Remote sensing based assessment of social vulnerability

Annemarie Ebert, Norman Kerle & Alfred Stein
Hengelosestraat 99, P.O. Box 6, 7500 AA Enschede
The Netherlands, {ebert; kerle; stein}@itc.nl

Abstract
Risk assessment requires knowledge about the present and future hazards, elements at risk, and the different types of vulnerability. This paper deals with the assessment of social vulnerability (SV), which in the past has been frequently neglected due to lack of data and assessment difficulties. Approaches for SV assessment that do exist are primarily based on community-based methods or on census data and have limited efficiency and transferability, thus a new method based on contextual analysis of air-and spaceborne image and GIS data was developed, and tested for parts of Tegucigalpa, Honduras. We employed an approach based on physical proxy variables derived from high resolution optical, laser scanning and GIS data. Object-oriented image analysis was applied for the estimation of those variables, focusing on the ones with physical characteristics: (i) settlement structure, (ii) topographic location, (iii) commercial development, and (iv) distance to service infrastructure and lifelines. A reference Social Vulnerability Index (SVI) was created from census data available for the study area on a neighbourhood level. For the evaluation of the proxy-variables' utility for explaining SV a regression model was applied to identify those variables that best explained the changes in the SVI. Eight out of 47 variables were significant and could explain almost 60 % of the variation, whereby the slope position and the proportion of built-up area in a neighbourhood were found to be the most valuable proxies.

1 Introduction

The number of people affected by natural disasters and the damage caused by such events has been increasing globally. Disaster management tools are available to reduce risk and to prepare for a hazardous event. A sound strategy, however, rests on comprehensive knowledge of the present risk, which is commonly expressed as Risk = Hazard * Vulnerability * Elements at risk. The risk equation is area-specific, which emphasizes that the risk is dependent on local conditions (Figure 1). For the quantification of both hazards and elements at risk generally accepted methodologies have been developed (Wisner et al., 2004; van Westen et al., 2005). However, assessment of vulnerability, expressed as the degree of loss, has been more troublesome (Villagrán de León, 2006). Physical vulnerability can be evaluated based on physical principles such as building structure. Social vulnerability, however, relates to the living circumstances of people and to their individual characteristics that make them susceptible to the impact of a hazardous event. While this makes SV very personal and spatially variable, and thus extremely hard to quantify, it is critical to include it in the risk equation (Cutter et al., 2003; Birkmann, 2005).

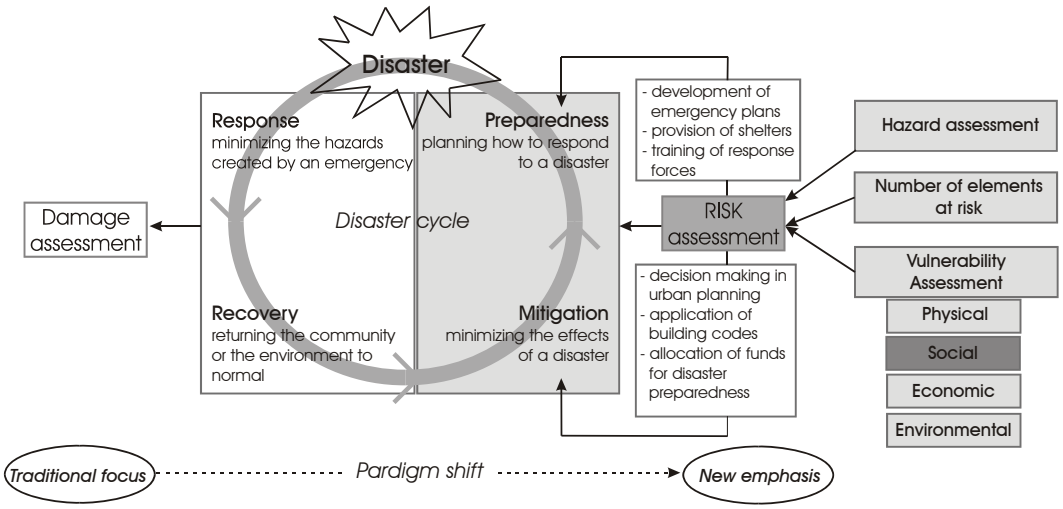

Figure 1. Disaster management cycle with the position of social vulnerability assessment.

