Validity and Reproducibility of Typhoon Damage Estimates using Satellite Maps and Images

Wenceslao C. Perante

Associate Professor V, Eastern Visayas State University (EVSU), Tacloban City, Leyte, 6500, PHILIPPINES

ARTICLE INFO
Received: May 08, 2016
Accepted: May 23, 2016
Published: Jun 06, 2016

ABSTRACT
The use of satellite maps and images in rapid assessments offers many advantages, but questions on validity, and reproducibility of data obtained is always a concern. This study used satellite maps and images to quickly assess the level of damage, the number of affected households, the number of affected individuals, and the status of roads and bridges after Typhoon Haiyan hit Tacloban City in November 8, 2013. Cohen's Kappa and Intraclass Correlations (ICC) were used to determine validity and reproducibility of visual damage assessment data derived from pre and post satellite maps. Findings suggest that the use of satellite maps and imagery is valid and reproducible; but for estimates on affected households and individuals, results are poor to moderate suggesting that assessors for this parameter should have a background, training, and experience on visual rapid assessments to yield better results. This method offers many advantages, but it also has many limitations. This study presents the processes employed in visual data interpretation using satellite maps and images; recommendations for improvements are also presented. The use of satellite maps and images may be employed in situations requiring quick data that are vital in creating an effective, and informed emergency response after a natural disaster, thereby, making delivery of intervention, relief, and services to the affected public faster and more efficient.

Key Words: Online applications, satellite maps, disaster response, satellite imagery, Cohen's kappa, intraclass correlations (ICC), Typhoon Haiyan, validity and reproducibility of data

Source of Support: EFN Engineering Works Consultancy and Research Services

No Conflict of Interest: Declared

INTRODUCTION
Natural disaster perennially struck the Philippines; the latest of which that caused major damage and death toll is Typhoon Haiyan that hit the central Philippines in November 8, 2013. As is usually the case, government agencies are confronted with the difficulty of quantifying damage impact due to limited resources hindering effective response and intervention. Six (6) months after the typhoon, report from the National Disaster Risk Reduction & Management Council (NDRRMC Sit. Rep. No. 108, 2014) showed that as of April 2014, there were 6,300 fatalities, 28,689 persons injured, and 1,061 missing individuals and 1,140,332 houses that were either partially or totally damaged.

Disasters of this magnitude that covered large impacted areas, quantifying level of damage and number of affected households, affected individuals, affected infrastructures is often difficult and limited, however, nowadays, with the availability of satellite maps and imagery, and related applications such as Geographical Information System (GIS) can hasten and quicken the process of data gathering, providing information that could enable government units, institutions, relief and AID agencies deliver fast, effective, and informed response (Wegscheider et al., 2013; Barnes et al., 2007; Gokon and Koshimura, 2012; Mas et al., 2012, 2014).

In Tacloban City, it took the city government 16 to 18 months after Super Typhoon Haiyan struck to distribute Cash Emergency Shelter Assistance to its beneficiaries. Although there are various reasons to the delay, one of the major reasons was the lack of data and information such as names, category of shelter damage of each beneficiary, and the total number of the beneficiaries, among other data requirements as government agencies such as the Department of Social Work and Development (DSWD) relied mostly on ground reports from the Local Government Units.

This study investigated the use of satellite maps and imagery to quickly assess the level of damage, number of affected households and individuals after Typhoon Haiyan hit Tacloban City in November 8, 2013. The validity and reproducibility of data obtained through visual assessment or manual visual interpretation of pre and post satellite maps and images were investigated so as to improve, possibly, the existing data gathering methodologies employed by various government agencies, government line agencies, aid agencies, NGOs, academe, and others.

STUDY SITE
This study was conducted in Tacloban City (Figure 1). The city is a first income class city and the capital of the province of Leyte, Philippines. It is 360 miles (580 km) southeast from Manila. It has a population of 221,174 as per 2010 population census, and it is the most populous city in Eastern Visayas. The city is the regional center of the Eastern Visayas, being the main gateway by air to the region.
Figure 1: Map showing Tacloban City and the eye of Typhoon Haiyan taken by Moderate Resolution Imaging Spectroradiometer (MODIS) of NASA’s Aqua satellite. Remote sensing data from the Joint Typhoon Warning Center, reported sustained winds approaching 315 kph (195 mph) just three hours before landfall, with gusts to 380 kph (235 mph). Haiyan hit land at around 6:30 AM Philippine Standard Time, November 8, 2013.

