RGNZ/MGPB

Physics-Informed Neural Networks for Offshore Tsunami Simulation of the 2006 Southern Java Earthquake

Rudarsko-geološko-naftni zbornik
(The Mining-Geology-Petroleum Engineering Bulletin)
DOI: 10.17794/rgn.2026.5.14

Original scientific paper

Sorja Koesuma[1]*, Arnida Lailatul Latifah[2], Angger Zufan Hanggara[1], Ahmad Fauzi Pohan[3], Sismanto Sismanto[4], Budi Eka Nurcahya[4], Sugeng Purwo Saputro[5], Rian Amukti[6]

[1]*Sebelas Maret University, Department of Physics, Surakarta, Indonesia.*
[2]*National Research and Innovation Agency (BRIN), Research Center for Computing, Cibinong, Indonesia.*
[3]*Andalas University, Department of Physics, Padang, Indonesia.*
[4]*Gadjah Mada University, Department of Physics, Yogyakarta, Indonesia.*
[5]*National Research and Innovation Agency (BRIN), Research Center for Geological Resources, Jakarta, Indonesia.*
[6]*National Research and Innovation Agency (BRIN), Research Center of Geological Disaster, Jakarta, Indonesia.*

Abstract

This study presents a comparative analysis between a conventional numerical model, COMCOT, and Physics-Informed Neural Networks (PINNs) for simulating tsunami waves generated by the July 17, 2006 Java earthquake (Mw 7.7). The mesh-free PINNs framework integrates the nonlinear shallow water equations directly into its loss function, enabling simulations without spatial discretization. The model was trained to forecast wave elevation fields using initial conditions derived from the Okada fault model. Results demonstrate that the PINNs model accurately replicates the COMCOT benchmark, successfully capturing the initial tsunami formation and subsequent propagation dynamics. Key quantitative metrics, including correlation coefficients (remaining above 0.98 for most of the simulation) and spatial error distribution (mean absolute percentage error <5% in the open ocean), confirm the model's high accuracy. Critically, once trained, the PINN model generates a complete wave elevation field in under one second, offering a substantial operational advantage over conventional grid-based solvers. Despite slight smoothing effects in high-gradient coastal regions, the study concludes that PINNs offer a reliable and accurate, mesh-free alternative to traditional methods, showing significant promise for future application in rapid tsunami hazard assessment and data-integration for early warning systems, with full real-time deployment identified as a key direction for future research.

Keywords:
modelling; COMCOT; PINNs; non-linear shallow water equation; machine learning

1. Introduction

Indonesia is one of the world's most disaster-prone countries due to its unique geographical position at the convergence of three major tectonic plates: the Indo-Australian, Eurasian, and Pacific plates. This complex tectonic setting results in frequent seismic activities, including high-magnitude earthquakes that often trigger destructive tsunamis (**Latifah and Masykuri, 2021**). Historical records from 2010 to 2018 documented nine significant tsunami events in Indonesia, underscoring the persistent threat to coastal communities and infrastructure (**Wuwungan et al., 2021**).

The southern coast of Central Java is particularly vulnerable to tsunami hazards. This region lies adjacent to the Indian Ocean, where the Java megathrust segment poses a substantial seismic risk. According to the National Center for Earthquake Studies (2017), this megathrust zone has the potential to generate earthquakes with magnitudes up to Mw 8.9. Coastal areas such as Purworejo, Kebumen, and Cilacap support thriving maritime economies and dense populations, and also feature complex subsurface geology. The relatively flat morphology of regions like Kulon Progo facilitates extensive tsunami inundation. The recent development of critical infrastructure, including the New Yogyakarta International Airport (NYIA) in Temon, located within the Aerotropolis area, further amplifies the urgency for accurate tsunami hazard assessment. Recent gravity modelling studies confirm that this Aerotropolis area is situated on rock layers with low to medium density contrast, associated with young sedimentary rocks and alluvium, making it potentially susceptible to shaking amplification during a major earthquake sourced from ei-

* Corresponding author: Sorja Koesuma
e-mail address: sorja@staff.uns.ac.id
Received: 19 January 2026. Accepted: 17 June 2026.
Available online: 26 October 2026

ther the Opak Fault or the megathrust zone (**Darmawan et al., 2024**; **Pohan et al., 2023**). Consequently, this region presents a critical area for further research on earthquake and tsunami hazards (**Hisan et al., 2020**).

Traditional numerical approaches, such as the Cornell Multi-grid Coupled Tsunami Model (COMCOT), have been widely applied in tsunami modelling by solving the nonlinear shallow-water equations (SWE) through finite-difference methods. However, these grid-based techniques are known to encounter significant limitations, including high computational costs when fine grids are used for coastal detail, as well as issues with numerical instability and oscillations in high-resolution scenarios (**Arifin et al., 2025; Li and Wang, 2024**). Such computational demands and stability concerns present major obstacles for real-time tsunami forecasting, where rapid and reliable simulations are crucial. These challenges highlight the growing necessity for more advanced modelling techniques that can provide increased accuracy while also reducing computational overhead. Recent innovations – such as unstructured grids, multi-resolution cell models, and improved time-stepping algorithms – offer promising solutions to make tsunami simulations both more efficient and robust (**Alkhadhr and Almekkawy, 2023**).

In recent years, Physics-Informed Neural Networks (PINNs) have emerged as a promising alternative for solving partial differential equations (PDEs) governing natural phenomena, including tsunami propagation. Unlike conventional methods, PINNs integrate physical laws directly into the neural network's loss function, eliminating the need for mesh discretization and enabling mesh-free, data-efficient simulations (**Tian et al., 2025; Raissi et al., 2019**). By embedding the shallow-water equations into the training process, PINNs ensure that predictions adhere to fundamental hydrodynamic principles while leveraging the flexibility of deep learning. This approach not only reduces computational expenses but also enhances the model's ability to handle complex boundary conditions and sparse observational data, making it particularly suitable for regions with limited historical tsunami records (**Alkhadhr and Almekkawy, 2023**).

The concept related to PINNs has spurred various advanced research, particularly in the fields of fluids and waves. For example, research by **Bihlo et al. (2022)** applied PINNs to the Shallow-water Equations in spherical coordinates, **Cho (2024)** utilized them to solve the Shallow-water Equations on grids with arbitrary geometry, and **Li and Li (2022)** used them to model Rogue waves from the nonlinear Schrodinger equation. These advanced studies have demonstrated the success of machine learning methods based on artificial neural networks and physical information in addressing complex problems (**Raissi et al., 2019; Moseley et al., 2020**).

This study presents a comparative analysis between PINNs and the conventional numerical model COMCOT v1.7 for tsunami wave modelling, using the 2006 Java earthquake (Mw 7.7) as a case study. The research aims to develop a PINNs framework for simulating tsunami generation and propagation based on seismic parameters, evaluate the accuracy of PINNs in reproducing tsunami dynamics compared to COMCOT simulations, and assess the computational efficiency and resource requirements of both approaches. The findings contribute to understanding the potential of machine learning methods in complementing traditional numerical approaches for tsunami hazard assessment, with implications for developing next-generation early warning systems.

2. Methods

2.1. Study Area

The study area encompasses the southern coastal region of Central Java, Indonesia, spanning from 108.347° to 111.539° East longitude and -7.670° to -10.828° South latitude (see **Figure 1**). This domain covers approximately 353.52 km × 137.53 km, incorporating both deep ocean areas and vulnerable coastal zones, including the densely populated regions of Purworejo, Kebumen, Cilacap, and Kulon Progo. Within this domain, four virtual automatic water level (AWL) evaluation stations were established at strategic locations. Point 1 is positioned in the deep offshore region to capture open-ocean propagation characteristics. Points 2 and 3 are located along the coastline of Kulon Progo adjacent to the New Yogyakarta International Airport (NYIA) aerotropolis zone, which represents the primary critical infrastructure of concern in this study given its location on low-to-medium density sedimentary formations susceptible to tsunami inundation. Point 4 is positioned near Cilacap to capture tsunami wave behaviour in a different coastal geometry and bathymetric environment, enabling assessment of the model performance across diverse nearshore conditions. The star symbol denotes the epicenter of the simulated earthquake.

2.2. Initial Condition

A tsunami is generated by the deformation of a submarine fault, which causes a displacement of the sea surface. The initial tsunami wave condition is obtained by calculating this vertical displacement of seawater. The vertical surface displacement can be calculated using the Steketee equation (**Tanioka and Satake, 1996**).

Where δ_{ij} is Kronecker delta, v_j is the direction cosine of the normal to the fault surface element $d\Sigma$, and λ and μ are the Lame constants. According to Tanioka and Satake (1996), μ_{ik} is the k^{th} component of the surface displacement brought about in the i^{th} direction by a point force of size F.

