Natural Hazards
and Earth System
Sciences

Detection of collapsed buildings from lidar data due to the 2016 Kumamoto earthquake in Japan

Luis Moya[1], **Fumio Yamazaki**[2], **Wen Liu**[2], **and Masumi Yamada**[3]

[1]International Research Institute of Disaster Science, Tohoku University, Miyagi, Sendai, 980-0845, Japan
[2]Department of Urban Environment Systems, Chiba University, Chiba 263-8522, Japan
[3]Disaster Prevention Research Institute, Kyoto University, Gokasho, Uji, 611-0011, Japan

**Correspondence:** Luis Moya (lmoyah@irides.tohoku.ac.jp)

Abstract. The 2016 Kumamoto earthquake sequence was triggered by an M_w 6.2 event at 21:26 on 14 April. Approximately 28 h later, at 01:25 on 16 April, an M_w 7.0 event (the mainshock) followed. The epicenters of both events were located near the residential area of Mashiki and affected the region nearby. Due to very strong seismic ground motion, the earthquake produced extensive damage to buildings and infrastructure. In this paper, collapsed buildings were detected using a pair of digital surface models (DSMs), taken before and after the 16 April mainshock by airborne light detection and ranging (lidar) flights. Different methods were evaluated to identify collapsed buildings from the DSMs. The change in average elevation within a building footprint was found to be the most important factor. Finally, the distribution of collapsed buildings in the study area was presented, and the result was consistent with that of a building damage survey performed after the earthquake.

1 Introduction

The detection of affected areas after an earthquake is very important for disaster response activities. Allocating resources such as relief forces, food, medicine, and shelter is crucial after a natural disaster strikes (Das and Hanaoka, 2014). Thus, proper information on the damage situation will improve the efficiency of distributing relief resources. The extent of the affected area also provides an idea of the scale of the disaster and an estimate of the relief demand. Damage assessment after an earthquake disaster is important for the scientific community as well. A significant amount of information has been

obtained from previous earthquakes and used to improve construction design codes to evaluate and mitigate damage to buildings and infrastructure in the event of future earthquakes. For instance, Whitman et al. (1973) provided earthquake damage probability matrices using data collected after the 1971 San Fernando, California earthquake. Yamazaki and Murao (2000) proposed vulnerability functions for Japanese buildings based on building inventory and damage data and the spatial distribution of strong motion (Yamaguchi and Yamazaki, 2001) during the 1995 Kobe earthquake in Japan.

Information gathered from field surveys is invaluable and very precise; however, the process requires significant time and effort, and access to affected areas is often hindered by road closures and secondary hazards. Remote sensing is an effective tool for detecting damaged areas because it can be used to document damage to large areas without direct access to the affected area (Yamazaki and Matsuoka, 2007; Rathje and Adams, 2008; Dell'Acqua and Gamba, 2012). Immense improvement to the accessibility of remote-sensing imagery data and geospatial data processing tools has been achieved over the last several years (Vuolo et al., 2016; Korosov et al., 2016). A dramatic increase in the number of satellite, aircraft, and unmanned aerial vehicle (UAV) sensors has been observed as well. One of the most successful approaches for assessing damaged areas is based on change detection between a pair of images taken before and after an earthquake (Meslem et al., 2011; Liu et al., 2013; Uprety et al., 2013). In addition, remote sensing has been used for long term urban recovery monitoring (Hoshi et al., 2014; Hashemi-Parast et al., 2017).
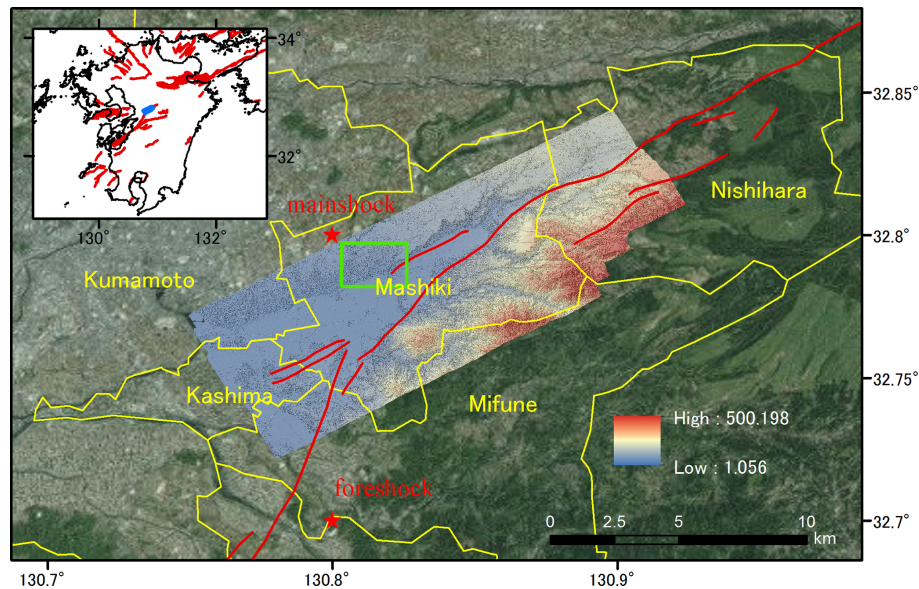
Figure 1. The post-event lidar data for the study area. Shaded colors represent the elevation. The green rectangle shows the locations of the area surveyed by Yamada et al. (2017). The inset shows Kyushu Island, and the blue polygon in the inset depicts the study area.

Schweier and Markus (2006) pointed out that airborne light detection and ranging (lidar) data can be used to classify collapsed buildings using the following geometrical features of a building extracted from lidar data: the height change from the initial one, the reduction of the total volume, the footprint borders, the inclination of the structure, the debris spread outside the footprint, the additional covered area outside the footprint, and the damage situation of the roof. They proposed a modification of the previous damage classification method (Okada and Takai, 2000) using these geometrical features. Although they suggested the use of airborne lidar data to analyze collapsed buildings, applications to real cases were not provided.

Applications of lidar for damage detection are still few compared with other remote-sensing technologies. The main reason is the lack of lidar data before a disaster. However, Aixia et al. (2016) performed a study on the possibility of detecting building damage using only a post-earthquake lidar digital surface model (DSM). Their results are promising for buildings with simple roof shapes, such as flat and pitched roofs. Rehor et al. (2008) proposed the use of a plane-based segmentation method to detect damaged buildings, wherein the number of unsegmented pixels in damaged buildings is larger than in undamaged buildings. Labiak et al. (2011) proposed an automated method to detect and quantify building damage using only a post-earthquake lidar DSM as well, but their results had low accuracy for heavily damaged and collapsed buildings. Hussain et al. (2011) combined lidar data with GeoEye-1 imagery to detect damaged buildings after the 2010 Haiti earthquake. They detected 190 damaged buildings out of 200; however, their procedure required manual intervention, and the damage level was not clearly classified.

Instead of lidar data, Maruyama et al. (2014) constructed two DSMs from two sets of aerial images: before and after the earthquake. Then, the collapsed buildings after the 2007 Niigata-Chuetsu-Oki earthquake, Japan, were identified using the difference in elevation between the DSMs.

An M_w 6.2 earthquake struck Kumamoto Prefecture, Japan, on 14 April 2016 at 21:26 JST. The event produced structural damage and resulted in nine human casualties (Cabinet Office of Japan, 2017). Then, 28 h later, a second earthquake with M_w 7.0 occurred close to the first one. Thus, the first event was designated the "foreshock" and the second the "mainshock". The epicenter of the foreshock was located at the end of the Hinagu fault, and the epicenter of the mainshock was located in the Futugawa fault. Both events were located in the town of Mashiki, which has a population of 33 000. The number of aftershocks following these events reached the largest number among recent inland earthquakes in Japan (Japan Meteorological Agency, 2017). The total number of deaths due to direct causes reached 50, and over 8000 residential buildings were severely damaged or collapsed due to the Kumamoto earthquake sequence.

