Transforming Census Data in Pakistan into Spatial Database for Earthquake Loss Estimation Models

M. Ahmed, S.H. Lodi, M.M. Rafi & N. Alam
NED University of Engineering and Technology, Pakistan

15 WCEE
LISBOA 2012

SUMMARY:
Earthquake loss estimation models are based on the information regarding the elements at risk. The built environment is the most important element at risk as lives of people are lost due to collapsed buildings. This paper presents the results of development of a geographical information system (GIS) for population and buildings in Pakistan. The district level census data was employed and was transformed into a GIS based 1x1 km grid format. The building to population ratios for the districts were calculated with the help of population and building statistics available in the census report. The random point generation technique was employed for population estimation in the cell of size 1x1 km. The cell population was multiplied with the building ratio to determine the estimated number of buildings in the cell.

Keywords: spatial database, loss estimation, element at risk, census, geographical information system

1. INTRODUCTION

Pakistan is situated at the triple junction of Arabian, Eurasian and Indian plates. The tectonic settings of Pakistan have made it a seismically active region. The devastations made first by the Quetta earthquake of magnitude 7.5 (which killed as many as 25,000 people in 1935) and recently by the Kashmir earthquake of magnitude 7.6 (which killed 73,000 people in 2005) can never be forgotten by the Pakistani nation. About 2.8 million people were also made homeless owing to the damage of 450,000 buildings due to the Kashmir earthquake (Rossetto and Peiris 2009). This earthquake indicated lack of capacity and preparedness in Pakistan to deal with the disaster and created the realization of the importance of earthquake mitigation efforts in the Country. Earthquake loss estimation is considered as one of the important components of disaster management and risk assessment models. The development of inventory of element at risk is an essential requirement for carrying out loss estimation. Since the objective of disaster management is to save lives of people, the built environment represents the most important element associated with risk. In this regard, building population, construction type, height and age are the data required in order to calculate the risk. Obtaining this data independently for a country is a mammoth task which requires considerable resources in terms of cost, time and manpower. The use of census data may provide a means of obtaining information on the aforementioned elements at risk. This information can be transformed into geographical information system (GIS) for multiple uses.

This paper describes the details of the work that was carried out to develop a GIS of population and building types in Pakistan on a grid of 1x1 km. The population census data (PCO 1998) was employed in order to obtain information of building typologies and population in the Country. This data was transformed into a GIS based grid format which contains information regarding habited land throughout Pakistan, and population and building distribution in the habited areas. The method employed and the results obtained are discussed in the forthcoming sections.

2. PHYSIOGRAPHY OF PAKISTAN

The physiography of Pakistan is heterogeneous comprising of grasslands, deserts, jungles, mountains, and plateaus stretching from the coastal areas of the Arabian Sea in the south to the mountains of the Karakoram range in the north. Pakistan geologically overlaps both with the Indian and the Eurasian tectonic plates where its Sindh and Punjab provinces lie on the north-western corner of the Indian plate while Balochistan and most of the Khyber-Pakhtunkhwa lie within the Eurasian plate which mainly comprises the Iranian plateau, some parts of the Middle East and Central Asia. The Northern Areas and Azad Kashmir lie mainly in Central Asia along the edge of the Indian plate and hence are prone to violent earthquakes where the two tectonic plates collide (Kazmi and Qasimjan 1997).

Extremely active faults are located in this region that makes a majority of population at risk. Northern and western sections of Pakistan are more sensitive to earthquake activity because they are surrounded by the micro plates of Afghanistan and Iranian and Indian plate. Chaman fault is located in the western section that goes along the northern Makran range, passes Quetta and then to Afghanistan. A fault also run along the Makran coast and it is found to be of the same nature as the west coast fault along the coast of Maharashtra, India (MoHW 2008).