Traditional methods to quantify SV for an area have been based either on statistical data, e.g. census-derived (Cutter et al., 2003), or on community-based approaches (Allen, 2006), though neither is well suited to capture SV in terms of its broad content and its dynamic nature. Here we investigate to what extent geoinformatics tools can also be of use in its assessment. We focus particularly on object-oriented analysis, which allows contextual and semantic image interpretation. Since SV is essentially a non-physical concept, we aim at identifying and quantifying physical proxies for different social aspects relevant to risk. We further propose a synergistic approach that maximises the use of remote sensing data whilst using supporting data from auxiliary map or field sources that cannot be extracted from image data.

2 Study area

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

3 Social vulnerability

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 censuses (Cutter et al., 2003; Moss et al., 2005). The former yields data at the level of an individual person with such information as age, gender, education, or disaster preparedness (Dwyer et al., 2004). Whilst being 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.

Census data cover larger areas and a wide range of variables, if available. The transferability of this approach is limited since census data are not always available for developing countries. 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. The possible repetition rate is determined by the frequency of the data collection (normally 5-10 years).

Besides a generally accepted definition a more efficient and comprehensive approach in terms of time and cost 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). Efficiency, transferability and repetition rate are important points that need to be considered, i.e. methods are needed that allow the assessment of SV in a sufficiently comprehensive way and that can be broadly applied in a sustainable fashion. A solution could be the combination of different traditional approaches with new methods, such as the analysis of remote sensing data.

4 Methodology: Object-oriented analysis of remote sensing data

So far, studies on vulnerability as one component of risk have mainly been focusing on the assessment of physical vulnerability (Müller et al., 2006). The concept of SV has so far been largely neglected in remote-sensing-based research because of its limited prominence within disaster research in general, but also because of limited technical means. The spatial resolution of image data has until recently been too low, and analysis has traditionally been on a pixel basis as opposed to an object-based approach a human analyst would follow in visual image interpretation. In this paper the utility of remote sensing data and analysis techniques for the assessment of SV indicators was explored by applying object-oriented image analysis on very high resolution satellite data (0.6 m ground resolution from Quickbird satellite). We chose object-oriented image analysis, implemented in eCognition, because it allows the consideration of

various digital data sets and generates important semantic information. This is a critical feature as the position of each image object in its natural and man-made environment partly determines its SV.

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 1). The evaluation for this study was based on the list of generally accepted criteria according to Cutter et al. (2003).

Table 1. Spatial expression of SV criteria

Criterion Spatial expression	**Spatial expression**
Socio-economic status	Indirectly through housing size & -condition
Commercial & industrial development	Possibly, with very high resolution (e.g. Quickbird or Ikonos
Residential property	Possibly, with very high resolution and clear physical outlines of property or additional cadastral data
Infrastructure and lifelines	Yes, with additional information
Rural/urban	Yes
Population growth	Possibly, using change detection
Gender, race and ethnicity, age, renters, employment loss, occupation, family structure, education, medical services, social dependence, special needs population	No

This study was a one-time assessment in which only urban areas were considered. Thus population growth and urban/rural distinction were not considered. However, substantial research has been done on tracking of population and urban growth with image data. The remaining broad indicators of SV that can possibly be mapped are:

- Socio-economic status;
- Commercial and industrial development;
- Residential property;
- Infrastructure/lifelines, and distance to those.

Table 2 shows how the original indicators for SV assessment were translated into proxy variables that can be delineated from remote sensing data and that express the local conditions in the study area. 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 socio-economic status of the people. However, field data would be necessary to associate actual roof materials with the roof types classified. The available service infrastructure was digitized from the 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).

