

JGR Solid Earth

RESEARCH ARTICLE

10.1029/2025JB031981

Earthquake Rupture Dynamics From Graph Neural Networks

Dunyu Liu¹  and Thorsten W. Becker^{1,2,3} 

¹Jackson School of Geosciences, Institute for Geophysics, The University of Texas at Austin, Austin, TX, USA,

²Department of Earth and Planetary Sciences, Jackson School of Geosciences, The University of Texas at Austin, Austin, TX, USA, ³Oden Institute for Computational Engineering & Sciences, The University of Texas at Austin, Austin, TX, USA

Key Points:

- We develop a Graph Network-based Simulator (GNS) for rupture dynamics
- The GNS generalizes to unseen hypocenters, pre-stress levels, and fault sizes with dozens to hundreds training scenarios
- The GNS achieves ~20–40 per-time-step speedup compared to physics-based EQdyna and may help with uncertainty quantification and Bayesian inference

Correspondence to:

D. Liu,
dliu@ig.utexas.edu

Citation:

Liu, D., & Becker, T. W. (2025). Earthquake rupture dynamics from graph neural networks. *Journal of Geophysical Research: Solid Earth*, 130, e2025JB031981. <https://doi.org/10.1029/2025JB031981>

Received 14 MAY 2025

Accepted 24 NOV 2025

Author Contributions:

Conceptualization: Dunyu Liu, Thorsten W. Becker

Data curation: Dunyu Liu

Formal analysis: Dunyu Liu, Thorsten W. Becker

Funding acquisition: Dunyu Liu, Thorsten W. Becker

Investigation: Dunyu Liu, Thorsten W. Becker

Methodology: Dunyu Liu, Thorsten W. Becker

Project administration: Dunyu Liu

Resources: Dunyu Liu, Thorsten W. Becker

Software: Dunyu Liu

Supervision: Dunyu Liu

Validation: Dunyu Liu

Visualization: Dunyu Liu

Writing – original draft: Dunyu Liu

Writing – review & editing: Dunyu Liu, Thorsten W. Becker

Abstract Earthquakes arise from tectonic loading of complex fault systems consisting of heterogeneous material parameters, geometry, rheology, and prestress. All of those are subject to uncertainties, and their interactions and sensitivities for the dynamic rupture problem are incompletely understood. Here, we apply Graph Neural Networks (GNNs) to approximate the behavior learned from more computationally intensive, physics-based (“high-fidelity”) computations to build a GNN-based simulator (GNS) for earthquake rupture dynamics. Given only a minimum input –the hypocenter location– our GNS can reproduce rate-weakening friction governed dynamic rupture behavior, from nucleation to propagation and termination. Outside the training set, the GNS can generalize well to different hypocenter locations, fault sizes, and pre-stress state levels while achieving a factor ~20–40 per-time-step computational speedup. This may allow for more efficient estimates of the mapping from pre-earthquake state, as might be inferred from geodesy, to expected rupture dynamics, for example. By extracting a coarse-grained version of the underlying dynamics, the GNS provides new perspectives to explore the physics of rupture. Further development of GNS may enable new kinds of parameter space exploration and provide surrogates for Bayesian model inference, uncertainty quantification, and optimal experimental design.

Plain Language Summary Large earthquakes pose major seismic hazards, but their source processes are hard to constrain and predict due to many uncertainties. Scientists generally use “high-fidelity,” full physics-based models to study earthquake dynamics, but the high computational cost can limit the exploration of parameters and uncertainties. With the advance of GPU-accelerated computing and deep learning, Graph Network-based Simulators (GNSs) have emerged as powerful tools to model complex physical dynamics. Here, we apply GNS to learn and predict earthquake source dynamics governed by rate-weakening friction. Our GNS can generalize to unseen hypocenters, pre-stresses, and fault geometries. It also runs faster than the traditional physics-based approaches, providing a potential avenue for uncertainty quantification and inferring rupture physics from observations.

1. Introduction

Even though major faults in plate boundaries are mapped relatively well in seismically active regions, large earthquakes continue to surprise us with their complexity. Ruptures arise from tectonic loading of fault systems that are characterized by heterogeneous material parameters, irregular and multi-strand geometry, and variable stress state. Despite progress in incorporating observationally constrained stress conditions and laboratory-derived friction laws for complex rupture scenarios (e.g., Jia et al., 2023; Ramos et al., 2021; Ulrich et al., 2019), many parameters influencing rupture dynamics remain poorly constrained. There are also remaining unknowns on the physical process side, such as to the role of fluids, frictional relationships, and rheology. Currently, there are few modeling frameworks capable of incorporating all relevant physics in a unified way to enable uncertainty quantification or Bayesian model inference for earthquake rupture.

Analysis typically must rely on a limited number of physics-based numerical simulations (PBNS) to explore aspects of the sensitivities of the nonlinear dynamics involved. For example, as for the basic rupture conditions, Andrews and Barall (2011) show how the combination of initial stress distributions and slip-weakening laws (e.g., Day, 1982; Ida, 1972) generate realistic rupture scenarios. Shi and Day (2013) show how fractal fault roughness and rate- and state-friction (Dieterich, 1979; Ruina, 1983) with normal stress dependence (Dieterich & Linker, 1992) reproduces realistic high frequency ground motion. Some of the additional effects explored with

© 2025. The Author(s).

This is an open access article under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

PBNS include the role of non-planar fault geometry and rate-weakening friction (e.g., Dunham et al., 2011; Luo & Duan, 2018) and the effects of fault valving and pore pressure evolution (e.g., Zhu et al., 2020). Moreover, Duan and Oglesby (2006), Liu et al. (2022), and Shaw et al. (2022) highlight how multi-fault geometries and earthquake sequences shape stress evolution and rupture segmentation over time.

PBNS approaches have also been used for Bayesian inversions for pre-stress and friction parameters for single ruptures (e.g., Gallovič et al., 2019; Suhendi et al., 2025; Taufiqurrahman et al., 2022), for example. However, integrating a range of plausible physics into a 3-D model with multiple fault segments while allowing for exploration of both structure and physics uncertainties remains a significant computational and technical challenge. Yet, this is what is ultimately required for hazard estimates, such as when seeking to map from inferred fault loading distributions to likely rupture scenarios.

Very recently, machine learning surrogates based on deep neural networks have started to be used to emulate dynamic ruptures. RuptureNet2D (Gong et al., 2025), trained on 300,000 2-D dynamic rupture simulations, provides a mapping from initial stress, frictional parameters, and hypocenter locations to rupture arrival times and final slip. A reported speedup of $\sim 1,000$ times compared to PBNS is partially achieved by targeting only the final output and bypassing the intermediate steps, and generalization is challenging when extrapolating to fault lengths and additional asperity patches beyond the training set. Tainpakdipat et al. (2025) present another end-to-end projection-type model using Fourier Neural Operators trained on 3,000 scenarios with fractal initial stresses, to map initial stress and frictional parameters to subsampled timeseries of slip rate. Their models achieve a reported $\sim 10^5$ speed up, with the limitation that the predictions are confined to trained nucleation locations and fixed rupture duration.

We explore an alternative deep learning approach using Graph Neural Networks (GNNs) here. GNNs have emerged as promising tools to capture the temporal evolution of high-dimensional physical systems simulations by emulating all intermediate system states (e.g., Lam et al., 2023; Pfaff et al., 2020; Sanchez-Gonzalez et al., 2020) and show strong generalization capabilities in terms of initial conditions, feature complexities, and model domain sizes. GNNs are particularly well suited for representing particle- and mesh-based systems, which many PBNS for rupture dynamics, such as finite-element and finite-difference methods, adopt (e.g., Harris et al., 2018; Liu & Duan, 2018; Premus et al., 2020; Ulrich et al., 2019; Wang & Day, 2020; Withers et al., 2018). GNN-based simulators (GNSs) can capture a wide range of physical processes, including turbulent flow and solid and granular mechanics (Choi & Kumar, 2024; Pfaff et al., 2020; Sanchez-Gonzalez et al., 2020) and other nonlinear systems such as weather (Lam et al., 2023).

Besides the typical finding that GNS can provide computational speedup compared to PBNS, GNS' inductive nature allows generalization to system sizes not seen during training (Thangamuthu et al., 2022). Moreover, incorporating physics-informed bias, such as inertia, in deep learning frameworks has also been shown to improve performance in the learning of dynamics (Choi & Kumar, 2024; Thangamuthu et al., 2022), and GNS show remarkable generalization and physics abstraction capabilities (Choi & Kumar, 2024; Pfaff et al., 2020; Thangamuthu et al., 2022).

Here, we develop a GNS for earthquake dynamics on a 2-D fault interface embedded in 3-D rock volume where the training data are generated from dozens to hundreds of 3-D PBNS rupture dynamics models and fault sliding is governed by friction with strong rate-weakening. We explore the GNS performance and generalization capabilities for three different types of setups –unseen hypocenters, pre-stress amplitudes and distributions, and fault sizes, and discuss GNS strengths, limitations, and potential future applications to earthquake physics and hazard.

2. Methods

2.1. The High Fidelity, PBNS Model

We use the CPU-parallelized finite-element software EQdyna (Duan, 2012; Duan & Oglesby, 2006; Liu & Duan, 2018; Liu et al., 2022) to generate earthquake rupture scenarios (sequences of fault sliding velocity states, or slip rates) at “high fidelity,” to train the GNS and as ground truth for testing and validation. EQdyna solves the 3-D elastodynamic equations of motion with a friction law on fault interfaces for rupture propagation including wave radiation. EQdyna has been benchmarked extensively, for example, within the *SCEC Spontaneous Rupture Code Verification Project* (Harris et al., 2009, 2018). Here, we choose the SCEC benchmark TPV104 as the

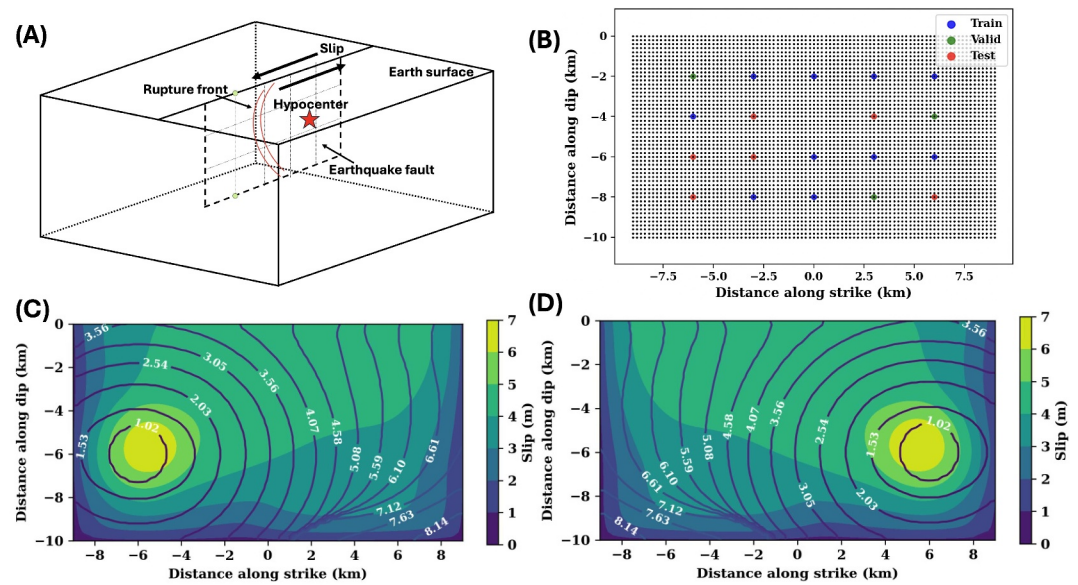


Figure 1. (a) Model setup for earthquake rupture dynamics. A fault interface is embedded in a 3-D elastic volume, and the fault walls move relative to each other, with relative displacement defined as slip. Slip evolution is influenced by initial stress conditions, frictional mechanisms and other parameters such as hypocenter location. (b) Hypocenter locations used to construct Data set D1 for GNS with training (blue), validation (green), and test (red) sets. Gray dots represent fault nodes in the reference model. (c–d). Final slip distribution and rupture time contours from two scenarios in the training set. White numbers indicate rupture arrival times in seconds.

reference, because the associated strong rate-weakening friction law produces significant slip rate changes during rupture. Moreover, this test case has been used for benchmarking by many rupture codes, ensuring comparability and easy reproducibility of our study. We then vary hypocenter locations and initial stress levels and stress anomaly distributions.

