



Intercomparison and validation of building damage assessments

G. Lemoine et al.

This discussion paper is/has been under review for the journal Natural Hazards and Earth System Sciences (NHESS). Please refer to the corresponding final paper in NHESS if available.

Intercomparison and validation of building damage assessments based on post-Haiti 2010 earthquake imagery using multi-source reference data

G. Lemoine¹, C. Corbane^{1,2}, C. Louvrier¹, and M. Kauffmann¹

¹European Commission, Joint Research Centre Via E. Fermi 2749, 21027 Ispra (VA), Italy

²Institut national de recherche en sciences et technologies pour l'environnement et l'agriculture – UMR TETIS Campus international de Baillarguet, 34398 Montpellier, France

Received: 1 March 2013 – Accepted: 23 March 2013 – Published: 22 April 2013

Correspondence to: C. Corbane (christina.corbane@gmail.com,
christina.corbane@teledetection.fr)

Published by Copernicus Publications on behalf of the European Geosciences Union.

| | |
|--------------------------|--------------|
| Title Page | |
| Abstract | Introduction |
| Conclusions | References |
| Tables | Figures |
| ⏪ | ⏩ |
| ◀ | ▶ |
| Back | Close |
| Full Screen / Esc | |
| Printer-friendly Version | |
| Interactive Discussion | |



Abstract

The Haiti 2010 earthquake is one of the first major disasters in which very high resolution satellite and airborne imagery was embraced to delineate the event impact. Several rapid mapping initiatives exploited post-earthquake satellite and airborne imagery to produce independent point feature sets marking the damage grade of affected buildings. Despite the obvious potential of the satellite remote sensing technology in providing damage figures, the scale and complexity of the urban structures in Port-au-Prince cause overall figures and patterns of the damage assessments to yield a rather poor representation of the true damage extent. The higher detail airborne imagery performs much better as confirmed by different validation studies carried out in the last two years. In this paper, in addition to the review and analysis of the different validation works, we investigate the quality of damage assessment derived by different activities through a simple intercomparison and a validation using a complete building ground survey. The results show that the identification of building damage from aerial imagery provides a realistic estimate of the spatial pattern and intensity of building damage.

1 Introduction

The use of Earth Observation in post-disaster rapid mapping and impact assessments has seen impressive take-up following the Indonesian tsunami of 2004 (Pesaresi et al., 2007; Voigt et al., 2007). In general, two categories of activities can be identified, (1) the so-called “rapid mapping” actions which aim at providing impact assessments as quickly as possible after a disaster has struck and (2) assessments to assist in the enumeration of economic impact of damage and losses. Major distinctions between the 2 categories are timeliness, completeness and accuracies. Rapid mapping outputs are meant to contribute to planning of emergency response actions, and hence, have to be delivered in a timely fashion, preferably within 24 h of the event. Furthermore, rapid mapping prioritizes information generation over the highest impact areas. Often

NHESSD

1, 1445–1486, 2013

Intercomparison and validation of building damage assessments

G. Lemoine et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Intercomparison and validation of building damage assessments

G. Lemoine et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



rapid mapping aims to portray the relative severity of local event impacts, for instance, low, medium and high damage areas, rather than absolute damage figures. Thus, some variation in absolute accuracy can be tolerated as long as this is not expected to lead to decisions that wrongly allocate response efforts. Although, especially in Europe, rapid mapping services are now used in an operational mode, little is known about error propagation of rapid mapping outputs into the decision making stage.

The second category of assessments typically requires significantly more time to generate than rapid mapping products, usually one week or more. These detailed assessments aim at being exhaustive, both in the geographical sense, i.e. covering the complete impact area as far as possible, and in the thematic sense, i.e. capturing the impact on specific built-up categories or land use classes as complete as possible. The intended use of detailed damage assessment goes beyond the immediate emergency response phase, and is typically aimed at generating an early estimate of economic damage and losses caused by the disaster. The latter points out that the Earth Observation analysis actually delivers only partial, though essential, elements in the overall economic assessment. Results from the damage enumeration need to be combined with other geospatial covariates (e.g. building size, building type) and economic parameters (e.g. cost per unit area, building material, relocation costs, etc.) to come to a total cost estimate. These cost estimates are important parameters in establishing financial and institutional support in the recovery and reconstruction phase that follows the disaster. The total number of buildings by damage grade is a weighted multiplier in the overall damage cost equation. Thus, errors in building enumeration and categorization propagate proportionally into the overall cost estimate. It is evident, therefore, that Earth Observation based enumeration of damage impact needs to be sufficiently accurate, and at least within the tolerable variance of the overall damage estimate.

Validation of Earth Observation output is generally a time- and effort intensive task, as it implies the collection of detailed and representative reference data, usual in-situ or from third party, expensive sources (e.g. higher resolution imagery, small scale up to date maps, etc.). The need for validation is particularly relevant in emergency response

and post-disaster needs assessment contexts: the risk that output leads to erroneous decision making may have strong negative consequences, either as wrongly directed emergency intervention or as an under- or over-estimation of economic recovery figures. Validation also provides an unbiased way to assess how new technological solutions contribute to methodological improvements in Earth Observation based assessments. Such technological improvements may include the access to higher spatial or spectral resolution imagery, enhanced viewing capabilities or advances in automatic and interactive processing of large image data sets.

The 12 January 2010 earthquake in Haiti caused widespread damage to buildings in the country's capital, Port-au-Prince, and surroundings towns. Post-event very high resolution GeoEye imagery was acquired over the impact area on 13 January, and, exceptionally, released as a 0.5 m spaced pan-sharpened natural color composite in the public domain, both integrated in Google Earth and accessible via FTP. Similarly, pre-event pan-sharpened QuickBird imagery was made available (of 26 August 2009 for much of Port-au-Prince, but older for other areas in the impact zone). This combination of pre- and post-event satellite images sets was used by diverse groups to generate rapid mapping output serving the emergency response user community. The satellite-based rapid mapping products were generated either through computer-aided visual interpretation (Voigt et al., 2011) or through automatic damage assessment approaches. Both mono-temporal (Hussain et al., 2011; Kaya et al., 2011; Li et al., 2011) and multi-temporal (Malinverni, 2011; Tiede et al., 2011) automatic damage assessment methods were developed as part of research works aimed at providing useful information to assist in damage assessment. However, since automatic damage assessment approaches are still experimental, most of the rapid satellite-based damage assessment maps produced shortly after the disaster relied on visual image interpretation.

From 17 January onwards, first Google Inc. and later the World Bank/ImageCAT/RIT teams began to release newly acquired aerial photography with a spatial resolution of 0.15 m. Since an initial analysis showed damage delineation to be much more

NHESSD

1, 1445–1486, 2013

Intercomparison and validation of building damage assessments

G. Lemoine et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



the imagery based damage assessment is able to reproduce the ground-based observations. Validation, therefore, provides quantitative indicators on the confidence in the representativeness of the final damage figures.

In this paper, our focus is on assessing the quality of damage assessment outputs derived by different activities through a simple intercomparison and a validation with multi-source reference data:

- we first compare two point sets (one set is from Open Street Map and the other one is from the Chinese Academy of Sciences) of building damage derived from pre- and post-event VHR satellite imagery with the PDNA point set. The latter has been jointly generated by UNOSAT, the World Bank and the Joint Research Centre based on the combined use of pre-event satellite and post-event aerial photography;
- then we review and analyze the results of different studies and approaches that were developed for the validation of the PDNA point set;
- and finally, we perform a full validation of the PDNA point set using as a reference an exhaustive field campaign and try to explain the observed underestimations or overestimations in building damage derived from the aerial imagery.

2 Data set description

In the following the building damage assessments derived from satellite and aerial imagery are presented. Then, we present, the different reference data sets used for the validation including oblique aerial imagery (pictometry) and field assessments.

2.1 Satellite-based damage assessments

Two point sets of damaged buildings derived from pre- and post-event VHR satellite imagery are analyzed. The sets are from Open Street Map (OSM) and the Chinese

Intercomparison and validation of building damage assessments

G. Lemoine et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Academy of Sciences (CAS), respectively (the datasets can be downloaded from this website: http://wiki.openstreetmap.org/wiki/WikiProject_Haiti/VectorAndMapData).

