International Journal on Advanced Science Engineering Information Technology

Vol.13 (2023) No. 5
ISSN: 2088-5334

Tsunami Database Development in the Sunda Arc Indonesia to Support Early Warning through Artificial Intelligence Technology

Mardi Wibowo[a*], Wahyu Hendriyono[a], Hanah Khoirunnisa[a], Reno Arief Rachman[a], Widjo Kongko[a] Gugum Gumbira[a], Shofia Karima[b], Reni Wijayanti[c], Eko Kustiyanto[a], Amalia Nurwijayanti[a], Destianingrum Ratna Prabawardani[a], Gumilang Ramadhan Pasma[a]

[a]*Research Center for Hydrodynamic Technology, National Research and Innovation Agency (BRIN), Yogyakarta, 55284, Indonesia*
[b]*Research Center for Geological Disaster, National Research and Innovation Agency (BRIN), Bandung 40135, Indonesia*
[c]*Research Center for Oceanography, National Research and Innovation Agency (BRIN), Jakarta 14430*
[*]*Corresponding author: mardi.wibowo@brin.go.id*

Abstract*--- The Sunda Arc-Indonesia is very vulnerable to tsunamis. There have been at least 55 tsunamis from 416–2018. Tsunami in the Sunda Arc is classified as a near-field tsunami with an arrival time of < 30 minutes after the earthquake. Meanwhile, the BMKG issued a warning within 5 minutes after the earthquake; therefore, speed in giving warnings is very vital. Artificial intelligence is an alternative technology that can quickly predict a tsunami's height and arrival time. For developing this technology, adequate quality and quantity of data and information on tsunamis are needed. Therefore, this study was conducted to build a tsunami database based on the results of simulations and numerical modeling of multiple scenarios from hypothetical and historical earthquake sources. This study used the open-source TUNAMI F1 model. This model simulates the propagation of tsunami waves using a linear equation. This study obtained 465 hypothetical earthquake sources, 534 historical earthquake sources, and 9,990 datasets from tsunami model simulation results. Each dataset contains ten information. Based on the 8.2 magnitude earthquake scenario, the potential tsunami hazard is 3–47 m with an estimated arrival time of < 30 minutes. An earthquake < 7 Mw can trigger a tsunami, especially an earthquake that is shallow and close to the coast, even though the tsunami height is < 0.5m. This data will be used to train an artificial intelligence-based tsunami prediction system. The artificial intelligence-based tsunami prediction system is expected to be used to strengthen the Indonesia tsunami early warning system (InaTEWS).

***Keywords*--- Artificial intelligence; InaTEWS; near-field tsunami; Sunda Arc; tsunami database; tsunami modeling**

Manuscript received 27 Jan. 2023; revised 2 Feb. 2023; accepted 13 May 2023. Date of publication 31 Oct. 2023.
IJASEIT is licensed under a Creative Commons Attribution-Share Alike 4.0 International License.

I. INTRODUCTION

The Sunda Arc is a zone where the Indo-Australian Plate is subducted under the Eurasian Plate, which extends from the Andaman, Sumatra, Java, and Nusa Tenggara to the Banda Sea (Figure 1). The movement of the two plates is part of the movement of the Indian Plate, which collides with the Asian Plate on the west side, and the movement of the Australian Plate, which collides with the Pacific Plate on the east side [1]. The active Sunda-Banda Arc stretches from the Andaman Islands to the Banda Sea for 6000 km. This arc was formed by the Indo-Australian Plate subducting under the Eurasian Plate [2]. The age of active subduction in the Sunda Arc is late Eocene or early Oligocene, which occurred after the collision between India and Asia collision about 45 Ma ago [3].

According to data from the Agency for Meteorology, Climatology, and Geophysics of Indonesia (BMKG), there have been at least 55 tsunamis in the Sunda Arc area from 416–2018, large and small [4]. In the Sunda Arc area, the areas most frequently hit by tsunamis are the west coast of Sumatra and East Nusa Tenggara [4]. A very large tsunami incident in the Sunda Arc and still remembered today is the Aceh tsunami of December 26, 2004, which claimed the lives of more than 227 thousand people [4]. Then the Nias's tsunami on March 28, 2005 (the death toll is around 1,000 people), the tsunami on July 17 at Pangandaran, 2006 (664 people), and the tsunami in Banyuwangi on June 2, 1994, which claimed the lives of more than 200 people) [4]. From 2004 until now, in the Sunda Arc area, there have been at least 11 tsunamis [5]. The last tsunami in this area occurred on December 25, 2018, due to a partial landslide of the body of Mount Krakatau during high waves [6], [7].

Meanwhile, based on historical data since 1900, in this study area, there have been at least 45 tsunamis [8] (Fig.1). The Sunda megathrust has generated numerous tsunamigenic earthquakes: in the past 120 years, 12 earthquakes along this boundary caused wave heights greater than 1 m [9]. Historical data on tsunamis in Indonesia can be retrieved and studied from world tsunami databases [10]–[12]. In addition to historical data, there is the potential for a tsunami that may occur in the future. South Java has the potential for a large earthquake, which may cause a tsunami with a height of up to 20 m in the south of Java [13].

Fig. 1 Tsunami source map (WinITDB with processing) 1800 - present, numbers indicate the year of occurrence [8]

Following the mandate of the Presidential Regulation of the Republic of Indonesia No. 93 of 2019 related to Tsunami Early Warning (TEW) in Indonesia [14], the National Research and Innovation Agency (BRIN) began in 2020 to develop a TEWS including the west Sumatra, south of Java and south of Bali. One of the supporting activities is developing a tsunami hazard prediction system that can estimate tsunamis' height and arrival time through numerical modeling and artificial intelligence. By using Near Real-Time (NRT) modeling and artificial intelligence tsunami prediction, it is hoped that when an earthquake with a potential tsunami occurs, an early prediction of tsunami height and arrival time can be made as part of a tsunami early warning. It is known that most tsunamis in Indonesia are classified as 'near filed tsunamis', where tsunamis generally arrive less than 30 minutes after the earthquake. Hence, data collection, processing, modeling, and analysis speed are vital [15].

For developing a tsunami prediction with artificial intelligence, it is very necessary to have sufficient tsunami data and information. However, historical tsunami data and information are very limited to training the intelligence system, while a large amount of tsunami data is needed. Therefore, this activity was carried out to develop a tsunami database based on simulations and numerical modeling of multiple scenarios from hypothetical and historical earthquake sources as input for tsunami prediction AI training.

