released as relatively continuous time series including at least the last 20 years. 2010 was for far the most complete year for a majority of indicators and it was chosen as base year.

2.1.5 Building structure typology

The World Agency of Planetary Monitoring & Earthquake Risk Reduction (WAPMERR) provided all the information concerning the construction types in each country, for three size categories

[3] UN WPP 2013 :World Population Prospects, the 2012 Revision. http://esa.un.org/wpp/Excel-Data/population.htm
[4] WDI:World Bank. World Development Indicators. http://databank.worldbank.org/data/databases.aspx
[5] ILO: International Labor Organization, LABORSTA. http://laborsta.ilo.org/
[6] UNSD : UN Statistical Division. http://unstats.un.org/unsd/environment/qindicators.htm
[7] UNESCO : Institut for Statistics http://stats.uis.unesco.org

of settlement size or "complexity", as percentage of peoples living in the determinate building structure (Wyss et al., 2013), it is important to underpin that, the distribution of the structural types was carried on the basis of the proportion of the population using each of them. The results are then expressed as percentage of people per structural type and not as a percentage of number of typologies per total number of buildings/dwellings

New data for GEG-2015 were gathered including the sub-national distribution of building types for 18 countries. Countries selected include the largest heterogeneous ones (i.e. China, India and Indonesia) and represent approximately 3.6 billion people, i.e. about 50% of the world's total population.

Housing census data and occupancy rates data were collected by WAPMERR from Statistics Office of each investigated country. When collected data were not satisfactory, supplementary information were searched from mainly research publications, reports and photos (Tolis et al., 2013). The quality and detail of information on the built environment varies greatly between countries. Most of collected data refers to the construction materials and in a few cases the age of the building, the number of floors or the characteristics of living quarters.

Successively, construction materials together with other collected parameters were interpreted and converted to the PAGER building typologies

The models for the building stock depend on the size of the settlements because villages and cities contain different building typologies. Therefore the building stock was divided into large urban, small urban, and rural areas; this subdivision is based on the thresholds defining the limits between rural and urban settlements proposed by Satterthwaite and Owen, (2006).

For GEG 2015 WAPMEER also provides an estimation of building typologies for residential use (dwellings) and non-residential use. Unfortunately published census data on the building stock are mostly constricted to residential housing excluding commercial or institutional occupancy, thus in several cases, assumptions and decisions were necessary to estimate the typologies of the non-residential building stock. These assumptions (together with the general methodology and data sources) are explained in detail in Tolis et al., (2013). This report is also an Input paper prepared for the Global Assessment Report on Disaster Risk Reduction 2015.

2.1.6 Produced capital stock: building stock evaluation

The produced capital stock (World Bank, 2006) tends to be the most readily understood form of capital due to its tangibility and the quality of data collected on investment levels. The World Bank (World Bank, 2011) published a dataset for 152 countries that provides broad estimates of the produced capital in urban areas.

The produced capital estimation is based on the Perpetual Inventory Method (PIM) and historical Gross Capital Formation (GCF) data (World Bank, 2006). Furthermore, the World Bank scaled-up this estimate by 24% to account for the economic value of urban land.

According to PwC (PwC (PricewaterhouseCoopers), 2013), currently the World Bank methodology appears to be the most consistent method of measuring produced and natural capital values at an international level. Other valid alternatives can be represented by the

United Nations University, (UNU 2012) and the Organization for Economic Co-operation and Development (OECD, 2011) but they have estimations only for a limited number of countries.

2.1.7 Gross Domestic Product (GDP) distribution at sub-national (regional) level

In order to capture the regional variations of economic activity within the country (and thus indirectly the capital stock), we geographically weight the capital stock by using a raster of the distribution of Gross Domestic Product (GDP) as proxy.

In the GAR 2009 and 2013 process we used a dataset originated by the World Bank Development Economics Research Group (DECRG) team for GAR, publicly available on the Preview data portal[8].

For the GEG 2015 the development of a new indicator has improved the distribution of the economic activity on the territory. This indicator fulfils two major requirements:

- Moderate the influence of population on urban capital stock estimation
- Distribute the wealth according to the sub-national variations of the GDP inside the country

The new indicator is the result of the integration of tabular information of GDP at regional level or Gross Regional Product (GPR), with the intensities of nighttime light VIIRS (Visible Infrared Imaging Radiometer Suite) recently published by NOAA[9]. The intensity of nighttime lights could represent a good proxy of human activities and this was already used at a global scale to map economic activity by Gosh et al., (2010). Nighttime lights are a class of urban remote sensing products derived from satellite sensors. VIIRS offers substantial improvements in spatial resolution when compared to the previous DMSP low-light imaging data (Elvidge et al., 2013).

GPR have been generated for 71 major countries using the following sources:

- Eurostat: 25 countries
- Beijing Normal University: 1 country (China)
- OECD: 1 country
- World Bank DECRG: 44 countries

For countries where the regional data was not available, national one were employed using data from World Development Indicators (World Bank 2013)

The process to find the best correlation between nighttime lights and economic activity is not straightforward (see the example in Figure 3) and needs some preprocessing such as light normalization and several statistical analyses including filtering, thresholds set-up, income regionalization which will be described below.