METHODS

Data sources: Pre and post Typhoon Haiyan satellite images

Five (5) pairs of randomly selected satellite maps provided by the European Space Company and published by ESRI (Environmental Systems Research Institute) an international supplier of Geographic Information System (GIS) software, web GIS and geo-database management applications were downloaded (Figure 3). Each map covers an area of approximately 800 meters x 1.3 kilometers section of Tacloban City, and has both pre and post-Super Typhoon Haiyan satellite images enabling the researchers to conduct visual comparisons.
Visual damage assessment using satellite images

Visual assessment is usually done through aerial surveys using helicopters and planes. However, online applications such as Google maps and Google Earth and other similar sites offer satellite maps that can be an alternative to costly aerial surveys. Visual surveys are usually rough estimates by the observer and questions on accuracy, validity, and reproducibility is always a concern.

For this study, visual assessment and evaluation of each of the pre and post satellite maps were undertaken following a list of parameters. Copies of the pre and post maps were provided to randomly selected respondents or map assessors and were tasked to make rapid visual estimates. The map assessors comprises of students (n=11) and professionals (n=27) with varying backgrounds, age, sex, and civil status. Ground data were obtained from the Local Government Unit of the City of Tacloban, or the Department of Social Work and Development.

Visual damage assessment guide

A guide of estimating ground cover of certain type of habitats as used in manta tow survey (Figure 2) (Miller, J.W. 1991; Rogers, Caroline S., et al., 1994) was presented to the observers as an adaptable method of visual assessment. The assessors were encouraged to adapt the process of visual estimation in determining the “level of damage” and related parameters using the satellite maps. The observers were encouraged to count houses, buildings, and similar structures (whenever visible and identifiable) so as to obtain data close to “ground reality” as possible.

Figure 2: Broad scale survey guide (manta tow) in estimating habitat type cover (Source: Miller, J.W., 1991).

Figure 3 (Next Page): The five pairs of satellite maps selected for detailed study. Maps were downloaded online provided by the European Space Company and published by ESRI (Environmental Systems Research Institute) an international supplier of Geographic Information System (GIS) software. (Source: http://www.esri.com/services/disaster-response/hurricanes/typhoon-haiyan-yolanda-sweep-map Accessed on January 13, 2016).
Data analysis

For this study, reproducibility is defined as the method of using satellite maps in measuring damage impact based on defined parameters and its ability to be duplicated, either by the same assessors, researchers, or someone else working independently; while validity is the level of correspondence of the measured parameters to the real world or the ground data collected through traditional means i.e. field interviews and surveys as sourced from Local Government Units or the Department of Social Work and Development.

Statistic used to measure validity and reproducibility employed Cohen's kappa (Cohen, 1960) and Intraclass correlation (ICC) (Fisher, 1954; Portney and Watkins, 2000; Landis and Koch, 1977). Kappa either Cohen’s or Fleiss measures inter-rater agreement for qualitative or categorical properties (Galton, 1892; Smeeton, 1985); and Intraclass correlation (ICC), measures agreement or consensus from two or more raters when data are assumed parametric (Portney and Watkins, 2000; Landis and Koch, 1977).

Kappa and ICC were calculated to determine the (a) validity of visual damage assessment from satellite maps by comparing it with ground data; and (b) reproducibility by comparing inter-assessors assessments of the satellite maps. For both statistics, Kappa and ICC were undertaken to test against the null hypothesis that the observed agreement (either validity or reproducibility) does not exceed the amount of agreement as expected by chance alone. As there is no standardized interpretation for the kappa statistic; this study used Landis and Koch (1977) range of interpretations and modified in this study as presented in Table 1.

