



Collaborative damage mapping for emergency response: the role of Cognitive Systems Engineering

N. Kerle¹ and R. R. Hoffman²

¹Faculty of Geo-Information Science and Earth Observation (ITC), Department of Earth Systems Analysis (ESA), University of Twente, Enschede, The Netherlands

²Institute for Human and Machine Cognition, Pensacola, FL., USA

Correspondence to: N. Kerle (n.kerle@utwente.nl)

Received: 4 May 2012 – Revised: 30 November 2012 – Accepted: 4 December 2012 – Published: 23 January 2013

Abstract. Remote sensing is increasingly used to assess disaster damage, traditionally by professional image analysts. A recent alternative is crowdsourcing by volunteers experienced in remote sensing, using internet-based mapping portals. We identify a range of problems in current approaches, including how volunteers can best be instructed for the task, ensuring that instructions are accurately understood and translate into valid results, or how the mapping scheme must be adapted for different map user needs. The volunteers, the mapping organizers, and the map users all perform complex cognitive tasks, yet little is known about the actual information needs of the users. We also identify problematic assumptions about the capabilities of the volunteers, principally related to the ability to perform the mapping, and to understand mapping instructions unambiguously. We propose that any robust scheme for collaborative damage mapping must rely on Cognitive Systems Engineering and its principal method, Cognitive Task Analysis (CTA), to understand the information and decision requirements of the map and image users, and how the volunteers can be optimally instructed and their mapping contributions merged into suitable map products. We recommend an iterative approach involving map users, remote sensing specialists, cognitive systems engineers and instructional designers, as well as experimental psychologists.

has included reconnaissance of the affected area with remote sensing data, in particular satellite imagery. Commercial satellite data have become indispensable in rapid post-disaster structural damage assessment (Zhang and Kerle, 2008). Until a few years ago access was virtually always restricted, even in the context of international disaster response. Fortunately, this has been changing in recent years, coinciding with increasingly routine post-disaster assistance based on remote sensing data, facilitated through the International Charter “Space and Major Disasters” (www.disasterscharter.org) (Ito, 2005; Stryker and Jones, 2009). In the context of Charter activations, ever growing numbers of satellite images are being acquired and damage maps generated for major disasters. This has led to increasing visibility of such support, as well as an expanding user base for the resulting maps. Traditionally, a small number of professional analysts at the Operational Satellite Applications Programme (UNOSAT) of the United Nations Institute for Training and Research (UNITAR), the Center for Satellite-based Crisis Information of the German Aerospace Center (DLR-ZKI), or the Service Régional de Traitement d’Image et de Télédétection (SERTIT; based at the University of Strasbourg, France) has been in charge of the image processing (Kerle, 2010).

Structural damage mapping based on image data has proved difficult despite advances in imaging technology in recent years (Kerle, 2010; Gerke and Kerle, 2011). At the same time, the number of remote sensing experts has been growing worldwide, as has the geoliteracy of lay people. This results from ubiquitous exposure to geodata and geotools (i.e., Google mapping tools, navigation systems, geotagging, etc.). Beginning in 2008, following Cyclone Nargis in Myanmar and the Wenchuan earthquake (China), first attempts

1 Introduction

Recent disasters, such as the 2010 earthquake in Haiti, have been met by rapid response assistance on part of the international community, both government and private sector. This

were made to apply collaborative mapping, a form of crowdsourcing (Goodchild and Glennon, 2010), in post-disaster situations. This has taken two principal forms: (i) mapping by a number of experts that share the burden of analyzing large data sets in a coordinated manner, and (ii) mapping efforts by thousands of largely laypersons in openly accessible mapping portals. However, given the inherent challenges of image-based damage mapping, as well as the high variability of structural damage (Kerle, 2013), there are a number of fundamental issues about how to best design a methodology for collaborative mapping in emergency response, which we address in this paper:

1. How can instructions to prospective damage mappers be conveyed in the most comprehensive and least ambiguous manner?
2. How can the organizations that coordinate the mapping know whether volunteers actually interpret instructions in the way they were intended?
3. How can it be insured that the instructions and examples translate into valid and useful mapping results?
4. How can the mapping contributions by many volunteers best be merged and evaluated?
5. What exactly is to be done with the geoinformation, and to what extent do different uses require adaptation in the mapping scheme?

Following a review of collaborative damage mapping, we address the five questions in order, but we converge on question 5.

2 State of the art in collaborative damage mapping

Geomapping activities depend crucially on an understanding of the goals of the individuals who have to make decisions using the geoproducts, their information and decision requirements that are entailed by the goals, and the goal-directed task activities that rely on the information and decision requirements. Cartographic theory has always emphasized how maps are made for particular purposes (Bertin, 1983) and, therefore, convey particular information. This is no less true for neogeography products (Turner, 2006) that are generated from modern remote sensing data or using novel collaborative mapping methods. They must be designed to allow the user to directly apprehend precisely the information that is crucial to their task and goals. This, however, does not appear to be the case for collaborative damage mapping. In fact we have insufficient knowledge about what crisis information the actual users need or understand, other than that they may not be especially trained map readers (Duchêne et al., 2011). Worse still, little effort is being made to understand the user better.

The principal source of the problem is the growing number of potential map users across many disciplines. For example, at the time of the 2010 Haiti disaster an estimated 10 000 non-governmental organizations were active in the country (Jobe, 2011). Together with the disaster relief agencies deployed after the earthquake, they all constitute potential users of geospatial information, including information about damage. Given that no clearly defined user exists, maps are currently distributed in PDF format, print-optimized for easy reproduction and dissemination among field staff. This approach ensures small file sizes and map context that does not change on different computer systems or can be easily modified by other users.

It also implies a one-fits-all approach for all prospective users. To our knowledge, the varied information needs of these stakeholders, including when and in what form they need it, have not been assessed, even though other researchers have previously identified this gap (Corbane et al., 2011a). Past research into the suitability of spatial disaster information has tended to focus on technical metrics. For example, Shimoni et al. (2011) assessed five different product types for flood disasters (e.g., flood water extent) generated as part of an international crisis response exercise, yet only focused on thematic and geometric accuracy aspects; to what extent user needs were met, or the map information was understood correctly, received no attention. Corbane et al. (2011a) developed an elaborate validation protocol for damage maps, but also acknowledged that generally user needs are poorly defined. With respect to structural damage mapping the trend has been towards an increasing number of map makers and map styles, rather than on creating products that suit specific needs, or that were designed together with the envisioned user (Fig. 1; Kerle, 2013).

The collaborative mapping that leads to damage map products is a process involving many people who are attempting to plan, make decisions, and take actions relying on a considerable amount of technology, including computers and communications systems. Collaborative mapping is complex for these reasons, and also because the situations that are of concern in the mapping are themselves complex, unfolding events that are rife with uncertainties and unpredictabilities. While the need to deal with the latter is a well-known requirement of the disaster responders (the map users), also the collaborative mappers need to respond flexibly to new data becoming available, new mapping tasks being added, or new information from the disaster scene becoming available.

Since its first release in 2005, Google Earth, like similar virtual globes, has had a multi-layer and flexible nature. Using Google's Application Programmable Interface, Google map information can be readily integrated into existing web sites, while also allowing customized map layers to be added to Google Earth. This has made it a tool of choice for data sharing, damage visualization etc. (Tomaszewski, 2011), and explains why Google Earth was selected as a dissemination vehicle by satellite operators willing to release post-disaster

