VOLUME 13 No. 1, 22 Desember 2023 Halaman 1-17

MICROSOFT BUILDING FOOTPRINT APPLICATION TO DETECT HUMAN EXPOSURE DUE TO TSUNAMI

PENGGUNAAN MICROSOFT BUILDING FOOTPRINT UNTUK MENDETEKSI KETERPAPARAN MANUSIA TERHADAP TSUNAMI

Andes Saragi*
Geo-Information for Disaster Management, Master of Environmental Science
The Graduate School, Universitas Gadjah Mada

Djati Mardiatno
Faculty of Geography, Universitas Gadjah Mada

Dyah Rahmawati Hizbaron
Faculty of Geography, Universitas Gadjah Mada

Submitted: 2022-11-29; Revised: 2023-08-31; Accepted:2023-09-03

ABSTRAK

Kejadian tsunami di malam hari lebih rentan menimbulkan korban jiwa karena manusia sedang beristirahat pada bangunan permukiman (rumah). Pada penelitian ini ekstraksi bangunan permukiman dilakukan menggunakan *Microsoft Building Footprint* (MBF) yang merupakan hasil penerapan teknologi *Artificial Intelligence*. Penelitian ini bertujuan untuk menganalisis jumlah jiwa manusia terpapar tsunami dimalam hari menggunakan MBF. Pemodelan tsunami dilakukan dengan metode *Berryman*. Citra Sentinel 2-A diekstraksi dari *Google Earth Engine*. Hasil analisis pemodelan inundasi menunjukkan total area tergenang seluas 717 Ha atau 17,34% dari total area. Hasil analisis akurasi MBF terhadap keseluruhan data yaitu: *Precision* 99,02%, *Recall* 98,40% dan *F1 score* 98,71%. Hasil analisis kesalahan MBF yaitu *False Positif* 0,97%, *False Negatif* 1,60% dan *Intersection of Union* 0,12%. Jumlah manusia terpapar yaitu 2.749 jiwa atau 6,32% total keseluruhan penduduk.

Kata kunci: *Microsoft Building Footprint*; Model Tsunami; Manusia Terpapar; Bangunan Permukiman; *Google Earth Engine;* Citra Sentinel 2-A.

ABSTRACT

Tsunami events at night are more prone to causing fatalities because humans are resting in residential buildings (houses). In this study, residential buildings were extracted using the Microsoft Building Footprint (MBF), which resulted from applying artificial intelligence technology. This study aims to analyze the number of people exposed to tsunamis at night using MBF. The tsunami modeling was carried out using the Berryman method. Sentinel 2-A Image extracted from Google Earth Engine. The results of the inundation modeling analysis show that the total inundated area is 717 Ha or 17.34% of the total area. The results of the MBF accuracy analysis on the entire data

*Corresponding author: andes.saragi@mail.ugm.ac.id

are a Precision of 99.02%, Recall of 98.40%, and F1 score of 98.71%. The results of the MBF error analysis are False Positive 0.97%, False Negative 1.60%, and Intersection of Union 0.12%. The number of people exposed is 2,749, or 6.32% of the total population.

Keywords: ***Microsoft Building Footprint; Tsunami Model; Human Exposed; Residential Buildings; Google Earth Engine; Sentinel 2-A Image.***

INTRODUCTION

Tsunamis are relatively rare but present the most dangerous and deadly hazard (Mardiatno et al., 2020). In general, the tsunami that occurred in Indonesia was caused by a tectonic earthquake (Løvholt, 2014). It is supported by Indonesia's coastline, which is one of the longest in the world, and Indonesia's position, which is located on three large tectonic plates that move relatively close to each other. The closer the plates are to the collision between the plates, resulting in tectonic earthquakes, and if the crash occurs on the seabed, it will cause the potential for a tsunami to occur. One area in Indonesia prone to tsunami disasters is the South Coast of Java Island. The southern coast of Java is prone to tsunamis because of its proximity to the subduction zone (Aris et al., 2019; Nurulloh, 2019). This area encountered tsunamis in 1994 and 2006, which resulted in massive casualties (Mardiatno, 2013).

This research was conducted in the coastal area of Bantul Regency. Mapping the population distribution in coastal areas is very important to inform the government in making the right decisions regarding the priority of evacuation and rescue operations (Mohamadi et al., 2019). Based on data from the Central Statistics Agency (BPS) in 2022, around 4.36% or 43,525 of the 998,647 residents of Bantul Regency live in coastal areas. Meanwhile, the distance of the closest settlement from the coastline is only less than 50 meters, so it is necessary to do tsunami modeling to analyze the exposure of human life based on data on settlements in the coastal area of Bantul Regency. Landslide is a term used to describe the extent to which a tsunami wave reaches land (Basith et al., 2016).

Tsunami events at night generally will result in many casualties. This is because humans naturally rest at night. When humans sleep, the house is at risk of being affected by a tsunami. Ensuring safe and sustainable development has become a hot topic of debate, especially in areas prone to hazards from natural disasters (Hizbaron et al., 2012). Settlement arrangements are made in disaster-prone regions to minimize loss of life. The tsunami risk assessment will be part of pre-disaster actions that need to be carried out in every planning, especially in coastal areas, as one of the important sectors for regional development (Mardiatno, 2015). It is important in informing decision-making on effective responses and for informing the public (Fakhruddin, 2021). The study results are expected to be input to the community and local governments to minimize casualties in the future.

METHOD

Sentinel-2A Image Extraction Using Google Earth Engine (GEE)

To obtain the sentinel-2A satellite image, this study was carried out using the GEE platform. The reason for using GEE is that this platform has many advantages compared to others. GEE archives all data sets and links them to cloud computing engines for open-source use. The data archive includes data from other satellites and vector datasets based on Geographic Information Systems (Mutanga, 2019). Although this platform has been launched long, it has yet to be widely used for remote sensing applications (Amani et al., 2020). Users can easily access archived remote sensing data via JavaScript. It allows users to select the desired data from millions of temporal images inherited from Google. One of the advantages that the author reasons is that the results obtained are adjusted to the predetermined Area of Interest (AOI) limit. Unlike other platforms, the selection of AOI is only used as an initial approach, and the final result is adjusted to the scene image.

Data filtering and sorting capabilities are also very fast in GEE.

Tsunami Inundation Modeling

The tsunami inundation modeling was carried out to analyze the area affected by the tsunami inundation at the initial wave height determined from the coastline (BNPB, 2019). This modeling was carried out using the Hloss method developed by Berryman (Berryman, 2006). This study focuses on inundation modeling in the worst-case scenario. The coastline used is the highest tidal shoreline obtained from BIG. The wave height used is based on the worst possible scenario around the coast of Bantul Regency, which is 11 meters (Widiantoro et al., 2020). The range of tsunami inundation on land is affected by surface roughness, symbolized as a surface roughness coefficient. The roughness coefficient is related to the tsunami energy submerged by the ground surface due to the contact between the ground surface and seawater (Fauzi, 2021). Surface roughness is the result of conversion from land use. Land use changes will affect the tsunami run-up modeling (Griffin et al., 2015; Wahyuni, 2020). The higher the coefficient value, the greater the potential to reduce the speed of the water to land and vice versa (Smart et al., 2016). This study uses the coefficient of surface roughness from previous researchers, which considers the density and conditions of the surrounding environment (Khomarudin, 2010).