The advantages of artificial intelligence technology for tsunami prediction have been widely stated, especially related to the tsunamis in Aceh in 2004 and Tohuku in 2011. The mean square error of applying Artificial Neural Networks (ANN) for predicting the arrival time of tsunamis in the Indian Ocean was 0.25 [16], [17]. The ANN forecasting model only takes a few seconds to provide data with accuracy similar to a typical tsunami propagation model, such as TUNAMI-N2, which requires 10 minutes of Central Processing Unit (CPU) time on a standard desktop PC [18], [19]. The correlation coefficients between the ANN prediction results and tsunami simulation results at Osaka Bay are 0.94 - 0.99. This indicates that this method is very reliable [20]. Rodriguez et al. [21] also used neural networks to predict the tsunami's maximum height & arrival time in the Spanish TEWS context.

Application of a deep learning algorithm for making tsunami inundation databases could reduce ± 90% of real-time computational effort [22]. Utilization of convolutional neural networks (CNNs) for TEW in Tohuku-Japan, verifying the feasibility of AI-supported tsunami forecasting to provide accurate and fast early warning [23]. Irfiani et al. classified tsunami potential in Indonesia based on earthquakes using the C4.5 algorithm [24]. Another AI algorithm, Bayesian Neural Networks (BNN), agrees with the numerical model for maximum tsunami heights over historical tsunami events confirmed in testing. BNNs could reasonably predict the maximum tsunami heights for virtual tsunamis [25].

The development of tsunami databases has been carried out for various purposes. Lin et al. [26] used equations of 1D nonlinear shallow water through a constant and uniform slope (analytical Green's function) to build a database of run-up. Setiyono et al. [27] built a pre-computed tsunami inundation database in Pelabuhan Ratu, where this method can shorten the time from 40 minutes to one minute. Harig et al. [5] describes the evolution of a pre-computed database scenario for the InaTEWS, especially in the tsunami modeling component. Ibtihaj et al. [28] built a WebGIS paleo tsunami database in Indonesia (there are 302 tsunami data) from the Late Miocene to the last tsunami in 2018.

BMKG has also developed a tsunami database with various scenarios. This database was compiled based on nonlinear shallow water theory, and the faults discretization as an input tsunami modeling was made by Geoscience Australia/DMInnovation [29]. The calculation of tsunami scenarios is based on the AWI-developed unstructured finite-element mesh numerical model of TsunAWI. Nowadays, InaTEWS uses the Tsunami Observation and Simulation

Terminal (TOAST) decision support system. The TOAST system provides precalculated tsunamigenic earthquakes to inform within TEW-1 and on-the-fly simulation of tsunami propagation for TEW-2 [30], [31]. With this system, BMKG can give warnings within 5 minutes after an earthquake[31].

Although using a different model, this study's results can enrich the BMKG's database. This study generated the tsunami database from the TUNAMI F1 model with historical and hypothetical sources of tsunami-generating earthquakes. The combination of historical and hypothetical data aims to obtain as much data as possible with the distribution of earthquake sources as evenly as possible.

II. Material and Method

A. Area of Study

The study area stretches from the west (in Aceh) to the east (in Wetar Island). In order to make the model run faster, the Sunda Arc area is divided into six regions (Fig. 3). The division of the region also considers the time of the tsunami wave movement from the earthquake source to the coastline, such that the tsunami arrival time is less than two hours. There is a large overlap area between adjacent regions so that any earthquake source can be located relatively far enough from the domain edges to avoid instability (*blow-up*) in the model simulation process.

B. Steps of Study

This study was carried out through a desk study by collecting secondary data on the source of an earthquake-causing tsunami and a tsunami model simulation with the open-source TUNAMI F1 model from Tohuku University, Japan. In general, the stages of implementing this activity are (Fig. 2):

- Develop a hypothetical earthquake source scenario.
- Compile historical earthquake data with complete parameters, including the mechanism of fault movement and moment tensor.
- Build a modeling domain based on national bathymetry data [32].
- Determine the observation locations on the shoreline.
- Generate an initial tsunami wave based on earthquake parameters using a multi-deform model.
- Build a multi-scenario tsunami modeling application to speed up the multi-scenario tsunami modeling process.
- Modeling tsunami propagation with the TUNAMI F1 model
- Build a dataset of the relationship between earthquake source parameters with arrival time and tsunami wave height at each observation location, which will be used to train an artificial intelligence system for tsunami prediction.

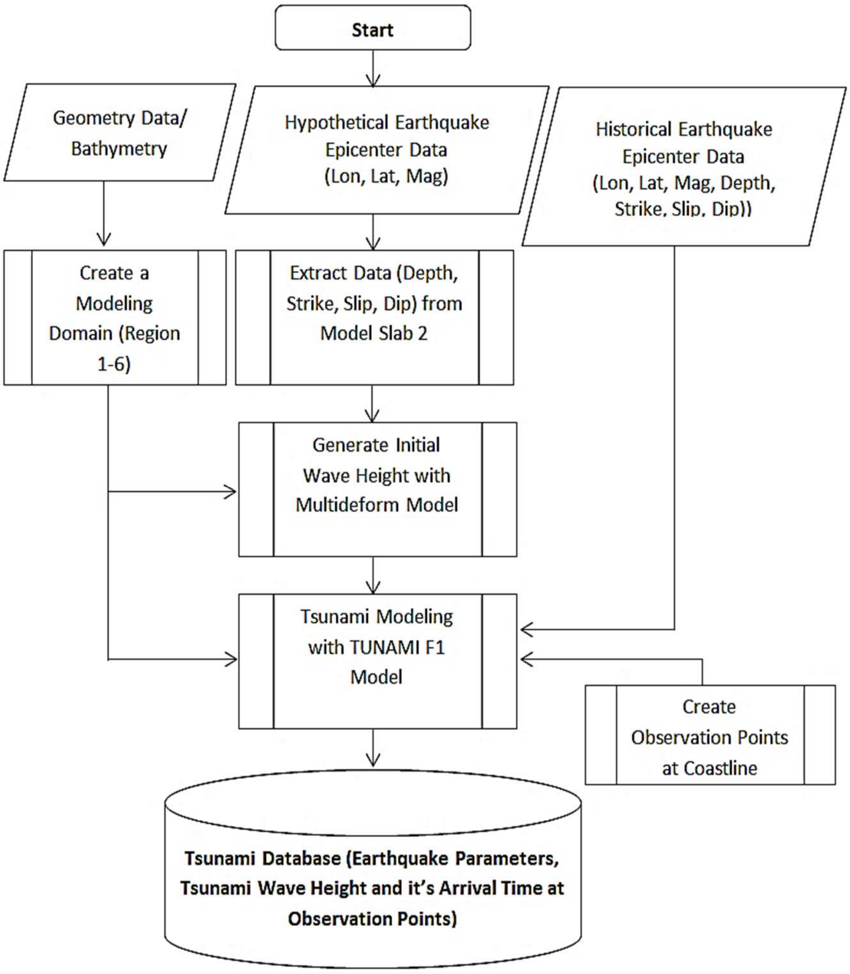


Fig. 2 Design of the study

C. Tsunami Propagation Modeling

The most important step in this study is tsunami propagation modeling from the source of the earthquake to the shoreline. This study used TUNAMI F1 model. This model simulates wave propagation in spherical coordinates and structured mesh (grid), based on linear equations and ignores nonlinear terms such as bed roughness [33]. A tsunami expert developed this model from Tohoku University, namely Imamura.

