

# URBAN SOCIAL VULNERABILITY ASSESSMENT USING OBJECT-ORIENTED ANALYSIS OF REMOTE SENSING AND GIS DATA. A CASE STUDY FOR TEGUCIGALPA, HONDURAS.

A. Ebert<sup>a</sup>, N. Kerle<sup>b</sup>

<sup>a</sup>UFZ – Helmholtz Centre for Environmental Research, Permoserstr. 15, D-04318 Leipzig, Germany - annemarie.ebert@ufz.de

<sup>b</sup>ITC – International Institute for Geo-Information Science and Earth Observation, PO Box 6, 7400 AA Enschede, The Netherlands - kerle@itc.nl

Commission VII, WG VII/7

**KEY WORDS:** Social vulnerability assessment, Remote sensing, Object-oriented analysis, Proxy variables, Tegucigalpa

## ABSTRACT:

This paper deals with the assessment of social vulnerability (SV) as a critical component of comprehensive disaster risk assessment. Indicators for SV relate to aspects on different scales. Individual characteristics, such as gender, age and education level, have to be assessed on a very local (individual) scale, whereas indicators such as living conditions, economic development and location of the household can be assessed on the scale of a building, building block, an administrative neighbourhood or city district. In turn, measures to reduce SV and thereby the disaster risk are taken on different levels. Information on SV is notoriously difficult to obtain, and traditionally either detailed field studies or census data have been used. This research, which was done in Tegucigalpa, Honduras, is not focused on individual people, but on the level of buildings and administrative neighbourhoods in the city, with the intention to analyse SV for a central area of 3\*3 km as a potential starting point for more detailed analysis if needed. The central novelty is the use of image-based contextual, object-oriented analysis and the focus on physical proxies as indicators for SV, whereby we focus on landslide and flood hazards. Very high resolution remote sensing data, as well as GIS data and city maps were applied to delineate proxy variables with the goal to analyse four indicators for social vulnerability: (i) socio-economic status, (ii) commercial and industrial development of a neighbourhood, (iii) abundance of infrastructure/lifelines, (iv) and distance to those. The validation of the results was done using a Social Vulnerability Index (SVI) created based on census data. A subsequent stepwise regression analysis showed that eight out of 47 proxy variables were significant and could explain almost 60 % of the variation of the SVI, whereby the slope position (i.e. location of a building) and the proportion of built-up area in a neighbourhood (i.e. neighbourhood composition) were found to be the most valuable proxies. To make the approach transferable to other study areas with different data availability we also indicate where data can potentially be substituted with lower quality information than applied in this study. This work shows that contextual segmentation-based analysis of geospatial data can substantially aid in SV assessment, and, when combined with field-based information, leads to an optimisation of the assessment in terms of assessment frequency and costs.

## 1. INTRODUCTION

The number of casualties and damage caused by natural disasters, while fluctuating strongly, has been increasingly in recent decades. Higher monetary values lead to large losses in developed countries, whereas in less developed countries the amount of people injured or killed and the loss of basic infrastructure are the main concerns after a disaster struck an area. Different living conditions and standards usually lead to different abilities of people to prepare for and cope with natural disasters, thus lead to a certain spatial pattern of vulnerability and resilience. This important fact needs to be considered in risk management. Knowledge about present and future hazards, elements at risk and different types of vulnerability are the place-dependent variables that need to be assessed in order to mitigate disaster risk. In practice, a lack of sufficient data is a main constraint in disaster risk assessment, especially in data scarce regions, such as less developed countries.

Tegucigalpa, the capital of Honduras, a city with more than 1 million inhabitants. Due to a lack of urban planning instruments many people are living in areas that are prone to hazards such as landslides and floods, which are often the only available and affordable spaces for building construction (Angel et al., 2004). The strong population growth leads to ecological and land use changes that increase the hazard and thus the risk. When Hurricane Mitch struck the city in 1998, numerous landslides and severe flooding caused thousands of fatalities and destroyed large parts of the city, including buildings and infrastructure.