Microsoft Building Footprint (MBF) Analysis

Microsoft has made significant investments in deep learning, computer vision, and Artificial intelligence (AI) has been applied to mapping. Microsoft used the Open Source CNTK Unified Toolkit and implemented Deep Neural Networks (DNNs) and ResNet34 with the RefineNet up-sampling layer to detect building traces from Bing images (Microsoft, 2022a). Extraction of buildings is carried out in two stages, namely Semantic Segmentation and Polygonization (Microsoft, 2022b). A model-based approach that predicts building footprints from satellite imagery can be easily applied on a large scale (Robinson et al., 2022). Microsoft Building Footprint is the result of extraction using Bing Image developed by the Microsoft and Bing Team computerized to produce building trace data (Huang et al., 2021). The use of high-resolution remote sensing imagery is increasingly being explored to obtain building footprint information automatically because manually identifying building geometry using high-resolution imagery takes much time and is expensive (Touzani & Granderson, 2021; Liu et al., 2019). In MBF, several conditions of building footprints need to be analyzed more deeply. These conditions are True Detection, Undetected, and False Detection. True Detection is a condition where MBF correctly detects buildings in the Bing image. Undetected is the condition in the Bing image that there is a building. At the same time, the MBF does not detect it, and false Detection is the condition of the MBF that there is a polygon of a building, but the appearance on the Bing image is not a building. The semantic segmentation accuracy of the building footprints model was measured using metrics: Precision, Recall, and F1 Score (Touzani & Granderson, 2021), which are shown in equations (2.1), (2.2), and (2.3). This metric is based on the predicted value of pixels represented by the number of True Positives (TP), False Positives (FP) and False Negatives (FN):

$$Precision = \frac{True\ Positif\ (TP)}{True\ Positif(TP) + False\ \ Positif(FP)} \quad \text{.....(2.1)}$$

$$Recall = \frac{True\ Positif\ (TP)}{True\ Positif(TP) + False\ \ Negatif(FN)} \quad \text{.....(2.2)}$$

$$F1\ Score = \frac{2\ (TP)}{2\ (TP) + (FP) + (FN)}$$

$$= 2\ \frac{Precision\ x\ Recall}{Precision + Recall} \quad \text{.....(2.3)}$$

Where precision is the percentage of detected building truly positive footprints; True

Positive is a condition if the prediction is positive and the result is correct; False Positive is a condition if the prediction is positive, but the result is negative; False Negative is a condition if the prediction is negative, but the result is positive; Recall is the detection percentage of all reference building footprints. F1 Score is a combination of Precision and Recall in one number (Touzani & Granderson, 2021; Liu et al., 2022; Data analysis, 2022).

Table 1.
Confusion matrix for MBF analysis

	Negatif MBF Predicted	Positif MBF Predicted
Actual Negatif	True Negatif (TN) = No Data	False Positif (FP) = False Detection
Actual Positif	False Negatif (FN) = Undetected	True Positif (TP) = True Detection

Source: Data analysis, 2022

The relationship of True Positive, False Negative, and False Positive to True Detection, Undetected, and False Detection is presented in Table 1. Table 1 shows a confusion matrix for MBF analysis. In the column in bold, there are MBF Predicted Negative and Positive. MBF Predicted is building Detection on MBF. Negative if not detected and positive if detected. On the line, there are negative and positive. Actual describes the actual conditions of the Bing image. Negative if there are no buildings and positive if there are buildings. Semantic segmentation is the process of categorizing all pixels in a digital image into a set of categories (Ps & Aithal, 2022). This segmentation process is prone to errors, resulting in several traces of overlapping buildings. In MBF, the overlapping part does not become a new polygon, so it does not affect the analysis of the number of buildings. It should be analyzed to determine the overall quality of the resulting data. The equation used is (Huang & Jin, 2022):

$$Intersection\ of\ Union\ (IoU) = \frac{Area\ of\ Overlape}{Area\ of\ Union} \quad \text{.....(2.4)}$$

Where the Intersection of Union (IoU) is the percentage of new polygons formed due to segmentation errors of the entire data, The Area of Overlap is the area that overlaps, and the Area of Union is the area that merges. The number of true Detection, false Detection, and undetected must be known to determine the distribution of buildings in an area. The equation to find out the actual number of buildings (fixed buildings) in an area using MBF is:

$$F_B = (MBF - F_D) + U_D = T_D + U_D \quad \text{.....(2.5)}$$

Where F_B is Fix Building, T_D is True Detection, F_D is False Detection, and U_D is Undetected (Data analysis, 2022). The distribution of building footprints effectively summarizes population patterns in both urban and rural areas (Huang et al., 2021). People live in buildings. Building footprints can outperform commonly used weighting variables in population sorting. Building statistics such as number and size have a stronger relationship with population than other weighting variables such as urban land use, light intensity, and distance to roads (Small & Nicholls, 2003).

Analysis of Settlement Buildings and Lives Exposed

Exposure analysis was conducted using a combination of Geographic Information Systems (GIS) and data analysis tools (Fakhruddin, 2021). The results of the tsunami inundation modeling were then analyzed using a spatial join with MBF to determine the exposed buildings. The exposed buildings were then checked in the field and interviewed to determine the use of buildings and the number of people in residential buildings. Checking is done by sampling using the Slovin equation:

$$n = \frac{N}{(1+(Ne^2))} \quad \text{.....(3.2)}$$

Where n is the number of samples, N is the number of population, namely the number of polygons of residential buildings, and e is the desired error rate. The results of exposure to settlements are then classified into hazard zone levels. Data collection and the process used to conduct the analysis are divided into two categories: residential buildings and population. The amount of exposure to the population is obtained based on the exposure to settlements (houses) (Irianti, 2017) after analyzing the number of permanent residents. Research flow chart can be seen in Figure 1.