Among the several remote-sensing technologies used to monitor the area affected by the Kumamoto earthquake (Yamazaki and Liu, 2016), a pair of lidar data sets taken before and after the mainshock were available (Moya et al., 2017). As mentioned before, this kind of data set is not often available. Therefore, this study explores the potential use of lidar data to identify collapsed buildings over the affected area. Building collapse is still the main cause of casualties and hence its prompt recognition is crucial for search and rescue operations. The difference in elevation, the standard devia-

tion, and the correlation coefficient are tested for this purpose.

2 Study area and data set

After the foreshock, a lidar-surveying flight was carried out during 15:00–17:00 (JST) on 15 April 2016 in order to record the effects of the earthquake (Asia Air Survey Co., Ltd., 2017). It produced point clouds with an average point density of 1.5–2 points m^{-2}. Subsequently, because the unexpected mainshock occurred, a second mission was set up during 10:00–12:00 (JST) on 23 April 2016, which produced point clouds with an average point density of 3–4 points m^{-2}. Both sets of lidar data were acquired using a Leica ALS50II instrument and the same pilot and airplane. After rasterization of the raw point clouds, two DSMs with a data spacing of 50 cm were created. The DSMs collected before and after the mainshock will hereafter be referred to as the BDSM and ADSM. Figure 1 shows the extent of the ADSM, which represents the entire study area. It covers the main part of Mashiki and some parts of Nishihara village, Mifune and Kashima towns, and Kumamoto city.

The study area is located in the near field of the Kumamoto earthquake sequence where significant permanent ground displacements were produced during the earthquake. A direct comparison of the BDSM and ADSM shows that the building coordinates do not match because the ADSM contains coseismic displacements. Therefore, the ADSM was shifted before detecting the damaged buildings based on the permanent crustal movement calculated by Moya et al. (2017). To do this, an automated procedure for calculating the permanent three-dimensional (3-D) displacement was implemented. The permanent ground displacement was calculated by 100 m grid size and then applied to the ADSM pixels within the grid size. Figure 2 illustrates the calculated permanent ground displacement of the common lidar data area. In the figure, the results of new field measurement carried out in August 2016 for surveying reference points after the Kumamoto earthquake are also shown (Geospatial Information Authority of Japan, 2017). The coseismic displacements estimated from the lidar data show good agreement with the survey results (Fig. 3). In Fig. 2, the causative fault is located in the areas where sudden changes in the direction of the permanent ground displacement are observed. Over the entire study area, a maximum horizontal displacement of approximately 2 m was observed.

3 Detection of damaged buildings

To focus on buildings, a geocoded building footprint data set, provided by the Geospatial Information Authority of Japan (GSI), was used. Only buildings with footprint areas greater than 20 m^2 were evaluated. Because the point densities of the BDSM and ADSM are different and the footprint
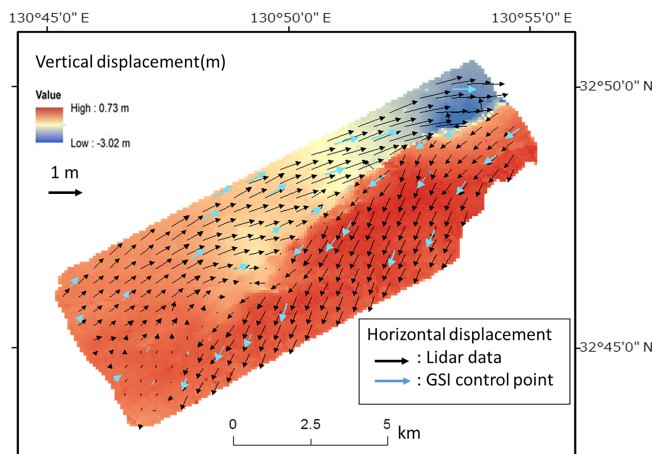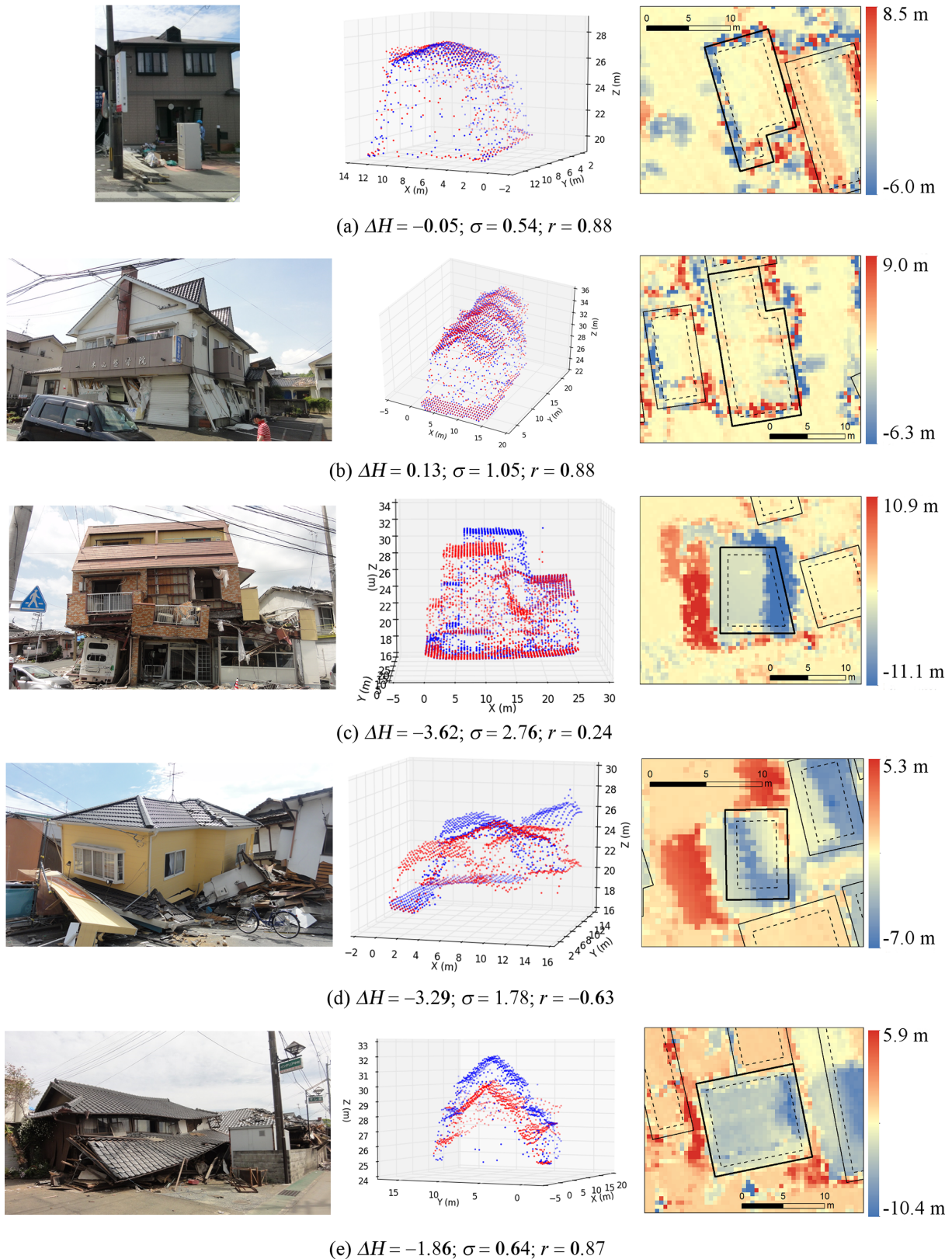

