GROUNDTRUTHING OPEN STREET MAP BUILDING DAMAGE ASSESSMENT

HAIYAN TYPHOON – THE PHILIPPINES

Final Assessment Report

April 2014

American Red Cross

USAID FROM THE AMERICAN PEOPLE

REACH Informing more effective humanitarian action

CONTENTS

SUMMARY

At 10:00 on 6 November 2013, Typhoon Haiyan (named Yolanda locally) entered the Philippines Area of Responsibility (PAR), bringing with it severe damage across the Central Philippines. As of 17 December, the Philippines Disaster Response Operations Monitoring and Information Centre (DROMIC) had reported 1,127,041 houses as having been damaged, of which 548,793 were totally destroyed by the typhoon. These figures were largely estimates, however, based on initial information provided by municipal officials.

Given the scale of the damage and the wide geographic scope, accurate damage figures were greatly sought after in the days following the typhoon to facilitate response area prioritization. In the initial days of the response, the Humanitarian OpenStreetMap Team (HOT OSM) rapidly mapped the base infrastructure (roads, buildings, populated places, etc.) in the typhoon-affected area. In response to the need for more accurate damage figures, the American Red Cross (ARC) mobilized HOT OSM and their volunteer network to build on this base data and crowd-source remote damage assessments based on post-disaster satellite imagery. This was conducted in parallel with the Shelter Cluster Rapid Needs Assessment being implemented by REACH to provide information on damage and needs at the household level.

To gauge the accuracy of the crowd-sourced damage assessments, REACH and ARC conducted a study comparing enumerated field damage assessments with the remote damage assessments conducted by the OSM community. The assessment found an overall accuracy rate of 36 per cent with general underrepresentation of damage when compared with field assessment data. A key outlier in these results was Tacloban City which was the main contributor to a 134 per cent overrepresentation of completely damaged buildings assessment-wide.

This assessment found that current satellite imagery does not allow for the detailed analysis required for damage assessments in humanitarian contexts. The current OSM platform also does not currently minimise these imagery limitations through its coordination or management processes, leaving room for improvements in the analysis methods used while the satellite imagery technology advances.

The conclusions and recommendations are intended to inform contributors and developers of crowd-sourcing platforms, as well as the humanitarian community at large, contributing to a dialogue about the ways to capitalize on the current methods and improve the way in which they are used in humanitarian settings. The results of this research will be shared with a wide range of stakeholders through the publication of this report as well as presentations at conferences. This research was conducted as a collaborative endeavour of REACH and ARC and financially supported by the United States Office for Foreign Disaster Assistance (OFDA).

ABOUT REACH

REACH is a joint initiative of two international non-governmental organizations – ACTED and IMPACT Initiatives – and the UN Operational Satellite Applications Programme (UNOSAT). REACH was created in 2010 to facilitate the development of information tools and products that enhance the capacity of aid actors to make evidence-based decisions in emergency, recovery and development contexts. All REACH activities are conducted in support to and within the framework of inter-agency aid coordination mechanisms. For more information, please visit: www.reach-initiative.org. You can write to us at: geneva@reach-initiative.organd follow us @REACH_info.

FIGURES AND MAPS

ACRONYMS

ARC	American Red Cross
EEFIT	Earthquake Engineering Field Investigation Team
GEO-CAN	A worldwide network of scientists tasked with assessing impacts of humanitarian disasters
HOT OSM	Humanitarian OpenStreetMap Team
JOSM	Java OSM Editor
OCHA	United Nations Organisation for the Coordination of Humanitarian Affairs
OFDA	United States Office of Foreign Disaster Assistance
OSM	OpenStreetMap
UAV	Unmanned Aerial Vehicle

GEOGRAPHIC CLASSIFICATIONS

Region	Highest form of governance below the national level
Province	Second highest form of governance comprised of multiple municipalities
Municipality	A collection of barangays that comprise a broader 'city'
Barangay	An area formed of 10,000 voters; the lowest administrative boundary
Sitio / Purok	Neighbourhood or area that is informal and not classified for administrative purposes

INTRODUCTION

At 10:00 on 6 November 2013, Typhoon Haiyan (named Yolanda locally) entered the Philippines Area of Responsibility (PAR). The typhoon intensified as it entered the Eastern Visayas region, first making landfall over Guiuan, Eastern Samar province, on 8 November, at 04:40. By 08:00 on 8 November the typhoon had made landfall six times across the Central Philippines and continued to weaken over the West Philippine Sea. Typhoon Yolanda left the PAR on 9 November at 15:30.

A total of 9,073,804 individuals, across 9,303 barangays, in 536 municipalities across the Central Philippines were identified by the Government of the Philippines as having been affected by Typhoon Yolanda. Of the affected population, a total of 1,910,547 individuals were displaced by Yolanda; with 422,290 people displaced to formal evacuation centres, and 1,488,257 to other locations. As of 17 December 2013, the Philippines Disaster Response Operations Monitoring and Information Centre (DROMIC) had reported 1,127,041 houses as having been damaged, of which 548,793 were totally destroyed by the typhoon. These figures were largely estimates, however, based on initial information provided by municipal officials.

