Deep Learning Approach for Rapid Tsunami Height and Arrival Time Prediction

Elizabeth Su[1]; Lingsen Meng[1]; Chao Liang[2]

[1] University of California: Los Angeles Department of Earth, Planetary, & Space Sciences

[2] Sichuan University-The Hong Kong Polytechnic University Institute for Disaster Management and Reconstruction (IDMR),

Abstract

Accurate and timely tsunami warnings are critical for mitigating loss of life and property. Traditional deterministic simulations of tsunami waveforms using numerical methods such as finite element models can provide detailed predictions but are computationally intensive, making them less suitable for rapid response. For operational warning systems, however, complete waveform details are often unnecessary—what is essential are reliable estimates of maximum coastal wave heights and arrival times.

This study presents a hybrid modeling framework that leverages deep learning to significantly reduce prediction time while maintaining accuracy. We first generate training data using stochastic slip distributions for Mw 7.5–9.0 earthquakes along the Japan Trench, with corresponding maximum wave heights and arrival times computed via high-fidelity numerical simulations. A convolutional neural network enhanced with transformer blocks is then trained to learn the mapping from earthquake slip distributions to coastal tsunami metrics, capturing both local spatial patterns and long-range dependencies.

Once trained, the model can produce near-instant predictions, enabling probabilistic tsunami hazard assessment and rapid damage estimation. This approach demonstrates the potential of AI-driven methods to complement traditional physics-based modeling, accelerating the delivery of actionable tsunami warning information and improving disaster preparedness.