The OSM set is a point set identifying collapsed buildings, which is a combination of an early damage assessment made by Information Technology for Humanitarian Assistance, Cooperation and Action (ITHACA) and extended with OSM contributions. It covers the complete impact area, including the cities outside greater Port-au-Prince, i.e. Gressier, Leogane and Jacmel. The downloaded OSM set of 15 May 2010 contains 8165 positions corresponding to heavily affected and collapsed buildings.

The CAS published a complete damage assessment set, collated from a collaborative satellite image interpretation exercise, on 17 February 2010. It contains vector outlines of heavily affected and collapsed buildings for 5587 locations, mostly in the greater Port-au-Prince and Carrefour areas.

2.2 Aerial-based damage assessments

The PDNA point data set designated here by UWJ (UNOSAT-World Bank-JRC) was generated by computer aided visual interpretation of pre-earthquake satellite imagery and post-earthquake airborne imagery with a spatial resolution of 0.15 m. The methodology consisted in assignment of different damage grades to individual buildings, using the European Macro-seismic Scale (EMS) 1998 (Grünthal and Levret, 2001). This scale defines 5 damage grades: grades 1 to 5 should ideally represent a progressive increase in the strength of shaking for different types of masonry and reinforced-concrete buildings. The damage descriptions for these two building types were specifically selected because they most closely represented the types of construction in Haiti. The main qualitative criteria used for recognizing damage grades on overhead or nadir imagery include, but is not limited to, the shifting of building perimeters, a lack of definition of perimeter walls, or obvious debris surrounding the damaged buildings. However, for some types of failure, using only nadir imagery will result in missed damage assignments. For example, many soft-story building failures were observed in the Haiti earthquake. These failures usually occur when the upper stories of a building collapse

NHESSD

1, 1445–1486, 2013

Intercomparison and validation of building damage assessments

G. Lemoine et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



directly on to the first or bottom story. They are common when the upper stories are constructed to be much stiffer than a more open or softer bottom story. In the nadir images, these types of failures are difficult to identify (Fig. 2a) without the use of oblique imagery or ground survey data (Fig. 2b). In addition, more minor or moderate damage to buildings that does not result in building collapse are also difficult to identify in the nadir images. Accordingly, the damage levels used for marking the individual buildings on the aerial photographs were limited to the higher grades of the EMS-98 scale, e.g. substantial to heavy damage (EMS grade 3), very heavy damage (EMS grade 4), and destruction (EMS grade 5). All building centroids were marked including those that did not exhibit visible damage (EMS grade 0, 1 and 2). The overall point data set consists of more than 300 000 labeled points.

2.3 Reference data sources

2.3.1 Field survey following a stratified random sampling design

One month after the earthquake, JRC, the World Bank, and UNOSAT deployed teams to carry out building-by-building ground surveys in the greater Port-au-Prince area including Carrefour, Petionville, and Delmas, as well as in Gressier and Leogane. The National Geospatial Capacity (Centre National de l'Information Géo-Spatiale CNIGS) of Haiti subsequently took over the ground survey work. The field survey that was carried out between 8 March and 7 May 2010 was done according to a stratified random sampling design with the purpose of surveying both high density and low density built-up areas as well as areas with high and low density of building damage. Four strata were accordingly defined:

- Medium to low density built-up zone/High damage density
- Medium to low density built-up zone/Low damage density
- High density built-up zone/High damage density

NHESSD

1, 1445–1486, 2013

Intercomparison and validation of building damage assessments

G. Lemoine et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



– High density built-up zone/Low damage density

The sampling unit consisted of a cell of 200 × 200 m corresponding to the cell size that was used during the photo-interpretation of the aerial images and that has been determined based on image resolution and on the size of detectable damage artifacts.

The total number of cells that fulfilled the stratification criteria was 2150. To determine the sample size (N) to be selected from the 2150 retained cells, we defined the confidence level and confidence interval as a priori criteria and calculated the sample size accordingly (Cochran, 1977).

$$N = \frac{Z^2 * (p) * (1 - p)}{c^2} \quad (1)$$

where: $Z = Z$ value (e.g. 1.96 for 95 % confidence level); $p =$ percentage picking a choice, expressed as decimal; $c =$ confidence interval, expressed as decimal.

For a 95 % confidence level, Z value is 1.96 and $c = 0.106$ and $p = 0.50$ were chosen. Hence the computed sampled size was equal to 86 representing 4 % of the population. Finally the 86 cells were selected randomly among the 2150 previously stratified cells in a number proportional to the stratum's size. In total, 6492 buildings were surveyed on the ground. Basic survey methods were used, comprising a detailed GPS-tagged photo and a description of the damage level based on the EMS-98 scale.

The field survey was specifically tailored to the validation of the aerial damage assessment. An atlas of building damage using aerial imagery was printed in A3 tiles and used in the field. Each surveyed building was numbered on the printed maps and its position and status recorded. For each day of field work one day of post-processing and quality control was necessary to verify the locations of the surveyed buildings and reposition the points collected in the field over the photo-interpreted points. Building status attributes were transcribed from the paper forms used in the field (e.g. damage state, number of floors, occupancy, etc.). Several mismatches and problematic points were found. This was mostly observed in shanty zones and heavily destroyed urban

Intercomparison and validation of building damage assessments

G. Lemoine et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



blocks and was mainly due to the fact that the points on the photo-interpreted damage assessment corresponded to the locations as observed on the pre-event imagery. Especially for completely destroyed areas it was practically impossible to define the positions of several collapsed buildings. A detailed description of the field campaign can be found in Taucer et al. (2010). The difficulty in directly relating remote sensing derived damage assessment to field observations is a typical problem of in-situ validation of damages to buildings. It is conditioned, on the one hand by the problem of defining the concept of damage and its manifestation in remote sensing imagery and, on the other hand, by the inherent limitations of field surveys such as accessibility, incomplete coverage conditioned by substantial cost in manpower and time or transcription errors (Kerle, 2010). In the study case, only 4329 of the 6492 surveyed buildings correspond to the buildings included in the aerial-based damage assessment.

2.3.2 Along roads track survey from Low Cost Mobile Mapping System (LCMMS)

A Low Cost Mobile Mapping System (LCMMS, Ajmar et al., 2011) developed by ITHACA was deployed in Port-au-Prince (Haiti) and tested during a five days survey from 24 February to 2 March 2010. The system allows acquiring video and single geo-referenced video-image frames by means of a transportable device easily installable on every type of vehicles. It is composed by four webcams with a total field of view of about 180 degrees and one GPS receiver, with the main aim to rapidly cover large areas for an effective usage in emergency situations. The system has been mounted on a car and used during a 5 days survey in Port-au-Prince and nearby localities. It allowed acquiring about 490 000 geo-referenced frames along the roads. A subset of the data acquired with LCMMS in Haiti has been used for the validation of assessments based on high resolution aerial images. Several blocks of Port-au-Prince covered by both the LCMMS survey and the high resolution aerial imagery were chosen (Fig. 1). Given the nature of LCMMS imagery, only 168 buildings facing the roads were taken into account. Those buildings have been analysed by an operator and classified according to

Intercomparison and validation of building damage assessments

G. Lemoine et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



the EMS-98 damage grades and used for assessing the accuracy of the aerial-based damage assessment (Ajmar et al., 2011).