Given the scale of the damage and the wide geographic scope, accurate damage figures were greatly sought after in the days following the typhoon to facilitate response area prioritization. In response to this initial gap in critical information, the Humanitarian OpenStreetMap Team (HOT OSM)[1] was mobilized to conduct crowd-sourced remote damage assessments based on satellite imagery. HOT OSM has been involved in creating high-quality geographic base data in Haiti, the Philippines and other smaller disasters over the past three years and the Typhoon Haiyan response provided an opportunity to continue building on this resource.

The potential utility of remote sensing imagery and rapid GIS-based mapping in humanitarian responses relies on the accuracy of these techniques. Recent studies from other emergencies have questioned the current capacity of these tools to deliver the levels of accuracy needed, but have acknowledged that these levels can be improved with further research, development and standardization for the humanitarian context[2].

This assessment sought to address some of these questions of accuracy by comparing remote damage assessment findings with field-level damage assessments and to identify any differences in accuracy. The assessment also aimed to assess the ability of crowd-sourced platforms to go beyond providing only base data by creating information about building-level damage. The conclusions and recommendations are intended to inform contributors and developers of crowd-source platforms as well as the humanitarian community at large, contributing to a dialogue about the how to capitalize on the present tools and improve the way in which they are used in humanitarian settings.

This research was conducted as a collaborative endeavour of the REACH Initiative and the American Red Cross (ARC). Building on its existing relationship with the Global Shelter Cluster and its expertise in field emergency damage assessments, REACH provided technical and methodological guidance and support in the field. In parallel, ARC coordinated activities with HOT OSM to crowd-source remote damage assessments and develop a data model for OSM. These efforts were financially supported by the United States Office for Foreign Disaster Assistance (OFDA) and conducted in parallel with the Shelter Cluster Rapid Needs Assessment[3].

[1] http://hot.openstreetmap.org/

[2] Shankar, Ravi. Accuracy of Post-Earthquake Building Damage Classification in Haiti. (Copenhagen 2010); EEFIT. The Haiti Earthquake of 12 January 2010: A Field Report. (Cambridge 2010); Cambridge Architectural Research Ltd. The Use of Remote Sensing for Building Damage Assessment Following the 22 Feb 2011 Christchurch Earthquake: the GEOCAN Study and its Validation – Draft for Comment. (Cambridge 2012)

[3] https://www.sheltercluster.org/Asia/Philippines/Typhoon%20Haiyan%202013/Documents/Haiyan%20Typhoon%20Shelter-WASH_assessment_Final%20Report_validated_formatted.pdf

ASSESSMENT METHODOLOGY

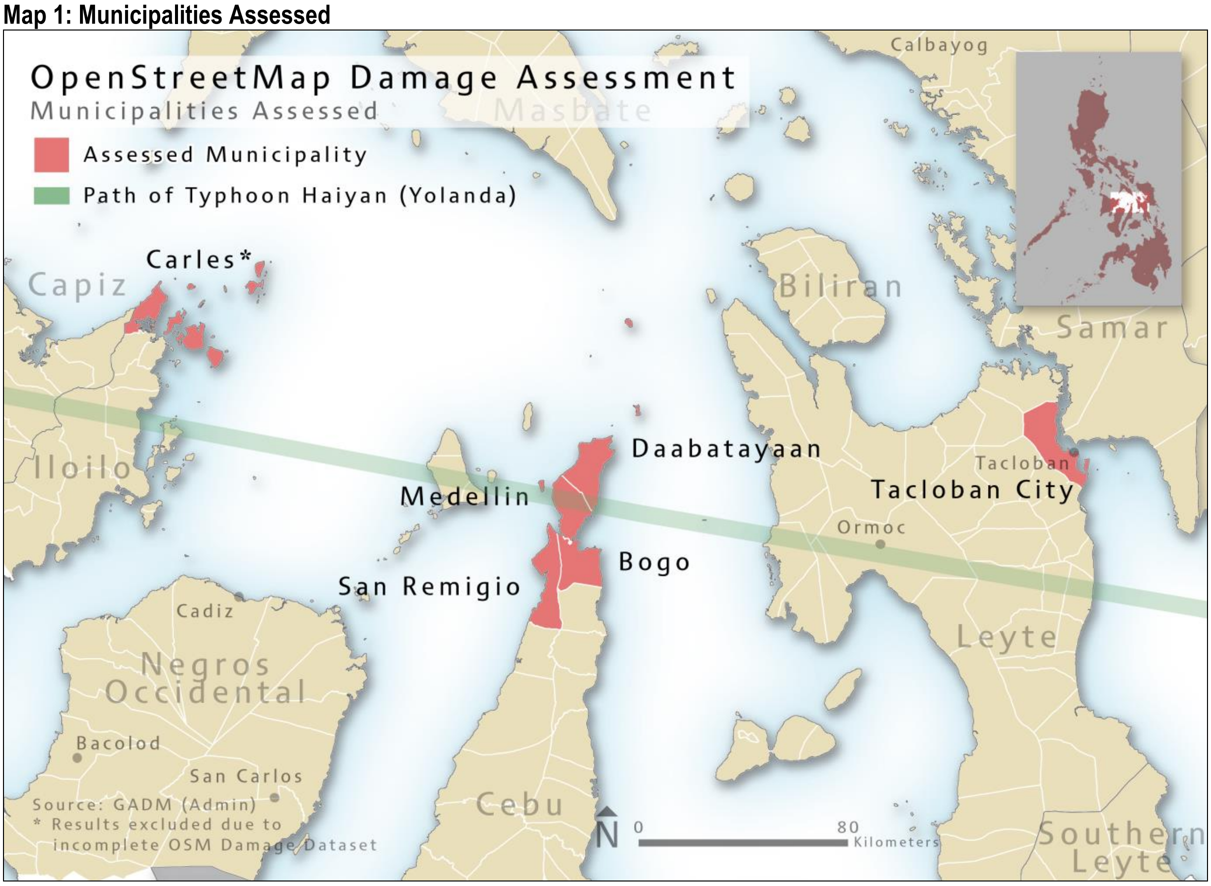
The following section will outline the methods and techniques used to collect and compare data from two datasets: (1) crowd-sourced damage assessments using satellite imagery from the OSM community and (2) observational field assessments using enumerators trained and managed by the REACH/ARC assessment team. Data collected using these methods were compared in order to quantify differences in validity rates. In addition, two rounds of satellite imagery desk reviews were conducted to understand the accuracy of building damage tags by OSM contributors in relation to the satellite imagery that was used for these remote assessments.