Figure 1a shows our reference setup, which models earthquake dynamics for a vertical, strike-slip, planar fault embedded in a 3-D isotropic, linear elastic half space, and breaking the surface. Slip evolution is governed by rate- and state-friction with strong-rate weakening (Table 1), following Equations 1–7 in the SCEC TPV104 benchmark description (strike.scec.org/cvws/download/SCEC_validation_slipLaw.pdf).

We use a fault dimension of 18×10 km, 200 m on-fault grid size, and there are a total of 4,500 fault nodes in each scenario. The 3-D model domain boundaries are extended 20 km away from the nearest fault edges, respectively, and radiated seismic waves are damped by absorbing boundary conditions (Liu & Duan, 2018). We start with homogeneous initial stresses and frictional parameters; at the hypocenter, nucleation is driven by imposing an additional shear stress over 1 s and within a 3 km radius, as per Equations 14–16 in the benchmark. Because the fault is uniformly velocity weakening, rupture front reflections from fault boundaries are expected. This setup is intended to test the GNS' capability to learn and predict the dynamics beyond simple, pulse-like extension of the rupture, such as the effect of boundaries.

Each computational scenario lasts 15 s, which is long enough for rupture to propagate over the whole fault and die out. With a time increment of 0.0167 s, each scenario contains 900 time steps. At the beginning of each scenario, the earthquake is under forced nucleation with minimal slip rate information on the fault. Therefore, we use rupture states after initiation at time $t = 1.2$ s to train the GNS models; this gives 826 time steps per scenario.

2.2. Graph Neural Networks and Message Passing

2.2.1. Graphs and Earthquake System Rupture State

Graph formalisms can be applied to emulate physical system interactions (e.g., Battaglia et al., 2016; Sanchez-Gonzalez et al., 2020). In an elastic solid, a frictional interface which governs earthquake rupture dynamics can be modeled as a graph $G = (V, E)$ where a set of vertex variables V ($v_i \in V$) captures the kinematics of

Table 1
Key Parameters for the Reference, Physics-Based Numerical Model

Fault length along strike (km)	18
Fault width along dip (km)	10
Grid size (m)	200
Initial shear stress (MPa)	40
Initial normal stress (MPa)	120
Density (kg/m ³)	2,670
S-wave velocity (km/s)	3.464
P-wave velocity (km/s)	6
Rate- and state- friction parameters	
Reference friction coefficient	0.6
Reference slip rate (m/s)	10 ⁻⁶
<i>a</i>	0.01
<i>b</i>	0.014
Characteristic length, <i>L</i> , (m)	0.4
Friction threshold for strong rate weakening, <i>f_w</i>	0.2
Slip rate threshold for strong rate weakening, <i>V_w</i> , (m/s)	0.1
Initial state variable value, <i>Ψ</i>	0.563591842632738
Additional peak shear stress for nucleation (MPa)	45
Nucleation radius (km)	3
Artificial nucleation duration (s)	1

discretized vertices on the interface, and a set of edges E ($e_{ij} \in E$) represents physical interactions between neighboring vertices through their relative locations (v_i and v_j).

In our finite-element approach (e.g., Liu & Duan, 2018; Liu et al., 2022), we discretize the fault interface into vertices V_{FE} which are connected by edges E_{FE} . We define a rupture state at time t as $S_t(\dot{u}, x)$, for all the vertices' slip rates $\dot{u}_i$ and their locations x_i on the fault interface ($i \in V_{FE}$). Slip rate is the relative velocity between two fault walls bordering the interface. We then use a graph $G = (V_{FE}, E_{FE})$ to represent the rupture state and its evolution.

2.2.2. Graph Neural Networks (GNNs) and Message Passing

A Graph Neural Network (GNN) is an essential element in GNS architecture. The GNN takes a graph $G = (V_{FE}, E_{FE})$ as the input and outputs a new graph $G' = (V'_{FE}, E'_{FE})$ with updated $\dot{u}'_i$ and e'_{ij} through message passing (Gilmer et al., 2017). Message passing is the process by means information is propagated through the graph structure, for example, capturing the transport of energy and momentum. In the dynamic rupture context, the GNN takes graph $G = (V_{FE}, E_{FE})$ as the input that describes the current system state S_t and updates to a new $G' = (V'_{FE}, E'_{FE})$, which after decoder and updaters, returns the next state S_{t+1} .

Message passing contains three steps:

$$1. \text{ Message construction } e'_{ij} = \phi_{\Theta_\phi}(v_i, v_j, e_{ij}) \quad (1)$$

$$2. \text{ Message aggregation } \bar{v}_i = \sum_{j \in N(i)} e'_{ij} \quad (2)$$

$$3. \text{ Update to vertex } v'_i = \gamma_{\Theta_\gamma}(v_i, \bar{v}_i) \quad (3)$$

here, Θ_ϕ and Θ_γ are a set of learnable parameters; ϕ_{Θ_ϕ} is the message construction function, a matrix operation, that takes in the attribute vectors of the connecting vertices (v_i and v_j) and those of the edge connecting the two

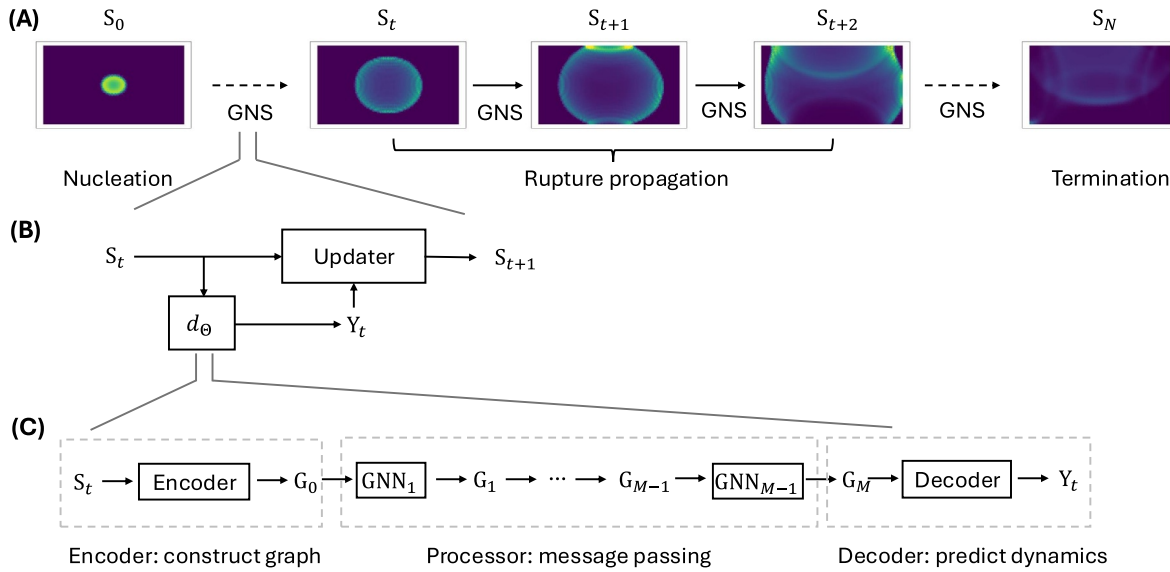


Figure 2. Concept of using Graph Network-based Simulator (GNS) to predict earthquake rupture dynamics with nucleation, rupture propagation, and termination, and the GNS architecture (modified from Choi and Kumar (2024) and Sanchez-Gonzalez et al. (2020)). (a) With an initial state S_0 as input, the GNS predicts the next state and autoregressively generate state predictions for entire ruptures. (b) The GNS predicts dynamics, Y_t , using the learned dynamics model, d_θ , and then uses an updater to predict next state, S_{t+1} . (c) The structure of the learned dynamics model d_θ , which includes an encoder to construct graph G_0 from input state S_0 , a processor using message passing through M Graph Neural Networks (GNNs) to update graphs, and a decoder to extract dynamics information Y_t from the last graph G_M .

vertices ($e_{i,j}$) and returns an updated edge feature vector $e'_{i,j}$. Then, messages are aggregated toward a single vector $\bar{v}_i$ where $N(i)$ is the set of sender vertices j related to vertex i . The last step is to update vertex features using the aggregated message $\bar{v}_i$ and the current attribute vector v_i , through the update function $\gamma_{\Theta_\gamma} = \Theta_\gamma(\bar{v}_i + v_i)$.

We refer readers to Figure 2 in Choi and Kumar (2024) for an illustration of the message passing concept in the GNN context. After message passing, the graph is updated to $G' = G(V'_{FE}, E'_{FE}) = G(v'_i, e'_{i,j})$. To propagate the information further into the network, multiple message passing steps can be used, where the involved “hyper-parameter” choices affect model prediction performance and computational speed. In a supervised learning setup, learnable parameters are optimized through training.

2.3. The Surrogate, GNS Model

On top of the GNNs is the GNS architecture to learn and predict earthquake rupture state evolution. We adapt an open-source, mesh-based GNS, *MeshNet* (Choi & Kumar, 2024; Kumar & Choi, 2023; Kumar & Vantasel, 2023) and proceed to present details on GNS input, as well as the encoder, processor, decoder, and updater, as shown in Figure 2.

2.3.1. Input

The input to the GNS, $s'_i \in S'$, is a vector containing current slip rate $\dot{u}'_i$ and on-fault location along strike and dip, $x'_i = [x'_{i,\text{strike}}, x'_{i,\text{dip}}]$, respectively, for all vertices. The current state is $s'_i = [\dot{u}'_i, x'_i]$. Sanchez-Gonzalez et al. (2020) show that for mesh-based GNS, only the current state is sufficient to predict the next one, without the need for memory of past states; this is different from the particle-based GNS examples for granular flow (Choi & Kumar, 2024). Vertex attributes can be expanded to include other features such as initial shear stress level, τ'_i , or frictional parameters, pore pressure, or different frictional laws.

2.3.2. Encoder

The vertex and edge encoders, ϵ^ν_θ and ϵ^e_θ converts s'_i into the vertex and edge attribute vectors v'_i and $e'_{i,j}$ and embed them into a latent graph $G_0 = (V_0, E_0)$ where $v'_i \in V_0$ and $e'_{i,j} \in E_0$.

$$v_i^t = \varepsilon_{\Theta}^v(\dot{u}_i^t), e_{i,j}^t = \varepsilon_{\Theta}^e(x_i^t) \quad (4)$$

A two-layered, 128-dimensional multi-layer perceptron (MLP) is used for the ε_{Θ}^v and ε_{Θ}^e , respectively. The vertex encoder ε_{Θ}^v uses only slip rates $\dot{u}_i^t$ to make predictions without the current position information. The updated positions of vertices are predicted using the last positions and slip rates by an explicit Euler step.

2.3.3. Processor

Starting from the initial graph G_0 , the processor handles message passing (Equations 1–3) for M steps (i.e., message passing steps; $G_0 \rightarrow G_1 \rightarrow \dots \rightarrow G_M$) and returns the updated graph G_M . We use two-layered 128-dimensional MLPs for both the message construction function, $\phi_{\Theta,\theta}$, and vertex update function, $\gamma_{\Theta,\gamma}$, and element-wise summation for the message aggregation in Equations 1–3.

2.3.4. Decoder

The decoder δ_{Θ}^v extracts the dynamics $y_i^t \in Y_t$ from the vertices v_i^t using the final graph G_M following $y_i^t = \delta_{\Theta}^v(v_i^t)$. A two-layer 128-dimensional MLP is used for δ_{Θ}^v .

2.3.5. Updater

The dynamics y_i^t , analogous to slip acceleration, are used to predict the next time-step slip rate, which is updated using Euler integration to get the next time-step location x_i^{t+1} . As noted by Choi and Kumar (2024), this updater imposes the inductive biases of an inertial frame to force the GNS to learn the interaction dynamics to improve learning efficiency. A traditional neural network must learn both the update function and the interaction dynamics. Ideally, this implies that the GNS learns the actual dynamics, that is, a coarse-grained version of local physical interactions, regardless of absolute vertex positions. This approach permits the GNS to generalize beyond the training fault plane size, hypocenter locations, stress levels and distributions, as will be shown below.