3 Block-sampling using Pictometry data

Pictometry imagery with a spatial resolution of better than 25 cm was acquired in early February 2010 from an aircraft flown at a height of around 1 km with multiple cameras at nadir and oblique angles. A set of 60 locations corresponding to urban blocks (defined by street intersections) were selected in the Port-au-Prince area so that approximately equal numbers of locations were in each of following land use categories: downtown; commercial, residential and informal (shanty) settlements. At each location, 20 buildings approximately (1247 buildings in total) were identified on the Pictometry images and assessed by Cambridge Architectural Research as part of an exercise funded by the World Bank (Spence and Saito, 2010). The Pictometry dataset has been validated in the field (Booth et al., 2011). The results demonstrated the tendency of Pictometry to underestimate the proportion of heavily affected buildings (grade 4) and destroyed buildings (grade 5) as well as the difficulty of making assessments of moderate or low damages (grades 2 and 3). The main reasons for the divergence between Pictometry and ground observations assignments were related to (i) the difficulty of spotting soft-story collapses and (ii) the problem of damage being obscured by buildings and trees. Other problems inherent to damage mapping in Pictometry data were also discussed in Gerke and Kerle (2011). The main issue is that the quality of visual damage mapping is subject to a number of constraints, principally image type and quality, viewing perspective, building type and density, as well as analyst experience. Besides, the individual indicators of structural impact, such as roof or façade damage, cannot always be unambiguously assigned to a damage category, especially for the lower damage classes (grades 1, 2 and 3). This uncertainty poses challenges also for the intercomparison of the Pictometry data, the field observations and the aerial based assessment.

Intercomparison and validation of building damage assessments

G. Lemoine et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



3.0.3 Exhaustive field survey

Between March and July 2010, the Haitian Ministry of Public Works (MTPTC), with support from United Nations Office of Project Services (UNOPS), the Pan America Development Foundation and Miyamoto International, conducted a “house-to-house” assessment of building damages in order to ensure the safety of each building and to estimate the rebuilding or repair costs of all the affected structures. A mapped inventory of the damage to 400 000 houses in the affected communes has been established with a color code for different levels of damage: slightly damaged but habitable (green); unstable and requiring repair (yellow); partly or completely destroyed or dangerous (red). In addition to the color code, each building was described in terms of percentage of damage based on the following class intervals: 0–1 %, 1–10 %, 10–30 %, 30–60 %, 60–100 % and 100 % (Miyamoto et al., 2011). This level of investigation was considered the most reliable and most accurate indicator of damage and thus represents the benchmark to which the remote-sensing assessment should be compared.

This inventory has however some quality problems that must be taken into account when using this dataset as a reference for validation purposes. Some typical examples of quality issues were encountered in the informal neighbourhoods where the inventory seems to have overlooked heavily damaged houses (Fig. 3a and b). Among the reasons for the underestimation of the damage we could possibly cite, the presence of rubble that prevented access to field visits. On some other occasions, several labels were assigned to the same building while neighboring buildings were not assessed (Fig. 3c). Apparently some visits were only conducted from the outside or from the road (Fig. 3d).

Another quality issue that was identified during the analysis of this dataset, was the presence of inconsistencies between the assignments of the color codes and the percentages of damages. Figure 4 shows that 7 % and 10 % of the buildings falling respectively in the red and yellow categories, have been labeled with a damage percentage value of 0–1 %. All these highlighted inconsistencies demonstrate the lack of

Intercomparison and validation of building damage assessments

G. Lemoine et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



a comprehensive quality control of the collected information before its on-line publication and the necessity to carefully handle this dataset especially when using it as a reference within a validation exercise.

4 Intercomparison of satellite and aerial based damage assessments

The purpose of the intercomparison is to provide an objective measure of the quality of the satellite based damage assessments and to develop an understanding of the validity and comparability of the satellite (OSM and CAS datasets) and aerial (UWJ) derived damages to buildings. Given that the three datasets have different coverage, it was decided to focus the analysis on a subset comprising all the three assessments and covering an area of approximately 25 × 15 km, which is enclosing the area depicted in Fig. 5.

Since the aerial based damage assessment represents the most detailed data, it was considered here as the reference against which the intercomparison was made. While the UWJ dataset consists of four different damage grades (Grade 1: no visible damage, Grade 3: substantial to heavy damage, Grade 4: very heavy damage, Grade 5: destruction), the OSM and CAS datasets contain only collapsed and heavily affected buildings that actually correspond to Grades 5 and 4, respectively according to the EMS-scale. Therefore, for the intercomparison, only points labelled with Grades 5 and 4 were considered from the UWJ dataset. The differences in the spatial resolutions between the satellite (50 cm) and the aerial (15 cm) imagery combined with displacement error between them do not allow an accurate point-to-point matching. To circumvent this problem, the point counts for grid cells of 50 × 50 m were taken as a reference for the subsequent analysis. The window size of 50 meters showed to be a reasonable compromise between (i) the displacement errors, (ii) the building sizes and the spacing between 16 them and (iii) the requirements for the following statistical analysis.

Overall statistics in Table 1 show that the detection of damage classes 4 and 5 in the aerial imagery is approximately a factor 8 greater than that in the satellite imagery.

Intercomparison and validation of building damage assessments

G. Lemoine et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Intercomparison and validation of building damage assessments

G. Lemoine et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



However, differences in mean and standard deviations already indicate that relative changes in the sets are not equally distributed. The study of the pattern of matching between the aerial and the satellite derived damage assessments (UWJ versus CAS and UWJ versus OSM) was performed using a 2-D cross-correlation (Gonzalez and Woods, 1992) on the gridded datasets. The point sets corresponding to grades 5 and 4 were aggregated for a direct comparison with the OSM and CAS datasets. For each 3×3 window of neighbouring grid cells, the normalised correlation coefficient is calculated as follows:

$$\gamma = \frac{\sum_{i,j=0}^2 (f(i,j) - \bar{f})(g(i,j) - \bar{g})}{\sqrt{\sum_{i,j=0}^2 (f(i,j) - \bar{f})^2 \sum_{i,j=0}^2 (g(i,j) - \bar{g})^2}} \quad (2)$$

where: $f(i, j)$, $g(i, j)$ is the point count for each grid sample in the window ($i, j = 0..2$ for a 3×3 window) for the first set (e.g. UWJ) and the second set (e.g. OSM); \bar{f} , \bar{g} is the average point count for all grid samples in the window.

The γ -value for each window is assigned to the central pixel of each 3×3 window. For two datasets showing a good matching of the point counts in the 50×50 m grid cell, the resulting cross-correlation (γ) coefficient will be close to 1. Inversely a low correspondence between the counting of damaged buildings will result in a value γ of close 0. Note that we assign $\gamma = 0$ to the central pixel of windows for which grid counts in either set are zero, and for which the normalised correlation coefficient is undetermined.

A graphical impression of the overall correlation is shown in Fig. 6. The images on the left provide a simple tool to inspect pattern matching directly. They are colour composites, in which the correlation coefficient is combined with the grid densities for the two sets that we compare, in the red, green and blue colour planes, respectively. Where green colours dominate, the UWJ damage densities are not matched with corresponding pattern in either the CAS or OSM data sets. Similarly, blue colours show

Intercomparison and validation of building damage assessments

G. Lemoine et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



where either CAS or OSM is not matched with the UWJ analysis, primarily because of the more limited size of the aerial imagery. High correlation shows up as red dots, where small patches show good matches, and yellow and white colours, primarily in the densely built-up areas. The histogram of the correlation coefficients for the OSM set shows that overall correlation is rather low, with more than half of all 3×3 windows showing correlation coefficient lower than 0.5. The CAS results show a slightly better overall correlation than the OSM set.

5 Validation of the aerial based assessment with multi-source reference data

In this section, we review and analyze the results of different sample-based validation methods that were developed at an early stage to assess the accuracy of the PDNA point set (UWJ), before presenting a full validation exercise that was conducted recently after the release of the exhaustive field survey conducted by the Haitian Ministry of Public Works (MTPTC).

5.1 Review of sample-based validation conducted in previous studies

Several sample-based validation activities were set up to review the quality of the visual interpretation of building damage on the aerial imagery. The results of these validation activities are also useful to extrapolate the damage for lower damage grades using a set of statistical models. These validation methods differ in terms of the sampling designs, the sample sizes and the type of reference data acquired. The purpose here is not to describe the full methodologies and the detailed results, but to establish common patterns in the results obtained from these different methods. Interested readers may refer to the following references for a detailed description of the sample-based validation studies: Corbane et al. (2011b) for the ground survey following a stratified random sampling design, Ajmar et al. (2011) for the along roads track survey from

Low Cost Mobile Mapping System (LCMMS) and Booth et al. (2011) and Spence and Saito (2010) for the block-sampling using Pictometry data.