[8] http://preview.grid.unep.ch/index.php?preview=data&events=socec&evcat=1&lang=eng

[9] http://ngdc.noaa.gov/eog/viirs/download_viirs_fire.html

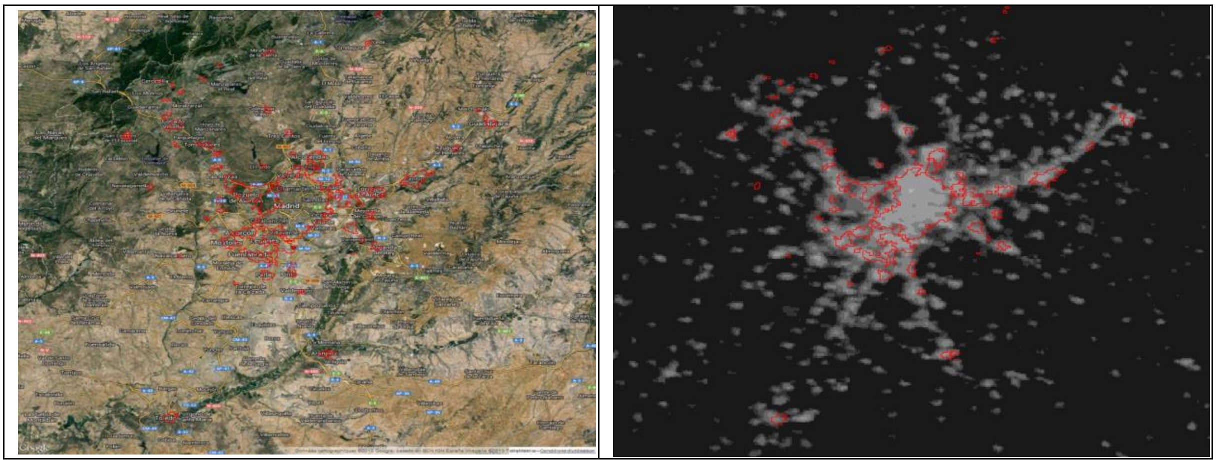

Figure 3: Comparison of Madrid (Spain) urban area using MODIS 500 (left) and VIIRS nighttime lights 2012 (right)

First each pixel of the original VIIRS nighttime light data was divided by its real area to compensate projections distortion in order to get a light density layer.

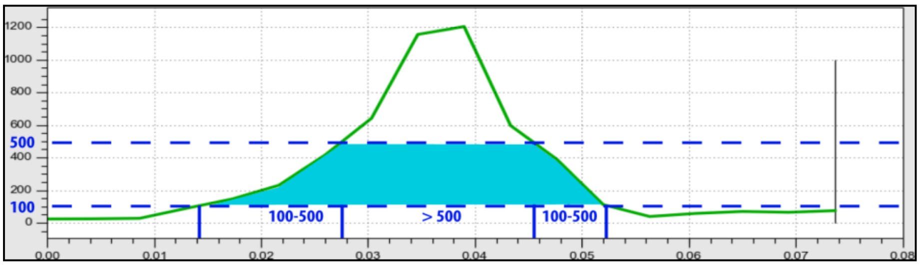

Figure 4: Example of light density slicing

Then light density was sliced (Figure 4) and the best combination of slices was iteratively tested (Figure 5) applying successive multiple linear regression.