The TUNAMI model has been very well validated with data from the Pangandaran tsunami 2006. TUNAMI model has been widely used to simulate several tsunami events in the world with good results [17], [34], [35]. The results of the TUNAMI modeling in the Pangandaran tsunami event are in line to previous research and the results of the BMKG's survey. Between the model and survey results, there is a difference in wave height of around 0.98m (20.74%) [34], [36]. Based on the 2004 Sumatera-Andaman tsunami modeling, it's known that TUNAMI and other tsunami models (such as MOST, COMCOT, TUNA) and ANN Tsunami Predict models provide quite good results [36]. Adriano et al. used TUNAMI model to simulate Sendai's tsunami on November 22, 2016, based on a model with two earthquake sources. The validation of this model results on the measurement data produces an NRMSE of 0.686-0.863 [37].

III. Results and Discussion

A. Earthquake Source Data

The hypothetical earthquake sources generating tsunamis are composed of 465 and are each about 50 km apart and evenly distributed throughout the Indian Ocean (Figure 3a). This hypothetical earthquake's depth, strike, dip and slip data were extracted from Model Slab 2. The Slab2 model calculated all active global subduction zones based on three-dimensional geometry. This model describes a uniform geometric analysis of all subducting plates [38].

Meanwhile, complete historical earthquake data with depth, dip, slip, and strike data were taken from the Global Centroid Moment Tensor (GCMT) [39], [40]. Since 1976, Sunda Arc has experienced at least 534 earthquakes [39], [40] (Fig. 3b). Historical data taken is an earthquake of magnitude more than 5.5 Mw to anticipate the possibility of an earthquake with a greater magnitude in that location.

Hypothetical and historical earthquake data will complement each other because historical earthquake data are not evenly distributed. Historical earthquake data shows several areas have never had an earthquake (seismic gap). The seismic gap areas include south of Timor Island, Bali, south of East Java, Sunda Strait, and Mentawai Island [41]. Based on the inversion global positioning system (GPS) data, it's clearly revealed that there is a seismic gap in the south of Java Island and the southwest coast of Sumatera. In the future, this seismic gap has the potential to become a source of megathrust earthquakes [13]. Whereas at that location, there is a possibility that energy accumulation is occurring. The accumulated energy will be released at a time as a strong earthquake [13], [41].

In the model simulation, historically and hypothetically, the earthquake magnitude was varied by ten values for each earthquake epicenter, namely Mw 6.4; 6,6; 6.8; 7.0; 7.2; 7.4; 7.6; 7.8; 8.0 and 8.2. The variation in the magnitude of Mw 6.4 is intended so that in the training of the AI model later, the system can automatically interpolate if a magnitude has never existed in the database. Meanwhile, BMKG has modeled an earthquake with a magnitude 7.2--9 Mw and interval 2 [29]. When an earthquake occurs, BMKG will immediately retrieve data from the existing database based on the magnitude and epicenter closest to the earthquake that occurred [30].

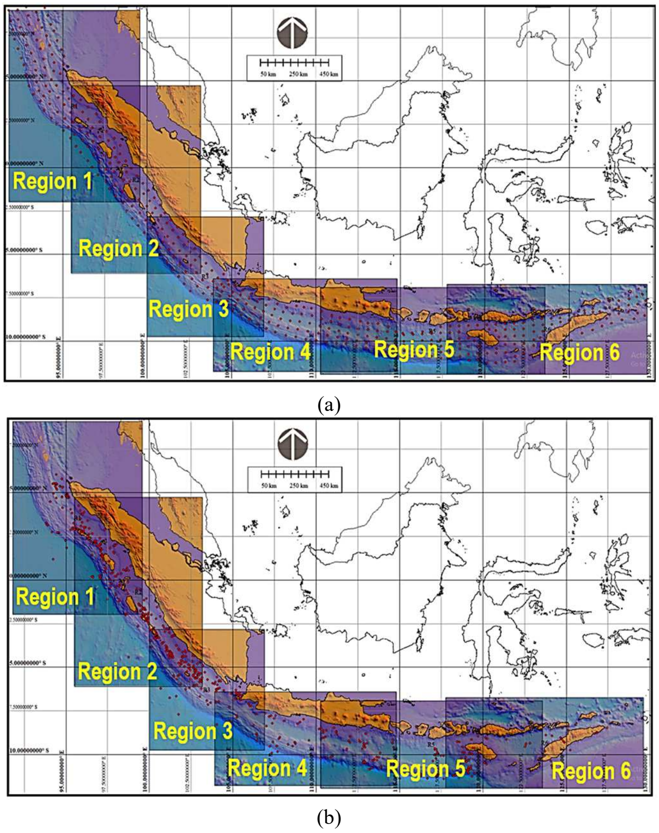


Fig. 3 Earthquake epicenter for tsunami modeling input (a) hypothetical (b) historical data

B. Observation Points

Observation locations on the coastline were made with an average distance of 15 km (Fig. 4). For complex coastlines, the observation locations can be closer, such as the coastline in a bay. The total number of observation locations is 477 points (Fig. 3). Based on the tsunami modeling, each observation point will obtain data on the tsunami's height and arrival time for each model scenario (Fig. 8a and 8b).

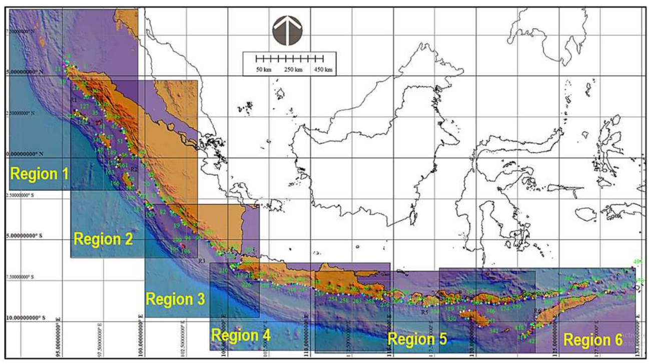


Fig. 4 Observation locations on the shoreline

C. Initial Tsunami Wave

Based on the existing earthquake parameters, the initial tsunami wave height data is generated at the earthquake source using a multi-deform model [42]. The *'multideform model'* was developed based on displacement data on inclined faults [43]. Tsunami modeling requires an input of the initial tsunami wave height. The initial tsunami wave height is calculated based on the parameters of the fault (length, width, slip angle, strike, and dip), the earthquake magnitude, and the depth of the epicenter. This earthquake generated an initial tsunami wave height, propagating in all directions [44].

D. Multi-Scenario Modeling Application

A simple multi-scenario modeling application was built to speed up and automate the building database of tsunami simulation results [45]. This application was built using MATLAB to not model per scenario (Fig. 5).