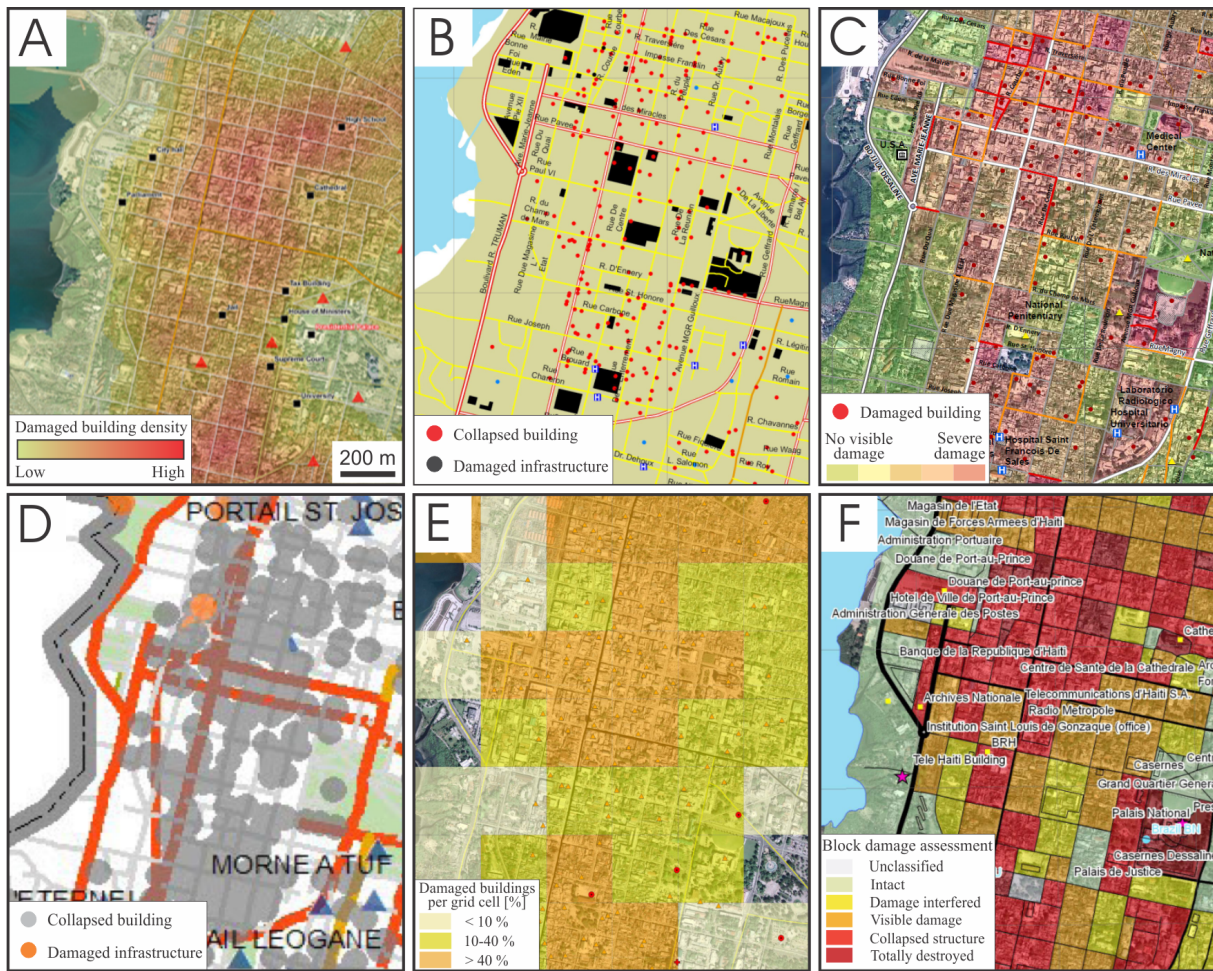


Fig. 1. Example damage maps of parts of Port-au-Prince (Haiti) following the 2010 earthquake, prepared by Service Régional de Traitement d’Image et de Télédétection (SERTIT; **A**), Information Technology for Humanitarian Assistance, Cooperation and Action (ITHACA; **B**), United Nations Cartographic Section (UNCS; **C**), Information Management & Mine Action Programs (iMMP; **D**), Crisis Information Center of the German Aerospace Center (DLR-ZKI; **E**) and the alliance of the Italian Space Agency and Telespazio e-GEOS (**F**). Legends adapted from original sources to show only elements related to structural damage. Note that map scales vary slightly.

imagery. This process has been actively fostered by Google, and in particular its philanthropic arm, Google.org, whose Crisis Response project has been providing support and coordination following major disasters and other humanitarian crises ever since the 2005 Hurricane Katrina event (van Aardt et al., 2011). Given Google Earth’s ubiquitous nature and flexibility, it has broadcast post-disaster satellite imagery and has invited efforts to add value to those data on the part of concerned individuals and organizations not officially involved in damage mapping. This has led to platforms that are open to any interested mapper, such as the proprietary Google Map Maker (GMM), or alternatives operating under open content license agreements, such as OpenStreetMap (OSM). Both tools are mainly used for rapid generation of base data, such as road infrastructure, and gazetteering.

In addition to collaborative mapping platforms that are open to all, such as GMM, there also exist closed systems

limited to registered image analysis experts. Following the 2008 Wenchuan earthquake, ImageCat (www.imagecatinc.com) established the Virtual Disaster Viewer (VDV; vdu.mceer.buffalo.edu, based on Microsoft Bing imagery), where 85 volunteers recruited in remote sensing image analysis circles mapped structural damage based on pre- and post-disaster high spatial resolution satellite imagery (Barrington et al., 2011). Detailed instructions on what and how to map were provided to the volunteers by ImageCat.

A similar effort was launched following the 2010 Haiti earthquake (Ghosh et al., 2011). Initially also based on the proprietary VDV platform, mapping was subsequently moved to Google Earth. More than 600 volunteers with remote sensing expertise and representing 131 private, governmental and academic institutions in 23 different countries participated in what became known as the Global Earth Observation-Catastrophe Assessment Network (GEO-CAN;



Fig. 2. Neighborhood in Port-au-Prince (Haiti) severely affected by the 12 January 2010 earthquake. Pre-event Geoeye-1 satellite image (26 August 2009; **A**), first post-event satellite scene (Geoeye-1, 13 January 2010; **B**), and airborne image acquired on 25 January 2010 (**C**). ©Google Earth.

Corbane et al., 2011b; Ghosh et al., 2011). Mapping was first carried out based on Geoeye-1 satellite data with 50 cm spatial resolution (down-sampled from the original 41 cm due to US government regulations), and later repeated on 15 cm aerial imagery (Fig. 2). GEO-CAN worked with a hybrid approach of an essentially open platform used by a closed community. The more recently established Tomnod platform for collaborative image-based mapping (tomnod.com) follows a similar approach. While mapping on open platforms such as GMM and OSM is largely restricted to base mapping and is controlled or guided via editorial and review processes, mapping on closed platforms or by a closed community is based on detailed instructions to the mappers, comprising of text-based descriptions of damage to be mapped, and supported by illustrations and example screen captures (see section on question 1).

Such collaborative approaches assume that all volunteers possess the perceptual and damage recognition skills required, which may be a misjudgement. With increasing sophistication of air- and spaceborne remote sensing systems the quality of data, especially in terms of spatial resolution, has been increasing. The most recent generation of civilian Earth observation satellites, Geoeye-1, has reached a resolution of 0.41/1.65 m (panchromatic/multispectral), while aerial imagery, such as of the type acquired after the Haiti disaster, easily reaches 15 cm. This increase in resolution implies that the mapping results based on those data must be gradually improving, but this has only materialized to a certain extent (Kerle, 2013). It was found in Haiti that the amount of building damage mapped in the 15 cm aerial imagery was approximately 10 times greater than what had been identified in the 50 cm satellite data (Lemoine, 2010),

demonstrating the value of greater detail. At the same time, work by Cambridge Architectural Research Ltd. on airborne Pictometry data (imagery that includes four oblique views and a vertical image) acquired shortly after the Haiti earthquake showed that despite the excellent data quality only about 63 % of the buildings mapped as heavily damaged or completely destroyed on the ground were identified as such in the visual analysis of the Pictometry imagery (K. Saito, personnel communication, 2011; Ghosh et al., 2011).

Figure 3 illustrates how the type of damage can lead to different scores on the ground, in Pictometry and based on satellite data. Complete destruction (D5 on the European Macroseismic Scale of 1998, EMS98; Grünthal, 1998) can often be readily identified in vertical satellite (a) and airborne Pictometry imagery (b), and confirmed as such on the ground (c). However, buildings suffering floor collapse while retaining an intact roof are often misclassified. The building in (d) showed no damage in the satellite image (damage score of 0), but received a D5 both in the Pictometry assessment (e) and in the field (f). The building in (g) also shows no visible damage in the image, and only minor cracks on the ground (i). However, it also illustrates that where site access is restricted no complete assessment (e.g., in this case to the ground floor) is possible. Oblique Pictometry data are particularly useful in such a case (h). Whatever limitations professional damage mappers face naturally also apply to volunteers. The Haiti damage mapping, including the GEO-CAN effort, the data validation work, and field mapping, were recently described by Ghosh et al. (2011).

A number of additional post-disaster response efforts based on crowdsourcing principles have been established. Several of them, such as Ushahidi (e.g., Roche et al., 2011),

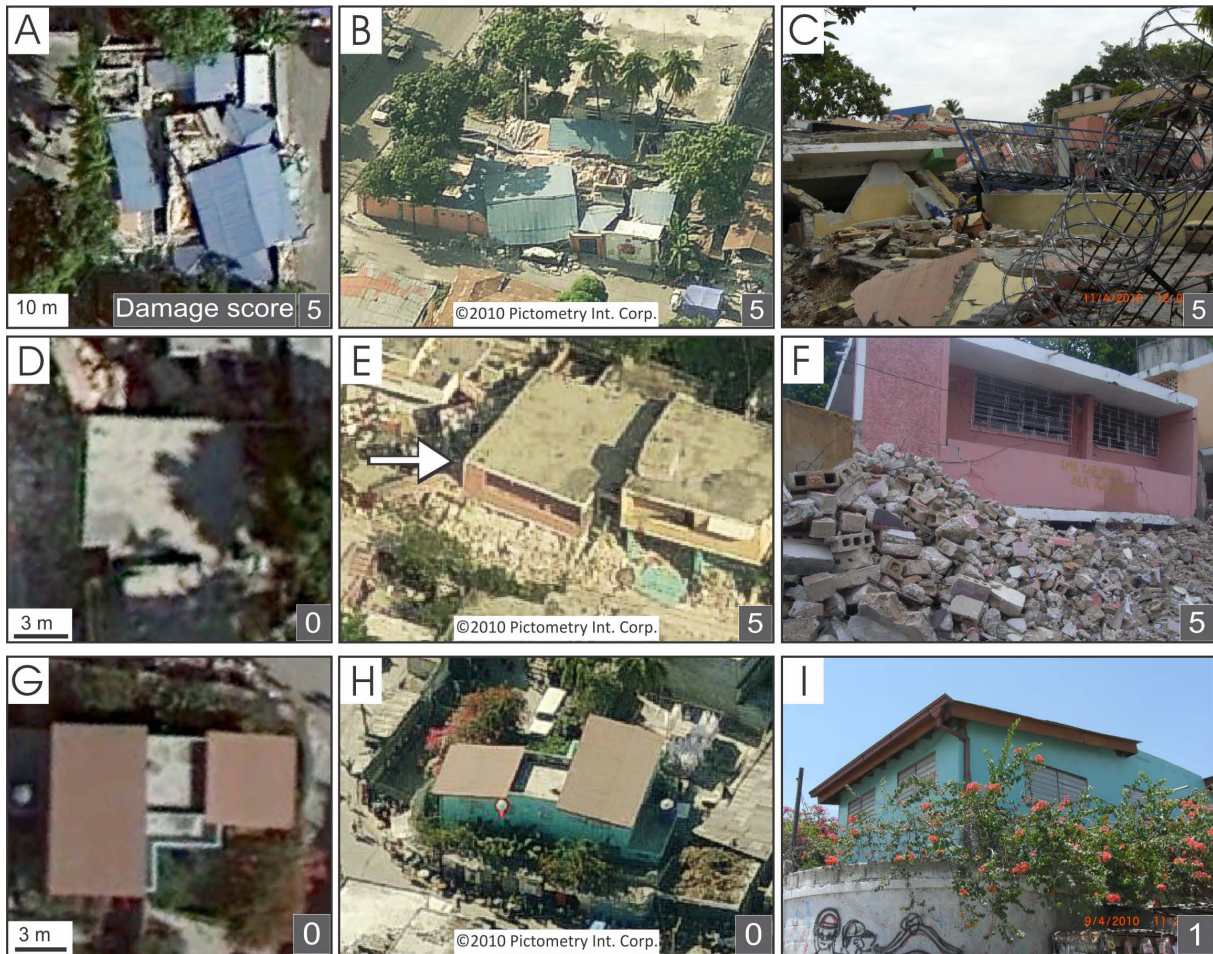


Fig. 3. Discrepancies in damage scores for image- and ground-based assessment as per EMS98 ranking (D0 = no damage, to D5 = complete destruction). Building in (D) corresponds to pancake collapsed structure indicated by the arrow in (E). See text for details. ©Google Earth. (B), (E) and (H) ©Pictometry International Corp. Field photos were provided by the UK-based Earthquake Engineering Field Investigation Team (EEFIT) and Cambridge Architectural Research Ltd.

or the USGS’s *Did you feel it?*, are also using cartographic tools, though only to visualize volunteered contributions such as incident reports.