Figure 2. Estimated three-dimensional coseismic displacement after the mainshock of the 2016 Kumamoto earthquake. The black arrows and the shaded colors indicate the horizontal and vertical displacements obtained from lidar (Moya et al., 2017). The blue arrows indicate the horizontal displacements at the control points measured by the Geospatial Information Authority of Japan (2016).
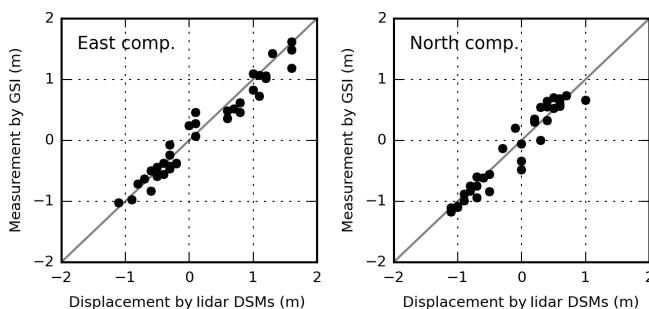

Figure 3. Comparison between the coseismic displacements estimated from the lidar data (Moya et al., 2017) and from field measurements (Geospatial Information Authority of Japan, 2016).

data include some errors, perfect matching of the DSMs with the building footprints could not be achieved. For this reason, the building footprints were reduced by 1 m (i.e., the reduced polygon is located inside a building footprint), and they were projected onto the same reference system as that of the DSMs (Fig. 4). The lidar data within the reduced building boundaries were then extracted and processed. The reason for using the reduced building boundaries was to discard the DSM data near the building boundaries in the subsequent analysis. The distance of the buffer (1 m) was decided based on a preliminary evaluation of the data (Moya et al., 2016).

Figure 4 illustrates five buildings located in the study area. For each case, the BDSM (blue dots), the ADSM (red dots), and the difference between the two DSMs are depicted. These buildings were selected in order to demonstrate different damage patterns: nondamaged, tilted, and collapsed buildings. It is worth noting that the difference between the DSMs for a nondamaged building (Fig. 4a) shows high val-

(a) $\Delta H = -0.05$; $\sigma = 0.54$; $r = 0.88$

(b) $\Delta H = 0.13$; $\sigma = 1.05$; $r = 0.88$

(c) $\Delta H = -3.62$; $\sigma = 2.76$; $r = 0.24$

(d) $\Delta H = -3.29$; $\sigma = 1.78$; $r = -0.63$

(e) $\Delta H = -1.86$; $\sigma = 0.64$; $r = 0.87$

Figure 4. Examples of collapsed buildings from the lidar data. The left column shows the photos taken after the mainshock by the authors. The middle column shows the lidar data, wherein the blue points depict the BDSM and the red points the ADSM. The right column shows the elevation differences between the two DSMs, wherein the solid lines depict the building footprints and the dashed lines depict the footprint reduced by 1 m.
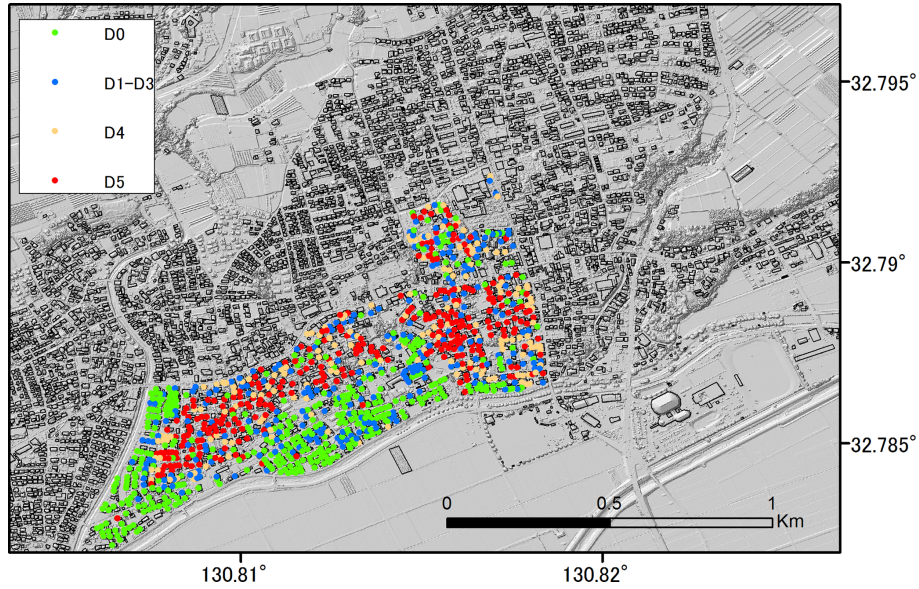
Figure 5. Building damage survey data from Yamada et al. (2017). The location of the survey area is shown in Fig. 1.

ues around the boundary of the building footprint, which was caused by the effect mentioned earlier. These errors are certainly present for tilted buildings as well and make damage detection very challenging (Fig. 4b). Figure 4c shows a typical collapsed steel-frame building with a well-known damage pattern that occurs with a soft story or a weak story, that is, a significant difference in the stiffness/resistance between one story and the rest. They show a significant horizontal/vertical movement, which is easier to detect by lidar data. Figure 4d shows a collapsed wooden building that was shifted significantly in the horizontal direction. Conversely, the collapsed wooden building shown in Fig. 4e does not exhibit a horizontal movement, only a vertical shift. Lateral spread of debris is an important issue when the building is located along a main road. For almost all the collapsed buildings, a clear decrease in building elevation was observed from the lidar DSMs.

The number of buildings within the study area is very large, so it is necessary to implement an automated procedure to evaluate the extent of their damage. In this study, three parameters were used: the average height difference between the two DSMs (ΔH) within the reduced building footprint, its standard deviation (σ), and the correlation coefficient (r) between the two DSMs. These parameters were calculated for each building using the following equations:

$$\Delta H = \frac{1}{N} \sum_{i=1}^{N} (\mathrm{Ha}_i - \mathrm{Hb}_i) \tag{1}$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N} ((\mathrm{Ha}_i - \mathrm{Hb}_i) - \Delta H)^2}{N}} \tag{2}$$

$$r = \tag{3}$$
$$\frac{N \sum_{i=1}^{N} \mathrm{Ha}_i \mathrm{Hb}_i - \sum_{i=1}^{N} \mathrm{Ha}_i \sum_{i=1}^{N} \mathrm{Hb}_i}{\sqrt{\left(N \sum_{i=1}^{N} \mathrm{Ha}_i^2 - \left(\sum_{i=1}^{N} \mathrm{Ha}_i\right)^2\right)\left(N \sum_{i=1}^{N} \mathrm{Hb}_i^2 - \left(\sum_{i=1}^{N} \mathrm{Hb}_i\right)^2\right)}}$$

where $i \in \{1, 2, \ldots, N\}$ and N is the number of elevation points inside a given reduced building footprint. Ha_i and Hb_i are the elevations from the ADSM and BDSM. The correlation coefficient ranges from -1.0 to 1.0 and has proven to be effective in detecting changes from a pair of satellite images (Liu et al., 2013; Uprety et al., 2013). A value of r close to 1.0 indicates no change.

Yamada et al. (2017) presented the distribution of building damage in the central part of Mashiki, wherein the damage was determined from aerial photos and field surveys. The damaged buildings were classified into four categories: no damage (D0), partial/moderate damage (D1–D3), severe damage/incline (D4), and story collapse (D5). Here, D1–D5 represent the degree of damage according to Okada and Takai (2000), which is similar to G1–G5 of the European Macroseismic Scale (EMS-98). Figure 5 shows the damage distribution over the surveyed area, which is located along the north side of the Akitsu River.