grp	from\to	0_3	3_5	5_8	8_10	10_20	20_30	30_40	40_50	50_60	60_70	70_80	80_90	90_100	100_200	200_300	300_400	400_500	500_1000	1000_5000	5000_10000	10000
grp_ind10	Zsnug_200_300															0.016	0.011	0.009	0.006	0.005	0.005	
grp_ind10	Zsnug_300_400																0.006	0.005	0.004	0.004	0.004	
grp_ind10	Zsnug_400_500																	0.005	0.004	0.003	0.003	
grp_ind10	Zsnug_500_1000																		0.003	0.003	0.003	
grp_ind10	Zsnug_1000_5000																			0.003	0.003	
grp_ind10	Zsnug_5000_10000																				0.003	
grp_ind10	Zsnug_10000_50000																					
rp_indserv10	Zsnug_0_3	0.026	0.032	0.043	0.051	0.089	0.124	0.150	0.168	0.179	0.186	0.190	0.193	0.194	0.197	0.195	0.193	0.191	0.186	0.176	0.173	
rp_indserv10	Zsnug_3_5		0.102	0.224	0.295	0.465	0.530	0.555	0.561	0.558	0.551	0.543	0.534	0.524	0.469	0.451	0.437	0.426	0.392	0.333	0.319	
rp_indserv10	Zsnug_5_8			0.435	0.482	0.560	0.591	0.601	0.598	0.589	0.577	0.566	0.554	0.543	0.478	0.457	0.442	0.430	0.391	0.328	0.313	
rp_indserv10	Zsnug_8_10				0.532	0.585	0.608	0.612	0.606	0.594	0.581	0.567	0.554	0.542	0.470	0.448	0.432	0.418	0.377	0.311	0.296	
rp_indserv10	Zsnug_10_20					0.595	0.614	0.617	0.608	0.595	0.580	0.565	0.552	0.538	0.463	0.439	0.423	0.409	0.367	0.299	0.283	
rp_indserv10	Zsnug_20_30						0.624	0.616	0.598	0.576	0.554	0.534	0.516	0.498	0.407	0.381	0.362	0.346	0.300	0.231	0.216	
rp_indserv10	Zsnug_30_40							0.595	0.565	0.533	0.504	0.479	0.456	0.435	0.333	0.304	0.285	0.268	0.223	0.161	0.149	
rp_indserv10	Zsnug_40_50								0.513	0.473	0.438	0.409	0.383	0.360	0.256	0.229	0.211	0.196	0.156	0.106	0.097	
rp_indserv10	Zsnug_50_60									0.412	0.375	0.344	0.317	0.292	0.195	0.171	0.154	0.141	0.108	0.070	0.064	
rp_indserv10	Zsnug_60_70										0.323	0.292	0.264	0.239	0.151	0.130	0.115	0.104	0.076	0.048	0.043	
rp_indserv10	Zsnug_70_80											0.251	0.223	0.199	0.121	0.102	0.089	0.079	0.056	0.034	0.031	
rp_indserv10	Zsnug_80_90												0.189	0.166	0.099	0.082	0.070	0.062	0.042	0.025	0.023	
rp_indserv10	Zsnug_90_100													0.140	0.085	0.069	0.057	0.049	0.033	0.019	0.017	
rp_indserv10	Zsnug_100_200														0.074	0.058	0.048	0.040	0.026	0.014	0.013	
rp_indserv10	Zsnug_200_300															0.010	0.006	0.005	0.003	0.002	0.002	
rp_indserv10	Zsnug_300_400																0.003	0.002	0.002	0.002	0.002	
rp_indserv10	Zsnug_400_500																	0.002	0.002	0.001	0.002	
rp_indserv10	Zsnug_500_1000																		0.001	0.001	0.001	
rp_indserv10	Zsnug_1000_5000																			0.001	0.001	
rp_indserv10	Zsnug_5000_10000																				0.002	
rp_indserv10	Zsnug_10000_50000																					
grp_serv10	Zsnug_0_3	0.006	0.008	0.012	0.015	0.030	0.045	0.057	0.066	0.072	0.076	0.078	0.080	0.081	0.085	0.084	0.084	0.084	0.082	0.078	0.077	
grp_serv10	Zsnug_3_5		0.037	0.104	0.151	0.308	0.383	0.418	0.430	0.432	0.428	0.423	0.417	0.410	0.367	0.352	0.341	0.332	0.304	0.248	0.231	
grp_serv10	Zsnug_5_8			0.306	0.360	0.454	0.491	0.504	0.503	0.496	0.486	0.475	0.465	0.454	0.395	0.377	0.363	0.352	0.318	0.253	0.234	
grp_serv10	Zsnug_8_10				0.431	0.491	0.518	0.524	0.519	0.508	0.495	0.483	0.470	0.458	0.392	0.372	0.358	0.346	0.309	0.241	0.221	
grp_serv10	Zsnug_10_20					0.504	0.527	0.531	0.523	0.511	0.496	0.483	0.469	0.456	0.387	0.366	0.351	0.338	0.300	0.231	0.211	
grp_serv10	Zsnug_20_30						0.543	0.536	0.518	0.498	0.477	0.458	0.440	0.424	0.339	0.316	0.299	0.285	0.244	0.175	0.157	
grp_serv10	Zsnug_30_40							0.518	0.489	0.460	0.432	0.408	0.387	0.367	0.273	0.248	0.231	0.217	0.178	0.118	0.103	

Figure 5: Iteration to find the best correlation sum of light per region vs. GPR

Different thresholds were tested in order to find the best regression parameters between light per region and GPR (dark green cells in Figure 5).

As a single global regression (meaning one regression for the whole world) was not statistically satisfying, the full dataset was split per OECD income region and a regional regression performed (Table 2).

Income group	Range	R2	lmin	lmax	Inters	xfactor
High OECD	75_500	0.548	75	500	21831606	71.306
High non-OECD	75_500	0.641	75	500	3583148	21.708
Upper middle	50_500	0.656	50	500	832353	47.695
Low s.l.	10_100	0.505	10	100	782460	32.937

Table 2: Results of multiple linear regressions per economical region

The Table 2 above presents the best regression parameters:

- Range: minimum and maximum thresholds applied
- R^2 : R-squared of the linear regression (a value of 1 means a perfect fit, a value of 0 means not correlation at all)
- lmin: minimum threshold applied
- lmax: maximum threshold applied
- Inters: intersection of the linear regression
- Xfactor: slope of the linear regression

Finally a correction factor was applied to each region in order to perfectly fit the simulated GDPR values with the originals.