This zone forms the boundary between the Arabian and the Iranian micro-plate, where the former subducts or dives beneath the latter. Thrust zones run along the Kirthar, Suleiman and Salt ranges. There are four faults in and around Karachi and other parts of deltaic Indus, and Makran coast. The first is the Allahbund fault that passes through Shahbundar, Jah, Pakistan Steel Mills, and runs through eastern parts of the city and ends near Cape Monz. This fault, in fact, has caused extensive damage in the past many centuries in the deltaic areas. The destruction of Bhanbhor in the 13th century and damage to Shahbundar in 1896 were caused by this fault (Pararas-Carayannis 2008). The other one emanates from the Rann of Kutchh. The third one is the Pubb fault which ends into Arabian sea near Makran coast and the last one is located in the lower Dadu district near Surajani and falls in the vicinity of Karachi. Tsunamis or tidal waves have also affected the coast of Pakistan. The worst case was in 1945 when an earthquake of magnitude 8 struck the Makran coast, waves as high as 12 meters were reported (Byrne et al. 1992). Being located close to the collision boundary of the Indian and Eurasian plates, Pakistan lies in a seismically active zone. Owing to high population density near seismically active areas, it is imperative that buildings should withstand the seismic hazard to which these may be exposed during their life time.

3. METHODOLOGY OF WORK

The data of building types in the report published by PCO (1998) is available in the form of the construction materials used in walls and roofs. Based on the available information, the wall and roof construction materials were classified into four categories each (Table 3.1). This work was carried out in view of utilization of this information, at a later stage, in carrying out building vulnerability analysis using their fragility curves based on damage grades, as defined by EMS-98 (Grünthal 1998). The building types identified and their equivalent EMS-98 nomenclatures are given in Table 3.2. The proportion of these building types in the overall building typology is illustrated in Figure 3.1.

Table 3.1. Possible Combinations based on Material of Construction

Material used in Walls	Material used in Roof
Baked Bricks/ Blocks/ Stones	RCC/RBC
Un-Baked Bricks/ Earth Bound	Cement/ Iron sheets
Wood / Bamboo	Wood/Bamboo
Other	Others

Table 3.2. Building types and their nomenclature

Building Type	EMS-98 Nomenclature
Brick Masonry Buildings	M5
Adobe Masonry Buildings	M2
Reinforce Concrete (RC) Buildings	RC1
Concrete Block Masonry Buildings	M8
Wood/Bamboo Reinforced Masonry Buildings	M7
Stone Masonry Buildings	M1
Other Masonry Buildings	OO