In the three sample-based validation studies, contingency matrices are used to measure the agreement/disagreement between the aerial based damage assessment (UWJ) and the different reference sources (ground-surveys, LCMMS georeferenced frames and Pictometry). Buildings with light or moderate damage states (grade 2) and substantial to heavy damage (grade 3) could not be established with any certainty from remotely sensed images as confirmed by Booth et al. (2011). Even with the Pictometry data, the interpreters encountered difficulties identifying grade 3 buildings. Therefore, it was decided to merge this damage grade with grades 0 and 1 for a more consistent comparison between the different datasets (i.e. field data versus remote sensing data).

Comparing the three accuracy assessments of the UWJ, it can be seen that the higher the sample size (N), the lower the overall accuracy (73% for $N = 4329$, 79% for $N = 1247$ and 88% for $N = 168$). However the contingency table using Pictometry data with a “moderate” sample size selected on the basis of an urban block sampling (Table 4) provides accuracy values close to those reached with a much larger sample size obtained from a ground-survey (Table 2). It is worth mentioning that both the sample collected from the ground survey and the sample used for pictometry interpretation were selected following a stratification by land use categories, whose purpose is to reduce the sampling variability by creating homogeneous groups of sampling units. The accuracy assessment using the LCMMS georeferenced frames is much more optimistic (Table 3). However, not only the sample size is very small (corresponding to 0.06% of the total number of buildings assessed on the aerial imagery) but also the sampling schema is very much biased: given the nature of LCMMS imagery, only buildings facing the roads were taken into account during the accuracy assessment.

When examining the accuracy per damage grade, it is possible to see a general agreement between the three sample-based validation studies in terms of user and producer accuracies for building damage grades 0–3. The main discrepancies are between the producer accuracy (error of omission) for grade 4 obtained with Pictometry

Intercomparison and validation of building damage assessments

G. Lemoine et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



data as a reference (64 %) and the values obtained with LCMMS and ground survey as references (25 and 30 %, respectively) . The biggest disagreement between the three accuracy assessments is for building damage grade 5, especially between the values obtained with the LCCMS as reference (user accuracy = 81 % and producer accuracy = 91 %) on one hand and those obtained with the ground survey and Pictometry on the other hand.

This comparative analysis between the three sample based validation works shows that Pictometry represents an interesting alternative for ground surveys for the validation of the detailed damage assessment given a reasonably distributed sample. It also suggests that, although the LCMMS is very practical tool for the collection of reference data on damaged buildings, the fact that only building facing accessible roads without fences or obstacles blocking the view from the car, could only be surveyed, limits its application for sample-based validation studies.

5.2 Validation based on exhaustive field survey

5.2.1 Damage statistics by type of land use

In this section, we extend the validation approaches discussed previously, with an analysis of the UWJ set versus the exhaustive field survey (MTPTC). According to the teams of structural engineers who led the exhaustive damage assessment program (Miyamoto et al., 2011), the majority of evaluated buildings (about 75 % of the total number of assessed buildings) were single-family residential structures. On the basis of this information, it is reasonable to consider that household-dwelling units are comparable to independent building units.

Due to the ambiguity of the color assignments in the MTPTC dataset (Sect. 2.3.4), it was decided to use instead the percentage of damage (0–1 %, 1–10 %, 10–30 %, 30–60 %, 60–100 % and 100 %) for the comparison with UWJ. In order to ensure the maximum semantic correspondence between the two point sets under comparison and due the low confidence in the assignments of damage grades 2 and 3 using remotely

Intercomparison and validation of building damage assessments

G. Lemoine et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



sensed data, only buildings assigned to damage levels 4 and 5 were considered from UWJ set and those assigned with damage percentages of 60–100 % and 100 % were selected from MTPTC. Hence, only buildings considered as heavily affected or destroyed were retained from each of two point datasets.

For the analysis a subset covering an area of 25 × 15 km and extending from Carrefour to Port-au-Prince was defined. A point-to-point comparison of the two survey results is only partly feasible, but does not result in very meaningful indicators. One of the issues is due to the placement of building labels: the UWJ aerial interpretation typically places the damage label at the center of what is perceived as the building roof. The MTPTC survey appears to put the damage label usually at the street-side, probably where the building entry is, and may have an intrinsic error due to the quality of the GPS, which should be in the order of, at least, 5 m. Since many building tend to be small, this leads to ambiguous matching between UWJ and MTPTC point location that refer to the same building. Furthermore, the MTPTC survey has more the character of a “household” survey. Thus, more than one household may be resident in a structure that is identified as a unique building (structure) in the UWJ analysis. Therefore, in many cases, there may not be a one-to-one relation.

To circumvent these problems, the point counts for grid cells of 50 × 50 m were taken as a reference for the validation: each grid cell of 50 × 50 m contained (i) the total count of buildings with damage percentages of 60–100 % and 100 % derived from the reference MTPTC ($\text{Count}_{\text{MTPTC}}$), and (ii) the total count of buildings assigned to damage levels 4 and 5 derived from the UWJ set ($\text{Count}_{\text{UWJ}}$), and (iii) the arithmetic difference in the counting between $\text{Count}_{\text{MTPTC}}$ and $\text{Count}_{\text{UWJ}}$ ($\text{Diff}_{\text{MTPTC-UWJ}}$).

Using the grid cells, it is relatively straightforward to inspect specific cases of relative under- and over-estimations between the two data sets. These are based on spatial queries in the data base in which all points and grid outlines have been stored. In this way, we can find grid cells that have a much larger (e.g. 5×) count of one set of points versus the other set. This analysis is very useful to provide an understanding of the underlying characteristics of the deviation between the sets.

Intercomparison and validation of building damage assessments

G. Lemoine et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Intercomparison and validation of building damage assessments

G. Lemoine et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Two examples are given in Fig. 7. On the left side of the figure, the example shows an overestimation of MTPTC surveyed points, with damage categories 60–100 % and 100 % with respect to UWJ damage points for categories 4 and 5. The main reason for this deviation is, apparently, that the damage artefacts registered in the MTPTC survey are likely to be related to internal damage, which is not shown in the aerial view, as little external signs of damage are visible. In the right figure, the opposite effect is shown. This is an area with extensive total destruction, visible as wide swaths of rubble. The MTPTC survey is aimed at an inventory of remaining structures, and has obviously not counted pre-existing structures that no longer exist.

The overall pattern is shown in Fig. 8 as a correlation analysis result, in a similar way as in Fig. 6. Overall, correlation is significantly higher, and more consistent than for the UWJ vs. CAS and UWJ vs. OSM analysis. Again, differences exist due to mutual exclusion, e.g. where UWJ grids contain significant point counts and MTPTC none (in green) and vice versa (in blue). However, a predominance of red and white color indicates good pattern matches.

The analysis of the differences in the number of heavily affected and destroyed between the two sets was disaggregated by type of land use. This land use layer was produced by ImageCat using pre- and post-earthquake imagery of Haiti, as part of the joint collaborative work effort by the UWJ team. Table 5a and b show the percentages of cells for which $\text{Count}_{\text{MTPTC}} = \text{Count}_{\text{UWJ}}$, $\text{Diff}_{\text{MTPTC-UWJ}} < 0$ (overestimation of the number of damaged buildings from the aerial assessment with respect to the field survey) and $\text{Diff}_{\text{MTPTC-UWJ}} > 0$ (underestimation of the number of damaged buildings from the aerial assessment with respect to the field survey). Table 5a represents the column percentages, while Table 5b represents the row percentages.

Overall on a total of 30 992 50 × 50 m cells that were analyzed, almost 53 % show a perfect agreement between the damage assessment obtained from aerial imagery and the field survey, while 21 % correspond to an overestimation of the damage from aerial imagery and 25 % to an underestimation. The highest disagreement was found mainly in shanty zones where it appears that the omission error was very large (for

almost 40% of cells, an underestimation of the heavily damaged and affected buildings has been detected). It is important to note that these results do not take into account the magnitude of the differences between the counting, which means that even if $\text{Diff}_{\text{MTPTC-UWJ}} = 1$, it is considered as an underestimation.