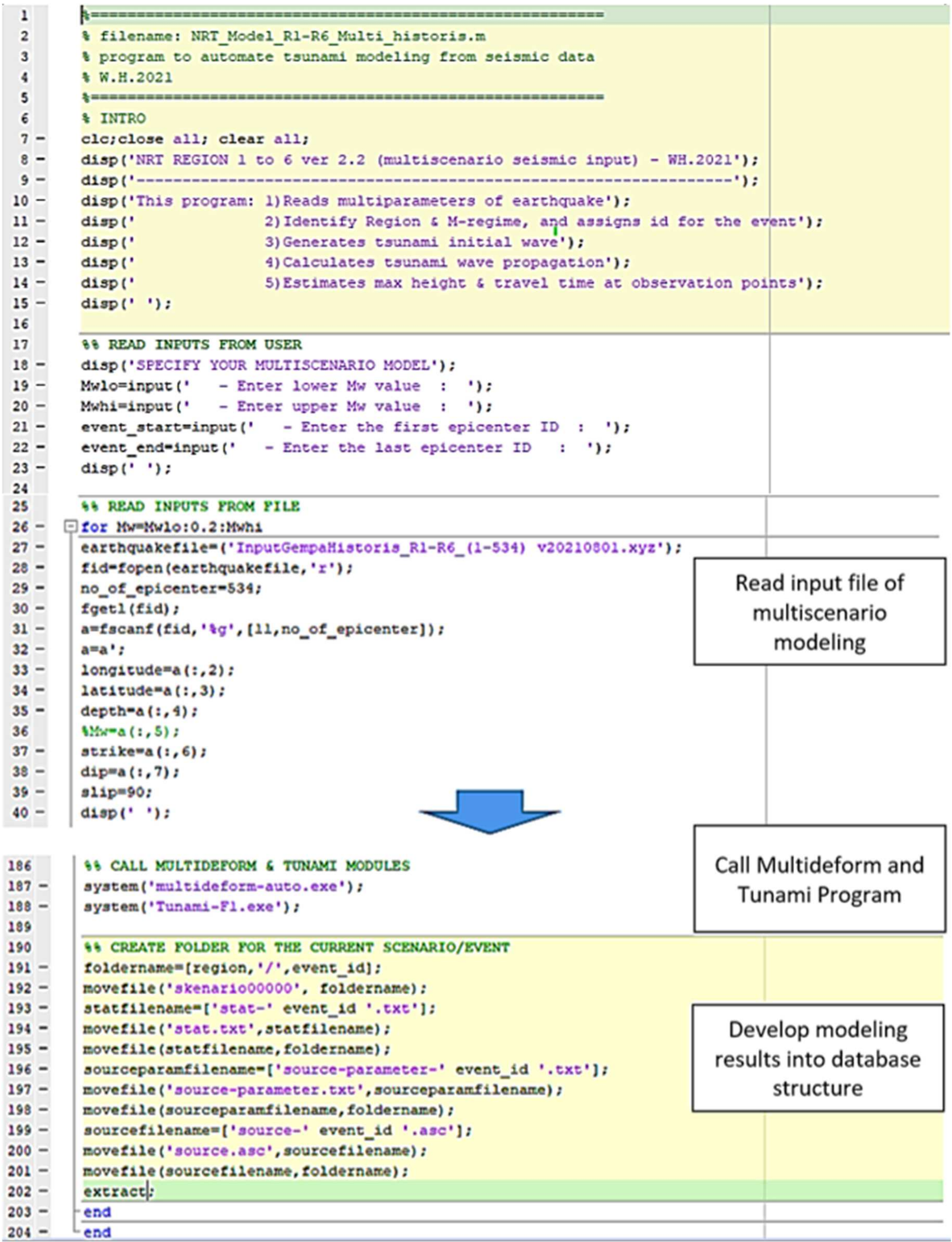


Fig. 5 MATLAB script snapshot for multi-scenario tsunami modeling

For a hypothetical earthquake scenario, this application reads the magnitude and coordinates of the earthquake and then uses the SLAB model to obtain all other earthquake parameters [46]. As for historical earthquake scenarios, this application automates the process, starting from reading earthquake parameters (longitude, latitude, depth, magnitude, strike, slip, and dip) from a single file that has been compiled containing hundreds of tsunami-generating earthquake scenarios.

Furthermore, based on the location of the earthquake's epicenter, both hypothetical and historical scenarios, this application identifies the location and inserts it into the right region to be modeled with the appropriate domain. The next step is to combine the *'scaling law model'*, the *'multideform model*', and the TUNAMI F1 model. The *'scaling law model'* calculates the fault dimensions and the magnitude of the dislocation based on the depth of the epicenter and the earthquake magnitude [47]. The *'multideform model'* generates the initial tsunami wave at the epicenter). TUNAMI F1 models the propagation of tsunami waves from the source to the coast. An additional script is written to calculate the arrival time and the maximum tsunami wave at observation points along the coastline.

E. Multi-scenarios Tsunami Modeling Database

The results of multi-scenario tsunami modeling, with historical and hypothetical earthquake inputs, are stored in a database of model results. This database contains thousands of folders, each folder containing the result of tsunami modeling from one scenario (Fig. 6). Inside each of these folders contains the source parameters, the Sea Surface Height (SSH) time-series matrix from generation to the end of the simulation, the maximum SSH matrix ever recorded at all points during the simulation, the wave arrival time matrix at all points, as well as a table of extracts of SSH results and time of arrival at the observation points along the coast (Fig. 7). In this study, the wave height threshold that is considered a tsunami is 0.1m. So Estimated Time of Arrival (ETA) is the arrival time of a 0.1m wave at the coastline. Based on this study, it turns out that an earthquake with a magnitude < 7 Mw can trigger a tsunami, especially an earthquake that is shallow and close to the coast, even though the tsunami height is < 0.5m.

Each folder representing each scenario is named based on the id-code 'magnitude-epicenter number'. This name was used to identify the results of multi-scenario modeling, which will later be used as the AI-tsunami dataset. The results of the multi-scenario modeling are stored in an earthquake-tsunami database. This dataset is created to construct AI-tsunami. Due to the very large area of the Sunda Arc, to speed up the tsunami model simulation process as well as more realistic results, especially regarding the arrival time of the tsunami, the Sunda Arc area is divided into six regions. In addition, the segregation of existing earthquake source data is also carried out by eliminating some earthquakes with epicenters on land or with a depth of more than 70 km.

After sorting the source data and existing earthquake parameters based on historical data for region one, there are 115 earthquake sources. For region two, there are 130 earthquake sources, region three has 104 earthquake sources, region four has 87 earthquake sources, region five has 83 earthquake sources, and region six has 15 earthquake sources. So overall, in the Sunda Arc area, there are about 534 earthquake epicenters.