Table 2. Physical proxies identified for the original SV indicators

Original indicator	**Parent proxy**	**Supporting proxy**
Socio-economic status	Settlement type	Proportion of built-up and vegetated area (4 proxies); Road conditions (1); Roof type (7); Available infrastructure (1); Texture (10)
	Topographic location	Slope position (12); Proportion of buildings in hazard zone (2)
Commercial and industrial development	Commercial development	Building heights (7)
Distance to lifelines	Distance to lifelines	Distance measures (3)

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 based on census data from 2000 that consist of seven variables (gender, literacy, wall material, roof material, water availability, waste disposal, building use) to calculate the SV for each neighbourhood. For the evaluation a stepwise regression model was applied. Input data for the statistical model were the 47 proxy variables as explanatory variables and a SV index (SVI) recently proposed by Haki et al. (2004).

Since knowledge about the importance of the variables was insufficient, only its sub-variables (such as male and female for gender) were weighted using PC (Saaty, 1980). To calculate the SVI for each neighborhood, the relative frequency of each sub-variable (e.g. relative amount of women per neighborhood) was multiplied with its weight derived from the PC. The resulting products of all sub-variables were summed up for each neighborhood using Equation 1.

$$SV = \sum_{i=1}^{m} v_i q_i \qquad \text{(Eq. 1)}$$

where SV is the vulnerability score for each neighbourhood, v_i the weight derived from the PC for each variable (values ranging from 0 to 1), and q_i the relative frequency of the variable per neighbourhood. Figure 2 shows the Quickbird image of the study area (A), and the vulnerability score SV as calculated from the census data for each neighborhood (B).

5 Results

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. A high proportion of built-up areas and of buildings on gentle slopes corresponds to low SV. 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. 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. 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).

For the lidar data the proportion of buildings with more than 6 floors, interpreted as an indicator for commercial development, was found to have a significant explanatory value (R^2=0.451), with higher proportions indicating lower SV.

6 Discussion

The physical proxies related to SV indicators explained close to 60% of the census-based vulnerability in the study area. This suggests that, while remote sensing data alone cannot replace traditional methods, in synergy with census data, field surveys and GIS information they can improve the assessment of SV in terms of efficiency, frequency and coverage of different scales. By extracting mappable factors in image data acquired at a higher frequency than field surveys or census data, the dynamics of SV can be addressed. The useful proxies identified and tested are the land use/land cover characteristics such as different roof types, the quality of roads and the abundance of vegetation, but also the slope angle, the distance to infrastructure and to the hazard zone, as well as building heights and the amount of service infrastructure per neighbourhood.

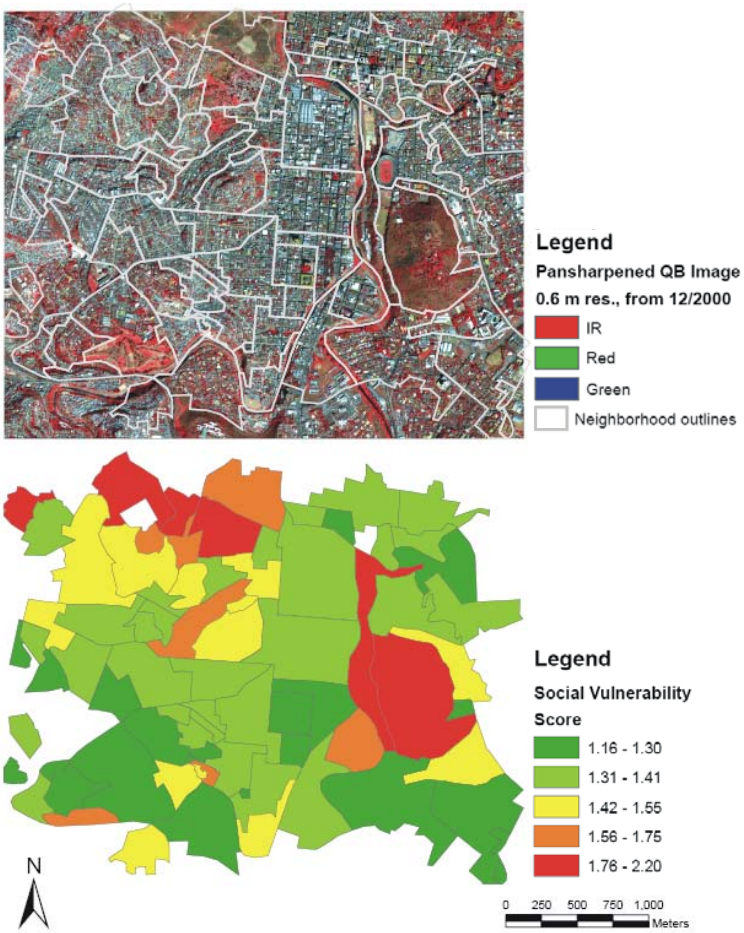

Figure 2. Satellite image of the study area and the results from the SVI per neighbourhood of the city.

