Characteristics of Tsunami Fragility Functions Developed Using Different Sources of Damage Data from the 2018 Sulawesi Earthquake and Tsunami

ERICK MAS,1 RYAN PAULIK,2 KWANCHAI PAKOKSUNG,1 BRUNO ADRIANO,3 LUIS MOYA,1,4 ANAWAT SUPPASRI,1 ABDUL MUHARI,5 ROKHIS KHOMARUDIN,6 NAOTO YOKOYA,3 MASASHI MATSUOKA,7 and SHUNICHI KOSHIMURA1

Abstract—We developed tsunami fragility functions using three sources of damage data from the 2018 Sulawesi tsunami at Palu Bay in Indonesia obtained from (i) field survey data (FS), (ii) a visual interpretation of optical satellite images (VI), and (iii) a machine learning and remote sensing approach utilized on multisensor and multitemporal satellite images (MLRS). Tsunami fragility functions are cumulative distribution functions that express the probability of a structure reaching or exceeding a particular damage state in response to a specific tsunami intensity measure, in this case obtained from the interpolation of multiple surveyed points of tsunami flow depth. We observed that the FS approach led to a more consistent function than that of the VI and MLRS methods. In particular, an initial damage probability observed at zero inundation depth in the latter two methods revealed the effects of misclassifications on tsunami fragility functions derived from VI data; however, it also highlighted the remarkable advantages of MLRS methods. The reasons and insights used to overcome such limitations are discussed together with the pros and cons of each method. The results show that the tsunami damage observed in the 2018 Sulawesi event in Indonesia, expressed in the fragility function developed herein, is similar in shape to the function developed after the 1993 Hokkaido Nansei-oki tsunami, albeit with a slightly lower damage probability between zero-to-five-meter inundation depths. On the other hand, in comparison with the fragility function developed after the 2004 Indian Ocean tsunami in Banda Aceh, the characteristics of Palu structures exhibit higher fragility in response to tsunamis. The two-meter inundation depth exhibited nearly 20% probability of damage in the case of Banda Aceh, while the probability of damage was close to 70% at the same depth in Palu.

Keywords: Fragility function, tsunami, 2018 Sulawesi, earthquake.

1. Introduction

Tsunami fragility functions (TFFs) are cumulative distribution functions that express the probability of a structure reaching or exceeding a particular damage state in response to a specific value of tsunami intensity measure or another engineering demand parameter. TFFs can be applied to estimate building damage in future scenarios together with economic losses (Wiebe and Cox 2013; Adriano et al. 2014; Rehman and Cho 2016); however, discussion of the accuracy of estimation when applying TFFs remains (Moya et al. 2017). Nevertheless, TFFs serve as a proxy for improvements of building codes where tsunami loading must be considered (Condori Uribe 2013; Chock et al. 2016), or when analyzing synthetic future damage scenarios (Moya et al. 2018c). The tsunami intensity measure can be represented by the tsunami inundation depth; however, in most cases it is not enough to have a unique explanatory variable for a complex phenomenon such as tsunami damage.
Tsunami flow velocity and hydrodynamic force have also been used independently (Song et al. 2017; Aránguiz et al. 2018) and combined (Charvet et al. 2017) to develop TFFs and fragility surfaces, respectively. To develop a TFF, it is necessary to compile a set of damage classification data samples and to correlate these with the tsunami intensity measures for sample locations following a particular statistical model. A review of early developed TFFs can be found in Tarbotton et al. (2015), Nanayakkara and Dias (2016). TFFs are generally empirical in nature, derived from the damage and inundation data collected from a field survey after a major disaster. Due to uncertainties in the real origin and cause of damage to buildings observed in their ultimate state during postdisaster surveys, TFF may incorporate damage to buildings produced by strong motion or tsunami alone, or the combination of both. Nevertheless, damage and destruction within an inundation area can be clearly related to a total or partial contribution of the tsunami hazard. To clearly separate the effects of earthquake and tsunami forces in a structure, analytical fragility functions can be used; however, this falls outside the scope of this paper. Efforts to develop analytical fragility functions can also be found in the literature (Alam et al. 2018; Medina et al. 2019).

In this study, we will concentrate only on empirical TFFs. In the effort to develop TFFs, several field teams may measure inundation depths and record later damage levels at affected asset locations. However, this activity is time-consuming and may place research teams at risk of harm in tsunami affected areas. Moreover, it may disrupt relief activities or increase social anxiety for external help if the survey tasks are not well-managed. Thus, other methods to obtain information for TFF development which avoid these issues are necessary. To that end, remote sensing technology, tsunami numerical simulation and geospatial analysis aid researchers to obtain tsunami damage information in a noninvasive manner (Koshimura et al. 2009a; Mas et al. 2012; Moya et al. 2018b; Adriano et al. 2019).

In this study, we developed TFFs using three types of building damage data obtained from different methods. First, we used conventional field survey data; then, we used a visual interpretation of optical satellite images, and finally, we used a machine learning approach utilized on multisensor and multi-temporal satellite images. Here, TFFs are developed using the damage data of the 2018 Sulawesi tsunami at Palu Bay in Indonesia. Sulawesi Island has experienced more than 15 tsunamis since 1820 (Pelinovsky et al. 1997), most of them from a seismic source. On September 28, 2018 at 18:02 local time, a shallow strike-slip earthquake occurred near Palu City (Heidarzadeh et al. 2019; Gusman et al. 2019; Socquet et al. 2019; Ulrich et al. 2019). The 7.5 Mw magnitude earthquake generated tsunami waves which arrived onshore within minutes with a short period, according to tidal gauge records (Muhari et al. 2018). Postevent analysis has found that in addition to the seismic generating source, tsunami waves were also caused by a series of submarine landslides that contributed to the generation of waves (Arikawa et al. 2018; Kijewski-Correa and Robertson 2018; Carvajal et al. 2019; Takagi et al. 2019; Nakata et al. 2020). To develop TFFs in this study, we will concentrate our analysis on tsunami damaged buildings in central Palu City, as shown in Fig. 1.