Figure 6 shows the scatter plots of the parameters calculated for the surveyed buildings, and Fig. 7 shows the histograms of the three parameters for the buildings with different damage levels, wherein the average (solid line) and the standard deviation (dashed line) are also included. Significant overlap of damage levels D0, D1–D3, and D4 was
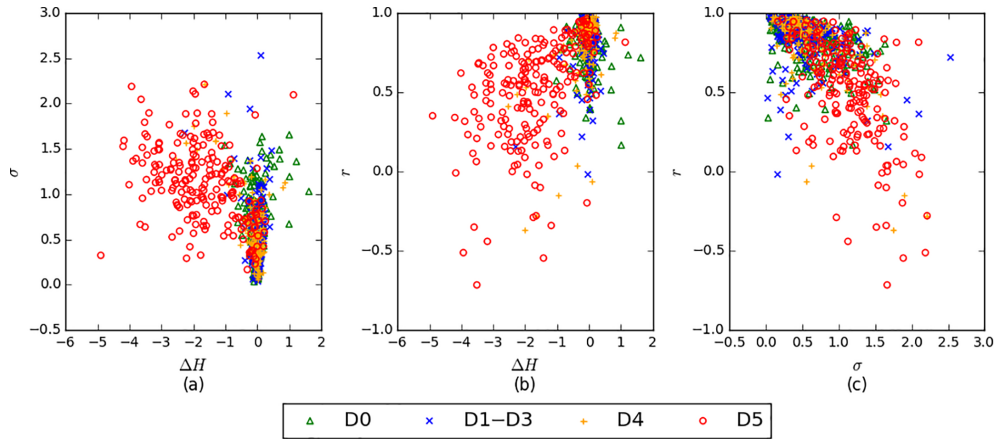
Figure 6. Scatter plots of the three parameters (ΔH, σ, r) calculated from the lidar DSMs for the buildings surveyed by Yamada et al. (2017).

observed regardless of which parameter was chosen. On the other hand, collapsed buildings (D5) tend to have large negative values of ΔH. Therefore, this paper focuses on the detection of collapsed buildings. It is important to note that few collapsed buildings show positive values of ΔH. A closer look showed that those buildings were covered by a neighboring building that had collapsed.

Although ΔH seems to be the dominant parameter for identifying collapsed buildings, the other two parameters (σ and r) can still provide additional information. For instance, if we observe the collapsed buildings from the scatter plot in Fig. 6c (red marks), a trend can be observed in which r trends to one when σ is close to zero. This trend is related to the collapse patterns and can be observed for the collapsed buildings shown in Fig. 4. Figure 4d shows a completely collapsed building in which the debris has spread laterally. For those cases, the values of r are low and the values of σ are large. On the other hand, Fig. 4e shows a collapsed building with a roof that remained almost the same shape while it collapsed almost vertically. This means that all the elevations inside the footprint decreased by about the same amount, thus leading to a high value of r and a low value of σ. This pattern is often difficult to detect from optical aerial and satellite optical images, because the sensor measures the landscape vertically. The histograms for collapsed buildings (Fig. 7) shows that several collapsed buildings have a value for r greater than 0.5, and it would be difficult to detect this from aerial or optical satellite imagery. Readers might notice that noncollapsed buildings also have a value of r close to 1 and σ close to zero; however, those can be first filtered using ΔH. Then, the pattern of collapse can be evaluated from the other two parameters (using a decision tree).

Within the study area, 26 128 building footprints were found. It is worth mentioning that few buildings were not well registered in the GIS map. Figure 8 shows the parameters calculated for each building, wherein the shaded color depicts the density of the dots. Most of the points are lo-

cated at approximately (ΔH, σ, r) = (0 m, 0.5 m, 0.9), which indeed represents noncollapsed buildings. Several buildings show positive values of ΔH. A closer look revealed two principle factors: (1) the collapse of a neighboring building and (2) plastic covers placed over the roof for protection from the rain.

The next concern was to define a criterion to set threshold values that can differentiate collapsed/noncollapsed buildings properly. A number of options were evaluated in this study. Since it is obvious that the buildings with clear negative values of ΔH correspond to collapsed buildings, we first analyzed the classification using a threshold for ΔH only. The buildings with ΔH values smaller than that threshold were classified as collapsed; the buildings with ΔH values greater than the threshold were classified as noncollapsed. The possible thresholds were tested on the buildings surveyed by Yamada et al. (2017). Figure 9 shows the Cohen's kappa coefficient and the overall accuracy calculated from the comparison between the estimated collapsed and noncollapsed buildings (i.e., using a given threshold) and the building damage classes based on the ground truth. For the comparison, the buildings with damage levels D0, D1–D3, and D4 were labeled noncollapsed buildings. The Cohen's kappa (k) coefficient and the overall accuracy (OA) are expressed as follows:

$$p_{no} = (p_{21} + p_{22})(p_{12} + p_{22}) \tag{4}$$
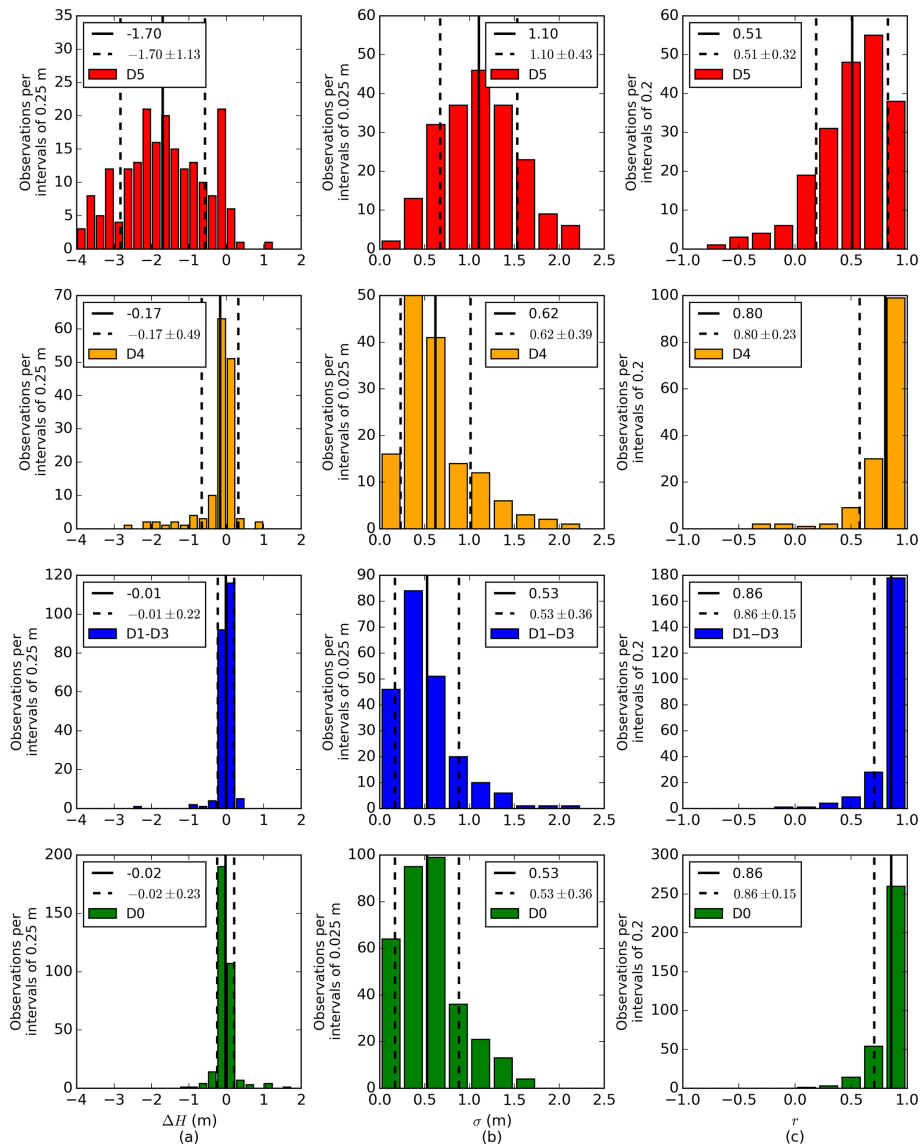
$$p_{yes} = (p_{11} + p_{12})(p_{11} + p_{21}) \tag{5}$$

$$p_o = p_{11} + p_{22} \tag{6}$$

$$p_e = p_{yes} + p_{no} \tag{7}$$

$$OA = p_o \tag{8}$$

$$k = \frac{p_o - p_e}{1 - p_e}, \tag{9}$$

where p_{11} and p_{12} are the ratio of noncollapsed buildings predicted as noncollapsed and collapsed buildings. p_{21} and p_{22} are the ratio of collapsed buildings predicted as noncol-

Figure 7. Histograms of the three parameters (ΔH, σ, r) calculated from the lidar DSMs for the buildings surveyed by Yamada et al. (2017), separated into four damage levels.

lapsed and collapsed buildings. From Fig. 9, it is observed that a threshold value of $-0.5\,\mathrm{m}$ gave the highest values for both the Cohen's kappa coefficient (0.80) and the overall accuracy (0.93).