Figure 9 shows the overall number of heavily affected and damaged buildings per category of land use for both datasets considering all the 50×50 m cells under analysis (30 922 in total). From this graph, it can be seen that despite the high level of disagreement detected in the previous table for shanty zones, especially for the cases of underestimation (2090 cells), in terms of absolute values, the number of heavily affected and damaged buildings in the aerial point set is slightly higher than that obtained from the field survey (13 449 against 13 334). Besides, in almost all major land use categories, except in the industrial areas, the assessment from aerial imagery detects more damaged building than the field inventory. In the downtown, there is an agreement in the absolute number of damaged buildings between the two datasets ($\Sigma\text{Count}_{\text{MTPTC}} = 1112$ and $\Sigma\text{Count}_{\text{UWJ}} = 1082$).

On the basis of the results shown in Table 5 and Fig. 6, we can detect a dependency between the counting of the damaged buildings in the two points datasets and the type of land use. To confirm these findings, we conduct a Tobit analysis with the purpose of identifying the land use categories that account for the marked differences observed between the damage assessments obtained from MTPTC and UWJ. Since we are interested in analyzing here the absolute differences in the counts between MTPTC and UWJ for cells of 50×50 m ($|\text{Diff}_{\text{MTPTC-UWJ}}|$) Tobit was chosen instead of common Ordinary Least Squares regression analysis because it accounts for the differences that are equal to 0 (i.e. $|\text{Diff}_{\text{MTPTC-UWJ}}| = 0$ is equivalent to an agreement between the two point sets and, hence, is a valuable piece of information). The results of the Tobit analysis are summarized in Table 6. They show that the type of land use accounts for a significant amount of the observed variance in the counting of damaged buildings among MTPTC and UWJ. Seven of the eight land use categories contributed consistently to the regression model (commercial, downtown, industrial, open land, residential high density,

NHESSD

1, 1445–1486, 2013

Intercomparison and validation of building damage assessments

G. Lemoine et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



residential low density, shanty). The coefficients of the Tobit model indicate the importance of the effect of each land use category on the differences in the counting. We note large significant coefficients mainly for shanty, downtown and residential high density categories (underlined land use classes in Table 6), while residential low density, open land and industrial had a lower contribution to the regression. A mutual analysis of Fig. 9 and the results of Tobit, indicates both the agreement (in downtown) and disagreement (in residential high density and shanty) between UWJ and MTPTC are influenced by the type of land use: in highly populated areas, especially in shanty zones, the overestimation of building damage from aerial imagery could be simply a result of an incomplete field survey while, in downtown, where field surveys seem to have been conducted in a more systematic way, the damage assessment from aerial imagery matches very well the one obtained from the field. Despite the differences observed between the two datasets, the damage inventory from aerial imagery demonstrates a pattern and has a total size which can be considered as a reasonable accurate in the field.

5.2.2 Comparison of the estimated total floor area and the associated replacement cost

The results of the joint UNOSAT-World Bank-JRC damage analysis (UWJ) were shared with the Haitian government in support of the PDNA framework and, more specifically, as an input to the Damage and Loss Assessment (DaLA) for the housing sector (Kemper et al., 2010). It was the first time that the PDNA was based on a damage assessment produced with remotely sensed data. It is known that the physical damage caused by the earthquake impact creates losses in asset values that can be measured by the cost of repair or replacement (Committee on Assessing the Costs of Natural Disasters, 1999). Accordingly, the damage figures derived from remote sensing were translated into estimates of total floor areas to be repaired or replaced. Building floor area was estimated from very high resolution pre-earthquake satellite imagery for typical building footprint sizes by type of land use. Number of stories information by type of land use was collected from field surveys. The total floor area was calculated by

Intercomparison and validation of building damage assessments

G. Lemoine et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Intercomparison and validation of building damage assessments

G. Lemoine et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



5 multiplying the building footprint by the number of stories (details on the estimation of the number of stories and the floor area by land use type are provided in Corbane et al., 2011b and Ghosh et al., 2011). Floor area figures can be used in combination with the damage counts to estimate a total cost based on affected area and with knowledge
 10 of recovery cost per unit floor space, possibly diversified by land use type. Unit repair costs ($\$ \text{m}^{-2}$) for different land use categories were obtained from the Haitian Ministry of Social Affairs (Social Housing Promotion and Planning Institute). These costs range from 40 \$ per m^2 , for moderate repairs, to 500 \$ per m^2 for complete replacement (Damage grades 4 and 5) (Ghosh et al., 2011). We expect that the differences in the damage statistics per land use category obtained from UWJ and MTPTC would have an
 15 impact on the overall repair/replacement cost to heavily affected and destroyed buildings. To verify this assumption, the total floor area and associated repair/replacement costs were computed using the MTPTC dataset and compared to the remotely sensed estimates based on the UWJ dataset Table 7 gives an overview of the number of heavily
 20 affected and damaged buildings per category of land use in the area of interest (extending from Carrefour to Port-au-Prince), the derived floor areas and the related replacement/repair costs for both datasets.

The results show that despite the differences in the damage counts observed in the two datasets, the overall cost estimated from remote sensing is a good approximation
 25 of the costs obtained from the detailed field survey (2425 versus 2258 Million US\$). Knowing that the primary objective of the PDNA is to support the elaboration of the Recovery Framework – normally within five to six weeks following the onset of a disaster, the joint damage assessment from aerial imagery currently represents the best possible compromise between timeliness and accuracy. The findings obtained from the joint damage assessment (UWJ), were indeed delivered to the DaLA team in time for the New York Donor Conference held on 31 March 2010.

6 Concluding remarks

In this paper, the quality of damage assessments derived from different sources following the 2010 Haiti Earthquake was analyzed. The following conclusions are made with respect to the cost-effectiveness of the different assessments:

- The intercomparison of the satellite and the aerial based damage assessments, showed that the damage derived from satellite imagery (CAS and OSM datasets) was underestimated by a factor of eight, compared to the damage derived from aerial imagery. These results suggest that despite the fast availability of very high resolution satellite imagery for rapid generation of damage assessment maps, the spatial resolution of 0.5 m is not sufficient for an accurate interpretation of building damage as already observed in Lemoine (2010), Corbane et al. (2011a) and Voigt (2011).
- In this work, for the first, time, different validation approaches of the aerial based assessment were compared. The results highlight the agreement between the outputs of sampling designs that take into account the characteristics of the built-up areas such as the density of building and the land use. For the validation of the aerial based damage assessment, pictometry can represent an interesting alternative for field surveys especially for extensive damage, as in the case of Haiti with all the obstacles that challenge in-situ data collection (presence of rubble, security issues, access to remote areas, absence of a quality control of the collected data, etc.). If the sensors can be flown, then pictometry is an interesting alternative source for the provision of validation data (Ghosh et al., 2011).
- Considering only heavily affected and destroyed buildings, the validation of the detailed aerial based damage assessment using the extensive field inventory, demonstrated that, overestimation of the damage from aerial imagery in the commercial and low density residential areas is compensated by underestimation in the industrial areas. This resulted in overall comparable damage figures with

Intercomparison and validation of building damage assessments

G. Lemoine et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Intercomparison and validation of building damage assessments

G. Lemoine et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



almost 40 000 affected buildings obtained from aerial data against 37 914 observed in the field. These results show that the identification of building damage from aerial imagery, although suffering from several limitations due to the nadir imagery, provides a realistic estimate of the spatial pattern and intensity of damage. While there is no doubt that such a detailed damage assessment depends very much on the timeliness of the aerial images acquisitions and requires a considerable effort of visual interpretation as an element of compromise; it remains the best trade-off in terms of required quality and timeliness for producing detailed damage assessments over large affected areas compared to satellite based assessments (insufficient quality) and exhaustive field inventories (too slow).