2. Posttsunami Surveys

Postevent field reconnaissance reports (Kijewski-Correa and Robertson 2018; Robertson 2018) have suggested that most of the structures in Palu City consist of light timber frames, followed by low-rise reinforced concrete structures with unreinforced masonry infill walls. Accordingly, in the same reports, the authors suggested that the earthquake shaking did not damage or only slightly damaged these structures. However, tsunami hydrodynamic and debris impact forces may have been the principal causes of failure and collapse in the waterfront area of Palu Bay. In contrast, other areas with liquefaction showed that lateral spreading was the main cause for collapse. Table 1 summarizes the available reports conducted and field survey data collected after the disaster and compiled in this study.

Muhari et al. (2018) first reported the observations from their field survey. Based on the tide gauge from Pantoloan port, they concluded that this event presented particularly short period tsunami waves
arriving very rapidly to the coast of Palu Bay with maximum 8-m height and 50-m inundation distance for the area surveyed. In addition, tsunami depth decreased rapidly for dense urban areas. Flow depths up to 1 m were found inside houses. For these reasons, the report also suggested that the event was possibly generated by additional sources other than the seismic seafloor deformation.

On the other hand, Putra et al. (2019) surveyed tsunami deposits along the coastline of Palu Bay. In this field survey, tsunami heights below 8 m were found at six locations where tsunami deposits were investigated. However, maximum inundation distance was reported as 310 m.

The next field survey, reported by Widiyanto et al. (2019), measured tsunami inundation and run-up across 18 sites in Palu Bay. A maximum run-up height of approximately 10.5 m was found. Similarly, a maximum inundation distance of 510 m was reported here. In addition, rapid arrival times for tsunami waves were reported by survivors approximately 3 to 8 min after the earthquake (Fig. 2). Widiyanto et al. (2019) also suggest some indicators of underwater landslides. During the same week of this survey, another group from Japan and Indonesia (Arikawa et al. 2018) was collecting information in the forms of geomorphological changes, tsunami traces and eyewitness testimonies. Landslide traces were found in six areas on both sides of the bay. Later, Kijewski-Correa and Robertson (2018) suggested a total of 13 locations for landslides based on survivor video and visual comparison of before and after satellite imagery of the area. These locations were used by Pakoksung et al. (2019) for tsunami simulations and verified with measurements provided by the Geological Agency of Indonesia (Fig. 2).
tsunami height measurements, Syamsidik et al. (2019) reported measurements from 3 to 5 m in the city of Palu (Fig. 2), while Arikawa et al. (2018)’s survey reported heights from 2 to 6 m in Palu Bay, with higher values observed towards the central and southern areas of the bay.

Similarly, Koshimura et al. (2019) conducted a field survey to identify tsunami impacts on buildings. Here, RTK-GPS measurements were obtained to delimit the tsunami inundation and inland penetration. In addition, flow depths and structural damage information were gathered. A maximum of 6 m of tsunami height was observed during this survey. Again, nonseismic sources were suggested in addition to the earthquake to explain short period waves and rapid arrival to the coast (Fig. 2).

Kijewski-Correa and Robertson (2018) presented a follow-up report to the prompt virtual assessment report published by Robertson (2018). The group visited the affected areas at the end of October 2018 and published their observations (Fig. 2) together with Mikami et al. (2019). In Kijewski-Correa and Robertson (2018), the interested reader may find general descriptions and observations of the event and its impact. In addition, the damage from the earthquake and liquefaction from a structural engineering point of view are presented. In contrast, Mikami et al. (2019) conducted an aerial photographic survey with UAV and a field survey of tsunami inundation and run-up heights. They strongly suggested that landslide sources were present; some of those were discussed in greater detail by Sassa and Takagawa (2019), Takagi et al. (2019) and Nakata et al. (2020).

According to Omira et al. (2019), ten sectors of large coastal collapse were identified during their field survey. In addition, tsunami run-up and inundation heights of 9.1 m and 8.7 m, respectively, were measured in Palu Bay.

Finally, another group that visited the affected area and conducted a field survey reported their observations in Paulik et al. (2019). A large number of data points of flow depths, buildings, roads and electricity infrastructure damage were shared with the scientific community (Fig. 2). The combination of these data and previous survey data represents the main input used in this study to construct the TFF of the Palu Bay area.

3. Data and Preprocessing

In this study, we have used the available field data surveyed by different groups within a common area
concentrated within the central Palu Bay area (see Figs. 1, 2).

The data required to develop TFFs are paired measurements corresponding to (i) a level of tsunami damage for a building and (ii) the tsunami inundation depth measured at the same location. Several methods can be used to obtain this information and have been described previously in the literature (Mas et al. 2012; Aránguiz et al. 2018). To summarize them, the following methods can be considered for each necessary parameter:

[1] Tsunami Inundation Depth
- Posttsunami field survey
- Tsunami inundation simulation
- Geospatial interpolation (surface from surveyed points)

[2] Tsunami Building Damage Classification
- Posttsunami field survey
- Visual interpretation using satellite remote sensing
- Machine learning algorithms using satellite remote sensing

3.1. Tsunami Inundation Depth

First, we compiled the field survey data published by five different groups to obtain a comprehensive tsunami inundation depth database for central Palu City area. To estimate the tsunami inundation depth, we used spatial interpolation methods based on all available field survey points. Next, we calculated the relative tsunami height of all surveyed depth points by adding the topographic height acquired from a digital elevation model (DEM) provided by the Geospatial Information Agency of Indonesia (Badan Informasi Geospasial, Indonesia - BIG). The DEM is a raster dataset of 15-m resolution with artificial zero elevations added towards a shoreline digitized from...
pre-event DigitalGlobe WorldView-2 satellite image data of February 20, 2018.