The task then is to predict the next state, $S_{t+1}(\dot{u}, x)$, from the current, $S_t(\dot{u}, x)$, using the GNS (Figure 2a). Following Sanchez-Gonzalez et al. (2020) and Choi and Kumar (2024), Figure 2b shows the structure of the GNS which uses a learned dynamic model, d_{Θ} , to predict dynamics, Y_t , analogous to slip acceleration, from an input state, S_t , and the updater to estimate the next state, S_{t+1} . Figure 2c shows the structure of d_{Θ} , which contains an encoder to construct a graph G_0 from the input state S_0 , a processor using message passing (Gilmer et al., 2017) through M GNNs to update graphs, and a decoder to extract dynamics information, Y_t , from the last graph, G_M .

2.4. Reference Data Sets

We prepare two data sets to train the GNS focusing on different objectives and a third data set to test GNS performance on more complex cases. Data set D1 has homogenous initial shear stress of 40 MPa on the fault but variable hypocenter locations. Our goal is to test whether the GNS can predict the right rupture dynamics even though hypocenter locations are different from those in the training data. Figure 1b shows the distribution of hypocenters for rupture scenarios in the training, validation, and testing sets, each of which contains 10, 3, and 6 scenarios, respectively. Figures 1c and 1d show final slip distributions and rupture front arrival time contours from two scenarios in the training set as an example.

Data set D2, in addition to various hypocenters, contains heterogeneous initial stress cases, here initially chosen for simplicity to be in the form of 4×4 km asperities at 35, 45, 50, and 55 MPa stress levels, representing 10% of all the fault vertices. Figures 3a–3e show the five hypocenter locations and associated locations of asperities with 35/55 MPa initial stress in the training set for M2, which totals at 30 scenarios. Figure 3f shows hypocenters and stress asperities of selected scenarios in the test set, whose asperity stresses (45/50 MPa) differ from those in the training set. We want to see if the GNS trained on D2 can generalize to unseen stress amplitudes and asperity locations, in combination with unseen hypocenter locations, and predict the associated modulations of rupture speed. To explore the effect of the size of training set, we also train model M3 on additional scenarios from data set D2; 37 scenarios are selected for each asperity stress level, which has five hypocenter locations. With four asperity stress levels, the total number of scenarios in the training set is 148 scenarios.

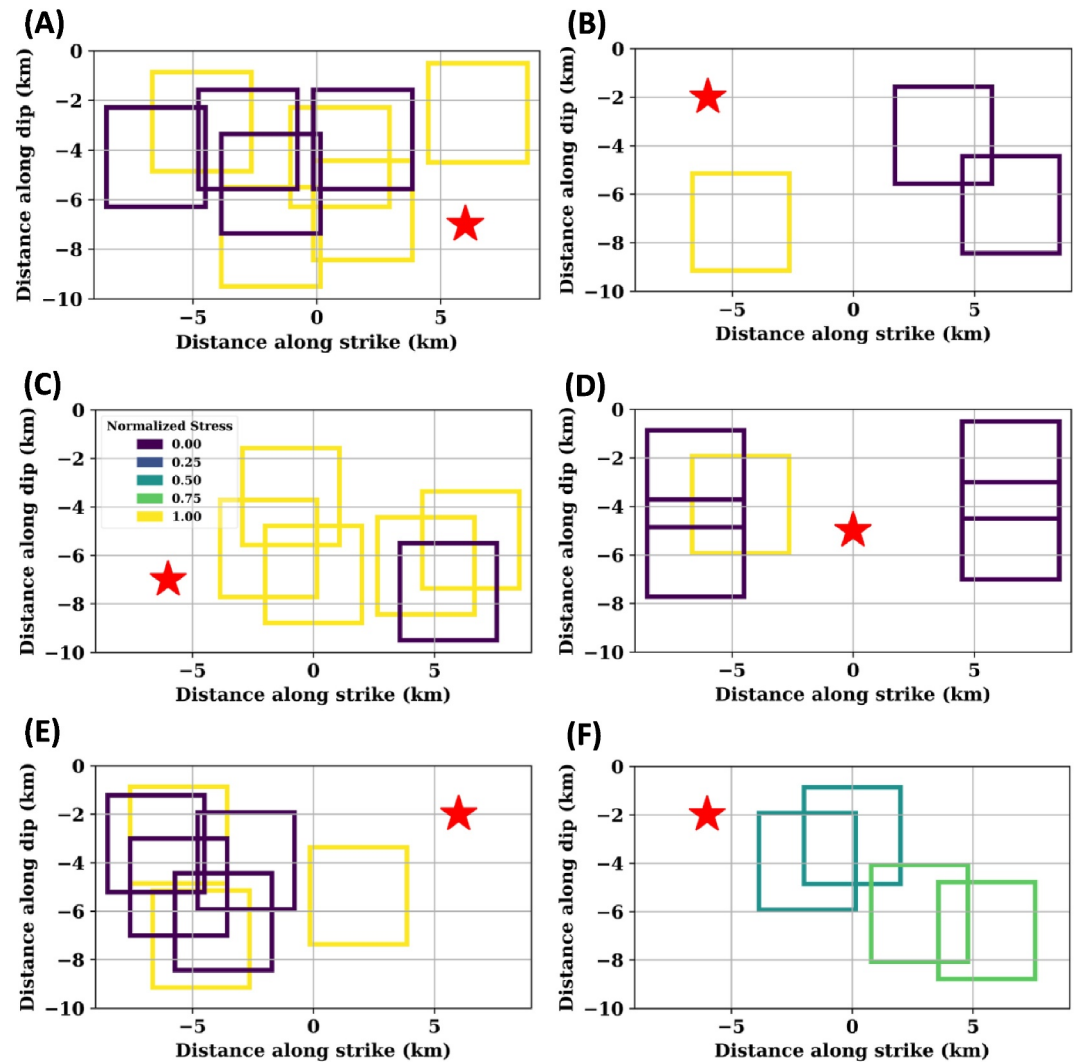


Figure 3. Hypocenter and stress anomaly (i.e., asperity) locations (boxes) and normalized asperity stress levels at 0, 0.25 (outside of asperity), 0.5, 0.75, and 1, which linearly corresponds to 35, 40 (outside of asperity), 45, 50, and 55 MPa. (a–e) 30 scenarios used in the training set, which contains only normalized asperity stress levels at zero and unity. (f) Initial stress setups in selected scenarios from the test set with normalized stress levels at 0.5 and 0.75, which are unseen in the training set.

To test the limits of the GNS' generalization capability, we build another Data set D3 with 15 scenarios using fractal-type, irregularly distributed, “rough” initial stress. There are five scenarios with different hypocenter locations for each roughness R of 0.1, 0.5, and 0.9, using the diamond square algorithm (Fournier et al., 1982), for initial stresses as shown in Figures 11a, 12a, and 13a. We test GNS M2 and M3 on D3 to check if the GNS can capture more complex stress distributions based on learning relatively simple, and different spatial character, features in the training set, and examine the impact of training set size.

2.5. Training and Prediction

We train the GNS' learnable parameter sets to minimize the mean square error (MSE) between predicted normalized slip acceleration proxy y_i^t and normalized ground truth slip acceleration $\hat{u}_i^t$, computed by derivatives of slip rate $\dot{u}_i^t$, for all the fault vertices based on a standard gradient-based optimizer, Adam (Kingma & Ba, 2014) run over training time steps and epochs. GNS performance is affected by the choices of hyperparameters, such as learning rate, number of message passing steps in GNNs, and noise levels. We start from hyperparameters optimized for granular dynamics (Choi & Kumar, 2024) including the number of message passing steps of 10, learning rate of 10^{-4} , batch size of 2, and Gaussian distributed noise with standard deviation of 0.02 m/s (Table 2).

Table 2
GNS Hyperparameters

Model	M1/M2	M3
Learning rate	10^{-4}	3×10^{-5}
Batch number	2	8
Number of message passing steps	5/10/15	5/10/15
Gaussian noise level	0.02/0.005	0.02/0.005
Number of hidden layers in multi-layer perceptron	2	2
Dimension of hidden layers in multi-layer perceptron	128	128

These hyperparameters achieve a good balance between predicting performance and training cost for our problem as well (cf. Pfaff et al., 2020). We also explore the effect of hyperparameter choices (Table 2) on training and prediction performance (Section 3.7).

During the prediction phase of the GNS, that is, *rollout*, we take an initial state which contains relatively small slip rate distributions (with peak ~ 1 m/s) inside the nucleation zone at $t = 1.2$ s simulation time, and the GNS then autoregressively generates the entire time sequence of earthquake rupture.

Figure 4a–4c show the training and validation loss curves for GNS M1, M2, and M3 trained on Data sets D1 (10 scenarios), D2 (30 scenarios), and D2 (148 scenarios), respectively. For M1, the training and valid loss show rapid descent until 1 million steps and a slower decrease afterward. There is no apparent overfitting (where we would expect validation loss increase and training loss to keep decreasing) over the whole training process. However, we do not expect significantly improved performance after 3 million steps. We thus consider the GNS for M1 trained at 3 million steps.

For M2, the training loss follows a similar trend to that in M1, with improvements diminishing after 1 million steps. Again, there is no sign of overfitting, and we choose the GNS for M2 trained at 3 million steps. For M3 with a lower learning of 3×10^{-5} and larger batch size of 8, the oscillations of losses over training steps are smaller. We choose M3 trained at 2.7 million steps as the training improvement diminishing with more steps. Interestingly, the strong oscillations appear not affect GNS performance. We also train M2 at a same slower learning of 3×10^{-5} , but it appears not affect the performance and oscillations. Adaptive learning rates or alternative optimizer choices may help suppress the oscillations and improve training stability or convergence speed.

3. Results

3.1. Predicting Rupture Dynamics From Unseen Hypocenters

We applied the M1 trained GNS to predict rupture propagation from unseen hypocenter locations. For this, we provide an initial state with slip rate distribution near the hypocenter and the GNS then generates the entire time sequence of rupture states autoregressively. Figure 5 shows rupture time contour comparisons between GNS predictions and PBNS ground truth from four scenarios. We report three metrics, root mean square error of slip

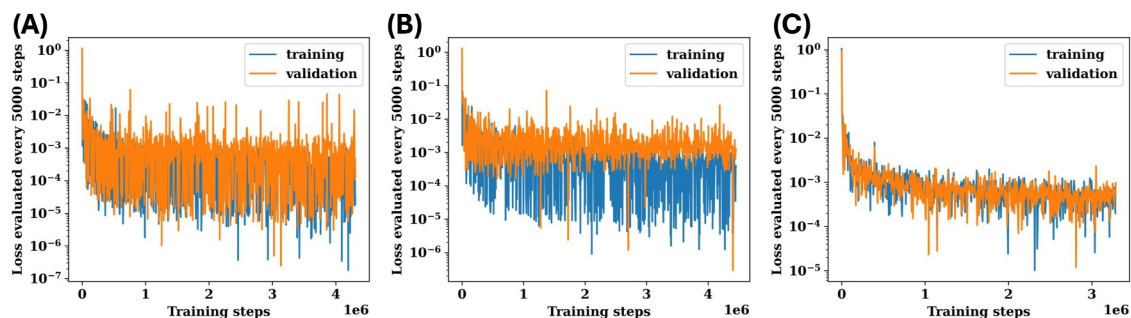


Figure 4. Training and validation loss for data set M1 (a), M2 (b), and M3 (c) tabulated every 5,000 steps.

