doi:10.26443/seismica.v5i1.2194

Hybrid Approach of Real-time Ground Motion Prediction for Earthquake Early Warning

Zhang Hongcai [1], **Liu Qifang** [*2], **Jin Xing** [1,3], **Cai Huiteng** [1], **Chen Zhiyong** [1], **Wang Shicheng** [1]

[1]Fujian Earthquake Agency, China Earthquake Administration, Fuzhou, China, [2]Suzhou University of Science and Technology, Suzhou, China, [3]Institute of Xiamen Marine Seismology, China Earthquake Administration, Xiamen, China

Author contributions: *Conceptualization*: Zhang Hongcai, Liu Qifang, Jin Xing. *Methodology*: Zhang Hongcai, Jin Xing, Cai Huiteng. *Writing - Original draft*: Zhang Hongcai, Liu Qifang, Wang Shicheng. *Writing - Review & Editing*: Jin Xing, Chen Zhiyong. *Funding acquisition*: Liu Qifang.

Abstract We propose a hybrid approach for ground motion field prediction in earthquake early warning (EEW) systems, combining the Propagation of Local Undamped Motion (PLUM) method with advanced machine learning (ML) algorithms. Our methodology integrates PLUM's seismic wavefield estimation with CONIP, a four-layer Convolutional Neural Network (CNN) model for onsite seismic intensity estimation, enabling real-time ground motion field predictions. Trained on 69,089 KiK-net seismic records, the ML-based model demonstrates a 28.8% reduction in median absolute error (MAE) in predicted Japan Meteorological Agency (JMA) seismic intensity compared to the traditional ground motion prediction equation (GMPE) implementation, with predictions available 1-5 seconds earlier during the initial 10 seconds after origin time. These improvements reflect enhanced accuracy in estimating peak ground motion intensities and greater stability in early-stage predictions. By employing a Monte Carlo simulation with Gaussian-perturbed inputs, the model yields well-calibrated uncertainty estimates that correlate with prediction errors. When benchmarked against the GMPE-based method, our approach shows consistent performance advantages in prediction reliability. Offline simulations of the 2005 West off Fukuoka earthquake confirm the hybrid approach's computational efficiency and enhanced predictive capability, particularly for destructive events. By improving prediction accuracy at critical thresholds, this solution provides an average lead time gain of 5.49 seconds over the original PLUM method for end-users and strengthens EEW effectiveness.

Production Editor: Andrea Llenos
Handling Editor: Pablo Heresi
Copy & Layout Editor: Hannah F. Mark

Signed reviewer(s): E. S. Cochrane, Debi Kilb

Received: December 6, 2025
Accepted: April 20, 2026
Published: May 24, 2026

1 Introduction

For end users of an earthquake early warning (EEW) system, estimated ground motion is more meaningful than the hypocenter and magnitude, which are currently common practice in many EEW systems despite the challenges in their rapid and accurate determination (Allen and Melgar, 2019). Reliable forecasting of ground shaking is crucial for making scientific decisions, such as slowing down high-speed trains, shutting down nuclear reactors, pausing elevators on a nearby floor, and warning people to evacuate dangerous areas. EEW systems today primarily rely on two methodological paradigms: source-based approaches, which first estimate earthquake source parameters (e.g., magnitude and location) and then predict shaking using ground motion models; and ground-motion-based approaches, which bypass source inversion and instead use real-time observed shaking to directly forecast intensities at nearby sites. Currently, most operation EEW systems are point-source algorithm-based (Allen, 2007; Brown et al., 2009; Satriano et al., 2011; Chen et al., 2015; Doi, 2011; Zhang et al., 2016). They estimate ground shaking using a ground motion prediction equation (GMPE), where ground motions (PGA, PGV, and seismic intensity) are simplified as a function of magnitude, hypocentral distance, fault mechanism, and site effects. This approach is straightforward but large scatters often occur since the source, site, and path conditions are spatially correlated and not fully included during ergodic GMPE estimation (Hoshiba and Aoki, 2015; Kuehn and Abrahamson, 2018; Spallarossa et al., 2019; Kuehn and Abrahamson, 2020). Finite source algorithms, such as FinDer (Böse et al., 2012, 2018) and GlarmS (Ruhl et al., 2017), use earthquake sources that have finite length rather than a single point, but the strategies adopted for ground motion estimation are similar. Therefore, the accuracy of ground motion predictions still relies on the accuracy of GMPEs. During the M_W 9.0 Tohoku earthquake, the national-wide EEW system in Japan succeeded in issuing warnings more than 15 seconds before the strong shaking arrival in the Tohoku district. At the time of the initial warning, the JMA EEW system reported a magnitude of 7.2, significantly lower than the true magnitude. However, it significantly under-predicted the ground shaking at large distances due to the large fault rupture, while it over-predicted the ground shaking during the intense aftershocks (Hoshiba and Ozaki, 2014).

To provide a more reliable estimation of ground shaking for Earthquake Early Warning (EEW), Hoshiba (2013) utilized the boundary integral equation to predict ongoing ground motion without the need for earth-

*Corresponding author: qifang_liu@126.com

quake source information such as location and magnitude, though the direction of wave propagation had to be assumed. Shortly after, Hoshiba and Aoki (2015) suggested a numerical shaking prediction technique with the help of a data assimilation technique for estimating ground motion, alongside the Radiative Transfer Theory (RTT) to minimize computation time. Regarding this, Kodera et al. (2018) then improved the algorithm and proposed the Propagation of Local Undamped Motion (PLUM) approach as a simpler wavefield prediction method. Seismic intensities of target sites are directly estimated through the usage of observed data during the PLUM process. Studies (Cochran et al., 2019; Minson et al., 2019; Kilb et al., 2021; Saunders et al., 2022, 2024, 2026) have exhibited that PLUM is a proficient and precise method that is also capable of dealing with complicated events, such as multiple simultaneous earthquakes. Therefore, PLUM is deemed to be a rather viable algorithm that has gained plenty of attention in the scientific community. In this paper, we demonstrate that combining the already capable PLUM approach with advanced Machine Learning (ML) data-driven techniques can advance the accuracy of ground motion predictions even further. This hybrid strategy specifically targets a fundamental constraint of PLUM, its reliance on observed strong motions from triggered stations, which limits warning lead time, particularly for sites near the epicenter. By integrating a ML model, we can predict final seismic intensity from only the initial seconds of waveform data, our method enables reliable early estimates before strong shaking arrives, thereby extending the effective warning window while maintaining spatial coherence through PLUM's propagation framework.

With the ever-increasing amount of seismic observation data, research into ML (a broad class of data-driven algorithms that learn patterns from data) and deep learning (DL, a subset of ML based on multi-layer neural networks capable of automatically extracting complex features from data) algorithms has become a hot topic in the seismological community (Kong et al., 2019; Mousavi and Beroza, 2022, 2023). Building on the pioneering work by Perol et al. (2018), numerous ML algorithms have been explored across various fields of seismology, including earthquake event detection (Meier et al., 2019; Ross et al., 2019a; Zhu et al., 2019), seismic phase identification (Ross et al., 2018; Wang et al., 2019; Zhao et al., 2019; Zhu and Beroza, 2019; Mousavi et al., 2020), earthquake location (Agius and Galea, 2011; Kriegerowski et al., 2019; Ross et al., 2019a,b; Münchmeyer et al., 2021a; Yang et al., 2024), event classification (Diersen et al., 2011; Linville et al., 2018; Woollam et al., 2019), seismic tomography (Adler et al., 2019; Araya-Polo et al., 2018, 2019a,b), EEW (Abdalzaher et al., 2024; Kong et al., 2016; Li et al., 2018; Allen et al., 2020; Münchmeyer et al., 2021b; Chitkeshwar, 2024; Kolivand et al., 2024; Noda, 2024; Zhang et al., 2024; Huang et al., 2025; Jiang et al., 2025), earthquake hazard mitigation (Khoshnevis and Taborda, 2018; Trugman and Shearer, 2018; Mousavi et al., 2020; Zhu et al., 2020; Joshi et al., 2024), and earthquake prediction (Al Banna et al., 2020; Li et al., 2020). Recently, Navarro-Rodríguez et al. (2025) systematically reviewed the study of ML and DL in EEW magnitude prediction and found these algorithms significantly outperform traditional methods like P_d (peak displacement) and τ_c (the characteristic period of the initial P-wave). The above research has once again proved the excellence performance of the ML/DL algorithms.

In order to understand what an ML model learns through extensive training with massive data, Zhang et al. (2022) developed an onsite ground motion prediction model using a 4-layer convolutional neural network (CNN) called the "CONIP" model, which is designed to estimate onsite seismic intensity using only the initial seconds of waveform data from a triggered station, and it continuously refines its prediction as more data become available over an expanding time window. This model was trained on over 32,000 seismic event records from California, spanning 2000-2019, with moment magnitudes (M_W) ranging from 2.0 to 7.1. We analyzed the prediction residuals of the modified Mercalli intensity (MMI) separately using the traditional GMPE-based approach and the CONIP method. We found that the MMI prediction residuals from the CONIP method were significantly smaller than those from the GMPE-based approach. Furthermore, in our previous study (Zhang et al., 2022), we discovered that the CONIP model could effectively learn site and source effects, but had difficulty learning path effects during the model training process. Given the exceptional performance of the CONIP model, we believe that it can be applied to EEW systems. However, since CONIP is a typical onsite prediction approach with limited ability to predict wavefields, the timeliness of warning information may be relatively short if relying solely on this method. To address this, we combined our CONIP module with the PLUM method, which provides practical solutions to this problem. Additionally, we discovered that the CONIP model has specific regional applicability. Hence, we collected historical data from the KiK-net network in Japan to create new training datasets. We then trained the unchanged network with this new data. Finally, we selected an event with a magnitude greater than 7.0 (the 2005 West off Fukuoka earthquake) and showed the feasibility of our hybrid approach by displaying the ground motion prediction using our approach second-by-second after the first station triggered. Our approach differs from Kodera et al. (2018), which estimates final S-wave amplitudes from initial P-wave parameters without ML, and from recent DL extensions such as Bloemheuvel et al. (2023) and Clements et al. (2024), which integrate neural networks into PLUM-like frameworks but rely on different input representations or network architectures. In contrast, CONIP directly predicts MMI from raw waveform segments at individual stations and is explicitly designed for seamless hybridization with PLUM to enhance both accuracy and warning timeliness.

This study adopts the same assumption as Kodera et al. (2018), that strong ground motion propagates without attenuation within radius R when R is sufficiently small (here, R=30 km). Building on this framework, we integrate the CONIP model into PLUM by replacing the

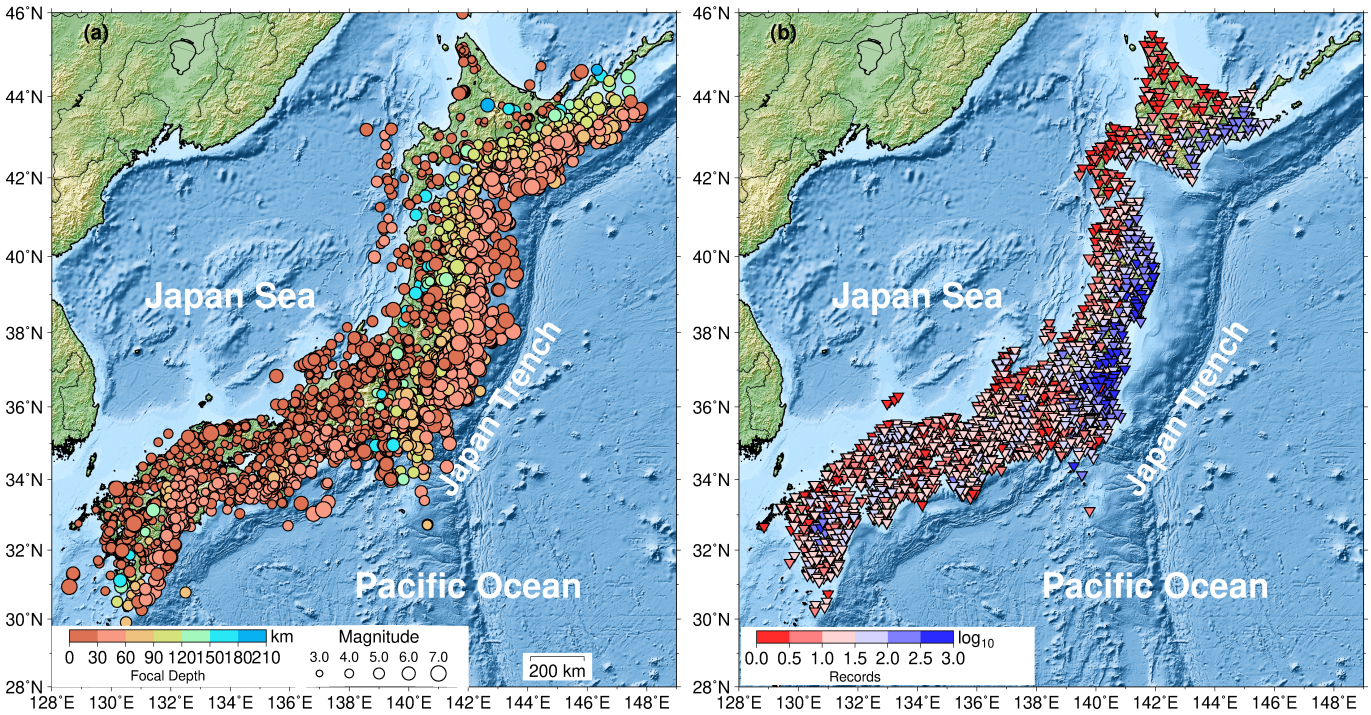

Figure 1 (a) Spatial distribution of hypocenters (denoted by dots) for all seismic events employed in the training of the CONIP model within this study. Hypocenter locations were determined by JMA, with dot size encoding event magnitude and color fill representing focal depth. (b) Spatial distribution of KiK-net stations (denoted by inverted triangles) used in training, color-coded by the number of ground motion records available per station in the training dataset.