We then created virtual points at the shoreline and run-up boundaries based on the survey data.

\[ Y_s = (y + aS_0)e^{\frac{x}{aS_0}} - aS_0 \]  
(1)

We applied Eq. 1, as suggested by Smart et al. (2016), to the estimation of water depth profiles to calculate the tsunami height at shoreline points \((Y_s)\) with respect to the nearest survey point \((y)\) obtained based on Voronoi cells from the survey data. We assumed a value of 90 for the roughness aperture parameter \((a)\), which is a value between the recommended values (Smart et al. 2016) corresponding to “undulating open ground” and “light buildings, coconut plantations,” which we judge to better describe the inland conditions of Palu Bay. Similarly, based on the analysis of slopes using the DEM data of the area of study, an average value of 0.1 was decided for \(S_0\).

Finally, the tsunami height values are interpolated within the inundation area. The inundation area was obtained based on the area reported by Pakoksung et al. (2019) and adjusted following the run-up data compiled from field surveys used in this study, along with a shoreline from pre-event satellite imagery.

The final correlation coefficient of points in the inundation surface and from field surveys was 0.98. Next, the DEM was subtracted from the resulting surface to obtain a tsunami inundation depth surface as shown in Fig. 3. We used the values from this surface at the respective locations of buildings classified by the two estimation methods that used remote sensing (i.e., VI and MLRS).

### 3.2. Tsunami Building Damage Classification

Building damage classification based on posttsunami field survey (Paulik et al. 2019), visual interpretation using satellite remote sensing (Copernicus EMS 2018) and a machine learning algorithm applied on satellite imagery (Adriano et al. 2019) are used to construct the TFF of the event. The original damage classification and reclassification for the three methods are summarized in Fig. 4 and Table 2.

In Fig. 4, the vertical line that divides the original classification bars defines the reclassification criteria. To the left side are ’damaged’ structures and to the right side are ’other’. We carefully balanced the rates of classes across methods.

### 4. Developing Tsunami Fragility Functions

Tsunami fragility functions provide a probabilistic relationship between either tsunami hydrodynamic features or tsunami intensity measure and building damage (Koshimura et al. 2009a; Macabuag et al. 2016). TFFs are developed using empirical data and linear regression models together with ordinary least squares (OLS) parameter estimation (Suppasri et al. 2012b, a; Mas et al. 2012). Other methods proposed in the literature (Charvet et al. 2017; Macabuag et al. 2016) use disaggregated data for generalized linear

---

**Figure 3**
The tsunami inundation depth surface calculated from the field survey data
Distribution and classification of the damage data from three different sources and their reclassification to construct TFFs

Table 2

| Reference               | No. of buildings in area of study | Original damage classification | Reclassified damage levels     |
|-------------------------|----------------------------------|--------------------------------|--------------------------------|
| Paulik et al. (2019)    | 307                              | a. Complete Damage             | Damaged (a, b)                 |
|                         |                                   | b. Partial Damage (Unrepairable)|                               |
|                         |                                   | c. Partial Damage (Repairable)  | Non-Damaged (c, d)             |
|                         |                                   | d. No Damage                    |                                |
| Copernicus EMS (2018)   | 1,772                            | a. Destroyed                    | Destroyed (a)                  |
|                         |                                   | b. Damaged                      | Other (b, c)                   |
|                         |                                   | c. Possibly Damaged             |                                |
| Adriano et al. (2019)   | 1,559                            | a. Destroyed                    | Destroyed (a)                  |
|                         |                                   | b. Damaged                      | Other (b, c, d)                |
|                         |                                   | c. Possibly Damaged             |                                |
|                         |                                   | d. No Damage                    |                                |
models. For a comprehensive review of TFFs in the literature, see Tarbotton et al. (2015), Charvet et al. (2017). Since the predictive power of these models has not yet been quantitatively assessed and remains in discussion among researchers (Macabuag et al. 2016; Aránguiz et al. 2018), in this study, we will develop TFFs using the OLS model methodology of Koshimura et al. (2009b). In the OLS model, two sets of data are necessary for correlation. As presented at the beginning of Sect. 3, tsunami inundation depth and building damage classification are the basic datasets required for TFF development. The details of the model methodology are described more precisely in Koshimura et al. (2009b), Mas et al. (2012), Suppasri et al. (2012b, a), Aránguiz et al. (2018). Herein, we summarize the steps as follows:

1. Compile damage classification and tsunami intensity measurement data by means of field survey, remote sensing or other methods.

2. Assemble a one-to-one dataset of damage level and tsunami intensity measurements.

3. Sort the sample data in descending order with respect to the tsunami intensity measure.

4. Divide the sorted data into groups with the same number of structures, where one group becomes one point for the regression model.

5. Calculate the damage rate in each group and the median value of tsunami intensity measurements within the group.

6. Consider that the cumulative probability density function of building damage follows the standardized normal or log-normal distribution, where the values of mean ($\mu$) and standard deviation ($\sigma$) are calculated by least-square fitting of the hydrodynamic features and the inverse of the probability on normal or log-normal papers.

7. Apply the OLS model to the median values and damage rates of data points to estimate the parameters of the statistical distribution’s cumulative density function (i.e., $\mu$ and $\sigma$).

Figure 5

a Map of the damage data compiled through field survey by Paulik et al. (2019) b histogram of two classes of structures (i.e., damaged and other) within the respective inundation depth intervals; c Tsunami Fragility Function developed from the field survey damage data
4.1. TFF Developed Only from Field Survey Data

In this case, damage data were collected from field survey of tsunami damaged structures. Based on observations, the damage classification is compiled together with maximum inundation depth measurements at structure locations. Adequate measurement techniques need to be followed as suggested by UNESCO-IOC (2014). Figure 5a shows the plot of the damage data within the study area gathered by Paulik et al. (2019) during field survey. We used this information to construct a TFF for the area of study. Following the OLS model methodology and the reclassification of the data as shown in Table 2 and Fig. 4, we plotted the histogram of number of structures per group of inundation depths represented by its median value (Fig. 5b). Finally, a TFF is developed through regression analysis of the discrete set of damage probabilities and median of inundation depths of groups. The fitting calculation resulted in a correlation coefficient of $R^2 = 0.92$, with 0.318 and 0.748 as the mean and standard deviations of the function, respectively. The resulting function and the points for fitting are plotted in Fig. 5c.

