

VLPNet: A unified deep learning model for progressive single-station magnitude estimation from variable-duration P-waves in earthquake early warning

Chunde Lu

Huazhong University of Science and Technology

Aihemaitijiang Tuerhong

Huazhong University of Science and Technology

Xiaopeng Gu

Huazhong University of Science and Technology

Yongtao Ma

Huazhong University of Science and Technology

Xiaohong Long

xhlong@hust.edu.cn

Huazhong University of Science and Technology

Research Article

Keywords: Earthquake early warning, Progressive magnitude estimation, Variable-duration P-waves, Single-station observations, Deep learning

Posted Date: October 14th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11082699/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

VLPNet: A unified deep learning model for progressive single-station magnitude estimation from variable-duration P-waves in earthquake early warning

Chunde Lu^a, Aihemaitijiang Tuerhong^{a,b}, Xiaopeng Gu^a, Yongtao Ma^a, Xiaohong Long^{a,c,*}

^a School of Civil and Hydraulic Engineering, Huazhong University of Science and Technology, Wuhan 430074, PR China

^b School of Transportation, Kashi University, Kashi 844000, PR China

^c Hubei Key Laboratory of Control Structure, Huazhong University of Science and Technology, Wuhan 430074, PR China

Abstract

Earthquake early warning requires magnitude estimates to be progressively refined as P-wave observations accumulate, yet many waveform-based estimators are restricted to predefined input windows and cannot directly accommodate changing warning times. Here we develop VLPNet, a unified deep learning model for progressive single-station magnitude estimation from variable-duration borehole P waves. Starting from an automatically identified P-wave arrival, VLPNet uses the downhole vertical acceleration component (UD2) to predict the complete-record peak horizontal displacement, which is then combined with hypocentral distance through a JMA displacement-magnitude relation. The dataset comprises 13,245 station-event records from 1,916 earthquakes with JMA magnitudes greater than 4.0 and hypocentral distances of 30-200 km, recorded at 550 KiK-net stations in Japan. The earthquakes were divided chronologically at the event level into training, validation, and test sets. A single trained model accepts P-wave observation durations from 0.5 to 6.0 s at a temporal resolution of 0.01 s without duration-specific retraining. Across five random seeds, the magnitude mean absolute error decreased from 0.461 ± 0.002 at 0.5 s to 0.317 ± 0.010 at 6.0 s, while the proportion of estimates within ± 1 magnitude unit increased from $89.2 \pm 0.7\%$ to $96.5 \pm 0.2\%$. At 3.0 s, VLPNet reduced the mean absolute error by 28.0-50.9% relative to conventional estimates based on peak displacement, maximum predominant period, and their combination. Accuracy improved most rapidly during the first 3 s and more gradually thereafter. At 3.0 s, more than 93% of the estimates were within ± 1 magnitude unit for each of the three site classes. These results show that VLPNet can progressively exploit newly arriving P-wave information using one model over the investigated warning-time range, while quantifying the trade-off between warning timeliness and magnitude accuracy.

Keywords: Earthquake early warning; Progressive magnitude estimation; Variable-duration P-waves; Single-station observations; Deep learning.

*Corresponding author.

E-mail address: xhlong@hust.edu.cn (X. Long).

1. Introduction

Earthquake early warning (EEW) uses the difference in propagation velocity between the earlier-arriving P waves and the subsequent, more damaging seismic waves to provide alerts before strong ground shaking reaches target areas (Kanamori, 2005). The available warning time, although generally limited to seconds or tens of seconds, can support personal protective actions, emergency control of critical infrastructure, and mitigation of secondary hazards (Allen et al., 2009; Satriano et al., 2011; Allen and Melgar, 2019). Earthquake magnitude is a key parameter in EEW, and its estimation accuracy directly affects the reliability of subsequent ground-motion prediction and warning decisions (Allen et al., 2009; Satriano et al., 2011). However, during the initial stage of an earthquake, only a short segment of the P wave is available, and the observed waveform evolves continuously as more signal information is received. Therefore, rapidly extracting stable magnitude-related information from limited and progressively accumulated P-wave observations remains an important challenge for EEW (Satriano et al., 2011; Allen and Melgar, 2019).