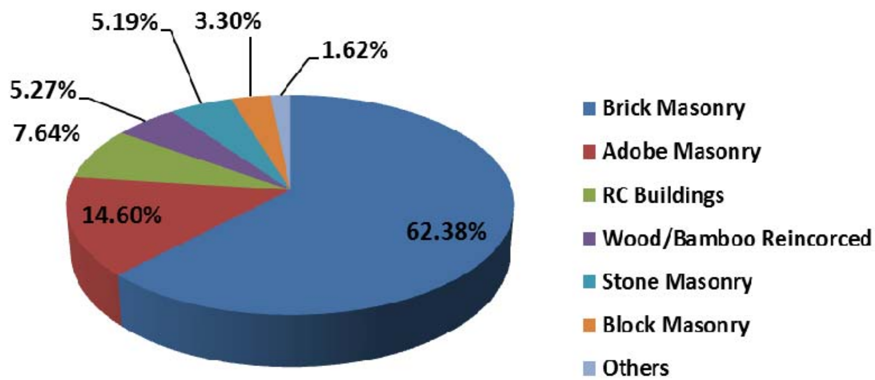

Figure 3.1. Proportion of different building types in Pakistan

Transforming Census data into GIS based grid format is generally based upon approximation techniques. In the method employed in the presented study, the GIS based district map of Pakistan was developed using the administrative regions in Pakistan, as given in census data (PCO 2001). Since this data was compiled on the district level the shapes of these districts on the GIS appeared in non-regular shapes or grids. On the other hand, for earthquake loss estimation layers consisting of smaller and regular patches are required. Therefore, a transformation of district level data into perfect geometry was carried out to overcome this shortcoming of the data. The inhabited land was identified within each district using high resolution satellite imagery available on Google Earth. The building and population data from the census report (PCO 2001) was incorporated within the inhabited land layer. The population in the inhabited land of the district was distributed using random point generation technique which employed Hawth's analysis tool (Beyer 2004) as available for ArcGIS (ESRI 2011); each generated point represents 1000 persons. The population distribution using the random points was based upon the statistical technique of computer generated pseudo numbers. The inhabited land in each district was divided into 1x1 km grid to achieve the uniqueness in geometry. The population points in each cell of 1x1 km size were counted to obtain the cell population. This was carried out using spatial join technique available in ArcGIS (ESRI 2011). The estimated number of buildings and their types were obtained by multiplying the cell population with building ratio of each building type within the district. Figure 3.2 shows the flow diagram of the aforementioned processes.

4. DISCUSSION ON RESULTS

4.1 Administrative Map of Pakistan

The total area of Pakistan is approximately 796,096 km2; whereas the population of the Country is around 140 million with 15 million housing units (PCO 1998). The printed district level map of Pakistan was obtained from the census report (PCO 2001). This map was electronically geo-referenced and digitized using ArcGIS (ESRI 2011) to obtain spatial extents of administrative boundaries of all

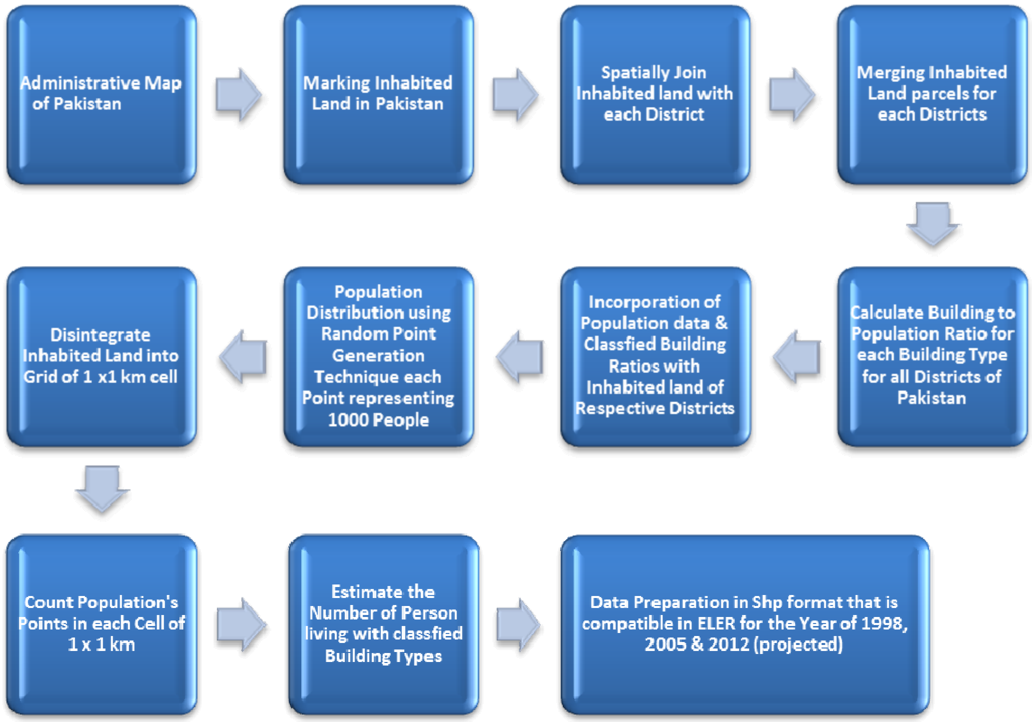

Figure 3.2. Flow diagram of the processes

Figure 4.1. Digitized Administrative Map of Pakistan

the districts in Pakistan. The total number of districts in Pakistan comes out to be 101 districts in addition to 13 tribal administrative regions. An image of this map is shown in Figure 4.1.