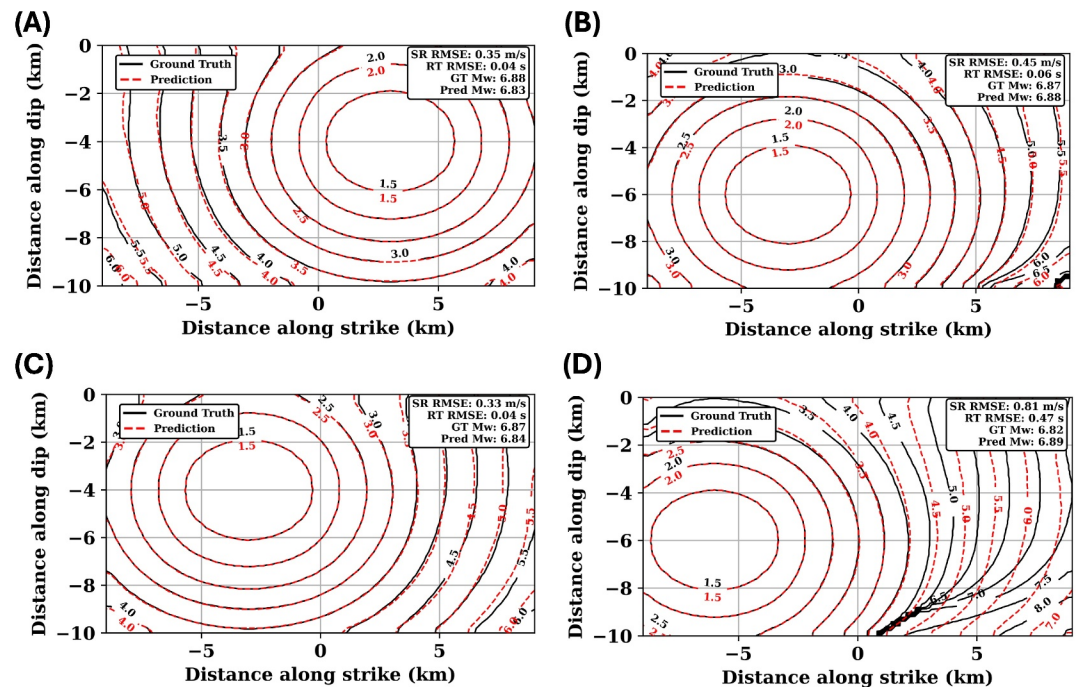


Figure 5. Rupture time contour comparisons between GNS predicted dynamic ruptures (red) and PBNS ground truth (black) for four cases with hypocenter locations unseen in the training set. Numbers are seconds since initiation when the rupture front arrives at that location and slip rate rises above 0.1 m/s. SR RMSE: Slip rate root mean square error for the slip rate time series on all fault vertices between GNS prediction and ground truth. RT RMSE: Rupture time root mean square error for rupture arrival times on all fault vertices between GNS prediction and ground truth. GT Mw: Ground truth event moment magnitude. Pred Mw: GNS predicted magnitude.

rate time series over all the fault vertices (SR RMSE), root mean square error of rupture arrival times over all the fault vertices (RT RMSE), and the difference between predicted and ground truth event moment magnitude, M_w , respectively.

The GNS predicts rupture front arrivals well overall, showing RT RMSE of 0.04–0.06 s for Figures 5a–5c, and SR RMSE of 0.33–0.45 m/s, where peak slip rates are in the range of 20–30 m/s. Figure 5d shows a scenario where rupture terminates early close to the bottom of the fault, where rupture time contours merge. The GNS prediction deteriorates where ruptures terminate early, showing RT RMSE of 0.47 s and SR RMSE of 0.81 m/s. Early rupture termination is a rare feature in the training set; this indicates that the GNS' performance gets worse for sparsely represented features. In terms of event magnitudes, all four models perform well, showing magnitude errors between 0.01 and 0.07. This means the GNS learns to preserve the total earthquake moment, although some rupture details for early termination scenarios are not well captured.

Figures 6a–6d show slip rate snapshots at times $t = 2.2, 3.2, 4.2$ and 5.2 s after rupture initiation for scenario Figure 5a. The GNS can predict the location and amplitude of the localized high slip rate rupture fronts well. Early into the rupture, the nucleation phase is captured accurately (Figures 6a), and Figure 6b shows that the reflection from the free surface and right boundary are also modeled accurately. Later into the rupture, the weaker reflection from the bottom of the fault and the surface reflection traveling to the middle depth of the fault are likewise captured (Figure 6c), along with the termination phase of rupture where slip rate drops to zero (Figure 6d). Moreover, the GNS can predict the termination of ruptures with zero slip rate across the fault (not shown).

3.2. Predicting Rupture Dynamics From Unseen Stress Heterogeneity

In a first step towards increased complexity, we explore how flexibly the structure of learned interactions can accommodate spatial variations in initial stress conditions. For this, we first apply the trained GNS M2 to scenarios with a single asperity of both unseen location and stress levels of 45/50 MPa (0.5/0.75 normalized stress). Figure 7 shows rupture time contour comparisons, together with SR RMSE, RT RMSE, and magnitude errors.

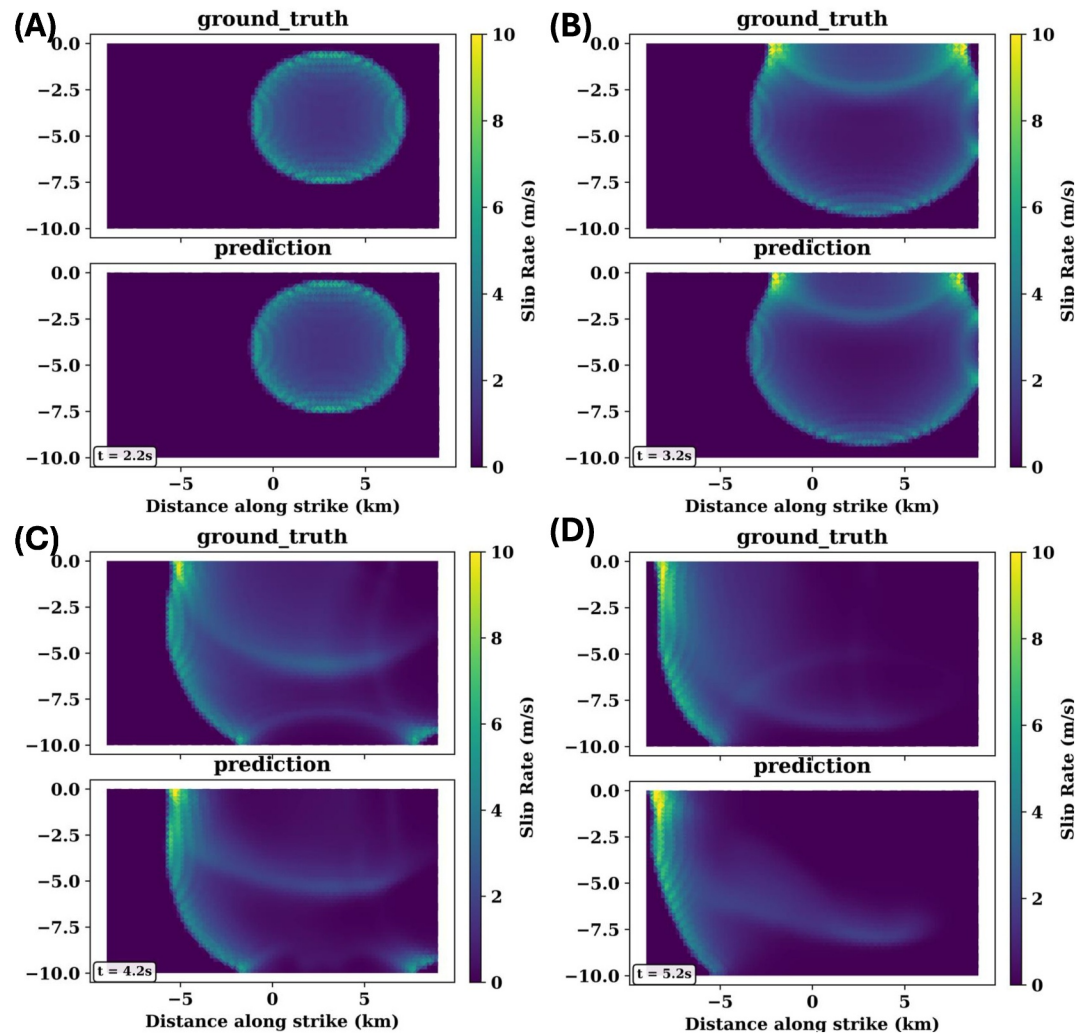


Figure 6. Slip rate snapshots of GNS prediction and PBNS ground truth at times $t = 2.2, 3.2, 4.2,$ and 5.2 s after rupture initiation.

Remarkably, to us, the GNS captures both the arrivals of rupture fronts at asperities and amplitudes of accelerations from these unseen stress levels and locations, showing SR RMSEs in the range of 0.42–0.63 m/s and RT RMSEs of 0.05–0.09 s, and magnitude errors no larger than 0.05. This indicates that heterogeneous material properties can be incorporated into the learned network. Should hold more broadly, a wide range of classes of heterogeneity, from fault roughness to friction parameters, could be considered within the GNS framework.

We then apply the GNS to scenarios with checkerboard-type multiple stress asperities, a more complex heterogeneity pattern, at 35/55 MPa stress with an unseen hypocenter. Figure 8 shows rupture time contour comparisons. The GNS captures rupture acceleration and deceleration caused by asperities reasonably well in both scenarios, showing SR RMSEs of 0.57–0.72 m/s and RT RMSEs of 0.11s and magnitude error no larger than 0.05.

Figures 9 and 10 show slip rate time series at asperity centers for models as in Figures 8a and 8b, respectively. The first major peaks in terms of amplitude and arrival time are well predicted by GNS. In Figure 9, stations along-strike and dip coordinates $(-1, -7)$ and $(3, -7)$ appear to miss the small, secondary peaks later in the rupture. However, a larger secondary peak is captured at station $(-1, -3)$. In Figure 10, all secondary peaks are well captured. This indicates that the GNS, as trained, may occasionally miss secondary peaks at specific sites, and derived products such as ground shaking from secondary phases would be underestimated. However, the

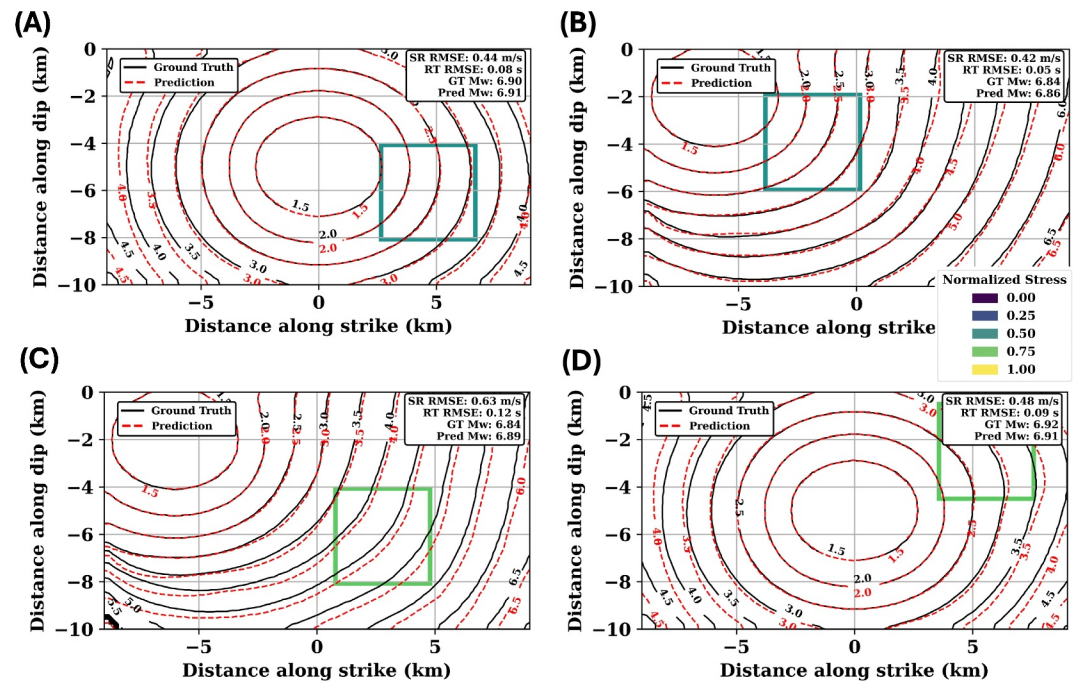


Figure 7. Rupture time contour comparisons between GNS predicted dynamic ruptures (red) and PBNS ground truth (black) for four scenarios with unseen stress level and location of asperities (boxes) in the training set. Labels and error metrics else as in Figure 5, see there for description.

dominant control on ground shaking, earthquake moments, are well predicted for both scenarios, as shown in Figure 8.