Each earthquake's epicenter will be modeled with ten variations in magnitude depending on the maximum potential of the existing earthquake magnitude, namely Mw 6.4; 6.6; 6.8; 7.0; 7.2; 7.4; 7.6; 7.8; 8.0 and Mw 8.2. So based on this historical data, 5,340 scenarios of tsunami modeling will be simulated. Meanwhile, based on the hypothetical earthquake source scenario, there are 465 earthquake sources. This hypothetical earthquake source is also modeled with ten variations of earthquake magnitude so that 4,650 data will be obtained. BMKG 2017 added a tsunami database of 2,570 scenarios in the Sunda Arc based on earthquake magnitudes of 7.0-9.0 [29] so that the data from this study can complement the BMKG data, especially for earthquakes with a magnitude < 7.0.

The problem that sometimes arises in the modeling process is when the running of a model scenario is not completed until the end (*blow-up*). *Blow-up* often occurs at earthquake sources close to the coastline and of large magnitudes. Running the model is repeated by reducing the time step (dt) to obtain the required courant number so running does not blow up [48].

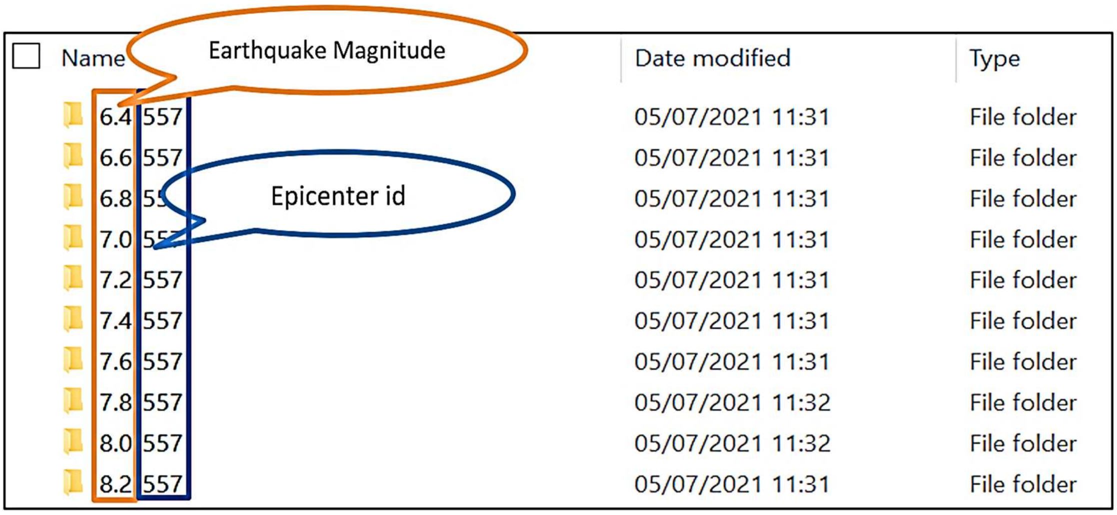
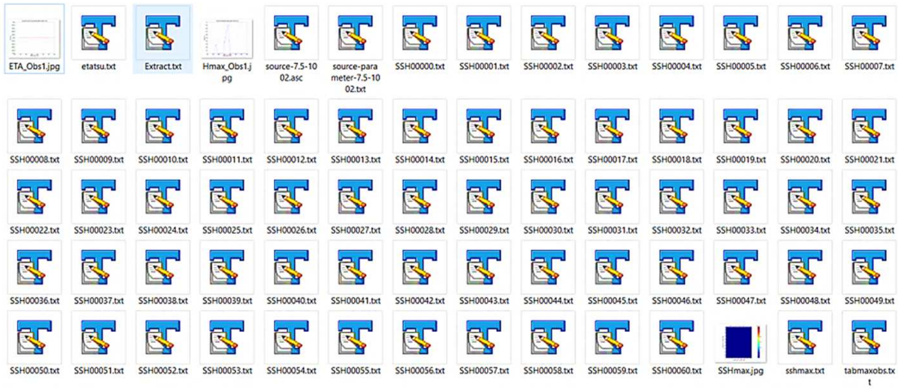


Fig. 6 Folder names in the database, based on magnitude and epicenter



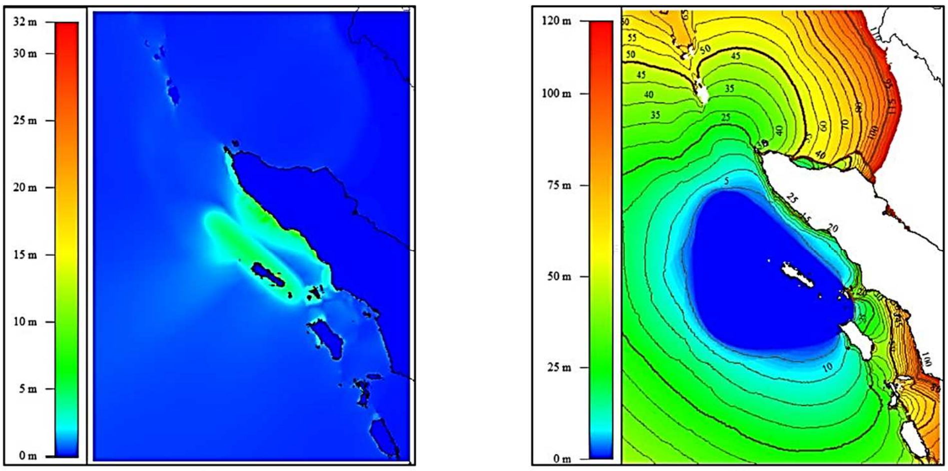
whereas

- ETA_Obs1.jpg: graphic of tsunami arrival time at the observation point
- etatsu.asc: tsunami arrival time matrix for each place
- Extract.txt: extract the height and arrival time of the tsunami waves at each observation point
- Hmax_Obs1.jpg: graphic of maximum tsunami height at the observation point
- Source-7.4-557.asc: initial tsunami wave for magnitude 7.4 dan epicenter 557
- Source-parameter-7.4-557.txt: tsunami generating earthquake source parameter values
- SSH00000.txt: tsunami height matrix at time step 0
- SSH00060.txt: tsunami height matrix at time step 60
- SSHmax.jpg: maximum tsunami height image across domains
- sshmax.txt: maximum tsunami height matrix across domains
- tabmaxobs.txt: table of maximum tsunami heights at observation points

Fig. 7 Files contained in a database folder

The following is an example of modeling results from an earthquake source location in the case of the December 2004 Aceh tsunami (epicenter at coordinates: 95.982 East Longitude and 3.295 North Latitude with a magnitude of Mw 8.8; depth 37.7 km; strike 312°; dip 15.6°; slip 90°.

