Port-au-Prince Earthquake Damage Assessment using Pictometry

A Report for ImageCAT Inc.

Cambridge Architectural Research Ltd
June 2010

25 Gwydir Street #6 Cambridge CB12LG UK info@carltd.com www.carltd.com tel +44 (0)1223 460475 fax +44 (0)1223 464142

Contents

Port-au-Prince earthquake damage assessment using Pictometry

Robin Spence, Keiko Saito, Cambridge Architectural Research Ltd
May 2010

1. Introduction

This report contains the results of the Pictometry-based damage assessment done by Cambridge Architectural Research Ltd, between 25.2.2010 and 15.3.2010, in association with ImageCat. The report also contains an assessment of the intersection of the Pictometry-based study with the GEO-CAN Phase 2 study, and a suggested extrapolation to estimated overall damage assessment based on this data. It also contains the results of a ground survey conducted by EEFIT in Haiti from 6.4.2010 to 13.4.2010 for comparison with the GEOCAN and Pictometry data.

The first phase of the work described below was carried out prior to completion of the joint JRC/UNOSAT/World Bank Building Damage Assessment PDNA report (JRC 2010), and was incorporated into that study. The second phase of the Pictometry work was carried out shortly after the submission of that report on March 11th, and serves to amplify the conclusions of that report, and to provide better statistical robustness to the breakdown of damage distributions by land-use class. The work was supported by the World Bank and the Global Facility for Disaster Reduction and Recovery (GFDRR).

2. *Damage assessment methoology*

In the first phase, (24.2.10- 1.3.10) a set of 25 randomly sampled survey locations was identified, all located within the PaP urban area, and all based around street intersections to facilitate subsequent street level photographic observation. At each location about 20 adjacent (or more or less adjacent) buildings were selected, a dataset of 523 buildings in all. For each building, the following information was assessed using the Pictometry images: number of stories; construction type (masonry or rc); use class (mainly residential or commercial); and damage level: D2, D3, D4, D5 or no visible damage (nvd). At ImageCAT this data was zoned according to a pre-existing land-use map produced by ImageCAT, which showed that the distribution of the selected survey locations among the principal land-use categories was uneven, and numbers of survey points were insufficient for a statistically robust sample of each land-use category.

A second phase of Pictometry analysis was therefore carried out (8.3-12.3.10), identifying a further 35 locations, so that approximately equal numbers of locations were in each of the following land-use categories: downtown; commercial; residential (high density and low density combined), and low-income or “shanty” settlements, all within the Port-au-Prince urban area. The relatively small industrial land-use category was not included. Table 1 shows the breakdown of the survey locations by landuse category, Figure 1 shows a map of all the survey locations and the land-use classification, and Figure 2 shows the identified buildings in a typical survey location. In this second phase a further 718 buildings were added to the dataset, recording the same data as in Phase 1.

Table 1. Breakdown of 60 survey locations by land-use class

Downtown	10
Commercial	18
Residential (high density)	11
Residential (low density)	4
Low-income (shanty)	17

Independently of the Pictometry assessment, the GEO-CAN Phase 2 damage assessment for each of the buildings in the dataset was identified (D5, D4, or not recorded). This allowed an error matrix to be assembled based on the combined dataset of 523+718 = 1241 buildings.

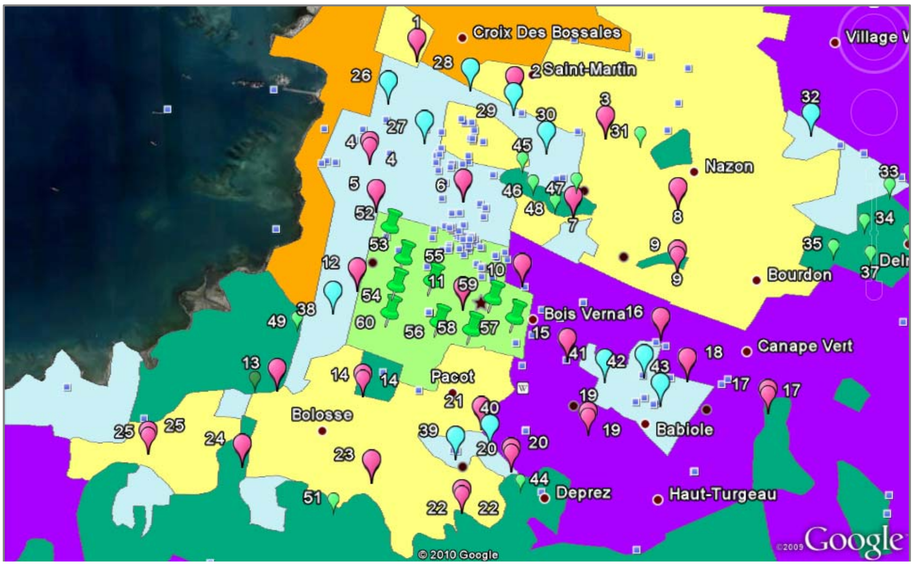


Figure 1 CAR's 60 damage assessment survey locations in Port-au-Prince, overlaid on ImageCAT landuse map

Figure 2 Building samples chosen at survey location 58 (Downtown area)

3 Results: Damage level assessment using Pictometry data

The overall damage assessment, and its breakdown by damage grade, is shown in Table 2.

Table 2 Overall damage assessment and breakdown by land-use class (nvd=no visible damage)

	Count					%			
	Commercial	Downtown	Res	Shanty	Total	Commercial (n=380)	Downtown (n=199)	Residential (n= 308)	Shanty (n=354)
D5	90	32	39	42	203	23.7%	16.1%	12.7%	11.9%
D4	36	24	20	35	115	9.5%	12.1%	6.5%	9.9%
D3	54	10	31	42	137	14.2%	5.0%	10.1%	11.9%
D2	34	24	24	43	125	8.9%	12.1%	7.8%	12.1%
nvd	166	109	194	192	661	43.7%	54.8%	63.0%	54.2%
total	380	199	308	354	1241	100.0%	100.0%	100.0%	100.0%
D4+D5						33.2%	28.1%	19.2%	21.8%

