[Slide 1]

Hello Everyone, my name is Georgiana Esquivias and today I will be presenting the method for developing a building exposure database for Nepal.

[Slide 2]

We will go step-by-step through a series of flowcharts that assists in organizing and determining the type of data that is required for a Level-3 building exposure.

[Slide 3]

A Level-3 Building Exposure improves upon country level data by identifying sub-national variations in construction patterns, such as building types or building density, based on climate or cultural region-specific norms. As well, a Level-3 building exposure identifies major urban areas and enhances building counts or structural mapping scheme.

[Slide 4]

To develop the building exposure, the METEOR team utilized various sources of Earth Observation imagery, national statistics and reports, and journals. Additional source reference will be found in the individual flowchart section.

[Slide 5]

To develop the building exposure, the METEOR team utilized various sources of Earth Observation imagery, national statistics and reports, and journals. Additional source reference will be found in the individual flowchart section.

[Slide 6]

Using Population and housing census data is the most common method for inferring number of buildings. However, because of the high-resolution Village Development Committee (or VDC) data provided by the National Society for Earthquake Technology (NSET) population was not required.

Another factor that played a key role in determining the structural distribution was the building height. This information allowed us to make a direct correlation between number of stories and structural types, as well as perform general heuristic checks for any outlier data.

There were two sources for the building height values; the Integrated Public Use Microdata Series (IPUMS) and the field survey data collected by Kathmandu Living Labs for the Humanitarian OpenSteetMap Team (HOTOSM) (Humanitarian OpenStreetMap Team, 2019).

In areas designated as rural development the height distribution was acquired from IPUMS, specifically heights for unreinforced brick masonry UFB-1 and UFB-5.

Based on discussion with Sharad Wagle (a Nepali Structural Engineer), cement mortar was assumed for those unreinforced masonry structures 5 stories or greater. The increased durability and bonding strength (in comparison to mud mortar) required for multi-story buildings was the logic behind such decisions. For low-rise, rural regions, mud mortar was assumed as the bonding agent.

For the remaining non-rural development pattern areas, the HOTOSM building survey was used to establish the relationship between the structural type and story height by development pattern.

[Slide 7]

The data collected by the Kathmandu Living Labs field team was used to characterize the height distribution in urban areas. In rural areas, the distribution of buildings by administrative level-2 was used to characterize the building stock, as gleaned from the IPUMS data.

Building height varies throughout the country and are typically dependent on the terrain and urbanity of the region. From our data collection exercise, we observed that most structures are primarily low-rise.

[Slide 8]

Here we observe the gradual shift of predominant structural types throughout the region, most likely driven by overabundance of or lack of construction materials and typical construction practices. Unreinforced masonry is predominant in most of the country, with the exceptions of wood-framed structures on the southern border and infilled reinforced concrete structures near the urban cores.

[Slide 9]

The distribution of buildings by structure type typically receives the most concerted attention in a risk study, and this project was no exception. The sources used to develop the exposures include infield surveys, micro-census data, and expert opinion.

[Slide 10]

Due to the active involvement from NSET and the high-quality census data, the structural distribution data was quite accurate for Nepal.

Steps for Estimating of structural distributions, or mapping schemes was two-fold. First, a structural engineer conducts a web reconnaissance of any available data regarding both typical construction materials and methodologies within the region, as well as any data that could infer the structural distribution within the country. Sources such as the World Housing Encyclopedia [WHE], Prompt Assessment of Global Earthquakes for Response [PAGER] (Wald, et al. 2008), and Global Earthquake Model [GEM] (Brzev et al., 2013) were reviewed to identify all known structural types within the country. These preliminary types were validated through the Nepal National Building Code and Google StreetView survey.

After the web reconnaissance the structural engineer begins to formulate the mapping scheme by development pattern.

Using the 2011 Nepal VDC-level census data the rural development pattern was assigned a mapping scheme using the building wall material type and number of household value. The building height values for UFB-1 and UFB-5 within the rural development pattern zones were obtained from IPUMS data set.

[Slide 11]

Kathmandu Living Labs conducted a building survey for HOTOSM and implemented a stratified sampling strategy and used a Bayesian updating approach.