FIELD ASSESSMENT

Sampling

In order to compare remote and field-level damage assessments, five of the most frequently remotely assessed municipalities were selected in the area most highly affected by the typhoon. These municipalities had the highest density of building damage assessments conducted by OSM contributors using imagery services provided by the U.S. State Department's Humanitarian Information Unit. A final municipality, Carles, was selected as a control due to the logistical advantages of conducting the survey in the same area as the concomitant REACH-implemented Shelter Cluster Rapid Needs Assessment.

Map 1: Municipalities Assessed



Groups of buildings within these municipalities were randomly selected from a hand-made polygon layer of a sufficiently dense[4] cluster of buildings within each municipality. Points were randomly created within this polygon layer using GIS software and then centred over the greatest number of buildings within approximately 50 meters of the original point location. An atlas of maps for each municipality was then created at approximately 1:2500 scale in rural municipalities (Bogo, San Remigio, Medellin, Carles) and approximately 1:1000 scale in Tacloban City. The adjustment for Tacloban City accounted for the greater density of buildings and avoided illegibility resulting from overlapping OSM IDs on the printed reference maps used by enumerators.

Table 1: Number of Buildings Assessed per Municipality

Municipality	Number of Buildings Assessed
Bogo	158
Carles	171
Daanbatayaan	293
Medellin	128
San Remigio	248
Tacloban City	342
TOTAL	**1340**

Building Classification

The field assessment consisted of a classification system based on the scale developed for the Shelter Cluster Rapid Needs Assessment. Enumerators were trained to use the four-level classification system for each building observed as part of the assessment. An enumerator guidance sheet detailed the specific criteria for each classification type to ensure consistency (see Annex A).

Table 2: Field Assessment Classification System

No damage	no observed structural damage
Partial damage	repairable damage to windows, foundations, ceilings inside buildings
Major damage	repairable damage to the roof rendering segments of the building uninhabitable
Completely destroyed	unrepairable structural damage

Maps containing OSM building data and high resolution satellite imagery were used by the field teams to identify OSM data points on the ground. Additionally, enumerators used these maps to record which buildings they assessed, noting the OSM ID on the Android based data collection system. Buildings previously identified in the OSM system were displayed as plain semi-translucent rectangles on maps, identified only by its OSM ID. This ensured enumerators were free from being influenced by the OSM damage assessment classification when conducting their own independent field assessments. Teams were sent to the field with these maps and a mobile data collection device. They were asked to assess all buildings in the map area and to note the OSM ID from the map when entering the damage assessment data on the mobile device.

The field assessment was conducted using an assessment tool built on the Android smartphone based Open Data Kit (ODK) platform. This platform significantly improves data quality by: (a) reducing human error as a result of loss of forms, data collection mistakes, and data entry mistakes thus improving the accuracy of collected data; (b) increasing the speed at which mapping products and analytical reports can be produced through reducing data cleaning time and removing the time for data entry; and (c) ensuring the protection of data as a result of completed forms being removed from the data collection tool upon upload to the centralised database.

[4] "Sufficiently dense" is defined as at least 20 buildings located within the viewframe of the maps provided to enumerators.

Limitations

There were three key limitations with the field assessment that may have affected the accuracy of the data collected. While these limitations did exist, every effort was made to minimize their effect by accounting for potential error in the results.

1. Limited technical capacity of enumerators: The enumerators were university students hired locally to conduct the assessment and did not necessarily have any background in damage assessments or technical expertise in construction or engineering. REACH and ARC provided comprehensive trainings in order to ensure that enumerators were able to correctly and consistently classify building damage.