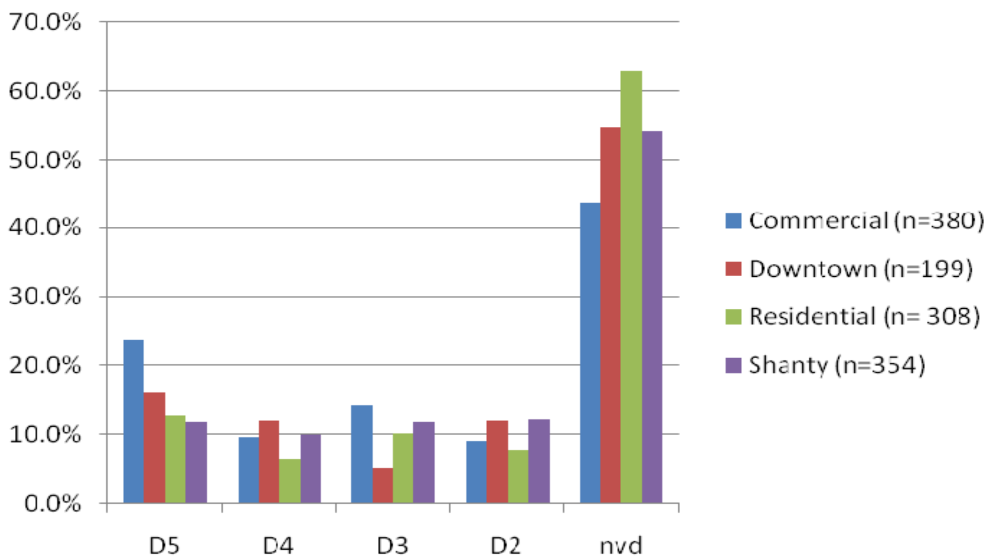


Figure 3. Damage distributions across 1241 buildings in PaP urban area from Pictometry survey, divided by land-use class.

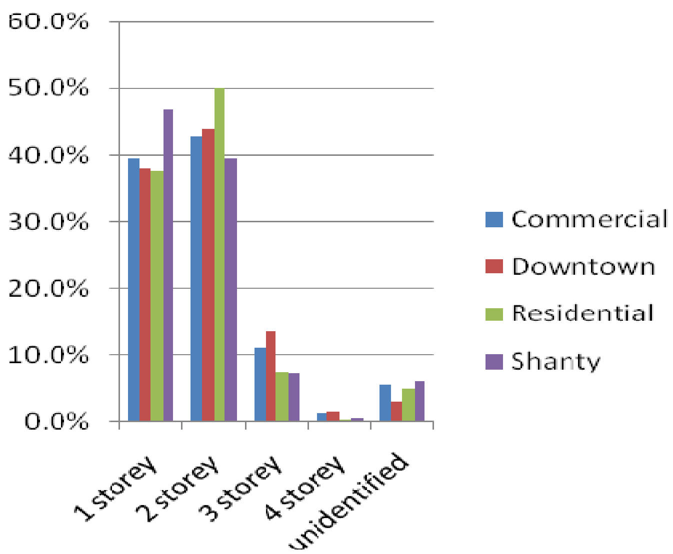
The proportion of buildings in damage states D5 and D4 is not dissimilar across all four land-use classes. The highest damage rate (33.2% D4 or D5) was found in the commercial area; the downtown area (28.2%) was also comparatively badly damaged. Damage proportions in the residential and shanty areas were significantly lower (19.2% and 21.8% respectively).

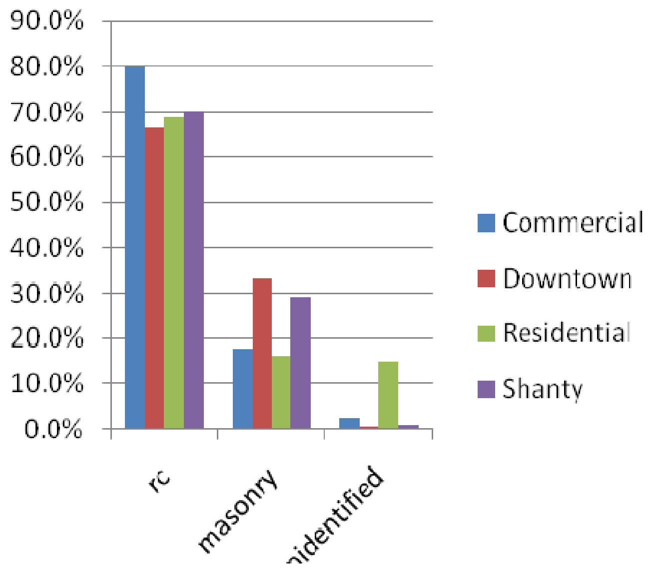
A comparison was also made between damage levels identified in this survey and the USGS Isoseismal map, and no consistent difference between the damage levels in the MMI=IX and the MMI=X areas was identified. The whole area could equally well be classified as MMI=IX.

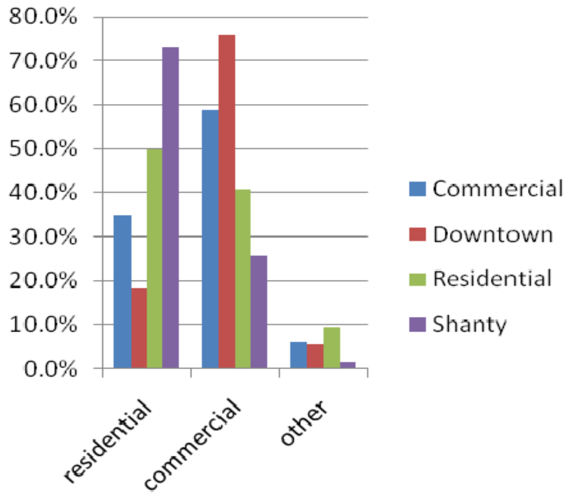
The Pictometry survey was able to clearly distinguish those buildings which had collapsed or partly collapsed. Thus the assignment of D5 or D4 could be made with some confidence, though the assignment of D5 rather than D4 was sometimes a matter of judgement. For a number of buildings damage at level D3 or D2 could be clearly seen (local wall or roof failures, parapet failures), etc, but it is likely that a closer view (e.g. from street level) would identify many instances of D3 or D2 in addition to those apparent from Pictometry (e.g. major wall cracking, failure of columns or column beam joints).