observed peak ground motion at triggered stations with CONIP-predicted modified MMI. The resulting hybrid system propagates these predicted intensities to surrounding locations in real time, thereby enabling earlier and more stable warning estimates than either approach alone. However, when integrating the PLUM algorithm with our CONIP model, especially for a dense observation network, we found that on-site seismic intensity can be estimated more rapidly and reliably than through conventional methods. Therefore, we focus on this specific application and propose a novel hybrid approach. The proposed hybrid framework uniquely integrates the spatial robustness of PLUM, known for its stable ground motion field estimation, with the early predictive capability of a DL-based onsite intensity estimator (CONIP). This synergy not only improves prediction accuracy but also extends actionable warning lead time, addressing a key operational limitation of original PLUM implementations.

2 Data

Similar to all ML algorithm models, a large amount of high-quality historical seismic data is needed to train our CONIP model; specifically, ∼70,000 records were used on in this paper. Our CONIP model has regional applicability, in that the ground motion, such as seismic intensity prediction accuracy, may decrease sharply when the trained model is applied to another region that was not included during the module training. More importantly, in our previous work (Zhang et al., 2022), the datasets used in the model training mainly consisted of small-to-moderate (mainly M_W2.0-6.0) events in Parkfield and Los Angeles, California, with few large-magnitude events ($M_W \geq 6.5$) involved. Thus, the trained model will have limited ability to accurately predict those large magnitude events.

We aim to demonstrate the efficacy of our hybrid seismic intensity prediction algorithm by utilizing a newly constructed dataset consisting of historical earthquake events from KiK-net of Japan (Okada, 2013). Japan is a region with high seismic activity, and the country has established dense seismic observation networks, which allowed us to collect a high-quality ML training dataset. Our selection criteria for records were as follows: earthquake magnitude greater than 2.0 (JMA scale), peak ground motion acceleration (PGA) greater than 10 cm/s^2, and surface records only, but no constraint was placed on the number of stations per event. These criteria were selected to isolate potentially damaging earthquakes from the vast amount of seismic data and provide valuable features for our CONIP model to enhance prediction accuracy. Our dataset includes major earthquake events, like the 2011 Tohoku earthquake and the 2003 Tokachi-Oki earthquake, but we excluded records of the Tohoku earthquake aftershocks collected during dataset construction (March 11, 2011 to June 30, 2011). This exclusion was due to the significant overlap of seismic records during dense aftershock sequences, which made it difficult to distinguish events in such complex situations. By focusing on significant events and excluding aftershocks, we aim to simplify our ML training and provide accurate predictions using our hybrid algorithm. As shown in Fig. 1(b), the seismic stations used in this study cover nearly the entire Japanese archipelago. However, the number of available train-

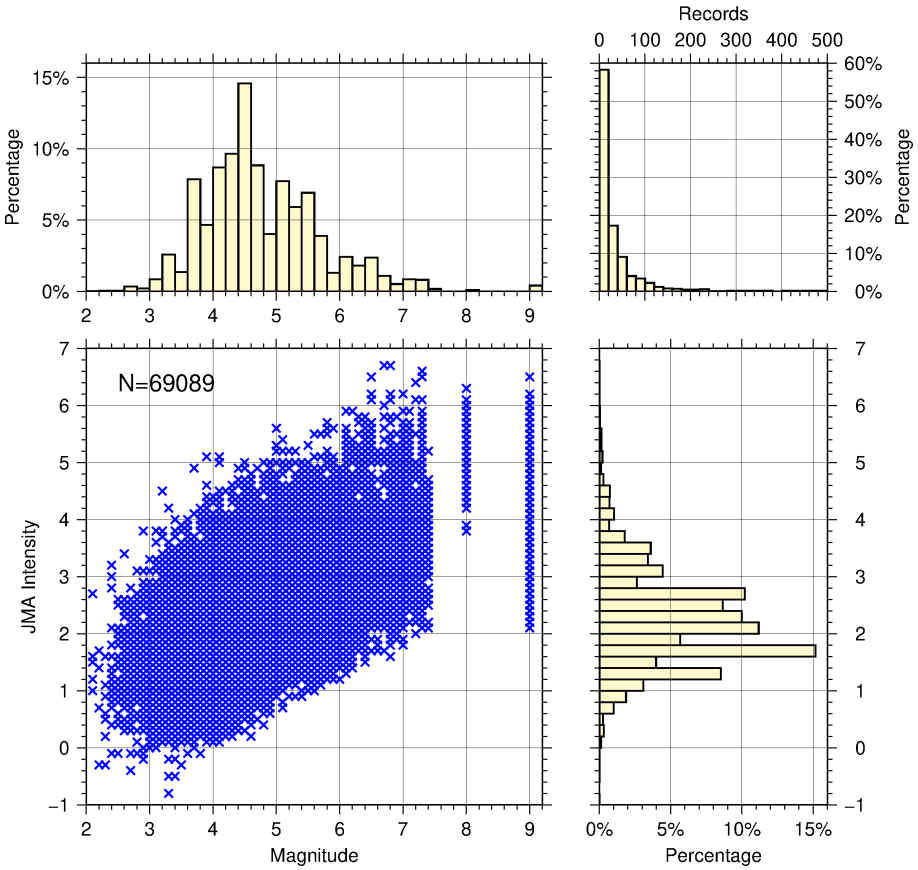

Figure 2 Statistics of the 69,089 records used in CONIP model training. The bottom-left panel displays a scatter plot of earthquake magnitude (M_j) versus JMA intensity measurements. Corresponding distribution histograms for magnitude and JMA intensity appear in the top-left and bottom-right panels, respectively. The top-right panel presents a histogram of the number of training records per station.

ing records per station is highly heterogeneous, ranging from just 1 to over 500. According to the statistics presented in Fig. 2, only ∼18% of stations have more than 50 records in the training dataset.

Following data acquisition, waveform preprocessing involved removing the mean value from each waveform and trimming 10-second three-component signals from the P-wave arrival time at each station. No bandpass filtering was applied, and no explicit signal-to-noise ratio (SNR) criterion was imposed. The preprocessed data were archived in HDF5 format (The HDF Group[1]) to facilitate efficient access. During manual waveform inspection, metadata timing errors were identified in KiK-net records. To correct these, we computed theoretical P-phase arrival times using the "tjma2001"[2] travel-time table and JMA hypocenter parameters (based on JMA bulletin), then aligned the observed P-phase picks (derived from STA/LTA+AIC detection) to these theoretical arrivals by waveform shifting. Our comprehensive training dataset consisted of 69,089 three-channel records from a total of 3,220 events. Readers can visualize the distribution of these events and histogram of the records in Fig. 1 and Fig. 2, respectively.

[1]https://www.hdfgroup.org/HDF5/

[2]https://www.data.jma.go.jp/eqev/data/bulletin/catalog/appendix/trtime/trt_e.html

3 Hybrid Approach

In our previous study (Zhang et al., 2022), we present the CONIP model, a four-layer convolutional neural network (CNN) that uses high-order feature extraction to improve onsite seismic intensity estimation from records. While it is challenging to determine the exact features learned during training, our seismic intensity prediction residual analysis confirmed that the model produces significantly smaller residuals, both in terms of mean absolute error and bias, compared to traditional GMPE-based methods, even with a short time window (TW=1 s). Unlike human-defined features like P_d, τ_c, and $IV2$ (squared velocity integral), CONIP learns more complex, autonomous features from the dataset, making them potentially more generalizable. In this study, we retain the CONIP model architecture, where the input tensor possesses dimensions of (N×3×(TW×100)), where N represents the number of events used during model training and 100 is the sampling rate of the record. The mean squared error (MSE) is adopted as the loss function. The dataset is partitioned in a 6:2:2 ratio (training:validation:testing). Optimization is performed using the Adam optimizer, with a learning rate of 0.0001 and we implement an "early-stop" strategy that stops model training if the validation loss does not decrease for eight training epochs. When doing EEW processing, we can continually update the onsite ground motion prediction in real-time as

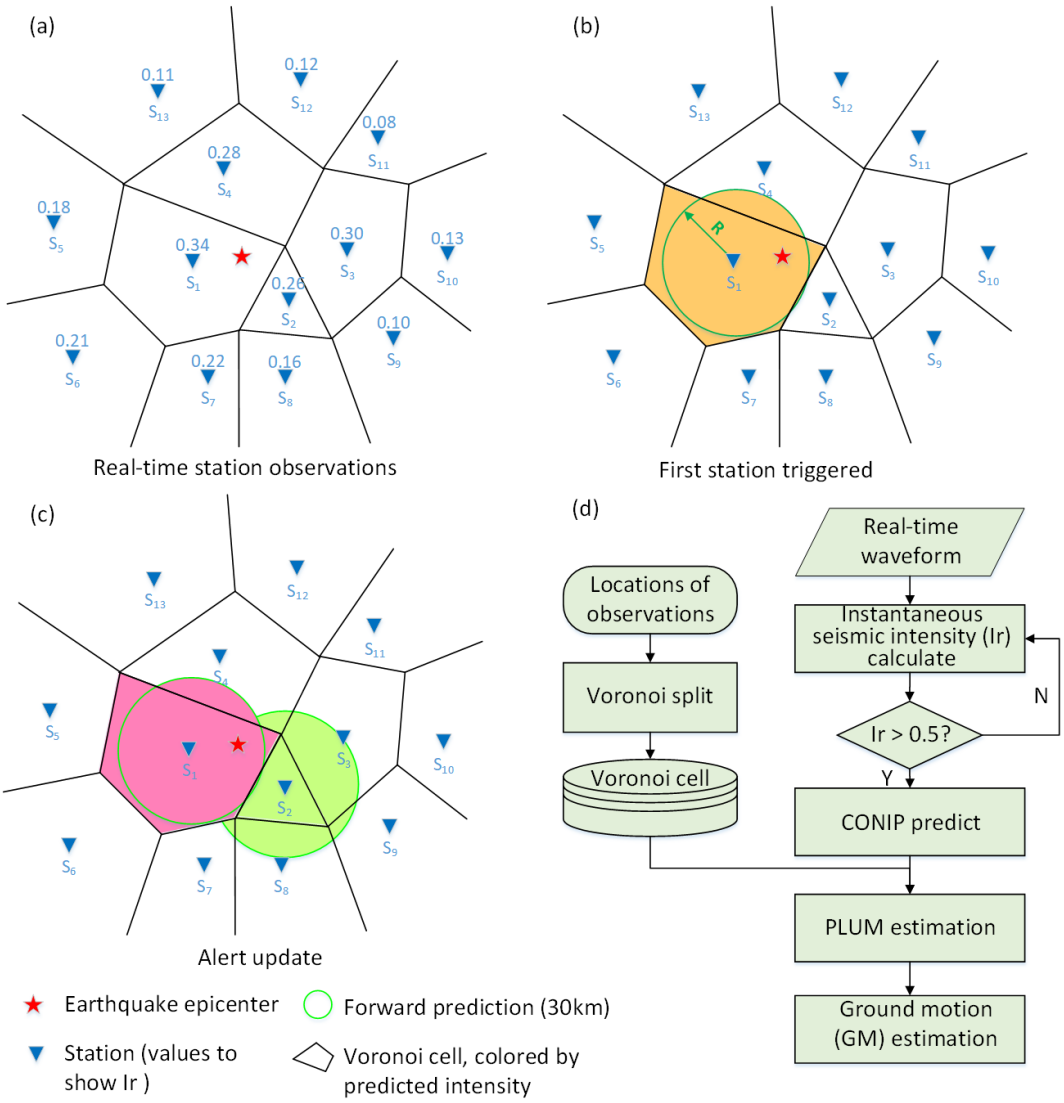

Figure 3 Illustration of the proposed hybrid methodology. Panel (a)-(c) depicts Voronoi tessellation within a seismic monitoring network, triggered stations (highlighted cells) initiate CONIP model predictions for seismic intensity at corresponding sites, with predicted values propagated to all cells associated with the triggered station through Voronoi analysis. Panel (d) presents the methodological workflow diagram.

the available waveform information grows with increasing TW. To accomplish this, we train separate models for each TW while keeping the same CONIP structure and training parameters. Interested parties looking for more information can visit our GitHub repository (see "Data Availability" section). For a more detailed explanation of the CONIP model and its performance, please refer to Zhang et al. (2022). Our findings suggest that the CONIP model enhances EEW systems by providing more accurate final seismic intensity estimates, and delivering these predictions earlier, thereby extending actionable warning lead time without compromising spatial reliability.