The Okada model, which is frequently used to produce the initial condition for tsunami modelling and

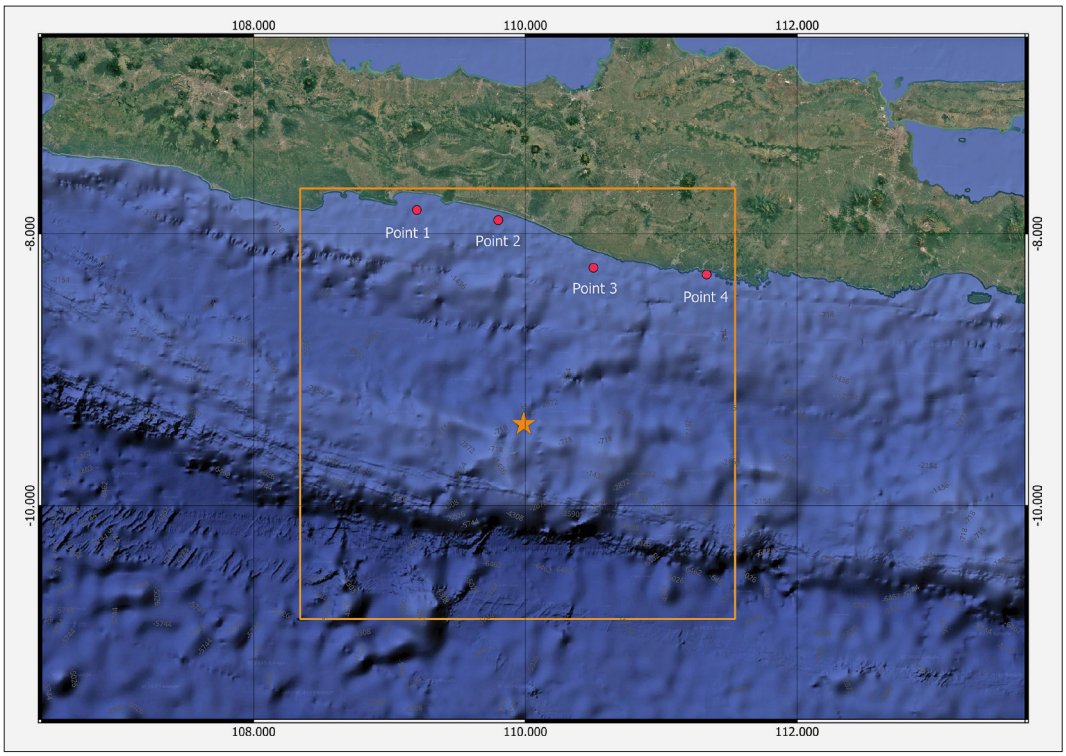


Figure 1: Spatial extent of the research domain covering the southern coastal region of Central Java, Indonesia, encompassing both deep-ocean areas and vulnerable coastal zones such as Purworejo, Kebumen, Cilacap, and Kulon Progo. The star marker represents the earthquake epicenter. The four dot points denote the evaluation locations.

simulation, is based on the Steketee equation. The region of seabed deformation and the aerial reach and shape of the initial tsunami wave are the same in the Okada model. The fault geometry parameters – fault length, dip, slip, strike, and depth – are necessary for this model. The sea surface displacement in an elastic half-space is computed using the Okada model (**Okada, 1985**).

$$\mu_k = \frac{1}{F}\Delta\mu_i\left[\lambda\delta_{ij}\mu_k^{n,n}+\mu\left(\mu_k^{i,j}+\mu_k^{j,i}\right)\right]v_j d\Sigma \qquad (1)$$

Using the empirical connections of Wells and Coppersmith (1994), regression analysis between the moment magnitude (M) and the base-10 logarithm of the parameter to be obtained is used to find the fault parameters. The following are some examples of these empirical relationships:

$$y=a+b^{*}log\ (x) \qquad (2)$$

$$log(y)=a+b^{*}log\ (x) \qquad (3)$$

Where the dependent and independent variables are denoted by x and y, respectively. The number of occurrences, regression coefficients (a and b), standard error of the coefficients, dependent variable standard deviation (s), correlation coefficient (r), and data range are among the regression descriptors (**Wells and Coppersmith, 1994**).

2.3. Governing Equation

The core of our PINNs approach encodes the nonlinear shallow water equations (SWE) - the fundamental PDEs governing tsunami propagation - directly into the neural network's loss function. The SWE system consists of the continuity and momentum equations. The continuity equation ensures mass conservation, expressing how the water surface elevation changes in response to horizontal fluxes P and Q (**Wijayanti et al., 2022**; **Xing, 2023**; **Wang, 2009**).

$$\frac{\partial\eta}{\partial t}+\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}=0 \qquad (4)$$

Where:

η= Water surface elevation (m)

$P=uH$ = x-direction flux ($H=\eta+h$, $total\ water\ depth$)

$Q=vH$ = y-direction flux

h= bathymetry (m)

The momentum equations track how these fluxes evolve under the influence of pressure gradients, nonlinear advection, and frictional effects.

$$\frac{\partial P}{\partial t}+\frac{\partial}{\partial x}\left(\frac{P^2}{H}\right)+\frac{\partial}{\partial y}\left(\frac{PQ}{H}\right)+gH\frac{\partial\eta}{\partial x}+F_x=0 \qquad (5)$$

$$\frac{\partial Q}{\partial t}+\frac{\partial}{\partial x}\left(\frac{PQ}{H}\right)+\frac{\partial}{\partial y}\left(\frac{Q^2}{H}\right)+gH\frac{\partial\eta}{\partial y}+F_y=0 \qquad (6)$$

Where:

g = gravitational acceleration ($9.8\ m/s^2$)

A critical aspect of these equations is their treatment of the nonlinear term, which becomes increasingly important as waves approach shallow coastal areas. The Manning friction terms (F_x and F_y) account for energy dissipation due to seafloor roughness, particularly crucial in near-shore regions where bathymetry variations are significant. Each is described in the equation (7) and (8) (**Wang, 2009**).

$$F_x = \frac{gn^2}{H^{7/3}} P(P^2 + Q^2)^{1/2} \tag{7}$$

$$F_y = \frac{gn^2}{H^{7/3}} Q(P^2 + Q^2)^{1/2} \tag{8}$$

2.4. Physics-Informed Neural Networks

The architecture used in modelling PINNs is generally a feedforward architecture with several hidden layers. Specifically, the implemented model is a fully connected network (FCN) with an input layer for the four features (*x, y, t, z*), six hidden layers each containing 100 neurons, and an output layer for three solution variables (η, u, v) (**Raissi et al., 2019**; **Lu et al., 2021**). The activation function used in this architecture is *tanh* because this function limits its output to the range between -1 and 1 (**Jagtap et al., 2020**). The input data is a combination of longitude coordinate data (*x*), latitude (*y*), time (*t*), and depth or bathymetry (*z*). Meanwhile, the output layer consists of wave height (η), velocity in the *x*-direction (u), and velocity in the *y*-direction (v).

The inclusion of bathymetric depth (z) as a fourth input feature is a deliberate and physically motivated design choice. The governing nonlinear shallow water equations (Equation 4–8) explicitly depend on the total water depth H = η + h, where h is the bathymetric depth. In a mesh-free PINN formulation, the network must evaluate physics residuals at arbitrary collocation points (x, y, t). Without z as an input, the network cannot access local depth information, making it impossible to correctly compute the depth-dependent terms P²/H, PQ/H, gH∂η/∂x, and the Manning friction terms (Equation 7, 8) that are central to the shallow water dynamics. The inclusion of z therefore reflects a physical necessity rather than a statistical shortcut, and is consistent with recent PINN formulations for depth-dependent wave equations (**Tian et al., 2025**). Importantly, the bathymetric data used as input to the network is identical to that used by COMCOT, ensuring a consistent physical basis for comparison.

In this study, the model's total loss is calculated by combining three distinct components: numerical data, initial condition data, and the physics rules represented by shallow water equations. The mean squared error (MSE) method is used to measure the difference between the predicted values and the true values for each component. The total loss, which is a weighted sum of these three components, is formulated in equation:

$$L_{total} = \lambda_{IC} L_{IC} + \lambda_{PDE} L_{PDE} + \lambda_{data} L_{data} \tag{9}$$

The parameters λ_{IC}, λ_{data} & λ_{PDE} are weighting factors that control the relative contribution of each component to the total loss function. Because gradient imbalance can impede the training process, proper loss balancing is essential for PINNs effectiveness (**Wang et al., 2021**). These parameters were consecutively set to 1, 1, and 0.0625 in this investigation in accordance with Cho's optimization technique (**Cho, 2024**).

The loss functions for the initial condition and the numerical data use supervised learning. The initial condition data is taken at t=0, while the numerical data is sampled across the entire domain. In both cases, data points are randomly selected. The proportion of data used for each component is 20,000 points out of 146,300 (13.67%) for the initial condition and 50,000 points out of 4,535,300 (1.11%) for the numerical data. These loss functions are defined by the following equations:

$$L_{IC} = \frac{1}{N_{IC}} \sum_{i=1}^{N_{IC}} \left| \eta(x_i, y_i, t_0) - \eta_{pred}(x_i, y_i, t_0) \right|^2 \tag{10}$$

$$L_{data} = \frac{1}{N_{data}} \sum_{i=1}^{N_{data}} \left| \eta(x_i, y_i, t_i) - \eta_{pred}(x_i, y_i, t_i) \right|^2 \tag{11}$$

Meanwhile, the physics-based loss component uses unsupervised learning. It is calculated on input data points (x,y,t,z) sampled across the entire domain using Latin Hypercube Sampling (LHS). It was chosen specifically due to its stratified sampling technique that ensures comprehensive domain coverage with significantly fewer points compared to traditional random sampling methods (**Iordanis et al., 2025**). Using LHS, we took a total of 400,000 points from 4,535,300 (8.82%). This loss function itself consists of three sub-components derived from the governing PDEs: the continuity equation, the *x*-momentum equation, and the *y*-momentum equation. The total physics loss is defined by the following equations:

$$L_{PDE} = L_{con} + L_{mom-x} + L_{mom-y} \tag{12}$$

$$L_{con} = \frac{1}{N_{con}} \sum_{i=1}^{N_{con}} \left| \frac{\partial \eta_i}{\partial t} + \left(\frac{\partial P_i}{\partial x} + \frac{\partial Q_i}{\partial y} \right) \right|^2 \tag{13}$$

$$L_x = \frac{1}{N_x} \sum_{i=1}^{N_x} \left| \frac{\partial P_i}{\partial t} + \frac{\partial}{\partial x}\left(\frac{P_i^2}{H} \right) + \frac{\partial}{\partial y}\left(\frac{P_i Q_i}{H} \right) + gH \frac{\partial \eta}{\partial x} + F_x \right|^2 \tag{14}$$

$$L_y = \frac{1}{N_y} \sum_{i=1}^{N_y} \left| \frac{\partial Q_i}{\partial t} + \frac{\partial}{\partial x}\left(\frac{P_i Q_i}{H} \right) + \frac{\partial}{\partial y}\left(\frac{Q_i^2}{H} \right) + gH \frac{\partial \eta}{\partial y} + F_y \right|^2 \tag{15}$$