References

- Ajmar, A., Balbo, S., Boccardo, P., Giulio Tonolo, F., Piras, M., and Prinic, J.: A Low-Cost Mobile Mapping System (LCMMS) for field data acquisition: a potential use to validate aerial/satellite building damage assessment, *Int. J. Digital Earth*, 1–21, doi:10.1080/17538947.2011.638991, 2011.
- Booth, E., Saito, K., Spence, R., Madabhushi, G., and Eguchi, R. T.: Validating assessments of seismic damage made from remote sensing, *Earthq. Spectra*, 27, S157–S177, 2011.
- Cochran, W.: *Sampling Techniques*, John Wiley & Sons, New York, 74–76, 1977.
- Committee on Assessing the Costs of Natural Disasters, N. R. C.: *The Impacts of Natural Disasters: A Framework for Loss Estimation*, National Academy Press, Washington DC, 80 pp., 1999.
- Corbane, C., Carrion, D., Lemoine, G., and Broglia, M.: Comparison of damage assessment maps derived from very high spatial resolution satellite and aerial imagery produced for the Haiti 2010 earthquake, *Earthq. Spectra*, 27, S199–S218, 2011a.
- Corbane, C., Saito, K., Dell’Oro, L., Gill, S., Piard, B., Huyck, C., Kemper, T., Lemoine, G., Spence, R., Krishnan, R., Bjorgo, E., Senegas, O., Ghesquiere, F., Lallemand, D., Evans, G., Gartley, R., Toro, J., Ghosh, S., Svekla, W., Adams, B., and Eguchi, R.: A Comprehensive Analysis of Building Damage in the January 12, 2010 Mw7 Haiti Earthquake using High-Resolution Satellite and Aerial Imagery, *Photogr. Eng. Remote Sens.*, 77, 977–1009, 2011b.

Intercomparison and validation of building damage assessments

G. Lemoine et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



- 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), *Earthq. Spectra*, 27, S179–S198, 2011.
- Gonzalez, R. C. and Woods, R. E.: *Digital Image Processing*, 3rd Edn., Addison-Wesley, Massachusetts, 954 pp., 1992.
- Group URD: Real-time evaluation of humanitarian action supported by DG ECHO in Haiti, European Commission, 72 pp., 2011.
- Grünthal, G. and Levret, A.: L'Echelle Macrosismique Européenne, European Macroseismic Scale 1998 (EMS-98), *Cahiers du Centre Européen de Géodynamique et de Séismologie* 19, Centre Européen de Géodynamique et de Séismologie, Luxembourg, 103 pp., 2001.
- Hussain, E., Ural, S., Kim, K., Fu, C. S., and Shan, J.: Building extraction and rubble mapping for city Port-au-Prince post-2010 earthquake with GeoEye-1 imagery and lidar data, *Photogr. Eng. Remote Sens.*, 77, 1011–1023, 2011.
- Kaya, G. T., Musaoglu, N., and Ersoy, O. K.: Damage assessment of 2010 Haiti earthquake with post-earthquake satellite image by support vector selection and adaptation, *Photogr. Eng. Remote Sens.*, 77, 1025–1035, 2011.
- Kemper, T., Lemoine G., Corbane C., Gill, S., Eguchi, R., Bjoro, E., and Dell'Oro, L.: Building Damage Assessment Report- Post Disaster Needs Assessment and Recovery Framework, *Ispra*, 38 pp., 2010.
- Kerle, N.: Satellite-based damage mapping following the 2006 Indonesia earthquake-How accurate was it?, *Int. J. Appl. Earth Observation Geoinf.*, 12, 466–476, 2010.
- Lemoine, G.: Validation of building damage assessments based on post-Haiti 2010 earthquake imagery using multi-source reference data, *Validation of Geo-Information Products for Crisis Management*, *Ispra*, 33–34, 2010.
- Li, P., Xu, H., and Song, B.: A novel method for urban road damage detection using very high resolution satellite imagery and road map, *Photogr. Eng. Remote Sens.*, 77, 1057–1066, 2011.
- Malinverni, E. S.: Change detection applying landscape metrics on high remote sensing images, *Photogr. Eng. Remote Sens.*, 77, 1045–1056, 2011.

Intercomparison and validation of building damage assessments

G. Lemoine et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Miyamoto, H. K., Gilani, A. S. J., and Wong, K.: Massive damage assessment program and repair and reconstruction strategy in the aftermath of the 2010 Haiti earthquake, *Earthq. Spectra*, 27, S219–S237, 2011.

Pesaresi, M., Gerhardinger, A., and Haag, F.: Rapid damage assessment of built-up structures using VHR satellite data in tsunami-affected areas, *Int. J. Remote Sens.*, 28, 3013–3036, 2007.

Shankar, R.: Accuracy Assessment of Post-Earthquake Building Damage Classification Haiti-Master thesis, International Health, Immunology & Microbiology Faculty of Health Sciences, University of Copenhagen, 2010.

Spence, R. J. S. and Saito, K.: Port-au-Prince Earthquake Damage Assessment using Pictometry, A report for ImageCat Inc. Cambridge Architectural Research Ltd., available at: <http://www.carltd.com/downloads/Haiti%20Pictometry.pdf> (last access: 12 March 2011), 2010.

Taucer, F., Corbane, C., and Gerhanrdinger, A.: Haiti earthquake of 12 January 2010: Field damage assessment, data collection and analysis, *Proceedings of the 8th Congresso De Sismologia E Engenharia Sísmica*, 2010, 2010.

Tiede, D., Lang, S., Fureder, P., Holbling, D., Hoffmann, C., and Zeil, P.: Automated damage indication for rapid geospatial reporting, *Photogr. Eng. Remote Sens.*, 77, 933–942, 2011.

Voigt, S., Kemper, T., Riedlinger, T., Kiefl, R., Scholte, K., and Mehl, H.: Satellite image analysis for disaster and crisis-management support, *IEEE Trans. Geosci. Remote Sens.*, 45, 1520–1528, 2007.

Voigt, S., Schneiderhan, T., Twele, A., Gähler, 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.

Intercomparison and validation of building damage assessments

G. Lemoine et al.

Table 1. Summary statistics for UWJ, CAS and OSM datasets based on 50×50 m grid cells. The second column shows the total number of points, graded 4 or 5, for the area of interest, the third column the maximum occurring in a grid cell of 50×50 m, the fourth and fifth column show mean and standard deviation, respectively, for those grid cells that have at least one damage point.

| u | $\sum u$ | $\max(u)$ | $\mu(u[u > 0])$ | $\sigma(u[u > 0])$ |
|-----|----------|-----------|-----------------|--------------------|
| UWJ | 39 814 | 68 | 3.75 | 4.57 |
| CAS | 4712 | 7 | 1.18 | 0.50 |
| OSM | 5100 | 11 | 1.32 | 0.74 |

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Intercomparison and validation of building damage assessments

G. Lemoine et al.

Table 4. Confusion Matrix Comparing the pictometry interpretation and the UWJ dataset of 1247 buildings in Port-Au-Prince (Table taken and modified from Corbane et al., 2011b).

| | | Aerial Photo interpretation (UWJ) | | |
|---------------------------|-----------------------|-----------------------------------|---------|---------|
| | | Grades 0–3 | Grade 4 | Grade 5 |
| Pictometry interpretation | Grades 0–3 | 87.84 | 49.49 | 21.65 |
| | Grade 4 | 7.34 | 23.23 | 11.34 |
| | Grade 5 | 4.82 | 27.27 | 67.01 |
| | Total | 100 | 100 | 100 |
| Overall Acc. = 79 % | User accuracy (%) | 90 | 20 | 68 |
| | Producer accuracy (%) | 88 | 64 | 32 |

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Intercomparison and validation of building damage assessments

G. Lemoine et al.

Table 5. The percentages of 50×50 m grid cells for which $\text{Count}_{\text{MTPTC}} = \text{Count}_{\text{UWJ}}$, $\text{Diff}_{\text{MTPTC-UWJ}} < 0$ (overestimation of the number of damaged buildings from the aerial assessment with respect to the field survey) and $\text{Diff}_{\text{MTPTC-UWJ}} > 0$ (underestimation of the number of damaged buildings from the aerial assessment with respect to the field survey). (a) shows the values in the form of column percentages while (b) shows them as row percentages.