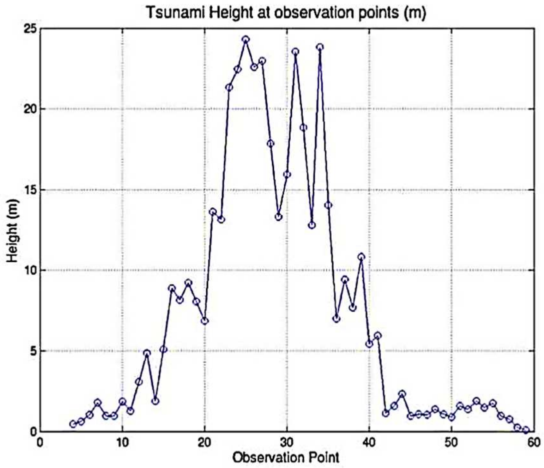
1) Tsunami Maximum Wave Height and Arrival Time: Based on Figure 8a, with the parameters mentioned above, it can be seen that most of the tsunami wave heights on the coast near the earthquake source reached 10 m, and the maximum tsunami height was around 30 m, especially on the bay on the island near the tsunami source (Fig. 8b). In the bay, the tsunami height will generally be higher than on the normal coast [49], [50]. These results are similar to the research by Rasyif et al. [51]. Other researchers stated that the maximum calculated tsunami height at Banda Aceh was about 5 m, and the first arrival time was ± 25 minutes after the earthquake [52]. The difference in tsunami height is caused by differences in the parameters of the earthquake generating the tsunami, and the bathymetry data was used. The tsunami propagation time is the time it takes for the first tsunami wave to move from the source to the observation point. Based on this model, the tsunami arrival time on the coastline was very fast. The tsunami directly hit some locations shortly after the earthquake.



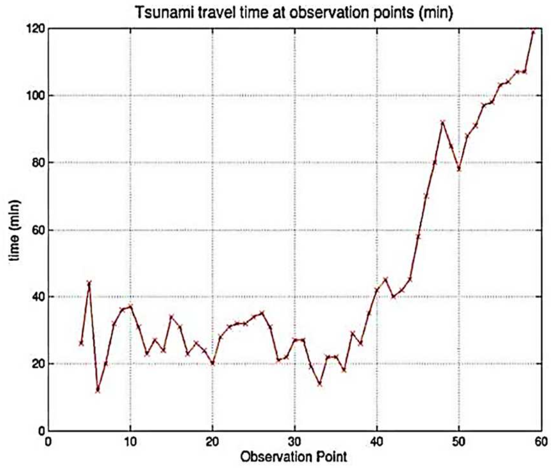
(a) (b)

Fig. 8 Example of model results in region 1: (a) maximum tsunami height (b) tsunami propagation time

2) Tsunami Wave Height and Arrival Time at Observation Points: Model results were extracted at designated observation points along the coastline to find out more precisely the tsunami wave height along the coastline. The results of the extraction of tsunami wave heights at observation points along the coastline are shown in Fig. 9a. In contrast, the tsunami wave's arrival time along the observation points on the coastline is shown in Fig. 9b.



(a)



(b)

Fig. 9 Extracted results at the observation point of region 1 (a) maximum tsunami wave height (b) tsunami arrival time

F. Potential Tsunami Threat along Sunda Arc

Multi-scenario tsunami modeling has been carried out along the Sunda Arc, both with historical earthquake sources (4,650 scenarios) and hypothetical (5,340 scenarios). Analysis of the model results can provide a glimpse of the distribution of the potential tsunami threat along the western coasts of Sumatra, southern Java, Bali, and Nusa Tenggara, as shown in the following figures.

Fig. 10a shows the distribution of potential tsunami threats along the Sunda Arc caused by earthquakes along the Sunda Arc subduction zone, with the same location (epicenter) and depth as those that have occurred in history (historical), assuming a maximum magnitude of 8.2 Mw. In this historical case estimate, the potential threat ranges from three to 35 meters, with an estimated arrival time between zero and 30 minutes. The areas with the greatest potential threat are the west coast of Aceh, West Sumatra, Bengkulu, the south coast of Banten, East Java, and East Nusa Tenggara.

Meanwhile, Fig. 10b shows the distribution of potential tsunami threats along the Sunda Arc caused by earthquakes along the Sunda Arc subduction zone, with a hypothetical epicenter distribution and depth according to the Sunda Arc slab model, assuming a maximum magnitude of 8.2 Mw. In the hypothetical case (Fig. 10b), assuming the epicenter is evenly distributed throughout the SLAB model area (which includes historical and unprecedented earthquake points), the potential threat ranges from 8–47 m, with an estimated arrival time of 0–28 minutes. The greatest potential threats are in locations almost the same as in historical cases. In both historical and hypothetical cases, the amplification of tsunami height occurred rapidly at these locations, possibly due to a bay's shoreline profile and the bathymetry changing the profile from the deep sea to the coast.

G. Development of Tsunami Early Warning based on Artificial Intelligence

Next time, the above data will be used to build a TEWS with artificial intelligence technology. The artificial intelligence algorithm used is an ANN. ANN toolbox is available in MATLAB. This program also reads the earthquake data as inputs and gives predicted heights and arrival times as the outputs. However, ANN does not contain modules to calculate fault displacement, initial wave, propagation, etc. Instead, this program consists of algorithms or formulae that can directly relate input parameters to corresponding outputs due to the training/ learning process using thousands of earthquake/tsunami model scenarios. This enables ANN to run much faster than the numerical modeling in an earthquake/tsunami and provides a preliminary hazard estimate. At the same time, more accurate values are still calculated by the model and available a few minutes later. The preliminary application of the ANN in generating predictions of height and travel time can predict high values and travel time of tsunami waves south of Java with an R2 value of 0.9948 for wave height and an R2 value of 0.9846 for travel time [17].

(a) historical earthquakes assuming a maximum magnitude of M8.2

(b) hypothetical earthquakes assuming a maximum magnitude of M8.2

Fig. 10 Potential tsunami threat along the Sunda Arc (tsunami height in meters)

IV. Conclusion

Based on this study, we obtained 9,990 datasets from tsunami model simulation results from 999 earthquake epicenters. Each dataset contains ten information. The main ones are maximum tsunami height and tsunami arrival time at the coastline. If an earthquake with magnitude 8.2 occurred, the potential tsunami threat ranges from 3–47 m with an estimated arrival time of < 30 minutes. Results from the multi-scenario modeling will be used to train an artificial intelligence-based tsunami prediction system and are expected to strengthen the InaTEWS. In addition, the database from this study is expected to enrich the BMKG database. Future development potential is automating database updates from the latest modeling results when a new earthquake occurs. It is also necessary to develop a tsunami database base on modeling for other regions in Indonesia, especially in eastern Indonesia.

Acknowledgment

We thank the management of ex-BTIPDP-BPPT, PRTH-BRIN, and all researchers in the Tsunami Modeling and Data Analysis for RDS System 2021-2022 activities.