3.3. Rupture Dynamics Predictions From Unseen Fractal Initial Stress

To explore the limit of the GNS generalization capability further, we apply GNS M2, which is only trained on two end-member stress amplitudes with a single rectangular asperity, to fractal-type distributed stresses, with amplitudes varying from 35 to 55 MPa. Figure 11 shows rupture time contour comparisons between M2 predictions and PBNS ground truth for ruptures with unseen hypocenter locations, and initial stress levels and patterns with roughness in the diamond square algorithm of $R = 0.1$. For Figures 11b and 11c, the GNS predicts RT RMSE of 0.11–0.16s and SR RMSE of 0.57–0.58 m/s, and magnitude errors of 0.07 and 0.01. For Figure 11d, where early rupture termination occurs, GNS prediction deteriorates to SR RMSE of 0.92 m/s and RT RMSE of 0.95 s, but moment magnitude error is still within 0.06.

Figure 12 shows the initial stress distribution and rupture time contour comparisons for a rougher initial stress with $R = 0.5$. Error metrics SR RMSEs are 0.68 and 0.66 m/s and RT RMSEs are 0.07 and 0.11s for B and C, respectively. Figure 12d shows a larger RT RMSE of 0.38 s, and SR RMSE of 0.82 m/s. Again, the magnitudes are well predicted with errors no larger than 0.07.

For initial stresses with an even higher roughness of $R = 0.9$ (Figure 13a), M2 still predicts reasonable rupture dynamics and magnitudes (Table 4, D3 cases T10–T14), but the SR RMSE, RT RMSE, and magnitude errors get consistently larger. The results indicate that the features in the relatively small training set for M2 are not comprehensive enough in their spatial character to train the GNS for more complex dynamics. Using M3, a GNS with an enlarged training set with 148 scenarios, but still constant amplitude, rectangular asperities, the predictions improve significantly, with RT RMSE no larger than 0.11 s and SR RMSEs of 0.46–0.76 m/s (13B–D and Table 4). This indicates that there are no fundamental limitations with the GNS approach but that the spectral character of fault zone heterogeneity affects the training process and strategy.

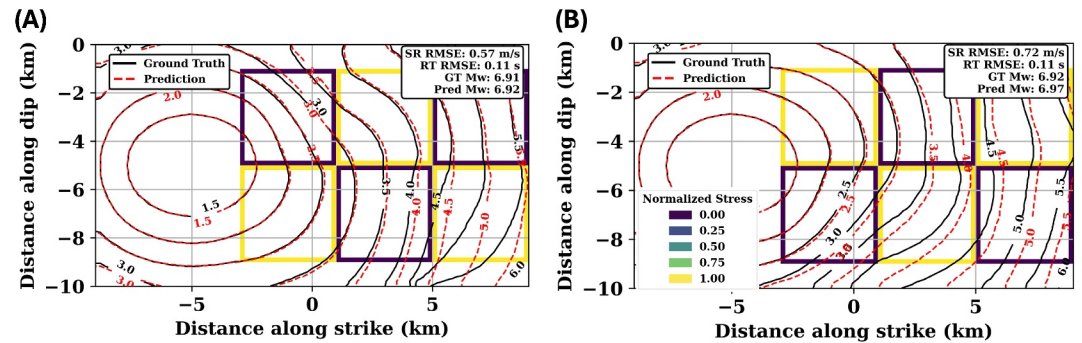


Figure 8. Rupture time contour comparisons between GNS predicted dynamic ruptures (red) and PBNS ground truth (black) for two scenarios with checkerboard multiple asperities not found in the training set. Labels and error metrics else as in Figure 5, see there for description.

3.4. Generalization to Larger Fault Geometry Size

To explore geometrical extrapolation capabilities, we apply GNS M1 to a fault of 40 km length along strike, more than twice the original 18 km training length. At the same spatial resolution, the larger fault model contains an additional 122% fault vertices. While prediction errors for rupture arrival times start to accumulate for long duration state prediction, the rupture time contours generally compare well with a RT RMSE of 0.12 s (Figure 14). The SR RMSE is 0.61 m/s, at a similar level to the scenarios with complex heterogeneous stress levels. The magnitude error is 0.09. This substantiates that the inductive nature of a GNS allows generalization to out of training model geometries (Pfaff et al., 2020). We also apply the GNS to a smaller fault size of 10 km by 5 km along strike and dip, respectively. The SR RMSE of 0.21 m/s, RT RMSE of 0.01 s, and magnitude error of 0.03 for D1.T_small (Table 3) show that the GNS can generalize to smaller fault geometry, confirming its geometrical interpolation capability.

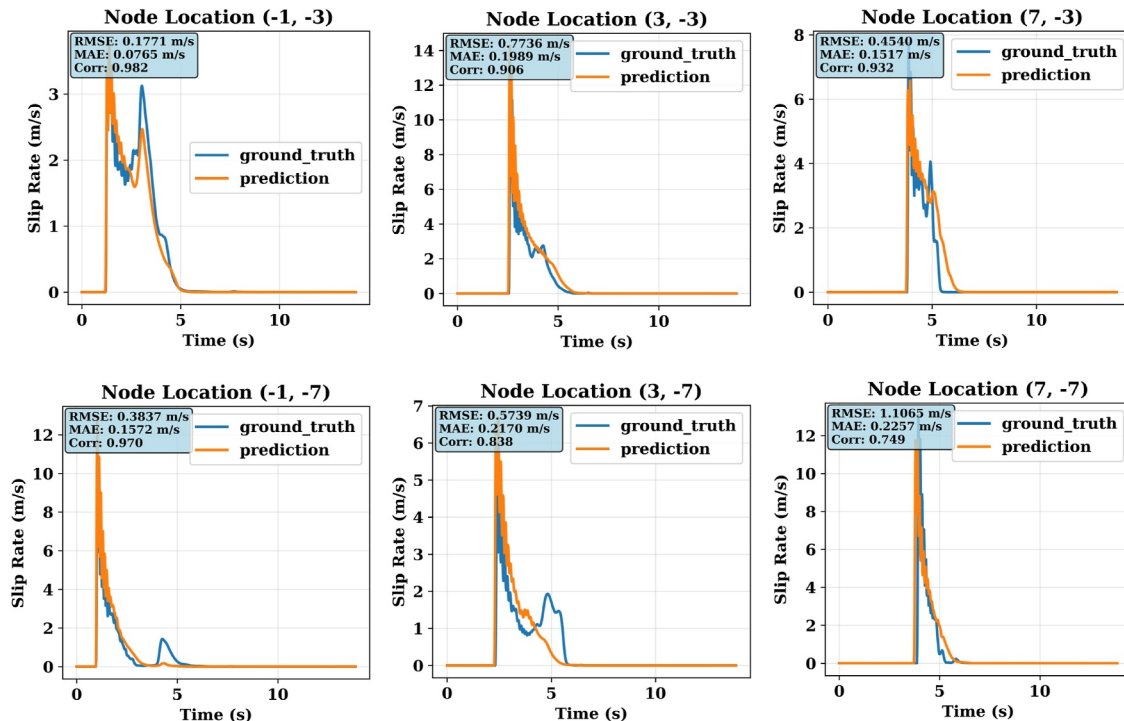


Figure 9. Comparisons of slip rate evolutions at the center of the asperities of the model shown in Figure 8a. Prediction is from the GNS, ground truth from the PBNS. RMSE: root mean square error for each time series; MAE: mean average error for each time series; Corr: correlation between the predicted and ground truth time series.

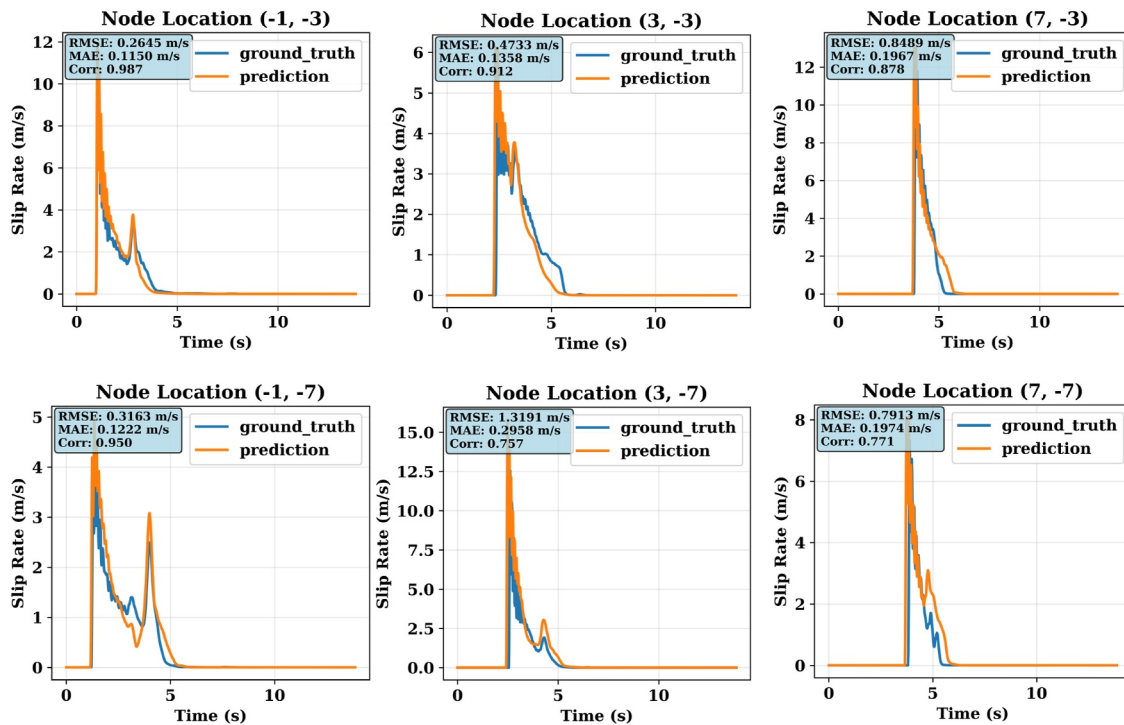


Figure 10. Comparisons of slip rate evolutions at asperity centers in the model in Figure 8b, else as in Figure 9.

3.5. Backward Compatibility

Model M2 is more general, or sophisticated, than M1 given the stress heterogeneity involved in training, and one might ask if M2 is overfit, or backward compatible with scenarios with less complexity, for example, homogeneous initial stresses but different hypocenter locations. We apply M2 to the D1 test data set and the larger fault test, and Table 3 shows misfit metrics for five testing scenarios. The errors are similar to the M1 case, which indicates that M2 is indeed backward compatible with models with less complexity.

3.6. The Impact of Enlarged Training Set

We intend to keep the training set size as small as possible, and as shown in Section 3.3, M2 with 30 scenarios approaches its limit in predicting dynamic ruptures with high fractal-type initial stresses. Table 4 shows misfit metrics for M2 and M3 predictions with initial stresses of different fractal roughness and hypocenter locations. For the scenarios with roughness $R = 0.1$, M2 is competitive to M3. However, with increased complexity of $R = 0.5$ and $R = 0.9$, M3 consistently performs better than M2. The results substantiate that the training set size needs to match the complexity of the prediction tasks.

3.7. Sensitivity Analyses of Hyperparameters

3.7.1. Message Passing Steps

We apply M2 to the D1 test case T_large using different message passing steps of 5, 10, and 15. Table 5 shows misfit metrics. The model with 5 message passing steps gives large SR RMSE and RT RMSE, which indicates that it is insufficient to capture the rupture dynamics. The predictions using models with 10 and 15 message passing steps match the ground truth well. Interestingly, performance of the model using 15 message passing steps is worse compared to that with 10 steps. This agrees with findings by Choi and Kumar (2024) and Sanchez-Gonzalez et al. (2020) for other physical systems. The results are consistent with the expectation that the physics for dynamic ruptures are dominated by local and short-range interactions, spanning around 10 grid cells for our experiments.