4.2. TFF Developed Through Visual Interpretation of Damage from Satellite Images

In this case, damage data are obtained through visual interpretation of pre- and postdisaster optical satellite images where, despite the top-view limitations in identifying front-view damages, this method can directly identify washed-away or destroyed structures. Here, we used the preliminary report of building damage classifications provided by the Copernicus Emergency Management System (EMS) and published on the 2nd of October, 2018 (Copernicus EMS 2018). Copernicus EMS damage data for the area of study are shown in Fig. 6a, with damage...
classified according to three levels: (i) destroyed, (ii) damaged and (iii) possibly damaged. We reclassified this information into (i) destroyed and (ii) other, following the criteria described in Table 2 and Fig. 4. The histogram in Fig. 6b shows a number of structures damaged at very low inundation depths; thus, in Fig. 6c, an initial damage probability at zero inundation level is observed. Several possible explanations for this result can be discussed as follows:

- Damage misclassification: Due to the limitation on optical imagery resolution and the bias of the user performing the classification, structures at lower inundation levels might have been interpreted as 'destroyed' when the actual level of damage was significantly lower.
- Inundation depth surface: In places where no clear field survey data on maximum run-up or flow depth were available, the interpolation algorithm might insert errors with considerably lower values regardless of the damage classification of the structures nearby.
- Explanatory variable: In this case, only tsunami inundation depth is used to correlate the damage to structures; however, because correlation does not imply causation, it is possible that the damage observed at lower depths could have been produced by other sources, such as the earthquake itself, the tsunami flow or tsunami debris (Song et al. 2017).

In addition, the empirical TFF related the structure damage to the maximum inundation flow depth measured after the disaster, while the damage might actually have occurred before the flow depth or flow velocity reached maximum values. In that case, the damage probability is underestimated. Unfortunately, methods to verify the time of damage of a large number of structures exposed to tsunami forces are very limited. A discussion on this matter can be found in Suppasri et al. (2019).

Despite the unsettling damage probability at depth zero, in this case the data were best fit to a normal cumulative distribution, with a high correlation coefficient $R^2 = 0.97$, and plotted as shown in Fig. 6c using the resulting mean (1.287) and standard deviation (1.210).

4.3. TFF Developed Using Machine Learning Classification

In this case, damage data were obtained through machine learning classification using a remote sensing approach proposed by Adriano et al. (2019). The work uses a multisensor and multitemporal approach for damage classification. A multisensor approach means, in this case, the use of synthetic aperture radar (SAR) imagery (i.e., Sentinel-1 and ALOS-2 PAL-SAR-2 datasets) and optical imagery (i.e., Sentinel-2 and PlanetScope datasets), while multitemporal refers to pre- and postevent acquisitions. The machine learning-based approach used a canonical correlation forest algorithm (Rainforth and Wood 2015) to classify building damage. Since the method used building data provided by OpenStreetMap, a larger number of structures were available here, though in some cases polygons that are not correlated to actual buildings might also be present. Figure 7a shows the damage classification in the area of study accomplished by applying Copernicus EMS classes. In Adriano et al. (2019), the same classes are used in order to enable training and verification of their method with the Copernicus data. The reclassification for TFF development is shown in Table 2 and Fig. 4. Similar to the visual interpretation method, the histogram shown in Fig. 7b and the TFF of Fig. 7c present an unexpected damage probability at zero inundation. We believe that this TFF is estimating damage probabilities at zero inundation depth for similar reasons described for the TFF produced from VI data. Nevertheless, in this case, a slightly higher correlation coefficient than before, $R^2 = 0.98$, was obtained for a normal cumulative distribution with mean of 1.612 and standard deviation of 1.126.

5. Discussion

5.1. TFFs of Three Types of Damage Data

Three sets of damage data were used to develop tsunami fragility functions (TFFs) for building structures affected by the 2018 Sulawesi tsunami at Palu in Indonesia. Figure 8 shows the results of developing TFFs with these data: (i) posttsunami field survey data (FS), (ii) Copernicus EMS visual damage
interpretation data (VI), and (iii) machine learning-based remote sensing (MLRS) classification of building damage data. The former uses tsunami inundation depth measured in the field at the same locations of the structures, while the latter two are correlated to a tsunami inundation surface constructed from the aggregation of several measurements in the field. Recall that since the survey points are not located exactly at the sites of the classified structures, the interpolation of an inundation depth surface becomes necessary.

First, we observe that the FS approach led to a more consistent function than did VI and MLRS, particularly due to the initial damage probability observed at the zero inundation depth in the latter two methods. One of the reasons that the FS approach does not present a damage probability at zero inundation depth might be the geospatial characteristics of the sample, where most of the surveyed points are located within the inundation area a certain distance from the inundation limit. The inundation limit is the zone of uncertainty and error of misclassification observed with the VI and MLRS methods. The reasons, as mentioned before, might be related to
the misclassification of damage near the tsunami inundation limit where flow depth is minimal. On the other hand, these structures might have been correctly classified as damaged buildings; however, the source of damage could have been the strong ground motion from the earthquake alone. Unfortunately, it remains a challenging task to clearly distinguish sources of the resulting damage during a field survey or a postevent remote-sensing-based survey. In addition, the use of an inundation depth surface, interpolated from points not located exactly where each structure with corresponding available damage information was, might have introduced errors into the dataset with respect to the regression. Additionally, there is a limitation in constructing a TFF with only one explanatory variable when other effects could have contributed significantly to the damage of structures near the inundation limit. This was also pointed out by Charvet et al. (2017), Aránzúig et al. (2018), Song et al. (2017).