| (a) | | | | |
|--------------------------|--|--|--|--|
| Landuse | %($\text{Count}_{\text{MTPTC}} = \text{Count}_{\text{UWJ}}$) | % $\text{Diff}_{\text{MTPTC-UWJ}} < 0$ | % $\text{Diff}_{\text{MTPTC-UWJ}} > 0$ | |
| Agriculture | 0.02 | 0 | 0.04 | |
| Commercial | 5.38 | 8.89 | 5.51 | |
| Downtown | 0.61 | 2.04 | 2.17 | |
| Industrial | 3.57 | 2.76 | 4.00 | |
| Openland | 0.37 | 0.10 | 0.34 | |
| Residential High Density | 8.79 | 17.97 | 18.13 | |
| Residential Low Density | 70.09 | 45.60 | 42.60 | |
| Shanty | 11.17 | 22.64 | 27.21 | |
| total | 100 | 100 | 100 | |

| (b) | | | | |
|--------------------------|--|--|--|-------|
| Landuse | %($\text{Count}_{\text{MTPTC}} = \text{Count}_{\text{UWJ}}$) | % $\text{Diff}_{\text{MTPTC-UWJ}} < 0$ | % $\text{Diff}_{\text{MTPTC-UWJ}} > 0$ | Total |
| Agriculture | 57.14 | 0.00 | 42.86 | 100 |
| Commercial | 46.55 | 31.22 | 22.28 | 100 |
| Downtown | 24.88 | 33.74 | 41.38 | 100 |
| Industrial | 54.47 | 17.05 | 28.48 | 100 |
| Openland | 64.89 | 7.45 | 27.66 | 100 |
| Residential High Density | 35.81 | 29.71 | 34.48 | 100 |
| Residential Low Density | 64.62 | 17.05 | 18.33 | 100 |
| Shanty | 33.80 | 27.78 | 38.42 | 100 |

Title Page

Abstract Introduction

Conclusions References

Tables Figures

⏪ ⏩

◀ ▶

Back Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Intercomparison and validation of building damage assessments

G. Lemoine et al.

Table 6. Results of the Tobit analysis performed on the absolute differences in the counts of damaged buildings between MTPTC and UWJ for cells of 50×50 m ($\text{Diff}_{\text{MTPTC-UWJ}}$) with landuse categories as explanatory variable.

| Landuse category | Coefficient | Std. Error |
|---------------------------------|-------------|------------|
| Agriculture | 5.117455 | 3.092541 |
| Commercial | 7.845463*** | 0.125140 |
| <i>Downtown</i> | 9.125413*** | 0.233125 |
| Industrial | 6.881543*** | 0.161008 |
| Open land | 5.892216** | 1.164927 |
| <i>Residential High Density</i> | 8.900220*** | 0.095819 |
| Residential Low Density | 5.852702*** | 0.068558 |
| <i>Shanty</i> | 9.288613*** | 0.088672 |

Significance codes: ** $P \leq 0.01$ – *** $P \leq 0.001$.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Intercomparison and validation of building damage assessments

G. Lemoine et al.

Table 7. Comparison between UWJ and MTPTC estimates of the total floor area and associated replacement/repair costs grouped by landuse category.

| | Number of heavily affected buildings | | Floor area per damage class (in m ²) | | *Replacement/repair costs (Million US\$) | |
|--------------------------|--------------------------------------|------------------------------|--|------------------------------|--|------------------------------|
| | UWJ (grades 4 and 5) | MTPTC (60–100% of damage) | UWJ (grades 4 and 5) | MTPTC (60–100% of damage) | UWJ (grades 4 and 5) | MTPTC (60–100% of damage) |
| Agricultural | 0 | 1 | 0 | 82 | 0.00 | 0.04 |
| Commercial | 3944 | 3002 | 615264 | 468312 | 307.63 | 234.16 |
| Downtown | 1082 | 1112 | 187727 | 192932 | 93.86 | 96.47 |
| Industrial | 747 | 1159 | 99949 | 155074 | 49.97 | 77.54 |
| Residential high density | 9108 | 8926 | 1204078 | 1180017 | 602.04 | 590.01 |
| Residential low density | 11548 | 10226 | 1893872 | 1677064 | 94694 | 838.53 |
| Shanty | 13449 | 13334 | 849977 | 842709 | 424.99 | 421.35 |
| Grand total | 39878 | 37760 | 4850867 | 4516190 | 2425.43 | 2258.10 |

*Unit Cost = 500 US \$ per m².

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



NHESSD

1, 1445–1486, 2013

Intercomparison and validation of building damage assessments

G. Lemoine et al.

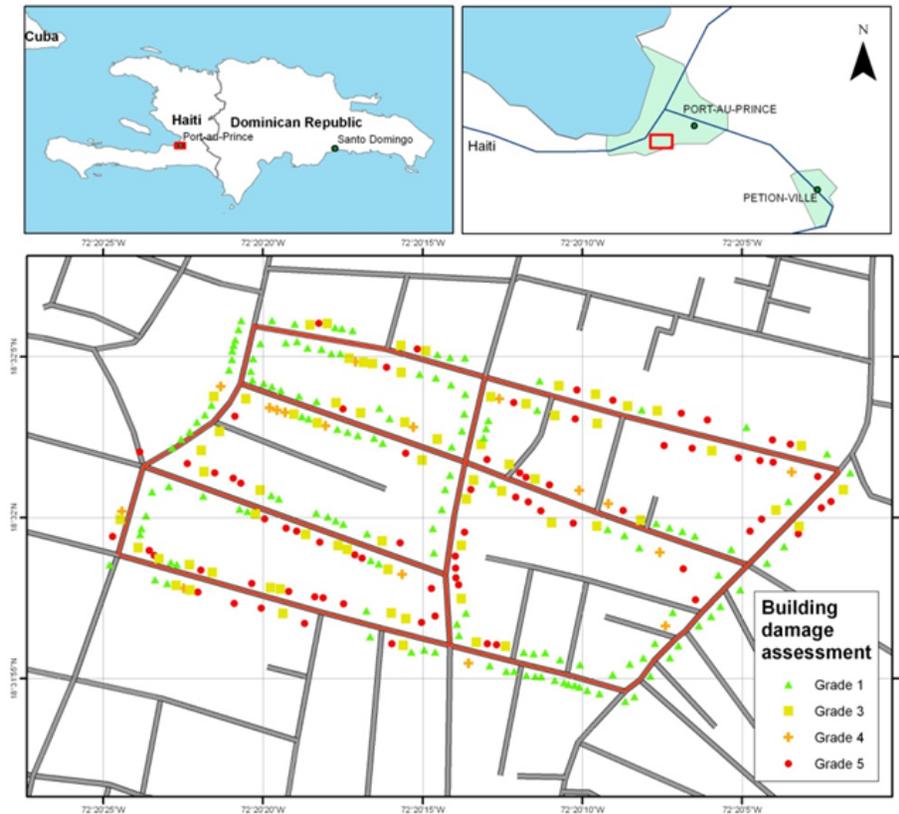


Fig. 1. Car survey ground track (red lines) and building damage assessment (based on high resolution aerial imagery) considered for the validation test (Ajmar et al., 2011).

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Intercomparison and validation of building damage assessments

G. Lemoine et al.



Fig. 2. Example of a building suffering from a soft storey failure observed on the aerial imagery **(a)** and on the oblique Pictometry image **(b)**. The lateral view provided by Pictometry data allows a better detection of this type of damage.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

Intercomparison and validation of building damage assessments

G. Lemoine et al.



Fig. 3. Examples illustrating the quality problems identified in the extensive field survey. Red squares correspond to partly or completely destroyed or dangerous buildings; yellow squares correspond to unstable buildings. **(a and b)** Overlooked heavily damaged houses; **(c)** Inconsistency in the labeling within the same area, with several labels assigned to the same building close to unlabeled buildings; **(d)** Buildings surveyed from the road. Each image fragment corresponds to an area of approximately 200 by 200 m (projection: UTM 18 North (WGS84)).