3 Challenges in collaborative mapping

Based on the challenges and potential shortcomings in current collaborative damage mapping approaches we address the 5 research questions stated earlier.

3.1 How can mapping instructions be conveyed in the most comprehensive and least ambiguous manner?

The degree of difficulty of collaborative mapping for emerging and emergency situations (e.g., in Haiti) depends on the actual mapping objectives. Comparatively simple and unambiguous tasks, such as image-based road mapping, one of the central aims of GMM and OSM, have proved rapid, compre-

hensive and robust, indicating that they require a lower degree of analytical ability than damage mapping. All edits are checked and approved in a review and moderation process. Independent assessments have found results to be very accurate, for example by comparison with maps published by national mapping organizations (Haklay et al., 2010; Haklay, 2010), though room for improvement has also been noted (Girres and Touya, 2010). Problems have also been found in the georeferencing accuracy of the source satellite data (Goodchild, 2007), which then propagate into mapping results.

However, structural damage mapping is an entirely different challenge. Numerous studies have explored the utility of image data (optical, radar, and airborne laser scanning [LiDAR] data, and both vertical and oblique imagery) (e.g., Kerle et al., 2008; Dell’Acqua and Gamba, 2012; Arciniegas et al., 2007). Many strategies to automate damage detection have been tested, but complex pattern recognition problems

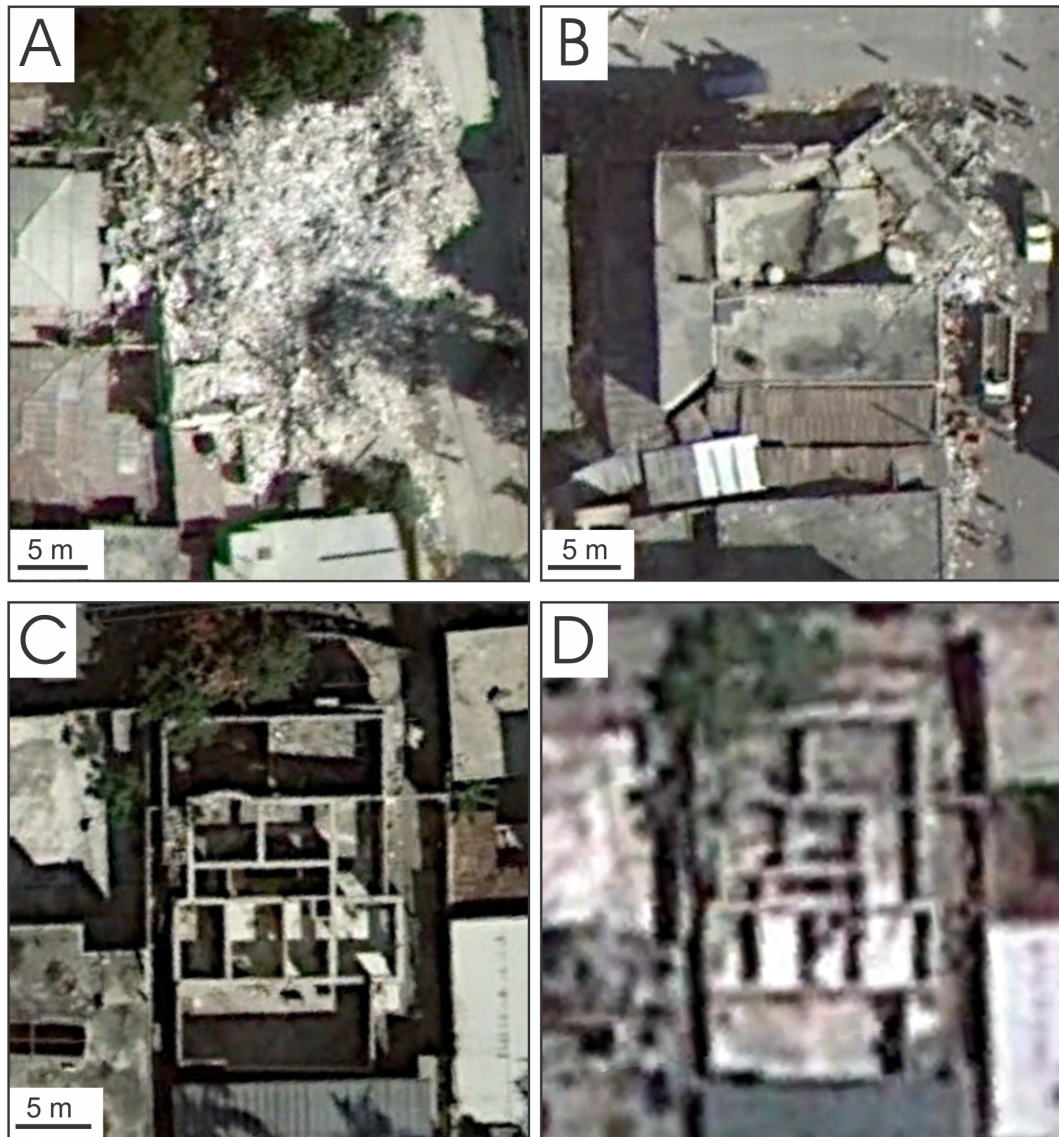


Fig. 4. Examples of different types of structural damage and their appearance in vertical image data in Port-au-Prince, Haiti. Rubble pile after disintegration of an adobe building (A), and severe structural damage of a reinforced concrete structure (B; aerial photographs from 25 January 2010). Both are examples of D5 damage (total destruction) on the EMS98 scale. Many structures in Port-au-Prince were already strongly decayed by the time of the earthquake, or had never been finished, which is easily mistaken for damage. The building in (C) and (D) (pre- and post-disaster Geoeye-1 satellite images of 26 August 2009 and 13 January 2010, respectively) is an example. ©Google Earth.

remain at the edge of what is possible via computational analysis. Reflecting the lack of a robust and flexible method, virtually all operational Charter-based mapping of structural damage continues to be based on manual, visual image interpretation (Voigt et al., 2011). This is precisely because of the difficulty of translating what is essentially a concept (“damage”) into (i) identifiable, physical parameters (Gerke and Kerle, 2011) and (ii) operational definitions or expressions of spatial elements or characteristics to recognize. Both must also be understandable for remote sensing professionals without specific damage mapping expertise.

The GEO-CAN mapping was aimed at very heavy damage (D4) and complete destruction (D5), respectively. Gerke and Kerle (2011) discussed the ambiguities of the EMS98 classification. In particular, individual damage elements (e.g., roofs or facades) do not add linearly to a given damage scale, and damage classifications vary for different building types, such as masonry or reinforced concrete. The damage in Port-au-Prince proved especially challenging given the great amount of decayed and unfinished buildings that shared similarities with damaged building stock (see Fig. 4c and d, as well as Castoldi et al., 2011). The EMS98 scale was also

designed for in situ damage assessment by structural engineers, not laypersons or remote sensing experts.

It is clear that an evaluation that includes a detailed view of different façades of a building, and possibly the interior, even by a layperson observing such a building on the ground, cannot be matched by an assessment based on largely vertical, single-view imagery that shows fewer perspectives and considerably less detail. This poses a considerable challenge for collaborative image-based damage mapping. Additionally, damage has a number of facets: it can be seen in structural or in economical terms, from a survivability perspective (as a function of structural integrity), or from a perspective of use and function. For example, a hospital only needs to sustain moderate structural damage for its utility to be compromised. These points need to be considered in collaborative damage mapping, and mappers provided with suitably comprehensive and unambiguous instructions. At the same time, these instructions are related to one specific definition of damage, typically physical destruction, and do not communicate anything about the actual information needs (or damage definitions) of the map user. No approaches to generate image-based damage information to satisfy such variation in damage definitions have been developed. All maps in Fig. 1 only show physical damage, thus are largely different visualizations of the same phenomenon.

From a cognition perspective it also remains to be assessed to what extent the difficulty in collaborative mapping is amplified by either the mapping software, or the communication tools used to convey mapping instructions to the volunteer. Devising damage mapping instructions is a challenge due to the complexity of the problem. Even with mapping focusing only on D4 and D5 damage (i.e., severe damage to total destruction) the way damage is expressed in imagery is highly variable. Some indicators are directly visible, such as clearly defined rubble piles (Fig. 4a) or large cracks or deformations in roofs (Fig. 4b) where the pre-event scene showed an intact structure. At other times proxies have to be used. For example, missing or smaller shadows compared to the pre-event situation, or different roof offsets between adjacent buildings (oval in Fig. 5b), indicate pancake collapse (intact roof but collapsed floors). However, sun angle differences between image acquisitions, and resulting shadow changes, can also cause confusion (compare shadows in a and b). The lack of shadow along the building in (b) can be mistaken for collapse of the upper part of the building. The roof of the building in (c) appears largely intact. However, there is evidence of floor collapse, with the overall roof angle changing with respect to the outer wall (compare c and d), and extensive blow-out debris on both sides (arrows). All of these indicators need to be evaluated:

- i. In imagery that may have poor quality (e.g., haze, dust);
- ii. where resolution, viewing angle or sun position (and hence building shadows) differ between pre- and post-event image (Fig. 5a and b);
- iii. where buildings are closely clustered or roofs overlap, effectively preventing individual structures from being identified, and;
- iv. where changes such as building construction or demolition have taken place between image acquisition dates.