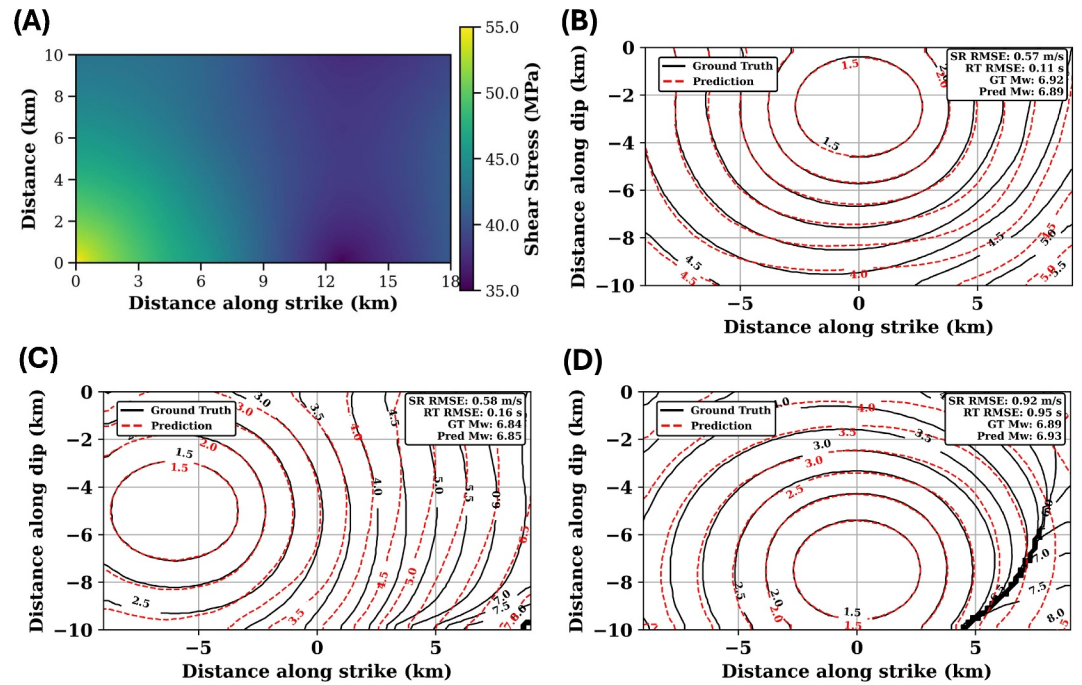


Figure 11. (a) Smooth fractal-type initial stress with roughness in the diamond square algorithm of $R = 0.1$. (b–d) Rupture time contour comparisons between GNS M2 predicted dynamic ruptures (red) and PBNS ground truth (black) for three scenarios with unseen stress levels and complex patterns and hypocenter locations in the training set. Labels and error metrics else as in Figure 5, see there for description.

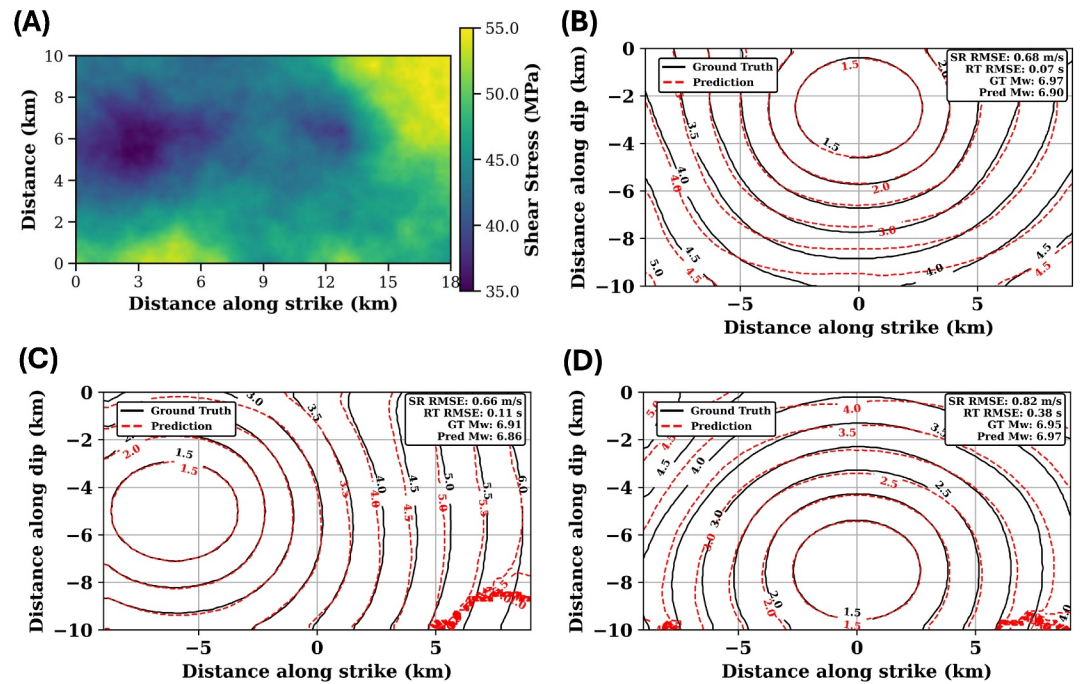


Figure 12. (a) Fractal initial stress with intermediate roughness of $R = 0.5$. (b–d) Rupture time contour comparisons between GNS M2 predicted dynamic ruptures (red) and PBNS ground truth (black) for three scenarios with unseen stress levels and complex patterns and hypocenter locations in the training set. Labels and error metrics else as in Figure 5, see there for description.

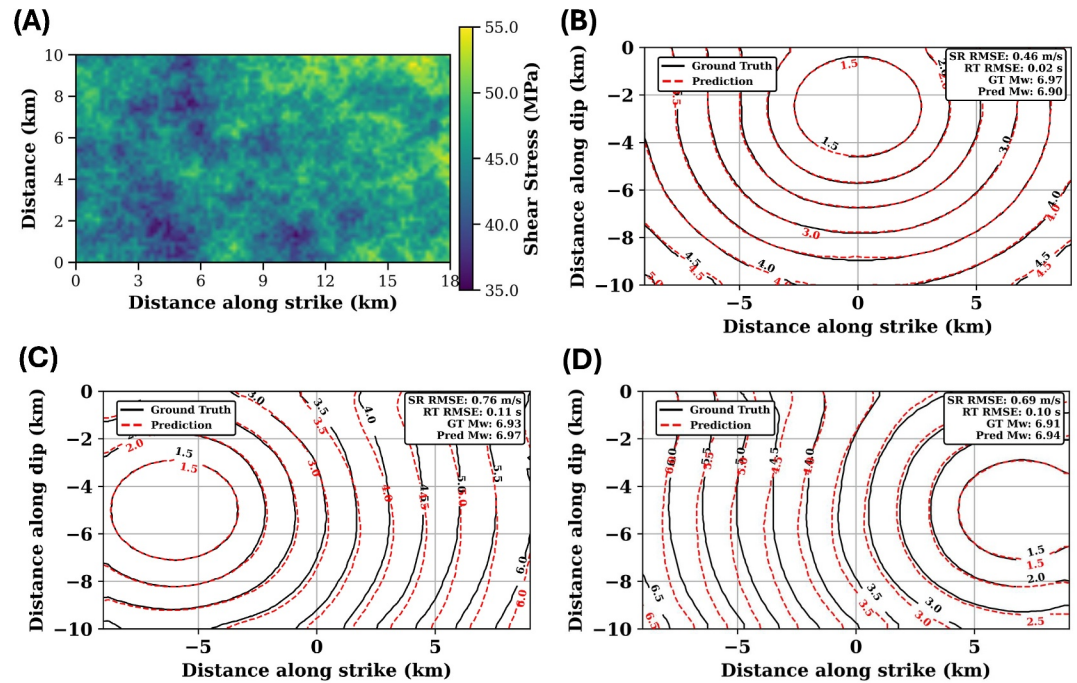


Figure 13. (a) Fractal type initial stress with high roughness $R = 0.9$. (b–d) Rupture time contour comparisons between GNS M3 predicted dynamic ruptures (red) and PBNS ground truth (black) for three scenarios with unseen stress levels and complex patterns and hypocenter locations in the training set. Labels and error metrics else as in Figure 5, see there for description.

3.7.2. Gaussian Noise Level Added in Training

We explore two Gaussian noise standard deviation levels, 0.02 (default) and 0.005 m/s (0.2% and 0.05%, respectively, in terms of ~ 10 m/s slip rate at rupture fronts in this study). Lam et al. (2023) show that a proper level of noise added in the training trajectories are necessary to stabilize the GNS predictions. For M2 with

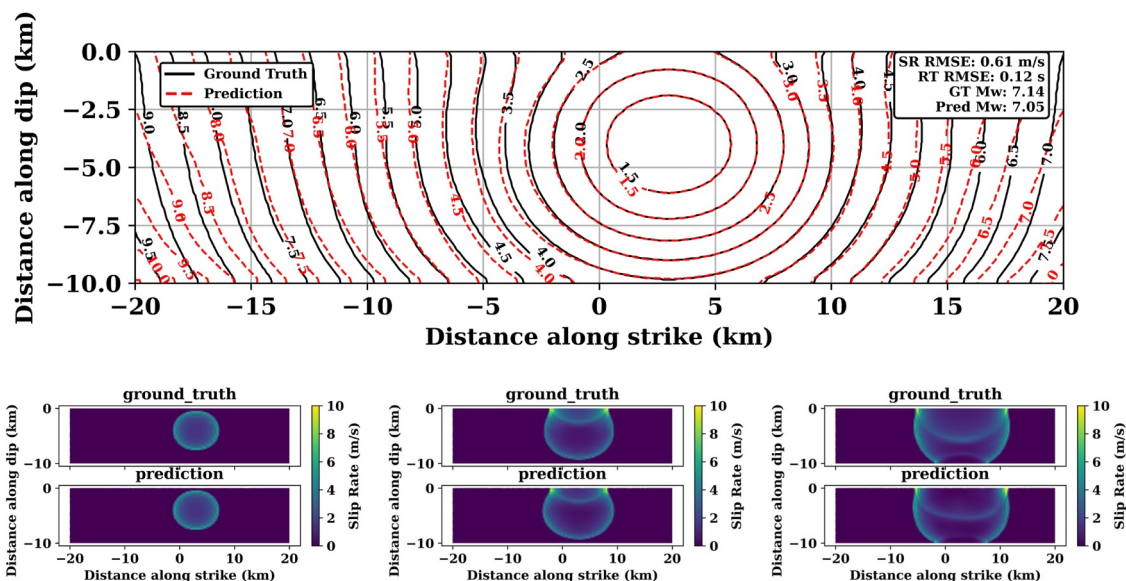


Figure 14. Rupture time contour comparisons between GNS predicted dynamic ruptures (red) and PBNS ground truth (black) for a larger fault exceeding the dimensions found in the training set. Contours and labeling as in Figure 5. Slip rate snapshot comparisons are shown for three time steps.

Table 3

Root Mean Square Error of Slip Rate Time Series Over All Fault Vertices (SR RMSE) and Root Mean Square Error of Rupture Time Arrivals (RT RMSE) and Moment Magnitude Error Between GNS Predictions Using M1 and M2 on Data Set D1 and Corresponding PBNS Ground Truths

Model Metrics	M1			M2		
	SR RMSE (m/s)	RT RMSE (s)	M_w error	SR RMSE (m/s)	RT RMSE (s)	M_w error
D1.T0 (Figure 5a)	0.35	0.04	0.05	0.39	0.05	0.0
D1.T1 (Figure 5b)	0.45	0.06	0.01	0.49	0.17	0.01
D1.T2 (Figure 5c)	0.81	0.47	0.07	0.93	0.79	0.09
D1.T3 (Figure 5d)	0.33	0.04	0.03	0.41	0.05	0.01
D1.T_large (Figure 14)	0.61	0.12	0.09	0.51	0.07	0.05
D1.T_small	0.21	0.01	0.03	0.33	0.02	0.06

0.005 m/s noise, the predictions indeed fail for half of scenarios generating SR RMSE at ~ 100 m/s; this confirms that a $\sim 0.2\%$ level of noise, or so, is necessary for stable GNS performance.

3.7.3. Learning Rate and Batch Size

For M1 and M2, which have 10 and 30 scenarios in the training sets, respectively, we use a fast learning rate of 10^{-4} and a small batch size of 2. Given the modest complexity involved, such learning rates and small batch size do not lead to overfitting. Although the training and validation losses oscillate widely, the model performance appears robust for most of the tasks. For M3, which has 148 scenarios involved, a batch size of 2 does not allow for enough complexity to be incorporated in each epoch, and the learning rate of 10^{-4} leads to overfitting and poor prediction. We thus use a learning rate of 3×10^{-5} and a batch size of 8, which significantly improves model performance (Table 4).