Very high resolution satellite data, a digital elevation model (DEM), a normalized digital surface model (nDSM), a set of different GIS data, census data and scattered ground truth information were available for the analysis (Table 1).

## 2. STUDY AREA AND DATA BASE

The hypothesis that remote sensing data can be used for SV assessment was tested on 87 neighborhoods (3 \* 3 km) of



Figure 1. Location of the study area Tegucigalpa in Honduras.

### 3. SOCIAL VULNERABILITY – DEFINITION AND ASSESSMENT

This paper deals with the assessment of social vulnerability, one component of risk that has frequently been neglected in this context. Not only because no consensus on its definition has been found yet, but also because existing assessment approaches are rather time and/or cost intensive.

We follow Clark et al. (1998) in defining SV as “people’s differential incapacity to deal with hazards, based on the position of the groups and individuals within both the physical and social worlds”, which has to be assessed with respect to the particular hazard or combination thereof (e.g. earthquakes and/or landslides). SV cannot be expressed in absolute values or losses. To quantify SV and to make it comparable between regions, indices containing different variables have been developed (Cutter et al., 2003), which are in most cases derived from data collected during community-based approaches or from census data.

While data collected using house-to-house surveys, community-based methods (Kienberger & Steinbruch, 2005) and house-to-house surveys (Palmiano-Reganit, 2005) are suitably detailed, only small areas can be covered. The method is also time-consuming, of low temporal resolution, and up-scalability of the results is questionable. On the other hand, census data based approaches (Cutter et al., 2003, Azar & Rain, 2007) are less time and cost intensive, but have a lower spatial resolution and can due to data availability only be repeated every 5 to 10 years. It has to be considered that SV is not only spatially, but also temporally highly dynamic. The main limitation, however, is that census data are collected for other purposes and that important components of SV, such as hazard perception, are not included.

Source	Date	Format	Resolution
Quickbird	12/2000	Raster	ms: 2.4 m; pan: 0.61 m
ResourceSat P-6	04/2006	Raster	5.8 m
Aerial Photographs	2001	Raster	0.4 m
DTM	n.a.	Vector	1.5 m
Lidar nDSM	03/2000	Raster	1 m
Hazard maps (floods & landslides)	2002	Vector	1:10 000
Main river network	2000	Vector	1:10 000
Various infrastructure	2002 & 2004	Vector	n.a.

Table 1. Data base available for the analysis.