The Pictometry survey was also used to identify, where possible, number of stories, form of construction (rc or masonry) and occupancy (mainly residential or commercial). These results are shown, by land-use class, in Figures 4, 5 and 6.

Results were surprisingly consistent across land-use classes. By storey height, numbers are fairly evenly divided between one and two storey buildings, with a very few 3 storey or higher buildings. By construction type, the vast majority (65% to 75%) of buildings appeared to be of reinforced concrete (rc) construction, with a rather higher proportion of non-rc (taken as masonry) buildings in the downtown and shanty areas. By occupancy, as expected the majority of buildings in the downtown and commercial districts were commercial, but a surprisingly high proportion of buildings even in the residential areas were identified as commercial. Mostly these were buildings with an apparent part-commercial function on the ground floor, which were identified as commercial; the numbers were probably inflated because the sampling of buildings tended to be along the more significant streets, to enable later comparison with ground-level photographic data.







(from top left) Figures 4, 5 and 6. Distribution of buildings across land use classes by number of stories, form of construction, and

The distribution of damage by construction type and number of stories was also investigated, by examining damage rates across the entire dataset. Table 3 shows the Pictometry-based damage distribution by construction type. Overall the rates of destroyed buildings (D4+D5) are not very different; masonry buildings performed slightly worse than reinforced concrete buildings, and the proportions damaged at lower levels (D3 and D2) were similar. However, the difference in the ratio of D4 to D5 is very striking. Among masonry buildings, the ratio is 1.61, whereas among rc buildings it is 0.39. Thus while masonry buildings are prone to partial collapse (a section of wall), this is much less likely with rc buildings. The fact that rc floors and roofs consist of continuous rc slabs which remain in one piece even if supporting structures fail is a sufficient explanation for this observed difference in behaviour.

Table 3. Pictometry-based damage distributions among masonry (n=299) and reinforced concrete frame buildings (n=914).

Construction type	D5	D4	D3	D2	nvd	D4+ D5
masonry	10.4%	16.7%	11.7%	11.0%	49.5%	27.1%
reinf conc	17.5%	6.9%	10.5%	10.1%	54.4%	24.4%

Table 4 shows the Pictometry-based damage distribution by number of stories. There is no obvious trend of increasing damage rate with number of stories. Although more two storey buildings collapsed than single storey, the rate of destroyed buildings (D4+D5) is fairly similar. The 16 4-storey buildings performed on average worse than the lower buildings, but the sample size is too small for this to be a sound conclusion from the data.

Table 4 Pictometry-based damage distribution by number of stories

Number of stories	D5	D4	D3	D2	nvd	D4+D5
1 storey (n=511)	11.0%	11.2%	11.7%	10.4%	55.2%	22.1%
2 stories (n= 548)	16.4%	8.4%	11.5%	10.8%	52.4%	24.8%
3 stories (n=117)	12.8%	6.0%	8.5%	8.5%	61.5%	18.8%
4 stories + (n=16)	25.0%	6.3%	12.5%	0.0%	56.3%	31.3%

4. *Comparison with the GEO-CAN Phase 2 damage assessment.*

By looking at the GEO-CAN Phase 2 map and the map of building samples together, it is possible to build an error matrix for comparison of the results of the two surveys, for the 1241 buildings included in both surveys, The error matrix is shown in Table 5, by numbers, and Table 6 by percentages.

Table 5 Comparison of Pictometry and GEO-CAN assessments by numbers of buildings. Nvd:non-visible damage

		GEOCAN			
		0	D4	D5	*Total*
Pict	nvd	631	16	14	*661*
	D2	103	11	11	*125*
	D3	104	22	17	*137*
	D4	70	23	22	*115*
	D5	46	27	130	*203*
	Total	*954*	*99*	*194*	*1241*

Table 6 Comparison of Pictometry and GEO-CAN assessment by percentages of Pictometry damage classes

GEOCAN	0	D4	D5	*Total*
nvd	95.5%	2.4%	2.1%	*100.0%*
D2	82.4%	8.8%	8.8%	*100.0%*
D3	71.5%	16.1%	12.4%	*100.0%*
D4	60.9%	20.0%	19.1%	*100.0%*
D5	22.7%	13.3%	64.0%	*100.0%*

From these tables the following conclusions can be drawn.

1. Of the 1241 data points, the proportions given as D5 and D4 were 16.4% and 9.3% by Pictometry, and 15.6% and 8.0% by GEO-CAN. Thus the overall estimate of the major levels of damage given by the two studies is quite close.
2. Of 203 individual buildings identified as D5 by Pictometry, 130 (64%) were identified as D5 by GEO-CAN, and 157 (77%) as D4 or D5.
3. Of the 661 buildings identified by Pictometry as having no visible damage, 95.5% were also not recorded as damaged in GEO-CAN.
4. Of 318 buildings identified as D4 or D5 by Pictometry, 202 (63.5%) were identified as D4 or D5 by GEO-CAN.
5. Of the 194 buildings identified by GEO-CAN as D5, 130 (67%) were also identified as such by Pictometry, and 152 (78%) were identified as either D4 or D5.
6. Of the 293 identified by GEO-CAN as D4 or D5, 69% were also identified as either D4 or D5 by Pictometry, a further 13% were identified as D3, 7.5% as D2, and 10% (30 buildings) had no visible damage in Pictometry.

Pictometry was therefore recognising a slightly larger proportion of D4 and D5 than GEOCAN, but the overall level of damage estimated by GEO-CAN was rather good. A number of those identified as collapsed by Pictometry but not by GEO-CAN were pancake or lower storey collapses where the roof shape was unchanged, and therefore not visible from the vertical satellite image.

The major mismatches are those 30 buildings identified as D4 or D5 in GEO-CAN but having no visible damage in Pictometry. A number of these are likely to be due to the area identified in GEO-CAN as being damaged at D4 or D5 being larger than the actual damaged area.

5. ***Investigating the large discrepancies between the GEO-CAN-Phase II results and Pictometry results***

Comparison between the GEO-CAN phase II and Pictometry damage assessments has been carried out. Of the discrepancies, the buildings with the largest discrepancies have been revisited and reasons for the discrepancy analysed. The main type of discrepancy that has been analysed are the cases where GEO-CAN assigned D5 or D4 whereas Pictometry assigned ‘nvd’(non-visible damage.