Instructions prepared for the Wenchuan and Haiti collaborative damage mapping tried to address this problem. The Virtual Disaster Viewer (VDV) guidelines for Wenchuan are a 12 page document that in 23 steps explains the concept and layout of the VDV, the registration procedure and navigation tools, an ideal mapping strategy to ensure completeness, as well as detailed descriptions of the different damage categories. Condensing the variety of direct and indirect damage expressions in a mapping instruction manual clearly poses an enormous challenge, raising the question of how accurate the resulting volunteer contributions can be.

3.2 Will the instructions and examples actually result in valid and useful mapping results?

According to the current collaborative mapping perspective, being able to provide a usable and useful mapping contribution has a number of preconditions:

- i. The instructions provided to the mapper need to unambiguously define and characterize the features to be mapped, be comprehensive enough to cover the entire spectrum of the feature's appearance in the data used, and provide a clear definition in case of subcategories (e.g., very heavy damage vs. complete destruction);
- ii. the volunteer has sufficient domain knowledge and technical skills to carry out the mapping;
- iii. the method to be used for the mapping must be made clear, including use of symbology, color codes and attributes, and;
- iv. the software environment must be easily learnable and highly usable, and needs to support an optimal interaction with the data (including easy use of multiple images if pre- and post-event data are available) and facilitate easy yet accurate mapping.

The validity of all of these assumptions needs to be assessed.

Assumption (i) constitutes the immediate challenge. As described above and by Gerke and Kerle (2011), image-based mapping of structural damage means assigning a single and meaningful label to a structural element (building, bridge, road etc.) that may be in a complex state. The mapper is provided with text descriptions and example imagery that are meant to be comprehensive and unambiguous (e.g., Fig. 5). However, ambiguity does not result from the number or choice of illustrations, but rather emerges in the interaction of instructions with what the reader knows and their context. It follows that ambiguity cannot be reduced simply

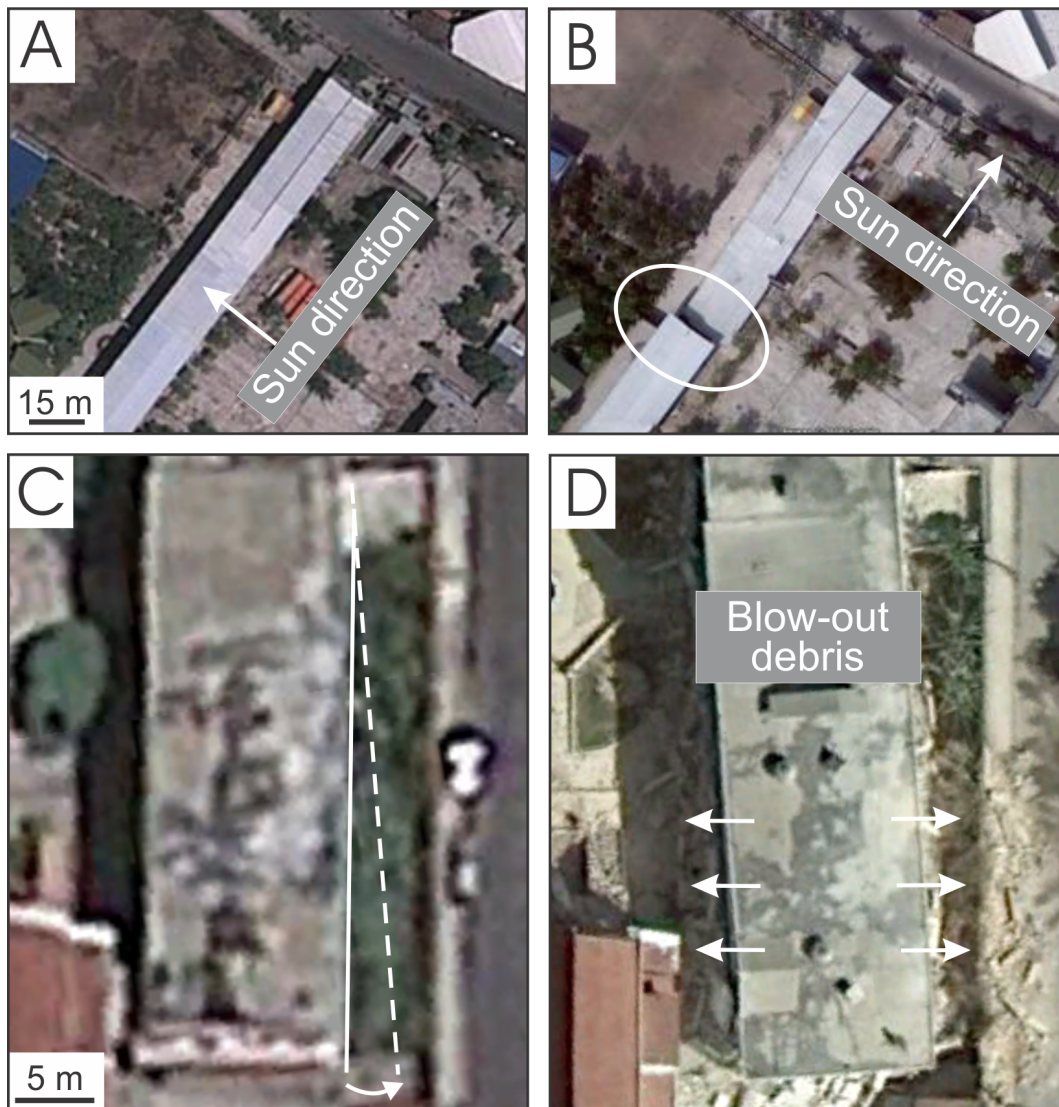


Fig. 5. Use of proxies in structural building damage assessment, such as shadow (A and B), and building offset and blow-out debris (C and D). See text for discussion. ©Google Earth.

by expanding the instruction manual. Rather, one must let the mapper practice the task and receive corrective feedback.

This entails an additional challenge, since in the post-disaster phase where the mapping takes place, the organizers typically possess no field information as a basis for validation and feedback. Also, the aim at comprehensive illustration of all damage types and forms of appearance may be counterproductive. While research in psychology has shown that learning to perceive complex patterns depends on having exposure to instances that span the range of critical variables, and instances that are “tough cases” as opposed to “clear cases” (Hoffman and Fiore, 2007; Klein and Hoffman, 1993), for initial learning of a complex task, comprehensiveness is not recommended. Insights from instructional design

and learning theory should be used to improve mapping instructions.

Assumption (ii) was indirectly addressed by GEO-CAN by prompting the volunteer to state their expertise in the registration process. According to ImageCat, volunteers were required to possess a MSc degree or equivalent, and at least 4–5 yr of remote sensing experience, information on which was to be provided during the registration. However, it is unclear if mapping results were tagged with some form of reliability attribute as a function of different experience levels, or whether results from less experienced volunteer factions received special validation. Ghosh et al. (2011) reported that experience level was matched against degree of complexity and difficulty of the tiles assigned to the volunteer. More fundamentally, however, it is not clear if or how general remote

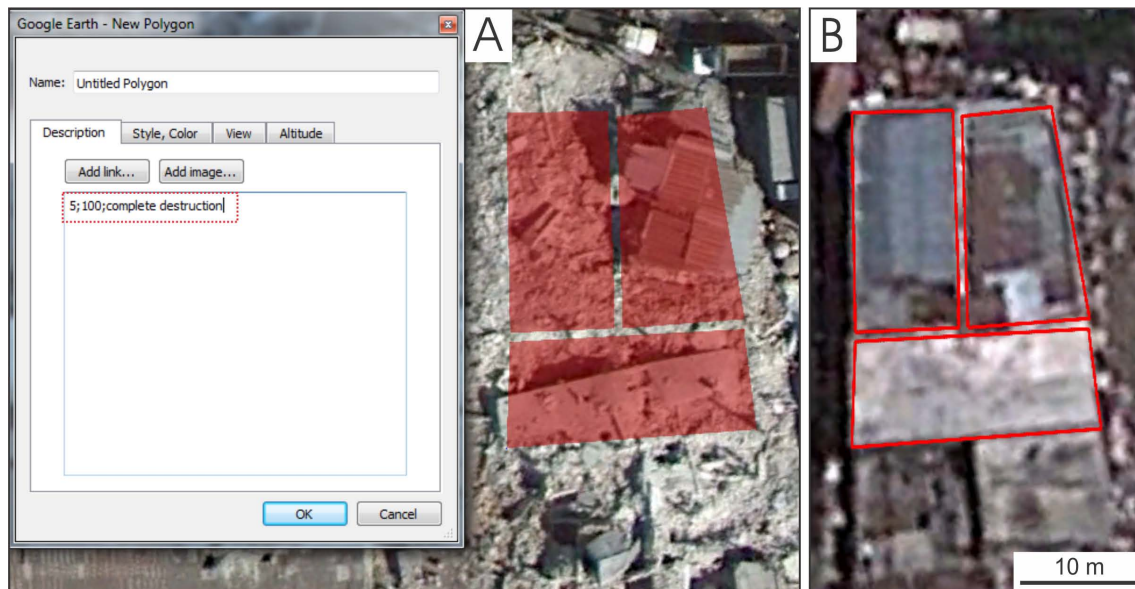


Fig. 6. Delineation of damaged buildings in Google Earth, and creation of kmz (Keyhole Markup Language) files that were later integrated by GEO-CAN. Identification of individual buildings, especially those with severe damage, is nearly impossible in post-event data alone (A); reference imagery (B) is needed. Volunteer mappers had to specify the damage level from 1 to 5 according to EMS98, confidence in the given score (in percent), and a comment such as on the damage level (dashed box). Screen captures from Google Earth. ©Google Earth.

sensing experience correlates with an ability to map structural damage accurately. Willingness to help and volunteer is strong following exceptional disasters, but it has been shown that volunteers are often not adequately skilled for the work they wish to support (for an account of inadequately prepared medical staff volunteering after the Haiti earthquake see Jobe, 2011). It is possible that collaborative damage mapping suffers from comparable limitations.