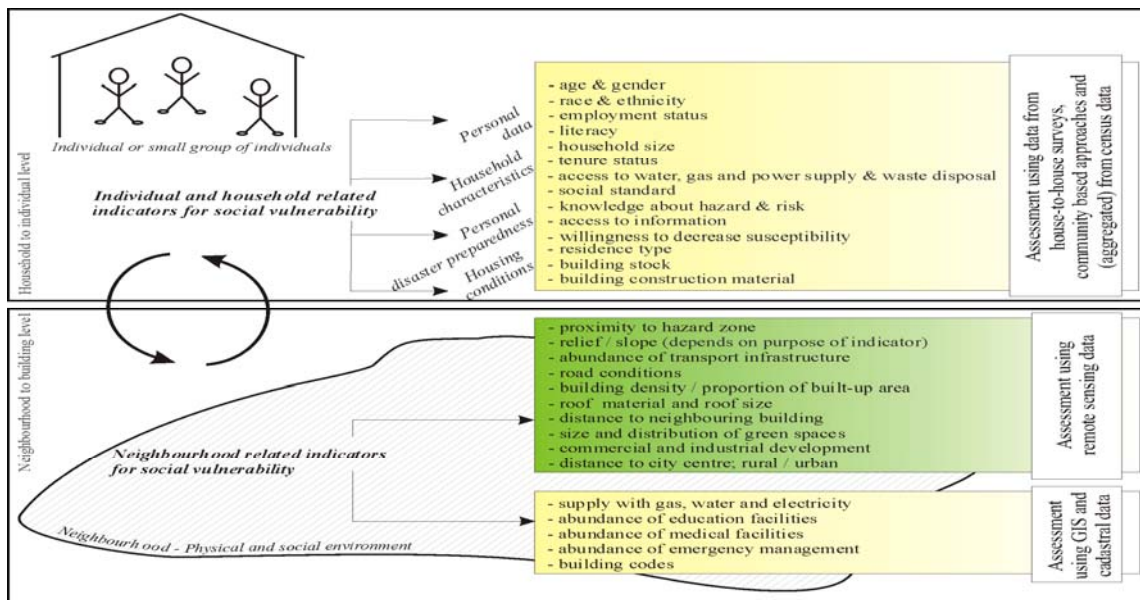


Figure 2. Indicators for social vulnerability assessment on individual/household and neighbourhood level and their proposed assessment methods.

A supplementary cost- and time-efficient approach that can be repeated frequently and that can be used in combination with traditional methods is thus needed. Satellite data have previously shown their high potential for the application in risk-

related topics and with the arrival of high resolution satellite sensors, enhanced image quality, and new image processing methodologies, a continuously growing number of information can be delineated from remote sensing data, particularly in

urban areas. In the field of SV research, on one level this opens up new possibilities. Even though SV tends to relate to personal and individual characteristics, physical indicators and spatial and contextual relationships that relate to the “physical and social world” of the individual (Clark et al., 1998) have previously been identified (Wu et al., 2002; Rashed & Weeks, 2003) and may well be observable in RS data. In Figure 2, indicators for SV assessment on two scales, the individual/household and the neighbourhood scale, are presented with their best-suited assessment methods.

Besides a generally accepted definition a more efficient and comprehensive approach for the assessment of SV is needed. Data availability and low efficiency have been pointed out as two of the most important constraints in current SV assessment (Birkmann, 2005).

#### 4. METHODOLOGY: OBJECT-ORIENTED ANALYSIS OF REMOTE SENSING DATA

So far, studies on vulnerability assessment using satellite data have mainly been focusing on the assessment of physical vulnerability (Müller et al., 2006). But by applying object-oriented image analysis, a wide range of additional, contextual information that is relevant for SV assessment can be extracted from remote sensing images. Very high resolution satellite data (0.6 m ground resolution from Quickbird satellite) have been used in combination with various digital data sets to extract important semantic information, e.g. the position of a building in relation to the hazard zone. This, for example, is a critical feature as the position of each image object in its natural and man-made environment partly determines its SV (Rashed & Weeks, 2003; Clark et al., 1998).

Given that SV is more of a non-tangible concept it has to be evaluated which criteria for its assessment potentially have a spatial expression in an image (Table 2). The selection of indicators for this study was based on the list of generally accepted criteria according to Cutter et al. (2003, Ebert et al., 2007). Selected according to assessment capabilities with remote sensing data and relevance for this study, four main indicators (Table 2) remained. Meaningful proxy variables (parent proxy, second column Table 2) have been found to describe the content of the indicators. These in turn had to be translated into proxy variables that can be directly delineated from remote sensing data and other available information (GIS data, city maps) using object-oriented analysis (supporting proxy, third column Table 2).

Original indicator	Parent proxy	Supporting proxy
<i>Socio-economic status</i>	Settlement type  Topographic location	Proportion of built-up and vegetated area (4 proxies); Road conditions (1); Roof type (7); Texture (10) Slope position (12); Proportion of buildings in hazard zone (2)
<i>Commercial and industrial development</i>	Commercial development	Building heights (7)

<i>Abundance of lifelines</i>	of	Number of lifelines	of	Number of lifelines (1)
<i>Distance to lifelines</i>	to	Distance to lifelines	to	Distance measures (3)

Table 2. Physical proxies identified for the original SV indicators. Numbers in parentheses indicate the actual number of proxy variables for each parent proxy.