Details of the PLUM method were discussed by Kodera et al. (2018), which will not be repeated in our paper. In short, it directly predicts the ground motion at a target site by using the maximum observed real-time seismic intensity within a predefined circular region around that site, after correcting for site amplification factors. This simple, non-parametric method operates without requiring real-time estimation of earthquake source parameters (e.g., hypocenter, magnitude). This makes it robust against challenges faced by conventional point-source models, such as the under prediction of shaking from large finite-fault earthquakes, missed events during intense seismic sequences, and over prediction for multiple simultaneous earthquakes. While providing accurate and robust ground-motion predictions in these complex scenarios, the trade-off is a shorter maximum available warning time compared to some point-source methods, as predictions are only possible after strong shaking is actually detected by nearby stations. Our approach, however, is different in that the seismic intensities at each triggered station are provided by our CONIP model, which predicts the final seismic intensity at each site in real time, instead of measuring it once it exceeds a trigger threshold. This difference allows us to estimate longer warning times than the original method, especially for large events. While our CONIP model is trained using triggered data, our hybrid approach uses the recursive filter (Kunugi et al., 2008), which consisting of six cascaded Infinite

Impulse Response (IIR) filter units, to obtain the instantaneous seismic intensity I_r (on the JMA scale, hereinafter) for non-triggered stations. We set a minimum I_r threshold of 0.5 to simplify the trigger discrimination algorithm, and if a station exceeds this threshold, our CONIP model generates the seismic intensity predictions. Following operational EEW practice, a prediction is considered accurate if the absolute difference between predicted intensities ($\hat{I}$) and observed intensities (I) is no greater than 0.5 units. We hereby define the recall rate as the fraction of events meeting this criterion.

In contrast to the original method, which considers the maximum observed real-time seismic intensity within a circular radius of R (e.g. 30 km), we used Voronoi diagrams (Aurenhammer, 1991) to identify nearby stations and estimate the maximum seismic intensity of all surrounding stations for each target site. To simplify the calculation, we did not account for site effects during ground motion prediction. Fig. 3 presents a schematic diagram of our hybrid method and shows the data process flow. In summary, our approach improves upon the original PLUM method by predicting final seismic intensities in real-time and leveraging Voronoi diagrams to estimate the maximum seismic intensity of surrounding stations. These improvements have the potential to increase warning times for large events, providing vital time to prepare and respond to earthquake impacts.

4 Result

4.1 Model training and validation

Our CONIP model is easy to train, requiring only a 16-core i7 CPU without any graphical processing units (GPUs). Each training epoch takes a few seconds (typically less than one minute) to complete, and we usually train for 80-100 epochs to obtain the "lowest loss" trained model. As a result, the entire training process is completed in approximately 30-35 minutes. Fig. 4(a) shows the training and validation loss for TW=3 s, while Fig. 4(b) demonstrates our trained model's ability to predict JMA seismic intensities on the KiK-net test set. As shown in Fig. 4, both training and validation losses drop quickly in the first few epochs. The model converges rapidly after about 40 to 50 epochs of training. The trained model exhibited excellent performance on the test set, with 66.8% and 92.7% of records reliably estimated with residuals less than $\pm$0.5 unit and $\pm$1.0 unit (JMA intensity scale, hereinafter), respectively. However, some records (7.3%) exhibited higher prediction errors, particularly for larger JMA seismic intensity records (e.g. $\geq$5.0), which we will further discuss in the "Discussion" section.

In our previous discussion (Zhang et al., 2022), we obtained separately trained models for each TW growth. This study defines a recall rate metric based on predicted seismic intensity values. To show the effects of TW growth, we compared the mean and standard deviations of the JMA seismic intensity prediction errors and the recall rate on the test set for each trained model in Fig. 5. The TW varies from 1 to 10 seconds. We observed that the mean prediction values for each TW were all around 0.0, indicating that the trained model could generate unbiased predictions even with a short TW. Moreover, as the TW increased, the standard deviations for each TW gradually decreased. For example, when the TW was 10 seconds, the prediction error was -0.02$\pm$0.39 unit, which was 37% smaller than the value in TW=3 seconds, which was -0.01 $\pm$0.62 unit. Similarly, a positive association emerges between longer TW and higher recall rate. The most pronounced improvement occurs during initial TW extension, the recall rate increases markedly from 65.5% at TW=1 s to 72.6% at TW=3 s. With further TW expansion to 10 s, the recall rate plateaus at 82.5%. This performance enhancement correlates directly with the greater waveform information available at larger TW durations, a relationship examined subsequently in this work.

This study employs the traditional GMPE method (Morikawa et al., 2010), which directly predicts JMA seismic intensity, as a benchmark to systematically evaluate the predictive accuracy and stability of our CONIP model for seismic intensity estimation. Results demonstrate that CONIP achieves significantly higher precision compared to the GMPE approach, maintaining robust performance even under extremely short time windows (e.g., TW=1 s). Further analysis reveals inherent systematic biases in GMPE predictions, with statistical errors quantified as 0.41$\pm$0.54 units, and a recall rate of 36.7%, representing substantially larger deviations than those observed for CONIP-based predictions (Fig. 5). The observed bias likely arises from a partial mismatch between the Morikawa et al. (2010) GMPE and the specific attributes of the KiK-net strong-motion records used in this study, such as local site conditions, magnitude-distance distribution, or event depth. Although a well-calibrated GMPE should in principle exhibit minimal systematic deviation, the persistent bias suggests that the adopted GMPE may not be fully optimized for real-time EEW applications or for the exact seismic context of our test events. Another possibility is that operational EEW systems, such as the national system in Japan, employ empirically adjusted attenuation models or incorporate additional corrections not captured by the standard equation. For these reasons, we treat the GMPE comparison here as an illustrative baseline rather than as a definitive reference. The observed enhancements in both accuracy (28.8% increase in recall rate for TW=1 s) and stability (27.8% lower variance for TW=10 s) further support the robustness of ML-based seismic intensity estimation frameworks, particularly their enhanced reliability in real-time early warning scenarios requiring rapid data processing within constrained temporal windows, where traditional methods struggle with systematic overestimation tendencies.

To assess the confidence of our CONIP model, we employ a Monte Carlo (MC) simulation approach. This method introduces random perturbations to the input data to quantify the model's uncertainty. Taking TW=3 s as an example, we randomly select 500 records to accelerate the calculation, for each test sample we generate 50 noisy variants by adding Gaussian noise (stan-

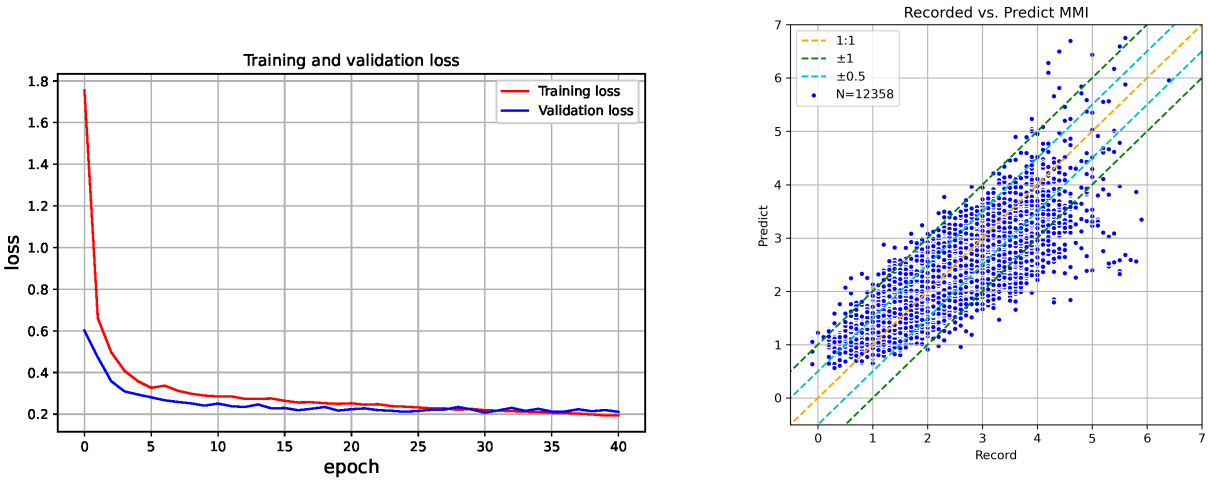

Figure 4 Evaluation of CONIP model training dynamics and seismic intensity prediction accuracy using a 3-second time window (TW=3 s). Training required approximately 14 minutes for 40 epochs and 25 minutes for 100 epochs. Panel (a) presents training and validation loss curves as a function of training epochs, while panel (b) compares predicted vs. observed JMA intensities for the test dataset using the optimized model.

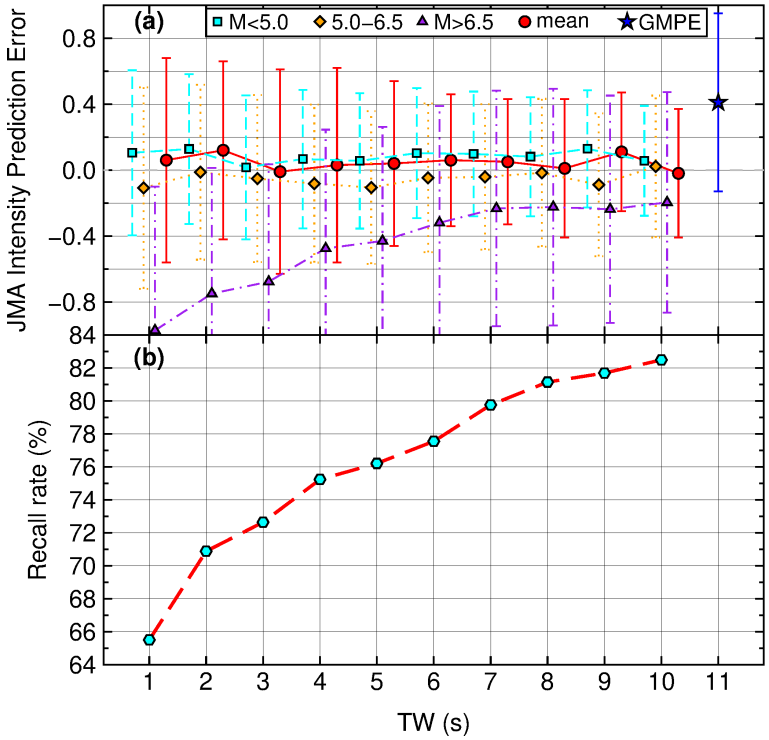

Figure 5 Statistical moments (mean $\pm 1\sigma$) of JMA intensity predictions stratified by magnitude bins (panel a) and model accuracy index (panel b) across varying TW in the test dataset. In panel a, cyan squares represent earthquakes with magnitudes less than 5.0, orange diamonds indicate events with magnitudes between 5.0 and 6.5, purple triangles denote earthquakes greater than 6.5 in magnitude. Red circles show the mean values per TW, while blue pentagrams indicate the prediction biases of the GMPE for comparison. Colored dots indicate mean values per TW, with vertical error bars denoting standard deviations.

dard deviation set to 0.1), and then perform repeated predictions using the trained model. By analyzing the distribution of predictions, we compute the variance for each sample, where a larger variance indicates higher uncertainty (lower confidence) in the model's prediction for that input. Furthermore, we compare the predicted absolute means with ground-truth labels to calculate prediction errors and examine the relationship between variance and error, thereby validating whether the model's uncertainty estimates align with its actual predictive performance.