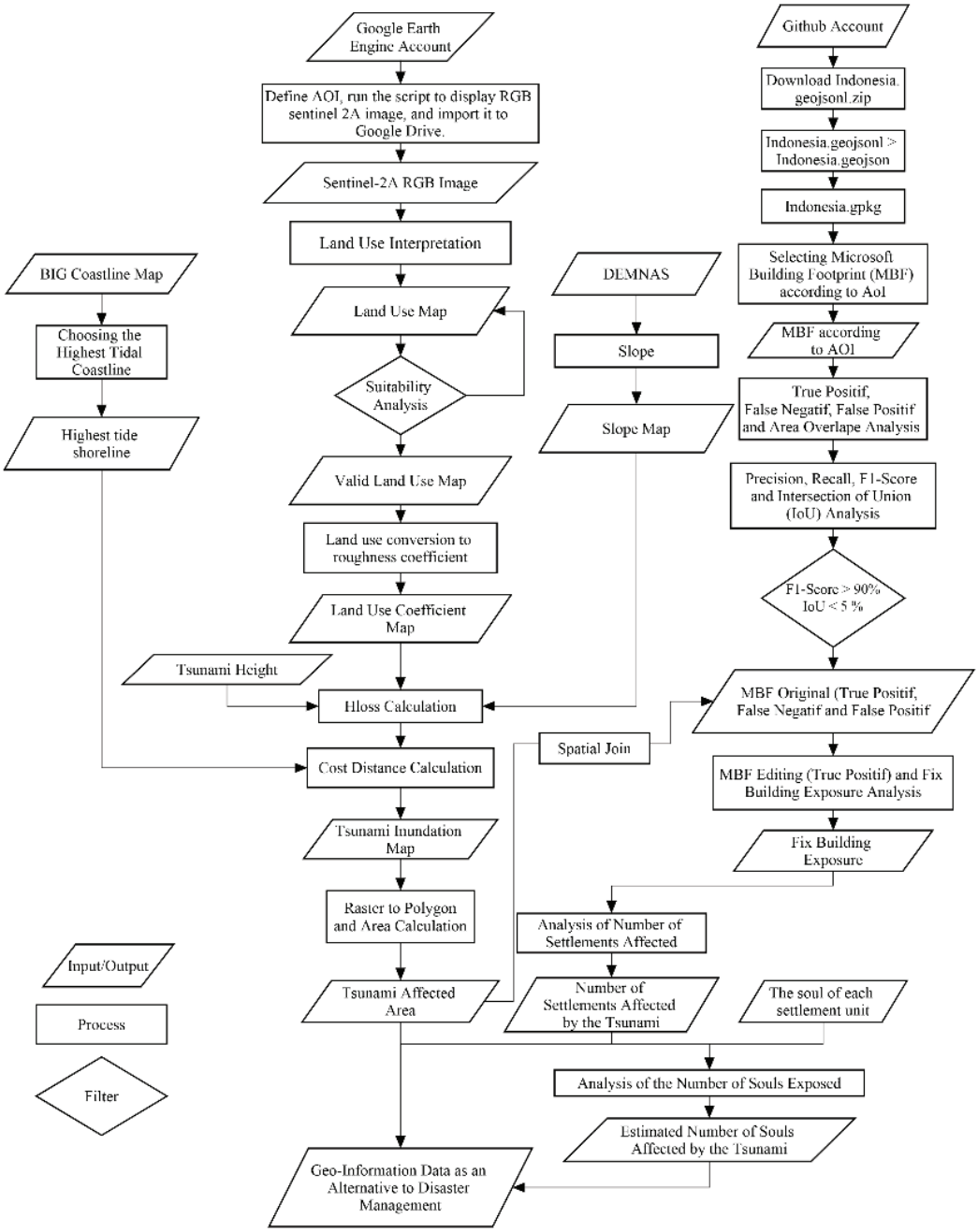

Figure 1.
Research flow chart
Source: Author (2022)

RESULTS AND DISCUSSION

Result of Tsunami Inundation Modeling Analysis

In this study, the inundation model was used to analyze human exposure to tsunamis in a scenario with a wave height of 11 meters. Analysis The tsunami inundation modeling was carried out in two stages. First, calculate Hloss using the raster calculator toolbox in ArcGIS with raster format input, namely the coefficient of roughness and slope in radians and the initial height of the wave. The initial height of the specified wave is 11 meters (Widiantoro et al., 2020). The wave height is the worst height around the study area based on the highest possible underwater earthquake scenario. The last step is calculating the inundation distance from the coastline using the cost distance toolbox. The coastline used is the highest tidal shoreline obtained from BIG. This scenario is to produce a wider coverage of inundation areas on land. The processing results are shown in Figure 2 and 3.

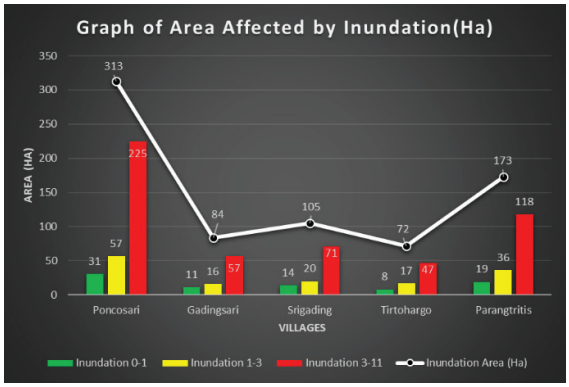

Figure 2.
Tsunami hazard map
Source: Data analysis (2022)

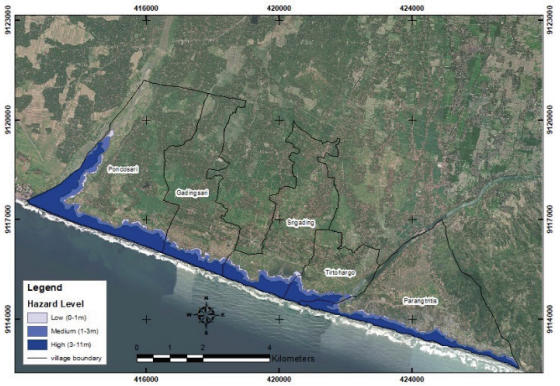

Figure 3.
Area affected by inundation
Source: Data analysis (2022)

In Figure 3, it can be seen that Poncosari Village is the village with the highest inundated area. This is because the Poncasari Village is directly adjacent to the river on the west side. The river has a surface roughness with the lowest coefficient value, so it is not too significant to inhibit water flow from the coastline to the mainland. The village with the second highest inundation area is Parangtritis. Based on Figure 3. This village has the longest coastline compared to other subdistricts, thus making the inundated area the second highest. This is inversely proportional to the village with the lowest inundation area, Tirtohargo. In Figure 2, it can be seen that Tirtohargo village is the village with the shortest coastline shared with another village. This affects the area that is inundated in the village.

MBF Analysis Results

In previous studies, building extraction was carried out through manual digitization using high-resolution imagery and aerial photographs (Khomarudin, 2010; Hadi & Astrid, 2017; Mardiatno et al., 2020; Alvianingsih et al., 2021; Paulik et al., 2021). However, manual digitization has its drawbacks: it is time-consuming and requires significant financial resources (Touzani & Granderson, 2021), making it inefficient for analyzing large study areas.