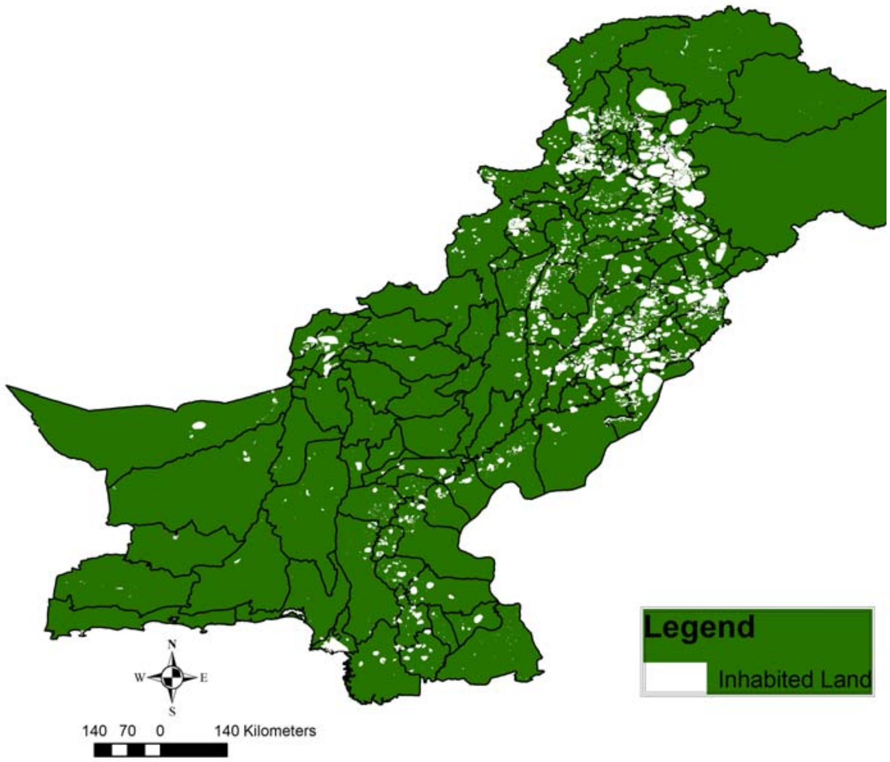

Figure 4.2. View of inhabited land in Pakistan

4.2. Identification of Inhabited land

The distribution of population in all the districts is not the same in Pakistan and the population was more concentrated in only few districts. For earthquake loss estimation model, identification of inhabited area within each district is important. As mentioned before, high resolution satellite data from Google Earth was employed in the development of inhabited areas. The boundaries of these inhabited areas were digitized and the layer representing these areas was incorporated with the administrative map of Pakistan (Figure 4.1). The results of this work are illustrated in Figure 4.2.

4.3. Random Point Generation

The inhabited region layer, which is overlying the administrative region layer in the GIS environment (Figure 4.2), is spatially join with the administrative region layer using ArcGIS (ESRI 2011). Spatially join is an innovative technique in GIS that can join the layers that are on the same spatial extent. Since the data was not available for each parcel of inhibited region and there were more than one inhabited region polygons in each district (Figure 4.3), it was necessary to integrate all parcels belonging to a district internally while leaving the spatial extent of these all parcels unchanged. "Dissolve" function is used to merge the parcels having the same district name. Once the parcels of inhabited region layer are merged and recognized with district name the inhabited land layer was joined with the already developed MS Excel table. This table contains census information in term of district population, population growth rate and the identified building type (Table 2) ratio for all the districts of Pakistan. The district population is then distributed over the inhabited land in a district using the random number generation technique in such a way that the generated points reflect the population in accordance to the census data (PCO 2001) (Figure 4.4).