Traditional single-station magnitude estimation generally relies on empirical relationships established using characteristic parameters extracted from the initial P-wave. Nakamura was among the first to use a characteristic-period parameter to characterize earthquake size, laying the foundation for rapid magnitude estimation based on the frequency content of P waves (Nakamura, 1988). Subsequently, Allen and Kanamori (2003) proposed the maximum predominant period, τ_p^{max} , to estimate magnitude from the frequency characteristics of the initial P-wave. Wu and Kanamori (2005) and Wu and Zhao (2006) further investigated the relationships of the initial P-wave peak displacement (P_d), with earthquake magnitude and

ground-motion intensity, demonstrating its capability to represent the amplitude information of early waveforms. Subsequent studies evaluated the applicability of these parameters in Japan, Taiwan, Italy, and other seismic regions and combined amplitude- and frequency-related parameters to improve the stability of magnitude estimation (Wu et al., 2007; Shieh et al., 2008; Olivieri et al., 2008; Zollo et al., 2010; Brown et al., 2011). Nevertheless, the empirical relationships based on these parameters may be affected by signal noise, propagation distance, site conditions, and observation-window duration. Moreover, conventional methods typically reduce a P-wave segment to one or a few manually defined scalar features, making it difficult to fully represent the local waveform patterns and their temporal evolution (Wolfe, 2006; Yamada and Mori, 2009; Sadeh et al., 2014; Spallarossa et al., 2019).

In recent years, advances in deep learning have provided new opportunities to extract source-related information directly from seismic waveforms. Convolutional neural networks (CNNs) have been successfully applied to earthquake detection, seismic-phase picking, magnitude estimation, and ground-motion prediction because of their ability to automatically extract hierarchical local patterns from waveform data (Mousavi et al., 2020; Zhu and Beroza, 2019; Mousavi et al., 2019). Recurrent neural networks (RNNs), particularly long short-term memory (LSTM) networks, are well suited to capturing temporal dependencies in seismic signals and have therefore been introduced into waveform-based earthquake characterization (Mousavi et al., 2019; Mousavi and Beroza, 2020; Chakraborty et al., 2022). Hybrid CNN-LSTM architectures further combine the local feature-extraction capability of CNNs with the sequential modeling capability of LSTMs, enabling nonlinear mapping from raw seismic waveforms to earthquake magnitude or other ground-motion parameters (Rafiei and Adeli,

2017; Zhu et al., 2025; Dai et al., 2024). Compared with empirical methods based on a small number of predefined parameters, these data-driven approaches can retain richer amplitude, frequency, and temporal-evolution information and reduce the information loss associated with manual feature selection.

Despite the promising performance of existing deep-learning methods, their adaptability to continuously varying P-wave observation durations remains insufficiently investigated. Many models have been developed and evaluated using waveforms with a predefined duration (Mousavi and Beroza, 2020; Chakraborty et al., 2022; Zhu et al., 2025). Consequently, models calibrated for a particular input duration may not be directly applicable at arbitrary stages of an evolving earthquake record, limiting their flexibility for continuously updated magnitude estimation. Moreover, when zero padding is used to standardize sequence lengths without explicitly retaining the valid length of each record, the padded portion may enter the recurrent computation and affect the resulting temporal representation. Although real-time deep-learning frameworks have been developed to update earthquake parameters as waveform information accumulates (Saad et al., 2022; Münchmeyer et al., 2021; Hou et al., 2024), unified single-station modeling of variable-length P-wave records has received comparatively limited attention. In addition, local site conditions can modify the amplitude and frequency characteristics of seismic signals (Sadeh et al., 2014; Spallarossa et al., 2019), whereas the consistency of deep-learning-based magnitude estimates across different site classes has received relatively limited attention. Therefore, a unified model capable of processing P-wave records of different lengths while maintaining stable performance under varying site conditions is still needed.

To address the above issues, this study proposes VLPNet, a unified framework for magnitude estimation from variable-length P-wave records. Rather than training separate networks for different observation durations, VLPNet combines CNN-based local feature extraction with LSTM-based temporal modeling and employs a packed-sequence strategy, allowing a single model to process P-wave segments with different observation durations. The main contributions of this study are threefold: (1) a unified model is developed to provide progressively updated estimates as P-wave information accumulates, without training separate models for different observation durations; (2) the model is systematically compared with conventional P_d and τ_p^{max} based methods over different P-wave observation durations; and (3) the model performance is evaluated across different site classes to assess its robustness under varying site conditions.

2. Ground-Motion Dataset and Processing

2.1 Dataset Selection and Partition

Ground-motion records were obtained from the Kiban Kyoshin Network (KiK-net), operated by the National Research Institute for Earth Science and Disaster Resilience (NIED). KiK-net stations are equipped with accelerometers installed at both the ground surface and borehole depth. For each station-event pair, the borehole vertical acceleration component (UD2) and two surface horizontal acceleration components (NS1 and EW1) were extracted. The UD2 component was used to construct the initial P-wave input, whereas the NS1 and EW1 components were used to determine the maximum horizontal displacement of the complete record.