2.5. Data Generation

We used supervised learning, which necessitates giving the network a dataset, to train the PINNs model. The dataset used in this study was divided into two parts: (1) initial condition data from the Okada fault model and (2) supervised data from tsunami propagation simulations using the COMCOT model. Although the source site was moved southward toward the southern coast of Java and Yogyakarta for the purposes of this simulation, the fault characteristics were based on the Pangandaran earthquake that occurred on July 17, 2006. This change is in line with the parameter configuration used by **Darmawan et al. (2024)**, who made a comparable modification for scenario-based tsunami modelling in the same area. **Table 1** contains a list of the adjusted fault parameters used in this research which were obtained from the historical earthquake catalog available on the United States Geological Survey (USGS) website (Search Earthquake Catalog (usgs.gov)).

Table 1: Fault parameters

Parameters	Values
Longitude (°)	109.988
Latitude (°)	-9.399
Length (km)	160
Width (km)	21
Depth (km)	20
Strike (°)	284
Dip (°)	9
Rake (°)	87
Dislocation (m)	3.95
Magnitude	7.7

For the supervised component, tsunami propagation was simulated using a Cartesian coordinate system on a computational grid of 385 × 380 over a time span of 0–30 minutes, with discretization parameters of 0.911 × 0.925 km for spatial resolution and 1 minute for temporal resolution. The entire initial state of the simulation was included in the dataset, and at t=0, representative points from the solution space were sampled and appended if they were located inside the solution domain.

For the initial condition component, the Okada model was applied to compute the coseismic sea surface deformation using the given fault parameters (**Okada, 1985**). This combined dataset ensured that the PINNs model was trained not only with supervised data derived from physics-based numerical modelling of tsunami propagation, but also with realistic initial conditions consistent with the physical characteristics of the July 17, 2006 tsunami event.

2.6. Model Evaluation

The performance of the PINNs model is rigorously evaluated by benchmarking its predictions against the results from a high-fidelity numerical model generated using the COMCOT software, following established practices in tsunami and wave modelling (**LeVeque et al., 2011**; **Synolakis et al., 2008**). This comparative analysis employs a multi-faceted validation approach based on several key metrics to quantify discrepancies in both time and space.

First, the absolute error is calculated at each time step to quantify the instantaneous point-wise discrepancy between the two models. Second, temporal correlation analysis is employed to assess the synchronization of the wave dynamics evolution over time; correlation coefficients provide critical insight into the phase agreement and the consistency of wave propagation patterns (**Synolakis et al., 2008**). Third, the mean absolute percentage error (MAPE) is computed to illustrate the spatial distribution of the relative error across the entire computational domain. This metric effectively evaluates the global spatial agreement and highlights regions with significant discrepancies, particularly in critical near-shore zones where nonlinear bathymetric interactions (**Madsen and Schaeffer, 2010**).

In addition to these field-wide metrics, a point-specific validation is conducted by establishing a network of imaginary automatic water level (AWL) stations at strategic coastal locations. A detailed time-series analysis at these virtual stations allows for a direct comparison of essential tsunami characteristics, including the wavefront arrival time, peak amplitude, and waveform shape, as predicted by PINNs against the COMCOT benchmark (**Løvholt et al., 2012**; **Post et al., 2009**). This multi-pronged methodology ensures a comprehensive assessment of the PINNs model's accuracy in replicating both the global wave field and local hydrodynamic features.

3. Results and Discussion

3.1. Initial Tsunami Generation and Source Characterization

The Okada model, which was based on the fault parameters of the July 17, 2006 earthquake, was used to create the initial tsunami wave field. With considerable vertical displacement, the deformation pattern (see **Figure 2**) displays a typical thrust fault mechanism. The fault's northwest segment saw the most uplift, which was around 0.75 meters, while the southeast segment experienced up to 0.45 meters of subsidence. A significant amount of water is efficiently displaced by this bipolar deformation pattern, which extends 160 km along strike and 21 km along dip, starting the tsunami wave propagation.

The Mw 7.7 scenario was found to be suitable for the fault characteristics determined from Wells and Coppersmith's (1994) empirical relationships, with an average dislocation of 3.95 meters and seismic moment of 4.107 $\times 10^{20}$ Nm. The fault parameters used included a strike of 284°, a dip of 9°, and a rake of 87°. The thrust mecha-

nism is compatible with the original wave condition's symmetrical pattern along the fault strike.

3.2. Tsunami Propagation PINNs

Reconstructing the initial wave field and early phases of propagation was a great feat of the PINNs model. After 100,000 iterations across two training stages, the training process reached convergence with a final loss value of 0.00018, as illustrated in **Figure 3a** and **3b**. While respecting the beginning and boundary conditions, the model was able to learn the physical limitations incorporated into the nonlinear shallow water equations.

The COMCOT numerical model, the PINNs model, and the absolute error at t = 0 s, which represents the initial condition of tsunami formation, are compared in **Figure 4.**

Sea surface elevation in the vicinity of the earthquake source location has the same spatial pattern in both COMCOT and PINNs. The raised and subsided zones created by the seafloor deformation brought on by the July 17, 2006 earthquake are represented by the red and blue areas. The initial wave field is accurately reproduced by the PINNs model, which closely resembles the reference COMCOT result. Only the transition zone between positive and negative elevations exhibits minor changes, which are probably the result of little variations in learning the spatial gradients. The PINNs architecture correctly captures the initial deformation pattern and satisfies the starting and boundary restrictions of the physical system, as evidenced by the absolute error being less than 0.01 m over the whole domain.

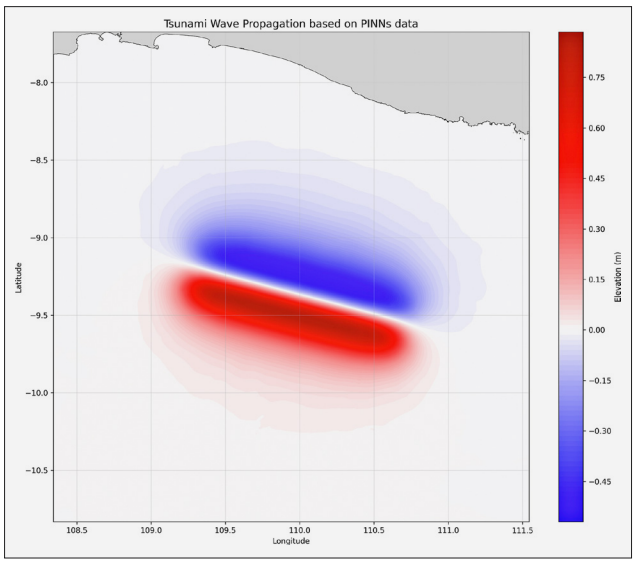


Figure 2: Initial tsunami wave field generated using the Okada dislocation model based on the fault parameters of the July 17, 2006 Java earthquake

The tsunami waves start to move northward toward Java's southern shore at t = 300 s (see **Figure 5**). In the middle of the computational domain, COMCOT and PINNs both show two primary wave fronts and a distinct interference pattern.

The tsunami's amplitude and propagation pattern are impressively well-represented by the PINNs model. The PINNs-generated wavefront appears somewhat smoother in the western portion of the domain (around 108.5°–109°E), where modest deviations are noted. This smoothing is attributed to three interacting factors: (1) the *tanh* activation function and the continuous function approximation inherent to fully connected networks, which tend to suppress high-frequency spatial gradients; (2) the sparser distribution of training collocation points in shallow near-shore regions where bathymetric gradients are steepest; and (3) the increased complexity of nonlinear wave interactions in shallow water, which places higher demands on the physics loss. The relatively poor representation of severe bathymetric gradients in the training data further contributes to this local disparity.

The tsunami wave has moved farther from the source area at t = 600 s (see **Figure 6**), interacting with the intricate bathymetry of the southern Java trench. Similar circular wave patterns are produced by COMCOT and PINNs, with wave amplitudes progressively diminishing as the energy disperses.