The last step will consist in transposing the results to the 5x5 grid and then calculating the coefficient of variation from national to per capita GDP values.

We have to keep in mind that the final indicator will not be a map of GDP distribution, but rather a map showing the variation between the GDP at national scale and its regional variations: in Figure 6, each pixel in the map shows its variation from the national average of the corresponding country.

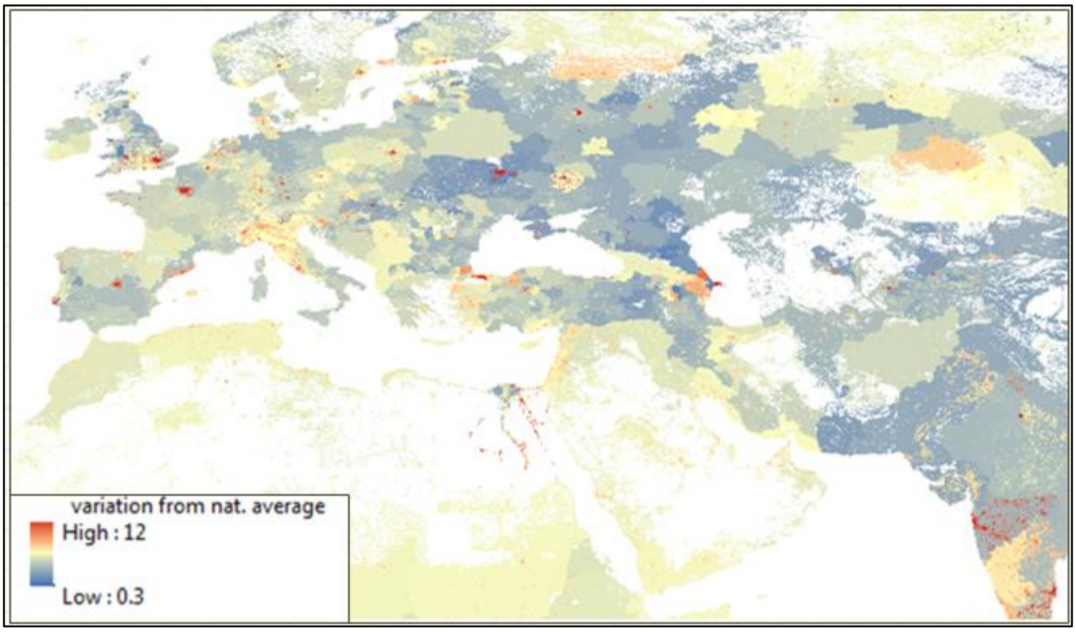

Figure 6: Coefficient of variation between the national values of GDP and those at subnational level, including regional GDP (or GPR) and GDP in high activity zones (derived from VIIRS)

2.2 Compilation and harmonization (missing data, assumptions)

Most socio-economic indicators, and the capital stock from the World Bank, do not cover the full list (216) of countries/territories used in the GAR assessment. In the case of the socio-economic indicators, in order to complete the missing values four different methodologies were employed (in order of preference):

1. Data were searched for through national statistical offices: this method was used particularly for France" DOM-TOM" by using the National Institute for Statistics (INSEE).
2. Data came from "global" unofficial databases such as the CIA Factbook[10] this was extensively used in various cases
3. Data were estimated using proxies (eg GDP versus capital stock).
4. Data were assumed to be equivalent to countries considered as "similar" in terms of geographic position, development, economy, sovereigns

The list of countries included in the World Bank capital stock exclude 66 economies. For 35 of these countries, the International Institute for Applied Systems Analysis (IIASA) provided a list of data extracted from its unpublished internal database. The missing data on produced capital for the rest of countries/territories were evaluated following the approach outlined by PwC (PricewaterhouseCoopers, 2013), as follows in (order of decreasing robustness):

a. Apply the World Bank algorithms[11] using World Bank Gross Capital Formation13 (GCF) data
b. Apply the World Bank algorithms using data on Gross Fixed Capital Formation (GFCF) from the IMF14 and EconoStats15
c. Apply the World Bank algorithms on GCF: GFC data are calculated by using a fixed ratio between GDP and GFC

The third methodology has been applied exclusively to smaller countries/territories where only GDP data were available. More detailed information and a list of estimated countries of this process was outlined in (De Bono, 2013).

2.3 Workflow: downscaling and data integration

Data on socio economic indicators (SEI) capital stock and building typologies are generally provided at the national scale by administrative units. In this section we will illustrate the methodology and assumptions in order to spatially disaggregate them onto the 5x5 reference grid.

Essentially our distribution model will follow a multi-layered, dasymetric, spatial approach where data by administrative units are converted to a regular finer surface (reference grid) by means of ancillary data constituted by the rasters of population and GDP variation.