Undetected/False Negative Analysis Results

Undetected analysis was performed using a grid system to focus the analysis area and minimize missed areas. Undetected on MBF is the inability to detect a building on the Bing Image. Various things can cause this

inability. One of the causes of undetected is the appearance of buildings covered with vegetation. This will not be easy at the segmentation stage. Buildings that were not detected in this study were then digitized on the screen manually. If the appearance of the building is not too much covered by other vegetation, then the segmentation stage can be carried out correctly to produce clear and clean Detection on the MBF. The village with the highest number of undetected is Parangtritis. The high distribution of vegetation causes the high undetected in Parangtritis Village, and there are areas in Bing Image that have poor pixels.

Undetected analysis was then performed on land use. Analysis of undetected buildings on the MBF of land use aims to determine the distribution of the highest undetected land use. To perform the analysis is done using the spatial join toolbox in ArcGIS. Based on the processing and analysis results, it can be seen that the highest number of undetected is in mixed gardens (Figure 4). This makes sense because, in mixed gardens, the roof of the building is blocked by trees, making it difficult for MBF to analyze the pixels of the building inside. The second highest is a medium-density settlement. This is because the surrounding of the settlement is also overgrown by trees, which obstructs the roof of the building in Bing's image. The initial working principle of MBF is the segmentation stage, which is to analyze the similarity of building pixels in the Bing image, which in this case is the roof of the building. If the appearance of the pixels of the roof of the building is blocked by other pixels, the roof of the building does not form a complete angle, so it is not detected. The third highest is low-density settlements. The occurrence of undetected is also influenced by if there is no significant difference in pixel color between the building and the surrounding environment (considered the same).

(a) (b)

Figure 4.
Highest undetected in mixed garden land use, (a) Undetected, (b) In the field
Source: Data analysis (2022)

False Detection / False Positive Analysis Results

False Detection is an error in detecting buildings on the MBF. In this study, False detection analysis was carried out manually by observing each research area using a grid system. False Detection needs to be analyzed because MBF focuses on building trace data, so non-building traces need to be sorted. This detection error can be caused by the similarity of pixels and shapes in the image, which is considered to resemble the roof of a building by MBF, and some pixels are not good in the Bing Image. Poncosari Village has the highest number of false detections. The high false Detection in Poncosari Village is because there are many fish/shrimp ponds around the coast. Another condition is that Poncosari Village has the broadest administrative area among other villages.

False detection analysis of land use is carried out using the spatial join toolbox in ArcGIS. This analysis was conducted to determine the distribution pattern of false Detection on land use. The table attributes resulting from ArcGIS processing are then processed using a pivot table in Excel. The analysis results show that many false detections are in paddy fields (Figure 5.) and fish/shrimp ponds. The high number of false detections in the land use is because both have the characteristic appearance of pixels and a shape that is almost the same as the roof of the building. The third highest is the use of volcanic beach sand. Based on field checks, this is because there is a freshwater swimming pool located in a sandy area, especially in Parangtritis Village.

(a) (b)

Figure 5.
Highest false detection in paddy field use, False detection (a), In the field (b).
Source: Data analysis (2022)

True Detection/True Positive Analysis Results

True Detection is a true positive (positive prediction and positive actual) in modeling, namely the condition of the Microsoft Building Footprint correctly detecting a building. True detection analysis is carried out together with false detection analysis. To find out the number of true detections in an area or region, the formula used is the number of MBF minus the false Detection. The fixed building formula can calculate the total number of buildings if the number of true detections is known by adding the number of true detections and the number of undetected. The number of true detections and the number of buildings are presented in Table 2.

Table 2.
Number of true detection and number of buildings

Villages	MBF	False Detection	Undetected	True Detection	Total of Buildings
Poncosari	4312	56	58	4256	4314
Gadingsari	3106	46	24	3060	3084
Srigading	3532	12	26	3520	3546
Tirtohargo	1332	4	34	1328	1362
Parangtritis	3654	38	114	3616	3730
Total	15936	156	256	15780	16036

Source: Data analysis (2022)

Table 2 shows that the total number of true detection buildings is 15,780, and the total number of buildings after analyzing false Detection and undetected is 16,036. Poncosari Village has the highest number of buildings, namely 4314. The number of buildings in Poncosari Village is directly proportional to the area owned by other Villages.

Undetected and False Detection Percentage Analysis Results

Undetected/false negative and false detection/false positive are the two main errors in MBF that should be analyzed. This error is closely related to the accuracy of modeling building footprints on the MBF. Analysis of the percentage of undetected and false Detection can be calculated after knowing the total number of buildings obtained by adding true Detection and undetected. This analysis was conducted to determine the percentage of undetected and false Detection of the overall data in the research area. To find out the number of undetected and false, Detection is divided into the total number of buildings and multiplied by 100%. The calculation results are presented in Table 3.

Table 3.
Percentage of undetected and false detection

Villages	Undetected (%)	False Detection (%)
Poncosari	1,34	1,30
Gadingsari	0,78	1,49
Srigading	0,73	0,34
Tirtohargo	2,50	0,29
Parangtritis	3,06	1,02
Overall Data	**1,60**	**0,97**

Source: Data analysis (2022)

In Table 3. The percentage of undetected and false Detection is shown in each village. The percentage of undetected in all urban villages is below 5%. It is relatively good, with an overall score of 1.68%. The percentage of false Detection is below 1.5% in five villages, with a value of 0.97% for the overall data. The predominance of errors on the MBF is undetected because the value in all data is close to double the false Detection. The value of false Detection is very small. It shows that MBF is very accurate in detecting buildings and non-buildings.

Precision, Recall, and F1 Score Analysis Results

Analysis of precision, Recall, and F1 scores can be performed if the number of false negatives, false positives, and true positives is known. False negatives were obtained by performing undetected analysis on MBF and Bing images. False Positive is obtained by performing a false detection analysis, and a true detection analysis obtains a true positive. Based on the undetected, false Detection, and true detection analysis that has been done, it is possible to calculate the three matrices. The calculation results are presented in Table 4.

Table 4.
Value of Precision, Recall and F1 Score

Villages	Precision (%)	Recall (%)	F1 Score (%)
Poncosari	98,70	98,66	98,68
Gadingsari	98,52	99,22	98,87
Srigading	99,66	99,27	99,46
Tirtohargo	99,70	97,50	98,59
Parangtritis	98,96	96,94	97,94
Overall Data	**99,02**	**98,40**	**98,71**

Source: Data analysis (2022)

Table 4 presents the results of processing precision, Recall, and F1 scores in the research area. The precision, Recall, and F1 score values, which are close to 100, indicate that MBF modeling in the research area is very suitable for use. The findings of this study align with the statement made by Microsoft. These results were also supported by research conducted in three major cities in the United States, which showed an average accuracy matrix value of 85% (Touzani & Granderson, 2021). The precision value is related to the false positives obtained from false Detection, so the smaller the percentage of false Detection, the higher the precision

value. It can be seen that Tirtohargo Village has the highest precision value, as well as in Table 3. The percentage of false Detection in Tirtohargo Village is the lowest. The recall value is related to the false negative obtained from undetected, so the smaller the percentage of undetected, the higher the recall value. The F1 score is the average accumulated value of precision and Recall.