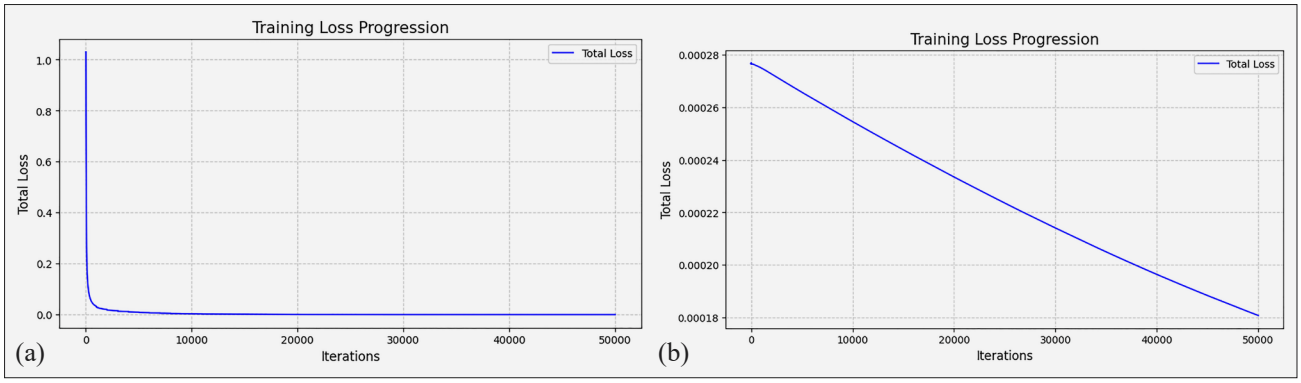


Figure 3: Training loss progression for PINNs, (a) Loss reduction during Stage 1 training, focusing on supervised learning from COMCOT data and initial condition constraints, (b) Loss reduction during Stage 2 training, emphasizing physics-based residual minimization using the shallow water equations.

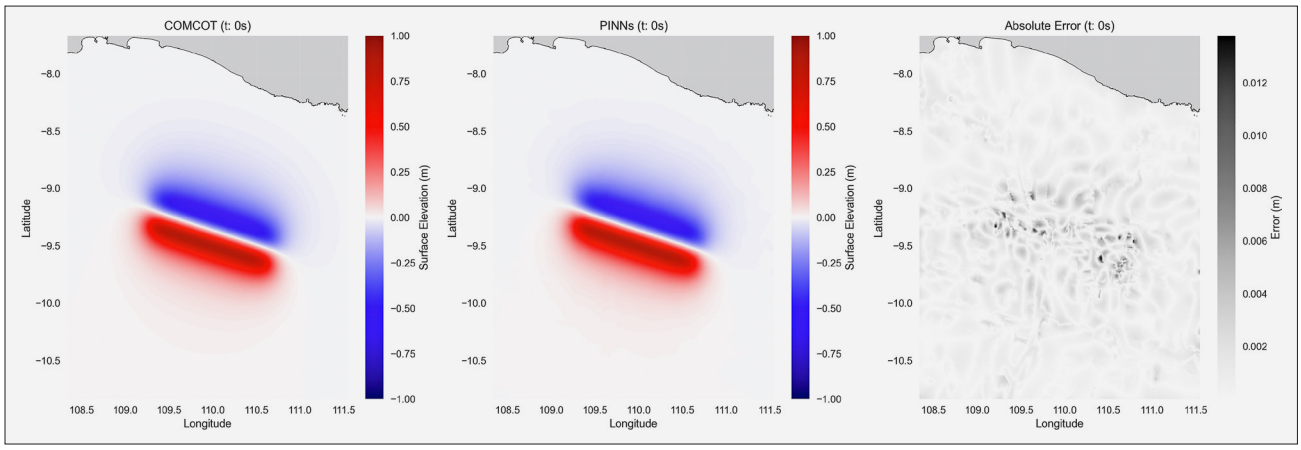


Figure 4: Comparison of surface wave elevation between COMCOT, PINNs, and their absolute error at t = 0 s

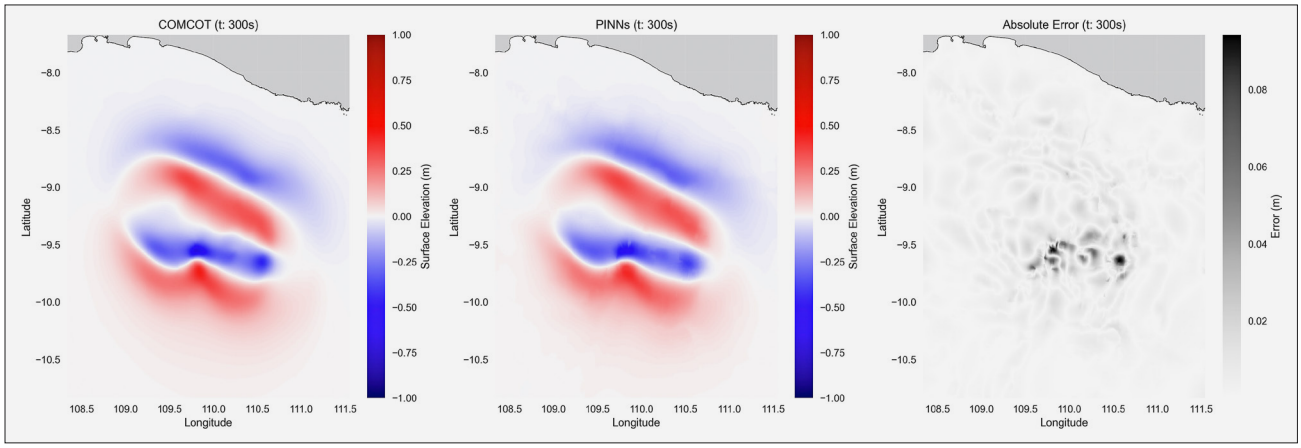


Figure 5: Comparison of surface wave elevation between COMCOT, PINNs, and their absolute error at t = 300 s

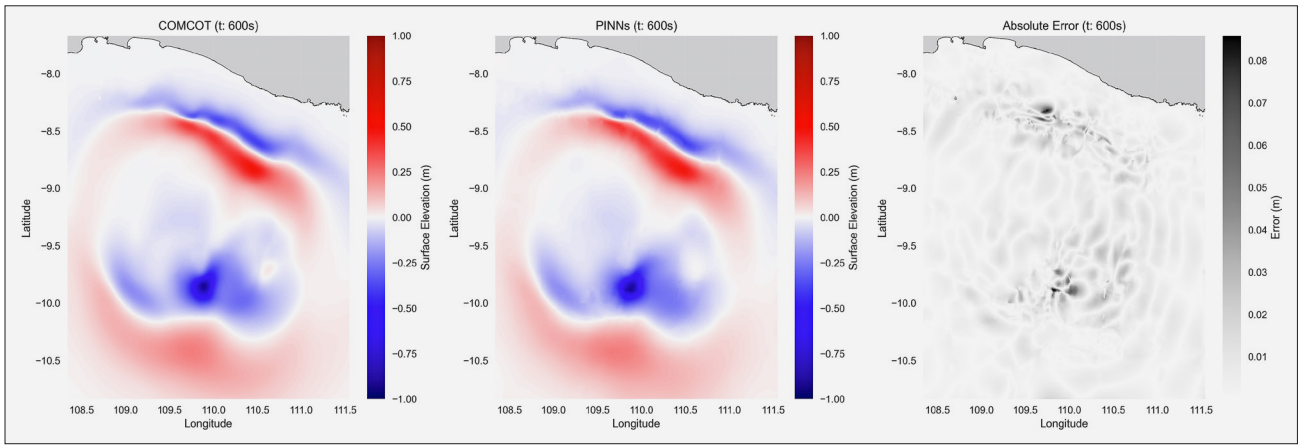


Figure 6: Comparison of surface wave elevation between COMCOT, PINNs, and their absolute error at t = 600 s

The general structure and direction of propagation are reliably captured by the PINNs model and the reference model. Near the northern boundary of the domain, where complex topography causes stronger wave refraction, slight amplitude variations are apparent. In these regions, PINNs typically exhibit a slightly smoother profile than COMCOT. This behaviour arises primarily from the *tanh* activation function and the continuous function approximation of the neural network, which inherently suppress abrupt spatial oscillations. Additionally, the collocation points sampled via Latin Hypercube Sampling are less dense in the steep bathymetric gradients of the near-shore zone, limiting the network's ability to resolve sharp wave fronts. The physics loss, while enforcing the shallow-water equations, does not fully compensate for this under-sampling, leading to locally

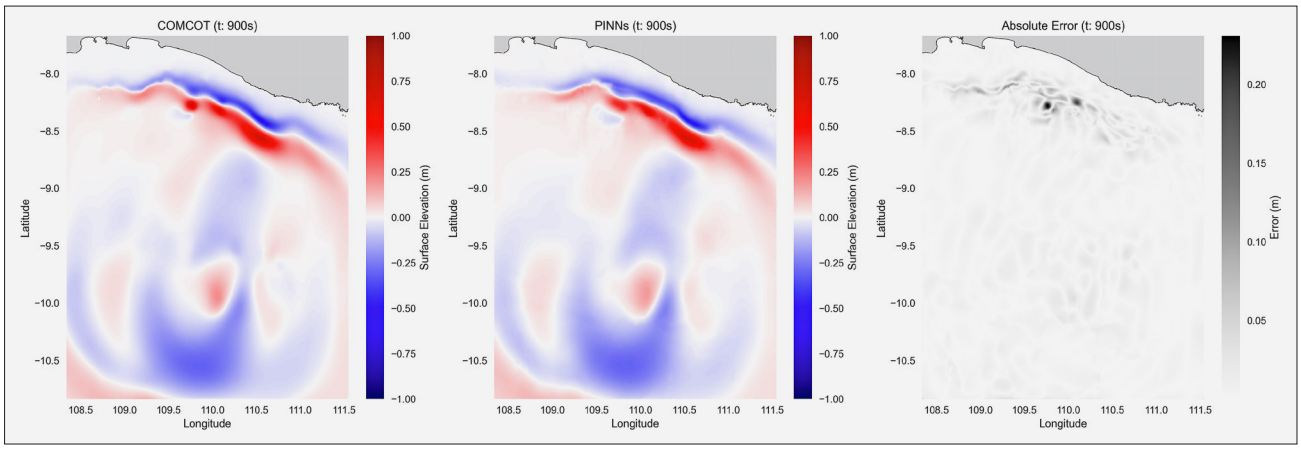


Figure 7: Comparison of surface wave elevation between COMCOT, PINNs, and their absolute error at t = 900 s