---

[10] https://www.cia.gov/library/publications/the-world-factbook/

[11] "World Bank algorithms » are provided in World Bank (2006, 2011)

2.3.1 Step 1: Define urban/rural population and aggregate data onto the 5x5 km reference grid

a) Extract urban population using the BUREF mask

The urban mask based on the BUREF dataset delimits the cells representing the built-up areas; it was employed to extract the population from the gridded population dataset (LandScan). Output includes people living in urban areas (Figure 7)

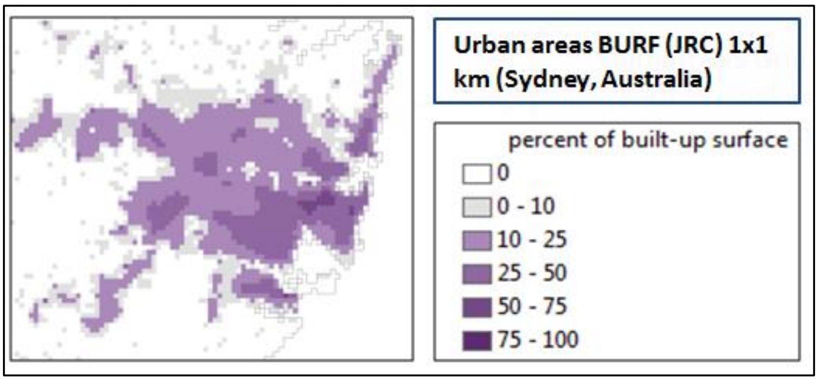

Figure 7: Urban area delimitation (BUREF)

b) Define the complex type

Once delimitated, the urban areas will be subdivided on the base of their complex type.

The complex type (cpx) characterizes the urban settlements based on their size according to the Satterthwaite classification (Satterthwaite 2006). It is important to remember that the classification is applied to a group of contiguous cells (at list a shared vertex) that represents an "agglomeration" and not to a single cell.

Therefore the different settlements have been subdivided into urban or rural classes as follows:

cpx1 = more than 20'000 people	(URBAN population)
cpx2 = between 20'000 and 2'000 people	(URBAN population)
cpx3 = less than 2'000 people	(RURAL population)

c) Define what is rural

Rural cells were defined as follows:

- All cells outside the urban mask (BUREF) independently of their population size
- Cells inside the urban mask but where the percentage of built-up area is < 10 % (BUREF)
- Cells inside the urban mask, belonging to an agglomeration having < 2'000 inhabitants (cpx3)

Figure 8 displays an example of subdivision into complex types for the city of Sydney.

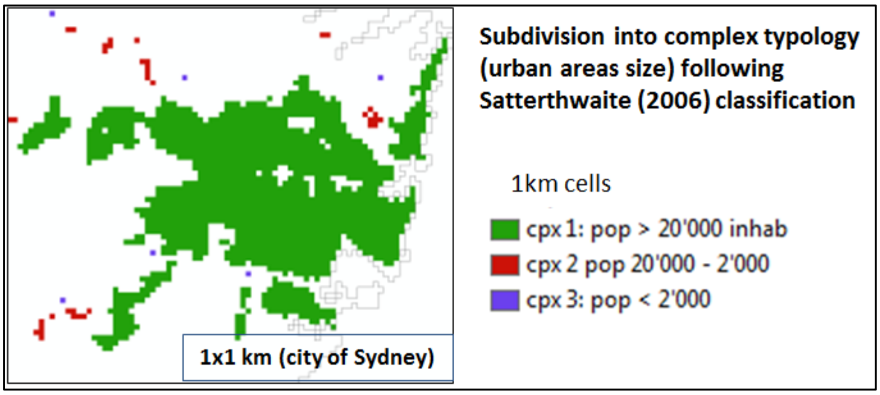

Figure 8: Subdivision into complex typology

In conclusion all cells that correspond to one of these three parameters have been automatically assigned to the complex type 3 (cpx3).

d) Aggregation onto the 5x5 km reference grid

Cells containing urban and rural population were successively aggregated onto the 5x5 km reference grid.

2.3.2 Step 2 Distribution of the population per socio-economic indicators (SEI)

The distribution of the population by level of income, employment, instruction and the number of beds in the health sector needs will be estimated with the objective to evaluate the building stocks related to residential and non-residential use.

According to the methodology already developed for GAR 2013, the following substeps are outlined:

1. Population distribution by income level following the four income classes established by the World Bank (World Bank's annual World Development Report (WDR), n.d.)
2. Evaluation of the labor force by occupation in industrial, governmental and commercial sectors
3. Estimation of the health service capacity in public and private sectors
4. Estimation of the capacity of the education services in public and private sectors.

Mostly, the algorithm to transpose a socio-economic parameter as a percentage of the total population to the grid cell using the population as proxy corresponds to (Equation 1):

$$SEI(x,y) = \frac{SEI(adm)}{\sum_{(adm)} Pop(x,y)} * Pop(x,y)$$

Equation 1: algorithm employed to downscaling a SEI to the grid cell

where:

- SEI(x,y) - population per socio-economic parameter per cell at x,y position;
- SEI(adm) - population per socio-economic parameter per administrative unit (general country level)
- Pop(x,y) - population living in the cell (extracted from LandScan).