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



NHESSD

1, 1445–1486, 2013

Intercomparison and validation of building damage assessments

G. Lemoine et al.

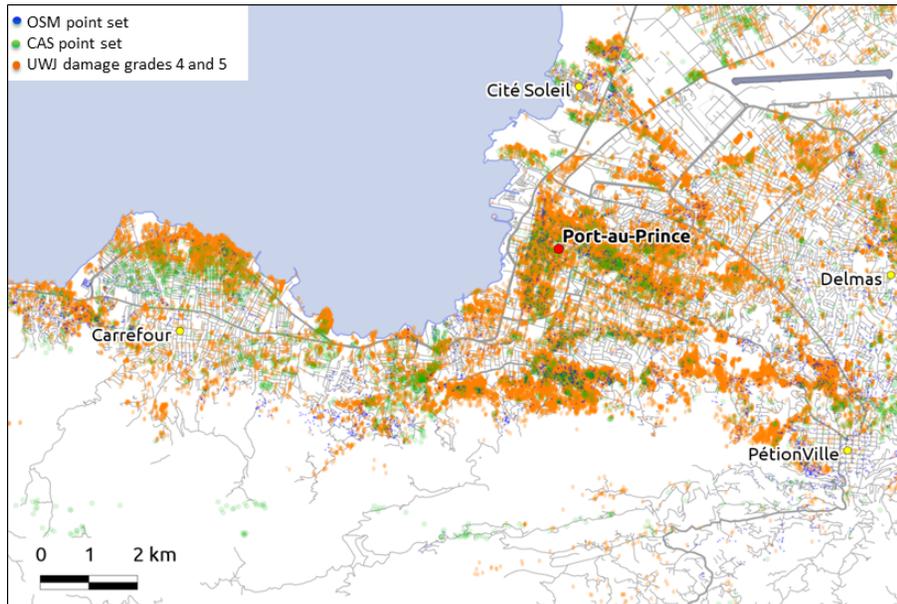


Fig. 5. Spatial distribution of the two satellite derived damage assessments (OSM and CAS point sets represented with blue and green dots respectively) and the aerial derived damage assessment (UWJ -damage grades 4 and 5 represented with orange dots) (© OpenStreetMap contributors, CC-BY-SA).

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



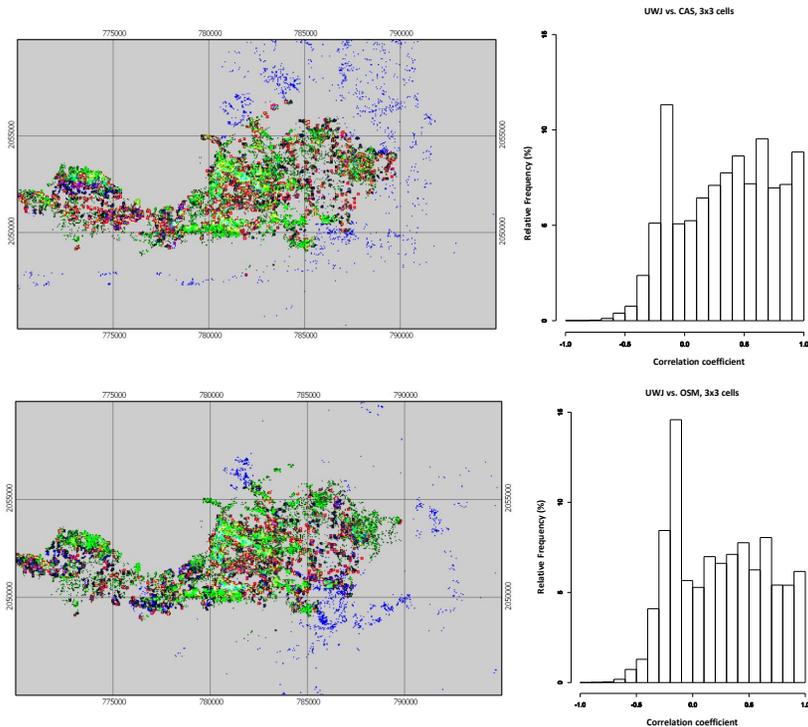


Fig. 6. Graphical representation of the correlation results. The images (projected to UTM 18 (WGS84) North) on the left overlay the local 3×3 normalised correlation coefficient (in red) with the UWJ density (in green) and (top) CAS density and (bottom) OSM density (in blue), respectively. Highly correlated grid counts show up as red (if isolated), white and yellow colours. Green colours show areas where UWJ damage points are not matched with either CAS or OSM damage points. Blue colours show the opposite. The histograms on the right show the distribution of the correlation coefficients for both set comparisons.

Intercomparison and validation of building damage assessments

G. Lemoine et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Intercomparison and validation of building damage assessments

G. Lemoine et al.

● UNOPS surveyed points (partly or completely destroyed or dangerous buildings)

▲ UWJ damage points (damage grades 4 and 5)



Fig. 7. Examples of (left) overestimation of MTPTC surveyed points (yellow circles) versus UWJ damage points (red triangles) for the highly damaged and destroyed categories and (right) underestimation. The background image is © Google Earth, 2010.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Intercomparison and validation of building damage assessments

G. Lemoine et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

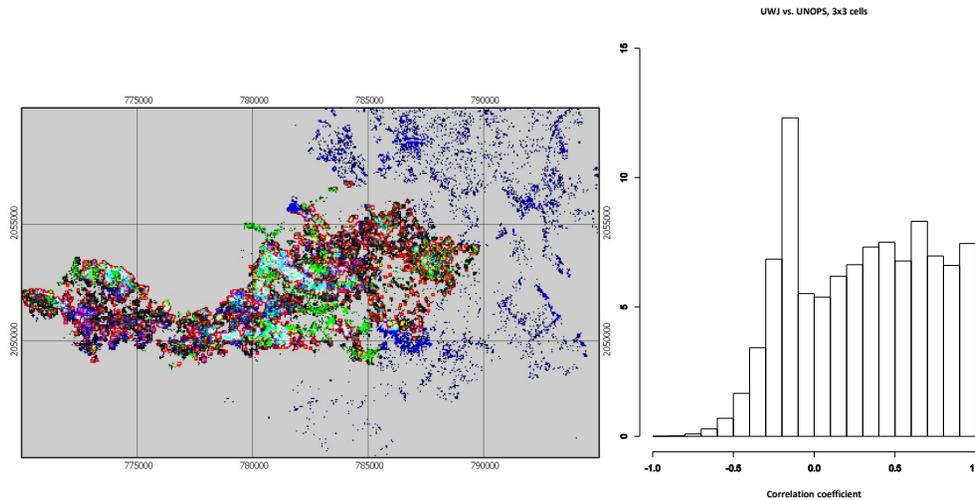


Fig. 8. Graphical representation of the correlation results between the UWJ and MTPTC sets. For an explanation, see Fig. 6.

Intercomparison and validation of building damage assessments

G. Lemoine et al.

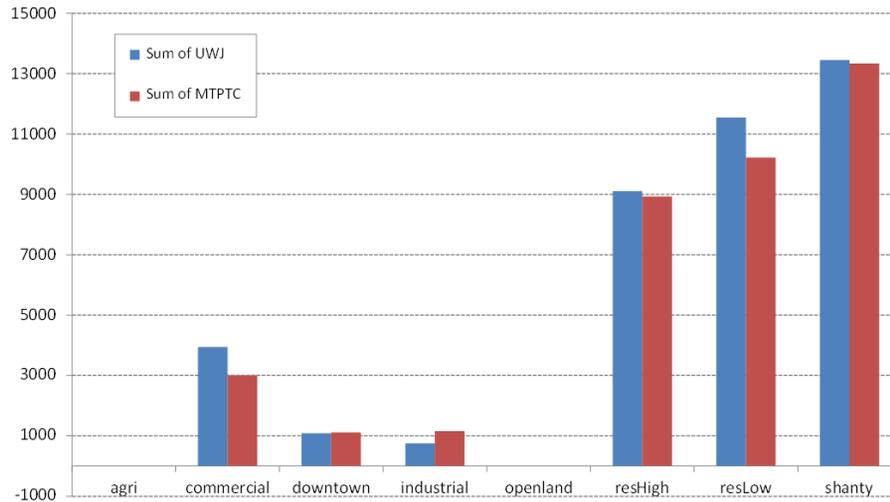


Fig. 9. Overall number of heavily affected and damaged buildings per category of land use for both datasets considering all the 50×50 m cells under analysis (30 922 in total).

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