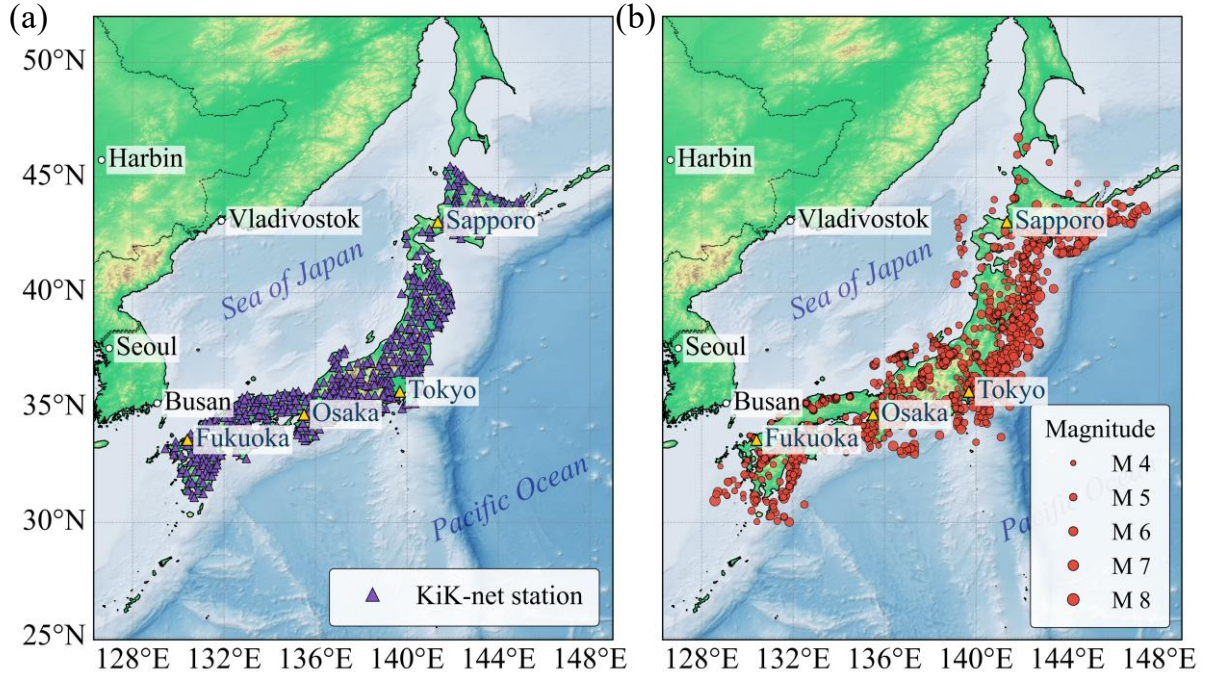


Fig. 1. Geographic distribution of KiK-net stations and earthquake epicenters.

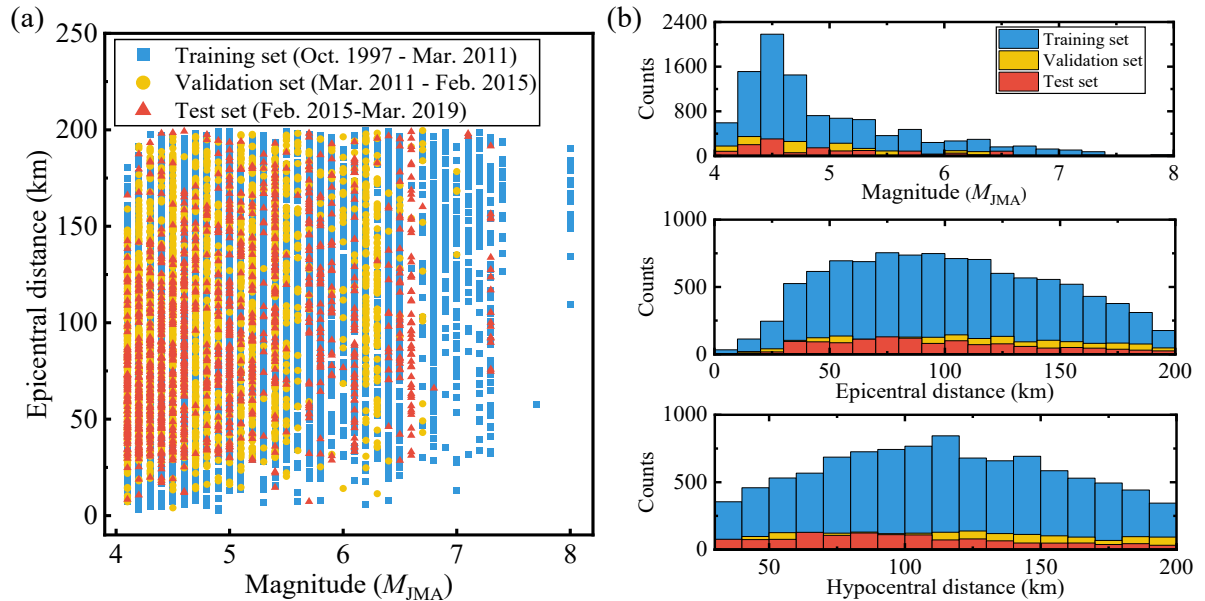


Fig. 2. Magnitude and distance distributions of the three datasets. (a) Magnitude versus epicentral distance; (b) magnitude, epicentral distance and hypocentral distance.