A land use/land cover classification was performed in two modes: with and without the usage of an additional normalized Digital Surface Model (nDSM) that was available for a part of the study area. In both modes, spectral as well as spatial and contextual information have been used for the class description. The classification accuracy for the classification with and without the nDSM was 89.1 % and 84.3 %, respectively. Based on the classification results, and by using additional GIS data and city maps, the supporting proxy variables were delineated.

The proportion of built-up and vegetated area was calculated for each neighbourhood, using spatial metrics (Herold et al., 2003). Similarly, the percentage of paved roads compared to the total road network per neighborhood was calculated. Seven different roof types and their neighbourhood proportion were also classified from the Quickbird image, which can give information about the socioeconomic status of the residents. However, field data would be necessary to associate actual roof materials with the roof types classified, as Quickbird’s spectral characteristics are insufficient for such determination. The available service infrastructure was digitized from city maps and quantified for each neighbourhood. Image texture measures were used to describe the distribution of grey values in an image and thus to characterize the heterogenic structures of a city (Tuceryan & Jain, 1998). To describe the topographic location, the slope position was calculated from a digital terrain model (DTM, 1.5 m resolution), to which each image object was then associated. Twelve slope classes in 5° intervals were delineated. All buildings in a hazard zone (flood and landslide) were further masked out, and the percentage compared to all buildings in the neighbourhood calculated. Building heights were further delineated from the lidar data set for a smaller part of the study area, and used to characterize commercial development. Finally, three possible distances to lifelines (less than 100 m, 100 to 250 m, more than 250 m) were defined that refer to the position of each segment in the image. A total of 47 supporting proxy variables were thus delineated (Table 2).

#### 5. RESULTS

The goal of the evaluation was to investigate if the proxy variables can be used to extract information from the satellite data that are relevant for the assessment of SV. We thus tested for a relation between the remote sensing derivatives (proxy variables) and a reference SV index (SVI, modified from Haki et al., 2004) based on census data from 2000 census in Tegucigalpa. The set consist of seven variables (gender, literacy, wall material, roof material, water availability, waste disposal, building use) to calculate the SV for each neighbourhood. The resulting values for each neighbourhood are shown in Figure 3.

We used a stepwise regression model to determine which of the 47 proxy variables (explanatory variables) can best explain the variation of the vulnerability score V (dependent variable) (Jain, 2005). The main observations were the following:

- (i) A high proportion of built-up areas and of buildings on gentle slopes corresponds to low SV.
- (ii) A high proportion of buildings on medium slopes, buildings exposed to landslide hazard, and the abundance of two specific roof types, corresponds to high SV values.
- (iii) A high amount of infrastructure corresponds to high SV, although at low significance. This is somewhat surprising and most likely this variable provides a correction on the other variables, compensating some of the variation in the variables with a higher significance, such as buildings at landslide hazard. It may also be the case that the amount of service infrastructure and lifelines is higher around buildings exposed to landslide hazard because of planning considerations by the city developers in the past.
- (iv) No significant influence was found for the selected texture measures, the amount of buildings at flood risk and the distance to infrastructure (e.g. lifelines).