Table 1: Range of interpretation used in this study following the ranges of Landis and Koch (1977); values of 0 to 0.40 were excluded for interpretation in this study as higher agreement is what is needed to make the use of satellite maps or images an effective tool for quick assessment

<table>
<thead>
<tr>
<th>Kappa / ICC Values</th>
<th>Strength of Validity</th>
<th>Strength of Reproducibility</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.81-1</td>
<td>Perfectly valid</td>
<td>Perfectly reproducible</td>
</tr>
<tr>
<td>0.61-0.80</td>
<td>Valid</td>
<td>Reproducible</td>
</tr>
<tr>
<td>0.41-0.60</td>
<td>Moderately valid</td>
<td>Moderately reproducible</td>
</tr>
<tr>
<td>0.21-0.40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-0.20</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

RESULTS

The visual damage assessment using satellite images (n=5 pairs) covered a total of approximately 5.3 km² square kilometers of the City of Tacloban. Results indicate that visual damage assessment of the satellite maps done by all map assessors (n=38) on the following parameters: level of damage (Likert scale 1-5, with 5 being the highest level); impact on roads and bridges (passable or not); showed relatively high reproducibility and validity coefficients; but for households and individuals’ affected, results indicate poor to moderate reproducibility.

Level of Damage

Based on the responses of 38 map assessors, 92% agreed that the level of damage on Sat Map-1 is, conservatively, at least level 3 or higher with 74% agreeing on level 3 and 4 (Figure 4). On Sat Map-2, where the impact is glaring as it wipe-out a near shore community, 81%
agreed on level 5 damage; the consensus of 100% among assessors’ on Sat Map-2 is a level of damage between 4 and 5. On Sat Map-3, the consensus of 93% is at level 4 and 5; and on Sat Map-4 is at level 3, while Sat Map-5, 70% agreed on Level 3 and 4.

Figure 4: Cumulative percentages of absolute agreement among assessors (n=38) on “level of damage” on the five (5) satellite maps.

Table 2: Indices and significance tests on “level of damage” for inter-assessors agreement (reproducibility); and assessors - ground data agreement (validity); $k$= Fleiss Kappa; ICC = interclass correlation coefficient.

<table>
<thead>
<tr>
<th>Satellite Map</th>
<th>Level of Damage Reproducibility (Assessor / Assessor)</th>
<th>Level of Damage Validity (Assessor / Ground Data)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$k$</td>
<td>ICC</td>
</tr>
<tr>
<td>Sat Map-2</td>
<td>5</td>
<td>0.69</td>
</tr>
<tr>
<td>Sat Map-3</td>
<td>5</td>
<td>0.53</td>
</tr>
<tr>
<td>Sat Map-1</td>
<td>3 &amp; 4</td>
<td>0.58</td>
</tr>
<tr>
<td>Sat Map-4</td>
<td>3 &amp; 4</td>
<td>0.62</td>
</tr>
<tr>
<td>Sat Map-5</td>
<td>3 &amp; 4</td>
<td>0.44</td>
</tr>
</tbody>
</table>

Table 2 shows the reproducibility and validity of assessors’ assessments on “level of damage” from the pre and post satellite maps; Sat Map-1 and Sat Map-3 had Fleiss Kappa values of 0.69 and 0.53, respectively. This study pooled values for level 3 and level 4 damage for Sat Map-1, Sat Map-4, and Sat Map-5, as most of the assessors; perhaps, where undecided as to which level the damage they observed should be categorized. Maybe this is due to the inexperience of the assessors on using a Likert scale and unfamiliarity with satellite map visual assessment. This is understandable because using a Likert scale in measuring parameters, the difference from one level to the next is not well defined; and perhaps not equidistant. What is consistent on the data, however, is that, for most assessors there is an observed damaged whose level is around 3 and 4 for Sat Map-1, Sat Map-4, and Sat Map-5; the pooled data for these satellite maps had Fleiss Kappa values of 0.58, 0.62, and 0.44, respectively.

Validity of the assessments employed Two Factor ANOVA without Replication following the procedure of Zaiontz, C. (2016) and Vanbelle, S. (2009) to obtain the ICC coefficients and p-values. Table 2 shows the entire assessors assessments on “level of damage” on the five (5) sets of satellite images showing results that are perfectly valid and reproducible (Landis and Koch, 1977).
Status of Roads and Bridges

The assessment on the visibility of “roads and bridges” including its “passability” (defined in this study as roads being passable) all showed moderate to high reproducibility and validity (Table 3). Results also indicate a relationship of the roads being “passable” when it is “visible” on the satellite map (correlation, r=0.77). Figure 5 shows the cumulative agreements of the assessors on the state of roads and bridges in the five satellite maps.