Assumptions (iii) and (iv) relate to the mapping implementation and the software environment. The VDV has similarities to platforms such as Google Earth but considerably less functionality. It does allow different image layers to be switched on and off (which is critical for comparison between pre- and post-event images), and map entries to be modified or deleted. This is important because mappers need to be able to correct their work; knowing that an entry is permanent can have a paralyzing effect on the mapper. In a sense, the capabilities of the software are “instructions” to the mapper (Norman, 1998). We argue that CTA would be needed to assess to what extent the instructions for how to map in VDV or Google Maps were properly designed, and matched the actual tasks that mappers had to perform. The ubiquity and broad user base, in particular of Google Earth, imply a self-explanatory nature. However, tools are very rarely self-explanatory, even when (experienced) users claim that they are (see Norman, 1998).

The first phase of the collaborative mapping following the Haiti earthquake was also carried out in the VDV. Within 48 h more than 133 km² were mapped based on 50 cm Geoeye-1 imagery, again using point signatures to map building

damage. For the second phase, using 15 cm aerial imagery and eventually covering more than 1000 km², Google Earth was used. Unlike in phase 1, buildings to be classified first had to be polygonized, and the resulting kmz (Google’s Keyhole Markup Language) files of building polygons with damage attribute sent to ImageCat for integration (Ghosh et al., 2011). More than 600 registered volunteers mapped 500 × 500 m grid cells, and received a document explaining the functionality of Google Earth, in particular the creation of polygons to be saved as kmz files. As for Wenchuan, a Damage Level Interpretation Guide was also provided to assist the volunteer, consisting of screen captures and a text-based characterization meant to be comprehensive and representative. With more than 1 billion users worldwide, Google Earth was likely a more familiar platform for most mappers and remote sensing specialists than the VDV. However, the technical challenge was greater. This was due to the need for the mapper to create actual building outlines (in very difficult urban terrain), the need for the mapper to discriminate buildings already decayed prior to the earthquake, and the effort of creating new map layers and computer files. The latter is a task that is imposed by the software, separate from the actual semantics of the mapping task, and adds another potential error source. Figure 6 illustrates both the challenge of identifying and outlining a destroyed building (red polygon), and the metadata to be provided for each polygon. Phase 2 also included an assessment of building floor numbers and total floor space (Ghosh et al., 2011), adding further complexity.

Both the VDV and the Google Earth damage mapping environments should be subjected to a usability evaluation,

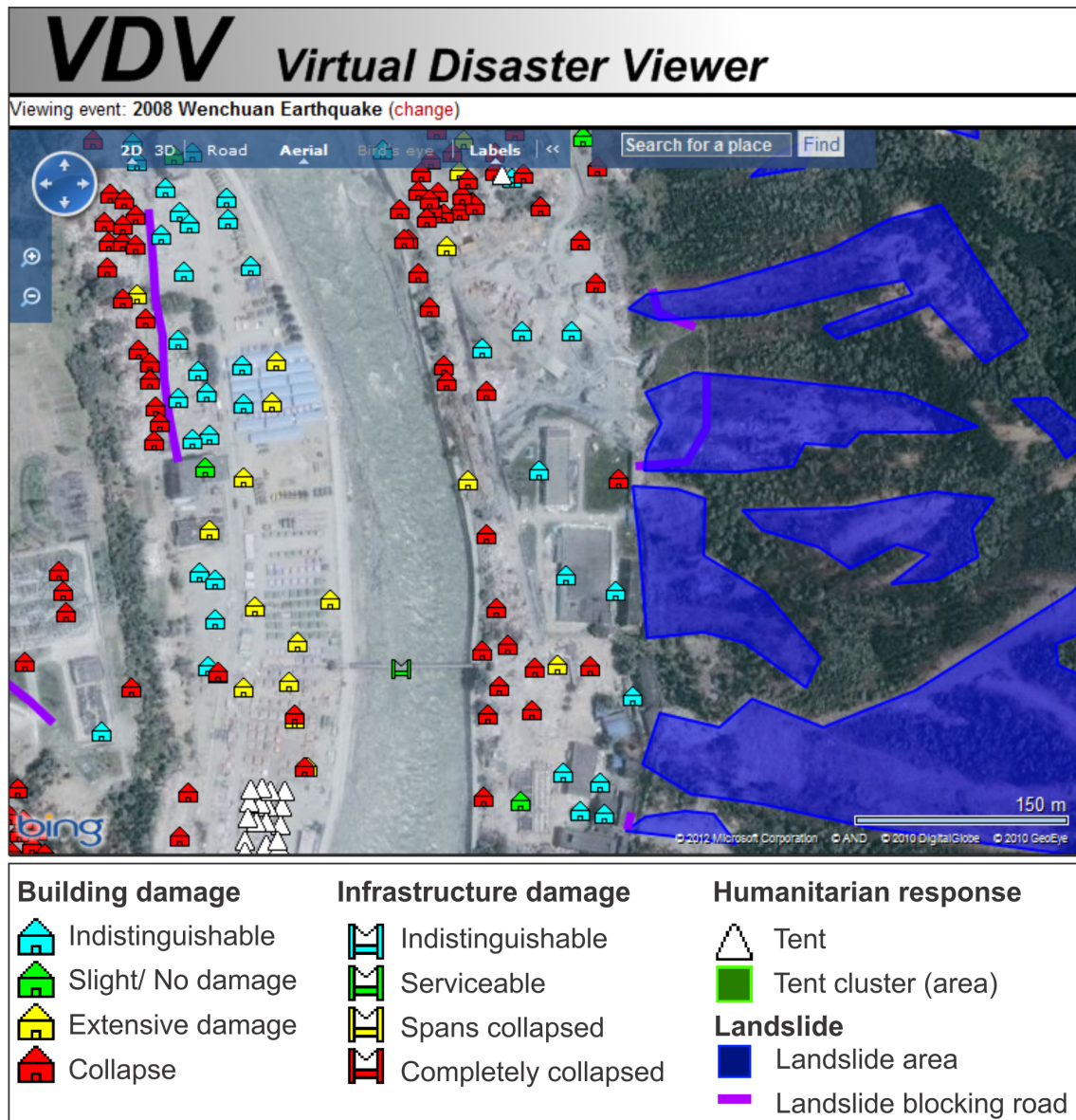


Fig. 7. ImageCat's Virtual Disaster Viewer, showing damage mapped following the 2008 Wenchuan (China) earthquake. Damage assessment was more difficult than in Haiti, given (i) the relatively low image detail (QuickBird with 2.4 m multispectral resolution), (ii) low image contrast and haze, and (iii) the larger number of damage types. For example, accurate detection of serviceable bridges or roads blocked by landslides requires additional distinct image analysis expertise.

such as the Cognitive Walkthrough as described by Polson et al. (1992) or Allendoerfer et al. (2005) to measure their performance, and to understand better what difficulties volunteers have with either working environment, or whether they are indeed used correctly.

Previous collaborative damage-mapping efforts have experimented with different types of damage demarcation, different degrees of mapping complexity, and different kinds of instructions. In the VDV-based Wenchuan case, mappers had to place symbols of different damage degrees on individual buildings, in addition to flagging infrastructure damage,

sites of humanitarian response (e.g., tents), and landslides caused by the earthquake (see Fig. 7), adding substantially to mapping complexity. The Haiti damage mapping only addressed building damage. The effect of complexity on mapping performance, both in terms of numbers of feature types and number of damage levels to be mapped, remains to be assessed.

3.3 How can the mapping coordinator know whether volunteers actually interpret instructions in the same way as they were intended?

This question presents a two-fold problem: the mapping coordinator makes assumptions about the instructions (i.e., that they are comprehensive and unambiguous), while the mapper assumes that the mapped features are clear to the coordinator. Neither is necessarily the case. No instruction manual can cover all eventualities and, as we mentioned earlier, such an aim may be counterproductive. Thus, while the mapper learns by doing and has to make choices for each building in the grid cell, the resulting data layer likely contains variations in mapping precision (building outlines) and accuracy (damage levels). Since the mapper is also not forced (and indeed should not be forced) to follow a specified decision making process, or to motivate or justify a given damage label, the mapping becomes a series of ad hoc decisions that reflect

- i. How the mapper understood (and mentally adjusted) the sample illustrations;
- ii. the image analysis experience of the mapper;
- iii. the learning process the mapper undergoes, and;
- iv. the time the volunteer allocated to the task.

Feedback from volunteers involved in the Haiti mapping has shown that not all of them clearly understood the objectives of the mapping (Barrington et al., 2011). Blunders and clearly erroneous submissions resulting from such misunderstanding were identified and removed by ImageCat (Ghosh et al., 2011).

Collaborative mapping efforts carried out to date were further complicated by the fact that the mappers did not actually know the purpose of the damage data. Thus, they had no clear guidelines for the requirement of the map users (and their understanding of damage), the required precision for the building outlines, nor for the cost of mapping errors (both false positives and negatives). The latter has become a particular problem in post-disaster damage mapping and can be summarized as accountability. Although mapping by registered volunteers means that all map submissions are traceable to the mapper, there are no penalties for inaccuracies. The term “penalty” is used here in the sense that the review process in open mapping platforms such as OSM leads to a reliability ranking of the mapper, with mapping errors having negative consequences. We hypothesize that corrective feedback more positively affects overall mapping performance, as well as the dedication of the volunteer to the task, than penalizing. Mappers currently receive no feedback at all on their work, even though it has been identified as critical in voluntary collaborative activities (MacEachren and Cai, 2006). Corrective feedback is crucial in all perceptual learning tasks (Hoffman, 1990).