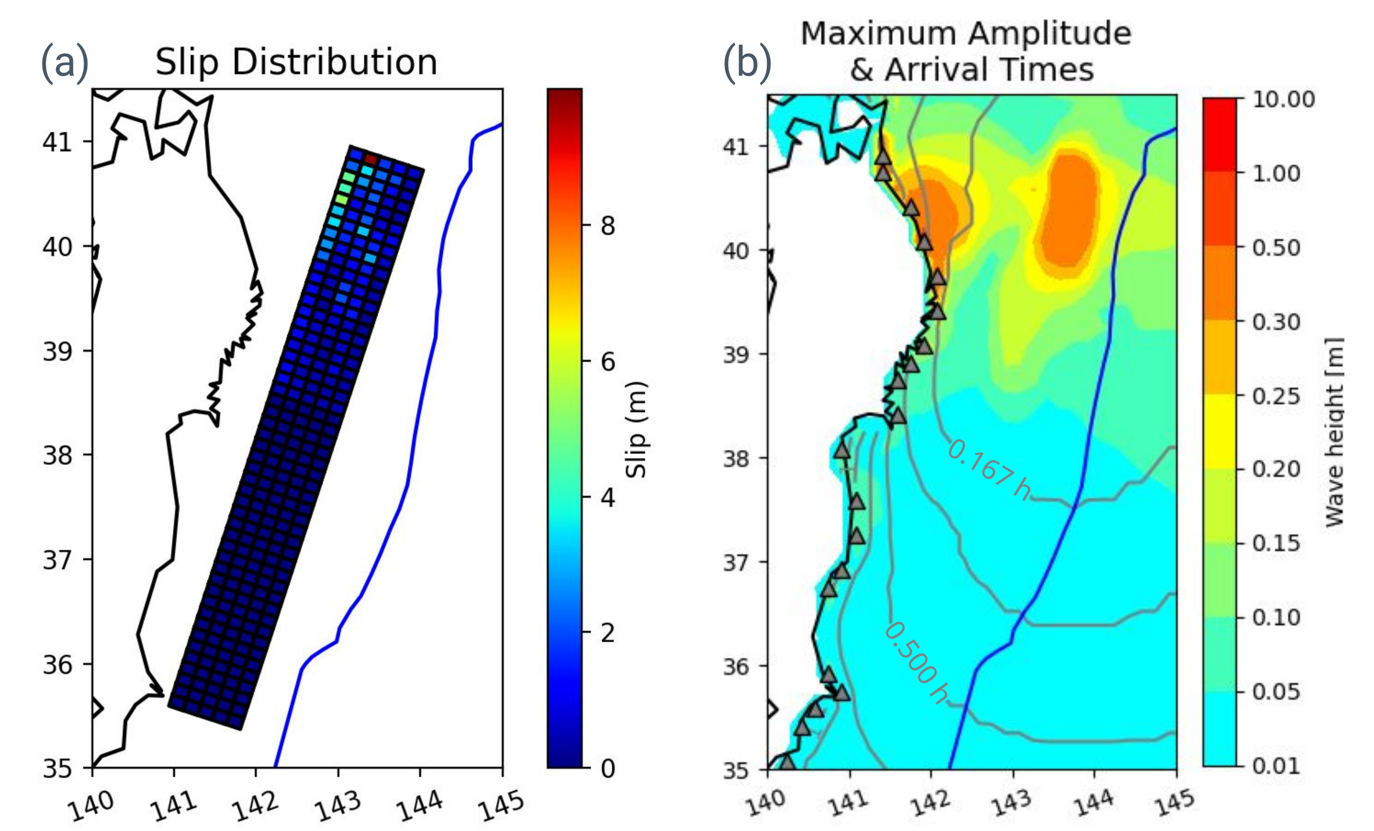

Figure 1: Sample slip distribution and tsunami propagation results used to train the neural network. (a) Stochastic slip model for a Mw 8.0 event along the predetermined fault geometry. Each subfault is colored according to the amount of slip. (b) Maximum amplitude and arrival times based on a numerical simulation for the same Mw 8.0 event. Arrival times are designated by the grey contours every 10 minutes. The output region is shaded according to maximum wave height. Grey triangles are located at each of the 20 coastal points used in training the neural network. The Japan Trench is indicated by a blue line in both figures.

Introduction

This study aims to improve tsunami early warning systems by training a convolutional neural network (CNN) to produce quick, accurate predictions of maximum tsunami wave height and arrival time at specific locations along Japan's coast. Previous studies have used a variety of input parameters (e.g. geodetic observations) within convolutional neural networks to improve the state of tsunami forecasting [1, 2]. Capitalizing on the growing capabilities for carrying out rapid earthquake source inversions, our model uses the slip across a fault plane to generate predictions [3]. By directly training the model with seismic source parameters, this approach has the potential to accelerate tsunami warning predictions for affected coastal populations.

Data Generation

Input slip distributions involved a fixed geometry (strike of 198°, dip of 15°, 250 subfaults) along a fault plane in the Japan Trench. This geometry was derived from the USGS finite fault model of the 2011 Tohoku earthquake [4]. 3350 stochastic slip models were generated with target magnitudes from Mw 7.5-9.0. The actual magnitude for each simulated earthquake varied from the target magnitude slightly, resulting in the distribution shown in Figure 2.