Figure 3 illustrates the cost-benefit concept of an integrated approach. The use of satellite data (solid lines) is efficient compared to house-to-house surveys, comparatively easy to repeat and relatively low in cost per mapping unit, but by itself not detailed enough for a comprehensive assessment of SV. Most detailed but also most expensive and time-consuming are house-to-house surveys, while census data are most efficient but least detailed (both in dashed line). In general, the higher the level of detail the higher the costs and the higher the time needed for data processing. The gap between the dashed and the solid line shows the trade-off between costs, efficiency and the level of detail. The width of the cost-benefit area is mainly dependent on the accepted trade-off at the considered scale. In general, the higher the level of detail required, the higher the costs and the lower the efficiency.

While the approach presented here is not valid for all SV indicators (see Table 1), it is helpful to update risk maps as soon as new space-borne information becomes available. Especially in areas undergoing rapid changes, such as population growth, mapping of the ever changing settlement extent and composition as indicators of SV is a valuable contribution to risk assessment. The more accurate, current and detailed the resulting risk maps, the better the framework for decision making for all stakeholders working with risk-related topics.

The synergistic approach proposed is most suitable for studies on neighbourhood to regional level (see Figure 3). From the remote sensing perspective its applicability is mainly dependent on the considered scale and the size of the study area. If the scale is too large (i.e. individual people), satellite data will not be able to provide relevant information. However, with satellite data available today (resolution less than 1 m), relevant information can be extracted from the scale of individual houses to

entire regions. Even if census data as a validation tool are not available, the use of remote sensing data can provide valuable information as long as the proxy variables reflect the indicators used to assess SV as shown in Table 2. The method to extract the SV indicators can be transferred to other regions if the content of the main proxy variables (see Tables 1 and 2), for instance of socio-economic status, is understood and can be applied to the area of interest. While the general proxy variables are given by the definition of SV, local knowledge or a good understanding of the problem is needed to identify the supporting variables that represent the characteristics of the study area, which are thus site-specific, both in urban and rural areas. If high-resolution data are available it is recommended to use object-oriented image analysis to make use of critical contextual information.

The general challenge of this work is that SV is a largely social, intangible concept. As was shown in this study though, physical proxies for SV assessment do exist and can explain the concept to a substantial extent. These physical proxies still offer many prospects for future research. Further emphasis should be laid on the creation and analysis of spatial metrics (Herold et al., 2003). As noted above, more research is also needed to understand better the link between building type and the SV of its occupants. Lastly, to increase efficiency more work in spatial statistics and extrapolation techniques is needed to assess vulnerability reliably with the least amount of data collected in house-to-house surveys.