Authors Contributions

In writing this article, MW and WH are the main authors. MW composed the manuscript, provided epicenter and earthquake parameters data, and supervised the modeling process; WH developed applications for multi-scenario tsunami modeling and ran tsunami models; RAR compiled earthquake data parameters; WK supervised this study; HK, GG, SK, and RW ran the tsunami model. All authors reviewed this manuscript. We declare no conflict of interest in preparing and publishing this article. We also ensure that this data and article are free from plagiarism.

References

[1] H. Harjono *et al.*, *Seismotektonik busur Sunda*, 1st ed. Jakarta, Indonesia: LIP Press, 2017.

[2] M. Kirchenbaur *et al.*, “Sub-arc mantle enrichment in the Sunda rear-arc inferred from HFSE systematics in high-K lavas from Java,” *Contributions to Mineralogy and Petrology*, vol. 177, no. 1, pp. 1–25, 2022.

[3] R. Hall, M. W. A. van Hattum, and W. Spakman, “Impact of India-Asia collision on SE Asia: The record in Borneo,” *Tectonophysics*, vol. 451, no. 1–4, pp. 366–389, Apr. 2008.

[4] R. Triyono *et al.*, *Katalog Tsunami Indonesia Per-Wilayah Tahun 416-2018*, 2nd ed. Jakarta, Indonesia: Agency for Meteorology, Climatology and Geophysics (BMKG), 2019.

[5] S. Harig *et al.*, “The Tsunami Scenario Database of the Indonesia Tsunami Early Warning System (InaTEWS): Evolution of the Coverage and the Involved Modeling Approaches,” *Pure and Applied Geophysics*, vol. 177, no. 3, pp. 1379–1401, 2020.

[6] A. Paris, P. Heinrich, R. Paris, C. Guerin, H. Hebert, and A. Gailler, “Numerical modeling of the December 22, 2018 Anak Krakatau landslide and the following tsunami in Sunda Strait, Indonesia,” OCEANS 2019 - Marseille, Jun. 2019, doi: 10.1109/oceanse.2019.8867270.

[7] J. C. Borrero *et al.*, “Field Survey and Numerical Modelling of the December 22, 2018 Anak Krakatau Tsunami,” *Pure and Applied Geophysics*, vol. 177, no. 6, pp. 2457–2475, Jun. 2020.

[8] WinITDB, “Integrated Tsunami Database for the World Ocean,” 2020. [Online]. Available: http://tsun.sscc.ru/WinITDB.htm. [Accessed: 01-Jul-2020].

[9] J. A. Reid and W. D. Mooney, “Tsunami Occurrence 1900–2020: A Global Review, with Examples from Indonesia,” *Pure and Applied Geophysics*, vol. May, p. 23, 2022.

[10] J. Goff, “Tsunami databases,” *Geological Records of Tsunamis and Other Extreme Waves*, pp. 75–93, Jan. 2020.

[11] C. Liu *et al.*, “Numerical Investigation of Sediment Transport of Sandy Beaches by a Tsunami-Like Solitary Wave Based on Navier-Stokes Equations,” *Journal of Offshore Mechanics and Arctic Engineering*, vol. 141, no. 6, Dec. 2019.

[12] T. Nacházel *et al.*, “Tsunami-related data: A review of available repositories used in scientific literature,” *Water (Switzerland)*, vol. 13, no. 16, pp. 1–31, 2021.

[13] S. Widiyantoro *et al.*, “Implications for megathrust earthquakes and tsunamis from seismic gaps south of Java Indonesia,” *Scientific Reports*, vol. 10, no. 1, pp. 1–11, Dec. 2020.

[14] Presiden Republik Indonesia, *Peraturan Presiden No. 93 Tahun 2019 Tentang Penguatan dan Pengembangan Sistem Informai Gempa Bumi dan Peringatan Dini Tsunami (Presidential Regulation of the Republic of Indonesia No. 93 of 2019 Concerning the Strengthening and Development of Earthquake*. Indonesia: Lembaran Negera Republik Indonesia Tahun 2019 Nomor 266, 2019.

[15] D. Inazu, N. Pulido, E. Fukuyama, T. Saito, J. Senda, and H. Kumagai, “Near-field tsunami forecast system based on near real-time seismic moment tensor estimation in the regions of Indonesia, the Philippines, and Chile 4. Seismology,” *Earth, Planets and Space*, vol. 68, no. 1, Dec. 2016.

[16] B. P. Kumar *et al.*, “Tsunami Early Warning System-an Indian Ocean Perspective,” *Journal of Earthquake and Tsunami*, vol. 2, no. 3, pp. 197–226, Nov. 2008.

[17] I. Mudita, W. Hendriyono, G. E. Putri, M. Wibowo, G. Gumbira, and N. D. Sulistyodarmayanti, “Preliminary Research in Tsunami Modelling-Leveraging Artificial Intelligence Technology,” in *Proceeding - 2021 IEEE Ocean Engineering Technology and Innovation Conference: Ocean Observation, Technology and Innovation in Support of Ocean Decade of Science, OETIC 2021*, 2021.

[18] M. Romano *et al.*, “Artificial neural network for tsunami forecasting,” *Journal of Asian Earth Sciences*, vol. 36, no. 1, pp. 29–37, 2009.

[19] S. Marras and K. T. Mandli, “Modeling and simulation of tsunami impact: A short review of recent advances and future challenges,” *Geosciences (Switzerland)*, vol. 11, no. 1, pp. 1–19, 2021.

[20] H. Mase, T. Yasuda, and N. Mori, “Real-Time Prediction of Tsunami Magnitudes in Osaka Bay, Japan, Using an Artificial Neural Network,” *Journal of Waterway, Port, Coastal, and Ocean Engineering*, vol. 137, no. 5, pp. 263–268, Sep. 2011.

[21] J. F. Rodríguez, J. Macías, M. J. Castro, M. de la Asunción, and C. Sánchez-Linares, “Use of Neural Networks for Tsunami Maximum Height and Arrival Time Predictions,” *GeoHazards*, vol. 3, no. 2, pp. 323–344, 2022.

[22] I. E. Mulia, A. R. Gusman, and K. Satake, “Applying a Deep Learning Algorithm to Tsunami Inundation Database of Megathrust Earthquakes,” *Journal of Geophysical Research: Solid Earth*, vol. 125, no. 9, p. e2020JB019690, Sep. 2020.

[23] F. Makinoshima, Y. Oishi, T. Yamazaki, T. Furumura, and F. Imamura, “Early forecasting of tsunami inundation from tsunami and geodetic observation data with convolutional neural networks,” *Nature Communications*, vol. 12, no. 1, pp. 1–10, 2021.

[24] E. Irfiani, A. Pius, P. Raben, V. S. Hadi, and M. Iqbal, “Classification of Tsunami Potential Based on Earthquakes in Indonesia Using the C4 . 5 Algorithm,” *Jurnal Mantik*, vol. 6, no. 3, pp. 3730–3736, 2022.