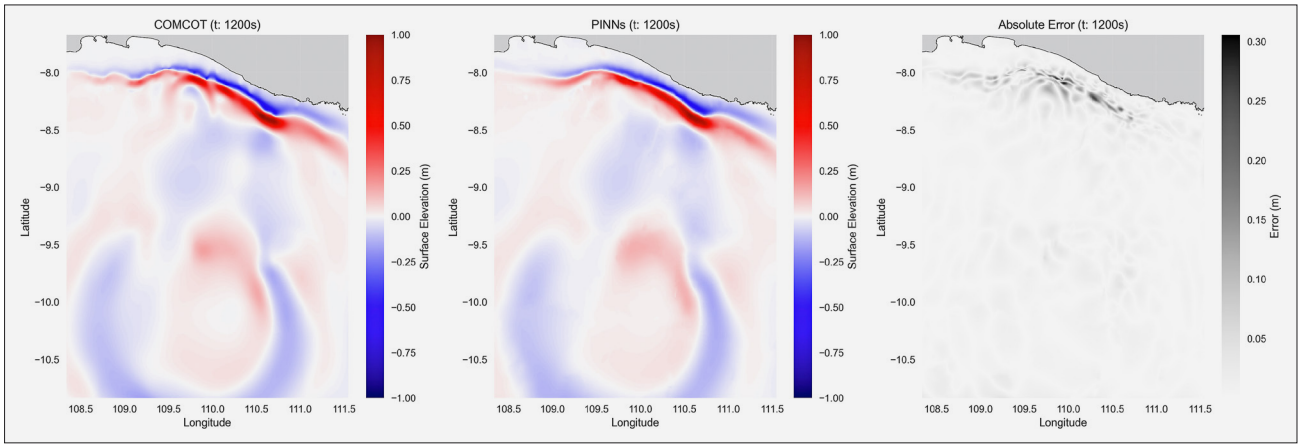


Figure 8: Comparison of surface wave elevation between COMCOT, PINNs, and their absolute error at t = 1200 s

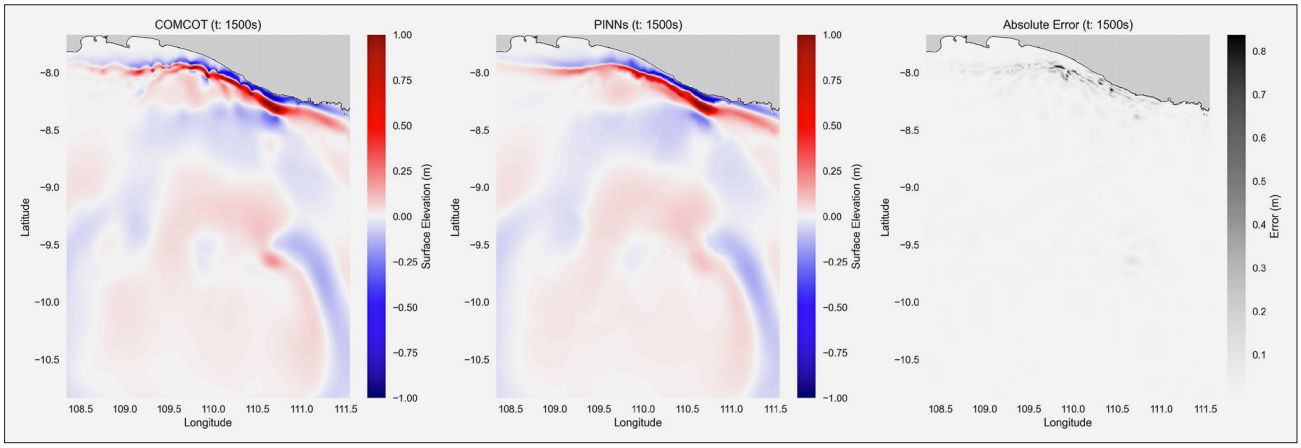


Figure 9: Comparison of surface wave elevation between COMCOT, PINNs, and their absolute error at t = 1500 s

smoother predictions. With most values below 0.05 m and a localized maximum of roughly 0.08 m near the interference zones, the absolute error is still rather small. This suggests that nonlinear wave dynamics can be reproduced by the PINNs framework without numerical divergence, even at intermediate propagation times.

The tsunami wave is still moving and interacting with the intricate bathymetric features south of Java at t = 900 s (see **Figure 7**). The most noticeable wavefront is seen close to the shore at 8.5°S, and both COMCOT and PINNs exhibit a regular pattern of alternating positive (red) and negative (blue) surface elevations.

Near the coastline, the COMCOT model exhibits sharper gradients and slightly higher amplitudes, whereas the PINNs results show smoother transitions. This is not a dissipation effect due to regularization but rather a

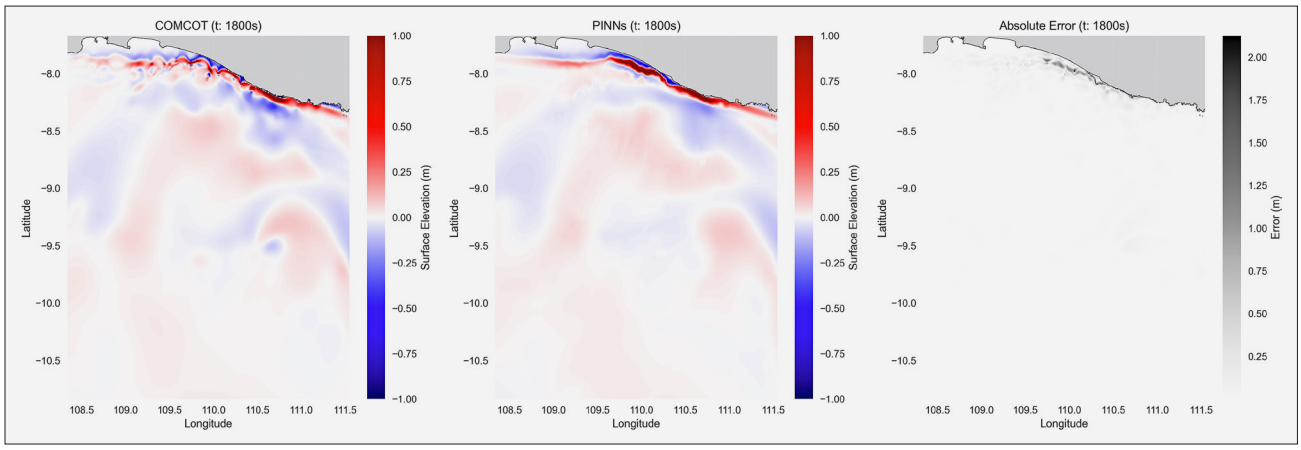


Figure 10: Comparison of surface wave elevation between COMCOT, PINNs, and their absolute error at t = 1800 s

consequence of three factors: the tanh activation function limiting high-frequency components, the lower density of collocation points in shallow regions with complex bathymetry, and the difficulty of the physics loss to exactly balance nonlinear advection and friction terms when gradients are extremely steep. Consequently, PINNs produce slightly smoothed nearshore waveforms while maintaining physically consistent propagation patterns. Despite these variations, both models show a high degree of consistency in the tsunami's basic propagation path and phase. The greatest disparities are located in the nearshore area, where bathymetric fluctuations are more complex, and the absolute error is generally less than 0.20 m throughout the computational domain. These findings show that even when nonlinear interactions between the incident and reflected waves grow more prevalent, PINNs retain both numerical stability and physical consistency.

The tsunami waves reach a more advanced stage of propagation around t = 1200 s (see **Figure 8**), with many reflected and refracted wavefronts evident across the southern Java Basin. Strong constructive interference is captured by both models throughout the coastal boundary, especially between 109.5°E and 110.5°E, where positive wave amplitudes are greater than 0.8 m.

This behaviour is precisely reproduced by the PINNs simulation, which has slightly finer amplitude peaks but a nearly identical spatial distribution to COMCOT. This smoothing effect is not caused by a pre-defined regularization term but arises from the intrinsic properties of the PINN architecture: the tanh activation function and the continuous function approximation across the spatio-temporal domain. When combined with the sparser collocation point distribution in steep-bathymetry regions and the high nonlinearity of the shallow-water equations near the coast, the network produces slightly smoothed amplitude peaks while still preserving the overall phase and propagation direction. At this point, the absolute error significantly rises and reaches a local maximum of roughly 0.30 m, which are mostly found in regions where several wavefronts interact. However, the majority of the domain displays errors smaller than 0.10 m, indicating that the PINNs framework is still dependable and strong for modelling long-duration tsunami propagation.

The tsunami wave has moved considerably closer to the coastline at t = 1500 s (see **Figure 9**), illustrating the event's changing dynamics. A well-defined system of wave crests and troughs characterizes the basic structure of the wave field, which is consistently captured by both the COMCOT and PINNs models. The PINNs model accurately reproduces the principal wave front's spatial location, closely matching the COMCOT benchmark between longitudes 109.5°E and 110.5°E.