Intersection of Union (IoU) Analysis Results

The IoU analysis is an analysis to find out how much overlap occurs at the MBF segmentation stage. The analysis is done by comparing the overlap area and the union area. The overlap area was analyzed using the intersect toolbox in ArcGIS. Union areas were analyzed using the union toolbox in ArcGIS. The results of data processing are presented in Table 5.

Table 5.
Value of Intersection of Union

Sub-district	IoU (%)
Poncosari	0,11
Gadingsari	0,09
Srigading	0,14
Tirtohargo	0,13
Parangtritis	0,13
Overall Data	**0,12**

Source: Data analysis (2022)

Table 5. shows the intersection of union values in the five villages and the overall data. The IoU values in the five villages are not significantly different. The highest IoU value is found in Srigading Village, which is 0.14%. This value shows that the percentage of overlap area to the union area is quite high compared to other urban villages. The total IoU value in all data is only 0.12%. This value is very small, so the overlapping error in the MBF is very small in the research area so that it can be ignored.

MBF Editing Exposure Analysis Results

In MBF, large polygons still detect several buildings in one footprint, so editing is needed. Editing on the Detection MBF was done to determine the number per unit of buildings exposed to the tsunami. Edits are only made on accurate detections with large polygons. Editing is done by dividing a large polygon into several smaller polygons according to the appearance of the roof in the image. The final processing results are presented in Figure 6.

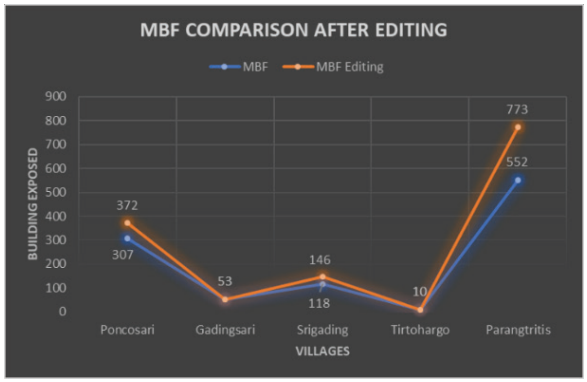

Figure 6.
Comparison graph of the number of buildings after MBF editing.
Source: Data analysis (2022)

Figure 6. shows a comparison graph of the number of buildings after and before editing the MBF. The graph shows that Gadingsari and Tirtohargo Villages are still the same. It shows that the density of buildings on the village coast is very low. Significant changes occurred in Poncosari and Parangtritis Villages. The significance of this change is very reasonable because many buildings with high density are on the coast of the village.

Final Result Analysis of Buildings Exposed to Tsunami Inundation from MBF

The final analysis of buildings exposed to tsunami inundation from MBF was carried out to determine the actual number of buildings exposed to tsunami inundation. MBF still has drawbacks. Namely, there are undetected and false Detection. MBF exposed to editing has not considered the presence of undetected and false Detection in the exposed area. The buildings exposed to MBF editing still need to be added to the number of undetected and subtracted from false detections to analyze the actual number. The results of data processing are presented in Table 6.

Table 6.
Number of False Detection, Undetected and Fix buildings exposed to tsunami inundation

Sub-district	MBF Editing	False Detection	Undetected	Fix Building
Poncosari	372	50	21	343
Gadingsari	53	15	1	39
Srigading	146	6	7	147
Tirtohargo	10	1	0	9
Parangtritis	773	23	29	779
Total				1317

Source: Data analysis (2022)

Table 6. Displays the final calculation of the number of fixed buildings or the actual number of buildings exposed to the tsunami inundation. After editing the original MBF in the inundation area, the fixed building value is obtained, then the false Detection and undetected in the exposed area. The fixed building value is obtained by reducing the number of MBF edits for false Detection and then increasing it by the number of undetected. The total number of buildings exposed to tsunami inundation as high as 11 meters reached 1,317. The highest number of fixed buildings is in Parangtritis Village, and the lowest is in Tirtohargo Village. The high number of buildings in Parangtiris Village is because, in the area, there are famous beach tourist destinations that are always crowded with visitors, thus attracting people to carry out development. The total fixed building exposed to the tsunami was then analyzed as a spatial analysis of the three predetermined inundation classes. It is to determine the distribution of buildings exposed to each class of tsunami inundation. The processing results are presented in Figure 7.

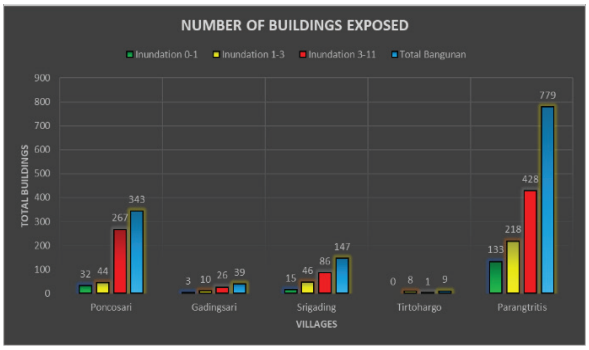

Figure 7.
Number of buildings exposed to tsunami inundation class
Source: Data analysis (2022)

In Figure 7. Shown the number of buildings exposed to each tsunami inundation at the height of 11 meters. At an altitude of 0-1 m inundation, a significant number of is found in Parangtriris Village compared to other villages, namely 133 buildings. Tirtohargo Village does not have buildings exposed to 0-1 m inundation. The same thing also happened to Parangtritis Village at 1-3 m Inundation. This village has almost five times the number of buildings from the second and third highest villages, namely Poncosari and Srigading. In inundation 3-11 m, each village's exposed

buildings differ considerably. Parangtritis is the village with the highest number of total buildings in the three inundation classes, 779 buildings, and the Tirtohargo Village is the lowest village, with nine buildings.

Analysis Result of Exposed Settlements

Exposed settlements were analyzed to find specific residential buildings (houses) that were affected by the tsunami inundation. This study focuses on buildings that humans use to stay at night. The scenario used is the worst-case scenario for a tsunami when humans are resting or sleeping in residential buildings. The analysis is carried out by means of a field survey and also analyzes the minimum area of a house that is livable. Based on the Decree of the Minister of Public Works and Public Housing Number 242 of 2020, the floor area of the lowest treaded public house is 21 M^2. The minimum building area analyzed in this study refers to the provisions of the Ministerial Regulation. The building area in the MBF smaller than 21M^2 is considered not a residence used for residence.