Reasons for GEO-CAN= D5/Pictometry= less than D3

There are three main reasons for this type of discrepancy. The first type is when the GEO-CAN footprints were inaccurate, merging two or more buildings into one footprint. Figure 8 shows two examples where the GEO-CAN footprint includes more than one Pictometry building point. If the Pictometry building is smaller and the building is not damaged or assigned nvd, and the GEO-CAN footprint inclues a building with D5, then the GEO-CAN footprint will be assigned D5. Half of the discrepancies where GEO-CAN=D5 and Pictometry=nvd (no visible damage) were of this type. When comparing the number of Pictometry building points that intersect with the GEO-CAN footprints against the number of intersecting GEO-CAN footprints, there are 253 GEO-CAN footprints that contain Pictometry building points, and within these footprints there are 292 Pictometry building points. This highlights the need for pre-defined building footprints that can be used for all damage assessments.

*Figure 8 (left) Four Pictometry buildings contained in one GEO-CAN building footprint drawn in red. Pictometry shows that Bdg 39.3 and Bdg 39.4 is one building. (Top right) Same four buildings seen in Pictometry from west to east, (bottom right) same set of buildings seen from east to west. *Note: in the GEO-CAN dataset dated 18th February 2010, the footprint for this building is smaller, with only Bdg 39.3 delineated as D5. The analysis carried out for this report is based on a GEO-CAN dataset dated 24th February 2010, which is shown in figure 8 above.*

The second reason for this type of discrepancy is due to the building being under construction. There were instances where GEO-CAN assigned D5 to these buildings under construction whereas Pictometry assigned nvd, based on the fact that the building does not show any sign of damage. This clearly highlights the need to put in place a rule that defines how buildings under construction should be dealt.

The third reason for the large discrepancies includes cases where the damage interpretation was clearly wrong (GEO-CAN assigning D5 to buildings that do not seem D5 in the vertical aerial photographs used for Phase II), or when buildings are set back from the main road at an angle i.e. front façade not in line with the neighbouring buildings.

Another observation made through the analysis of these discrepancies is that the definition of D5 needs to be made clear. Does D5 include partial collapses, or is it only used for complete collapses? According to the EMS98 scale, D5 only applies to buildings that have completely collapsed (heap of rubble), however in some instances in GEO-CAN, buildings with partial collapse were assigned D5. The definition of partial collapse/complete collapse will be affected by the definition of the footprint of a building.

Reasons for GEO-CAN=D4/Pictometry= less than D2

Most of the reasons listed above for GEO-CAN=D5/Pictometry= less than D3 also apply here. In addition, there is a tendency in the GEO-CAN damage assessment where if there is collateral damage from neighbouring buildings and/or rubble (or objects resembling rubble) is visible on the roof, the building is classified as D4 in GEO-CAN. Similar to these collateral damages, if there is material (e.g. timber) visible in the back yard or on roof tops, these buildings are likely to be classified as D4 in GEO-CAN. In developing countries, it is quite common to see materials on the roofs of buildings as seen in figure 9 below, which makes it difficult to assess whether the building is damaged or not.

(Google vertical aerial photograph)

(Pictometry)

Figure 9 An example of a building that was classified as D4 in GEO-CAN(left) but nvd (no visible damage) in Pictometry. The objects seen on the rooftop give the impression of damage to the roof/building.

6. Approximate extrapolation to assess overall building damage

Based on the Pictometry data, and assuming this is a good sample of all Port au Prince's buildings, the overall damage distribution is D5 16.4%, D4 9.2%, D3 11.0%, the remainder D2 or less. Similar proportions are known for the 4 separate use classes. These estimates may be used to assess the likely distribution of overall damage in Port-au-Prince, using a plausible extrapolation to make an estimate of the proportions of buildings at lower damage states. Appendix 1 proposes a method to make such an extrapolation based on the assumption of that the overall damage distribution follows the binomial form. The extrapolation involves determining the binomial coefficient which best corresponds to the observed combined proportion damaged at levels D4 and D5, and using this value to estimate the proportions damaged at levels D0, D1, D2 and D3. The proportions at damage level D4 and D5 are assumed to be as observed. This process leads to the overall damage distributions shown in Figure 7 and Table 7.

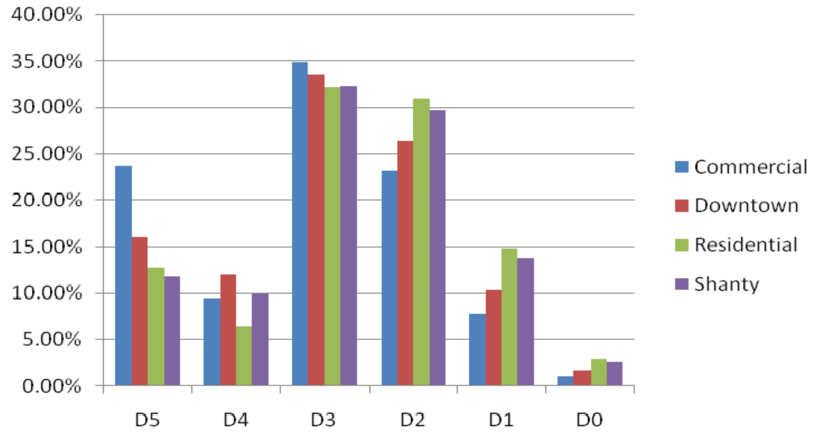


Figure 7 Estimated overall damage distributions assuming binomial distribution governs lower damage levels

Table 7 Estimated overall damage distributions as shown in Figure 7

	Commercial	Downtown	Residential	Shanty
D5	23.68%	16.08%	12.66%	11.86%
D4	9.47%	12.06%	6.49%	9.89%
D3	34.84%	33.53%	32.20%	32.23%
D2	23.23%	26.35%	30.93%	29.75%
D1	7.74%	10.35%	14.86%	13.73%
D0	1.03%	1.63%	2.86%	2.54%

Note that this procedure results in estimates of the proportions of buildings at damage levels D3 and D2 higher than actually observed by a factor of 2-3. This is in accordance with the expectation that (as noted above), many instances of these lower damage levels would not be visible in the Pictometry survey. A better assessment of the error in the assignment of damage levels D3 and D2 will be possible on the basis of the analysis of ground photos of damage to the same buildings, which will form the next part of this study.