The survey data provided the structural mapping scheme for each development pattern**.**

[Slide 12]

The HOTOSM building survey data was reviewed and verified by a local in-country engineer from NSET and in-house engineer from ImageCat. All of the mapping schemes were mapped to the PAGER standard structural types and then to the GEM Taxonomy. These structure types were overlaid with the manually delineated development pattern sample polygons to create a refined mapping scheme. The distribution of structural classes closely reflected known patterns of construction practices with respect to elevation, proximity to India, and urbanity. A final round of sanity checking is conducted by ImageCat.

[Slide 13]

Development patterns are patterns of construction in a given country that typify the building structure development and density as much as possible. They sometimes correspond with land use, but not always. The development patterns are determined by a structural engineer working with GIS analysts to conduct a web reconnaissance exercise using Google Earth, and structural distribution web searches to characterize the urbanity density and development patterns for each country. For Nepal, the ImageCat engineer characterized 8 development pattern types.

[Slide 14]

Here we see the final results of the mapping schemes as percentage of households by structural class mapped by administrative VDC units in Nepal.

[Slide 15]

Typically, the number of buildings is estimated from populations and average number for persons per household. As previously mentioned in the case of Nepal, NSET was able to provide the estimated number of households by Village Development Committee (VDC) which is approximately 700 administrative zones. However, Kathmandu was covered by a single zone, and exposure still needed to be converted into the number of buildings.

[Slide 16]

The total number of building was inferred using a combination of datasets that include IPUMS, VDC census data, OpenStreetMap survey data, and aggregated OSM building count raster. These were used to estimate the average number of households per building type for each VDC census unit.

[Slide 17]

Using the Nepal census number of household per VDC by wall material type a relationship between the average total building area using structural type was created to calculate total number of buildings for the rural zones.

[Slide 18]

For each development pattern, a project analyst reviewed 15-arcsecond grid cells selected using a random number generator and identified cells where at least 90% of the buildings were represented in the OpenStreetMap building footprint data.

(Humanitarian OpenStreetMap Team, 2019).

[Slide 19]

These selected grids were used to train a regression model using 3-arcsecond Earth observation data which includes the Sentinel-1 SAR mosaic with dual polarization bands, Facebook Connectivity Lab & CIESIN's High Resolution Settlement Layer, (Facebook Connectivity Lab & Center for International Earth Science Information Network [CIESIN], 2016), and the Global Urban Footprint (GUF) from DLR.

[Slide 20]

In addition, the model was constrained using OpenStreetMap building footprints with an area of 10 meters square to 5,000 meters square to create an aggregated 3-arcsecond building count raster as a minimum value.

We have found that Sentinel-1 SAR mosaic proved more effective in highly developed regions, whereas GUF was better suited for rural areas.

[Slide 21]

Dasymetric mapping is the process of spreading the number of buildings from the VDC census data to the 3-arcsecond grid by way of statistical assessment of moderate resolution Earth Observation (EO) data.

To collect Earth Observation indicators of settlements and density of buildings, various remote sensing data sets were used. These included:

[Slide 22]

The follow remotely sensed data products were used for the dasymetric mapping. The data was obtained from the US National Oceanic & Atmospheric Agency (NOAA), NASA, the US Oak Ridge National Laboratory, the European Commission Joint Research Centre, the German Aerospace Center, WorldPop, and OpenStreetMap.

[Slide 23]

These remotely sensed earth observation products and building footprint aggregates establish the distribution statistics for dispersing buildings by urban density and determining the development pattern types throughout the country. Each of the Earth Observation products individually, and in a combination, act as weights to disperse the known number of households per VDC by structural type and development pattern to 3-arcsecond grid cells.

For example, with development patterns 1 or 2 (which resembling the rural or single-family residential communities) the even building distribution of the VDC is reallocated only to grid cells within the VDCs associated to human settlement.

When determining the weights for distribution within a given VDC, several machine learning algorithms were run using the Earth Observation to develop a prediction model. In the case of Nepal which has very complex terrain, GUF was highly weighted, and support vector was determined to be the most effective AI tool. For developments with higher populations and building density, the reallocation of buildings becomes more complex and requires a more detailed examination of the structural types and mapping schemes.