The processing results show that, in total, there are 23 types of designation for building use. The number of types of designation for this building is because the flooded area is a tourism area, so various economic activities are carried out. The highest type of designation is houses, with 449 buildings. The second highest is the house and place to eat. Many locals build a place to eat on the front side of their house as their main livelihood. The third highest is the shop. The number of stalls is a natural thing because most people working in the coastal area are traders. Based on the building designation data, an analysis was conducted between residential and non-residential buildings. The designation of buildings included in the residential category is a house, a house and a place to eat, a house and a shop. The determination of the category is based on the function of the settlement, namely the place of residence or place of residence. Building designations other than those three categories are considered non-residential buildings. Based on the separation results, the number of non-residential buildings is 601, or 46%, and residential buildings are 716, or 54%. The percentage of residential buildings is relatively low. The low percentage of residential buildings is normal in tourist areas, especially the coast. In coastal tourism, trade, and service activities are very high, developed by just living permanently. The vulnerability also causes low settlements to high waves and tsunamis in coastal areas. The number of settlements exposed to the tsunami was then analyzed against the inundation model using ArcGIS and a pivot table in Excel to produce an output, as shown in Figure 8.

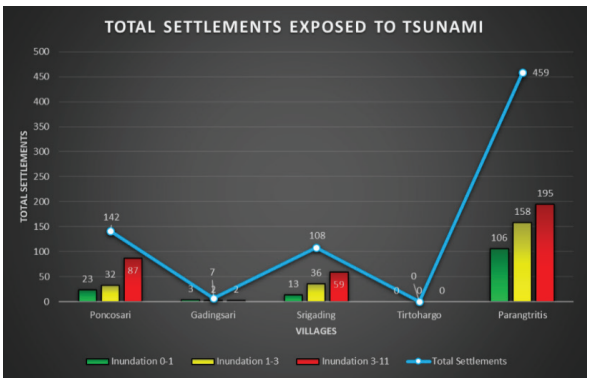

Figure 8.
Number of settlements exposed to the tsunami
Source: Data analysis (2022)

Figure 8. shows each village's number of settlements exposed to the tsunami based on the inundation class. Based on the figure in total, Parangtritis Village has the highest number of 459 residential buildings. The second highest is Poncosari Village, which has 142 residential buildings. Tirtohargo Village is the only village that has yet to have exposed settlements. It is because there are no settlements in that area. The pattern of the number of buildings exposed to 0-1m, 1-3m, and 3-11m inundations always increases in Poncosari, Srigading, and Parangtritis Villages. However, this pattern only occurs in Gadingsari Village. The total number of settlements affected by the tsunami reached 716 buildings. The spatial distribution of settlements is shown in Figure 9.

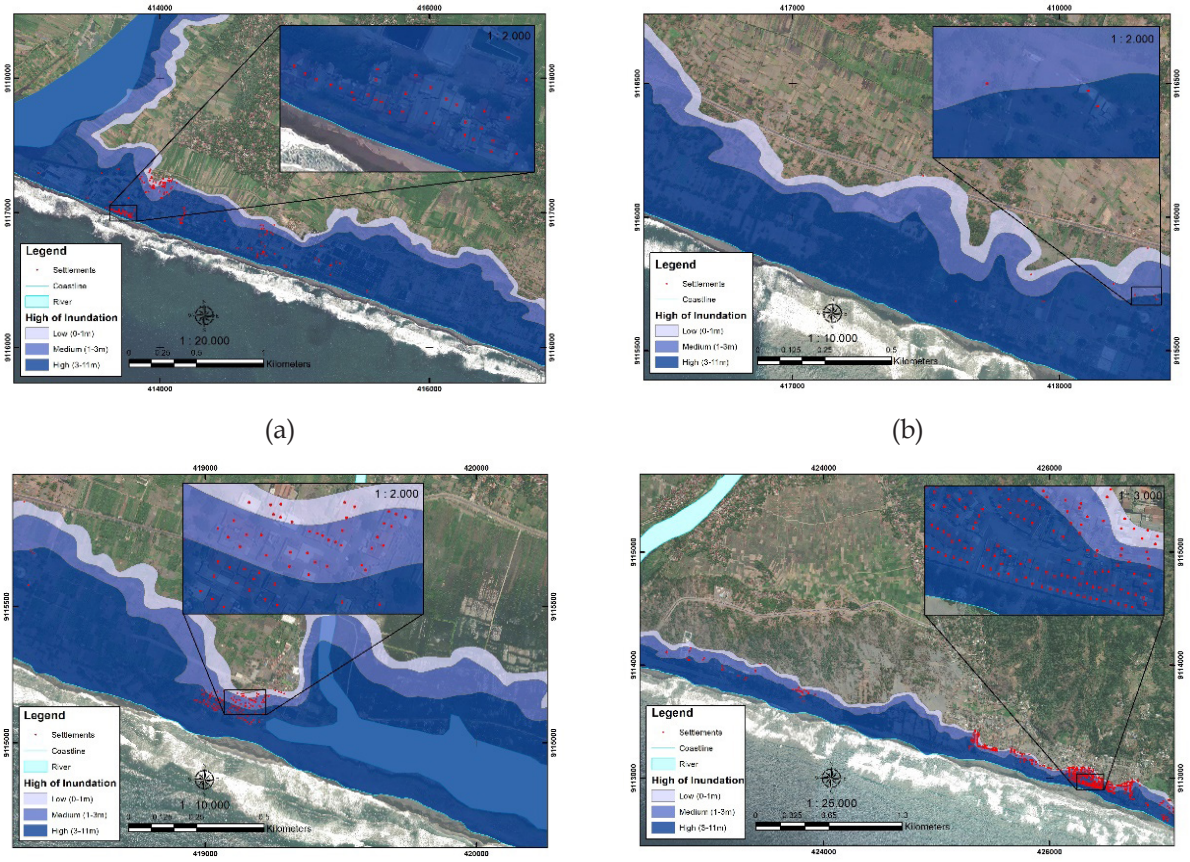

Figure 9.
Distribution of settlements exposed to the tsunami: (a) Poncosari Village; (b)Gadingsari Village; (c) Srigading Village; (d) Parangtritis Village
Source: Data analysis (2022)

Figure 9 (a) shows that the distribution pattern of settlements in Poncosari Village is spread out at several points, and the density between settlements is quite dense. It is different in Gadingsari Village in Figure 9 (b). There are few settlements with very long distances between settlements. In Figure 9 (c)., it can be seen that the distribution of settlements in the Srigading Village gathers at one point with a medium density level. In Figure 9 (d)., the pattern of settlements in Parangtritis Village can be seen that many are clustered on the east with high density, and the rest are spread out on the west with low distribution and density. It can be concluded that the distribution pattern of settlements follows the most famous beach tourism in each village.

Result of Analysis of Estimated Number of Lives Affected by Tsunami.

Analysis of the estimated number of people exposed to the tsunami was carried out in the worst-case scenario, namely at night. The night scenario was chosen because humans spend much time resting at home. The analysis of the distribution of human life is carried out based on the distribution of settlements (houses). To find out the number of people exposed, multiply the number of settlements exposed in the inundation model by the number of residents in each housing unit. The processing results are presented in Figure 10.