One important finding is that the PINNs solution shows a smoother depiction of these wave contours, whereas the COMCOT model resolves slightly sharper gradients and more complex elements within the wave field This smoothing effect is primarily due to the continuous function approximation of the fully connected network and the use of the *tanh* activation function, which inherently limit high-frequency spatial variations. Furthermore, the collocation points sampled in the shallow coastal zone are sparser relative to the complexity of the bathymetry, and the physics loss, while enforcing the shallow-water equations, cannot entirely compensate for this limited resolution. Adaptive collocation sampling, identified as a future research priority, would help mitigate this issue. This smoothing behaviour, while beneficial for numerical stability, may result in a conservative underestimation of peak inundation heights at the shoreline, which constitutes a limitation that warrants further investigation in future work. Over the great majority of the domain, the absolute error between the two models is still remarkably small, typically less than 0.3 meters. As might be predicted, the most noticeable errors are concentrated along the wave fronts' steep gradients, where minute variations in phase or amplitude are amplified. This outcome demonstrates that the PINNs model retains a reliable and physically consistent simulation at this intermediate propagation stage, successfully captur-

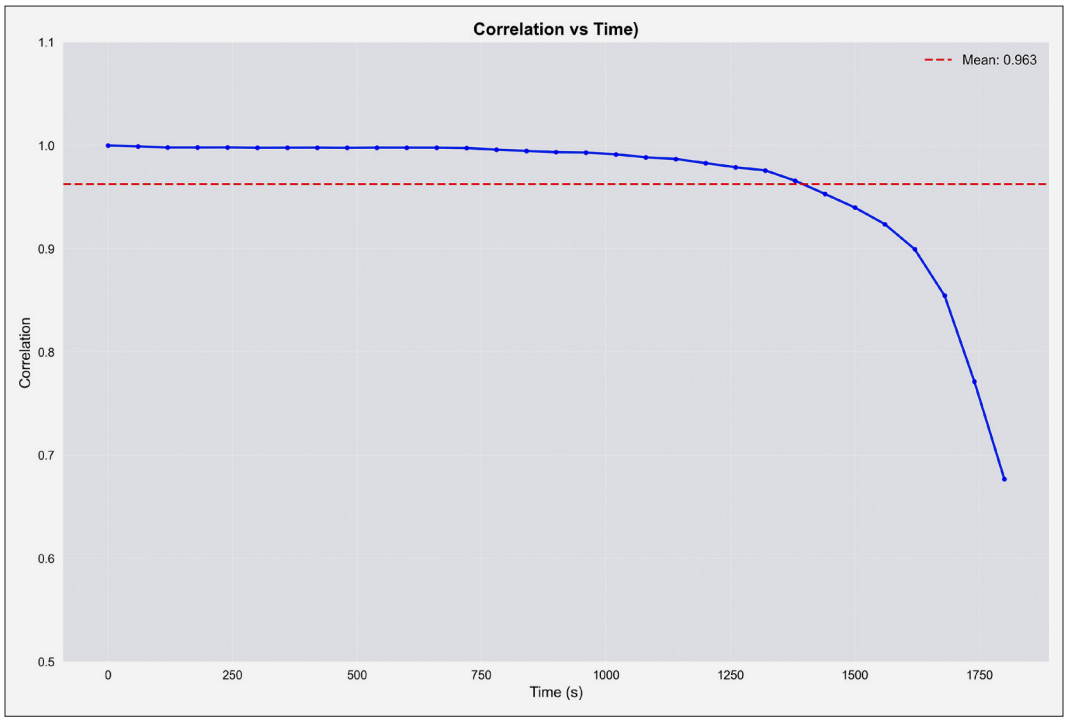


Figure 11: Time-dependent correlation coefficient between COMCOT and PINNs wave height fields.

ing the key propagation dynamics while generating only small, localized discrepancies.

The tsunami wave has moved closer to Central Java's shoreline at t = 1800 s (see **Figure 10**). With distinct wave crests (red) and troughs (blue) throughout the domain, the COMCOT and PINNs models both show a cohesive and intricate wave field structure. Both models predict that the main wave front will approach the southern coast between roughly 109.5°E and 110.5°E.

The PINNs solution captures the general pattern with slightly smoothed contours, whereas the COMCOT simulation preserves sharper wave fronts and fine-scale characteristics. This smoothing is not a deliberate regularization effect but rather the outcome of three interacting factors: the *tanh* activation function and continuous network approximation, the relatively sparse collocation points in the steep-bathymetry near-shore region, and the increased demands placed on the physics loss by strongly nonlinear shallow-water interactions. Improving near-shore resolution through adaptive collocation sampling is identified as a key direction for future work. With errors mostly below 0.25 m throughout the majority of the open ocean, the absolute error map demonstrates a high degree of agreement between the two models. It is most difficult to capture the most dynamic and non-linear physics in the leading wave fronts and gradients, where the biggest mistakes are concentrated. These results show that even as the wave moves into shallower coastal waters, the PINNs framework effectively predicts the progressive wave propagation and preserves physical fidelity.

3.3. *Comparative Analysis*

3.3.1. Error Metrics and Spatial-Temporal Accuracy

For most of the simulation time, the correlation coefficient between the COMCOT and PINNs outputs stays strong (see **Figure 11**). The correlation remains above 0.98 from the beginning of the simulation until about 1300–1400 s, suggesting that the two models' temporal evolution of wave height is almost perfectly in agreement. The great overall consistency is further supported by the mean correlation value of 0.962 (red dashed line).

After 1400 s, there is a discernible fall, and as the simulation gets closer to 1800 s, the connection drops even more dramatically. This decrease coincides with the main wave crest's arrival in a number of coastal locations close to the simulation window's edge, where the short simulation period may result in insufficient waveform capture. Because of this, the PINNs model accurately replicates the overall structure but shows somewhat diminished temporal synchronization in the vicinity of the last time steps.

The PINNs framework effectively reflects the dynamic behaviour of tsunami propagation, as evidenced by the high correlation over the majority of the time domain. Deviations mostly occur close to the simulation border, where physical information becomes sparse.

The spatial distribution of the mean absolute percentage error (MAPE) over all time steps (see **Figure 12**) shows that errors are primarily concentrated around the coast, with a maximum MAPE of 38.6% at particular shoreline sites. Open ocean regions, on the other hand,

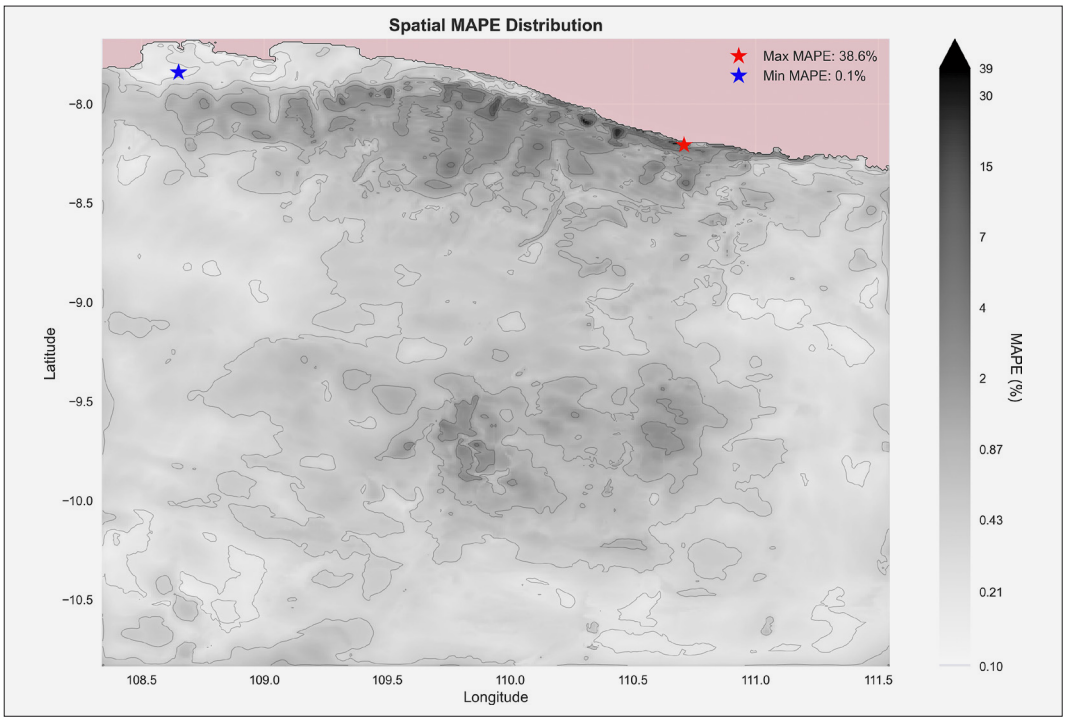


Figure 12: Spatial distribution of mean absolute percentage error (MAPE) over all time steps.

continuously maintain low errors (<5%), indicating the effectiveness of PINNs in simulating deep-water propagation where nonlinear effects are less noticeable.

3.3.2. Wave Characteristics and Travel Time Analysis

The PINNs model reproduces the COMCOT outputs with strong temporal and amplitude consistency, according to the time series comparison over four coastal sites (Point 1-4) (see **Figures 13 & 14)**. Both models record a low-energy signal with nearly equal flat-to-gentle fluctuations at Point 1, where wave activity is negligible. With a maximum difference of just 0.001 m, the peak appears close to the end of the 1800 s window.

Both models depict the gradually increasing wave amplitude near the shoreline at Point 2, which is marked by a positive wave crest that rises gradually. Although PINNs slightly overestimates the amplitude, the peak appears near 1800 s, resulting in the biggest difference (0.145 m). The general waveform, which has a late-arriving peak, a mild upward slope, and an initial quiescent condition, is consistent among models.