2. Different classification system from OSM: The classification system used for the field assessment consisted of four categories, whereas the classification system used by OSM assessors was three categories. The methodology section for the imagery desk review below discusses how these two systems were reconciled, but there is a possibility that given the different number of steps in the scale, buildings may have been classified differently. Some studies have concluded that the fewer levels within a classification system, the more accurate the classification becomes[5]. Thus, the four category classification system used by the field assessment may have led to less accurate findings. REACH and ARC attempted to minimise this effect through close management and oversight of enumerators in the field.

3. Timing of field assessment: The field assessment began 20 days after the typhoon, while the OSM contributors remotely assessed the building damage 7-10 days after the disaster. An unidentified proportion of initially damaged building could have been repaired or reconstructed in the 10-13 day window between the OSM remote assessment and the field assessment.

IMAGERY DESK REVIEW

Two rounds of imagery desk reviews were conducted. The first round was conducted prior to the release of the interim report for this assessment[6], while the second review was conducted prior to the release of this report to confirm the findings of the first review with a larger sample. Both desk reviews used three layers of data to review the imagery used during the remote damage assessments:

1. Building polygons that had been remotely assessed, styled in a red-yellow-green “traffic light” pattern indicating which were remotely tagged destroyed, damaged or without damage, respectively;
2. Bing imagery showing the affected area prior to the disaster;
3. Digital Globe imagery from the U.S. State Department Humanitarian Information Unit showing the affected area after Typhoon Haiyan. One image showed the Tacloban metropolitan area within five days of the typhoon making landfall and one image showed Northern Cebu within two weeks of the typhoon making landfall.

Buildings were analysed visually to better understand the comparison across all three datasets. Visual observations were recorded and integrated into this report.

[5]Shankar, Ravi. Accuracy of Post-Earthquake Building Damage Classification in Haiti. (Copenhagen 2010).
[6] http://americanredcross.github.io/OSM-Assessment/

Sampling

Buildings located in the areas assessed during the field assessment were part of both desk reviews. In the first desk review, buildings were selected by ordering the buildings being reviewed by OSM ID, filtering by buildings marked with “no damage”, randomly generating ten numbers and selecting buildings whose place in the order corresponded to these numbers. This process was employed because OSM IDs have no correlation to location or tagging. Hence they are a neutral means of assigning numbers to buildings.

The “no damage” tag was reviewed because field findings for that category mismatched to a high degree with the remote damage assessment data. The same process was used for the second desk review, increasing the sample size to 25.

Building Classification

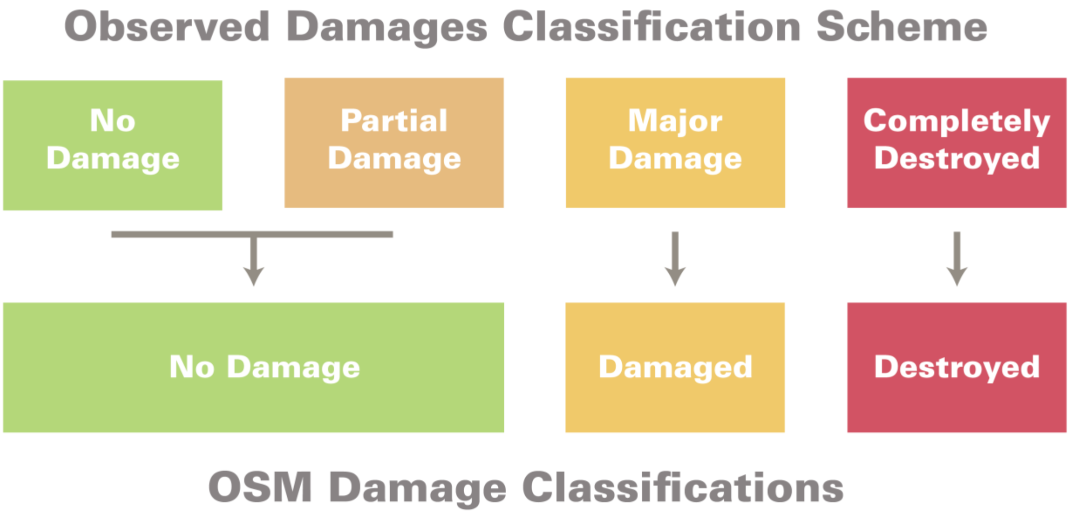
As mentioned above, the OSM community used three classification levels of building damage, compared with the four levels of the Shelter Cluster Rapid Needs Assessment. The OSM community adopted this three category classification scheme because minor damage and levels of damage as a whole could not be perceived using existing satellite imagery. Contributors were unable to view “partial” damage to the sides or insides of buildings due to the inherent limitations of overhead satellite imagery. Therefore, this report compares the number of undamaged buildings recorded in OSM with the undamaged or partially damaged buildings observed during enumeration. Under the OSM system, buildings were tagged as “major”, “damaged”, “destroyed” or “collapsed” and then categorized within the three categories of “undamaged”, “damaged” and “destroyed or collapsed”. Table 3 illustrates how the tags and classification categories align.