3.4 How can the mapping contributions by many volunteers best be merged and evaluated?

Collaborative mapping has the great advantage of parallel processing, leading to enormous time saving while keeping the mapping load per individual manageable. On the other hand it means that many hundreds of tiles should be checked (ideally, for accuracy, completeness and blunders). This is not an impossible goal. Using appropriate sampling, a representative statistical evaluation of the above 3 parameters is well possible. This was attempted by ImageCat staff, who visually checked submissions for obvious mistakes, checked all GIS layers automatically to search for polygon overlaps and spurious slivers, and subjected 5 % of all submissions to a detailed validation (Ghosh et al., 2011). Welinder et al. (2010) developed quantitative methods to assess the competence, expertise and potential bias of crowdsourcing contributors that could also be applied in the GEO-CAN context. The tiles also have to be integrated, implying grid cell border conflicts and ambiguities that also need to be resolved. In total some 2000 tiles of 500 × 500 m were mapped, and > 30000 building outlines submitted.

A number of additional specific questions about collaborative mapping emerge that have not yet been scientifically addressed:

- i. Should more than one person map a given grid cell to highlight variability in damage estimation or mapping mistakes, and simultaneously provide a validity and accuracy check? This was also recommended by Barrington et al. (2011), and already used by Tomnod in 2011 to map temporary refugee shelters in Somalia. Their CrowdRank approach used redundant entries to gauge mapper reliability and for the statistical estimation of the total number of shelters;
- ii. should mapping be more collaborative (using electronic communication and discussion means among volunteers)? For example, Fiore et al. (2010) showed how the knowledge of individuals enters the team decision making process and is used to build group knowledge, which can lead to more accurate and confident results. Tomnod used a shared training and feedback document where mappers could post questions and suggestions, which led to a discussion about mapping strategies, problems encountered, and to revised instructions;
- iii. should groups of volunteer mappers be assigned a professional responsible for a number of grid cells?

Dickinson et al. (2010) found that volunteers/amateurs work best when accompanied by a professional, though this may also be counterproductive by undermining confidence of the mapper. In the GEO-CAN-type situation only people with a remote sensing background participated, thus perhaps no mentoring/guiding is needed. Given that image-based damage mapping also challenges remote sensing experts, likely

the role of such a professional would be that of a moderator and discussion leader. On the other hand very few of the volunteers will have had actual experience (let alone substantial expertise) with damage mapping, and some coaching by someone more experienced could well be useful. If the experienced damage mapper serves as a back-stop rather than someone imposed on the volunteer, their input and guidance may be more effective. Given the likely absence of actual ground damage information, the mapping organizers, together with selected experts, could also create a consensus-based damage dataset based on visual image interpretation for a small test area to provide training and corrective feedback to the volunteers (see also Barrington et al., 2011). However, it must be considered that such expert data are themselves “noisy estimates” (Smyth et al., 1994). Castoldi et al. (2011) investigated the effect of image pre-processing on visual damage mapping, hypothesizing that appropriately enhanced images better highlight salient damage features. Using eye-tracking equipment they found that images enhanced using a morphological segmentation approach (Pesaresi and Benediktsson, 2001) led to improved mapping of D4 and D5 damage, and also reduced time between the identification of two damaged buildings. However, also unfinished structures (e.g., Fig. 4c and d) were highlighted in the pre-processing, leading to an increased number of false positives. The study further revealed that the value of image enhancement differed for groups with different remote sensing expertise. In particular expert mappers benefited from the pre-processing for the identification of D4 damage. This implies that such image enhancement increases the saliency of damage features, but does not reduce the semantic complexity of the task.

Another possibility is a technical solution aiming at plausibility checking of submitted results, or possibly automatic feedback during mapping such as giving suggestions or alerting to likely mapping errors (e.g., overlapping building polygons, excessively large number of damaged features per area unit, etc.). While some of these measures are very context-specific (e.g., building shape), basic plausibility checks, such as non-overlap of pre-disaster building polygons (see Fig. 5b), or flagging of polygons too small or too large to correspond to a building, can be performed by a computer. Within limits, the statistical plausibility of the damage level of an individual building with respect to the state of structures surrounding it can be assessed. In the Haiti work, for example, D4 and D5 damage was already statistically extrapolated to estimate damage at lower levels of severity (Ghosh et al., 2011). Such assessment, however, is strongly affected by both seismic site affects and the state and construction material of a given building, all of which are spatially highly variable.

3.5 What exactly needs to be done with the geoinformation, and to what extent do different users require adaptation in the mapping scheme?

Maps of any kind need to be useable, useful and understandable. But these three “u’s” must be understood as being relative to the person who is using the maps, and trying to make sense of the information they contain for their particular purposes and goals. A disaster manager has a variety of information needs, such as the number of buildings affected to various degrees, the structural integrity of public buildings that may serve as shelters or support posts, and suitable sites for emergency shelters. A government worker needs to begin estimating the amount of debris that needs to be removed, and how much concrete will be needed to rebuild a transportation infrastructure. The floor space estimation in Haiti was the basis for the subsequent Post Disaster Needs Assessment (PDNA), i.e., determination of financial requirements for rebuilding and recovery.

These and numerous other questions need to be answered through the use of software support systems and tools that are usable, useful, understandable and that lead to information (maps) that directly aids the decision maker. A given map might be elegant or informative from the standpoint of a remote sensing scientist or the cartographer who devised the map categories and symbology, but that same map might or might not make immediately perceptible the information that the emergency responder needs (Fig. 1; for a discussion on radically different damage map styles for the same disaster area in Haiti see Kerle, 2013).

Morss et al. (2005) studied the potential mutual misunderstandings between scientists and decision makers in flood risk management. They showed how perception and understanding of essential terms (e.g., uncertainty) differed in the two groups, and also how scientists had to reconsider their abstract definition of “the” decision maker. This led to an understanding of the specific needs of different user groups, which get further modulated by interaction between groups, and how those reach decisions using scientific information. It was shown how decision making needs must be included in the scientific research and development process in an iterative manner, with emphasis on multidisciplinary. Whether such a process leads to more useful map products in the post-disaster response context cannot be presumed, and must be determined empirically. This is where cognitive task analysis comes in.

4 The role of Cognitive Systems Engineering

Collaborative mapping and emergency response are both complex cognitive work systems. The scientific study and design of such systems is the focus of the field of cognitive systems engineering (CSE; Hoffman and Woods, 2000), which has its origins in the disciplines of human-computer

interaction and ergonomics. The primary methodology of CSE is referred to as cognitive task analysis (CTA; Crandall et al., 2006). The effectiveness of a technology intervention can be compromised by rapid changes in technology, the dynamic nature of the work that technology must support, and scaffolding for the worker to develop robust mental models. CTA is a suite of empirical and experimental methods that support technology developers in identifying cognitive challenges and understanding cognitive demands. CTA identifies, documents, and communicates the reasoning processes underlying performance so that the decision requirements can be incorporated into the design and implementation of information technology. CTA enables an understanding of the cognitive landscape that permits operators to use technology effectively, and developers to create information systems that support and enhance human performance. CTA methods have evolved as a necessary response to the increasingly complex sociotechnical systems that exist at the intersection of people, technology and work (Hoffman and Militello, 2008; Hoffman and Woods, 2000).

CTA studies aim to capture what people know and how they reason, including what they pay attention to, the strategies they are using to make decisions or detect problems, what they try to accomplish, and what they know about the way a process or system works. They can reveal the risks, time elements, opportunities, and mistakes that confront people as they work. They can help us understand the physical surroundings of work, including the technologies, tools, artifacts, work conditions, stressors, and types of team interaction and communication patterns that impact cognitive processes.

A variety of observational methods can be used to understand the flow, pace, climate, and activity of the workplace. This is often a critical component of CTA and would be expected to be central in the analysis of the cognitive work of emergency responders as they rely on mapping products. CTA can also help the designers of mapping methods to understand how volunteer mappers can be better instructed, especially in a more group-based approach. While only limited CTA has been applied to the challenges of rapid mapping for emergency response or issues of crowdsourced mapping (e.g., the use of thinking aloud – semi-structured interviews by Castoldi et al., 2011), it has been applied in the field of remote sensing more broadly (Hoffman and Markman, 2001; Hoffman et al., 2002).

Several of the methods mentioned here have been successfully adapted to study team cognitive processes: the way the teams communicate, draw inferences, share information, make judgments and decisions, see patterns, and make errors. Virtually all modern sociotechnical systems have a team aspect, and CTA methods can provide important insights into understanding cognitive processes and functions at the team and organizational levels. Some of the data analysis techniques and representational formats that appear particularly useful include creating narrative accounts of incidents and

examples, cataloguing cues and patterns, identifying themes, coding of conceptual and perceptual categories, and describing cognitive sequences and chronologies, and creating synthesized/integrated concept maps (Crandall et al., 2006).

We assert that the eventual purpose of the map information, and to what extent mapping strategies have to be adjusted to different user needs (question 5), play a central role in answering the questions related to the mapping process itself. Therefore, first the information and decision requirements of the people who will use the mapping products must be determined in as multidisciplinary a process as possible (Morss et al., 2005). This includes an assessment of their goals, and the information they need to navigate their problem space, how they navigate their trade-offs, and how they make decisions and take actions to achieve their goals. This determines what has to be shown to them, and how. Thus, it is formative of the work of the volunteer mappers, the instructions they are given, etc.

Software systems and displays that are intended to support cognitive work are, in effect, hypotheses about how the work will be conducted, based on some model of the work that is inside the head of the person (or team) that is writing the software, building the mapping interface and making decisions about the map features, symbology, etc. When the model of the work held by the designer (i.e., the remote sensing scientist) differs from the work process and requirements of the actual end-users of the software, or forces some designer-centred process on them, problems always result. It will lead to software with limited utility (Koopman and Hoffman, 2003), as has been seen repeatedly in small-scale and large-scale software procurements (Hoffman et al., 2009; Neville et al., 2008; Hollnagel and Woods, 2005). The technology to be used must thus be suitable for the available data and the questions to be answered. It also needs to strike a balance between guiding the mapper to a concrete decision (e.g., a specific identification and severity indicator for a given damaged feature from a fixed set of options), while not forcing a decision in ambiguous situations.