7. Validating the Pictometry damage assessment using EEFIT ground survey data

A field mission to Haiti was launched by the British Earthquake Engineering Field Investigation Team (EEFIT) between the 6th and 13th of April, 2010, funded by EPSRC, UK. One of the main objectives of the field work by the EEFIT team, of which co-author of this report Keiko Saito was a member, was to carry out surveys to assign damage levels to a number of sampled buildings using the EMS98 scale. The EEFIT survey results have been used here to assess the accuracy of the Pictometry damage assessment results.

Brief summary of the field survey methodology

Since one of the main intended use of the field survey data was to assess the accuracy of the Pictometry damage assessment, it was decided during the planning of the mission that the field team should survey the same buildings on the ground as those included in the Pictometry data. Of the 60 pictometry survey locations indicated in Figure 1, 15 were chosen as candidate locations, using the parameters listed in the headings of the columns in Table 8, making sure to select locations with diverse characteristics. Of the 15 locations, the EEFIT team was able to visit 8 locations in the 3.5 days that was spent doing the building surveys, and in total 124 buildings were surveyed in Port au Prince. Table 8 shows the characteristics of the 8 locations visited, figure 8 shows the 8 locations in Google earth.

Table 8. The characteristics of the 8 locations visited by the EEFIT team, identified by desktop study before the field visit.

ID	*Land_use*	*Topography*	*Soils*	*Comment*
L1	Residential	Flat	Alluvial	Near port
L22	Residential	Sloping	Rock?	Near edge of very steep slope
L3	Residential-likely shanty	Flat; 40m	?	On a major street intersection; close to camp and to 31
L55	Downtown	Flat and low	Sandy?	Large govt buildings expected
L9	Residential?	Sloping; 100m		Difficult access?
L33	Shanty edge	Flat 105	Alluvial?	Close to major factory complex
L41	Commercial?	Sloping; 85m	?	Close to 43 and to steeper slope
L17	Shanty edge	Sloping; 135m	Rocky	Close to very steep slope up; could be done with 41 & 43

Figure 8 The eight locations visited in Port au Prince, Haiti, by the EEFIT team to carry out a systematic building survey. The number of buildings surveyed in each location varied between 10 to 20 buildings, with the exception of location 41 where 42 buildings were surveyed.

Each building was visited on foot, and the damage assessment was carried out mainly by Edmund Booth, team leader of the EEFIT team and earthquake engineer with 30 years of experience. For 17 out of the 124 buildings, the survey team was given access to the inside of the property, which allowed the team to assign a damage level with confidence. The parameters recorded for each building followed the format of the Pictometry damage assessment, i.e. number of storeys, construction type, use type, damage level (EMS98) and comments. The results were tabulated and compared against the Pictometry damage assessment results.

Comparison between Pictometry damage assessment and EEFIT ground survey

Figure 9 shows the comparison of the damage level distribution from the two surveys. The results from the Phase II GEO-CAN damage assessment using Aerial Photography has also been added for reference. Since GEO-CAN only identified D4 and D5s, comparison has been made for those two damage levels only.

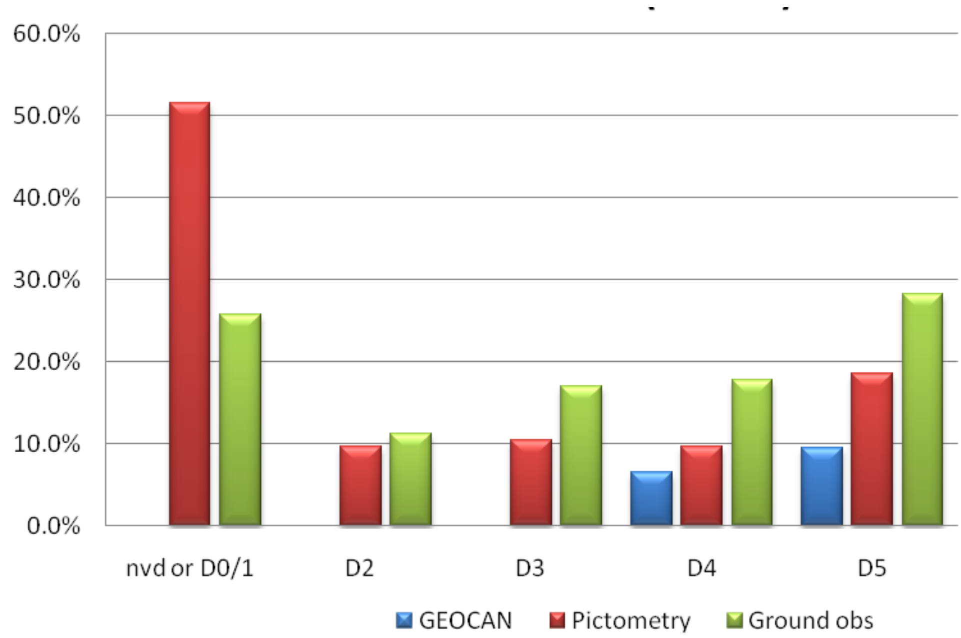


Figure 9. Comparison of GEO-CAN, Pictometry damage assessment and EEFIT ground survey results. N=124.

The differences between the three approaches in the proportions of buildings at damage levels D4 and D5 are striking. Among the 124 buildings included in the ground observation survey, the observed proportions at damage level D4 and D5 were 18% and 28% respectively. The proportions seen in the Pictometry survey were 10% and 19% at D4 and D5, ie only 63% of the actually observed number, while those observed through the GEOCAN study were 7% and 10% at D4 and D5, ie only 40% of the actually observed number.

Another notable difference between the Pictometry and ground observation results is seen in class *nvd or D0/1*. The difference is almost two fold. This can be explained by the fact that nvd (Pictometry) includes the buildings where the view was obscured due to trees or other factors, as well as the true D0/1s. Table 9 shows the breakdown of the number of buildings assigned nvd in Pictometry that was assigned a damage level of more than D2 on the ground by the EEFIT team.

