Poser, K., Kreibich, H., Dransch, D. (2008): Humans as sensors: assessing the quality of information from the public for rapid flood estimation - In: Pebesma, E., Bishr, M., Bartoschek, T. (Eds.), GI-Days 2008, Proceedings of the 6th Geographic Information Days, June 16–18, Münster, (IfGI-prints , 32), Institut für Geoinformatik, 117-122

Humans as Sensors: Assessing the quality of information from the public for rapid flood loss estimation

Kathrin Poser[1,2], Heidi Kreibich[1] & Doris Dransch[1]

[1]GeoForschungsZentrum Potsdam, Telegrafenberg, 14473 Potsdam
kathrin.poser@gfz-potsdam.de

[2]Center for Disaster Management and Risk Reduction Technology (CEDIM)

Abstract. Observations from the affected population can be an important source of information for disaster management and rapid loss estimation. New Internet technologies facilitate fast and easy data collection from the public. A major obstacle for using this information is its unknown quality. This research will develop methods to assess the quality of observations from the affected population for rapid loss estimation after flood events. It will yield an assessment of the fitness for use of flood observations from the public for empirical loss modeling and an automated procedure to evaluate the quality of these observations.

1 INTRODUCTION

For disaster management and rapid loss estimation, a broad overview of the situation is required. Information from many different sources can contribute to this overview, including different sensors, aerial and satellite images and observations of the emergency forces. Recent disasters, most notably hurricane Katrina, have shown that the affected population can be a very important source of information for disaster management.

1.1 Humans as sensors

Humans observe their environment and can act as intelligent sensors (Goodchild 2007). For flood disasters, they can observe different phenomena such as water levels, flow velocity or water contamination to supplement sensor networks, and even some parameters that cannot be measured by any sensors such as damage levels.

Until recently, collecting information from the public by telephone, mail, or personal meetings was tedious and costly. New Internet technologies have made this task easier and faster.

So far, however, information provided by the affected population has only rarely been taken into account systematically for disaster management

or event and damage assessment. A notable exception are the community Internet intensity maps that the US Geological Survey produces based on information collected from the public via a web interface (Wald and Dewey 2005). The main limitations of data collected from the affected population that impede their use are:

- Availability: It is unknown, how much and which and from where information will be supplied. Unlike a sensor network, information collection from the public cannot be planned in advance so as to yield an optimal configuration of observations for the phenomenon of interest.
- Data quality: Unlike physical sensors, humans cannot be calibrated and do not comply with standards. Mostly, they are not trained for specific observations. When affected by a disaster, humans can be very emotional which may impair their judgment.

To tap the potential of humans as sensors for disaster management and rapid loss estimation, particularly the question of the quality of such information is essential.

1.2 Aim of the Research

The overall aim of this research is to develop methods to systematically make observations of the affected population usable for disaster management and rapid loss estimation using floods as an example. The main questions to be addressed are:

- Which required information can be supplied by the affected population with sufficient quality? How can this information be collected?
- How can the quality of this information be assessed and controlled?

2 RESEARCH APPROACH

The research comprises two parts (see Figure 1). In the first part, the quality of observations from the public for flood events will be assessed using existing data from telephone interviews (Thieken et al. 2007). The results of this study will be used in the second part of the research to develop an automated procedure for assessment of observations from the public to be used in a prototypical implementation of web-based data collection for flood events.