[25] M. J. Song and Y. S. Cho, “Modeling maximum tsunami heights using bayesian neural networks,” *Atmosphere*, vol. 11, no. 11, pp. 1–13, 2020.

[26] J. H. Lin, Y. F. Chen, C. C. Liu, and G. Y. Chen, “Building a pre-calculated quick forecast system for tsunami run-up height,” *Journal of Earthquake and Tsunami*, vol. 8, no. 3, p. 1440002, 2014.

[27] U. Setiyono, A. R. Gusman, K. Satake, and Y. Fujii, “Pre-computed tsunami inundation database and forecast simulation in Pelabuhan Ratu, Indonesia,” *Pure and Applied Geophysics*, vol. 174, no. 8, pp. 3219–3235, Aug. 2017.

[28] I. Ibtihaj, S. Suparno, M. Rizqy Septyandy, G. Abdul Jabbar, and T. Rani Puji Astuti, “Indonesia paleotsunami database as an effort to reduce the tsunami disasters in Indonesia,” in *E3S Web of Conferences 13th AIWEST-DR 2021 - Aceh International Workshop and Expo on Sustainable Tsunami Disaster Recovery - Post-the 2004 Indian Ocean Tsunami Challenges: Strengthening Community Resilience from Compound Disasters in Pandemic Situations*, vol. 340, p. 01001, 2022.

[29] T. Handayani, R. M, S. Harig, A. Immerz, N. Rakowsky, and J. Grifin, “Extending the database of pre-computed tsunami simulations for the Indonesian tsunami early warning system (InaTEWS) | EPIC,” in

International Tsunami Symposium Bali-Flores 21-25 August 2017, 2017.

[30] D. H. F. Pradana, E. A. Wiguna, R. Amien, A. N. Saputri, M. Wibowo, and W. Hendriyono, "Development of Real-Time Tsunami Early Warning System Dashboard Based on Tunami-F1 and Machine Learning in Sunda Arc , Indonesia," in *The 2022 IEEE Ocean Engineering Technology and Innovation Conference: Management and Conservation for Sustainable and Resilient Marine and Coastal Resources*, 2022.

[31] Daryono, "Peringatan Dini Tsunami di Selatan Jawa." BMKG & Badan Geologi, Bandung, p. 66, 2020.

[32] Indonesian Agency for Geospatial Information (BIG), "Peta BATNAS dan DEMNAS," Jakarta, 2018.

[33] F. Imamura, A. C. Yalçiner, and G. Ozyurt, *Tsunami Modelling Manual*, 2nd ed., no. April. Sendai Japan, 2006.

[34] S. H. Rahmawan, G. Ibrahim, M. A. Mustofa, and M. Ahmad, "Studi Potensi Bahaya Tsunami di Selatan Jawa," Bandung, 2012.

[35] M. Wibowo, "Modeling the Potential of Tsunami Hazard in Labuan Bajo Towards A Disaster-Resilient Tourism Area," *Indonesian Journal of Geography*, vol. 54, no. 1, 2022.

[36] N. H. Mardi, M. A. Malek, M. S. Liew, and H. E. Lee, "A Conceptual Review of Tsunami Models Based on Sumatera-Andaman Tsunami Event," in *InCIEC 2014*, 2015.

[37] B. Adriano, Y. Fujii, and S. Koshimura, "Tsunami source and inundation features around Sendai Coast, Japan, due to the November 22, 2016 M w 6.9 Fukushima earthquake," *Geoscience Letters*, vol. 5, no. 1, p. 12, Dec. 2018.

[38] G. P. Hayes *et al.*, "Slab2, A Comprehensive Subduction Zone Geometry Model," *Science*, vol. 362, no. October, pp. 58–61, 2018.

[39] A. M. Dziewonski, T. A. Chou, and J. H. Woodhouse, "Determination of earthquake source parameters from waveform data for studies of global and regional seismicity.," *Journal of Geophysical Research*, vol. 86, no. B4, pp. 2825–2852, 1981.

[40] G. Ekström, M. Nettles, and A. M. Dziewoński, "The global CMT project 2004-2010: Centroid-moment tensors for 13,017 earthquakes," *Physics of the Earth and Planetary Interiors*, vol. 200–201, pp. 1–9, Jun. 2012.

[41] A. Rahman, A. N. Vita, and S. Rohadi, "Seismic Gap Identification along Sunda Trench using Mapping Analysis of Energy Accumulation Zone," in *AIP Conference Proceedings*, vol. 2320, no. March, pp. 5–10, 2021.

[42] W. Kongko and T. Schlurmann, "The Java tsunami model: Using highly-resolved data to model the past event and to estimate the future hazard," in *Proceedings of the Coastal Engineering Conference*, 2011.

[43] L. Mansinha and D. E. Smylie, "The Displacement Fields of Inclined Faults," *Bulletin of the Seismological Society of America*, vol. 61, no. 5, pp. 1433–1440, Oct. 1971.

[44] F. Imamura, "Review of tsunami simulation with a finite difference method." World Scientific, pp. 25–42, 1996.

[45] Research Center for Hydrodynamic Technology, "Final Report Tsunami Modeling and Data Analyze for Read Down Station," Yogyakarta, 2021.

[46] G. P. Hayes, "USGS Subduction Zone Geometry (Slab 1.0 & Slab 2) and their use in Probabilistic Seismic Hazard Analysis," 2018.

[47] D. L. Wells and K. J. Coppersmith, "New Empirical Relationships among Magnitude, Rupture Length, Rupture Width, Rupture Area, and Surface Displacement," 1994.

[48] A. Cipta and F. Imamura, "Study on tsunami numerical modeling for making tsunami hazard maps in indonesia," *Bulletin of the International Institute of Seismology and Earthquake Engineering*, vol. 43, no. May 2017, pp. 127–132, 2009.

[49] R. R. R. Alam, M. B. Adityawan, M. Farid, A. Chrysanti, Widyaningtias, and M. A. Kusuma, "Tsunami-induced inundation on the coast of Palu City," in *IOP Conference Series: Earth and Environmental Science*, vol. 708, no. 1, 2021.

[50] G. Pasau, G. H. Tamuntuan, and A. Tanauma, "Numerical modelling for tsunami wave propagation (case study: Manado bays)," in *IOP Conference Series: Materials Science and Engineering*, vol. 567, no. 1, 2019.

[51] T. M. Rasyif, S. Kato, Syamsidik, and T. Okabe, "Numerical simulation of morphological changes due to the 2004 tsunami wave around Banda Aceh, Indonesia," *Geosciences (Switzerland)*, vol. 9, no. 3, 2019.

[52] L. H, P. N. T, and I. F, "Tsunami catalog and zones in indonesia," *J. of Natural Disaster Science*, vol. 22, p. 25, 2000.