In Pictometry, the view for most of the buildings is somewhat obscured, meaning it is very likely that at least one of the facades is not visible due to its closeness to the neighbouring buildings. The presence of a tall tree in the immediate vicinity would also obscure the view of a building. In the former case, it may still be possible to see the damage on the visible sides of the buildings. Even when there are trees in the vicinity, in some cases some damage may be visible and some not. In future studies, it may be beneficial to describe in more detail the reasons for assigning nvd to a building so that nvd can be analysed in more detail.

Table 9 Statistics on the number of buildings assigned nvd using Pictometry, of which were assigned D2 or above by the EEFIT ground survey team. GO=ground observation. Nvd=non visible damage

	Total	Of which view of building seriously obscured in Pictometry	Of which soft-storey collapses
Nvd in Pictometry but D2 in GO	11	0	N/A
Nvd in Pictometry but D3 in GO	12	0	N/A
Nvd in Pictometry but D4 in GO	11	0	N/A
Nvd in Pictometry but D5 in GO	5	2	2
	39		

Out of the 124 buildings surveyed by the EEFIT team, 72 of them were assigned nvd (non visible damage) using Pictometry, of which 39 (54%) were identified as D2 or above on the ground. The main reason for the discrepancies here is that the damage is truly not visible in Pictometry. Two questions arise from the table above:

1. Can we assume that approximately 50% of nvd will in reality be D2 or above in other earthquakes? Will this proportion change in areas with different intensities?
2. Can we apply the proportion of D2, D3, D4 and D5 in the table above to reassign half of the nvd in Pictometry in future events?

More case studies are needed to answer these questions.

If we consider a difference of more than two damage levels to be a serious error for the Pictometry damage assessment when compared to the EEFIT survey, the breakdown of the serious errors is shown in table 10.

Table 10 Breakdown of the errors with more than two damage level difference between ground observation by EEFIT and Pictometry assessment

	Total	Of which view of building seriously obscured in Pictometry	Of which soft-storey collapses
D5 in GO but nvd or D3 in Pictometry	12	4	6
D4 in GO but nvd or D2 in Pictometry	13	0	N/A
D3 in GO but nvd in Pictometry	10	0	N/A
Pictometry assessment worse than GO	11	0	N/A
Total	46	4	6

Some examples from table 10 are shown in Figure 10.

Figure 10. Examples of D5 (EEFIT) – nvd or D3 in Pictometry

These buildings were assigned D5 by the ground survey, whereas in the Pictometry survey they were assigned either nvd or D3, and represent the largest discrepancies. Pictures (1) and (3) show soft-storey collapses typically seen in a residential area. Picture (2) shows a building where one of the columns has been damaged, causing one corner of the roof to fall. The building in picture (4) has partial collapse where one of the corners, including a column, has completely failed. The tree in front obscured the view in Pictometry, hiding the failed section of the building. The view of buildings (1) and (3) in Pictometry is show in Figure 11.

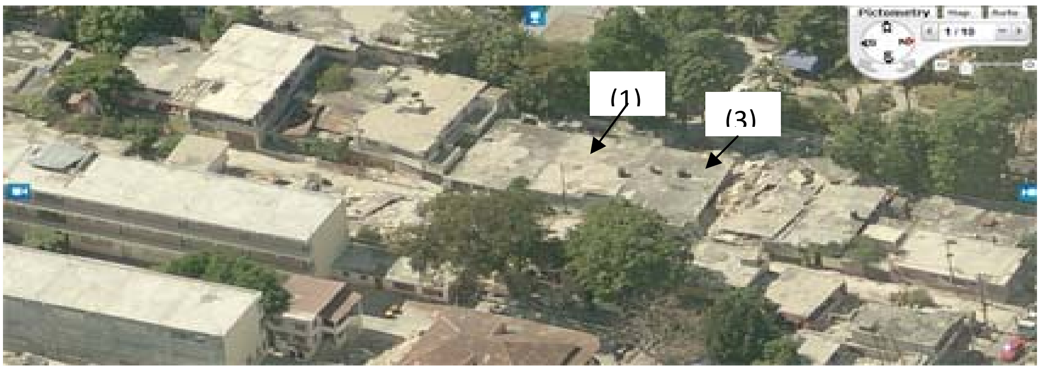


Figure 11 Buildings (1) and (3) in Figure 10 as seen in Pictometry.

It remains to be seen whether the number of damage interpretations with these large discrepancies can be reduced. Reducing this type of error will be essential if the accuracy of the overall damage assessment using Pictometry, or any other imagery, is to be improved.

The question of sample size

As seen in Figure 9, a surprising amount of discrepancy between the number of buildings assigned to D5 and D4 is observed between the three methods. In particular, the discrepancy observed between the D5 and D4 of the Pictometry and Geocan results is unexpected, since the two assessments produced comparable ratios of D5 and D4 in the previous comparison described on page 6 (bullet point 1) of this report. One possible explanation is that the sample of buildings included in the ground survey is not adequately representative of the dataset included in the Pictometry survey; another possible explanation for the discrepancies is in the sample size. Here the sample size is 124, whereas the sample size for the Pictometry study is 1241. It is clearly important to consider the effect of sample sizes on the assessment of the overall proportion of the damage levels.

Standard sampling theory (Hammond, R and McCullagh, P. S., 1978 has been used here to obtain an indication of the likely achievable accuracy using various sample sizes for the ground data. To make the approach applicable to this dataset, the data has been regrouped from the five damage levels into two, i.e. D4-5 and D3-2-1. Table 11 shows the proportion of each regrouped damage level found in the sample buildings surveyed by the EEFIT team.

Table 11 Proportion of D5/4 against D3/2/1 in the ground validation dataset surveyed by the EEFIT team. N=124

D5/4	D3/2/1	total
46.0%	54.0%	100.0%

Assuming that the 124 buildings in the EEFIT dataset are representative of the entire building stock, and considering the above assessment result as a pilot survey, at the 95% confidence level the error margin of the **proportion of the D5/4 for the total statistical population** is ±8%. This error margin of the proportion of D4/5 in the sample can be derived using the formula below:

$$N = p\% \times q\% \times (z/d)^2$$

Where N is the sample size, p is the proportion of category 1 (in this case D4/5), q is the proportion of category 2 (in this case D3/2/1), z is the z value at the 95% confidence level and d is the tolerable error margin included in the results when using sample size N. Given the sample size of the EEFIT survey (n=124) and solving for d, the error margin d is approximately 8% (i.e. ±8% of 46% for D5/4). If the EEFIT sample size was 1241, then the error margin is reduced to ±2.7% (i.e. between 43.2% and 48.7%). The advantage of having a larger sample size is obvious. However when planning for a field survey, practical considerations such as man power, time, accessibility should be considered.