In the Wenchuan mapping no flexibility was allowed, while for the Haiti event mappers assigned confidence levels for each building polygon created (see Fig. 6a), effectively reducing rigidity of the process. However, complete mapping coverage (every building in a grid cell) is expected in collaborative mapping, and usually only a limited number of hard choices is provided (e.g., no visible damage all the way up to complete destruction). Sometimes there is an option for “unclear/uncertain” or “suspected damage” and the like. An improved process would allow the expression and sharing of rich judgments rather than the use of check boxes or reductive ratings scales, as is currently the case. The Tomnod platform, by fostering discussion and feedback, is beginning to address this issue.

It also needs to be considered that the eventual map user typically has a limited appreciation for the difficulties damage mappers face, and might take results for granted. Thus

ideally the mapping system should not only include the possibility to indicate the mapping confidence of a given mapped feature, but also a clear communication of its meaning in a collaborative development process with the map user. While a damage map user does not need to understand all technical details of the remote sensing and image analysis process, the “competence envelope” of the procedures applied in the map making should be clear.

CTA is a rich palette of multiple methods, including various forms of structured interviews and a number of experiment-like tasks that evaluate human performance (Hoffman and Militello, 2008). Retrospective analyses, such as the Critical Decision Method (CDM) (CDM; Hoffman et al., 1998) provide rich case studies and are informative of how the decision maker reasons and strategizes about difficult problems and tough cases. Combinations of methods, such as the combination of the CDM with the creation of Conceptual diagrams (Moon et al., 2011) are an efficient way to reveal the knowledge and information requirements of decision makers, as well as how volunteer mappers operate. CTA has made significant contributions to work design and software design in many venues in government, industry and the private sector (Cooke and Durso, 2007; Hollnagel and Woods, 2005). We emphasize that the CTA methods introduced in this article are in principle suitable to increase both the understanding of the decision making processes of emergency responders and the volunteer mappers.

5 The way forward

In this article we address some principal weaknesses in current collaborative damage mapping approaches, which are also the two endpoints of an information chain: the potential users of post-disaster information and their diverse needs that are insufficiently understood, and the way the volunteers’ knowledge is elicited and aggregated into a collaborative map product. The chain itself comprises three main actors: the volunteer community, the mapping organizers, and the users of the products. We identify a number of misconceptions that affect the effectiveness of this work chain:

- i. All volunteer mappers possess the skills to map structural damage accurately;
- ii. the same mapping instructions are equally understandable to all volunteers;
- iii. a map showing damage (in whatever form) is per se useful to the decision maker;
- iv. the map users will understand the damage information shown, as well as the uncertainties affecting its generation, and;
- v. a hardcopy damage map fulfils the needs of the disaster responder.

We posit that the creation of a consensual approach to collaborative mapping must include a continuing process of CTA to understand the information and decision requirements of the people who will use the map and image products (i.e., emergency responders), and to understand how the volunteers can be optimally instructed and their mapping contributions merged into user-needs specific and flexible map products with a known confidence indication. In particular, the mapping organizers (e.g., GEO-CAN), as connecting element between volunteer and map user, must address these issues. The following is a brief example outline of how CTA tools can be used practically to address our question 1 (“How can mapping instructions be conveyed in the most comprehensive and least ambiguous manner?”):

Step 1. Conduct Recent Case Walkthroughs with experienced volunteer mappers, resulting in a list of specifications of those aspects of the mapping task that were made easy and those made difficult.

Step 2. Conduct the Envisioning Desirements Method (Hoffman and McCloskey, 2013). This version of the CDM scaffolds the participant in describing work features that would make the work easier or better. Comparison of the outcome of the two methods would result in requirements for improved instructions and training.

Step 3. An evaluation study would take the standard form for training evaluation, involving the control of task demand characteristics and subject characteristics (experienced versus inexperienced volunteers). Re-designed and legacy training could then be compared in terms of usability and performance in a small-scale simulated emergency response task using a scenario based on archival data.

We recognize that it is a challenge to understand the specific needs of the many thousands of potential maps users in a post-disaster scenario, and to translate those needs into a portfolio of flexible information elements the volunteer community can generate. As emphasized by Duchêne et al. (2011), users must be allowed to select the information they require in a form that suits their needs and preferences, including customization of color schemes and symbology used, scaling or generalization of information, and use as hard- or softcopy. At the same time there must be some regularization or standardization.

It is apparent that current collaborative mapping methods and software have been created largely on the basis of “designer-centered design” rather than “work-centered design”. The questions we raise in this paper are not about the technology of mapping, but rather about the need for an approach that embraces the contributions of instructional designers into the creation of instructions. Specifically, we assert that both mapping instructions and software tools can be

best designed on the basis of results from CTA. All three actors are engaged in complex cognitive work. Unless all the efforts of the collaborative mappers and the map makers are oriented to making products that aid the cognitive work of the map user, then the enterprise is not going to work optimally. As was done by Morss et al. (2005) for flood risk management, we recommend an iterative, end-to-end-to-end approach that includes not only the remote sensing specialists, but also cognitive systems engineers and instructional designers, as well as experimental psychologists.

Following the Haiti earthquake, collaborative damage mapping was also applied to the 2011 Christchurch (New Zealand) earthquake. The event led to the creation of the Tomnod platform, with the intention of creating a simpler environment for collaborative damage mapping. While the site does include some online training, this primarily concerns the mapping tools in the software environment, but beyond that addresses few of the concerns we raise in this article (see Sect. 3.4). The platform and the approach it employs, including the CrowdRank method and the use of a dynamic discussion document, appear useful for the mapping of objects that have relatively clearly identifiable (and trainable) salient characteristics (e.g., tents in refugee camps). However, we consider its objective to "...open the initiative to a larger group of untrained analysts drawn from the general Internet public" (Barrington et al., 2011; p. 684) for semantically far more complex structural damage mapping to be misguided.

Acknowledgements. The authors thank the Earthquake Engineering Field Investigation Team (EEFIT) and Cambridge Architectural Research Ltd. for permission to reproduce field photographs, and Pictometry International Corp. for access to Pictometry Online. We also acknowledge the comments by two anonymous reviewers that helped to improve the manuscript.

Edited by: M. E. Contadakis

Reviewed by: M. Tsakiri-Strati and one anonymous referee

References

- Allendoerfer, K., Aluker, S., Panjwani, G., Proctor, J., Sturtz, D., Vukovic, M., and Chen, C.: Adapting the cognitive walk-through method to assess the usability of a knowledge domain visualization, Proceedings of INFOVIS 05: IEEE Symposium on Information Visualization, Minneapolis, MN, US, 2005, WOS:000234641900026, 195–202, 2005.
- Arciniegas, G., Bijker, W., Kerle, N., and Tolpekin, V. A.: Coherence- and amplitude-based analysis of seismogenic damage in Bam, Iran, using Envisat ASAR data, IEEE T. Geosci. Remote Sens., 45, 1571–1581, doi:10.1109/TGRS.2006.883149, 2007.
- Barrington, L., Ghosh, S., Greene, M., Har-Noy, S., Berger, J., Gill, S., Lin, A. Y. M., and Huyck, C.: Crowdsourcing earthquake damage assessment using remote sensing imagery, Ann. Geophys., 54, 680–687, doi:10.4401/ag-5324, 2011.
- Bertin, J.: Semiology of graphics, University of Wisconsin Press, Madison, WI, US, 1983.
- Castoldi, R., Broglia, M., and Pesaresi, M.: Image interpreters and interpreted images: an eye tracking study applied to damage assessment, Proceedings of VALGEO 2011 – 3rd International Workshop on Validation of Geo-information Products for Crisis Management, Ispra, Italy, 2011, 83–91, 2011.
- Cooke, N. J. and Durso, F.: Stories of Modern Technology Failures and Cognitive Engineering Successes CRC Press, Boca Raton, FL, US, 136 pp., 2007.
- Corbane, C., Carrion, D., and Broglia, M.: Development and implementation of a validation protocol for crisis maps: reliability and consistency assessment of burnt area maps, Int. J. Digital Earth, 8–24, doi:10.1080/17538947.2010.512153, 2011a.
- Corbane, C., Saito, K., Dell’Oro, L., Gill, S. P. D., Piard, B. E., Huyck, C. K., Kemper, T., Lemoine, G., Spence, R. J. S., Shankar, R., Senegas, O., Ghesquiere, F., Lallemand, D., Evans, G. B., Gartley Joaquin Toro, R. A., Ghosh, A., Svekla, W. D., Adams, B. J., and Eguchi, R. T.: A comprehensive analysis of building damage in the 12 January 2010 Mw7 Haiti earthquake using high resolution satellite and aerial imagery, Photogr. Eng. Remote Sens., 77, 997–1009, 2011b.
- Crandall, B., Klein, G. A., and Hoffman, R. R.: Working Minds: a Practitioner’s Guide to Cognitive Task Analysis, MIT Press, Cambridge, MA, US, 2006.
- Dell’Acqua, F. and Gamba, P.: Remote sensing and earthquake damage assessment: Experiences, limits, and perspectives, Proceedings of the IEEE, 100, 2876–2890, doi:10.1109/jproc.2012.2196404, 2012.
- Dickinson, J. L., Zuckerberg, B., and Bonter, D. N.: Citizen Science as an ecological research tool: Challenges and benefits, in: Annual Rev. Ecol., Evolution, and Systematics, 41, edited by: Futuyma, D. J., Shafer, H. B., and Simberloff, D., Annual Review of Ecology Evolution and Systematics, Ann. Reviews, Palo Alto, CA, US, 149–172, 2010.
- Duchêne, C., Christophe, S., and Ruas, A.: Generalisation, symbol specification and map evaluation: feedback from research done at COGIT laboratory, IGN France, Int. J. Digital Earth, 4, 25–41, doi:10.1080/17538947.2010.510305, 2011.
- Fiore, S. M., Rosen, M. A., Smith-Jentsch, K. A., Salas, E., Letsky, M., and Warner, N.: Toward an understanding of macrocognition in teams: predicting processes in complex collaborative contexts, Human Factors, 52, 203–224, doi:10.1177/0018720810369807, 2010.
- Gerke, M. and Kerle, N.: Automatic structural seismic damage assessment with airborne oblique Pictometry ©imagery, Photogr. Eng. Remote Sens., 77, 885–898, 2011.
- Ghosh, S., Huyck, C. K., Greene, M., Gill, S. P., Bevington, J., Svekla, W., DesRoches, R., and Eguchi, R. T.: Crowdsourcing for rapid damage assessment: the Global Earth Observation Catastrophe Assessment Network (GEO-CAN), Earthquake Spectra, 27, S179–S198, doi:10.1193/1.3636416, 2011.
- Girres, J. F. and Touya, G.: Quality assessment of the French OpenStreetMap dataset, Transactions in GIS, 14, 435–459, doi:10.1111/j.1467-9671.2010.01203.x, 2010.
- Goodchild, M.: Citizens as sensors: the world of volunteered geography, GeoJournal, 69, 211–221, doi:10.1007/s10708-007-9111-y, 2007.