Table 3: OSM Classification System

Category	Tag
Undamaged	building=yes
Damaged	building=yes AND damage=major OR building=damaged
Destroyed or collapsed	building=yes AND damage-destroyed OR building=destroyed OR building=collapsed

Given the two classification systems used by the field assessment and the OSM community, the assessment team used the logic model shown in Figure 1 to compare building classifications between the field assessment and the OSM damage assessment. The figure illustrates that the “no damage” and “partial damage” categories from the field assessment were consolidated to equate with the “no damage” category of the OSM damage assessment.

Figure 1: Field and OSM Damage Assessment Classification Schemes



Limitations

There were two key limitations to the imagery desk review that had an impact on the accuracy of the assessment and the impact of its findings. The assessment team has developed recommendations to inform future research on this topic to ensure these limitations are addressed in the future.

1. Comparison with aerial imagery: In some other similar research, there is a comparison with aerial imagery in order to further explore limitations of satellite imagery. Due to a lack of appropriate aerial imagery, this assessment did not use aerial imagery as a point of comparison, limiting identification of the specific deficiencies in satellite imagery.

2. Satellite imagery resolution: The resolution of existing satellite imagery sources was too low to reliably differentiate between destroyed and merely damaged buildings. Buildings with major damage in particular may be mistaken for destroyed; habitable buildings with heavily damaged roofs can appear destroyed at a one square meter pixel resolution. Buildings that were swept from their foundations may not appear at all, confusing inexperienced OSM contributors. The use of aerial imagery from unmanned aerial vehicles (UAVs) has significant potential in decreasing these resolution issues.

OSM Damage Assessment Compared to Field Assessment Findings

Overall, when compared to the field assessment**, the OSM remote damage assessment overrepresented the "destroyed" category by 134 per cent**, while the other two categories – "major damage" and "no damage" – were underrepresented by 25 per cent and 18 per cent, respectively. Furthermore, when grouping all categories together**, the proportion of buildings that were accurately tagged by OSM contributors was a mere 36 per cent**. This proportion was generally consistent for all municipalities, except Tacloban City, which will be discussed in the next section.

Damage Classification	OSM (%)	Observed (%)	Difference	Percent under/over-represented in OSM
Destroyed	32.76	14	18.76	Overrepresented 134%
Major Damage	29.19	39.24	-10.05	Underrepresented 25%
No / Partial Damage	38.05	46.77	-8.72	Underrepresented 18%

This lack of accuracy was even more pronounced in urban areas, with only 26 per cent of buildings accurately categorised compared to around 40 per cent for rural and peri-urban areas[7]. Feeding into this over and under representation is the reality that buildings were categorised incorrectly. Notably, of buildings tagged as "collapsed" or "destroyed" through OSM, only 16 per cent were actually destroyed: 43 per cent had major damage, 25 per cent were partially damaged and 15 per cent were undamaged. Conversely, buildings tagged as "undamaged" actually had major damage or were destroyed 50 per cent of the time. Of buildings tagged as "damaged," eight per cent were actually destroyed, 21 per cent were undamaged and 71 per cent were actually damaged (majorly or partially).

Analysis of the Imagery

Overall, the comparative OSM and field assessment results clearly point to a critical lack of accuracy in the damage assessment conducted by OSM contributors. The reasons for this are supported both within the analysis in this assessment as well as similar studies conducted after other humanitarian emergencies. Investigating the imagery used by OSM contributors to classify buildings, it is clear that the imagery did not provide enough detail, nor was it at a high enough resolution for contributors to accurately classify the buildings.

Both imagery desk reviews found that remote assessment damage tags largely matched observable damages in the available imagery. By randomly selecting buildings marked in OSM as "undamaged" and flagged by enumerators as "major damage" or "totally destroyed" and analysing the imagery provided to OSM contributors, in nearly all cases, the imagery was the limiting factor in accurately classifying the buildings[8]. The imagery seemed to show undamaged buildings, when in reality they had sustained some damage, according to the field assessment. In a few other cases, the imagery was unclear due to low resolution, so OSM contributors appeared to have left the buildings' tags untouched as per the instructions provided by HOT OSM Activation Leads for unclear buildings. In a final two cases, two clearly damaged buildings had been incorrectly labelled as "building=yes" (or "undamaged"). They also lacked the review tags indicating whether a building was assessed for damages or not. This suggests that the buildings may have been accidentally overlooked during OSM editing.

[7]This should be explored further, but is outside the scope of this report. Rural, urban and peri-urban are highly complex terms that may not fully correspond to the type of analysis intended. Instead of using these classifications, a post-assessment density analysis should be conducted to determine the building density of assessed areas and, thus, conduct analysis that could then be used for further classification and analysis.
[8]OSM contributor error not explained by the limitations of the imagery was found in 20 per cent of reviewed tags.

The desk review found that there was a lack of comprehensive pre-disaster building data and a potential tendency for OSM contributors to digitise visibly damaged buildings with the exclusion of undamaged areas. Without pre-disaster data for the location of existing buildings, OSM contributors could have omitted areas that had been completely damaged without knowing a building had existed in that location. This could have led to a large omission error, something this study does not explore.