The EEFIT survey results can also be compared to the UNOPS damage assessment data[1] that includes the assessment of more than 11,000 buildings carried out on the ground. The UNOPS data indicates that 35 % of the buildings surveyed so far are red tagged (equivalent of D5/4?) as of 14th May 2010. When comparing this figure to the EEFIT survey result in table 11, the UNOPS value falls just outside of the estimated confidence interval (i.e. 37.2%-54%) derived using the EEFIT data. Hence a sample size of 124 would be appropriate to estimate the proportion of D5/4 with an error margin of ±8%. However to reduce the possible error margin, it is necessary to have a larger sample. In addition, if the current PDNA loss estimate method is to be used, assuming different damage rates in different land-use classes, a sample of 130 buildings or so would be required from each land cover class used in the PDNA loss estimate to estimate the proportion of the D5/4 in each land cover class within an 8% error margin. In addition, considering that the level of ground shaking is the principal cause of damage, samples from area of different ground motion intensity will also be needed for the overall assessment of damage for the entire affected area. Appendix B includes a table that shows, using sampling theory the sample size required to achieve a certain error of margin in the assessment of the proportion of buildings in damage classes D4 and D5..

8. Discussion: future work

Of all the earthquake damage survey methods, field damage surveys will likely remain the most accurate in the coming years. However it will always be difficult to carry out a systematic survey amidst the chaos in the immediate aftermath of the event. For this reason, remote damage surveys using imagery, using either satellite or airborne data (whatever is available), will be able to provide us with a valuable insight into the distribution of damage on the ground in the early days after the event. The Haiti earthquake has provided us with an opportunity to further investigate the errors likely involved in damage assessments done using imagery.

Statistical methods to reassess the imagery based damage assessment results using the field survey results need to be developed. There are potentially two approaches; the first being using maximum likelihood, and the second to model the distribution using Bayesian statistics, although the latter will require more case studies and damage distribution from historic earthquakes to set the prior.

For PDNA loss estimation, the assessment of the damage level distribution is only one factor. Other factors include identifying the average floor space for each land cover (or other categories) class, estimation of the (re)construction cost for various building types. There is a need to identify the other factors that need to be taken into account e.g. ground motion, geology, topography.

The PDNA process itself, as was discussed during the JRC meeting on the 20-21st May, 2010, is not yet fully finalised. The time is right to develop a damage assessment methodology using remote sensing that is transferable, taking into account the operational time constraints within the context of the PDNA and available resources, so that the use of RS for damage assessment can be fully incorporated into the PDNA process.

There is a need for a review of the damage level definitions to be used for remote sensing damage assessment. The Italian system (Goretti et al 2002) in which the damage definition is provided for both the

[1] Personal communication with Ian Gough, UNOPS, Haiti on 15th May 2010.

horizontal and vertical components is one possibility that may be beneficial to the remote sensing community.

And finally, more studies need to be done to assess whether the Binomial distribution method can be used to estimate the lower damage levels, given the D4 and D5s. The Haiti EEFIT survey revealed that the D4 and D5s assessed using Pictometry potentially has a significant omission error. To improve the accuracy of the overall damage assessment, firstly the omission errors in the higher damage levels must be reduced.

9. Conclusions

- Pictometry has been shown to be a highly effective tool for quickly observing and recoding post-event damage data.
- The Pictometry data reveals a significant amount of serious damage (levels D4 and D5) which is not visible in the vertical aerial photographs used for the GEO-CAN study. Much of this is damage to lower stories which can be seen in the oblique Pictometry images but not in a vertical view.
- Pictometry data also enables some but not all damage at lower damage levels (D2 and D3) to be identified, which cannot be seen at all in the vertical aerial imagery.
- Ground observation has shown, however, that even the Pictometry images are not good enough to spot all the damage at the most serious levels. In a relatively small sample of buildings examined in detail on the ground, 46% of the buildings were observed to have damage at levels D4 or D5, only 63% of which were identified as having these damage levels in Pictometry .
- Studies of individual buildings show that the principal causes of these discrepancies are lower-storey collapses which were not visible even in Pictometry, and cases where the key Pictometry image was obscured either by trees or adjacent buildings.
- It seems probable that proportions determined from Pictometry should be increased by about 50% for a good estimate of the proportions of buildings damaged at level D4 and D5; and that proportions determined from the GEOCAN approach should be doubled for a good estimate of buildings damaged at levels D4 and D5.
- A binomial distribution approach has been proposed here to assess the proportions of buildings in lower damage levels based on better estimates of the proportions damaged at levels D4 and D5.
- However, the ground observation dataset used for this study is too small to be a sound basis for proposing reliable "scaling factors" to apply to aerial observations. Further study will be needed using additional ground observations from Haiti (which may become available from the UNOPS damage assessment currently being carried out), and from earthquake damage elsewhere.

References

Hammond, R and McCullagh, P. S., 1978, Quantitative Techniques in Geography: an introduction. Second edition. ClarendonPress, Oxford.

JRC, 2010. *Building Damage Assessment Report, Haiti Earthquake of 12th Jan 2010, Post Disaster Needs Assessment and Recovery Framework*, JRC, UNOSAT, World Bank GFDRR, and CNIGS, Version of March 11th, 2010.

Goretti, A and Di Pasquale, G, (2002). An overview of post-earthquake damage assessment in Italy, *EERI Invitational Workshop*, Pasadena.