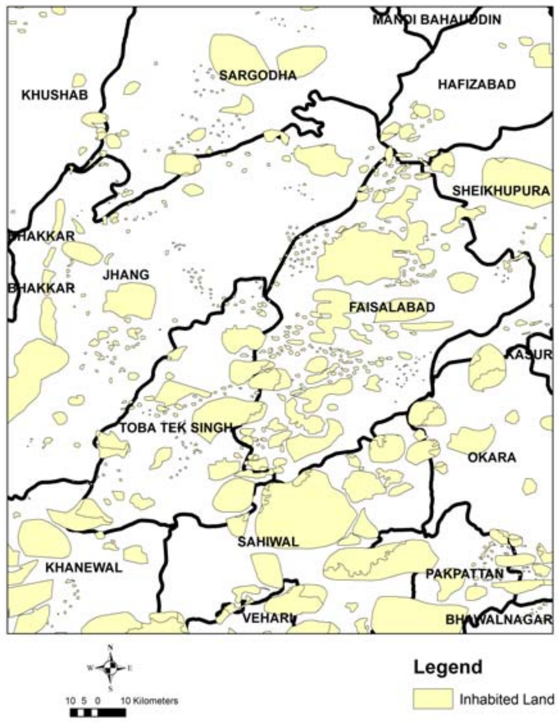

Figure 4.3. Inhabited land parcel belonging to Karachi District

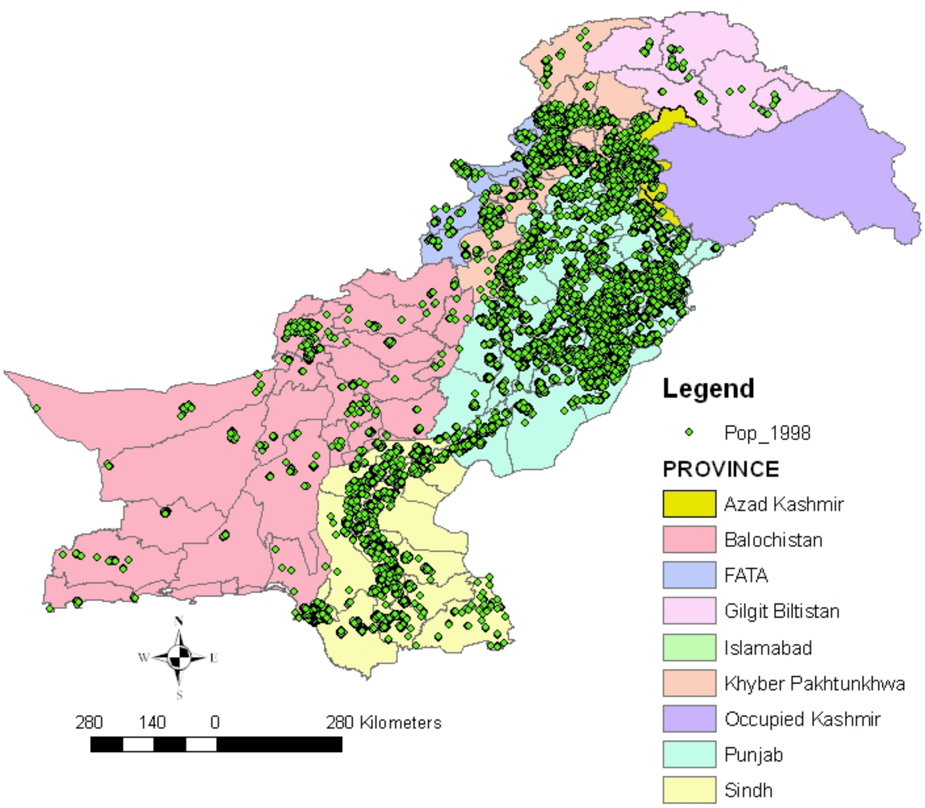

Figure 4.4. View of Population Distribution using Points

4.4. Grid Overlay of 1x1 KM

As mentioned before, grids of cell size 1x1 km were developed using Hawth's analysis tool (Beyer 2004). The total number of cells came out to be 2,476,827 for the entire Country. All the cells that were located either in unpopulated regions or outside the geographical boundaries of Pakistan were removed. Figure 4.5 shows an example of regular pattern of inhibited land in few districts of Pakistan. Similar patterns were obtained for the entire Country and are available in the GIS format.