After this operation, each cell represents the number of exposed persons for each of the eleven socio-economic sectors (income, labor, instruction, and health, in Table 1) in a portion of an urban or rural area.

The whole residential population corresponds to the sum of the four income classes. Non-residents include workers in industry and service sectors, students, represented by the sum of peoples included in the labor and education sectors.

2.3.3 Step 3 Distribution of the capital stock per socio-economic sectors

This downscaling process consists in transferring the produced capital from the administrative unit (country) to the 5x5 grid cells using the already estimated population per cell and sector as the main proxy. In order to refine the process and keep a more realist snapshot of the distribution of economic pattern of the country, the results were further weighted by two different sets of variables:

1. The regional variation of Gross Regional Product (GRP) as calculated in section 2.1.7 Gross Domestic Product (GDP) distribution at sub-national (regional) level.
2. The "unitary values" at national level and by socio-economic sector.

Essentially the process involving GRP consists in evaluating a coefficient of variation between the national values of GDP (capita) and those at subnational (regional) level as already showed in paragraph 2.1.7 Gross Domestic Product (GDP) distribution at sub-national (regional) level. For each 5x5 cell a coefficient of variation from national values is calculated. This coefficient is further multiplied by the urban capital stock of the cell. Capital stocks are successively normalized in order to match with national totals. In other words the coefficient indicates how much a cell will differ from the national average of GDP.

The "unitary values" give an indication about cost and surface of the built-up surface related to the building usage; they have been processed by CIMNE on the bases of data coming from the Global Construction Cost and Reference Yearbook (Compass International Inc., 2012).

The unitary values are successively used as weight factor and integrated in the process: basically they consist of an evaluation per country and socio-economic sector of the building surface and its unitary cost. It is important to underline, that surface and unitary costs, are relative values across the country; and they have to be interpreted as factors, used only to differentiate the surface and its cost from one socio-economic class to another. In other words, for a determinate country a cost value of 2 for low income buildings and 8 for high income buildings, only means that the latter have cost four times those of lower income.

The produced capital is downscaled to the cell level using population data and then multiplied by the already calculated coefficient GDPcv%. This will results in a database that moves from a pure population distribution to a goods type distribution (population per buildings use) according to the following equation (Equation 2):

$$PCSEI = [PCc(adm) * PopSEI(x,y)] * GDPcv\%(x,y) * UVs$$

Equation 2: Algorithm employed to calculate the produced capital per cell (PCSEI) per socio-economic indicator (SEI)

where:

- PCSEI (x,y) is the produced capital at cell size per SEI, PCc(adm) is the produced capital per capita at national level,
- PopSEI(x,y) the population of the cell per sector,
- GDPcv% the coefficient of variation for the cell located at x,y coordinates, and
- UVs the coefficient related to unitary costs/surfaces related to the socio economic sector.

2.3.4 Step 4 Distribution of the capital stock per socio-economic sector and building type

Once the population and capital stock are estimated for residential and non-residential usage, it is necessary to distribute them in the different building typologies present in the country, integrating the information provided by WAPMERR that contains the population distribution by level of complexity and by building type (or structural system). This last operation is simply performed by multiplying the already calculated sectorial capital stock (PCSEI) per capita by the corresponding population living in a certain structural system.

Finally, after this last step, each record will correspond to a building type, based on the level of income/sector, specific to a localization of 5km x 5km of an urban area with an associated level of complexity.

3. Results

The main objective of GEG-2015 was to generate a global evaluation of exposed assets in urban and rural areas in order to provide a specific exposure input data to be used in the CAPRA platform.

Two distinct dataset were produced (Figure 9):

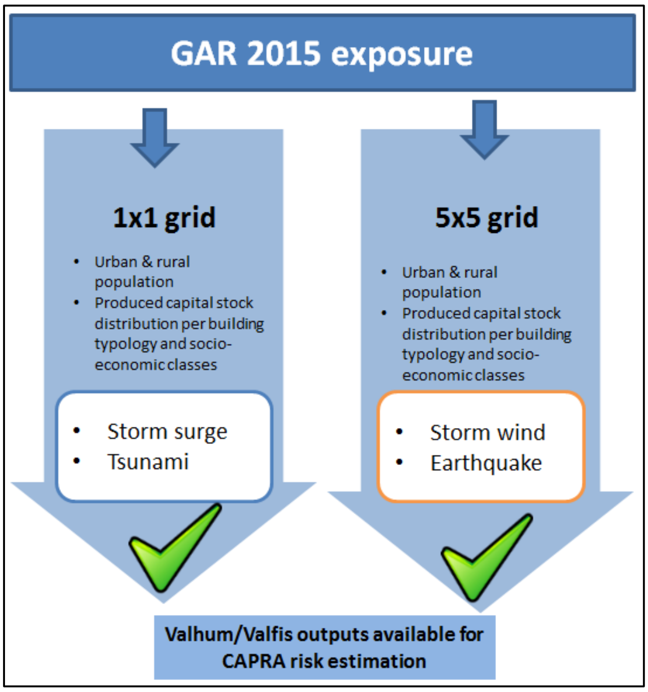

Figure 9: Main outputs for GEG-2015

- A 5x5 global exposure grid suitable per Risk calculation mainly in case of earthquakes and cyclones (wind).
- A 1x1 global coastal (10 km) exposure grid suitable per Risk calculation in case of tsunamis and storm surges.