As shown in Fig. 6, which displays the relationship between prediction variance and absolute error, we observed that when the model reports low uncertainty (small variance), the predictions are very accurate (small errors). Quantitatively, the Spearman

rank correlation between variance and absolute error is ρ=0.32 (p<0.01), suggesting a modest but statistically significant tendency for higher uncertainty to coincide with larger errors. More critically, we evaluate calibration by computing the fraction of observations falling within the 95% prediction interval derived from the MC samples. We find a prediction interval coverage probability (PICP) of 96.7%, indicating that the model's uncertainty estimates are slightly conservative but well-calibrated, a desirable property for operational EEW systems where underconfidence is safer than overconfidence. These results support the use of CONIP's uncertainty output as a practical confidence metric in real-time applications.

To evaluate the statistical significance of our CONIP model, we conducted a permutation test with 1000 randomized iterations (for TW=3 s, Fig. 7). The model demonstrated meaningful predictive capability on the test dataset, achieving a mean squared error (MSE) of 0.706 and coefficient of determination (R^2) of -0.069. To establish whether this performance exceeded chance levels, we generated a null distribution by systematically shuffling the relationship between predictions and true values while preserving feature data. For each permutation cycle, we computed the same evaluation metrics used for the true model, creating comprehensive reference distributions for both MSE and R^2. The true model's performance metrics showed clear separation from the permutation distributions. These p-values were 0.001 for MSE and 0.001 for R^2, both falling below the 0.05 significance threshold. Visual inspection of the distribution plots showed that the true model's MSE value was in the lower tail of the permuted MSE distribution, while its R^2 value was in the upper tail of the permuted R^2 distribution. This directional alignment is favorable because MSE quantifies the average squared prediction error, where lower values indicate a better fit, and thus a true MSE significantly lower than the permuted values reflects non-random predictive accuracy. Conversely, R^2 measures the proportion of variance explained relative to a naive mean predictor, with higher values indicating greater explanatory power, so a true R^2 significantly higher than the permuted values confirms the presence of meaningful predictive structure. This pattern indicates that our model detects meaningful patterns in the data rather than random noise. The statistically significant results (p<0.05 for both metrics) provide strong evidence that the model performs much better than random chance, confirming its practical usefulness for the regression task at hand.

4.2 A test event of our hybrid ground motion approach

To showcase the effectiveness of our hybrid ground motion estimation approach, we analyzed the West off Fukuoka earthquake that was not incorporated in the CONIP model training. This earthquake struck on March 20, 2005 (2005-03-20, 01:53:42.33 UTC), with a magnitude of M_j 7.0 (M_W 6.6) and a focal depth of 9.0 km. The hypocenter is located at 33.7392° N, 130.1763° E, as determined by JMA. The event occurred offshore, approximately 9 km west of Fukuoka City, a major urban center with a population exceeding 1.5 million, and generated strong shaking (JMA intensity 6-) in densely populated areas of northern Kyushu. The earthquake caused severe damage, including 1 death, over 500 injuries, and multiple buildings damaged in the vicinity of the epicenter (Nishimura et al., 2006). This event serves as a compelling test case due to its offshore location near the densely populated Fukuoka metropolitan area, the availability of high-quality strong-motion records from KiK-net and K-NET, and its established use in previous EEW research as a benchmark for assessing real-time ground motion prediction. For this reason, we deliberately excluded it from the training dataset and reserved it exclusively as an independent test example, enabling direct and consistent comparison with other published EEW methodologies that have also evaluated performance on this well-documented event.

We conducted an offline simulation using KiK-net strong-motion records time-aligned to the earthquake's source origin time. Fig. 8 illustrates the epicentral geometry and station layout used in this evaluation. The results demonstrate that our hybrid ground motion estimation approach effectively captures the spatiotemporal evolution of shaking intensity, highlighting its potential to enhance real-time seismic hazard assessments during early warning.

Next, we utilize the hybrid ground motion approach as outlined earlier to estimate the ground motions of this event. Unfortunately, we are unable to use our trained model for the first few seconds, as no triggered data is available. Therefore, we begin our estimation 6.84 seconds after the event occurs. Fig. 9 displays snapshots of the ground motion predictions for 5, 10, 15, and 20 seconds after the start of the estimation, corresponding to 11.84, 16.84, 21.84, and 26.84 seconds after the earthquake origin time, respectively. To illustrate the potential impact of the event, our method generates spatially continuous estimates of seismic intensity, in contrast to the Japan national EEW system, which issues a single JMA intensity value per county-level administrative division. This high-resolution output enables differentiated hazard assessments within the same jurisdiction, reflecting local variations in ground motion. We utilize the Kriging interpolation algorithm to produce the contour map based on the estimated JMA intensity on each observation site. Interestingly, our hybrid approach may underestimate the ground motion at the beginning of the event, and parts of the potential damage areas may not have enough time to take action. However, our hybrid method can still provide useful response times for those in further areas. For example, 10 seconds after the first trigger, according to our theoretical calculation results, only 8 KiK-net stations are triggered, but we can still provide warnings to surrounding areas that have suffered severe shaking, such as in Sasebo city. However, the warning times provided by our hybrid method are relatively short, typically only a few seconds, likely due to the nature of the onsite warning approach, which we will discuss later. To evaluate the reliability of the onsite seismic intensity prediction results generated by our hybrid method,

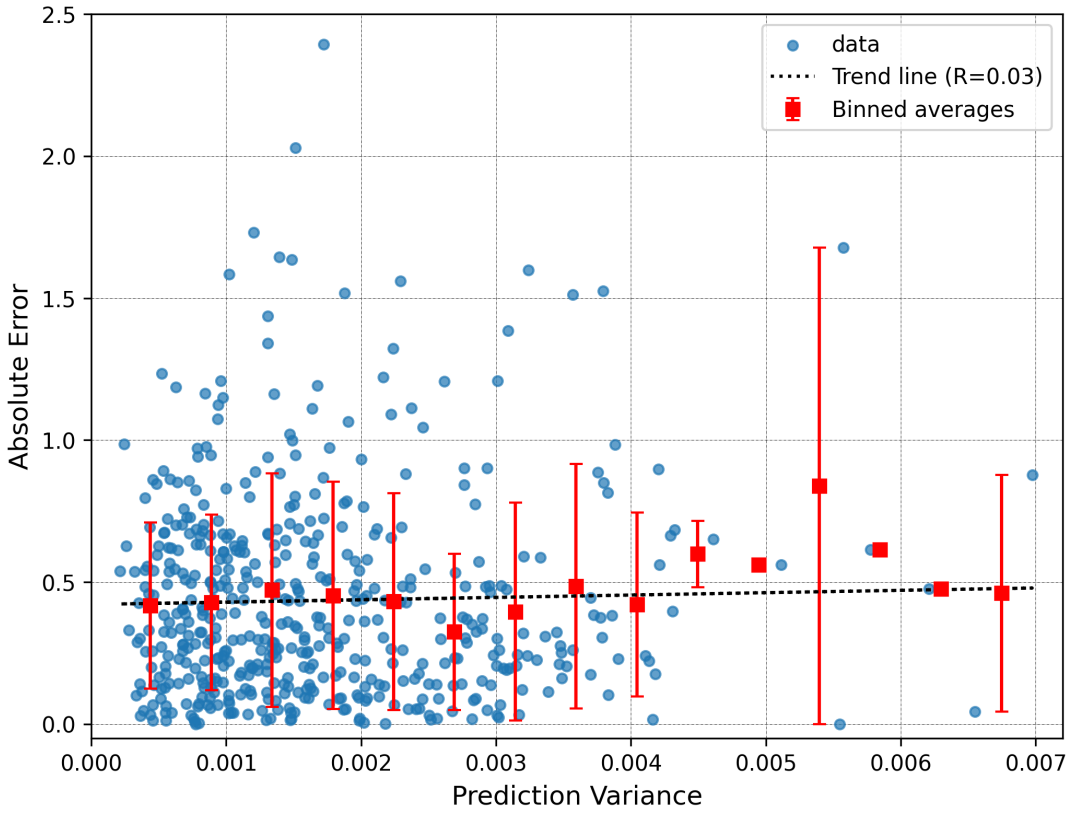

Figure 6 Monte Carlo simulation analysis of prediction confidence (TW=3 s). Scatter plot illustrating the relationship between prediction variance (uncertainty) and absolute error for onsite seismic intensity estimates. Each circle represents a test sample, with higher variance indicating greater model uncertainty. The red squares with error bars denote the binned averages of absolute error and their corresponding standard deviations within each bin of prediction variance. The black dashed line is the linear trend line fitted to all data points, indicating an overall positive probabilistic relationship between uncertainty and error.

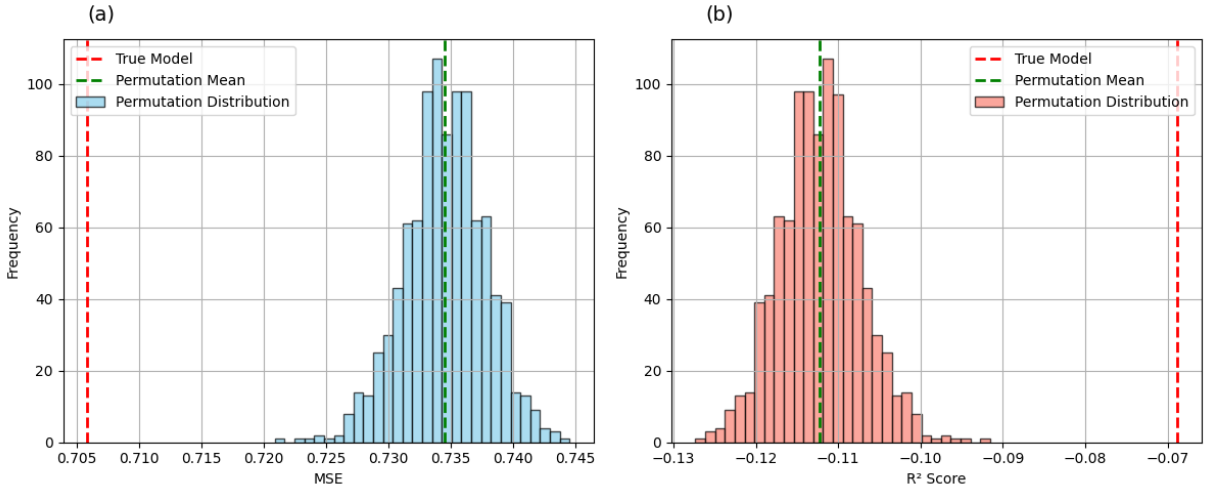

Figure 7 Permutation test results for the CONIP model (TW=3 s). Panel (a) displays the null distribution of MSE from 1000 random permutations, with the red line indicating the model's true MSE and the green line showing the mean of permuted MSE values. Panel (b) presents equivalent results for the R2 metric, demonstrating statistical significance relative to random guessing.

we randomly select three stations in this event, namely SAGH01, FKOH01, and KMMH02 (shown in Fig. 8), and compare the onsite seismic intensity predictions of the original PLUM method and our hybrid method for each station, as shown in Fig. 10. It is evident from this figure that our new method estimates seismic intensity more rapidly, and eventually converges to the real observations. Our limit of ten seconds during our CONIP model training means that we cannot predict the seismic intensity beyond this point, whereby we choose the