Other studies have been conducted supporting the finding that current remote sensing techniques do not provide clear enough images for accurate damage assessments. In a comparative analysis of aerial imagery, satellite imagery and field observations following the 2010 Haiti earthquake[9], the Earthquake Engineering Field Investigation Team (EEFIT) found a lower level of agreement between results from the satellite imagery and field assessments when compared with damage assessments done using aerial imagery. Using a kappa statistic[10], the study found a kappa of 0.22 for the comparison between the field assessment and satellite imagery. While still low, the study found a higher kappa (0.31) for the comparison between the field assessment and analysis of aerial imagery. One of the study's conclusions is that the higher level damages were not able to be classified using satellite imagery, as damage to the sides and foundation of the building could not be seen. This was also the case for aerial imagery, but less so. Overall, damage levels were underrepresented in the assessments using satellite and aerial imagery when compared to the field assessment.

A similar study was conducted by Cambridge Architectural Research following the 22 February 2011 earthquake in Christchurch, New Zealand[11]. The study compared damage classifications conducted by GEO-CAN[12] and a field assessment. The study found an overall accuracy of 36 per cent for the GEO-CAN assessment and further found that there were a number of omission errors – buildings that were classified as damaged by the field assessment, but that had not been identified as affected by GEO-CAN. The omission error proportion was quite large – 64 per cent – leading the researchers to conclude that satellite imagery limited the utility of remote damage assessments, as it did not allow analysts to identify the actual extent of damage.

A Possible "Media Effect"

The previous analysis has focused on the underrepresentation of damage in remote assessments that has been found both within this assessment as well as others. This does not account for the large overrepresentation of buildings classified as 'totally destroyed" by OSM contributors when compared with the field assessment.

One possible explanation for this is the focus on certain high impact areas at the exclusion of others by the media. The hypothesis would stand that if there were a significant focus by the media on a certain geographic area, overestimates would be higher in this area given the high publicity and an OSM contributor's tendency to classify a building as damaged even when the image may not have been clear enough to ascertain this.

In order to explore this bias, overestimates of damage in the most highly publicized area assessed in this study – Tacloban City – were compared with overestimates in other, less publicized areas. In support of the hypothesis of "media effect", Tacloban City had by far the highest percentage of overestimated cases of damage – 92 per cent of all overestimation. This accounts for nearly all of the overall "totally damaged" overestimation outlined above. All other municipalities had relatively substantially less damage overestimation, ranging from 76 per cent to 56 percent.

[9] EEFIT. The Haiti Earthquake of 12 January 2010: A Field Report. (Cambridge 2010).
[10] A measurement of interrater agreement that is commonly used to assess the accuracy of classification assessments. The kappa index has a range of values between zero and one – zero being agreement purely by chance and one being perfect agreement.
[11] Cambridge Architectural Research Ltd. The Use of Remote Sensing for Building Damage Assessment Following the 22 Feb 2011 Christchurch Earthquake: the GEOCAN Study and its Validation – Draft for Comment. (Cambridge 2012).
[12] A worldwide network of scientists tasked with assessing impacts of humanitarian disasters using remote methods.

REACH Informing more effective humanitarian action

One municipality in particular – Carles – had zero instances of overestimation. It is important to note that Tacloban City was prioritised by the OSM community, whereas Carles was not assessed at all until researchers on this assessment requested it be. Another possible explanation for the highly divergent results in Carles could be that the OSM community had the opportunity to conduct a thoroughly validated pre-disaster review using Bing imagery before proceeding with the post-disaster damage assessment. While this may have affected the overrepresentation, it does not seem to have affected underrepresentation, as results from the remote damage assessment in Carles are only 39 per cent accurate when compared to the field assessment.

Given that the general consensus among researchers assessing the validity of remote damage assessments is that satellite imagery tends to lead to underrepresentation of damage levels, this is a key finding that would be important to explore further in reference to crowd-sourced remote damage assessments.

CONCLUSION

The results of this research, as well as other studies mentioned in this report confirm the inherent limitations of current satellite imagery to accurately assess building damage in a humanitarian context. These inherent limitations are further amplified by the business processes, lack of coordination mechanisms for crowd-sourced data and the limited use of pre-disaster imagery in post-disaster analysis.

As satellites improve and unmanned aerial vehicles proliferate, imagery resolution will increase to a point where remote volunteers can assess with confidence building damages for even the smallest of dwellings. In the meantime, refinements to the mechanisms used to deliver imagery, improvements to the guidance and training of OSM contributors and standardized validation review procedures for contributor data would greatly boost the accuracy of this data. Most importantly, stronger efforts to create pre-disaster base data layers will enhance the speed and reliability of remote damage assessments while yielding immediately useful data.

Humanitarian agencies and donors should invest now in the disaster preparedness approaches and technology needed to make OSM more operationally useful for disaster preparedness and response. Communities, cities, provinces and countries hosting these activities and agencies adopting these technologies will benefit when disaster strikes and they can use OSM for damage analysis and post-disaster planning. OSM is strong because it's an ecosystem; a collection of tools and approaches designed to support a central database and map. That makes it powerful, resilient and adaptable, but also difficult to improve or utilize in a piecemeal fashion.