- Goodchild, M. F. and Glennon, J. A.: Crowdsourcing geographic information for disaster response: a research frontier, *Int. J. Digital Earth*, 3, 231–241, doi:10.1080/17538941003759255, 2010.
- Grünthal, G.: European Macroseismic Scale 1998 (EMS-98), *Cahiers du Centre Européen de Géodynamique et de Séismologie* 15, Centre Européen de Géodynamique et de Séismologie, 99 pp., 1998.
- Haklay, M.: How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets, *Environ. Planning B-Planning & Design*, 37, 682–703, doi:10.1068/b35097, 2010.
- Haklay, M., Basiouka, S., Antoniou, V., and Ather, A.: How many volunteers does it take to map an area well? The validity of Linus' Law to volunteered geographic information, *Cartographic Journal*, 47, 315–322, doi:10.1179/000870410x12911304958827, 2010.
- Hoffman, R. R.: Remote perceiving: A step toward a unified science of remote sensing, *Geocarto Int.*, 5, 3–13, doi:10.1080/10106049009354254, 1990.
- Hoffman, R. R. and Fiore, S. M.: Perceptual (Re)learning: a leverage point for human-centered computing, *IEEE Intelligent Systems*, 22, 79–83, doi:10.1109/mis.2007.59, 2007.
- Hoffman, R. R. and Markman, A. B.: *Interpreting Remote Sensing Imagery: Human Factors*, Lewis Publishers, Boca Raton, FL, US, 289 pp., 2001.
- Hoffman, R. R. and McCloskey, M.: *Envisioning desirements*, *IEEE Intelligent Systems*, in press, 2013.
- Hoffman, R. R. and Militello, L. G.: *Perspectives on Cognitive Task Analysis: Historical Origins and Modern Communities of Practice*, CRC Press/Taylor and Francis, Boca Raton, FL, US, 516 pp., 2008.
- Hoffman, R. R. and Woods, D. D.: Studying cognitive systems in context: Preface to the special section, *Human Factors*, 42, 1–7, doi:10.1518/001872000779656633, 2000.
- Hoffman, R. R., Crandall, B., and Shadbolt, N.: Use of the critical decision method to elicit expert knowledge: A case study in the methodology of cognitive task analysis, *Human Factors*, 40, 254–276, doi:10.1518/001872098779480442, 1998.
- Hoffman, R. R., Coffey, J. W., Carnot, M. J., and Novak, J. D.: An empirical comparison of methods for eliciting and modeling expert knowledge, *Proceedings of the 46th Meeting of the Human Factors and Ergonomics Society*, Santa Monica, CA, US, 482–486, 2002.
- Hoffman, R. R., Neville, K., and Fowlkes, J.: Using cognitive task analysis to explore issues in the procurement of intelligent decision support systems, *Cognition, Technology, and Work*, 11, 57–70, doi:10.1007/s10111-008-0120-5, 2009.
- Hollnagel, E. and Woods, D. D.: *Joint Cognitive Systems: Foundations of Cognitive Systems Engineering*, *Ergonomics in Design: The Quarterly of Human Factors Applications*, CRC Press Boca Raton, FL, 2005.
- Ito, A.: Issues in the implementation of the International Charter on Space and Major Disasters, *Space Policy*, 21, 141–149, 2005.
- Jobe, K.: Disaster relief in post-earthquake Haiti: Unintended consequences of humanitarian volunteerism, *Travel Medicine and Infectious Disease*, 9, doi:10.1016/j.tmaid.2010.10.006, 2011.
- Kerle, N., Heuel, S., and Pfeifer, N.: Real-time data collection and information generation using airborne sensors, in: *Geospatial Information Technology for Emergency Response*, edited by: Zlatanova, S. and Li, J., Taylor & Francis, London, 43–74, 2008.
- Kerle, N.: Satellite-based damage mapping following the 2006 Indonesia earthquake—How accurate was it?, *Int. J. Appl. Earth Observation Geoinform.*, 12, 466–476, doi:10.1016/j.jag.2010.07.004, 2010.
- Kerle, N.: Remote sensing based post-disaster damage mapping with collaborative methods, in: *Intelligent Systems for Crisis Management*, edited by: Zlatanova, S., Peters, R., Dilo, A., and Scholten, H., *Lecture Notes in Geoinformation and Cartography*, Springer, Berlin Heidelberg, 13 pp, 2013.
- Klein, G. A. and Hoffman, R. R.: *Seeing the invisible: perceptual-cognitive aspects of expertise*, in: *Cognitive Science Foundations of Instruction*, edited by: Rabinowitz, M., Lawrence Erlbaum Associates, Inc, Hillsdale, NJ, US, 203–226, 1993.
- Koopman, P. and Hoffman, R. R.: Work-arounds, make-work, and Kludges, *IEEE Intelligent Systems*, 18, 70–75, doi:10.1109/mis.2003.1249172, 2003.
- Lemoine, G.: Validation of building damage assessments based on post-Haiti 2010 earthquake imagery using multi-source reference data, *Proceedings of VALgEO 2010 – 2nd International Workshop on Validation of Geo-information Products for Crisis Management*, Ispra, Italy, 2010, 33–34, 2010.
- MacEachren, A. M. and Cai, G. R.: Supporting group work in crisis management: visually mediated human-GIS-human dialogue, *Environment and Planning B-Planning & Design*, 33, 435–456, doi:10.1068/b3188, 2006.
- Moon, B., Hoffman, R. R., Cañas, A. J., and Novak, J. D.: *Applied Concept Mapping: Capturing, Analyzing and Organizing Knowledge*, Taylor & Francis, Boca Raton, FL, US, 380 pp., 2011.
- Morss, R. E., Wilhelmi, O. V., Downton, M. W., and Grunfest, E.: Flood risk, uncertainty, and scientific information for decision making – Lessons from an interdisciplinary project, *Bull. Am. Meteorol. Soc.*, 86, 1593–1601, doi:10.1175/bams-86-11-1593, 2005.
- Neville, K., Hoffman, R. R., Linde, C., Elm, W. C., and Fowlkes, J.: The procurement woes revisited, *IEEE Intelligent Systems*, 23, 72–75, doi:10.1109/mis.2008.15, 2008.
- Norman, D. A.: *The Invisible Computer*, MIT Press, Cambridge, MA, US, 320 pp., 1998.
- Pesaresi, M. and Benediktsson, J. A.: A new approach for the morphological segmentation of high-resolution satellite imagery, *IEEE T. Geosci. Remote Sens.*, 39, 309–320, doi:10.1109/36.905239, 2001.
- Polson, P. G., Lewis, C., Rieman, J., and Wharton, C.: Cognitive walkthroughs – a method for theory-based evaluation of user interfaces, *Int. J. Man-Machine Studies*, 36, 741–773, doi:10.1016/0020-7373(92)90039-n, 1992.
- Roche, S., Propeck-Zimmermann, E., and Mericskay, B.: GeoWeb and crisis management: issues and perspectives of volunteered geographic information, *GeoJournal*, 1–20, doi:10.1007/s10708-011-9423-9, 2011.
- Shimoni, M., Crosetto, M., Lang, S., Bally, P., and Boubila, F.: The independent service validation in GMES RESPOND: the flood validation exercise, *Int. J. Digital Earth*, 4, 58–76, doi:10.1080/17538947.2010.510298, 2011.
- Smyth, P., Burl, M. C., Fayyad, U. M., and Perona, P.: Knowledge discovery in large image databases: dealing with uncertainties in ground truth, *KDD Workshop*, Seattle, Wa, US, 1994.

- Stryker, T. and Jones, B.: Disaster response and the International Charter program, *Photogr. Eng. Remote Sens.*, 75, 1342–1344, 2009.
- Tomaszewski, B.: Situation awareness and virtual globes: Applications for disaster management, *Comp. Geosci.*, 37, 86–92, doi:10.1016/j.cageo.2010.03.009, 2011.
- Turner, A. J.: *Introduction To Neogeography*, O'Reilly Media, Inc., Sebastopol, CA, US, 54 pp., 2006.
- van Aardt, J. A. N., McKeown, D., Faulring, J., Raqueno, N., Castlerline, M., Renschler, C., Eguchi, R., Messinger, D., Krzaczek, R., Cavillia, S., Antalovich, J., Philips, N., Bartlett, B., Salvaggio, C., Ontiveros, E., and Gill, S.: Geospatial disaster response during the Haiti earthquake: a case study spanning airborne deployment, data collection, transfer, processing, and dissemination, *Photogr. Eng. Remote Sens.*, 77, 943–952, 2011.
- Voigt, S., Schneiderhan, T., Twele, A., Gahler, M., Stein, E., and Mehl, H.: Rapid damage assessment and situation mapping: learning from the 2010 Haiti earthquake, *Photogr. Eng. Remote Sens.*, 77, 923–931, 2011.
- Welinder, P., Branson, S., Belongie, S., and Perona, P.: The multidimensional wisdom of crowds, in: *Advances in Neural Information Processing Systems 23*, edited by: Lafferty, J., Williams, C. K. I., Shawe-Taylor, J., Zemel, R. S., and Culotta, A., MIT Press, Cambridge, MA, US, 2424–2432, 2010.
- Zhang, Y. and Kerle, N.: Satellite remote sensing for near-real time data collection, in: *Geospatial Information Technology for Emergency Response*, edited by: Zlatanova, S. and Li, J., Taylor & Francis, London, 75–102, 2008.