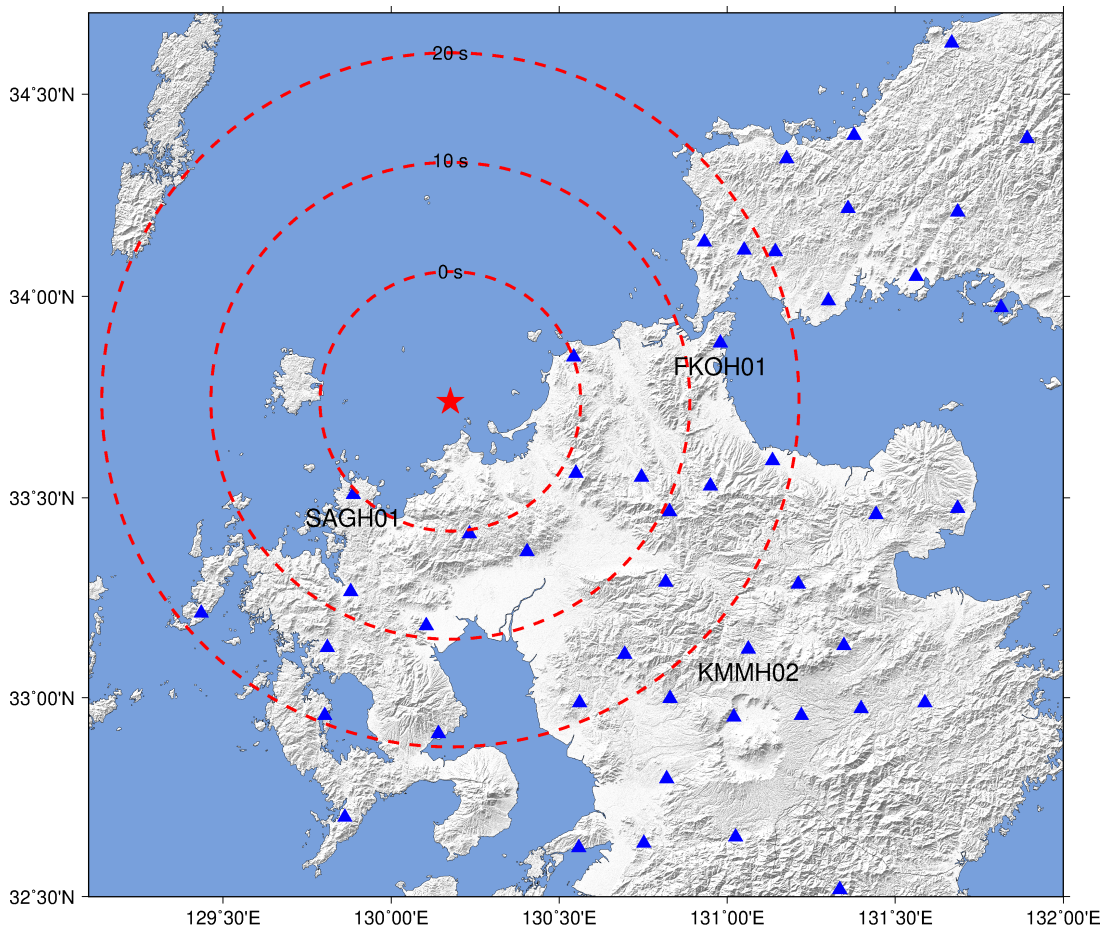

Figure 8 Spatial distribution of the 2005 West Off Fukuoka Earthquake hypocenter (red star) and KiK-net stations (blue triangles) employed in this study. Three concentric circles centered at the epicenter denote key temporal milestones in the EEW timeline: the innermost circle ("0 s") marks the theoretical P-wave front location at the moment of the first station trigger, which occurred 6.84 s after the origin time; the middle and outer circles ("10 s" and "20 s") represent the P-wave front positions 10 and 20 seconds after the initial trigger (16.84 and 26.84 seconds after the origin time), respectively, assuming an apparent P-wave velocity of 6.0 km/s . A total of 8 and 21 KiK-net stations were triggered within 10 s and 20 s after the first detection, respectively. Stations labeled with their codes (below) are referenced in Fig. 10.

value of Ir for this situation. The result will be the same for a long TW. To quantitatively assess the performance advantage of the proposed hybrid framework over the original PLUM method, we compare the times required to achieve a final JMA intensity estimate within ± 0.1 units of the observed value across all 58 stations used in this study during this event (Fig. 11). As shown in Fig. 11(a), the hybrid method consistently delivers final predictions earlier than PLUM, particularly at larger hypocentral distances. The difference in lead time is further illustrated in Fig. 11(b), where each data point represents the time advantage of the hybrid approach. For stations located within 100 km of the hypocenter, the mean additional lead time is 5.49 seconds, with individual gains ranging from 2 to over 10 seconds. This improvement reflects the enhanced early estimation capability of the ML-based model, which leverages spatial correlations and learned patterns more effectively than traditional single-station triggering logic.

In summary, through the demonstration of this event, we illustrate the practicality of our proposed hybrid ground motion estimation approach. However, it can also be seen from this example that the warning times provided by our method are relatively short. We believe this is related to the nature of the onsite warning approach and will discuss it further.

5 Discussion

We are endeavoring to integrate advanced ML algorithms into the earthquake early warning (EEW) process to enhance the accuracy of seismic intensity predictions and the practicality of the EEW system. Our results reaffirmed the superior performance of ML models like CONIP in rapid seismic intensity estimation. Although our CONIP model is only a four-layer CNN-based network and not as complex as other ML models, such as the LSTM or transformer models (e.g., Münchmeyer et al., 2021b), it can still provide more reliable ground motion predictions than traditional GMPE-based methods even with short time window records (e.g. TW=1 s). As mentioned earlier, the CONIP model has limitations in estimating ground motion fields and can only predict onsite seismic intensities, reducing the effectiveness of the method for off-shore events and events in regions of minimal station coverage. To integrate the CONIP model into an actual EEW processing system,

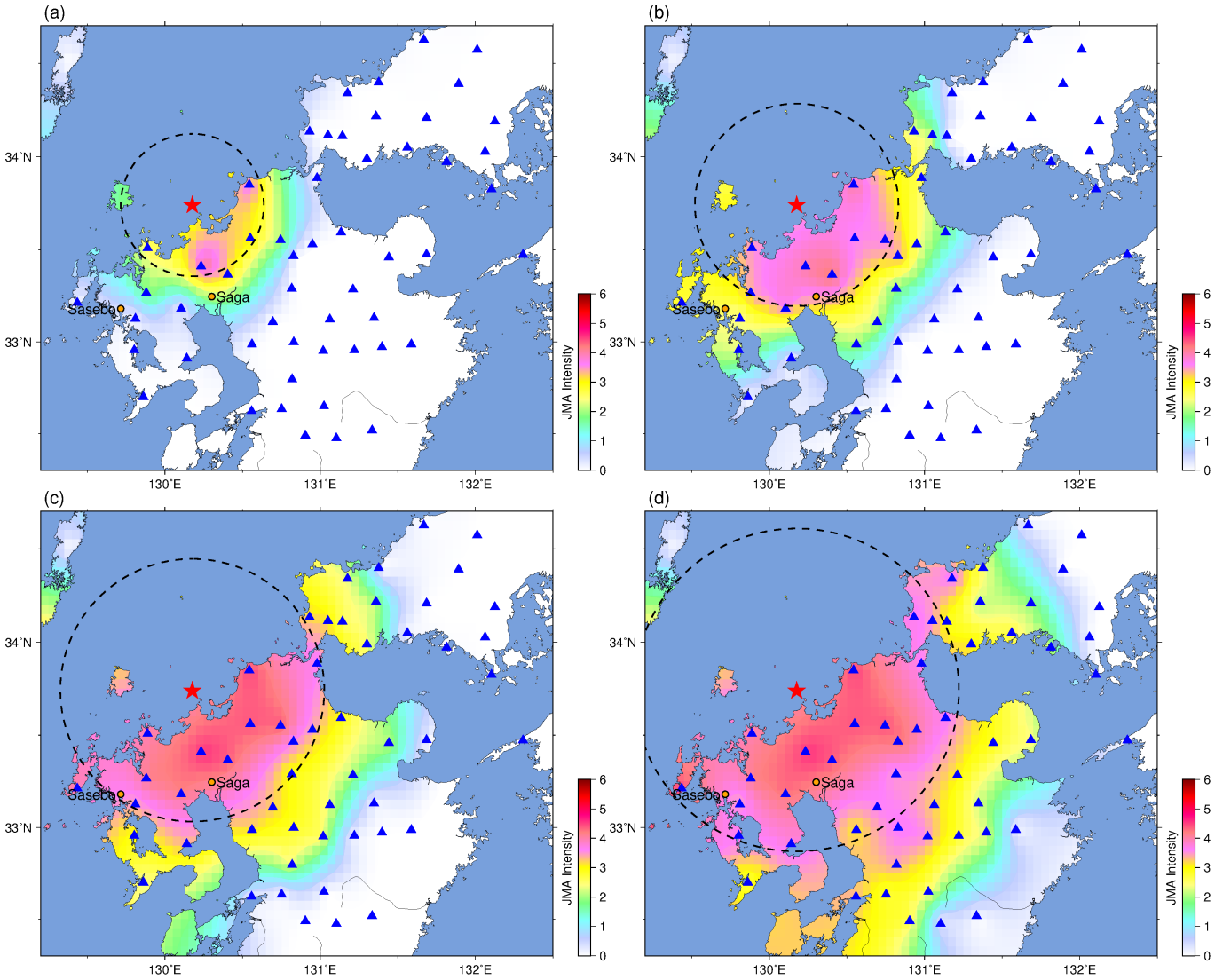

Figure 9 Temporal snapshots of ground motion predictions for the 2005 West Off Fukuoka Earthquake. Subfigures (a)-(d) depict predicted seismic intensity distributions at 5, 10, 15, and 20 seconds after initial trigger detection, corresponding to 11.84, 16.84, 21.84, and 26.84 seconds after the earthquake origin time, respectively. Blue triangles denote KiK-net stations as shown in 8, while labeled orange dots mark the locations of Saga and Sasebo cities. Dashed black circles illustrate theoretical S-wave propagation fronts at each time step. Colored contour maps represent estimated JMA intensity levels across the region, with color shading indicating intensity values on the JMA scale.

we have proposed a new hybrid approach inspired by the practical PLUM method (Kodera et al., 2018). In our hybrid approach, instead of real-time seismic intensity calculations at each site, we use the CONIP model to estimate final ground motion after the station is triggered, while still using the original PLUM processing strategy for ground motion field estimation. This combined approach can not only maintain the high reliability of the present PLUM algorithm prediction results but also promote the efficiency of seismic intensity predictions. Since our CONIP mode is a lightweight model, the time required for on-site intensity prediction is negligible, thus this hybrid approach introduces no additional computational delay. Our study shows this hybrid approach can provide end-users with longer warning times.

While ML models like CONIP demonstrate clear advantages in computational speed and prediction accuracy, their performance is inherently constrained by two critical factors: physical limits and data dependency. First, ML predictions cannot transcend the physical boundaries of seismic wave propagation, predictions cannot precede the arrival of seismic waves at a site, and accuracy improves with longer time windows (e.g., TW=10 s vs. TW=1 s) as more waveform information becomes available. Second, the reliability of ML models is highly sensitive to the quality and representativeness of training datasets. For instance, our CONIP model exhibits reduced generalizability when applied to regions with sparse historical seismic records (e.g., California or Western China), as highlighted in later sections. These limitations underscore the necessity of hybrid approaches that integrate data-driven models with physics-based principles to balance computational efficiency and predictive reliability.

During CONIP model training, it was observed that approximately 13% of the records' magnitude exceeded M_j 6.0. As mentioned in the previous section, we

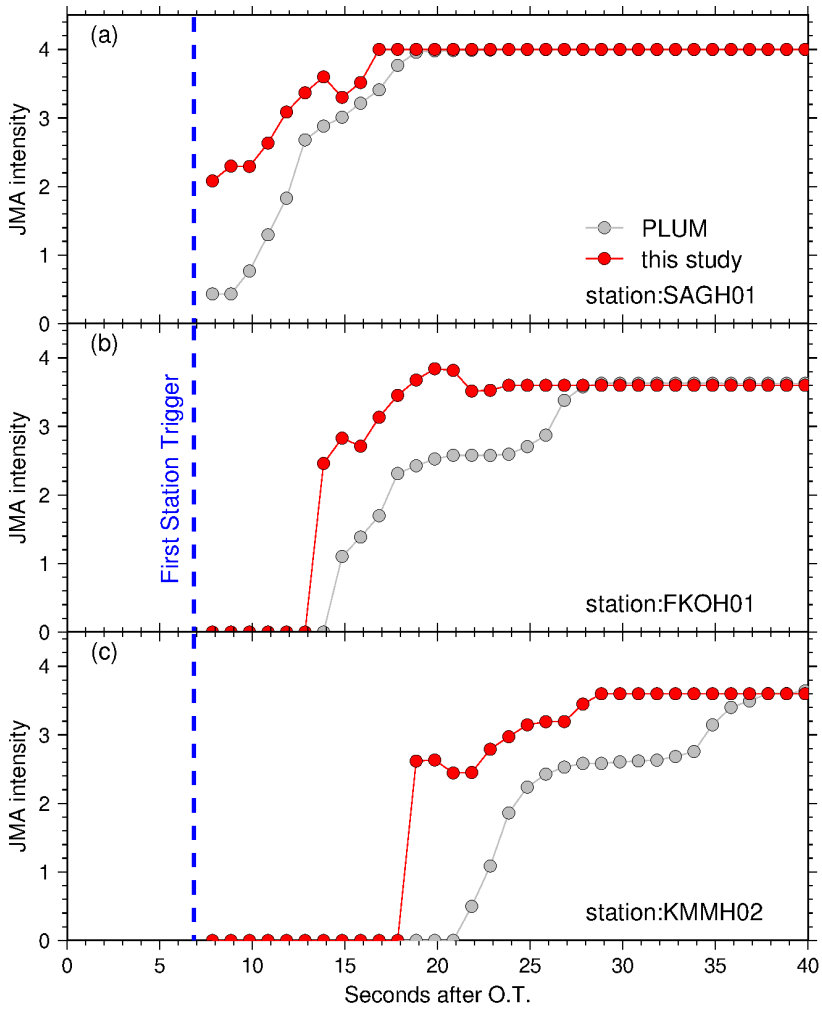

Figure 10 Comparison of JMA seismic intensity predictions for stations (a) SAGH01, (b) FKOH01, and (c) KMMH02 using the original PLUM method (gray markers) and the proposed hybrid approach (red markers). Station locations are indicated in Fig.8.