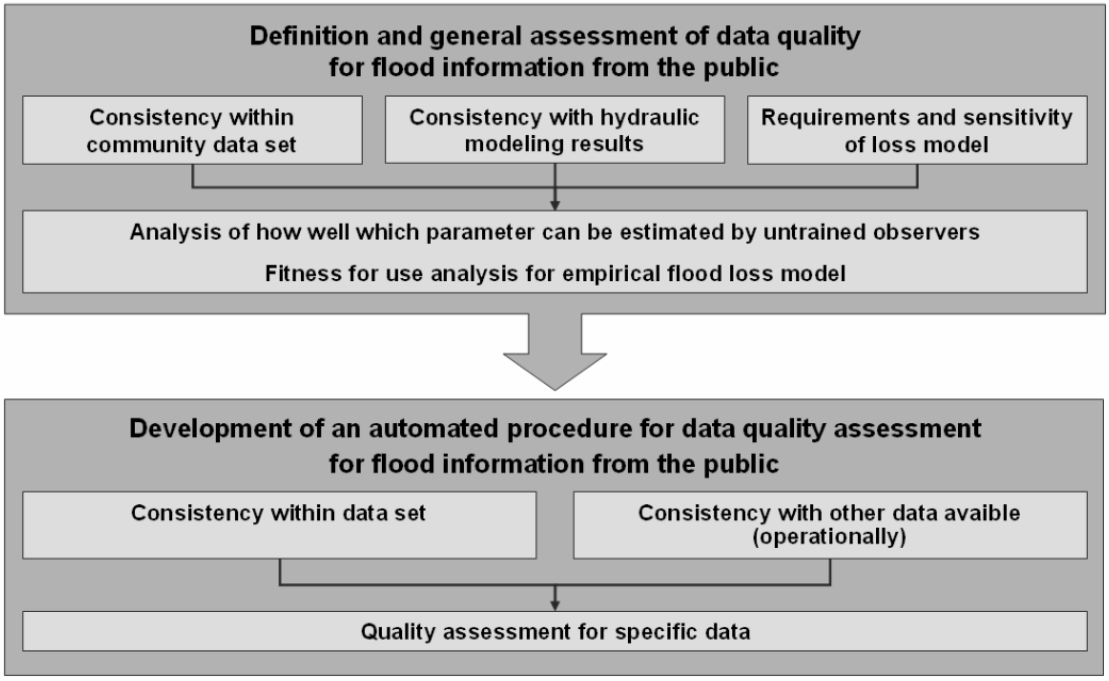

Figure 1: Overview of the research approach

Data quality has many aspects. The general definition of quality as "the totality of characteristics of a product that bear on its ability to satisfy stated and implied needs" (ISO 2002) needs to be broken down into more detailed quality elements in order to be useful. For spatial data, ISO 19113 proposes the following data quality elements (ISO 2002):

- Completeness
- Logical consistency
- Positional accuracy
- Temporal accuracy
- Thematic accuracy

These elements constitute the internal data quality, i.e. focus on the characteristics of the data independent of possible use. A complementary concept is external data quality or "fitness for use" which assesses the fitness of certain data for a specific task in a specific area (Devillers and Jeansoulin 2006)

In this study, the main focus will be on attribute accuracy for water level and flow velocity observations from the public as well as fitness for use of these data for empirical flood loss modeling using the model FLEMOps (Büchele et al. 2006, Thieken et al. submitted).

To assess the attribute accuracy of water level and flow velocity observations from the public, data from telephone interviews after the flood in the Elbe River catchment in 2002 will be used (Thieken et al. 2007). These observations will be compared with results of hydraulic models for selected sections of the Elbe River and its tributaries, with radar-derived inundation maps as well as with water marks at buildings (see Figure 2). As flow velocity is difficult to observe directly, several indicators such as a verbal description and the type of debris transported were used in the telephone interviews to get an estimate. These will be evaluated individually and in combination for their accuracy.