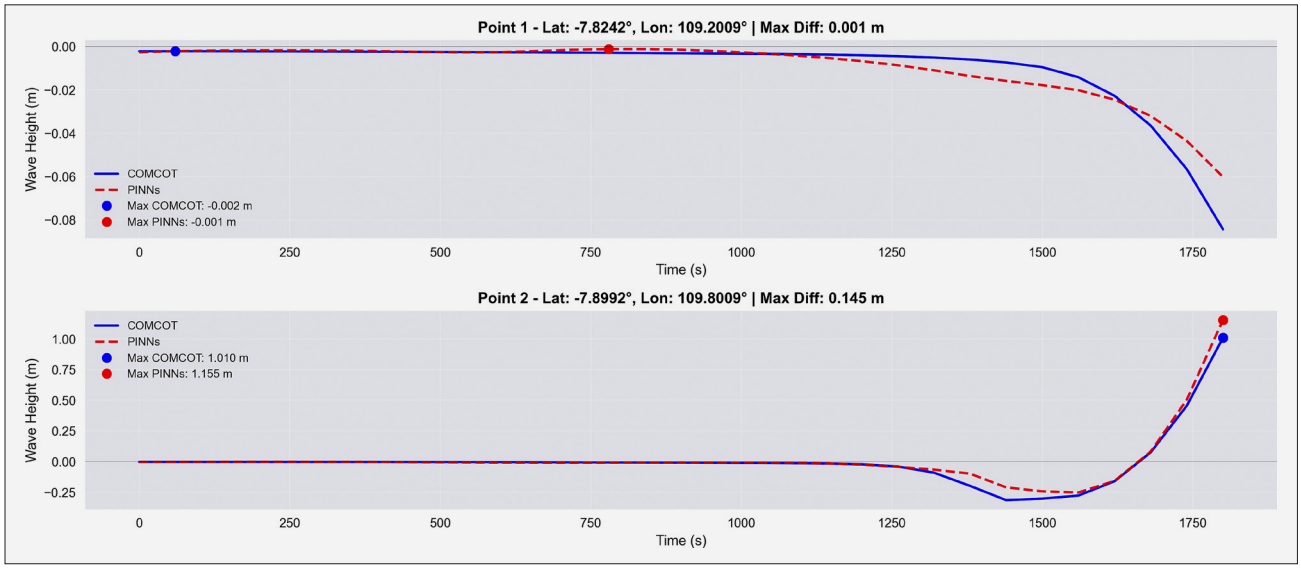


Figure 13: Time series analysis of wave elevation between COMCOT and PINNs at Point 1 and Point 2

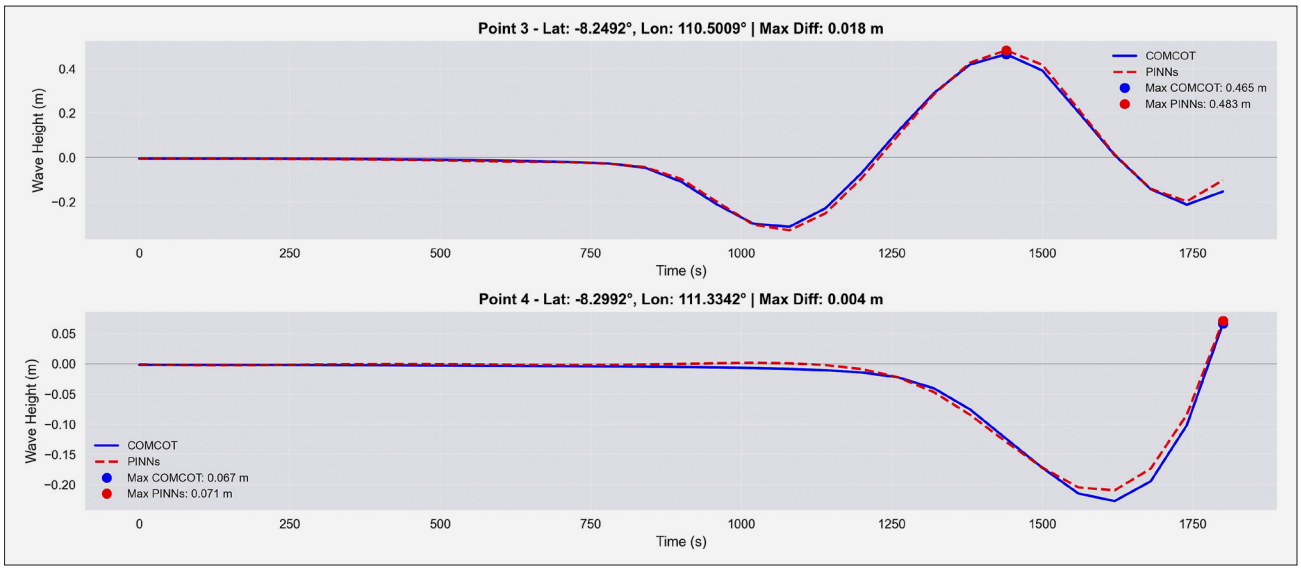


Figure 14: Time series analysis of wave elevation between COMCOT and PINNs at Point 3 and Point 4

The most complete waveform is shown at Point 3, where the distinctive trough–crest–decay sequence is clearly resolved by both COMCOT and PINNs. The peak appears between 1650 and 1700 s, much before the simulation limit, after a sharply negative trough and a steeply ascending crest. Unlike other sites, this early arrival makes it possible to catch the entire wave pattern without truncation. The close match (maximum difference 0.018 m) shows that PINNs are able to accurately reconstruct the wave's dynamic structure, including the post-peak decay and the transition from trough to crest, in addition to its amplitude.

Both models show a smooth, low-amplitude waveform at Point 4, where wave energy drops once more. The profile is partly shortened due to the crest being extremely near to 1800 s, but the agreement is still high with a 0.004 m difference. Both simulations consistently replicate the wave pattern, which is a slow decay followed by a delayed tiny crest.

Overall, the time series comparison demonstrates that both COMCOT and PINNs reflect the key features of tsunami wave history, such as peak arrival timing, amplitude growth toward narrowing coastline geometries, and trough-crest sequences. Point 3 has the strongest waveform completeness since its peak occurs well within the simulation window, and PINNs consistently follow the physics-based COMCOT patterns.

4. Conclusions

Using the 2006 Java tsunami as a case study, this paper effectively illustrates the use of PINNs for predicting tsunami formation and propagation. The physical dynamics controlled by the nonlinear shallow water equations were successfully represented by the PINNs model, which was trained using a combination of supervised data from COMCOT simulations and initial condition data from the Okada model. PINNs can accurately simulate tsunami wave fields at different phases of propagation, from initial deformation to near-shore dynamics, according to a comparison with the traditional COMCOT model. For the majority of the simulation, the model showed minimal absolute errors in open ocean regions and maintained a strong temporal correlation (>0.98). Neural network regularization's smoother gradient representation caused minor differences in coastal regions, but overall performance was quite consistent. The results highlight the potential of PINNs as an adjunct to conventional numerical techniques, including benefits in mesh-free simulation, computational efficiency, and integration of physical laws. A critical operational finding is that, although the offline training phase required approximately 11 hours, the trained model generates a complete tsunami wave elevation map in under one second. This sub-second inference capability represents a transformative advantage for rapid hazard assessment, far exceeding the minute-scale computation times of conventional numerical solvers such as COMCOT. Full operational real-time deployment, where fault parameters are estimated in near-real-time and fed into a pre-trained PINN ensemble, remains an important direction for future research. Recent advances in PINN-based data assimilation suggest a promising pathway for extending this framework to fully autonomous early warning scenarios in the Indonesian context. Future work should also prioritize validation against observed tide gauge records from historical events, adaptive collocation sampling strategies to improve near-shore resolution, and systematic sensitivity analysis of the loss function weighting scheme.

Acknowledgement

The authors would like to thank the reviewers for their constructive comments and suggestions that improved the

quality of this manuscript. We also acknowledge the technical support provided by the Department of Physics, Sebelas Maret University, and the National Research and Innovation Agency (BRIN) during the data processing and computational stages of this research. Additional gratitude is extended to Gadjah Mada University as the host of the Indonesian Collaborative Research (RKI) program under Scheme B for the 2022 fiscal year (Decree No. 490/UN1.P.III/KPT/HUKOR/2022, Contract No. 1543/UN1/DITLIT/Dit-Lit/PT.01.03/2022), and to Sebelas Maret University as the partner institution (Contract No. 872.1/UN27.22/PT.01.03/2022).

Funding

This research was supported by the Indonesian Collaborative Research (RKI) Program under Scheme B for the 2022 fiscal year, through a contract between Gadjah Mada University (Contract No. 1543/UN1/DITLIT/Dit-Lit/PT.01.03/2022) and Sebelas Maret University (Contract No. 872.1/UN27.22/PT.01.03/2022).

5. References

Alkhadhr, S., & Almekkawy, M. (2023). Wave equation modeling via physics-informed neural networks: Models of soft and hard constraints for initial and boundary conditions. Sensors, 23(5), 2792. https://doi.org/10.3390/s23052792.

Arifin, A. Z., Joesidawati, M. I., Herlambang, T., Mizan, S., & Suryanto, A. A. (2025). Simulation of Tsunami Wave Propagation Using the Finite Difference Method for Disaster Early Warning System. Nonlinear Dynamics & Systems Theory, 25(1).

Bihlo, A., & Popovych, R. O. (2022). Physics-informed neural networks for the shallow-water equations on the sphere. Journal of Computational Physics, 456, 111024. https://doi.org/10.1016/j.jcp.2022.111024.

Cho, H. (2024, March). Physics Informed Neural Network for Solving the Shallow Water Equations on Grids of Arbitrary Geometry. In SoutheastCon 2024 (pp. 405-411). IEEE. https://doi.org/10.1109/SoutheastCon52093.2024.10500298.