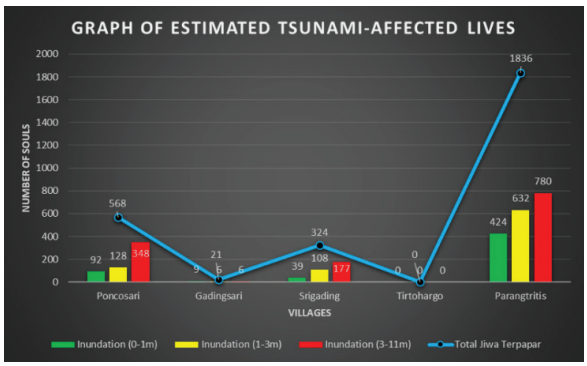

Figure 10.
Graph of estimated tsunami-affected lives
Source: Data analysis (2022)

Figure 10. presents a graph of estimated lives exposed to the tsunami. It can be seen that Parangtritis Village has the highest number of exposed lives, namely 1,836 human lives. The second highest village is Poncosari, with 568 inhabitants, followed by the Srigading village, with 324 inhabitants, and Gadingsari, with 21 inhabitants. There are no exposed lives in Tirtohargo Village because there are no settlements. The total number of people exposed in the research area is 2,749 human lives or 6.32% of the total population in five urban villages.

CONCLUSION

Inundation tsunami modeling using the highest scenario of 11 meters resulted in an inundation area of 747 hectares or about 17.34% of the total area studied. The analysis of the use of MBF to obtain building trace data in this study resulted in a precision value of 99.02%, Recall of 98.40%, and an F1 score of 98.71% in all data. This value is a representation of the reliability of the MBF in detecting buildings. The high value is because the buildings in the research area are in the low to medium range with a medium density average. It is to the advantage of MBF, which is very good at detecting medium density building areas or rural areas. The results of the MBF error analysis are 0.97% false positive, 1.60% false negative, and 0.12% Intersection of Union (IOU). The lack of MBF is that in buildings with high density, there is one polygon that detects several buildings.

Based on the settlements extracted from MBF, which were then analyzed against the tsunami inundation model, it was found that 716 residential buildings were exposed to inundation or about 54% of the total buildings. Parangtritis is the village with the highest number of exposed residential buildings, namely 459 buildings. The high residential buildings are exposed in Parangtritis Village because Parangtritis Beach is a famous and leading tourist destination in Bantul Regency. The total number of exposed human lives reached 2,749 people or 6.32% of the total population in five urban villages. Parangtritis Village is the village with the highest total number of people exposed, namely 1,836 people. The number of people exposed in each village is directly proportional to the number of settlements exposed. In Tirtohargo Village, no people were exposed because there were no exposed settlements in the area affected by the inundation.

Due to the high exposure of tsunamis to residential buildings and people, it is necessary to carry out spatial planning for the development of settlements on the coast. These settlements should be safe and free from exposure to tsunami waves. In addition, the Regional Government (Pemda) and related stakeholders need to periodically carry out early preparedness and evacuation simulations for residents in affected areas. It will help make the community resilient to tsunami hazards.

Acknowledgments

The author thanks Dr. Djati Mardiatno, M.Sc., and Dr. Dyah Rahmawati Hizbaron, M.T., M.Sc., as a supervisor who has provided a lot of knowledge and advice with input in this research. Thanks to Bappenas, which has sponsored master's education.

BIBLIOGRAPHY

Alvianingsih, M., Pradipta, W. I., Hayatiningsih, I., & Hanifa, N. R. 2021. Exposures of Building and Population to Tsunami Hazard in Pangandaran Beach, Indonesia.

IOP Conference Series: Earth and Environmental Science, 925(1). https://doi.org/10.1088/1755-1315/925/1/012037.

Amani, M., Ghorbanian, A., Ahmadi, S. A., Kakooei, M., Moghimi, A., Mirmazloumi, S. M., Moghaddam, S. H. A., Mahdavi, S., Ghahremanloo, M., Parsian, S., Wu, Q., & Brisco, B. 2020. Google Earth Engine Cloud Computing Platform for Remote Sensing Big Data Applications: A Comprehensive Review. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13, 5326–5350. DOI: 10.1109/JSTARS.2020.3021052.

Aris Marfai, M., Fatchurohman, H., & Cahyadi, A. 2019. An Evaluation of Tsunami Hazard Modeling in Gunungkidul Coastal Area using UAV Photogrammetry and GIS. Case Study: Drini Coastal Area. *E3S Web of Conferences*, 125(201 9). DOI: 10.1051/e3sconf/201912509005.

Basith Abdul, Kongko Widjo, O. N. 2016. Pemodelan Spasial Landaan Tsunami Menggunakan Variasi Lokasi Sumber Dan Magnitud Gempa Studi Kasus Kota Padang. *Conference on Geospasial Information Science and Engineering*.

Badan Nasional Penanggulangan Bencana. 2019. *Modul Teknis Penyusunan Kajian Risiko Bencana Tsunami*. Direktorat Pengurangan Risiko Bencana. BNPB Jawa Barat.

Badan Pusat Statistik. 2022. Kabupaten Bantul Dalam Angka. *BPS Kabupaten Bantul*. Accessed on 12 September 2022. Retreived from https://bantulkab.bps.go.id.

Berryman, K. 2006. *Review of Tsunami Hazard and Risk in New Zealand*. New Zealand: Lower Hutt: Institute of Geological & Nuclear Sciences.

Fakhruddin, B., Kintada, K., & Tilley, L. 2021. Probabilistic tsunami hazard and exposure assessment for the pacific islands- Fiji. *International Journal of Disaster Risk Reduction*, 64, 102458. DOI: 10.1016/j.ijdrr.2021.102458.

Fauzi. Y. 2021. Pemodelan Potensi Genangan Tsunami Berbasis Pada Ancaman Gempa Bumi Di Zona Subduksi Selatan Jawa. *Dissertation*, Indonesia: Universitas Gadjah Mada.

Hadi, F., & Astrid, D. (2017). Aplikasi SIG Untuk Pemetaan Zona Keterpaparan Permukiman Terhadap Tsunami Studi Kasus: Kota Pariaman, Sumatera Barat. *Seminar Nasional Geomatika*, 317–324.

Hizbaron, D. R., Baiquni, M., Sartohadi, J., & Rijanta, R. 2012. Urban Vulnerability in Bantul District, Indonesia—Towards Safer and Sustainable Development. *Sustainability*, 1, 2022–2037. DOI: 10.3390/su4092022.

Huang, X., Wang, C., Li, Z., & Ning, H. 2021. A 100 m population grid in the CONUS by disaggregating census data with open-source Microsoft building footprints. *Big Earth Data*, 5(1), 112–133. DOI: 10.1080/20964471.2020.1776200.