However, while the TFF developed with the FS data resulted in an apparently better description of the damage in the area, with lower uncertainty of damage classification, this method requires high levels of management and organization and is time-consuming. Moreover, survey teams may be at risk when entering recently devastated areas and might contribute to social anxiety and false expectations of disaster relief. In contrast, VI and MLRS are remote-based approaches that pose no risk and are noninvasive. The first, VI, has shown higher accuracy in the past (Mas et al. 2015), but it is also time-consuming, and multiple biases are inserted when it involves multiple users. Moreover, when weather conditions are not favorable, optical images might be covered by clouds and VI cannot be conducted.

On the other hand, MLRS is a novel approach that is gaining momentum; throughout its improvement, higher accuracy comparable to VI has been achieved (Moya et al. 2018a, 2020). We compared the damage interpretation from the VI and MLRS methods for the same locations of structures available by FS. The summary of accuracy assessments for both methods are in Tables 3 and 4. The tables suggest that the overall and the balanced accuracy are similar in cases of both VI and MLRS. Thus, considering the field survey data as the baseline for accuracy assessment, both methods exhibit satisfactory classification accuracy. With such high accuracy and the increasing development of radar sensors, remote sensing and machine learning algorithms offer a promising future for rapid damage mapping (Syifa et al. 2019). Recall that the machine learning classifier used in MLRS was calibrated with the data from the VI; therefore, it is expected that both the VI and the MLRS results should have the same level of accuracy. The relevance of machine learning algorithms is that they can be rapidly applied to cover the entire tsunami-affected area, given that the inundation depth is also available; thus, much more data could be available to construct a more reliable TFF. This damage assessment is critical during the first stages of the disaster response. A clear topic for future discussion would be the dependency on training data to build a reliable model. Some efforts to avoid the dependence of this information have been reported elsewhere (Moya et al. 2018b).

In addition, since remote sensing approaches are limited by the top-view for damage interpretation, it

| Table 3 | Accuracy assessment of damage data from visual interpretation with respect to the field survey |
|---------|--------------------------------------------------------------------------------------|
| Visual interpretation | Recall | Precision | F1 score |
| Destroyed | Other | Destroyed | Other | Destroyed | Other |
| 0.71 | 0.69 | 0.46 | 0.87 | 0.56 | 0.77 |
| Overall accuracy: 0.70 | Balanced accuracy: 0.66 |

| Table 4 | Accuracy assessment of damage data from machine learning remote sensing with respect to the field survey |
|---------|--------------------------------------------------------------------------------------|
| Machine learning and remote sensing | Recall | Precision | F1 score |
| Destroyed | Other | Destroyed | Other | Destroyed | Other |
| 0.81 | 0.68 | 0.38 | 0.94 | 0.52 | 0.79 |
| Overall accuracy: 0.71 | Balanced accuracy: 0.66 |
is expected that underestimation of damage might occur. However, in this case, overall overestimation was observed. Looking deeper into this contradiction to our hypothesis, we found that among the remaining 30% of non-accurate classification in the VI, 22% was due to underestimation (i.e. the VI classified a structure on a lower damage level than the field survey), and 8% was due to overestimation. The hypothesis thus holds for the point-by-point comparison of methods; however, it still does not explain the overall overestimation in the final TFF. Upon examining Fig. 6, one notices that a highly damaged area to the east side of the bay is concentrated at lower inundation depths (i.e. closer to the inundation limit). Here, Fig. 9 shows that the VI method from Copernicus EMS classified this area as damaged for this event; however, inspecting photographs from Google Street View revealed that

Figure 9

a Reclassified visual interpretation of damage provided by Copernicus EMS and the tsunami inundation surface in the background; b damage data from (a) with a Google satellite image as background; c Palu Bay. The yellow inset shows the locations of areas in (a) and (b); d Google Street View screenshot from September 2017, where an apparently abandoned structure is present; e Google Street View screenshot from August 2018, where the structure is no longer present before the tsunami event
the conditions of apparent damage were present even before the event. It is most certain that including all these data into developing the TFF yielded the overestimation observed in Fig. 8. While this misclassification affected the lower limits of the damage probability, the upper limits are still consistent with the TFF developed from survey data. The TFF from FS was constructed from data focused on higher to intermediate levels of inundation; therefore, the uncertainties occurring at lower levels of inundation are not present and the TFF holds as consistent and reliable to represent the damage probability of this event.

Another interesting finding within this particular case of misclassification was that although the MLRS classification used the VI data for training the classifier, it also included side-looking SAR images, which improved the classification as shown in Fig. 10. MLRS classified the areas misclassified by the VI method as nondamaged. This verifies the remarkable advantage of the machine learning and remote sensing approach, where an automatic classifier surpassed a human classifier.

Overall, as one can notice in Fig. 8, TFF results of three types of damage data are similar above the 1-m inundation depth. Since tsunami hazards and damage are very limited at lower inundation depths, one can still argue that the VI and MLRS methods are highly advantageous for fair estimation of tsunami damage probability in higher ranges of inundation depth. This is understandable since higher damage conditions can be better observed from satellite images.

5.2. Sulawesi–Palu TFF Compared to Other TFFs

Now, let us concentrate on the TFF developed with the FS data. We found that the shape of the function is consistent with TFFs developed from previous tsunami events. In particular, we cite the TFF developed after the 1993 Nansei-oki Hokkaido earthquake using the damage data from the Okushiri area (Koshimura and Kayaba 2010) (Fig. 11). Similarities might be due to the condition and characteristics of structures, which were primarily wooden in Okushiri (Koshimura and Kayaba 2010) and composed of light timber framed with sheet metal walls and roofing in the case of Palu City (Kijewski-Correa and Robertson 2018). Nevertheless, a slightly lower probability of damage was observed in the area of study compared to that observed in Okushiri. On the other hand, Koshimura et al. (2009b) developed TFF after the 2004 Indian Ocean tsunami with the damage data collected from Banda Aceh. Figure 11 shows the comparison of the Banda Aceh and the Palu TFF. At the time of the tsunami in 2004, Banda Aceh presented mainly nonengineered
reinforced concrete, confined masonry, and timber-framed buildings (Saatcioglu et al. 2006), which may explain the differences in probability damage at low level inundation depths. For instance, at 2-m inundation depth, Banda Aceh TFF shows a 0.19 damage probability, while a 0.69 rate of damage can be expected at the same level in Palu City. The differences in ‘fragility’ between materials and construction on these two areas are reflected in the TFF. Interestingly, though, both TFFs reach nearly 100% damage probability around the 5-to-6-m inundation depth. In Palu, a maximum of 5-to-6-m tsunami height was observed, while 7 to 9-m maximum height was estimated for Banda Aceh.