![Cumulative percentages of absolute agreement among assessors (n=38) on category “roads and bridges visibility” and “roads and bridges passability” on the five (5) satellite maps.](image)

Table 3: Indices on “roads and bridges visibility” and “roads and bridges passability” for inter-assessors agreement (reproducibility) and the ratio of assessors /ground data agreement (validity); k= Fleiss Kappa.

<table>
<thead>
<tr>
<th>Satellite Map</th>
<th>Road and bridges visibility (reproducibility)</th>
<th>Road and bridges visibility (validity)</th>
<th>Road and bridges passability (reproducibility)</th>
<th>Road and bridges passability (validity)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sat Map-1</td>
<td>0.64</td>
<td>0.78</td>
<td>0.52</td>
<td>0.63</td>
</tr>
<tr>
<td>Sat Map-2</td>
<td>0.64</td>
<td>0.78</td>
<td>0.57</td>
<td>0.70</td>
</tr>
<tr>
<td>Sat Map-3</td>
<td>0.60</td>
<td>0.74</td>
<td>0.52</td>
<td>0.64</td>
</tr>
<tr>
<td>Sat Map-4</td>
<td>0.52</td>
<td>0.63</td>
<td>0.49</td>
<td>0.55</td>
</tr>
<tr>
<td>Sat Map-5</td>
<td>0.68</td>
<td>0.81</td>
<td>0.64</td>
<td>0.78</td>
</tr>
</tbody>
</table>

Number of Households and Individuals Affected

The number of households and individuals’ affected (range: 000), not all the five (5) maps showed acceptable reproducibility and validity results; and those showing agreements, the strength of values are poor and moderate; thus this study cautions that in estimating households and populations using satellite maps; only assessors that have experience, background, and training on visual rapid assessments should be considered to obtain better accuracy, reproducibility, and validity.
Table 4: Indices on the “number of households and individuals affected” inter-assessors’ agreement (reproducibility); validity were not calculated as the assessment were not reproducible; $k =$ Fleiss Kappa.

<table>
<thead>
<tr>
<th>Satellite Map</th>
<th>Reproducibility $k$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Households</td>
</tr>
<tr>
<td>Sat Map-1</td>
<td>0.55*</td>
</tr>
<tr>
<td>Sat Map-2</td>
<td>0.23</td>
</tr>
<tr>
<td>Sat Map-3</td>
<td>0.24</td>
</tr>
<tr>
<td>Sat Map-4</td>
<td>0.40</td>
</tr>
<tr>
<td>Sat Map-5</td>
<td>0.27</td>
</tr>
</tbody>
</table>

*significant results

Figure 6: Absolute agreement among assessors on the “number of households and individuals affected”; Sat Map-1 (n=27), Sat Map-2 (n=27), Sat Map-3 (n=27), Sat Map-4 (n=38), Sat Map-5 (n=34). Answers not in absolute numbers were excluded from analysis resulting in varying n values.

**DISCUSSION**

The magnitude of damage caused by Typhoon Haiyan was difficult to quantify through traditional method of field survey and data gathering. Time during disasters is always a limiting factor i.e. the urgency of relief and aid to reach the ground is vital as it decides the survival or demise of many disaster victims. Fortunately, time limitations had been overcome around the world by many institutions and research teams of varying disciplines through the use of satellite maps and images (Belen, 2015; Liu et al., 2014, 2013; Mas et al., 2014, 2013; NDRRMC, 2014, 2013) that provides preliminary quick assessment of impacted areas thus assisting agencies, institutions, and NGOs bring aid and relief to the ground rapidly. Although, questions on accuracy, validity, and reproducibility of data obtained using satellite maps and imagery are always a concern; the advantages it offers, perhaps, outweighs the disadvantages.