Table 4

Root Mean Square Error of Slip Rate Time Series Over All Fault Vertices (SR RMSE) and Root Mean Square Error of Rupture Time Arrivals (RT RMSE) and Moment Magnitude Error Between GNS Predictions Using M2 and M3 for Data Set D3 With Different Fractal Initial Stresses and Corresponding PBNS Ground Truths

Model Metrics	M2			M3			Better model
	SR RMSE (m/s)	RT RMSE (s)	M_w error	SR RMSE (m/s)	RT RMSE (s)	M_w error	
Low stress roughness $R = 0.1$							
D3.T0	0.57	0.11	0.03	0.35	0.04	0.01	M3
D3.T1 (Figure 11b)	0.58	0.16	0.01	0.71	0.27	0.08	M2
D3.T3 (Figure 11c)	0.92	0.95	0.04	0.82	0.88	0.05	M3
D3.T4 (Figure 11d)	0.53	0.08	0.01	0.73	0.15	0.05	M2
Intermediate stress roughness $R = 0.5$							
D3.T5 (Figure 12b)	0.68	0.07	0.07	0.46	0.03	0.07	M3
D3.T6 (Figure 12c)	0.66	0.11	0.05	0.63	0.09	0.02	M3
D3.T7	1.08	0.37	0.01	0.88	0.13	0.07	M3
D3.T8 (Figure 12d)	0.82	0.38	0.02	0.88	0.12	0.06	M3
D3.T9	0.9	0.21	0.07	0.93	0.33	0.03	M3
High stress roughness $R = 0.9$							
D3.T10 (Figure 13b)	0.65	0.06	0.1	0.46	0.02	0.07	M3
D3.T11 (Figure 13c)	0.93	0.17	0.02	0.76	0.11	0.04	M3
D3.T12 (Figure 13d)	3.2	0.28	0.19	0.69	0.1	0.03	M3
D3.T13	0.99	0.13	0.09	0.82	0.11	0.02	M3
D3.T14	0.94	0.16	0.07	0.66	0.24	0	M3

Table 5

Root Mean Square Error of Slip Rate Time Series Over All Fault Vertices (SR RMSE) and Root Mean Square Error of Rupture Time Arrivals (RT RMSE) and Magnitude Error Between GNS Predictions Using M2, With Different Message Passing Steps, for the Large Homogeneous Test Rupture and Corresponding PBNS Ground Truths

Model message passing steps/Metrics	SR RMSE (m/s)	RT RMSE (s)	M_w error
5	3.66	1.17	0.25
10	0.51	0.07	0.05
15	0.70	0.17	0.07

3.8. Computational Performance

Table 6 shows computational cost per timestep of the GNS on a single Nvidia A100 GPU and the PBNS implemented by means of the EQdyna code, using 8 CPU cores on a workstation with AMD EPYC 7543, both architectures were released in 2020. The GNS run times are ~ 29 –41 times faster than the ground truth simulations by EQdyna. This PBNS approach should be very efficient given the explicit time integration, under-integrated hexagons with hourglass control, and no solving of system of equations (e.g., Liu & Duan, 2018). The GNS still achieves an order of magnitude, ~ 19 -fold speedup compared to the PBNS for 200 m resolution models, and a ~ 41 -fold for 100 m resolution, that is, finer resolved models should benefit from the GNS even more.

However, a fair comparison between ML and high-fidelity type approaches is difficult. This is true in general, and here in particular. For example, we compare a PBNS EQdyna run using eight CPUs with the GNS which uses a single GPU. The GNS predicts dynamics on a 2-D surface, the fault interface, only, while EQdyna needs to solve the full physics equations for the 3-D volume surrounding the fault. A GPU-accelerated PBNS might achieve 10 fold speedup compared to single CPU (e.g., Premus et al., 2020), for example, a suggestion we cannot directly test as EQdyna currently only runs on CPUs. Either way, accepting what we suggest is roughly an order of magnitude speedup of the GNS compared to the PBNS for what might be a typical application is certainly not as striking an advantage as the performance reported for the mapping-type ML approaches. However, we note that we try to be conservative when stating speedup because we consider all time steps of the system evolution as provided by our GNS approach.

Tainpakdipat et al. (2025) use a single CPU PBNS computation as a benchmark and predict 76 time steps, subsampled 20-fold from the time steps of their PBNS. Gong et al. (2025) use 8 CPUs to run their PBNS benchmarks but the ML surrogate only predicts the final time step for the rupture sequence, skipping 1,500–4,000 timesteps from the PBNS. Adjusting for the time steps predicted, and the CPU cores used to run the PBNS, our GNS speedup would nominally be $\sim 3,200$ or $\sim 30,000$ –80,000 if computed in a similar way to Tainpakdipat et al. (2025) or Gong et al. (2025), with the caveats from above.

Moreover, the GNS appears quite general, with strong generalization capability, and requires relatively smaller training sets, compared to other ML approaches, where training costs can be significant. There are still ample opportunities to further optimize the GNS approach for enhanced speedup. For example, one could consider batch rollout of multiple scenarios and/or using mixed precision computing. Moreover, for GNS or other surrogate models, there is a trade-off between information yield and computational efficiency. Additional speedup could be achieved by sacrificing the information yield by, say, training the GNS on subsampled data sets. However, the advantage of the GNS is not just raw speedup for some end results, but the information yield of capturing the full spatiotemporal evolution of system states. Such outputs are valuable for uncertainty quantification and Bayesian inference, and the GNS approach is different from surrogate models that directly map input parameters to certain, restricted outputs, as noted.

Table 6

Computational Cost for GNS Rollout and Comparison to Traditional PBNS

Model grid size	GNS rollout per timestep (ms)	CPU run time per timestep (ms)	Speedup
200 m	~ 11	~ 322	~ 29.3
100 m	~ 35	$\sim 1,449$	~ 41.4

Table 7
Computational Cost for Training the GNS Models

Model	GNS training time per million steps using one Nvidia A100
M2/M3	14.33 hr

Therefore, the GNS approach may still offer an efficient pathway to explore the high dimensionality of rupture dynamics parameter space and work well in real world applications. The GNS implementation is parallelized (Kumar & Vantassel, 2023) and can thus take advantage of the on-going investments in massively parallel GPU architectures. Table 7 shows the training cost of M2 and M3 per million training steps at 14.33 hr using a single Nvidia A100.

4. Discussion

While the overall generalization performance of our GNS is quite good, there are, of course, limitations when seeking to predict physical behaviors that are incompletely represented or absent from the training set. As expected, if the training set only contains scenarios with homogenous initial shear stress, as for D1, the GNS fails to predict rupture acceleration/deceleration caused by stress asperities, and we cannot expect to capture other effects, such as a different friction law, without dedicated training. For early termination of ruptures, scenarios which are sparsely represented in our training set, the GNS still predicts event magnitudes well, but performs poorly for rupture arrest as might be expected (e.g., Appendix A, testing scenario D3.T2). To allow the GNS to learn early rupture termination, or any physical features critical for future applications, the training data set should be augmented with scenarios containing those desired features, or larger weights in loss functions could be applied when rare scenario samples are trained.

However, our experiments also show that the predictive power of a GNS can be incrementally improved by adding features such as heterogeneous initial stress to the training set. The GNS abstracted rules for propagating rupture appear to be automatically adapted in a suitable way by additional learning to capture the right sorts of regional interaction dynamics. We only train the GNS on modest complexity in selected parameters, such as at two asperity stress levels on a single rectangular patch, with only 30/148 scenarios. The GNS then generalizes to strongly spatially variable stress distributions, in combination with unseen hypocenter locations, and produces an estimate of the full spatiotemporal evolution of fault states.

Our GNS approach can be compared to the ML approaches of Gong et al. (2025) and Tainpakdipat et al. (2025) which are very efficient, but only provide end-to-end mappings, and rely on thousands or hundreds of thousands of training scenarios. Gong et al. (2025) show that their networks show larger errors for longer faults and more heterogenous stress patches beyond the training set. This indicates that the GNS's generalization advantages may come with the cost of less speedup compared to the other approaches. More generally, different ML approaches show complementary strengths and weaknesses; it remains to be seen which use cases are suitable for what sort of surrogates, and how the GNS performance might be improved for specific applications.

Our results highlight how the GNS can extract a type of interaction “stencil” akin to a coarse-grained, local abstraction of the key physical relationships governing rupture (i.e., conservation of momentum and mass plus constitutive law). This implies the potential to expand this analysis to a wider set of parameters and applications, such as how variations in pre-stress levels and spatial distribution affect the local interaction stencil. Since there is a link between the neural network and the physics through the interaction stencil, we might consider the GNS a candidate for a robust surrogate approach.

Additional features can be added to the GNS at selected levels, and extensions to a range of heterogeneity such as variations in frictional parameters, different types of frictional evolution laws, and fault roughness appear feasible using our framework. Further exploration of an expanded rupture dynamics parameter space is interesting, and the interactions captured by enhanced learning of a more general GNS might allow for model generalization to general classes of heterogeneities and constitutive laws.

Another pathway forward is to extend the current 2-D planar representation of fault geometry to fully networked fault systems by incorporating more fault geometric complexity of variations in fault dip and strike as vertex features. Representing 3-D fault network structures through a GNS 2-D graph projection may help reduce computational cost. In addition, opportunities also exist in understanding and then improving how the stencils are represented in the weights of the GNS networks, and what physical relationships results from different combinations of hyperparameters, such as message passing steps, entail to better emulate the real physical processes.

Currently, our GNS only predicts slip rate evolution on the fault interface, restricting its utility for seismic hazard assessment to source characteristics such as magnitude, final slip distribution, and rupture duration. However,

there is an opportunity to additionally incorporate wave propagation on the free surface for system states and use the GNS to predict rupture and ground shaking concurrently. Broadening the GNS to predict surface motions will further enhance its utility for seismic hazard assessment of ground shaking.

Efforts to build enhanced GNS surrogates, with parallelization, should reduce the computational costs. We speculate that applications may include surrogates to relate probabilistically represented fault zones to seismicity, and a more broadly trained GNS might represent one version of a much-needed comprehensive representation of our best guesses for the fundamental sets of multi-scale fault zone physics.

The GNS can then be a general, physics informed operator, generating likely rupture characteristics with quantified uncertainties based on incomplete, variable resolution, time-dependent, heterogeneous initial conditions and material parameters. While we are only concerned with the rupture dynamics problem here, a similar GNS approach should lend itself readily to other transient earthquake system problems such as fault network, post-seismic, and earthquake cycle stress heterogeneity evolution.

5. Conclusions

We present a Graph Network-based Simulator (GNS) for dynamic earthquake rupture. The GNS can predict rupture dynamics in comparable accuracy to high-fidelity physics-based simulations, in particular during the nucleation and rupture front propagation stage, with ~ 20 – 40 times per-time-step speedup compared to the physics-based modeling. The generalization capability of the GNS extends to scenarios with unseen hypocenter locations, stress asperity strength and geometry variations, and fault sizes. The GNS appears to learn a coarse-grained version of the physics governing local dynamics on the fault plane. This physics-informed structure of the GNS implies good robustness, as well as providing new perspectives on how to potentially explore different types of physical interactions of relevance for earthquake dynamics. With improved computational efficiency and expanded training, the GNS may be suited to serve as a comprehensive surrogate for future Bayesian model inference, uncertainty quantification, or optimal experimental design tasks. Those could cover rupture dynamics applications within both a fundamental exploration of physics and an operational, time-dependent seismic hazard assessment context.

Appendix A: Predicting Rare Physical Feature of Early Rupture Termination

Figure A1A shows the prediction of M1 on testing scenario D1.T4 where the rupture terminates early in the middle of the fault, where dense rupture curves merge, and the GNS, which is not trained to learn such rare physical dynamic pattern, fails to capture the termination. In this case, the prediction of magnitude errors at ~ 0.3 . Figure A1B shows the prediction of M3 on testing scenario D3.T2, where rupture terminates prematurely but ruptures coming from the top half of the fault keeps breaking the whole fault. The GNS predicts continuous ruptures where the ground truths show terminations. However, in this case, the magnitude is still well predicted.