Figure 11
Comparison of TFFs. Left: Okushiri vs. Sulawesi–Palu events; Right: Banda Aceh vs. Sulawesi–Palu events

Table 5
Summary of observations from the comparison of TFFs developed from three different sources of damage data

| Source of data            | Benefits                                                                 | Limitations                                                                 |
|--------------------------|--------------------------------------------------------------------------|----------------------------------------------------------------------------|
| Field survey (FS)        | • Shows higher consistency at lower inundation depth levels where little to no damage is expected. | • Requires high levels of management, organization and coordination, thus becoming a time-consuming task. |
|                          | • Lower uncertainty and higher detail of damage classification.           | • May disrupt disaster relief or create social expectations in affected areas. |
| Visual interpretation    | • Large coverage with high classification accuracy (approximately 70% accuracy in this study). | • May place research teams at risk of harm.                                |
| (VI) from satellite      | • Noninvasive approach. Conducted remotely so that no direct social stress was exerted in affected areas. | • Uncertainty and inaccuracies, in particular at lower inundation depth levels, of damage classification. |
| images                   |                                                                          | Satellite images taken by optical remote sensors can be affected by clouds or aerosol optical thickness in bad weather. |
| Machine learning         | • Large coverage with high classification accuracy (approximately 71% accuracy in this study). | • Damage classification is biased by the visual interpretation of the producer. |
| remote sensing (MLRS)    | • Noninvasive approach. Conducted remotely so that no direct social stress was exerted in affected areas. | • It is time-consuming and resource intensive.                             |
|                          | • May surpass human-based classifications when combining multiple sources of satellite imagery data. | • Uncertainty and inaccuracies, in particular at lower inundation depth levels, of damage classification. |
|                          | • Nonbiased classification (at least with respect to the producer or human interpreter). |                                                                            |
|                          | • Rapid and requires limited resources.                                   |                                                                            |
Regardless of the differences in maximum values surveyed in Palu and Banda Aceh, both TFFs reach the highest probability at 5-to-6-m depth. Thus, at least for Indonesian structures, it is recommended that 5 to 6 m of inundation depth should be considered as the critical level for any kind of structure exposed to tsunami. With this TFF, local authorities can generate damage scenarios of future possible tsunami disasters (Moya et al. 2017, 2018c) and improve preparedness and disaster risk reduction activities.

6. Conclusions

Different sources of building damage data were used to develop tsunami fragility functions (TFFs) for the case of the 2018 Sulawesi tsunami at Palu in Indonesia. TFFs from (i) posttsunami field survey data, (ii) Copernicus EMS visual damage interpretation data and a tsunami inundation depth surface constructed from the posttsunami field measurements, and (iii) machine learning-based remote sensing classification of building damage data combined with the tsunami inundation depth surface. Based on the analysis and comparison of the developed tsunami fragility functions, we summarize our observations in Table 5.

For this tsunami damage observed after the 2018 Sulawesi event in Indonesia, the FS TFF exhibits a similar shape to the function developed after the 1993 Hokkaido Nansei-oki tsunami, with a slightly lower damage probability between 0-to-5-m inundation depth. However, in comparison to other TFFs developed after events in Indonesia, for instance in Banda Aceh after the 2004 Indian Ocean tsunami, the characteristics of Palu structures exhibit higher fragility in response to tsunamis. A two-meter inundation depth correlates with nearly 20% probability of damage in the case of Banda Aceh, while approximately 70% probability of damage was evaluated at the same depth in Palu. Furthermore, 5-to-6-m inundation depth might be considered as the level of maximum tsunami damage, at least for low-rise unreinforced masonry buildings.

Acknowledgements

This study was partly funded by the Japan Science and Technology Agency (JST) J-Rapid project number JPMJJR1803; the JST CREST project number JP-MJCR1411; the Japan Society for the Promotion of Science (JSPS) Kakenhi Programs (17H06108, 17H02050, and 17H01293); the Core Research Cluster of Disaster Science at Tohoku University, Japan (a Designated National University); the MEXT Next Generation High-Performance Computing Infrastructures and Applications R&D Program; and the National Fund for Scientific, Technological and Technological Innovation Development (Fondecyt - Peru) [contract number 038-2019]. The collaboration of Suppasri, A. and Pakoksung, K. was supported by Tokio Marine & Nichido Fire Insurance Co., Ltd.; Willis Research Network (WRN); Pacific Consultants Co., Ltd.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

Adriano, B., Mas, E., Koshimura, S., Estrada, M., & Jimenez, C. (2014). Scenarios of earthquake and tsunami damage probability in Callao region, Peru using tsunami fragility functions. Journal of Disaster Research, 9(6), 968–975. https://doi.org/10.20965/jdr.2014.p0968.
Adriano, B., Xia, J., Baier, G., Yokoya, N., & Koshimura, S. (2019). Multi-source data fusion based on ensemble learning for rapid building damage mapping during the 2018 Sulawesi
earthquake and tsunami in Palu, Indonesia. Remote Sensing, 11(7), 886. https://doi.org/10.3390/rs11070886.

Alam, M. S., Barbosa, A. R., Scott, M. H., Cox, D. T., & van de Lindt, J. W. (2018). Development of physics-based tsunami fragility functions considering structural member failures. Journal of Structural Engineering, 144(3), 1–17. https://doi.org/10.1061/(ASCE)ST.1943-541X.0001953.