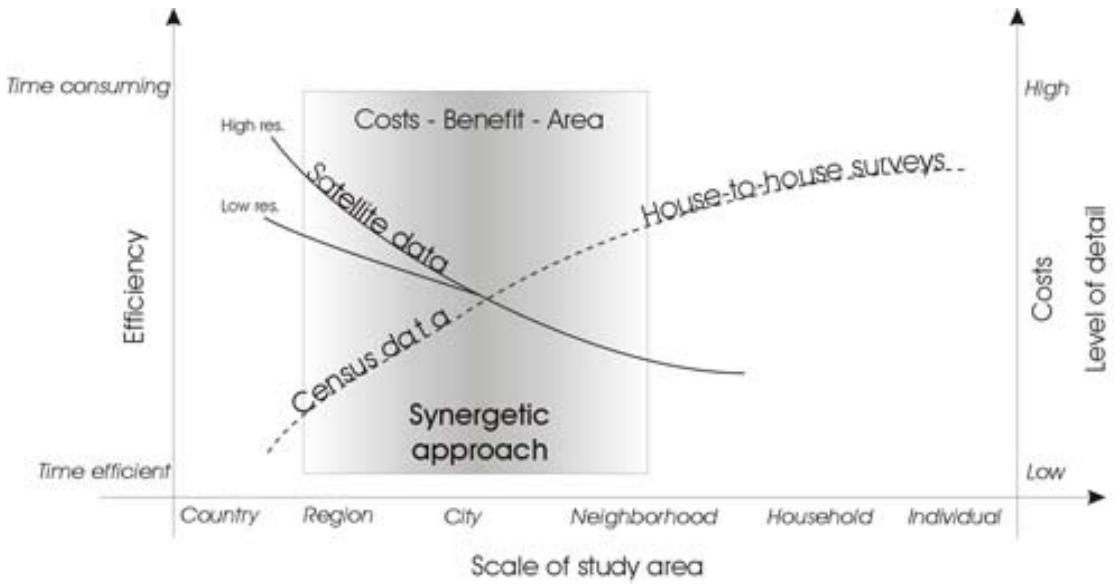

Figure 3. Cost-benefit area of an integrated approach for the assessment of social vulnerability implementing both traditional field surveys and remote sensing technology.

References

Allen, K.M., 2006. Community-based disaster preparedness and climate adaptation: local capacity-building in the Philippines. Disasters, 30, 81-101.

Angel, S., Bartley, K., Derr, M., et al., 2004. Rapid Urbanization in Tegucigalpa, Honduras. Preparing for the doubling of the City's Population in the next twenty-five years. Woodrow Wilson School of Public and International Affairs, Princeton University.

Birkmann, J., 2005. Measuring Vulnerability. Report on the 1st meeting of the expert working group "Measuring Vulnerability" of the United Nations University Institute for Environment and Human Security (UNU-EHS). Bonn.

Clark, G.E. et al., 1998. Assessing the vulnerability of coastal communities to extreme storms: the case of Revere, MA., USA. Mitigation and Adaptation Strategies for Global Change, 3(1): 59-82.

Cutter, S. L., Boruff, B. J., Shirley, W. L., 2003. Social Vulnerability to Environmental Hazards. Social Science Quarterly 82 (2) 242-260.

Dwyer, A., Zoppou, C., Nielsen, O., Day, S., Roberts, S., 2004. Quantifying Social Vulnerability: A methodology for identifying those at risk to natural hazards. Geoscience Australia.

Haki, Z., Akyuerek, Z., Duezguen, S., 2004. Assessment of Social Vulnerability Using Geographic Information Systems: Pendik, Istanbul Case Study. In Proceedings of the 7th AGILE Conference on Geographic Information Science. Heraklion, Greece.

Herold, M., Liu, X., Clarke, K. C., 2003. Spatial Metrics and Image Texture for Mapping Urban Land Use. Photogrammetric Engineering & Remote Sensing 69 (9) 991-1001.

Jain, S., 2005. System evolution using high resolution satellite data for urban regimes. Indian Institute of Technology Roorkee, Department of Architecture and Planning. PhD thesis.

Moss, R., Brenkert, A., Malone, E.L., 2000. Measuring Vulnerability. A trial indicator set. Pacific Northwest National Laboratory.

Mueller, M., Segl, K., Heiden, U., Kaufmann, H., 2006. Potential of High-Resolution Satellite Data in the Context of Vulnerability of Buildings. Natural Hazards 38 247-258.

Saaty, T. L., 1980. The analytic hierarchy process. McGraw-Hill International Book Company.

Tuceryan, M., Jain, A. K., 1998. The Handbook of Pattern Recognition and Computer Vision. Chen, C. H., Pau, L. F., Wang, P. S. P. (ed.), World Scientific Publishing Co. 207-248.

van Westen, C. J., Kumar Piya, B., Guragain, J., 2005, Geoinformation for urban risk assessment in developing countries; the SLARIM project. In Proceedings of the 1st International Symposium on Geo-Information for Disaster Management. Delft.

Villgrán de León, J. C., 2006. Vulnerability. A conceptual and methodological review. UNU-EHS. Bonn.

Wisner, B., Blaikie, P., Cannon, T., Davis, I., 2004, At Risk. Natural hazards, people's vulnerability and disasters. Routledge. London.