Appendix A: Use of the binomial distribution for defining damage probability matrices: application to Haiti

Background

Braga et al (1982) , analysing the damage data for 41 towns in Southern Italy affected by the 1980 Irpinia earthquake, showed that the damage distributions across the 6 damage levels D0...D6 approximate well to binomial distributions. In this approach a single parameter p defines the entire distribution, with the proportion of the population in each of the 6 damage levels being defined by
the terms of the binomial expansion of the expression $((p + (1\text{-}p)))^5$. Thus the proportion in D0 is $(1\text{-}p)^5$; in D1 is $5.p.(1\text{-}p)^4$; in D2 is $10.p^2.(1\text{-}p)^3$; in D3 is $10.p^3.(1\text{-}p)^2$, in D4 is $5.p^4.(1\text{-}p)$; and in D5 is p^5. The sum of these proportions is necessarily 1.0.

If the binomial distribution was exact, then if the proportion in any damage state were known, this could be used to define the binomial parameter. But of course all real distributions deviate to some degree from the binomial, so choosing to determine the binomial parameter from any one part of the distribution involves an error, which will be greater the smaller is the known proportion of the distribution.

Haiti damage assessment needs

For use in the Haiti situation, an estimate of the total distribution is required to be made based on damage assessments derived from satellite image interpretations, in which only the proportion of buildings in damage states D4 and D5 can be estimated with confidence. With the use of Pictometry data, it is possible that the proportion in damage state D3 can additionally be estimated. Can good extrapolations be made from this data to estimate the total damage rapidly enough to inform the PDNA?

To test this, an analysis has been made of some damage data assembled after the Umbria Marche earthquake. The data comes from 5 locations with intensities (MCS scale) from VI to VIII in which some 5000 masonry buildings were surveyed (Dolce, Masi and Goretti, 1999). The total distribution and the true best-fit binomial parameter, p, was known for each of these datasets. For each of the 5 datasets, estimates of p were derived from just the D5, then from the D4 +D5, and then from the D3+D4+D5 parts of the dataset. The results are plotted in Fig A1.

Fig 1 shows that the estimate using D5 only is poor; it is in error on average by 12.7%, and the regression coefficient (R^2) of estimated p on actual p is less than 10%. When the D4 data is added in, the estimate improves greatly; the average error is 5%, and R^2 increases to 56%. When D3 data is added, the estimate becomes good; the average error is only 3%, and R^2 increases to 95%.

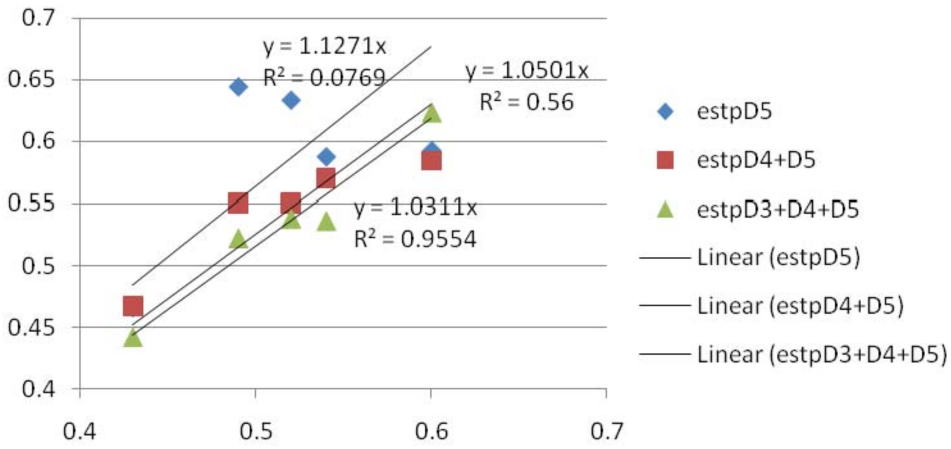


Figure A1. Estimates of binomial parameter p plotted against "true" binomial parameter for 5 damage datasets derived from surveys after the 1997 Umbria Marche earthquake (Dolce et al 1999)

This suggests that estimates of the total damage distribution based only on the totally collapsed buildings will be poor. If we can confidently determine the D4s and D5s, the estimate may be good enough for an initial assessment of the total damage. With a knowledge of the proportion of D3s as well, a rather good estimate would be possible. The actual ground shaking intensity does not need to be known for this analysis. Indeed the damage data should be used to define the intensity distribution afterwards. The quality of the estimate will of course depend on the proportion of heavily damage buildings in the dataset. In Haiti it is higher than it was in Umbria Marche, which suggests that the estimate of p based on a part of the distribution should be better for Haiti.

Conclusion.

It is worth putting some effort into making an estimate of the proportions of buildings damaged at D3. With this and a good estimate of the proportions in D4 and D5, an extrapolation to estimate the total damage can be made with an accuracy which is probably acceptable for the PDNA.

References

Braga, F, Dolce, M and Liberatore, D, 1982 A statistical study on damaged buildings and an ensuing review of the MSK 1976 scale, Ch 5 in *The Southern Italy 23.11.1980 earthquake,* CNR Publication 503, Rome

Dolce, M, Masi, A and Goretti, A, 1999 Damage to buildings due to the 1997 Umbria-Marche Earthquake, in Bernardini ed *Seismic Damage to Masonry Buildings*, Balkema, Rotterdam ISBN 90 5809 115 5

Appendix B: Required sample size according to the margin of error allowable and estimated proportion of the two categories D5/4 and D3/2/1. Taken from Hammond and McCullagh, 1978.

D5/4(%)	D3/2/1(%)	z (95%)	d (margin of error%)	S (sample size)
25	75	1.96	1	7203
10	90	1.96	1	3457
20	80	1.96	1	6147
30	70	1.96	1	8067
40	60	1.96	1	9220
50	50	1.96	1	9604
10	90	1.96	2	864
20	80	1.96	2	1537
30	70	1.96	2	2017
40	60	1.96	2	2305
50	50	1.96	2	2401
10	90	1.96	5	138
20	80	1.96	5	246
30	70	1.96	5	323
40	60	1.96	5	369
50	50	1.96	5	384
10	90	1.96	6	96
20	80	1.96	6	171
30	70	1.96	6	224
40	60	1.96	6	256
50	50	1.96	6	267
10	90	1.96	7	71
20	80	1.96	7	125
30	70	1.96	7	165
40	60	1.96	7	188
50	50	1.96	7	196
10	90	1.96	10	35
20	80	1.96	10	61
30	70	1.96	10	81
40	60	1.96	10	92
50	50	1.96	10	96