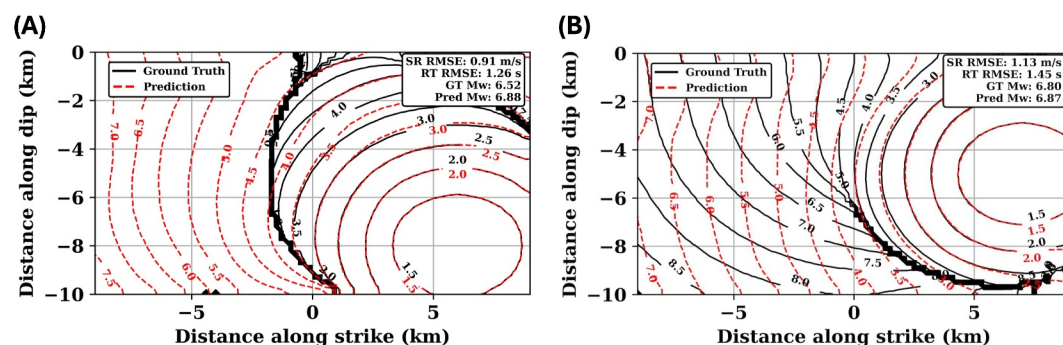


Figure A1. (a) Rupture time contour comparisons between GNS M1 predicted dynamic ruptures (red) and PBNS ground truth (black) for testing scenario D1.T4 with unseen hypocenter locations in the training. (b) Rupture time contour comparisons between GNS M3 predicted dynamic ruptures (red) and PBNS ground truth (black) for testing scenario D3.T2 with unseen stress levels and complex patterns and hypocenter locations in the training set. Labels and error metrics else as in Figure 5, see there for description.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

The code of the GNS for earthquake rupture dynamics, EQGNS, and the finite element software EQdyna are publicly available on GitHub and the Zenodo repository (Liu & Becker, 2025). Necessary scripts and configuration files used to generate EQdyna ground truth scenarios, training, validation, and testing data sets for GNS, and rollout results are archived under the Zenodo repository (Liu & Becker, 2025).

Acknowledgments

DL and TWB were partially funded by the National Science Foundation under Grant EAR-2121666. DL was also supported by the Southern (now *Statewide*) California Earthquake Center (SCEC), and this is SCEC contribution no. 14197. SCEC is funded by National Science Foundation under Cooperative Agreement EAR-1600087 and the United States Geological Survey under Cooperative Agreement G17AC0047. We thank Dr. Mousavi and an anonymous reviewer for their comments on an earlier version which helped to improve the manuscript.

References

- Andrews, D. J., & Barall, M. (2011). Specifying initial stress for dynamic heterogeneous earthquake source models. *Bulletin of the Seismological Society of America*, 101(5), 2408–2417. <https://doi.org/10.1785/0120110012>
- Battaglia, P., Pascanu, R., Lai, M., & Jimenez Rezende, D. (2016). Interaction networks for learning about objects, relations and physics. *Advances in Neural Information Processing Systems*, 29.
- Choi, Y., & Kumar, K. (2024). Graph neural Network-based surrogate model for granular flows. *Computers and Geotechnics*, 166, 106015. <https://doi.org/10.1016/j.compgeo.2023.106015>
- Day, S. M. (1982). Three-dimensional simulation of spontaneous rupture: The effect of nonuniform prestress. *Bulletin of the Seismological Society of America*, 72(6A), 1881–1902. <https://doi.org/10.1785/bssa07206a1881>
- Dieterich, J. H. (1979). Modeling of rock friction. I. experimental results and constitutive equations. *Journal of Geophysical Research*, 84(Nb5), 2161–2168. <https://doi.org/10.1029/JB084iB05p02161>
- Dieterich, J. H., & Linker, M. F. (1992). Fault stability under conditions of variable normal stress. *Geophysical Research Letters*, 19(16), 1691–1694. <https://doi.org/10.1029/92GL01821>
- Duan, B. (2012). Dynamic rupture of the 2011 Mw 9.0 Tohoku-Oki earthquake: Roles of a possible subducting seamount. *Journal of Geophysical Research*, 117(B5), B05311. <https://doi.org/10.1029/2011JB009124>
- Duan, B., & Oglesby, D. D. (2006). Heterogeneous fault stresses from previous earthquakes and the effect on dynamics of parallel strike-slip faults. *Journal of Geophysical Research*, 111(B5), B05309. <https://doi.org/10.1029/2005jb004138>
- Dunham, E. M., Belanger, D., Cong, L., & Kozdon, J. E. (2011). Earthquake ruptures with strongly rate-weakening friction and off-fault plasticity. part 2: Nonplanar faults. *Bulletin of the Seismological Society of America*, 101(5), 2308–2322. <https://doi.org/10.1785/0120100076>
- Fournier, A., Fussell, D., & Carpenter, L. (1982). Computer rendering of stochastic models. *Communications of the ACM*, 25(6), 371–384. <https://doi.org/10.1145/358523.358533>
- Gallovič, F., Valentová, L., Ampuero, J. P., & Gabriel, A. A. (2019). Bayesian dynamic finite-fault inversion: 1. Method and synthetic test. *Journal of Geophysical Research: Solid Earth*, 124(7), 6949–6969. <https://doi.org/10.1029/2019jb017510>
- Gilmer, J., Schoenholz, S. S., Riley, P. F., Vinyals, O., & Dahl, G. E. (2017). Neural message passing for quantum chemistry. In P. Doina & T. Yee Whye (Eds.), *Proceedings of the 34th international conference on machine learning, PMLR, proceedings of machine learning research* (pp. 1263–1272).
- 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(5), e2024JB030069. <https://doi.org/10.1029/2024jb030069>
- Harris, R. A., Barall, M., Aagaard, B., Ma, S., Roten, D., Olsen, K., et al. (2018). A suite of exercises for verifying dynamic earthquake rupture codes. *Seismological Research Letters*, 89(3), 1146–1162. <https://doi.org/10.1785/0220170222>
- Harris, R. A., Barall, M., Archuleta, R., Dunham, E., Aagaard, B., Ampuero, J. P., et al. (2009). The SCEC/USGS dynamic earthquake rupture code verification exercise. *Seismological Research Letters*, 80(1), 119–126. <https://doi.org/10.1785/gssrl.80.1.119>
- Ida, Y. (1972). Cohesive force across the tip of a longitudinal-shear crack and Griffith's specific surface energy. *Journal of Geophysical Research*, 77(20), 3796–3805. <https://doi.org/10.1029/JB077i020p03796>
- Jia, Z., Jin, Z., Marchandon, M., Ulrich, T., Gabriel, A. A., Fan, W., et al. (2023). The complex dynamics of the 2023 Kahramanmaraş, Turkey, Mw 7.8–7.7 earthquake doublet. *Science*, 381(6661), 985–990. <https://doi.org/10.1126/science.adi0685>
- Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- Kumar, K., & Choi, Y. (2023). Cylinder flow with graph neural network-based simulator. In *DesignSafe-CI*. <https://doi.org/10.17603/ds2-fzg7-1719>
- Kumar, K., & Vantassel, J. (2023). GNS: A generalizable graph neural Network-based simulator for particulate and fluid modeling. *Journal of Open Source Software*, 8(88), 5025. <https://doi.org/10.21105/joss.05025>
- Lam, R., Sanchez-Gonzalez, A., Willson, M., Wirsberger, P., Fortunato, M., Alet, F., et al. (2023). Learning skillful medium-range global weather forecasting. *Science*, 382(6677), 1416–1421. <https://doi.org/10.1126/science.adi2336>
- Liu, D., & Becker, T. (2025). Source code and dataset for research article earthquake rupture dynamics from graph neural networks. In *Earthquake rupture dynamics from graph neural networks (1.0)*. [Software][Dataset]. <https://doi.org/10.5281/zenodo.17095311>
- Liu, D., & Duan, B. (2018). Scenario earthquake and ground-motion simulations in north China: Effects of heterogeneous fault stress and 3D basin structure. *Bulletin of the Seismological Society of America*, 108(4), 2148–2169. <https://doi.org/10.1785/0120170374>
- Liu, D., Duan, B., Schärer, K., & Yule, D. (2022). Observation-constrained multicycle dynamic models of the southern San Andreas and the northern San Jacinto faults: Addressing complexity in paleoearthquake extent and recurrence with realistic 2D fault geometry. *Journal of Geophysical Research: Solid Earth*, 127(2), e2021JB023420. <https://doi.org/10.1029/2021JB023420>
- Luo, B., & Duan, B. (2018). Dynamics of nonplanar thrust faults governed by various friction laws. *Journal of Geophysical Research: Solid Earth*, 123(6), 5147–5168. <https://doi.org/10.1029/2017jb015320>
- Pfaff, T., Fortunato, M., Sanchez-Gonzalez, A., & Battaglia, P. W. (2020). Learning mesh-based simulation with graph networks. *arXiv preprint arXiv:2010.03409*.
- Premus, J., Gallovič, F., Hanyk, L., & Gabriel, A.-A. (2020). FD3D-TSN: A fast and simple code for dynamic rupture simulations with GPU acceleration. *Seismological Research Letters*, 91(5), 2881–2889. <https://doi.org/10.1785/0220190374>

- Ramos, M. D., Huang, Y., Ulrich, T., Li, D., Gabriel, A.-A., & Thomas, A. M. (2021). Assessing margin-wide rupture behaviors along the cascadia megathrust with 3-D dynamic rupture simulations. *Journal of Geophysical Research: Solid Earth*, 126(7), e2021JB022005. <https://doi.org/10.1029/2021JB022005>
- Ruina, A. (1983). Slip instability and state variable friction laws. *Journal of Geophysical Research*, 88(Nb12), 359–370. <https://doi.org/10.1029/JB088iB12p10359>
- Sanchez-Gonzalez, A., Godwin, J., Pfaff, T., Ying, R., Leskovec, J., & Battaglia, P. W. (2020). Learning to simulate complex physics with graph networks. In *Proceedings of the 37th international conference on machine learning* (p. 784). JMLR.org.
- Shaw, B. E., Fry, B., Nicol, A., Howell, A., & Gerstenberger, M. (2022). An earthquake simulator for New Zealand. *Bulletin of the Seismological Society of America*, 112(2), 763–778. <https://doi.org/10.1785/0120210087>
- Shi, Z., & Day, S. M. (2013). Rupture dynamics and ground motion from 3-D rough-fault simulations. *Journal of Geophysical Research: Solid Earth*, 118(3), 1122–1141. <https://doi.org/10.1002/jgrb.50094>
- Suhendi, C., Li, B., Vasyura-Bathke, H., Liu, J., Jónsson, S., & Mai, P. M. (2025). Bayesian inversion and quantitative comparison for bilaterally quasi-symmetric rupture processes on a multisegment fault in the 2021 Mw7.4 Maduo earthquake. *Geophysical Journal International*, 240(1), 673–695. <https://doi.org/10.1093/gji/ggae398>
- Tainpakdipat, N., Abdelmeguid, M., Zhao, C., Azzadenesheli, K., & Elbanna, A. (2025). Fourier neural operators for accelerating earthquake dynamic rupture simulations.
- Taufiqurrahman, T., Gabriel, A. A., Ulrich, T., Valentová, L., & Gallovič, F. (2022). Broadband dynamic rupture modeling with fractal fault roughness, frictional heterogeneity, viscoelasticity and topography: The 2016 Mw 6.2 amatrice, Italy earthquake. *Geophysical Research Letters*, 49(22), e2022GL098872. <https://doi.org/10.1029/2022gl098872>
- Thangamuthu, A., Kumar, G., Bishnoi, S., Bhattoo, R., Krishnan, N., & Ranu, S. (2022). Unravelling the performance of physics-informed graph neural networks for dynamical systems. *Advances in Neural Information Processing Systems*, 35, 3691–3702.
- Ulrich, T., Gabriel, A. A., Ampuero, J. P., & Xu, W. (2019). Dynamic viability of the 2016 Mw 7.8 Kaikoura earthquake cascade on weak crustal faults. *Nature Communications*, 10(1), 1213. <https://doi.org/10.1038/s41467-019-09125-w>
- Wang, Y., & Day, S. M. (2020). Effects of off-fault inelasticity on near-fault directivity pulses. *Journal of Geophysical Research: Solid Earth*, 125(7), e2019JB019074. <https://doi.org/10.1029/2019JB019074>
- Withers, K. B., Olsen, K. B., Shi, Z., & Day, S. M. (2018). Validation of deterministic broadband ground motion and variability from dynamic rupture simulations of buried thrust earthquakes. *Bulletin of the Seismological Society of America*, 109(1), 212–228. <https://doi.org/10.1785/0120180005>
- Zhu, W., Allison, K. L., Dunham, E. M., & Yang, Y. (2020). Fault valving and pore pressure evolution in simulations of earthquake sequences and aseismic slip. *Nature Communications*, 11(1), 4833. <https://doi.org/10.1038/s41467-020-18598-z>