To determine whether the use of all the parameters could produce better accuracy in detecting collapsed buildings, the support vector machine (SVM) method was selected to construct a plane that separates collapsed and noncollapsed buildings in the three-dimensional database (ΔH, σ, r). The plane has the largest distance from the nearest training data (ground truth data). Using kernel functions, SVM can be used to construct a nonlinear function as well. However, in this study we only evaluated linear functions (i.e., a plane or linear kernel function). Figure 10 shows the plane

constructed using SVM, wherein the red and blue marks depict the collapsed and noncollapsed buildings based on the ground truth. The plane was constructed using the same amount of data for the two classes. Thus, 205 noncollapsed buildings were selected randomly from the surveyed data. The analysis was performed several times, and although the plane obtained showed small variations due to the random selection of the training data, the Cohen's kappa coefficient produced in each analysis was almost constant, with minor fluctuations around 0.80. The accuracy produced by SVM is very similar to the accuracy obtained when only a ΔH threshold is used. For a linear kernel SVM, the vector $\boldsymbol{w}$ perpendicular to the decision plane is defined by the following
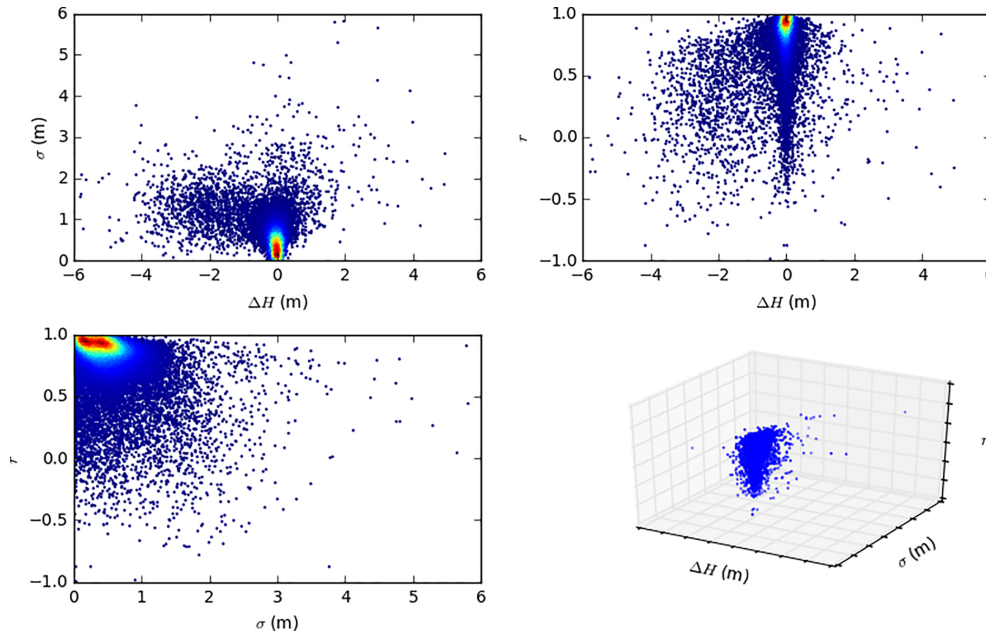
Figure 8. Scatter plots of the three parameters (ΔH, σ, r) calculated for all the buildings in the study area. The color represents the density of points, wherein red shows the area with the highest density.
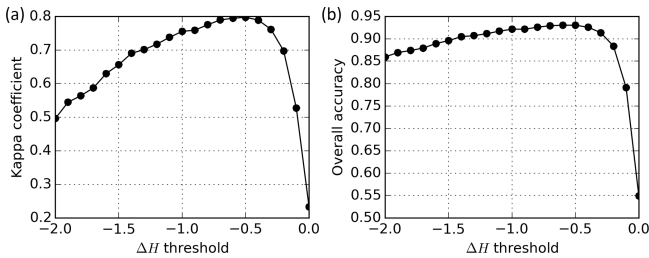
Figure 9. Kappa coefficient **(a)** and overall accuracy **(b)** obtained from the comparison between the data surveyed by Yamada et al. (2017) and the estimated collapsed buildings based on different ΔH threshold values.

expression:

$$\boldsymbol{w} = \sum_i \alpha_i y_i \boldsymbol{x}_i, \tag{10}$$

where $\boldsymbol{x}_i$ is a training vector that contains the three parameters (ΔH, σ and r), y_i represents the class that can be either 1 or -1, and the coefficients α_i are obtained by solving the following problem:

$$\min_\alpha \left(\frac{1}{2} \alpha^T \mathbf{Q} \alpha - \boldsymbol{e}^T \alpha \right) \tag{11}$$

$$Q_{ij} = y_i y_j \left(\boldsymbol{x}_i \cdot \boldsymbol{x}_j \right) \tag{12}$$

$$0 \le \alpha_i \le C, \; i = 1, \ldots, n, \tag{13}$$

where $\boldsymbol{e}$ is a vector with elements that are all ones. C is the upper bound and is used as a regularization parameter.

The parameter C trades off misclassification of training examples against simplicity of the decision surface. A low C value makes the decision surface smooth and a high C value aims to classify all training examples correctly (Skitlearn, 2017a). In this study a value C equals to 1 was used. In order to evaluate its effects, a cross-validation procedure was performed. For each C value, 80 % of the surveyed data were selected randomly and were used to calibrate the SVM classifier. The rest of the surveyed data were used to calculate a score that represents the accuracy. The overall accuracy was chose as the score. The procedure was repeated five times and the average was stored. Figure 11 shows the cross-validation accuracy. It is observed the accuracy remains mainly constant with small fluctuations at lower values. However, a difference of approximately 3 % is observed between the worst and the best accuracy. Therefore, it is concluded that the C value did not affect the SVM classifier in our study.