For that reason, policymakers that take engagement with OSM seriously enough to make investments for the long-term will reap much greater operational rewards. Thoughtful, sizeable and sustained investments will make the difference between OSM's currently limited utility to disaster damage assessments and a more robust geographic open data platform that can be the foundation for understanding and implementing disaster preparedness and response activities at a household level.

Enhancing OSM engagement in disaster risk preparedness and management directly benefits all aid actors involved in the operational response to humanitarian emergencies. The potential uses of crowd-sourced OSM data and damage assessments could provide aid actors with a powerful tool for response with the proper investment and coordination. REACH and ARC are committed to supporting and facilitating this engagement and working to understand how to make the process more effective.

RECOMMENDATIONS

Given the limitations of satellite imagery in remote damage assessments found both within this study and others as well as the possibility of external influence on crowd-sourced contributor classifications, the following recommendations focus on minimising these effects and limitations by improving process and management. The recommendations are geared toward the OSM community and the larger humanitarian community.

Pre and Post Imagery Comparison

As mentioned before, the lower overrepresentation of damage in Carles municipality could have been due to the fact that a full pre-disaster imagery review was conducted before assessing post-disaster damage. Having a point of reference to assess whether a building is damaged or not could lead to higher accuracy rates when remotely assessing damage in humanitarian contexts.

In the OSM context, this could be achieved by upgrading the Java OSM Editor (JOSM) to include a pre and post imagery interface that would allow contributors to flip between pre and post images for rapid comparison. This requires for a location to already be digitised; something that the OSM community is already working on with ARC for natural disaster hotspots.

Timely and Targeted Imagery Provision

Compounding the limitations inherent in the use of satellite imagery was the issue of timely imagery delivery and use of this imagery by OSM contributors. Imagery for some areas was provided in a timely manner – just days following the storm. For others, however, imagery was not provided until well over a week after the storm, making accurate assessments of damage increasingly more difficult.

The rate of recovery within a disaster zone can easily outpace the rate at which satellite imagery is acquired, analysed, and shared; where imagery is slow, response activities are not. These delays in data acquisition ultimately slow the adoption and use of created data. The OSM community's reliance on dated imagery reduces the accuracy of damage assessments and applicability of the data. Imagery not released quickly can become obsolete or inaccurate within a matter of days, leading to equally inaccurate damage assessments and less timely information. This is compounded by the existing slow distribution of imagery and imagery derived products to the field caused by bandwidth limits.

Furthermore, specifically in the Philippines, satellite imagery coverage areas were driven by popular media accounts rather than actual needs or requests from the ground. In the Philippines, many affected areas went uncovered while Tacloban City was targeted dozens of times, possibly leading to the overrepresentation of damage levels discussed above.

Satellite imagery contributors must work to ensure that post-disaster imagery is provided to contributors within 24-48 hours of a disaster. This increases the likelihood of more accurate damage assessments by ensuring that images accurately portray the current reality on the ground. Humanitarian response agencies should also build strong relationships and information sharing processes with satellite imagery providers and humanitarian coordination structures to better target affected areas.

Better Crowd Coordination

This assessment demonstrated the continued responsiveness and diligence of the crowd when well directed. For example, the entire municipality of Carles was mapped and validated within 48 hours after the request of ARC. Impressively, this occurred three weeks after Typhoon Haiyan made landfall when media attention no longer drove OSM contributors towards the HOT OSM listserv and web platform.

Humanitarian agencies that build technical expertise and cultivate relationships with OSM should be able to direct remote OSM contributors toward priority mapping tasks not just during the initial response period but well into the recovery phase of an operation. Quality communications and transparency about goals, products, and successes are the key inputs to ensure quality outputs.

Additionally, aid actors should work to link the OSM community with the humanitarian coordination system during a crisis response in order to ensure that priority areas are assessed and information is shared with agencies even without a direct relationship with the OSM community. This could be a role for the United Nations Organisation for the Coordination of Humanitarian Affairs (OCHA) through the cluster framework, and in partnership with the United Nations High Commissioner for Refugees (UNHCR) in contexts of cross-border displacement of populations.

Guidance Materials

One potential contributing factor to the low levels of accuracy of damage is the lack of guidance materials for contributors new to OSM or to humanitarian damage assessments in general. When analysing the imagery that was used by OSM contributors, researchers on this study found a number of errors that may have been avoided with more guidance from the OSM community as a whole.

Creating generic and disaster-specific damage assessment guidance materials for remote mappers would improve the accuracy of results and reduce the number of potentially inaccurate judgment calls contributors are asked to make. The OSM community should strive to build documents of lessons learned and best practices for disaster assessments in order to mitigate the impact of OSM community members leaving the community and depleting the institutional memory of the platform. Tailored materials featuring disaster-specific imagery could help to better identify damage patterns common to local construction types and improve the accuracy of the data created for a given disaster. OSM contributors could be required to watch a short YouTube video or skim a 5-10 page visual guide to damage tagging before beginning a task.