wanted to assess how our CONIP model performs on these events, which is a critical aspect of the hybrid algorithm's evaluation. Unlike traditional feature calculation methods such as P_d and τ_c, our CONIP model automatically learned deeper features during its training. However, we did notice seismic intensity underestimation, particularly when a short TW was used for larger events (i.e., M_j>6.5). This limitation stems in part from the pronounced class imbalance present in the training dataset. As illustrated in Fig. 2, high-intensity ground motion records ($\geq$4.0) account for merely a small proportion ($\sim$3.5%) of the total data. In preliminary experiments, we investigated mitigation strategies such as oversampling and loss reweighting to prioritize high-intensity samples; however, these measures delivered only marginal improvements in the prediction accuracy for strong shaking, arguably due to the inherently scarce representation of such events. To preserve the statistical fidelity of ground motion occurrence rates, we retained the original data distribution during final model training, while explicitly acknowledging this constraint as a factor influencing high-intensity performance. Moreover, according to prior research (Wu and Teng, 2004; Wu et al., 2006; Wu and Zhao, 2006; Zollo et al., 2006; Olson and Allen, 2005; Yih-Min et al., 2007; Crowell et al., 2009; Wurman, 2010), such phenomena could be linked to a lack of information when restricted TWs are employed, particularly for large magnitude earthquakes with long durations. This aligns with the physical limitations discussed above, as seismic intensity predictions for large events will likely be underestimated if the source duration is significantly longer than the present TW. In brief, even though an ML algorithm can derive more useful features from waveform data, it cannot overcome the boundaries of physics. As a consequence, the initial ground motion predictions of our hybrid technique will likely be underestimated for large events, but with TW expansion, they will be rapidly updated to approach actual values. On the other hand, an alternative explanation for this phenomenon may be that the scarcity of large-magnitude events during model training limits the learning of effective event-specific features, thereby compromising the model's generalization performance for large-magnitude scenarios. Recent studies have demonstrated that applying ground motion simulation algorithms can effectively mitigate the adverse effects of sparse large-magnitude data (Cui et al., 2024; Lin et al., 2023), consequently enhancing the applicability of machine learning methods. This constitutes one of the primary directions for future

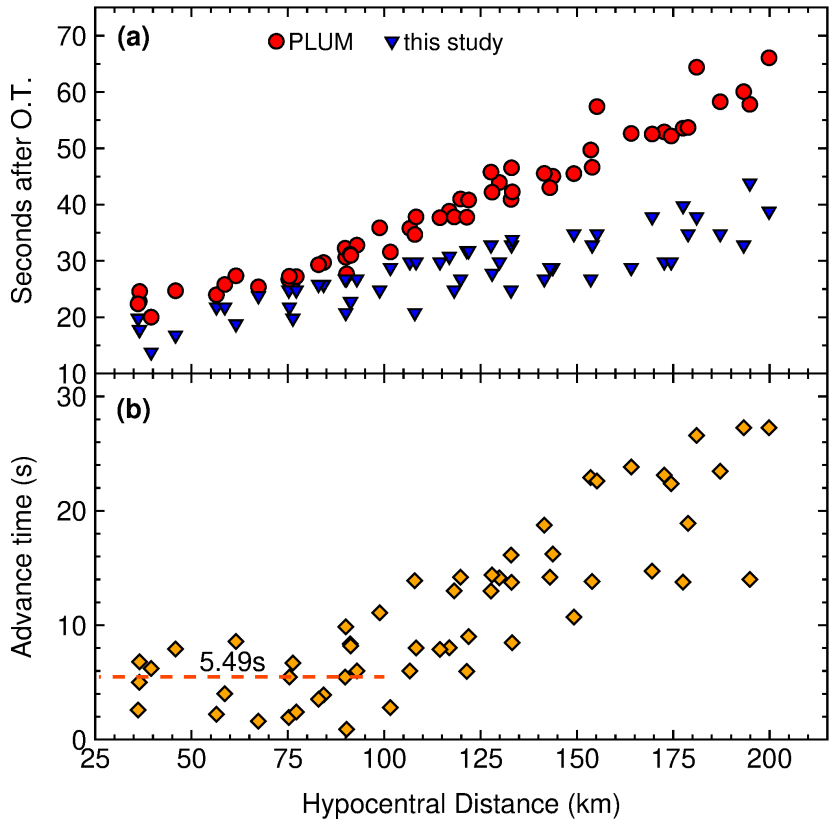

Figure 11 Comparison of final JMA intensity estimation time between the original PLUM method and the hybrid approach during the 2005 West Off Fukuoka Earthquake (58 stations). Panel (a) shows the time (in seconds after origin time, O.T.) at which each station achieved a final intensity prediction within ±0.1 unit of the observed value, where red circles denote PLUM results, blue triangles represent the hybrid approach. Panel (b) displays the difference in final intensity estimation time between the two approaches (hybrid minus PLUM), with orange diamonds indicating the additional advance time gained by the hybrid model. A horizontal dashed line marks the average gain of 5.49 seconds for stations within 100 km of the hypocenter.

research.

As a data-driven subject, the reliability and applicability of ML models are directly affected by the quality and quantity of training datasets. Luckily, KiK-net provides diverse historical data of a range of magnitudes and depths, and their quality is high. Hence, we did not intentionally remove any data during dataset construction. However, as previously mentioned, our CONIP model relies heavily on specific data features, and its accuracy in predicting ground motion, such as seismic intensity, reduces significantly when applied to different regions. Specifically, we conducted tests using historical earthquake data from western China; however, due to the scarcity and highly uneven spatial distribution of available strong-motion records, unlike the dense and uniformly instrumented KiK-net network, the model's predictive performance was notably inferior. This regional dependency further highlights the data-driven nature of ML models, as discussed in the limitations section above. Through seismic prediction residual analysis (Zhang et al., 2022), we determined that our model learned more site and source features from the present training dataset. Therefore, we conclude that our trained CONIP model does not exhibit great generalizability. A possible solution to ensuring reliable and stable seismic intensity predictions for specific application areas would be to retrain the model using datasets from the area. However, this is challenging when the area has insufficient historical records. In practice, CONIP model training typically requires tens of thousands of high-quality, spatially diverse strong-motion records, and regions with significantly fewer observations often yield unstable or overfitted models. To address this issue, we aim to construct a region-mixed training dataset by merging data collected from different areas such as California, Japan, and China to enhance model performance. We are grateful for notable works and published ML datasets like STEAD (Mousavi et al., 2019), INSTANCE (Michelini et al., 2021) and DiTing (Zhao et al., 2023), which facilitate our research.

During the CONIP model training, we limited the TW to 10 seconds. This was mainly due to the fact that the majority of the seismic records in the training dataset were close to the epicenter, within a hypocentral distance of less than 100 km. Therefore, most stations would have recorded the S wave within the TW=10 s, making it less meaningful to try to "predict" the maximum seismic intensity after the S wave had already arrived. Depending on the study region, the length of the cut-off TW may be adjusted accordingly based on the actual dataset. We believe that such adjustments will improve the applicability of our hybrid approach. It is important to note that we did not eliminate records that had already captured the S wave as the TW length increased. One reason for this is that it would have been difficult to do so, and another reason is that it may be

more realistic to include such cases in practical applications. It is often challenging to accurately estimate source parameters in short TWs, especially for large events.

One issue worth mentioning is that in the original PLUM method, site effects of different stations needed to be considered for ground motion real-time estimation. However, in our hybrid method, there is no such step. One of the important reasons is that we select only the stations adjacent to the triggering station for "undamped propagation" instead of all stations within 30 km. Although this may reduce the warning time provided, we assume that the difference in site effects between stations will be relatively small. Therefore, we do not consider site effects correction in our approach. However, we do recognize that doing so is not fully justified, as existing studies have shown that site effects are quite complex even within a small area. Therefore, we are trying to incorporate the spatial distribution of site features into CONIP models to provide more accurate estimates.

Evaluating processing performance for dense aftershocks following a significant earthquake is also a crucial aspect of our hybrid approach. As such, our hybrid algorithm largely relies on the processing ability of the PLUM method, as the CONIP onsite seismic intensity prediction is only engaged when the real-time seismic intensity calculation exceeds 0.5. Therefore, regarding dense aftershocks, the primary function of our ML model is to improve onsite seismic intensity prediction. However, according to previous analyses (Kodera et al., 2018; Cochran et al., 2019), the PLUM method has an excellent performance in event detection, allowing for no missed significant events. Hence, we anticipate that our hybrid approach will also perform well in scenarios where dense aftershocks are present. A collection of aftershock sequences has been acquired to further test this hybrid method in a future study.

6 Conclusion

Our research introduces a hybrid approach for real-time ground motion prediction in EEW systems. This novel methodology integrates ML algorithms with the original PLUM method. The combination leverages the strengths of both techniques: ML's adaptive pattern recognition capabilities complement PLUM's robust wavefield estimation. This synergistic design represents an advancement beyond standalone implementations of either approach. Test results demonstrate the improvements in prediction accuracy and timeliness. The hybrid approach achieves higher accuracy, both in JMA intensity amplitude and spatial distribution, in short-term predictions compared to the original PLUM algorithm. This translates to extended lead times (1-5 seconds earlier within in the first 10 seconds post-origin) for end-users, providing critical additional seconds for protective actions. Specifically, the method reduces instances of underestimation during large seismic events while maintaining computational efficiency.

It should be noted that the CONIP model was trained exclusively on high-quality strong-motion data from Japan, primarily sourced from the KiK-net. Consequently, its performance is optimized for the seismic characteristics, site conditions, and station geometry typical of this region. The core of the hybrid approach proposed in this study is to develop a locally tailored CONIP model that leverages region-specific data to achieve accurate and timely ground motion estimation. Direct application of the trained model to other tectonic settings, such as California or western China, may yield suboptimal results due to differences in source mechanisms, crustal structure, and, critically, the sparsity or uneven distribution of observational data. This regional dependency underscores the importance of local data availability and the need for retraining or adaptation when deploying such data-driven approaches in new areas.

The framework demonstrates practical deployment potential in operational EEW environments. Its computational efficiency ensures processing delays of only a few millisecond when generating on-site intensity predictions. The architecture leveraging CONIP's timeliness when a few stations' data are available while preserves PLUM's stability when sufficient stations are triggered. This balance makes the hybrid approach suitable for integration into existing EEW infrastructure. By resolving limitations in conventional methods while maintaining real-time performance, the hybrid approach offers a viable enhancement for earthquake early warning capabilities. The methodology has been validated using real seismic data, confirming its potential to improve warning reliability in active seismic regions.

Data and code availability

All the strong motion data used in this study can be downloaded online (http://www.kyoshin.bosai.go.jp, National Research Institute for Earth Science and Disaster Resilience, 2025, last accessed July 2025). JMA unified earthquake catalog and JMA travel time-table are available online (https://hinetwww11.bosai.go.jp/auth/JMA, last accessed July 2025). Our CNN model and python code are available at https://github.com/ZhangHCFJEA/HybridGroundMotion.git. Figures in this article were prepared using the Generic Mapping Tools (GMT, Wessel et al., 2019) and Matplotlib (https://matplotlib.org). The TensorFlow libraries (Abadi et al., 2016, https://www.tensorflow.org, last accessed July 2025) were used to train and build our ML model. The ObsPy (Beyreuther et al., 2010; Krischer et al., 2015) toolbox was used for data reading, writing, and trimming the strong motion records.

Competing interests

The authors declare that they have no competing interests. All authors have no personal, financial, professional, or other relationships that could inappropriately influence (or be perceived to influence) the work reported in this manuscript.