Darmawan, D. C. S., Sasmi, A. T., & Khoirunnisa, H. (2024, June). Geospatial Information Analysis Based on Numerical Modelling of Tsunami Inundation in Coastal Area of Temon District, Kulon Progo Regency. In IOP Conference Series: Earth and Environmental Science (Vol. 1357, No. 1, p. 012043). IOP Publishing. https://doi.org/10.1088/1755-1315/1357/1/012043.

Hisan, N. K., Prakoso, S. S., Chania, D. F. A., & Arhananta, A. (2020, July). Kontrol Geomorfologi Dan Karakteristik Pantai Terhadap Kerentanan Tsunami Di Yogyakarta International Airport, Kabupaten Kulonprogo. In Prosiding Seminar Teknologi Kebumian dan Kelautan (SEMITAN) (Vol. 2, No. 1, pp. 485-487). https://doi.org/10.31284/j.semitan.2020.1060.

Iordanis, I., Koukouvinos, C., & Silou, I. (2025). On the efficacy of conditioned and progressive Latin hypercube sampling in supervised machine learning. Applied Numerical Mathematics, 208, 256-270. https://doi.org/10.1016/j.apnum.2023.12.016.

Jagtap, A. D., Kawaguchi, K., & Karniadakis, G. E. (2020). Adaptive activation functions accelerate convergence in deep and physics-informed neural networks. Journal of Computational Physics, 404, 109136. https://doi.org/10.1016/j.jcp.2019.109136.

Latifah, A., & Masykuri, A. F. (2021). Pemodelan Tsunami pada Zona Megathrust Pantai Selatan Jawa menggunakan Community Model Interface for Tsunami (ComMIT). In Prosiding Seminar Nasional Fisika dan International Physics Seminar (Vol. 1, No. 1, pp. 318-324).

LeVeque, R. J., George, D. L., & Berger, M. J. (2011). Tsunami modelling with adaptively refined finite volume methods. Acta Numerica, 20, 211-289. https://doi.org/10.1017/S0962492911000043.

Li, J. G., & Wang, P. (2024). Global tsunami modelling on a spherical multiple-cell grid. Ocean Modelling, 192, 102461. https://doi.org/10.1016/j.ocemod.2024.102461.

Li, J., & Li, B. (2022). Mix-training physics-informed neural networks for the rogue waves of nonlinear Schrödinger equation. Chaos, Solitons & Fractals, 164, 112712. https://doi.org/10.1016/j.chaos.2022.112712.

Løvholt, F., Glimsdal, S., Harbitz, C. B., Zamora, N., Nadim, F., Peduzzi, P., & Smebye, H. (2012). Tsunami hazard and exposure on the global scale. Earth-Science Reviews, 110 (1-4), 58-73. https://doi.org/10.1016/j.earscirev.2011.10.002.

Lu, L., Meng, X., Mao, Z., & Karniadakis, G. E. (2021). DeepXDE: A deep learning library for solving differential equations. SIAM review, 63(1), 208-228. https://doi.org/10.1137/19M1274067.

Madsen, P. A., & Schaeffer, H. A. (2010). Analytical solutions for tsunami runup on a plane beach: single waves, N-waves and transient waves. Journal of Fluid Mechanics, 645, 27-57. https://doi.org/10.1017/S0022112009992485.

Moseley, B., Markham, A., & Nissen-Meyer, T. (2020). Solving the wave equation with physics-informed deep learning. arXiv preprint arXiv:2006.11894. https://doi.org/10.48550/arXiv.2006.11894.

National Center for Earthquake Studies. (2017). Buku Peta Sumber dan Bahaya Gempa Indonesia Tahun 2017. Kementerian Pekerjaan Umum dan Perumahan Rakyat, Jakarta.

Okada, Y. (1985). Surface deformation due to shear and tensile faults in a half-space. Bulletin of the Seismological Society of America, 75(4), 1135-1154. https://doi.org/10.1785/BSSA0750041135.

Pohan, A. F., Sismanto, S., Nurcahya, B. E., Lewerissa, R., Koesuma, S., Saputro, S. P., ... & Adhi, M. A. (2023). Utilization and modeling of satellite gravity data for geohazard assessment in the Yogyakarta area of Java Island, Indonesia. Kuwait Journal of Science, 50(4), 499-511. https://doi.org/10.1016/j.kjs.2023.05.016.

Post, J., Wegscheider, S., Mück, M., Zosseder, K., Kiefl, R., Steinmetz, T., & Strunz, G. (2009). Assessment of human immediate response capability related to tsunami threats in Indonesia at a sub-national scale. Natural Hazards and Earth System Sciences, 9(4), 1075-1086. https://doi.org/10.5194/nhess-9-1075-2009.

Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational physics, 378, 686-707. https://doi.org/10.1016/j.jcp.2018.10.045.

Synolakis, C. E., Bernard, E. N., Titov, V. V., Kânoğlu, U. T. K. U., & Gonzalez, F. I. (2008). Validation and verification of tsunami numerical models. Pure and Applied Geophysics, 165(11), 2197-2228. https://doi.org/10.1007/s00024-004-0427-y.

Tanioka, Y., & Satake, K. (1996). Tsunami generation by horizontal displacement of ocean bottom. Geophysical research letters, 23(8), 861-864. https://doi.org/10.1029/96GL00736.

Tian, Y., Ding, S., Huang, L., Su, G., & Chen, J. (2025). Physics-informed neural networks for solving the two-dimensional shallow water equations with terrain topography and rainfall source terms. Water Resources Research, 61(9), e2025WR040052. https://doi.org/10.1029/2025WR040052.

Wang, S., Teng, Y., & Perdikaris, P. (2021). Understanding and mitigating gradient flow pathologies in physics-informed neural networks. SIAM Journal on Scientific Computing, 43(5), A3055-A3081. https://doi.org/10.1137/20M1318043.

Wang, X. (2009). User manual for COMCOT version 1.7 (first draft). Cornell University, 65, 480.

Wells, D. L., & Coppersmith, K. J. (1994). New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. Bulletin of the seismological Society of America, 84(4), 974-1002. https://doi.org/10.1785/BSSA0840040974.

Wijayanti, R., Pasma, G. R., Nurwijayanti, A., & Rachman, R. A. (2022, December). Tsunami Numerical Modelling in Ujung Kulon National Park with the Earthquake Magnitude Scenario of 6.5 and 6.9 Mw. In IOP Conference Series: Earth and Environmental Science (Vol. 1118, No. 1, p. 012070). IOP Publishing. https://doi.org/10.1088/1755-1315/1118/1/012070.

Wuwungan, C., Pasau, G., & Tongkukut, S. H. J. (2021). Pemodelan Perambatan Gelombang Tsunami di Laut Banda Berdasarkan Skenario Gempa 8.0 dan 9.0 Mw. Jurnal MIPA, 10(2), 55-58. https://doi.org/10.35799/jmuo.10.2.2021.34006.

Xing, S. (2023). Analysis of Shallow Water Equation Based on Tsunami Simulations: Evidence from the Pacific Ocean. Highlights in Science, Engineering and Technology, 72, 906-911. https://doi.org/10.54097/1ygn2714.

SAŽETAK

Fizički potaknute neuronske mreže za simulaciju cunamija na otvorenome moru izazvanih potresom na južnoj Javi 2006.

Ovo istraživanje prikazuje usporednu analizu između konvencionalnoga numeričkog modela COMCOT i fizički potaknutih neuronskih mreža (PINN) za simulaciju valova cunamija izazvanih potresom na Javi 17. srpnja 2006. godine (Mw 7,7). Okvir PINN-a bez mreže (*mesh-free*) integrira nelinearne plitkovodne jednadžbe izravno u funkciju gubitka omogućujući simulacije bez prostorne diskretizacije. Model je treniran za predviđanje polja visine valova primjenjujući početne uvjete dobivene iz Okadina modela rasjeda. Rezultati pokazuju da PINN model točno reproducira referentni COMCOT model uspješno hvatajući početno formiranje cunamija i kasniju dinamiku širenja. Ključne kvantitativne metrike, uključujući koeficijente korelacije (iznad 0,98 tijekom većine simulacije) i prostornu distribuciju pogreške (srednja apsolutna postotna pogreška <5 % u otvorenome oceanu), potvrđuju visoku točnost modela. Posebno je važno istaknuti da nakon treniranja PINN model generira cijelo polje visine valova za manje od jedne sekunde, što predstavlja znatnu operativnu prednost u odnosu na klasične numeričke metode temeljene na mreži. Unatoč blagom učinku zaglađivanja u obalnim područjima s velikim gradijentima, studija zaključuje da je PINN pouzdana i točna alternativa bez mreže, s velikim potencijalom za buduću primjenu u brzoj procjeni rizika od cunamija i integraciji podataka u sustave ranoga upozoravanja, pri čemu je potpuna primjena u stvarnome vremenu istaknuta kao ključni smjer budućih istraživanja.

Ključne riječi:
modeliranje, COMCOT, PINN, nelinearna jednadžba plitke vode, strojno učenje

Author's contribution

Sorja Koesuma: conceptualization, methodology, writing - review & editing, project administration, validation, supervision. **Arnida Lailatul Latifah**: conceptualization, methodology, writing - review & editing, validation, formal analysis, investigation. **Angger Zufan Hanggara**: methodology, software, resources, data curation, writing - original draft, visualization. **Ahmad Fauzi Pohan**: writing - review & editing. **Sismanto**: writing - review & editing. **Budi Eka Nurcahya**: writing - review & editing. **Sugeng Purwo Saputro**: writing - review & editing. **Rian Amukti**: writing - review & editing. All authors have read and agreed to the published version of the manuscript.