These two major outputs were released in January 2014 and they will be publicly available in 2015, through the Preview platform[12]

Both outputs satisfy the five essential elements necessary for further probabilistic risk assessments:

- ID;
- Geographic location;
- Construction type for vulnerability classification;
- Exposed economic value;
- Human occupation.

Each record (exposed value) in the GEG-2015 represents a certain structural building typology of a certain income level/sector in a certain area with a point representation in the centroid of the 5x5/1x1 cell.

The GEG-2015 is an (off-line) geo-database developed using open source PostgreSQL/PostGIS software. The database supports several kinds of queries including extraction, aggregation/disaggregation of both GIS and tabular data.

3.1 Improvements from GAR 2013

Several essential improvements have been done since the 2013 edition of GEG they include (Figure 10):

- The addition of the rural population, while 2013 editions only focus on urban settlements. Thus the number of exposed assets has been multiplied of around 18 times.
- Utilization of an improved built-up mask, the BUREF, provided by JRC; enhancements were considerable especially in those areas where MODIS 500 showed clearly its limitations (e.g. highly populated regions of Southeast Asia).
- Updated base layers including most of the SEI and the population distribution grid (LandScan)
- Utilization of construction types, provided by WAPMEER, based on sub-national information per 18 countries, representing about 50% of the world's total population.
- Improvement of the geographic distribution of the produced capital stock. It has been weighted by means of a new indicator: the coefficient of variation between the national values of GDP and those at subnational level.

[12] http://preview.grid.unep.ch/

- Generation of a non-aggregated grid (1x1 coastal grid) for coastal exposure to small scale hazards

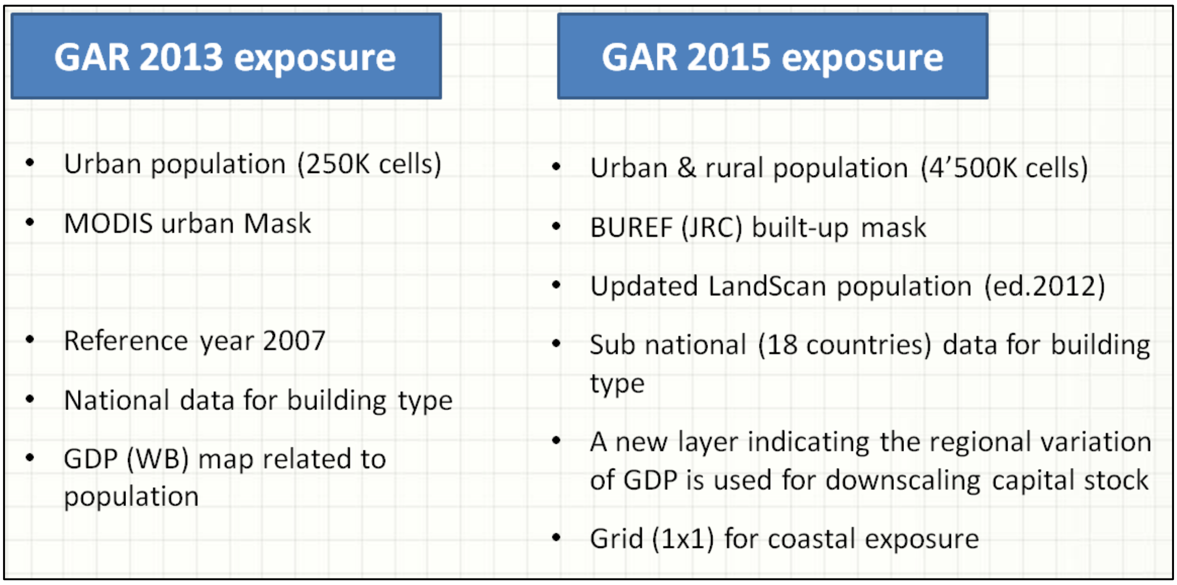

Figure 10: Differences between GEG-2013 and 2015

4. Conclusions

The purpose of GEG 2015is to give an order of magnitude of the potential economic value of exposed assets, in the perspective of a comparison between countries. The consistency of the methodological approach used in the development of GEG, as well as the choice of the best data currently available for its implementation, have produced a product fully adapted to the needs of the global model of the evaluation of probabilistic risk.

The format and resolution appears to be appropriate for the scope. The 5 x 5 km grid covers the need for those hazards that are affecting wide areas. However, the dataset would need to be improved for hazards which are affecting smaller areas (e.g. landslides, tsunamis, floods). As a result, in the case of tsunamis and storm surges a fine resolution 1x1 global coastal (10 km) grid has been developed. This dataset uniquely allows having better information on the spatial distribution of the exposed assets, because it consists of a non-aggregated version of the 5 x 5 km grid generated using the same inputs. In order to improve it, an approach with bottom up (critical facilities) contribution for coastal areas will be highly advisable.