Objective methods of contributor evaluation could be considered as well. A short “test” for new contributors could be required in order to begin contributing. This could consist of a series of images from past responses requiring the new contributor to correctly classify buildings according to the specific damage classification scale outlined in the guidelines mentioned above.

Comprehensive Validation Processes

Time is of the essence during a humanitarian emergency and validation of remote damage assessment results may not be a top priority for crowd-sourced damage assessors. Given the wide range of skills that individuals possess when contributing to OSM, there must be a comprehensive validation process in order to minimise errors.

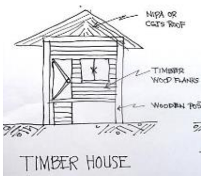
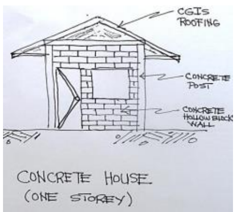
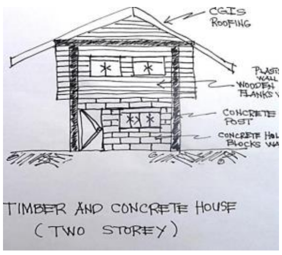
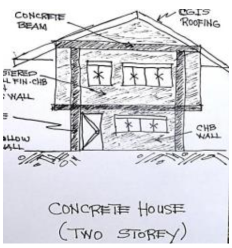
One possible solution is to develop two categories of assessor based on the results of the evaluation mentioned above. By making the evaluation a regular measurement of skills, OSM contributors could be placed into different categories based on scores on the test. A score on the test above a certain threshold would move a contributor into a “validator” status that would require and allow them to validate a certain percentage of contributions proportional to their own level of contribution.

Another crowdsourcing mapping platform called Tomnod[13] uses contributed data to triangulate results automatically. By leveraging the power of crowdsourcing, its CrowdRank algorithm triangulates all contributions for a specific area to eliminate outliers and find agreement on the results. This is a method that should be further explored in future studies.

[13] www.tomnod.com

Annex A - Damage Level Classification Guide for Enumerators

Damage Categories
No Damage
Minor Damage
Major Damage
Totally Destroyed

Name	Dwelling Type	Photo example	Damage Type	Category
Hut			1. Collapsed totally	totally
			2. Building Tilting sideways (right or left)	major
			3. Wooden Posts/beams bent/cracked/ dislocated	major
			4. Walls missing/damaged	major
			5. Roof missing/damaged	major
			6. Doors and windows damaged	minor
			7. Floors – collapsed/broken	minor
			8. Stairs / collapsed/missing	minor
			9. Foundation off line from wooden posts	major
Timber Frame			1. Collapsed totally	totally
			2. Building Tilting sideways (right or left)	major
			3. Wooden Posts/beams damaged - dislocated	major
			4. Walls missing/damaged	major
			5. Roof missing/damaged	major
			6. Doors and windows damaged	minor
			7. Stairs / collapsed/missing	minor
			8. Foundation off line from wooden posts	major
Timber and Concrete (one storey)			1. Collapsed totally	totally
			2. Tilting sideways (right or left)	major
			3. Concrete columns/beams damaged/bent/cracks/tilt	major
			4. Timber Walls/dislocated/broken/missing	major
			5. Concrete Hollow Block work /collapsed/tilt/cracks	major
			6. Roof damaged/missing	major
			7. Doors and windows damaged	minor
			8.Plaster/damaged/cracks/removed	minor
Concrete House (one Storey)			1. Collapsed totally	totally
			2. Tilting sideways (right or left)	major
			3.Concrete columns /beams/ damaged/bent/cracks/tilt	major
			4. Concrete Hollow Block work/collapsed/tilt/cracks	major
			5. Ceiling damaged/missing	minor
			6. Roof damaged/missing	minor
			7. Doors and windows damaged	minor
			8. Floor Slab / broken/cracks/split	minor
			9. Plaster/damaged/cracks/split	minor
Timber and Concrete House (two Storey)			1. Collapsed totally	totally
			2. Tilting sideways (right or left)	major
			3.Concrete/Timber columns /beams/ damaged/bent/cracks/tilt	major
			4. Concrete Hollow Block work/collapsed/tilt/cracks	major
			5. Ceiling damaged/missing	minor
			6. Roof damaged/missing	minor
			7. Doors and windows damaged	minor
			8. Floor Slab / broken/cracks/split	minor
			9. Plaster/damaged/cracks/split	minor
			10. First Floor Failed /Collapsed	major
Concrete House Two Storey			1. Collapsed totally	totally
			2. Building Tilting sideways (right or left)	major
			3.Concrete/Timber columns /beams/ damaged/bent/cracks/tilt	major
			4. Concrete Hollow Block work/collapsed/tilt/cracks	major
			5. Ceiling collapsed (inside)	minor
			6. Roof damaged/missing	major
			7. Doors and windows damaged	minor
			8. Floor Slab / broken/cracks/split	minor
			9. Plaster/damaged/cracks/split	minor
			10. First Floor Failed /Collapsed	major