Aránguiz, R., Urra, L., Okuwaki, R., & Yagi, Y. (2018). Development and application of a tsunami fragility curve of the 2015 tsunami in Coquimbo, Chile. Natural Hazards and Earth System Sciences, 18(8), 2143–2160. https://doi.org/10.5194/nhess-18-2143-2018.

Arikawa, T., Muhari, A., Okumura, Y., Dohi, Y., Afriyanto, B., Sujatmiko, K. A., et al. (2018). Coastal subsidence induced several tsunamis during the 2018 Sulawesi earthquake. Journal of Disaster Research, 13, sc20181201. https://doi.org/10.20965/jdr.2018.sc20181201.

Carvajal, M., Araya-Cornejo, C., Sepúlveda, I., Melnick, D., & Haase, J. S. (2019). Nearly-instantaneous tsunamis following the Mw 7.5 Palu earthquake. Geophysical Research Letters, 46(10), 5117–5126. https://doi.org/10.1029/2019GL082578.

Charvet, I., Macabuag, J., & Rossetto, T. (2017). Estimating tsunami-induced building damage through fragility functions: Critical review and research needs. Frontiers in Built Environment. https://doi.org/10.3389/fbuil.2017.00036.

Chock, G., Yu, G., Thio, H. K., & Lynett, P. J. (2016). Target structural reliability analysis for tsunami hydrodynamic loads of the ASCE 7 standard. Journal of Structural Engineering, 142(11), 04016092. https://doi.org/10.1061/(ASCE)ST.1943-541X.0001499.

Condori Uribe, J. W. (2013). Análisis y Diseño Estructural de Evacuación Vertical Resistente a Sismos y Tsunamis. (In Spanish). Tech. rep., SENCICO, Lima, Peru. http://sencico.gob.pe/publicaciones.php?id=501.

Copernicus EMS (2018). Copernicus Emergency Management Service (European Union), EMSR317. https://emergency.copernicus.eu/mapping/list-of-components/EMSR317.

Gusman, A. R., Supendi, P., Nugraha, A. D., Power, W., Latief, H., Sunendar, H., et al. (2019). Source model for the tsunami inside coastal communities. Pure and Applied Geophysics, 176(8), 3291–3304. https://doi.org/10.1007/s00024-019-02258-5.

Moya, L., Mas, E., & Koshimura, S. (2017). Evaluation of tsunami fragility curves for building damage level allocation. Research Report of Tsunami Engineering, 34, 33–41.

Moya, L., Marval Perez, L., Mas, E., Adriano, B., Koshimura, S., & Yamazaki, F. (2018a). Novel unsupervised classification of collapsed buildings using satellite imagery, hazard scenarios and fragility functions. Remote Sensing, 10(2), 296. https://doi.org/10.3390/rs10020296.

Heidarzadeh, M., Muhari, A., & Wijanarto, A. B. (2019). Insights on the source of the 28 September 2018 Sulawesi tsunami, Indonesia based on spectral analyses and numerical simulations. Pure and Applied Geophysics, 176(1), 25–43. https://doi.org/10.1007/s00024-018-2065-9.

Kijewski-Correa, T., & Robertson, I. (2018). STEER: Structural Extreme Event Reconnaissance network: Palu earthquake and tsunami, Sulawesi, Indonesia Field Assessment Team 1 (FAT-1) Early Access Reconnaissance Report (EARR). Tech. rep., STEER.

Koshimura, S., & Kayaba, S. (2010). Tsunami fragility inferred from the 1993 Hokkaido Nansei-oki earthquake tsunami disaster. Journal of Japan Association for Earthquake Engineering, 10(3), 87–101. https://doi.org/10.5610/jaee.10_3.87. (in Japanese with English abstract).

Koshimura, S., Namegaya, Y., & Yanagisawa, H. (2009a). Tsunami fragility — A new measure to identify tsunami damage –. Journal of Disaster Research, 4(6), 479–488. https://doi.org/10.20965/jdr.2009.p0479.

Koshimura, S., Oie, T., Yanagisawa, H., & Imamura, F. (2009b). Developing fragility functions for tsunami damage estimation using numerical model and post-tsunami data from Banda Aceh, Indonesia. Coastal Engineering Journal, 51(3), 243–273. https://doi.org/10.1142/S0578783409002004.

Koshimura, S., Muhari, A., Adriano, B., Moya, L., Ayunda, D., Afriyanto, B., et al. (2019). Field survey of the 28 September earthquake tsunami of Sulawesi, Indonesia. Geophysical Research Abstracts, 21. https://meetingorganizer.copernicus.org/EGU2019/EGU2019-12100-2.pdf.

Macabuag, J., Rossetto, T., Ioannou, I., Suppasri, A., Sugawara, D., Adriano, B., et al. (2016). A proposed methodology for deriving tsunami fragility functions for buildings using optimum intensity measures. Natural Hazards, 84(2), 1–29. https://doi.org/10.1007/s11069-016-2485-8.

Mas, E., Koshimura, S., Suppasri, A., Matsuoka, M., Matsuyama, M., Yoshii, T., et al. (2012). Developing tsunami fragility curves using remote sensing and survey data of the 2010 Chilean tsunami in Dichato. Natural Hazards and Earth System Science, 12(8), 2689–2697. https://doi.org/10.5194/nhess-12-2689-2012.

Moya, L., Bricker, J. D., Kure, S., Adriano, B., Yi, C. J., Suppasri, A., et al. (2015). Field survey report and satellite image interpretation of the 2013 Super Typhoon Haiyan in the Philippines. Natural Hazards and Earth System Science, 15(4), 805–816. https://doi.org/10.5194/nhess-15-805-2015.

Mikami, T., Shibayama, T., Esteban, M., Takabatake, T., Nakamura, R., Nishida, Y., et al. (2019). Field survey of the 2018 Sulawesi tsunami: Inundation and run-up heights and damage to coastal communities. Pure and Applied Geophysics, 176(8), 3291–3304. https://doi.org/10.1007/s00024-019-02258-5.