Funding

This work was supported by National Natural Science Foundation of China under Grant No. 52378523. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Acknowledgements

We sincerely thank the Editor, Dr. Pablo Heresi, and the reviewers, Dr. E.S. Cochran and Dr. Debi Kilb, for their thoughtful comments and constructive feedback, which have greatly improved the quality and clarity of this manuscript. The authors would like to express their sincere gratitude to National Research Institute for Earth Science and Disaster Resilience (2019) for providing access to the strong ground motion observation data used in this study, which laid a solid foundation for the validation of the proposed hybrid model. The support provided by the China Scholarship Council (CSC) during a visit of Zhang Hongcai (201904190007) to the University of Oregon is also acknowledged.

References

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., et al. {TensorFlow}: a system for {Large-Scale} machine learning. In *12th USENIX symposium on operating systems design and implementation (OSDI 16)*, pages 265–283, 2016. doi: 10.48550/ARXIV.1605.08695.

Abdalzaher, M. S., Soliman, M. S., Krichen, M., Alamro, M. A., and Fouda, M. M. Employing Machine Learning for Seismic Intensity Estimation Using a Single Station for Earthquake Early Warning. *Remote Sensing*, 16(12), 2024. doi: 10.3390/rs16122159.

Adler, A., Araya-Polo, M., and Poggio, T. Deep recurrent architectures for seismic tomography. In *81st EAGE conference and exhibition 2019*, volume 2019, pages 1–5. European Association of Geoscientists & Engineers, 2019. doi: 10.3997/2214-4609.201901512.

Agius, M. R. and Galea, P. A single-station automated earthquake location system at Wied Dalam Station, Malta. *Seismological Research Letters*, 82(4):545–559, 2011. doi: 10.1785/gssrl.82.4.545.

Al Banna, M. H., Taher, K. A., Kaiser, M. S., Mahmud, M., Rahman, M. S., Hosen, A. S., and Cho, G. H. Application of artificial intelligence in predicting earthquakes: state-of-the-art and future challenges. *IEEE Access*, 8:192880–192923, 2020. doi: 10.1109/ACCESS.2020.3029859.

Allen, R. M. The ElarmS earthquake early warning methodology and application across California. In *Earthquake early warning systems*, pages 21–43. Springer, 2007. doi: 10.1007/978-3-540-72241-0_3.

Allen, R. M. and Melgar, D. Earthquake early warning: Advances, scientific challenges, and societal needs. *Annual Review of Earth and Planetary Sciences*, 47(1):361–388, 2019. doi: 10.1146/annurev-earth-053018-060457.

Allen, R. M., Kong, Q., and Martin-Short, R. The MyShake platform: A global vision for earthquake early warning. *Pure and Applied Geophysics*, 177(4):1699–1712, 2020. doi: 10.1007/s00024-019-02337-7.

Araya-Polo, M., Jennings, J., Adler, A., and Dahlke, T. Deep-learning tomography. *The Leading Edge*, 37(1):58–66, 2018. doi: 10.1190/tle37010058.1.

Araya-Polo, M., Adler, A., Farris, S., and Jennings, J. Fast and accurate seismic tomography via deep learning. In *Deep learning: Algorithms and applications*, pages 129–156. Springer, 2019a. doi: 10.1007/978-3-030-31760-7_5.

Araya-Polo, M., Farris, S., and Florez, M. Deep learning-driven velocity model building workflow. *The Leading Edge*, 38(11): 872a1–872a9, 2019b. doi: 10.1190/tle38110872 a1.1.

Aurenhammer, F. Voronoi diagrams—a survey of a fundamental geometric data structure. *ACM computing surveys (CSUR)*, 23(3): 345–405, 1991. doi: 10.1145/116873.116880.

Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y., and Wassermann, J. ObsPy: A Python toolbox for seismology. *Seismological Research Letters*, 81(3):530–533, 2010. doi: 10.1785/gssrl.81.3.530.

Bloemheuvel, S., van den Hoogen, J., Jozinović, D., Michelini, A., and Atzmueller, M. Graph neural networks for multivariate time series regression with application to seismic data. *International Journal of Data Science and Analytics*, 16(3):317–332, 2023. doi: 10.1007/s41060-022-00349-6.

Böse, M., Heaton, T. H., and Hauksson, E. Real-time finite fault rupture detector (FinDer) for large earthquakes. *Geophysical Journal International*, 191(2):803–812, 2012. doi: 10.1111/j.1365-246X.2012.05657.x.

Böse, M., Smith, D. E., Felizardo, C., Meier, M.-A., Heaton, T. H., and Clinton, J. F. FinDer v.2: Improved real-time ground-motion predictions for M2–M9 with seismic finite-source characterization. *Geophysical Journal International*, 212(1):725–742, 2018. doi: 10.1093/gji/ggx430.

Brown, H. M., Allen, R. M., and Grasso, V. F. Testing elarms in Japan. *Seismological Research Letters*, 80(5):727–739, 2009. doi: 10.1785/gssrl.80.5.727.

Chen, D.-Y., Hsiao, N.-C., and Wu, Y.-M. The Earthworm based earthquake alarm reporting system in Taiwan. *Bulletin of the Seismological Society of America*, 105(2A):568–579, 2015. doi: 10.1785/0120140147.

Chitkeshwar, A. The Role of Machine Learning in Earthquake Seismology: A Review: A. Chitkeshwar. *Archives of Computational Methods in Engineering*, 31(7):3963–3975, 2024. doi: 10.1007/s11831-024-10099-2.

Clements, T., Cochran, E. S., Baltay, A., Minson, S., and Yoon, C. GRAPES: earthquake early warning by passing seismic vectors through the grapevine. *Geophysical Research Letters*, 51(9): e2023GL107389, 2024. doi: 10.1029/2023GL107389.

Cochran, E. S., Bunn, J., Minson, S. E., Baltay, A. S., Kilb, D. L., Kodera, Y., and Hoshiba, M. Event detection performance of the PLUM earthquake early warning algorithm in southern California. *Bulletin of the Seismological Society of America*, 109(4): 1524–1541, 2019. doi: 10.1785/0120180326.

Crowell, B. W., Bock, Y., and Squibb, M. B. Demonstration of earthquake early warning using total displacement waveforms from real-time GPS networks. *Seismological Research Letters*, 80(5): 772–782, 2009. doi: 10.1785/gssrl.80.5.772.

Cui, W., Chen, K., Wei, G., Lyu, M., and Zhu, F. Simultaneous magnitude and slip distribution characterization from high-rate GNSS using deep learning: case studies of the 2021 M w 7.4 Maduo and 2023 Turkey doublet events. *Geophysical Journal International*, 238(1):91–108, 2024. doi: 10.1093/gji/ggae140.

Diersen, S., Lee, E.-J., Spears, D., Chen, P., and Wang, L. Classification of seismic windows using artificial neural networks. *Procedia computer science*, 4:1572–1581, 2011. doi: 10.1016/j.procs.2011.04.170.

Doi, K. The operation and performance of earthquake early warnings by the Japan Meteorological Agency. *Soil Dynamics and Earthquake Engineering*, 31(2):119–126, 2011. doi: 10.1016/j.soildyn.2010.06.009.

Hoshiba, M. Real-time prediction of ground motion by Kirchhoff-Fresnel boundary integral equation method: Extended front detection method for earthquake early warning. *Journal of Geophysical Research: Solid Earth*, 118(3):1038–1050, 2013. doi: 10.1002/jgrb.50119.

Hoshiba, M. and Aoki, S. Numerical shake prediction for earthquake early warning: Data assimilation, real-time shake mapping, and simulation of wave propagation. *Bulletin of the Seismological Society of America*, 105(3):1324–1338, 2015. doi: 10.1785/0120140280.

Hoshiba, M. and Ozaki, T. Earthquake early warning and tsunami warning of the Japan Meteorological Agency, and their performance in the 2011 off the Pacific Coast of Tohoku Earthquake (9.0). In *Early warning for geological disasters: Scientific methods and current practice*, pages 1–28. Springer, 2014. doi: 10.1007/978-3-642-12233-0_1.

Huang, Y.-M., Chen, K.-Y., Lin, W.-W., and Chen, D.-Y. An attention-based framework with multistation information for earthquake early warnings. *IEEE Transactions on Geoscience and Remote Sensing*, 2025. doi: 10.1109/TGRS.2025.3557404.

Jiang, M.-Y., Chen, D.-Y., and Chin, T.-L. Ground-Shaking Intensity Prediction for Onsite Earthquake Early Warning Using Deep Learning. *Seismological Research Letters*, 96(1):526–537, 2025. doi: 10.1785/0220230263.

Joshi, A., Raman, B., Mohan, C. K., and Cenkeramaddi, L. R. Application of a new machine learning model to improve earthquake ground motion predictions. *Natural Hazards*, 120(1):729–753, 2024. doi: 10.1007/s11069-023-06230-4.

Khoshnevis, N. and Taborda, R. Prioritizing ground-motion validation metrics using semisupervised and supervised learning. *Bulletin of the Seismological Society of America*, 108(4): 2248–2264, 2018. doi: 10.1785/0120180056.

Kilb, D., Bunn, J. J., Saunders, J. K., Cochran, E. S., Minson, S. E., Baltay, A., O'Rourke, C., Hoshiba, M., and Kodera, Y. The PLUM earthquake early warning algorithm: A retrospective case study of West Coast, USA, data. *Journal of Geophysical Research: Solid Earth*, 126(7):e2020JB021053, 2021. doi: 10.1029/2020JB021053.

Kodera, Y., Yamada, Y., Hirano, K., Tamaribuchi, K., Adachi, S., Hayashimoto, N., Morimoto, M., Nakamura, M., and Hoshiba, M. The propagation of local undamped motion (PLUM) method: A simple and robust seismic wavefield estimation approach for earthquake early warning. *Bulletin of the Seismological Society of America*, 108(2):983–1003, 2018. doi: 10.1785/0120170085.

Kolivand, P., Saberian, P., Tanhapour, M., Karimi, F., Kalhori, S. R. N., Javanmard, Z., Heydari, S., Talari, S. S. H., Mousavi, S. M. L., Alidadi, M., et al. A systematic review of Earthquake Early Warning (EEW) systems based on Artificial Intelligence. *Earth Science Informatics*, 17(2):957–984, 2024. doi: 10.1007/s12145-024-01253-2.

Kong, Q., Allen, R. M., Schreier, L., and Kwon, Y.-W. MyShake: A smartphone seismic network for earthquake early warning and beyond. *Science advances*, 2(2):e1501055, 2016. doi: 10.1126/sciadv.1501055.

Kong, Q., Inbal, A., Allen, R. M., Lv, Q., and Puder, A. Machine learning aspects of the MyShake global smartphone seismic network. *Seismological Research Letters*, 90(2A):546–552, 2019. doi: 10.1785/0220180309.

Kriegerowski, M., Petersen, G. M., Vasyura-Bathke, H., and Ohrnberger, M. A deep convolutional neural network for localization of clustered earthquakes based on multistation full waveforms. *Seismological Research Letters*, 90(2A):510–516, 2019. doi: 10.1785/0220180320.

Krischer, L., Megies, T., Barsch, R., Beyreuther, M., Lecocq, T., Caudron, C., and Wassermann, J. ObsPy: A bridge for seismology into the scientific Python ecosystem. *Computational Science & Discovery*, 8(1):014003, 2015. doi: 10.1088/1749-4699/8/1/014003.

Kuehn, N. and Abrahamson, N. A Non-Ergodic Ground-Motion Model for Japan. 2018. https://api.semanticscholar.org/CorpusID:165156673.

Kuehn, N. M. and Abrahamson, N. A. Spatial correlations of ground motion for non-ergodic seismic hazard analysis. *Earthquake Engineering & Structural Dynamics*, 49(1):4–23, 2020. doi: 10.1002/eqe.3221.

Kunugi, T., Aoi, S., Nakamura, H., Fujiwara, H., and Morikawa, N. Real-time calculation of seismic intensity used in Japan. In *AGU Fall Meeting Abstracts*, volume 2008, pages S13C–1832, 2008. doi: 10.4294/zisin.60.243.

Li, R., Lu, X., Li, S., Yang, H., Qiu, J., and Zhang, L. DLEP: A deep learning model for earthquake prediction. In *2020 international joint conference on neural networks (IJCNN)*, pages 1–8. IEEE, 2020. doi: 10.1109/IJCNN48605.2020.9207621.

Li, Z., Meier, M.-A., Hauksson, E., Zhan, Z., and Andrews, J. Machine learning seismic wave discrimination: Application to earthquake early warning. *Geophysical Research Letters*, 45(10): 4773–4779, 2018. doi: 10.1029/2018GL077870.