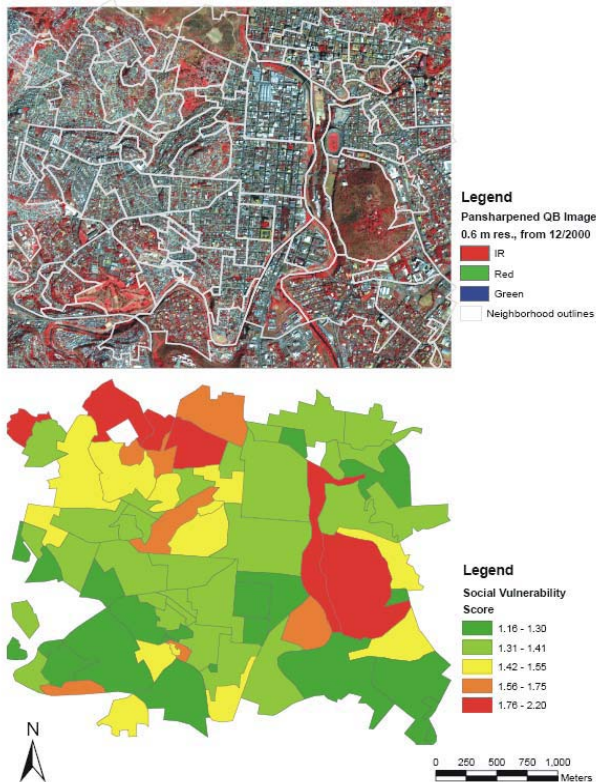


Figure 3. Scores from the social vulnerability index per neighborhood (bottom) with the pansharpened Quickbird image as reference (top).

## 6. VALIDATION, CONCLUSION, DISCUSSION AND OUTLOOK

Eight out of 47 variables explained almost 60 % of the variance of the SVI. While this is from the statistic perspective a satisfactory result, it has to be noted that the applied SVI was calculated based on a small amount of variables from census data not ideally suited for SV assessment. The data only include a limited amount of indicators relevant for SV assessment, and also allow validation only at the level of a census unit (neighbourhood). If the validation could have been performed

on a larger scale (e.g. using data from house-to-house surveys), the proxy variables could have been delineated on a smaller and more detailed scale as well. To validate the results obtained in this study and to quantify their contribution to risk assessment better, the validation could be improved by the comparison with results from detailed traditional methods, e.g. from house-to-house surveys in the same area. Unfortunately, they were not available for this research.

Nevertheless, the potential of the method becomes obvious: relevant information, consisting to a large proportion of valuable contextual information, can be delineated from remote sensing data in combination with other digital data sets (Table 1). The information that can be extracted ranges from building-specific information (location, roof material, etc.) to neighbourhood information (neighbourhood composition, living standards, etc.).

In terms of transferability, the delineated proxies can be adapted and evaluated to fit different places and different hazards. The main precondition of this approach is the initial understanding of SV in a particular place, mainly based on the specific local conditions and the understanding of those factors. For example, if it is not the case that poor people live in areas with unfavourable terrain, the consequent evaluation of this proxy variable can be changed accordingly or combined with other information such as neighbourhood composition.

For the assessment of physical vulnerability, it is important to refer to a certain natural hazard (e.g. flood, earthquake, landslide, storm, etc.) as physical characteristics of buildings imply different degrees of vulnerability to different hazards. As discussed in the literature (e.g. Coburn et al., 1994) indicators for SV do not necessarily relate to a certain hazard, but do express an intrinsic incapability and lack of resilience to cope with natural hazards in a more general way. This enables the application of this method for SV assessment for hazard-specific studies as well as for multi-hazard studies.

Some indicators, e.g. construction material, are used for both the assessment of physical and social vulnerability. On the one hand, it can be used to express the resilience or the lack thereof to withstand the power of a hazardous event (physical vulnerability). On the other hand, it is being used to describe indirectly the social standard and wealth of a household and thereby their vulnerability and coping capacities in case of a disastrous event (social vulnerability).

Another advantage is that medium scale approaches such as ours consider spatial relations of objects and their environment, and thereby provide a spatial framework in which information can be interpreted, whereas individual, very small scale information lacks this context and topology information. Thus, a combination of approaches at different levels is needed.

In terms of data interchangeability, previous studies showed that the use of lidar data for the extraction of building heights can be substituted by the use of stereo- or orthophotos (Fraser et al., 2002).

In this paper we present a new method for SV assessment based on contextual analysis of remote sensing and GIS data. An approach based on physical proxy variables and spatial metrics that were derived from high resolution optical and laser scanning data, in combination with elevation information and existing hazard data. Object-oriented image analysis was applied for the definition and estimation of proxy variables that describe the living environment and living standards, such as