Findings showed that the data obtained from satellite maps are reproducible and valid particularly for parameters “level of damage” and “status of roads and bridges”; however for estimates on “number of households and individuals affected”, results showed poor reproducibility. The latter result could perhaps be attributed to (1) the assessors lack of training and experience on visual rapid assessments, indeed, significant technical expertise is required for satellite imagery interpretation in habitat mapping (Smith, G. et.at, 2011);
(2) the lack of background information on the area and locality, the demographics, and the population density of the target satellite maps for assessment; and (3) the low resolution and poor clarity of the satellite maps used (each map were color printed on an 8.5” x 11” bond paper (n=10) per assessor); had these issues been addressed beforehand, perhaps, estimates on “number of households and individuals affected” would yield significant reproducibility and validity values.

Satellite maps and imagery provides quick data and cost savings, among other advantages; but has many limitations (Ramsey E. III et al, 2012, van Westen, K. et al, 1996) i.e. the subjective interpretation of the data afforded by the satellite maps which can vary from one observer to the next, reproducibility of assessment are possibly dependent on observers work experience and training; the limited information that can be derived from satellite images such as storm surges and its magnitude which cannot be observed (Mas et al, 2012, 2014), resolution and clarity of the map available, and weather constraints as satellites may not be able to provide clear maps and images due to cloud cover in affected areas for a period of time (Liu et al, 2014).

The cost of acquiring satellite maps and images is no longer a hindrance. Google Earth and Google Maps offers free resources on the web, and had covered almost every region of the planet which can serve as pre-event maps. Post event maps and its acquisition could be a hindrance in quickly assessing damage impact, but could be remedied if prior arrangement are done with satellite image providers, or perhaps a collaboration with foreign institutions, for example, Japan International Cooperation Agency (JICA) which provide Pleiades satellite data (Belen, 2015) or with Copernicus Emergency Management Service of the European Commission DG Joint Research Centre that provides free, timely and accurate geo-spatial information derived from satellite remote sensing (NDRRMC,2013) among other institutions. Domestically, possibly with National Mapping and Resource Information Authority (NAMRIA), or Department of Science and Technology (DOST) as NAMRIA has its Unified Mapping Project (UMP), and DOST has its Project NOAH (Nationwide Operational Assessment of Hazards) for flood modeling activities (Belen, 2015), but this study is uncertain if a post event satellite maps can be obtained at a time when it is needed from these institutions.

In summary, rapid assessment using satellite maps offers numerous advantages and can provide reproducible and valid data assessments that its use in recent years had become universal; furthermore, the cost of post event satellite maps and images are now affordable and in some agencies even free. Additionally, the GIS software that allows the capture, manipulation, storing, checking, analysis, integration and display of spatial data (Burrough and McDonnell, 1998) is widely available on shops selling software and online, also the software requires minimal learning curve for most computer technicians.

Finally, this study recommends the following: (1) a training should be designed and provided to satellite map assessors to develop technical expertise and experience on visual satellite map and imagery assessment; (2) map with the best resolution should be used, better if it is a zoom-able digitized map, as it is perhaps, ideal for visual satellite map assessments; (3) background data on target maps for assessment should be at hand (population demographics, or habitat cover depending on the objectives of satellite map assessment); ( 4) the assessors should collaborate with allied agencies and institution i.e. National Statistics Office (NSO) for its population census data, perhaps encouraged the agency to plot the data on a GIS, or with the Department of Environment and Natural Resources (DENR) if the assessment is about ecology and environment); (5) satellite map
assessment should be done by a team with varied technical expertise and capabilities; and lastly, (6) satellite map assessment result may be posted online so as to encourage participation and input from people with direct knowledge or years of ground experience on the target site/s or expertise on the parameter being assessed.

ACKNOWLEDGEMENTS

This study thanks the BSCE students from Eastern Visayas State University (EVSU) and Samar State University (SSU); and the groups of professionals in Eastern Visayas, Philippines who participated in the manual visual interpretation of pre and post satellite maps and images used in this research. The maps used in this study were courtesy of European Space Company as published by ESRI (Environmental Systems Research Institute). Funding was provided by EFN Engineering Works Consultancy and Research Services to the author. Thanks are due to three colleagues in the academe: Dr. Fely Gomba of SSU, Dr. Norma Abolencia of EVSU, and NC Perante of EFN Engineering Consultancy who patiently reviewed this paper and provided their comments and inputs towards its improvement.

REFERENCES


Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977; 33:159-74.