Moya, L., Mas, E., & Koshimura, S. (2017). Evaluation of tsunami fragility curves for building damage level allocation. Research Report of Tsunami Engineering, 34, 33–41.

Moya, L., Marval Perez, L., Mas, E., Adriano, B., Koshimura, S., & Yamazaki, F. (2018a). Novel unsupervised classification of collapsed buildings using satellite imagery, hazard scenarios and fragility functions. Remote Sensing, 10(2), 296. https://doi.org/10.3390/rs10020296.

Moya, L., Mas, E., Adriano, B., Koshimura, S., Yamazaki, F., & Liu, W. (2018b). An integrated method to extract collapsed buildings from satellite imagery, hazard distribution and fragility curves. International Journal of Disaster Risk Reduction, 31(March), 1374–1384. https://doi.org/10.1016/j.ijdrr.2018.03.034.

Moya, L., Mas, E., Koshimura, S., & Yamazaki, F. (2018c). Synthetic building damage scenarios using empirical fragility functions: A case study of the 2016 Kumamoto earthquake. International Journal of Disaster Risk Reduction, 31(October 2017), 76–84. https://doi.org/10.1016/j.ijdrr.2018.04.016.

Moya, L., Muhari, A., Adriano, B., Koshimura, S., Mas, E., Marval-Perez, L. R., et al. (2020). Detecting urban changes using phase correlation and L1-based sparse model for early disaster
response: A case study of the 2018 Sulawesi Indonesia earthquake-tsunami. Remote Sensing of Environment, 242 (July 2019), 111743. https://doi.org/10.1016/j.rse.2020.111743.

Muhari, A., Imamura, F., Arikawa, T., Hakim, A. R., & Afriyanto, B. (2018). Solving the puzzle of the September 2018 Palu, Indonesia, tsunami mystery: Clues from the tsunami waveform and the initial field survey data. Journal of Disaster Science, 13 (September), 2–4. https://doi.org/10.20965/jdr.2018.sc20181108.

Nakata, K., Katsuma, A., & Muhari, A. (2020). Submarine landslide source models consistent with multiple tsunami records of the 2018 Palu tsunami. Sulawesi, Indonesia: Earth, Planets and Space. https://doi.org/10.1186/s40623-020-01169-3.

Nanayakkara, K. I. U., & Dias, W. P. S. (2016). Frailty curves for structures under tsunami loading. Natural Hazards, 80 (1), 471–486. https://doi.org/10.1007/s11069-015-1978-1.

Omira, R., Dogan, G. G., Hidayat, R., Husrin, S., Prasetya, G., Annunziato, A., et al. (2019). The September 28th, 2018, tsunami in Palu-Sulawesi, Indonesia: A post-event field survey. Pure and Applied Geophysics, 176 (8), 1379–1395. https://doi.org/10.1007/s00024-019-02145-z.

Pakoksung, K., Suppasri, A., Imamura, F., Athanasius, C., Omang, A., & Muhari, A. (2019). Simulation of the submarine landslide tsunami on 28 September 2018 in Palu Bay, Sulawesi island, Indonesia, using a two-layer model. Pure and Applied Geophysics, 176 (8), 3323–3350. https://doi.org/10.1007/s00024-019-02235-y.

Paulik, R., Gusman, A., Williams, J. H., Pratama, G. M., Lin, S.-L., Prawirabhati, A., et al. (2019). Tsunami hazard and built environment damage observations from Palu City after the September 28 2018 Sulawesi earthquake and tsunami. Pure and Applied Geophysics, 176 (8), 3305–3321. https://doi.org/10.1007/s00024-019-02254-9.

Pelinovsky, E., Yuliated, D., Prasetya, G., & Hidayat, R. (1997). The 1996 Sulawesi tsunami. Natural Hazards, 16 (1), 29–38. https://doi.org/10.1023/A:1007904610680.

Putra, P. S., Aswan, A., Maryunani, K. A., Yulianto, E., & Kongko, W. (2019). Field survey of the 2018 Sulawesi tsunami deposits. Pure and Applied Geophysics, 176 (6), 2203–2213. https://doi.org/10.1007/s00024-019-02181-9.

Rainforth, T., & Wood, F. (2015). Canonical correlation forests. arXiv, pp 1–51, arXiv:1507.05444.

Rehman, K., & Cho, Y.-S. (2016). Building damage assessment using scenario based tsunami numerical analysis and fragility curves. Water, 8 (3), 109. https://doi.org/10.3390/w8030109.

Robertson, I. (2018). Palu earthquake and tsunami, Sulawesi, Indonesia Preliminary Virtual Assessment Team (PVAT) report. Tech. Rep. October, StEER. https://www.designsafe-ci.org/data/s00024-019-02290-5.

Saatioglu, M., Ghobarah, A., & Nistor, I. (2006). Performance of structures in Indonesia during the December 2004 great Sumatra earthquake and Indian Ocean tsunami. Earthquake Spectra, 22 (3 suppl), 295–319. https://doi.org/10.1193/1.2209171.

Sassa, S., & Takagawa, T. (2019). Liquefied gravity flow-induced tsunami: first evidence and comparison from the 2018 Indonesia Sulawesi earthquake and tsunami disasters. Landslides, 16 (1), 195–200. https://doi.org/10.1007/s10346-018-1114-x.
Widiyanto, W., Santoso, P. B., Hsiao, S.-C., & Imananta, R. T. (2019). Post-event field survey of 28 September 2018 Sulawesi earthquake and tsunami. *Natural Hazards and Earth System Sciences, 19*(12), 2781–2794. https://doi.org/10.5194/nhess-2019-91.

Wiebe, D. M., & Cox, D. T. (2013). Application of fragility curves to estimate building damage and economic loss at a community scale: a case study of Seaside, Oregon. *Natural Hazards, 71*(3), 2043–2061. https://doi.org/10.1007/s11069-013-0995-1.

(Received March 27, 2020, revised April 30, 2020, accepted May 4, 2020, Published online June 3, 2020)