This study also evaluated the potential use of unsupervised classification for collapsed buildings. Specifically, k-means cluster analysis was applied to all the data in the study area. Unlike SVM, k-means clustering does not require training data. Therefore, the database of all the buildings in the study area (Fig. 8) was used. The method clusters the data and separates them into two groups, which represent the collapsed and noncollapsed buildings. The objective of the method is to minimize the inertia of each group, that is, the summation of the squared distance between all the data points of a group and its centroid. The result is highly dependent on the initialization of the centroids. Here, a k-means++ initialization scheme was used. K means++ initializes the centroids to a
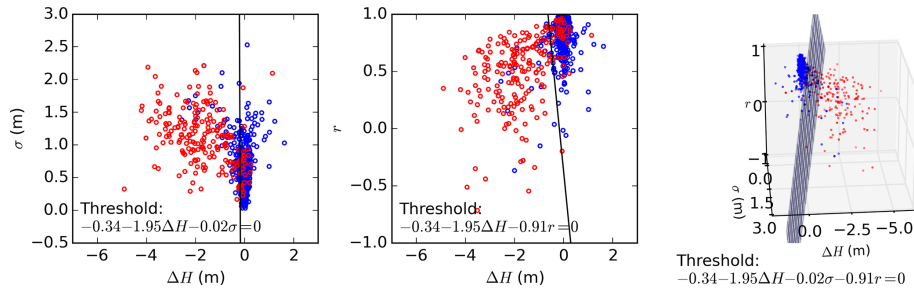
Figure 10. Classification of collapsed (red) and noncollapsed (blue) buildings using the three parameters based on SVM.
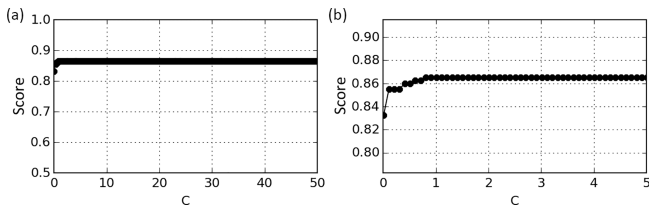
Figure 11. The classifier's cross-validation accuracy as a function of C. **(a)** Overall evaluated range of C. **(b)** A closer look at values lower than 5.

distance from each other (Scikit-learn, 2017b). Figure 12 represents the predicted collapsed and noncollapsed buildings using the k-means clustering method, for which the Cohen's kappa coefficient obtained was 0.76. Figure 13 shows the confusion matrix calculated from the comparison between the ground truth data and the predicted results from the three methods explained above, applying a ΔH threshold, SVM, and k-means clustering. The first two methods show the same level of accuracy, while k-means clustering shows a lower accuracy.

Figure 14 illustrates the spatial distribution of collapsed buildings estimated using a ΔH threshold of -0.5 m. A large number of collapsed buildings were observed in the study area (Fig. 14a). The red and black polygons represent the collapsed (D5) and noncollapsed (D0–D4) buildings. The color of the pixels represents the difference in elevations between the ADSM and BDSM. Blue pixels depict differences of elevations less than -0.5 m, and yellow pixels represent differences greater than 0 m. Figures 14b and 15 provide a closer look of the areas where the collapsed buildings are concentrated. Figure 14b also depicts the location of the collapsed buildings surveyed by Yamada et al. (2017) as black triangles. Within the study area, a total of 26 128 buildings were evaluated, and 1760 buildings were classified as collapsed (ΔH less than -0.5 m).

It was observed that some buildings that collapsed during the foreshock (14 April event) were also detected by the lidar methods. In order to be detected, the debris of those buildings should be either severely disturbed by the mainshock (16 April event) or removed before the ADSM was

recorded. For instance, Fig. 16 shows two buildings that collapsed during the foreshock (Fig. 16a). However, because the mainshock produced a significant reduction in their elevations (Fig. 16b), it was also detected from the pair of lidar.

4 Discussion

This paper evaluated the use of lidar data to detect damaged buildings by means of three parameters: ΔH, σ, and r. It was found that collapsed buildings can be identified precisely from the average difference in height, ΔH. However, the other two parameters can provide additional information about the collapse pattern. The collapsed patterns are correlated with the failure mechanism of buildings, which might highlight some deficiencies in the design codes that were used in the construction process. A detailed understanding of the failure mechanism is important to the practice of forensic engineering, the investigation of failures and other performance problems. Moreover, with further evaluation, the collapsed pattern might contribute to future improvements of the building design codes. Unfortunately, it was not possible to calibrate a threshold that can properly classify different collapsed patterns. The main reason is because there was no information related to the collapse patterns in the survey data. Perhaps this task can be done in future research after new survey data are released.

Some words regarding sources of error that were present in this study should be mentioned. The footprint data provided by the Geospatial Information Authority of Japan (GSI) is rather precise but not perfect. Three drawbacks were observed: (1) a few buildings were not included in the database, (2) a slight shift between the building footprint and corresponding lidar data is sometimes observed, and (3) in some cases, a group of buildings, consisting mostly of two or three buildings, were registered within one building footprint. These uncertainties may have produced errors in the detection of collapsed buildings. However, they did not have a significant impact on the overall results, which is confirmed in Fig. 9, wherein the Cohen's kappa coefficient and the overall accuracy are significantly high. This problem can be solved by performing a manual inspection or automatic detection of buildings from the BDSM in order to update
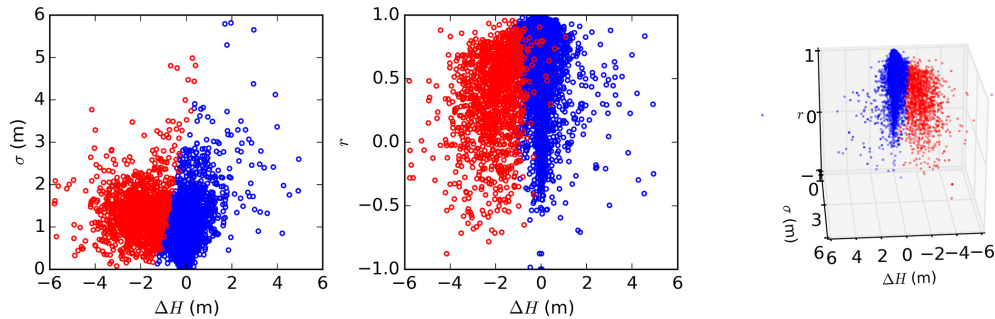
Figure 12. Classification of collapsed (red) and noncollapsed (blue) buildings using the three parameters based on the k-means clustering method.
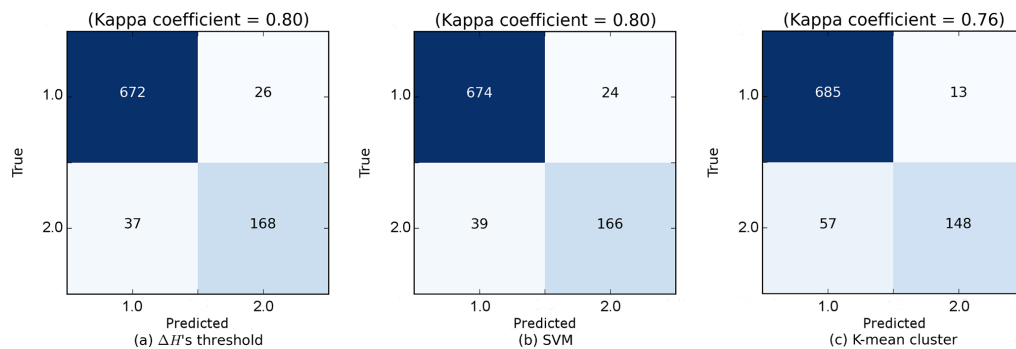
Figure 13. Confusion matrix calculated from the comparison of the ground truth data and the predicted damage levels based on the ΔH threshold **(a)**, SVM **(b)**, and k-means clustering **(c)**. Two damage levels, noncollapsed (1.0) and collapsed (2.0), were employed.

the data set. However, the authors decided to work with the data in its current state because this uncertainty is likely to be present in other real situations in which a quick report on damage extent is required.

Of the three methods evaluated here, the k-means clustering exhibits the lowest accuracy. The main reason is that, unlike the SVM method, the k-means clustering does not use any truth data. However, it produced a kappa coefficient of 0.76 and overall accuracy of 92 %, which is still quite good. The k-means clustering method is useful for taking a first glance at the distribution of collapsed buildings because the method does not require any training data. The procedure is well-known and robust, with several efficient algorithms with proven fast convergence.