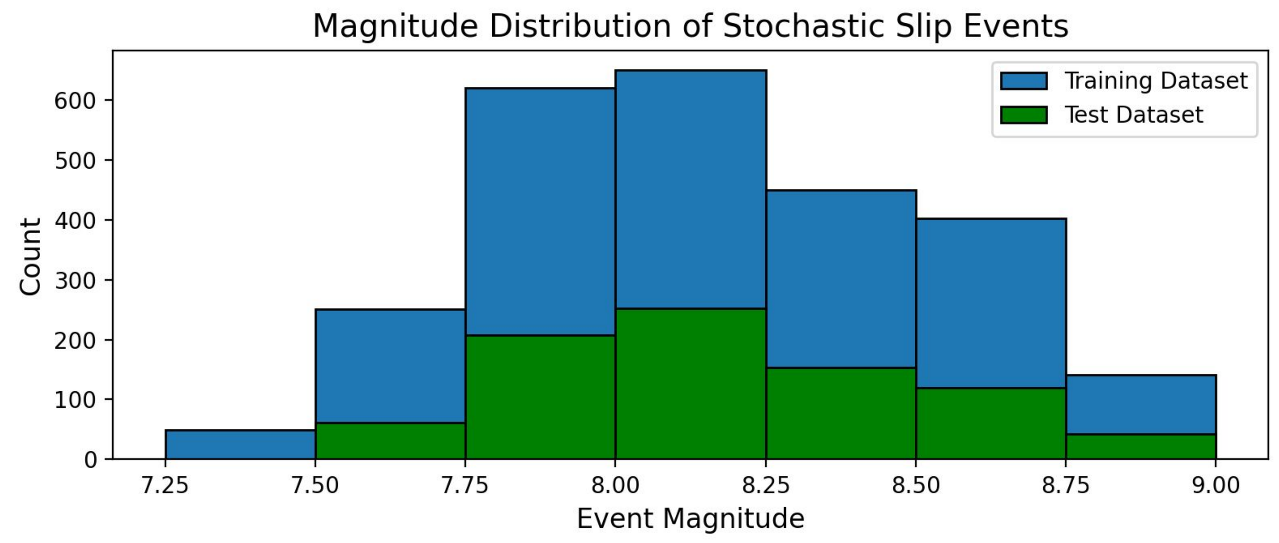

Figure 2: Distribution of earthquake magnitudes in the training and test datasets. Test data comprised of 25% of the entire dataset. This project did not evaluate the model's extrapolation capacity for events outside the range of magnitudes during training.

Each slip distribution was used to create an Okada model of seafloor deformation and subsequently simulate tsunami propagation using depth-averaged shallow water equations with 3 levels of Adaptive Mesh Refinement [5]. The simulation calculated the maximum wave height and arrival time across a regular fixed grid with a spatial resolution of 1/12°.

Input data for the neural network consists of the latitude, longitude, and depth (km) of the subfault centerpoint as well as the slip (m) and rake for each of the 250 subfaults. The output data contained the maximum wave height and arrival time for 20 fixed grid points along the coast of Japan (Figure 1b).