The development of GEG-2015 is based on some assumptions and shows different points that can be improved especially if new data become available.

- The GEG-2015 was fundamentally constructed using national indicators that were successively disaggregated onto a 5x5 grid: this does not take into account important local variations across countries. This limitation is partially attenuated, in the case of wealth by using a regional distribution raster of GRP as weight when distributing the national capital stock.
- Comparisons with actual insurance information indicate that the values from produced capital provided by the World Bank are substantially underestimating the true values of properties in the countries. As the underestimation seems to be a

general feature it might not have a significant impact when the total risk is compared between countries but it becomes an issue if absolute loss figures are concerned. Further improvements could include information coming from insurance commissioners, census data and real estate statistics.

- The capital stock in each cell is distributed on the basis of the number of persons living in that cell and does not take into account the real value of the assets of the cell. In this case a partial improvement is done by using a factor (unitary costs and surfaces) during the calculations of capital stock for each asset.

Figure captions

References

Bhaduri, B., Bright, E., Coleman, P., Dobson, J., 2002. LandScan: Locating people is what matters. Geoinformatics 5, 34–37.

Bright, E.A, 2002. LandScan Global Population 1998 Database.

Compass International Inc., 2012. Global construction cost and reference yearbook (2012).

De Bono, A., 2013. The Global Exposure Database for GAR 2013. UNISDR, Geneva, Switzerland.

De Bono, A., Mora, M.G., 2014. A global exposure model for disaster risk assessment. Int. J. Disaster Risk Reduct. doi:10.1016/j.ijdrr.2014.05.008

Elvidge, C.D., Baugh, K.E., Zhizhin, M., Hsu, F.-C., 2013. Why VIIRS data are superior to DMSP for mapping nighttime lights. Proc. Asia-Pac. Adv. Netw. 35, 62. doi:10.7125/APAN.35.7

Gosh, T., Elvidge, C.. D., Powell, R.L., 2010. Shedding Light on the Global Distribution of Economic Activity. Open Geogr. J. 3, 147–160. doi:10.2174/1874923201003010147

Jaiswal, K., Wald, D., Porter, K., 2010. A Global Building Inventory for Earthquake Loss Estimation and Risk Management. Earthq. Spectra 26, 731–748. doi:10.1193/1.3450316

OECD, 2011. Land valuation in the National Accounts questionnaire. (No. STD/CSTAT/WPNA(2011)7).

Pesaresi, M., Blaes, X., Ehrlich, D., Ferri, S., Gueguen, L., Haag, F., Halkia, M., Heinzel, J., Kauffmann, M., Kemper, T., Ouzounis, G.K., Scavazzon, M., Soille, P., Syrris, V., Zanchetta, L., European Commission, Joint Research Centre, Institute for the Protection and the Security of the Citizen, 2012. A global human settlement layer from optical high resolution imagery concept and first results. Publications Office, Luxembourg.

PwC (PricewaterhouseCoopers), 2013. UNISDR and PwC – Working together to reduce disaster risks. ISDR, Geneva, Switzerland.

Satterthwaite, D., Owen, D.L., 2006. Outside the large cities: the demographic importance of small urban centres and large villages in Africa, Asia and Latin America. IIED, London.

Schneider, A., Friedl, M.A., Potere, D., 2009. A new map of global urban extent from MODIS satellite data. Environ. Res. Lett. 4, 044003. doi:10.1088/1748-9326/4/4/044003

SEDAC (Socioeconomic Data and Applications Center), 2012. Global Rural-Urban Mapping Project (GRUMP). Palisades, NY 10964 USA.

Tolis, Rosset, Wyss, 2013. Detailed building stock at regional scale in three size categories of settlements for 18 countries worldwide. Geneva, Switzerland.

UNISDR, 2009. Global Assessment Report on Disaster Risk Reduction: Risk and poverty in a changing climate., United Nations International Strategy for Disaster Reduction. ed, GAR. Geneva, Switzerland.

United Nations University, 2012. Inclusive Wealth Report 2012: Measuring Progress Toward Sustainability. Cambridge University Press, Cambridge.

World Bank, 2006. Where is the wealth of nations?: measuring capital for the 21st century. The World Bank, Washington, D.C.

World Bank, 2011. The changing wealth of nations: measuring sustainable development in the new millennium, Environment and development. World Bank, Washington, D.C.

World Bank's annual World Development Report (WDR) [WWW Document], n.d. URL http://wdronline.worldbank.org/worldbank/a/incomelevel (accessed 4.6.14).

Wyss, M., Tolis, S., Rosset, P., Pacchiani, F., 2013. Approximate Model for worldwide building stock in three size categories of settlements. Technical report prepared for GAR 2013. ISDR, Geneva, Switzerland.