Finally we would like to discuss the building damage rate in the Kumamoto earthquake in comparison to other recent M7-level crustal earthquakes in Japan. In the 1995 Kobe earthquake (M_w 6.9), about 49 000 buildings were collapsed (G5 in EMS-98 scale) or severely damaged (G4) out of 560 000 buildings in the affected urban area (Building Research Institute, 1996). The recorded strong motion distribution in the Kobe earthquake was at a similar level to that of the Kumamoto earthquake (Yamaguchi and Yamazaki, 2001). However, in 1995, the number of strong motion accelerometers was much lower, about only 10 in the hard-hit zone. The building density of the affected area was much

higher in the Kobe region than that of the Kumamoto region. Considering these differences, it is difficult to conclude which earthquake was more destructive. From our experiences (Yamazaki and Murao, 2000; Yamaguchi and Yamazaki, 2000), the severity of damage for timber-frame houses in Mashiki was at a similar level to that in the hard-hit zone of the Kobe region. There were a few more recent M7-level crustal earthquakes in Japan, such as the 2004 Niigata-Chuetsu earthquake (M_w 6.6) and the 2007 Niigata-Chuetsu-Oki earthquake (M_w 6.6). However, the population density of the affected urban areas was lower in these events, and thus it is again difficult to compare their damage situations (Nagao et al., 2011) with that of Kumamoto, although their strong motion levels were again comparable with that of the Kumamoto event.

Another point of discussion is whether or not the M_w 6.2 foreshock influenced the overall damage situation of Mashiki. Our preliminary conclusion is "partially yes, but not so much". Based on numerical analyses of the behavior of collapsed typical timber-frame buildings (Building Research Institute, 2015), those built by the old seismic code (before 1982) had mostly collapsed only from the mainshock's excitation, even without being preshaken by the foreshock (Suto et al., 2017). However, in some cases, building models had collapsed only from the sequence of the foreshock and main-
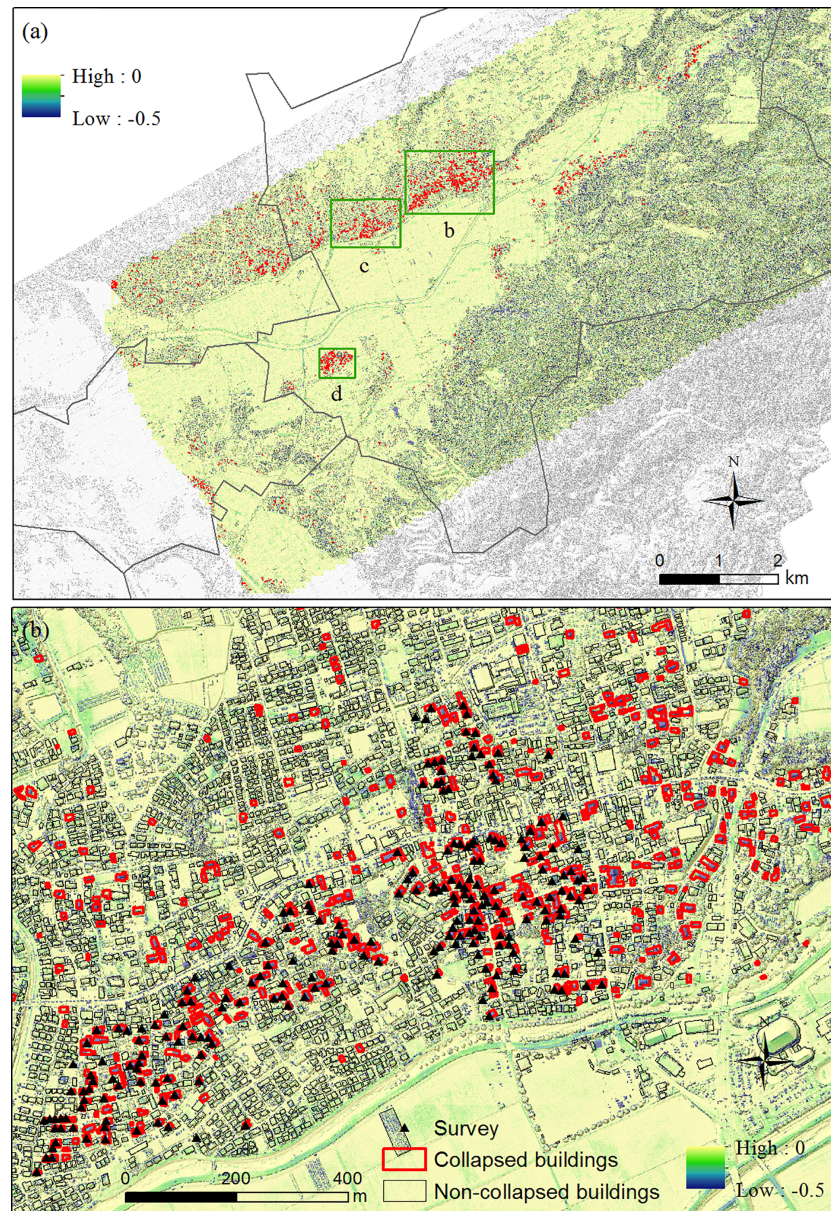
Figure 14. (a) A map showing the distribution of collapsed ($\Delta H < -0.5$ m) buildings, shown as red polygons, in the study area. The pixel color represents the difference in elevation between the BDSM and ADSM. The green squares show the locations of areas shown in panel **(b)** and Fig. 15. Close-up view of area **(b)** in which the collapsed buildings were concentrated. The black triangles show the D5 buildings from Yamada et al. (2017).

shock excitations. More detailed results on this matter will be presented in the near future.

5   Conclusions

In this study, the spatial distribution of collapsed buildings was extracted from a pair of lidar data sets taken before and after the 2016 M_w 7.0 Kumamoto earthquake. For this purpose, geographic information on building footprints was employed. Three parameters were used: the average (ΔH) and

standard deviation (σ) of the height differences between the two DSMs and the correlation coefficient between them (r). The parameters were evaluated using the building damage survey data set obtained by Yamada et al. (2017); ΔH was found to be very efficient for identifying collapsed buildings. However, the other parameters provided insights into the collapse pattern. After evaluating different methodologies with which to identify collapsed buildings, buildings with ΔH less than -0.5 m were considered to have collapsed. The distribution of collapsed buildings obtained by Yamada et al. (2017)

Figure 15. Close-up view of areas **(c)** and **(d)** in Fig. 14a where collapsed buildings are concentrated. The red and green polygons are the collapsed and noncollapsed buildings estimated using the threshold ($\Delta H < -0.5$ m).

Figure 16. (a) Aerial image taken on 15 April; **(b)** aerial image taken on April 23. The thick red polygons show buildings that collapsed after the foreshock and were detected using the ΔH threshold.

was illustrated together with the height difference between the two DSMs, and good agreement was observed. From a total of 26 128 evaluated buildings, 1760 collapsed buildings were identified. To our knowledge, this result may be the first case in which a large number of collapsed buildings were identified from a pre- and post-event lidar DSM pair.

It is expected that the use of lidar data to identify damaged areas will eventually increase in the near future. However, because of the current lack of data, the implementation of a method to identify collapsed building using only post-event lidar data is important and will be considered in a future study. Additional future studies related to the use of these lidar data are the quantification of debris related to road blockage, and the identification of landslides.

Competing interests. The authors declare that they have no conflict of interest.

Edited by: Oded Katz
Reviewed by: two anonymous referees

References

Aixia, D., Zongjin, M., Shusong, H., and Xiaoqing, W.: Building damage extraction from post-earthquake airborne LiDAR data, Acta Geol. Sin.-Engl., 90, 1481–1489. 2016.

Asia Air Survey Co., Ltd.: The 2016 Kumamoto earthquake, available at: http://www.ajiko.co.jp/article/detail/ID5725UVGCD/, last access: 1 April 2017.

Building Research Institute: Final report of damage survey of the 1995 Hyogoken-Nanbu earthquake, available at: http://www.kenken.go.jp/japanese/research/iisee/list/topics/hyogo/pdf/h7-hyougo-jp-all.pdf (last access: 1 September 2017), 1996 (in Japanese).