Neural Network Architecture

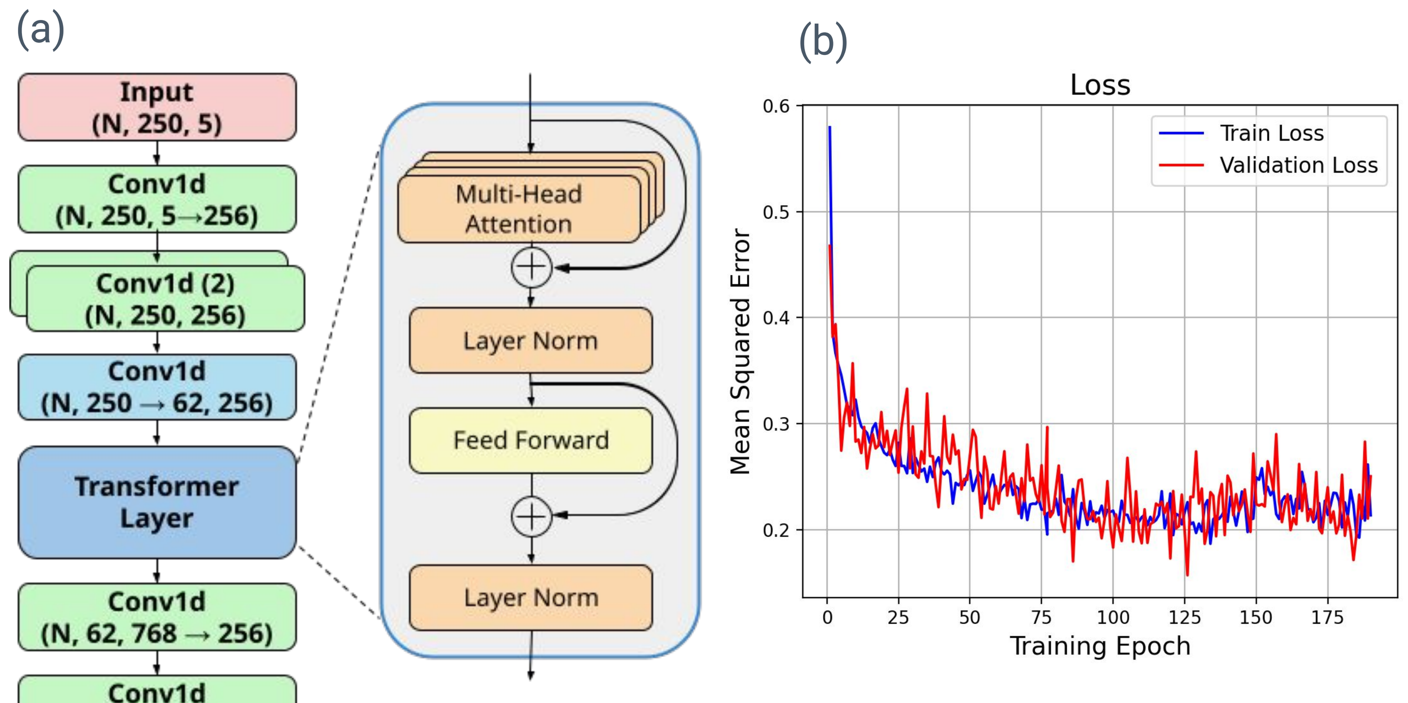

Figure 3: (a) Convolutional neural network architecture. To downsample the data, 1D convolution is applied with a kernel size of 5, stride of 4, and padding of 2 prior to the transformer layer. All other convolution layers use kernels of size 1. A linear interpolation is performed after the data passes through all other layers to output exactly 20 points. ReLu was used as an activation function. (b) Training and validation loss curve over 200 training epochs.

Our model modifies the convolutional neural network structure introduced by Gong & Wang et al in the development of RuptureNet2D, a CNN designed to predict earthquake dynamic rupture parameters [6]. In particular, a convolution layer and linear interpolation layer were added to downsample the input data (information from 250 subfaults) to produce predictions for 20 specific locations along the coastline. The original surrogate model employs 1-dimensional convolution layers and transformer blocks with multi-head self-attention, making this CNN structure well-equipped to identify the non-local trends present throughout our dataset.

Results

The model was trained for 200 epochs using a batch size of 64, dropout value of 0.1, and optimizing with Adam at a learning rate of 1e-4. Mean squared error defined the loss function for the training data (65% of the dataset) and validation data (20% of the dataset), excluding NaN values that may have been output by the tsunami propagation simulation.

Once trained, the model's near-instant evaluation capacity presents a significant reduction in time needed for early warning predictions, compared the computational time of each numerical simulation (~1-6 minutes).

The best model was evaluated on an independent test dataset using 25% of the original observations. Pearson's correlation coefficient was calculated to quantify the error across the testing dataset for each output variable, yielding $R_{time} = 0.8470$ for the arrival time and $R_{height} = 0.7945$ for the maximum wave amplitude. These results suggest that while the model is capable of generating accurate predictions, further training is necessary prior to increasing the complexity of the problem.