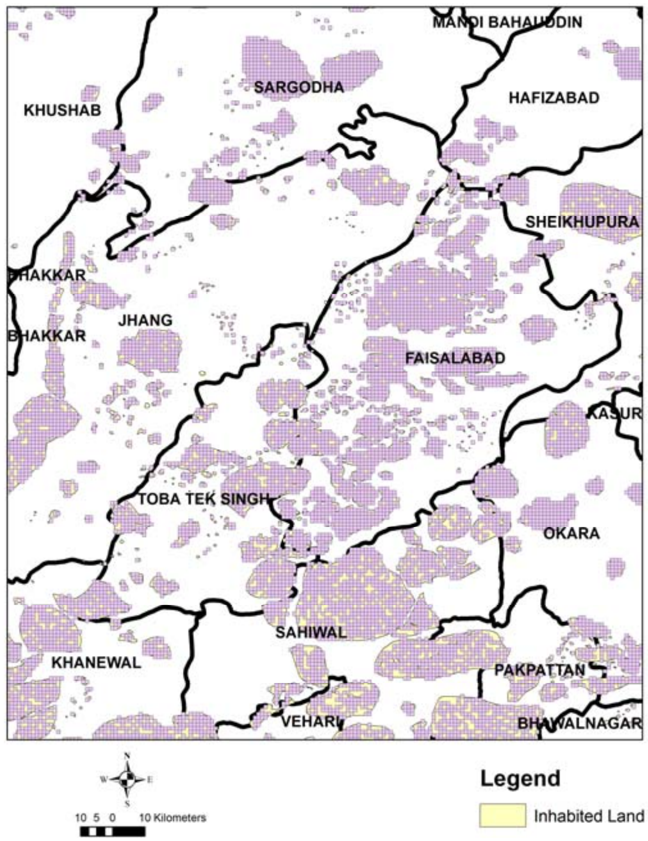

Figure 4.5. A view of grids for inhibited land developed on GIS

5. CONCLUSIONS

This paper presented the results of a study which was conducted to estimate the population, and buildings and their types all over Pakistan. This data may be employed in earthquake loss estimation models to determine possible life and property loss during an earthquake. The census data at the district level was transformed into GIS based 1x1 km grid format. The census report was published in 2011. The building to population ratios for the districts were calculated using the data available in the census report. High resolution images from Google Earth were employed for identification of inhabited land. Random point generation method was employed for population distribution in each cell of the inhabited land. The cell population was multiplied with the building ratio to obtain total buildings in the cell.

AKCNOWLEDGEMENT

The authors acknowledge the support provided for this study by Earthquake Model for Middle East (EMME). The first author is also thankful to Dr Shahid Ali for his guidance during the study.

REFERENCES

Beyer, H.L. (244). Hawth's Analysis Tool for ArcGIS 2004. Available from http://www.spatialecology.com/htools.

Byrne, D.E., Sykes, L. and Davis D.M. (1992). Great thrust earthquakes and aseismic slip along the plate boundary of the Makran subduction zone. Journal of Geophysical Research **97:B1**,449-478.

ESRI. (2011). ArcGIS Desktop: Release 10. Redlands, CA: Environmental Systems Research Institute.

Grünthal, G. (1998). European Macroseicmic Scale 1998 (EMS-98), Cahiers du Centre Européen de Géodynamique et de Séismologie 15, Centre Européen de Géodynamique et de Séismologie, Luxembourg.

Kazmi, A.H. and Qasimjan, M. (1997). Geology and Tectonics of Pakistan, Graphic Publishes, Pakistan.

Ministry of Housing & Works (MoHW). (2008). Building Code of Pakistan – Seismic Provisions, Government of Pakistan.

Pararas-Carayannis, G. (2008). Future Tsunami Disaster in the Indian Ocean. *14th World Conference on Earthquake Engineering*.

Population Census Organization (PCO). (2001). Population and Housing Census 1998, Government of Pakistan, Statistics Division.

Rossetto, T. and Peiris, N. (2009). Observations of damage due to the Kashmir earthquake of October 8, 2005 and study of current seismic provisions for buildings in Pakistan. Bulletin of Earthquake Engineering **7**,681–699.