Huang, Y., & Jin, Y. 2022. Aerial Imagery-Based Building Footprint Detection with an Integrated Deep Learning Framework: Applications for Fine Scale Wildland–Urban Interface Mapping. *Remote Sensing*, 14(15). DOI: 10.3390/rs14153622.

Irianti, G. L. 2017. Determining Tsunami Evacuation Building Location and Evacuation Routes Based on Population Dynamic And Human Behaviour In Disaster Evacuation In Pacitan Sub-District Area. *Thesis*, Indonesia and Nederland: Double Degree M.Sc. Programme Gadjah Mada University and University of Twente.

Kementerian Pekerjaan Umum dan Perumahan Rakyat. 2020. Keputusan Menteri nomor 242/KPTS/M/2020. *Jaringan Dokumentasi dan Informasi Hukum Kementerian PUPR*. Accessed on 10 September 2022. Retreived from https://jdih.pu.go.id/detail-dokumen/2746/1#div.

Khomarudin, M. R. 2010. Tsunami Risk and Vulnerability: Remote Sensing and GIS Approaches for Surface Roughness Determination, Settlement Mapping and Population Distribution Modelling. *Dissertation*, Germany: Universitat Munchen: Fakultat Fur Geowissenschaften Der Ludwig.

Griffin, J., Latief, H., Kongko, W., Harig, S., Horspool, N., Hanung, R., Rojali, A., Maher, N., Fuchs, A., Hossen, J., Upi, S., Edi Dewanto, S., Rakowsky, N., & Cummins, P. 2015. An evaluation of onshore digital elevation models for modeling tsunami inundation zones. *Frontiers in Earth Science*, 3(June), 1–16. DOI: 10.3389/feart.2015.00032.

Liu, P., Liu, X., Liu, M., Shi, Q., Yang, J., Xu, X., & Zhang, Y. 2019. Building footprint extraction from high-resolution images via spatial residual inception convolutional neural network. *Remote Sensing*, 11(7). DOI: 10.3390/rs11070830.

Liu, T., Yao, L., Qin, J., Lu, N., Jiang, H., Zhang, F., & Zhou, C. 2022. Multi-scale attention integrated hierarchical networks for high-resolution building footprint extraction. *International Journal of Applied Earth Observation and Geoinformation*, 109(February), 102768. DOI: 10.1016/j.jag.2022.102768.

Løvholt, F., Glimsdal, S., Harbitz, C. B., Horspool, N., Smebye, H., de Bono, A., & Nadim, F. 2014. Global tsunami hazard and exposure due to large co-seismic slip. *International Journal of Disaster Risk Reduction*, 10(PB), 406–418. DOI: 10.1016/j.ijdrr.2014.04.003.

Mardiatno, D. 2013. A proposal for tsunami mitigation by using coastal vegetations: Some finding from southern coastal area of Central Java, Indonesia. *Journal of Natural Resources and Development*, 3(7), 85–95.

Mardiatno, D., Sunarto, Wf, L. R., Saptadi, G., & Ayuningtyas, E. A. 2015. Risk mapping and tsunami mitigation in Gunungkidul area, Yogyakarta. *AIP Conference Proceedings*, 1658 (May). DOI: 10.1063/1.4915043.

Mardiatno, D., Malawani, M. N., & Ma, R. 2020. International Journal of Disaster Risk Reduction The future tsunami risk potential as a consequence of building development in Pangandaran Region, West Java, Indonesia. *International Journal of Disaster Risk Reduction*, 46(January), 101523. DOI: 10.1016/j.ijdrr.2020.101523.

Microsoft. 2022. AI Assisted Mapping. *Microsoft*. Accessed on 20 March 2022. Retreived from https://www.microsoft.com/en-us/maps/building-footprints.

Microsoft. 2022. "Microsoft/IdMyPhBuildingFootprints". *Github*. Accessed on 20 March 2022. Retreived from https://github.com/microsoft/IdMyPhBuildingFootprints.

Mohamadi, B., Chen, S., & Liu, J. 2019. Evacuation priority method in tsunami hazard based on DMSP/OLS population mapping in the Pearl River estuary, China. *ISPRS International Journal of Geo-Information*, 8(3), 1–15. DOI: 10.3390/ijgi8030137.

Mutanga, O., & Kumar, L. 2019. Google earth engine applications. *Remote Sensing*, 11(5), 11–14. DOI: 10.3390/rs11050591.

Nurulloh, U. I. 2019. Arahan Tata Guna Lahan Berdasarkan Pemodelan Bahaya Tsunami Dan Dinamika Penggunaan Lahan. *Thesis*, Indonesia: Universitas Gadjah Mada.

Paulik, R., Williams, S., Simi, T., Bosserelle, C., Chan Ting, J., & Simanu, L. 2021. Evaluating building exposure and economic loss changes after the 2009 South Pacific Tsunami. *International Journal of Disaster Risk Reduction*, *56* (February), 102131. https://doi.org/10.1016/j.ijdrr.2021.102131.

Ps, P., & Aithal, B. H. 2022. Building footprint extraction from very high-resolution satellite images using deep learning. *Journal of Spatial Science*, 00(00), 1–17. DOI: 10.1080/14498596.2022.2037473.

Robinson, C., Ortiz, A., Park, H., Gracia, N. L., Kaw, J. K., Sederholm, T., Dodhia, R., & Ferres, J. M. L. 2022. Fast building segmentation from satellite imagery and few local labels. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 2022-June, 1462–1470. DOI: 10.1109/CVPRW56347.2022.00152.

Small, C., & Nicholls, R. J. 2003. A global analysis of human settlement in coastal zones. *Journal of Coastal Research*, 19(3), 584–599. DOI: 10.2307/4299200.

Smart, G. M., Crowley, K. H. M., & Lane, E. M. 2016. Estimating tsunami run-up. *Natural Hazards*, 80(3), 1933–1947. DOI: 10.1007/s11069-015-2052-8.

Touzani, S., & Granderson, J. 2021. Open data and deep semantic segmentation for automated extraction of building footprints. *Remote Sensing*, 13(13). DOI: 10.3390/rs13132578.

Wahyuni, L. 2020 Analisis Risiko Bencana Tsunami Untuk Mendukung Tata Ruang Pesisir Di Kabupaten Kulonprogo. *Dissertation*, Indonesia: Universitas Gadjah Mada.

Widiyantoro, S., Gunawan, E., Muhari, A., Rawlinson, N., Mori, J., Hanifa, N. R., Susilo, S., Supendi, P., Shiddiqi, H. A., Nugraha, A. D., & Putra, H. E. 2020. Implications for megathrust earthquakes and tsunamis from seismic gaps south of Java Indonesia. *Scientific Reports*, 10(1), 1–11. DOI: 10.1038/s41598-020-72142-z.