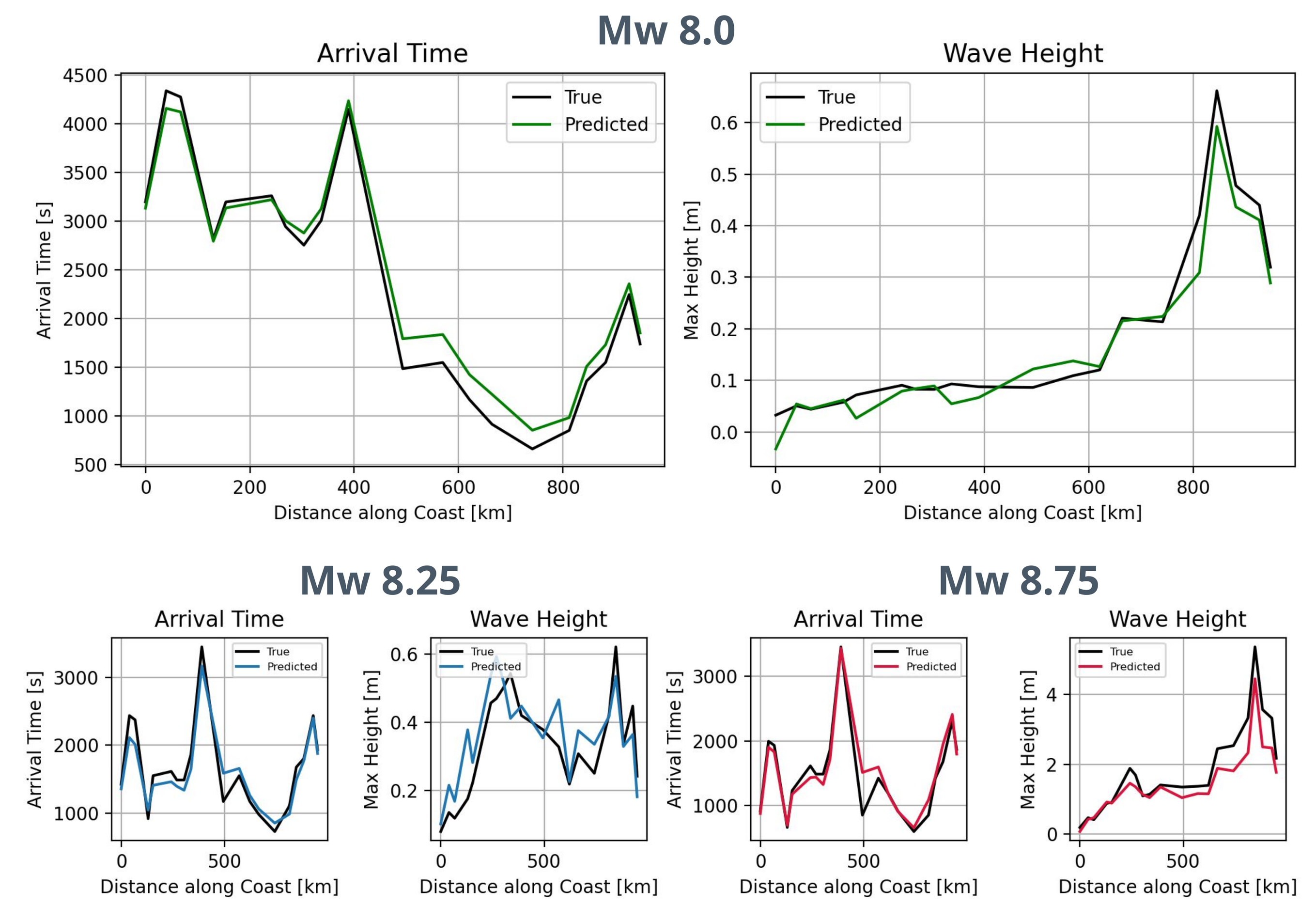

Figure 4: Example outputs from the trained model. The upper charts correspond to the Mw 8.0 event depicted in Figure 1. Distance along the coast is measured from the southernmost location included in the analysis.

Conclusion & Future Work

This work highlights the utility of machine learning in geophysical settings, while underscoring the need for balancing data availability and model complexity. Complex convolutional neural networks are often trained effectively with hundreds of thousands of observations. In this case, the time-intensive data simulation process hindered our ability to maximize the neural network's performance. Additional investigations may also consider increasing the number of predicted outputs to generate more specific and thorough tsunami warnings.

References

1. Min-Jong Song, Yong-Sik Cho, 2024. Early warning for maximum tsunami heights and arrival time based on an artificial neural network, Coastal Engineering, Volume 192, ISSN 0378-3839, https://doi.org/10.1016/j.coastaleng.2024.104563.
2. Makinoshima, F., Oishi, Y., Yamazaki, T. et al, 2021. Early forecasting of tsunami inundation from tsunami and geodetic observation data with convolutional neural networks. Nat Commun 12, 2253. https://doi.org/10.1038/s41467-021-22348-0.
3. Melgar, D., and Y. Bock, 2015. Kinematic earthquake source inversion and tsunami runup prediction with regional geophysical data. J. Geophys. Res. Solid Earth, 120, 3324–3349. doi: 10.1002/2014JB011832.
4. USGS, 2018. M 9.1 - 2011 Great Tohoku Earthquake, Japan. https://earthquake.usgs.gov/earthquakes/eventpage/official20110311054624120_30/finite-fault.
5. Mandli, K.T., Ahmadia, A.J., Berger, M.J., Calhoun, D.A., George, D.L., Hadjimichael, Y., Ketcheson, D.I., Lemoine, G.I., LeVeque, R.J., 2016. Clawpack: building an open source ecosystem for solving hyperbolic PDEs. PeerJ Computer Science. doi:10.7717/peerj-cs.68
6. Gong, Z., Wang, Z., Liang, C., Nienkötter, A., Wang, J., Ren, C., & Peng, X. (2025). RuptureNet2D, a deep neural network based surrogate for dynamic earthquake rupture simulation in two dimensions. Journal of Geophysical Research: Solid Earth, 130, e2024JB030069. https://doi.org/10.1029/2024JB030069

Acknowledgements

This work used computational and storage services associated with the Hoffman2 Cluster which is operated by the UCLA Office of Advanced Research Computing's Research Technology Group. Figures were produced using Python. Stochastic slip models were produced with MudPy and tsunami simulations were carried out with Clawpack Version 5.11.0.

UCLA

U.S. National Science Foundation

U.S. DEPARTMENT OF ENERGY