Lin, J.-T., Melgar, D., Sahakian, V. J., Thomas, A. M., and Searcy, J. Real-time fault tracking and ground motion prediction for large earthquakes with HR-GNSS and deep learning. *Journal of Geophysical Research: Solid Earth*, 128(12):e2023JB027255, 2023. doi: 10.1029/2023JB027255.

Linville, L., Tibi, R., Young, C. J., and Brogan, R. Classification of Local Seismic Events in the Utah Region: A Comparison of Amplitude Ratio Methods with a Machine Learning Approach. Technical report, Sandia National Laboratories (SNL-NM), Albuquerque, NM (United States), 2018. doi: 10.1785/0120190150.

Meier, M.-A., Ross, Z. E., Ramachandran, A., Balakrishna, A., Nair, S., Kundzicz, P., Li, Z., Andrews, J., Hauksson, E., and Yue, Y. Reliable real-time seismic signal/noise discrimination with machine learning. *Journal of Geophysical Research: Solid Earth*, 124 (1):788–800, 2019. doi: 10.1029/2018JB016661.

Michelini, A., Cianetti, S., Gaviano, S., Giunchi, C., Jozinović, D., and Lauciani, V. INSTANCE–the Italian seismic dataset for machine learning. *Earth System Science Data*, 13(12):5509–5544, 2021. doi: 10.5194/essd-13-5509-2021.

Minson, S. E., Bunn, J., Cochran, E. S., Kilb, D. L., Saunders, J. K., Parker, G. A., Baltay, A., Hoshiba, M., and Kodera, Y. Real-Time Performance of the PLUM Earthquake Early Warning Method During the 2019 Ridgecrest Earthquake Sequence. In *AGU Fall Meeting Abstracts*, volume 2019, pages S44C–06, 2019. doi: 10.1785/0120200021.

Morikawa, N., Kanno, T., Narita, A., Fujiwara, H., and Fukushima, Y. A new attenuation relation of seismic intensity for Japan based on recent strong-motion records (in Japanese with English abstract). *Geophys. Bull. Hokkaido Univ.*, pages 149–158, 2010.

Mousavi, S. and Beroza, G. Deep-learning seismology: Science, 377, eabm4470, doi: 10.1126/science. abm4470, 2022. doi: 10.1126/science.abm4470.

Mousavi, S. M. and Beroza, G. C. Machine learning in earthquake seismology. *Annual review of earth and planetary sciences*, 51: 105–129, 2023. doi: 10.1146/annurev-earth-071822-100323.

Mousavi, S. M., Sheng, Y., Zhu, W., and Beroza, G. C. STanford

EArthquake Dataset (STEAD): A global data set of seismic signals for AI. *IEEE Access*, 7:179464–179476, 2019. doi: 10.1109/ACCESS.2019.2947848.

Mousavi, S. M., Ellsworth, W. L., Zhu, W., Chuang, L. Y., and Beroza, G. C. Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. *Nature communications*, 11(1):3952, 2020. doi: 10.1038/s41467-020-17591-w.

Münchmeyer, J., Bindi, D., Leser, U., and Tilmann, F. Earthquake magnitude and location estimation from real time seismic waveforms with a transformer network. *Geophysical Journal International*, 226(2):1086–1104, 2021a. doi: 10.1093/gji/ggab139.

Münchmeyer, J., Bindi, D., Leser, U., and Tilmann, F. The transformer earthquake alerting model: A new versatile approach to earthquake early warning. *Geophysical Journal International*, 225(1):646–656, 2021b. doi: 10.1093/gji/ggaa609.

National Research Institute for Earth Science and Disaster Resilience. NIED K-NET, KiK-net [dataset], 2025. doi: 10.17598/NIED.0004.

Navarro-Rodríguez, A., Castro-Artola, O. A., García-Guerrero, E. E., Aguirre-Castro, O. A., Tamayo-Pérez, U. J., López-Mercado, C. A., and Inzunza-Gonzalez, E. Recent advances in early earthquake magnitude estimation by using machine learning algorithms: a systematic review. *Applied Sciences*, 15(7):3492, 2025. doi: 10.3390/app15073492.

Nishimura, T., Fujiwara, S., Murakami, M., Suito, H., Tobita, M., and Yarai, H. Fault model of the 2005 Fukuoka-ken Seiho-oki earthquake estimated from coseismic deformation observed by GPS and InSAR. *Earth, planets and space*, 58(1):51–56, 2006. doi: 10.1186/BF03351913.

Noda, S. Deep learning estimating of epicentral distance for earthquake early warning systems. *Bulletin of the Seismological Society of America*, 114(4):2054–2064, 2024. doi: 10.1785/0120230112.

Okada, Y. Recent progress of seismic observation networks in Japan. In *Journal of Physics: Conference Series*, volume 433, page 012039, 2013. doi: 10.1088/1742-6596/433/1/012039.

Olson, E. L. and Allen, R. M. The deterministic nature of earthquake rupture. *Nature*, 438(7065):212–215, 2005. doi: 10.1038/nature04214.

Perol, T., Gharbi, M., and Denolle, M. Convolutional neural network for earthquake detection and location. *Science Advances*, 4(2): e1700578, 2018. doi: 10.1126/sciadv.1700578.

Ross, Z. E., Meier, M.-A., and Hauksson, E. P wave arrival picking and first-motion polarity determination with deep learning. *Journal of Geophysical Research: Solid Earth*, 123(6):5120–5129, 2018. doi: 10.1029/2017JB015251.

Ross, Z. E., Trugman, D. T., Hauksson, E., and Shearer, P. M. Searching for hidden earthquakes in Southern California. *Science*, 364 (6442):767–771, 2019a. doi: 10.1126/science.aaw6888.

Ross, Z. E., Yue, Y., Meier, M.-A., Hauksson, E., and Heaton, T. H. PhaseLink: A deep learning approach to seismic phase association. *Journal of Geophysical Research: Solid Earth*, 124(1): 856–869, 2019b. doi: 10.1029/2018JB016674.

Ruhl, C. J., Melgar, D., Grapenthin, R., and Allen, R. M. The value of real-time GNSS to earthquake early warning. *Geophysical Research Letters*, 44(16):8311–8319, 2017. doi: 10.1002/2017GL074502.

Satriano, C., Elia, L., Martino, C., Lancieri, M., Zollo, A., and Iannaccone, G. PRESTo, the earthquake early warning system for Southern Italy: Concepts, capabilities and future perspectives. *Soil Dynamics and Earthquake Engineering*, 31(2): 137–153, 2011. doi: 10.1016/j.soildyn.2010.06.008.

Saunders, J. K., Minson, S. E., Baltay, A. S., Bunn, J. J., Cochran, E. S., Kilb, D. L., O'Rourke, C. T., Hoshiba, M., and Kodera, Y. Real-time earthquake detection and alerting behavior of PLUM ground-motion-based early warning in the United States. *Bulletin of the Seismological Society of America*, 112(5):2668–2688, 2022. doi: 10.1785/0120220022.

Saunders, J. K., Cochran, E. S., Bunn, J. J., Baltay, A. S., Minson, S. E., and O'Rourke, C. T. Incorporating intensity distance attenuation into PLUM ground-motion-based earthquake early warning in the United States: The APPLES configuration. *Earth's Future*, 12(2):e2023EF004126, 2024. doi: 10.1029/2023EF004126.

Saunders, J. K., Cochran, E. S., and Bunn, J. J. Refinements to the attenuated propagation of local earthquake shaking (APPLES) ground-motion-based earthquake early warning algorithm. *Geophysical Journal International*, 244(2):ggaf418, 2026. doi: 10.1093/gji/ggaf418.

Spallarossa, D., Kotha, S. R., Picozzi, M., Barani, S., and Bindi, D. On-site earthquake early warning: a partially non-ergodic perspective from the site effects point of view. *Geophysical Journal International*, 216(2):919–934, 2019. doi: 10.1093/gji/ggy470.

Trugman, D. T. and Shearer, P. M. Strong correlation between stress drop and peak ground acceleration for recent M 1–4 earthquakes in the San Francisco Bay area. *Bulletin of the Seismological Society of America*, 108(2):929–945, 2018. doi: 10.1785/0120170245.

Wang, J., Xiao, Z., Liu, C., Zhao, D., and Yao, Z. Deep learning for picking seismic arrival times. *Journal of Geophysical Research: Solid Earth*, 124(7):6612–6624, 2019. doi: 10.1029/2019JB017536.

Wessel, P., Luis, J., Uieda, L., Scharroo, R., Wobbe, F., Smith, W., and Tian, D. The generic mapping tools version 6. *Geochemistry, Geophysics, Geosystems20*, pages 5556–5564, 2019. doi: 10.1029/2019GC008515.

Woollam, J., Rietbrock, A., Bueno, A., and De Angelis, S. Convolutional neural network for seismic phase classification, performance demonstration over a local seismic network. *Seismological Research Letters*, 90(2A):491–502, 2019. doi: 10.1785/0220180312.

Wu, Y.-M. and Teng, T.-l. Near real-time magnitude determination for large crustal earthquakes. *Tectonophysics*, 390(1-4):205–216, 2004. doi: 10.1016/j.tecto.2004.03.029.

Wu, Y.-M. and Zhao, L. Magnitude estimation using the first three seconds P-wave amplitude in earthquake early warning. *Geophysical research letters*, 33(16), 2006. doi: 10.1029/2006GL026871.

Wu, Y.-M., Yen, H.-Y., Zhao, L., Huang, B.-S., and Liang, W.-T. Magnitude determination using initial P waves: A single-station approach. *Geophysical Research Letters*, 33(5), 2006. doi: 10.1029/2005GL025395.

Wurman, G. *Earthquake early warning and the physics of earthquake rupture*. University of California, Berkeley, 2010.

Yang, C., Zhang, K., Chen, G., Pan, Y., Zhang, L., and Qu, L. Application of machine learning to determine earthquake hypocenter location in earthquake early warning. *IEEE Geoscience and Remote Sensing Letters*, 21:1–5, 2024. doi: 10.1109/LGRS.2023.3348107.

Yih-Min, W., Kanamori, H., Allen, R. M., and Hauksson, E. Determination of earthquake early warning parameters, τ c and Pd, for southern California. *Geophysical Journal International*, 170(2): 711–717, 2007. doi: 10.1111/j.1365-246X.2007.03430.x.

Zhang, H., Jin, X., Wei, Y., Li, J., Kang, L., Wang, S., Huang, L., and Yu, P. An earthquake early warning system in Fujian, China. *Bulletin*

of the Seismological Society of America, 106(2):755–765, 2016. doi: 10.1785/0120150143.

Zhang, H., Melgar, D., Sahakian, V., Searcy, J., and Lin, J.-T. Learning source, path and site effects: CNN-based on-site intensity prediction for earthquake early warning. *Geophysical Journal International*, 231(3):2186–2204, 2022. doi: 10.1093/gji/ggac325.

Zhang, K., Lozano-Galant, F., Xia, Y., and Matos, J. Intensity prediction model based on machine learning for regional earthquake early warning. *IEEE Sensors Journal*, 24(7):10491–10503, 2024. doi: 10.1109/JSEN.2024.3354857.

Zhao, M., CHEN, S., Fang, L., and David A, Y. Earthquake phase arrival auto-picking based on U-shaped convolutional neural network. *Chinese Journal of Geophysics*, 62(8):3034–3042, 2019. doi: 10.6038/cjg2019M0495.

Zhao, M., Xiao, Z., Chen, S., and Fang, L. DiTing: A large-scale Chinese seismic benchmark dataset for artificial intelligence in seismology. *Earthquake Science*, 36(2):84–94, 2023. doi: 10.1016/j.eqs.2022.01.022.

Zhu, H., Sun, Y., Zhao, W., Zhuang, F., Wang, B., and Xiong, H. Rapid learning of earthquake felt area and intensity distribution with real-time search engine queries. *Scientific reports*, 10(1):5437, 2020. doi: 10.1038/s41598-020-62114-8.

Zhu, L., Peng, Z., McClellan, J., Li, C., Yao, D., Li, Z., and Fang, L. Deep learning for seismic phase detection and picking in the aftershock zone of 2008 Mw7. 9 Wenchuan Earthquake. *Physics of the Earth and Planetary Interiors*, 293:106261, 2019. doi: 10.1016/j.pepi.2019.05.004.

Zhu, W. and Beroza, G. C. PhaseNet: a deep-neural-network-based seismic arrival-time picking method. *Geophysical Journal International*, 216(1):261–273, 2019. doi: 10.1093/gji/ggy423.

Zollo, A., Lancieri, M., and Nielsen, S. Earthquake magnitude estimation from peak amplitudes of very early seismic signals on strong motion records. *Geophysical Research Letters*, 33(23), 2006. doi: 10.1029/2006GL027795.