Building Research Institute: Wallstat version 3.1, collapsing simulation program for timber structures, available at: http://www.nilim.go.jp/lab/idg/nakagawa/wallstat.html (last access: 1 September 2017), 10 September 2015.

Cabinet Office of Japan: Summary of damage situation in the Kumamoto earthquake sequence, available at: http://www.bousai.go.jp/updates/h280414jishin/index.html, last access: 1 September 2017 (in Japanese).

Das, R. and Hanaoka, S.: An agent-based model for resource allocation during relief distribution, Journal of Humanitarian Logistics and Supply Chain Management, 4, 265–285, 2014.

Dell'Acqua, F. and Gamba, P.: Remote sensing and earthquake damage assessment: Experiences, limits, and perspectives, Proceedings of the IEEE, 100, 2876–2890, 2012.

Geospatial Information Authority of Japan: New measurement for survey reference points after the 2016 Kumamoto Earthquake, available at: http://www.gsi.go.jp/sokuchikijun/sokuchikijun60019.html, last access: 1 April 2017.

Hashemi-Parast, S. O., Yamazaki, F., and Liu, W.: Monitoring and evaluation of the urban reconstruction process in Bam, Iran, after the 2003 M_W 6.6 earthquake, Nat. Hazards, 85, 197–213, 2017.

Hoshi, T., Murao, O., Yoshino, K., Yamazaki, F., and Estrada, M.: Post-disaster urban recovery monitoring in Pisco after the 2007 Peru earthquake using satellite image, Journal of Disaster Research, 9, 1059–1068, 2014.

Hussain, E., Ural, S., Kim, K., Fu, C., and Shan, J.: Building extraction and rubble mapping for city Port-au-Prince post-2010 earthquake with GeoEye-1 imagery and Lidar Data, Photogramm. Eng. Rem. S., 77, 1011–1023, 2011.

Japan Meteorological Agency: The number of aftershocks of recent inland earthquakes in Japan, http://www.data.jma.go.jp/svd/eqev/data/2016_04_14_kumamoto/kaidan.pdf, last access: January 2017 (in Japanese).

Korosov, A. A., Hansen, M. W., Dagestad, K., Yamanaka, A., Vines, A., and Riechert, M.: Nansat: a scientist-orientated python package for geospatial data processing, Journal of Open Research Software, 4, 11 pp., 2016.

Labiak, R. C., Aardt, J. A. N., Bespalov, D., Eychner, D., Wirch, E., and Bischof, P.: Automated method for detection and quantification of building damage and debris using post-disaster Lidar data, Proc. SPIE 8037, Laser Radar Technology and Applications XVI, Vol. 8037, 8 pp., 2011.

Liu, W., Yamazaki, F., Gokon, H., and Koshimura, S.: Extraction of tsunami-flooded areas and damaged buildings in the 2011 Tohoku-oki earthquake from TerraSAR-X intensity images, Earthq. Spectra, 29, S183–S200, 2013.

Maruyama, Y., Tashiro, A., and Yamazaki, F.: Detection of collapsed buildings due to earthquakes using a digital surface model constructed from aerial images, J. Earthq. Tsunami, 8, 1450003 (13 pp.), 2014.

Meslem, A., Yamazaki, F., and Maruyama, Y.: Accurate evaluation of building damage in the 2003 Boumerdes, Algeria earthquake from Quickbird satellite images, J. Earthq. Tsunami, 5, 1–18, 2011.

Moya, L., Yamazaki, F., Liu, W., Chiba, T., and Mas, E.: Detection of collapsed buildings due to the 2016 Kumamoto earthquake from Lidar data, World Engineering Conference on Disaster Risk Reduction, Lima, Peru, 5–6 December, 8 pp., 2016.

Moya, L., Yamazaki, F., Liu, W., and Chiba, T.: Calculation of coseismic displacement from lidar data in the 2016 Kumamoto, Japan, earthquake, Nat. Hazards Earth Syst. Sci., 17, 143–156, https://doi.org/10.5194/nhess-17-143-2017, 2017.

Nagao, T., Yamazaki, F., and Inoguchi, M.: Analysis of building damage in Kashiwazaki city due to the 2007 Niigata-ken Chuetsu-oki earthquake, Proc. 32nd Asian Conference on Remote Sensing, Taipei, Paper No. 228, 6 pp., 2011.

Okada, S. and Takai, N.: Classifications of structural types and damage patterns of buildings for earthquake field investigation, Proceedings of the 12th World Conference on Earthquake Engineering, paper 0705, Auckland, New Zealand, 2000.

Rathje, E. and Adams, B. J.: The role of remote sensing in earthquake science and engineering, opportunities and challenges, Earthq. Spectra., 24, 471–492, 2008.

Rehor, M., Bahr, H., Tarsha-Kurdi, F., Landes, T., and Grussenmeyer, P.: contribution of two plane detection algorithms to recognition of intact and damaged buildings in lidar data, The Photogrammetric Record, 23, 441–456, 2008.

Schweier, C. and Markus, M.: classification of collapsed buildings for fast damage and loss assessment, B. Earthq. Eng., 4, 177–192, 2006.

Scikit-learn: Support Vector Machines, available at: http://scikit-learn.org/stable/modules/svm.html#svm-classification, last access: 6 August 2017a.

Scikit-learn: Clustering, available at: http://scikit-learn.org/stable/modules/clustering.html#k-means, last access: 6 August 2017b.

Suto, T., Yamazaki, F., and Liu, W.: Numerical simulation of wooden buildings in the Kumamoto earthquake based of Wall-stat, Annual Meeting of JSCE, Fukuoka, Japan, 2 pp., 2017 (in Japanese).

Uprety, P., Yamazaki, F., and Dell'Acqua, F.: Damage detection using high-resolution SAR imagery in the 2009 L'Aquila, Italy, earthquake, Earthq. Spectra, 29, 1521–1535, 2013.

Vuolo, F., Żółtak, M., Pipitone, C., Zappa, L., Wenng, H., Immitzer, M., Weiss, M., Baret, F., and Atzberger, C.: Data service platform for sentinel-2 surface reflectance and value-added products: system use and examples, Remote Sens., 8, 16 pp., 2016.

Whitman, R. V., Reed, J. W., and Hong, S.: Earthquake damage probability matrices, Proceedings of the Fifth World Conference on Earthquake Engineering, Rome, 2531–2540, 1973.

Yamada, M., Ohmura, J. and Goto, H.: Wooden Building Damage Analysis in Mashiki Town for the 2016 Kumamoto Earthquakes on April 14 and 16, Earthq. Spectra, 33, 1555–1572, 2017.

Yamaguchi, N. and Yamazaki, F.: Fragility curves for buildings in Japan based on damage surveys after the 1995 Kobe earthquake, Proceedings of the 12th World Conference on Earthquake Engineering, paper 2451, Auckland, New Zealand, 2000.

Yamaguchi, N. and Yamazaki, F.: Estimation of strong motion distribution in the 1995 Kobe earthquake based on building damage data, Earthq. Eng. Struct. D, 30, 787–801, 2001.

Yamazaki, F. and Liu, W.: Remote sensing technologies for post-earthquake damage assessment: A case study on the 2016 Kumamoto earthquake, Keynote Lecture, 6th ASIA Conference on Earthquake Engineering, Cebu City, Philippines, 8 pp., 2016.

Yamazaki, F. and Matsuoka, M.: Remote sensing technologies in post-disaster damage assessment, J. Earthq. Tsunami, 1, 193–210, 2007.

Yamazaki, F. and Murao. O.: Vulnerability functions for Japanese buildings based on damage data from the 1995 Kobe earthquake. Implications of recent earthquakes on seismic risk, 91–101, Imperial College Press, London, 2000.