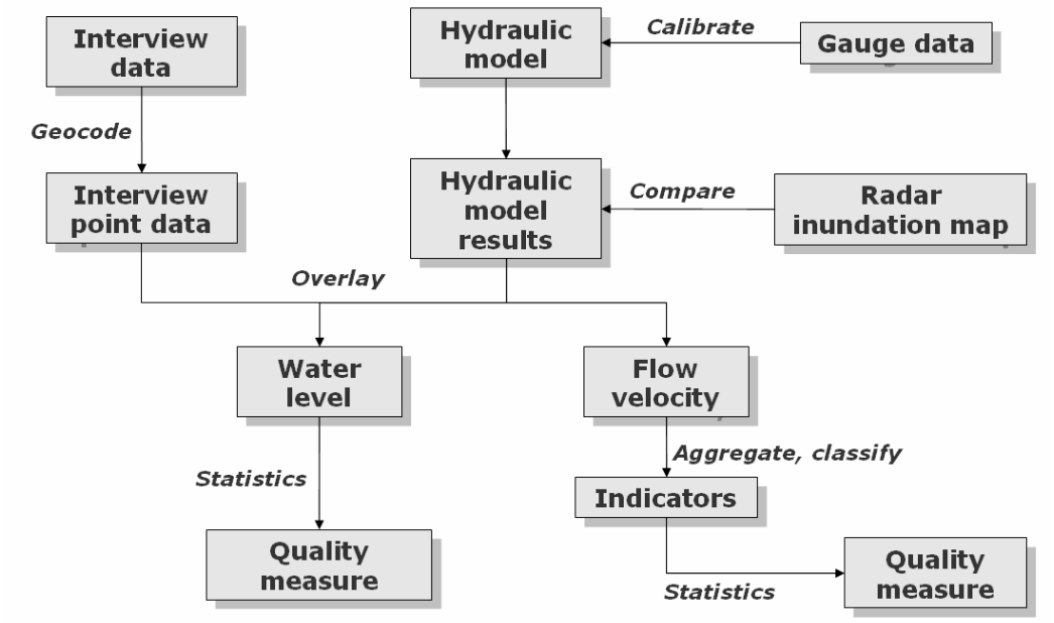

Figure 2: Workflow of the assessment of data quality using data from telephone interview after the flood in the Elbe catchment in 2002

After assessing the accuracy of the observations, the fitness for use of these data for empirical damage modeling will be assessed. For this, a sensitivity analysis of the damage model with respect to the input parameters that are observed by the public will be performed.

The results of the quality assessment of existing data will be used to develop an automated procedure for the assessment of information of the affected population after flood events to be used in realtime. The challenge is that only very little data is available operationally and in near-realtime to aid the assessment.

3 EXPECTED RESULTS

The research will yield the following results:

- A description of the quality of estimates of flood parameters by untrained observers
- An evaluation of the fitness for use of the information from the affected population for empirical flood loss modeling
- A method for automated integration and assessment of observations from the affected population together with other (sensor) data
- A prototypical implementation of such an automated method

4 DISCUSSION AND OUTLOOK

New web technologies enable the general public to create geographic information. In some areas, a great number of people are putting a lot of effort and enthusiasm into geographically referencing existing and creating new spatial information. For this information to be useful, in particular for scientific or operational use such as disaster management and event analysis, the assessment of its quality as will exemplarily be performed in this study is an important aspect. It is, however, also very important to understand what motivates people to contribute and how they can be motivated to join and to continue these efforts so that the concept of using humans as sensors will not be a hype, but a long-term trend.

5 REFERENCES

Büchele B., H. Kreibich, A. Kron, A. Thieken, J. Ihringer, P. Oberle, B. Merz and F. Nestmann (2006) Flood-risk mapping: contributions towards an enhanced assessment of extreme events and associated risks. NHESS - Natural Hazards and Earth System Sciences, 6: 485-503

Devillers, R. and R. Jeansoulin (2006). Spatial Data Quality: Concepts - Fundementals of Spatial Data Quality. R. Devillers and R. Jeansoulin. London, ISTE. 31-42.

Goodchild, M.F. (2007). Citizens as voluntary sensors: spatial data infrastructure in the world of Web 2.0. International Journal of Spatial Data Infrastructures Research 2: 24–32.

ISO (International Organization for Standardization) (2002). ISO 19113:2002 Geographic information – Quality principles.

Thieken, A.H., H. Kreibich, M. Müller and B. Merz (2007). Coping with floods: A survey among private households affected by the August 2002 flood in Germany. Hydrological Sciences Journal, 52(5) 1016-1037

Thieken A.H., H. Kreibich, T. Neltchinova and B. Merz (submitted) FLEMOps - A new model for the estimation of flood losses on the meso-scale. Journal of Hydrology.

Wald, D.J. and J.W. Dewey (2005). Did You Feel It? Citizens Contribute to Earthquake Science. US Geological Survey. Available at http://pubs.usgs.gov/fs/2005/3016/pdf/FS-2005-3016.pdf.