[Slide 24]

To prevent unpopulated areas from being considered as settlements, especially in the highly mountainous terrain, a mask was created by combining the extents of night-time light (VIIRS), ambient population (LSCAN), and Landsat based Global Human Settlement Layer. These data sets were reclassified into inhabited vs. uninhabited masks using a minimum threshold value determined by visual inspection. These minimum values correspond to even the sparsest human presence. This mask was used to subset the high-resolution Global Urban Footprint (GUF), Sentinel-1 based GHSL product, Sentinel-1 SAR mosaic, and the High-Resolution Settlement Layer (HRSL) that go into the machine learning process to come up with the development patterns; subsetting the data sets decreases the processing time.

[Slide 25]

To characterize building density in more populated areas, analysts digitized development patterns training polygons in the top 10 most populated cities in the country.

[Slide 26]

For these cities, the digitized vector data was used directly, rather than the classified development pattern grid. The sample vectors were digitized using the Google Earth area tool and saved as KMLs. The basemap vintage and source used during digitization vary by region and zoom level. However, the most current high-resolution satellite images are used. The training polygons and the moderate resolution Earth Observation products described above are used in a machine learning process (CART algorithm (Breiman et al., 1984; Khaled et al., 2014), Random forest (Breiman, 2001), and Support Vector Networks (Cortez et al., 1995)) for assigning the development patterns throughout the country, which informs the estimated building density. The intensity of urbanity correlates to both the building density and the structural distribution. For Nepal, the ImageCat engineer characterized 8 development pattern types.

[Slide 27]

The process of dasymetric mapping in Nepal was challenging due to the prevalence of mountainous terrain. Both the Sentinel-1 SAR mosaic and optical Earth Observation data often interpret steep terrain as associated to built environment. To remedy this situation, the project team reviewed the results on a regional basis and made adjustments.

Here is an example of the SAR mosaic built-up indicators. In the image to the left, the bright blue strips of higher reflectance can clearly be seen to correlate with the OSM building footprints, seen in red in the image on the right. It can also be seen, that regardless of the relatively pervasive footprints data, many urban areas are missed and cannot be inferred.

In the delta in the southern half of the image, the SAR mosaic is highly effective. As we move farther north into more hilly terrain detecting built-up area becomes more difficult, but in the high terrain areas in the Northeast of the images the sensor is no longer effective.

[Slide 28]

The replacement cost values were provided by National Society of Earthquake Technology (NSET).

[Slide 29]

The replacement cost values were derived by an in-country survey conducted by the Kathmandu Living Labs and HOTOSM in the surrounding urban and rural regions of Kathmandu.

[Slide 30]

There are the follow replacement cost values for Nepal in USD.

[Slide 31]

Here we have the estimated replacement cost values for Kathmandu and the surrounding region.

[Slide 32]

For the building area, data was collected from multiple sources and applied in separate geographies to estimate building area.

[Slide 33]

As with the building height, the total building area was calculated using a combination of IPUMS, Nepal census data, Humanitarian OpenStreetMap Team (HOTOSM) survey data (Humanitarian OpenStreetMap Team, 2019), and 3-arcsecond aggregated OSM building area raster created by ImageCat (ImageCat, Inc., 2019).

For the rural development pattern type, the height distribution from IPUMS, specifically heights for unreinforced brick masonry was used along with the average building footprint area from the aggregated building area raster to determine the average total building area.

For all other non-rural development pattern designated zones the Humanitarian OpenStreetMap Team (HOT) building survey was used to establish the total building area by structure type per development pattern using the surveyed building footprint and height values.

[Slide 34]

There is a great deal of information that goes into creating a building exposure database. The aim of these flow charts are to:

1) Assist future researchers in update and refine the exposure database as a whole or finely tune an individual section with improved data.

2) Help assist in organizing metadata collection and

3) Help work through the logic of the processing steps.

[Slide 35]

That concludes the flowchart section. We hope that these flowcharts assist you and future teams in conducting, updating or refining a level-3 building exposure.

[Slide 36]

Thank you and have a good day.