The records were selected according to the following criteria: (1) JMA magnitude $M > 4.0$; (2) hypocentral distance between 30 and 200 km; and (3) peak ground acceleration exceeding 3 Gal on each of the UD2, NS1, and EW1 components. Records sampled at 100 Hz were retained directly, whereas those sampled at 200 Hz were downsampled to 100 Hz using zero-

phase decimation with an eighth-order Chebyshev type I IIR anti-aliasing filter. The resulting dataset comprised 13,245 station-event records from 1,916 earthquakes recorded at 550 KiK-net stations. As shown in **Fig. 1(a)**, the stations provide broad spatial coverage across the Japanese archipelago. The selected earthquakes include both inland and offshore events, with their epicentral locations and magnitudes shown in **Fig. 1(b)**.

The earthquakes were chronologically divided at the event level into training, validation, and testing sets at a ratio of 70:15:15. The training set comprised 1,341 earthquakes occurring between 11 October 1997 and 16 March 2011. The validation set contained 287 earthquakes occurring between 16 March 2011 and 6 February 2015, and the testing set comprised 288 earthquakes occurring between 26 February 2015 and 27 March 2019. All records associated with the same earthquake were assigned to the same subset to prevent information leakage. The validation set was used for model selection, whereas the testing set was used only for the final performance evaluation.

2.2 *P-Wave Arrival Identification*

The P-wave arrival was automatically identified from the borehole vertical acceleration component (UD2) using a combined STA/LTA-AIC procedure. Each waveform was first demeaned and filtered using a fourth-order zero-phase Butterworth band-pass filter between 1 and 15 Hz to suppress low-frequency drift and high-frequency noise. The short-term average to long-term average characteristic ratio was then calculated as:

$$r(i) = \frac{\text{STA}(i)}{\text{LTA}(i)} \quad (1)$$

where i denotes the sampling point. The STA and LTA windows were set to 0.2 and 4.0 s, corresponding to 20 and 400 sampling points at a sampling frequency of 100 Hz, respectively.

153 A preliminary trigger was identified when $r(i)$ exceeded 2.5 and remained above 2.0 for at least
154 0.1 s.

155 To exclude records with weak initial arrivals, the signal-to-noise ratio was calculated as:

$$156 \quad \text{SNR} = 10 \log_{10} \left[\frac{\frac{1}{N_s} \sum_{i=i_t}^{i_t+N_s-1} a_0^2(i)}{\frac{1}{N_n} \sum_{i=1}^{N_n} a_0^2(i)} \right] \quad (2)$$

157 where $a_0(i)$ is the demeaned UD2 acceleration, i_t is the preliminary STA/LTA trigger point, and
158 N_s and N_n denote the numbers of sampling points in the signal and noise windows, respectively.
159 The first 1.0 s of each record was used as the noise window, while the 0.3-s waveform following
160 the preliminary trigger was used as the signal window. Only records with an SNR greater than
161 8 dB were retained.

162 The preliminary trigger was subsequently refined using the Akaike Information Criterion
163 (AIC). A 3.0-s search interval centered on the preliminary trigger was extracted from the
164 demeaned UD2 record, and the AIC value was evaluated at each candidate point according to
165 the variance change between the waveform segments before and after that point. The candidate
166 corresponding to the minimum AIC value was selected as the final P-wave arrival.

167 2.3 Station Site Classification

168 To evaluate the robustness of the proposed model under different site conditions, the 550
169 KiK-net stations were classified using three site-characterization parameters: the time-averaged
170 shear-wave velocity in the upper 30 m (V_{s30}), the predominant frequency derived from the
171 horizontal-to-vertical spectral ratio ($f_{0,HV}$), and the site amplification factor (AF). These
172 parameters describe the near-surface velocity structure, predominant site frequency, and overall

amplification characteristics, respectively. Stations with missing values for any of the three parameters were excluded during the initial data selection.

Before clustering, the three parameters were standardized to eliminate differences in their numerical scales. A Gaussian mixture model (GMM) with three components was then applied to the standardized station-level feature vectors. The probability distribution of a feature vector $\mathbf{z}$ was expressed as:

$$p(\mathbf{z}) = \sum_{k=1}^3 \pi_k N(\mathbf{z} | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \quad (3)$$

where π_k , $\boldsymbol{\mu}_k$, and $\boldsymbol{\Sigma}_k$ denote the mixture weight, mean vector, and covariance matrix of the k -th component, respectively. Each station was assigned to the component with the highest posterior probability.

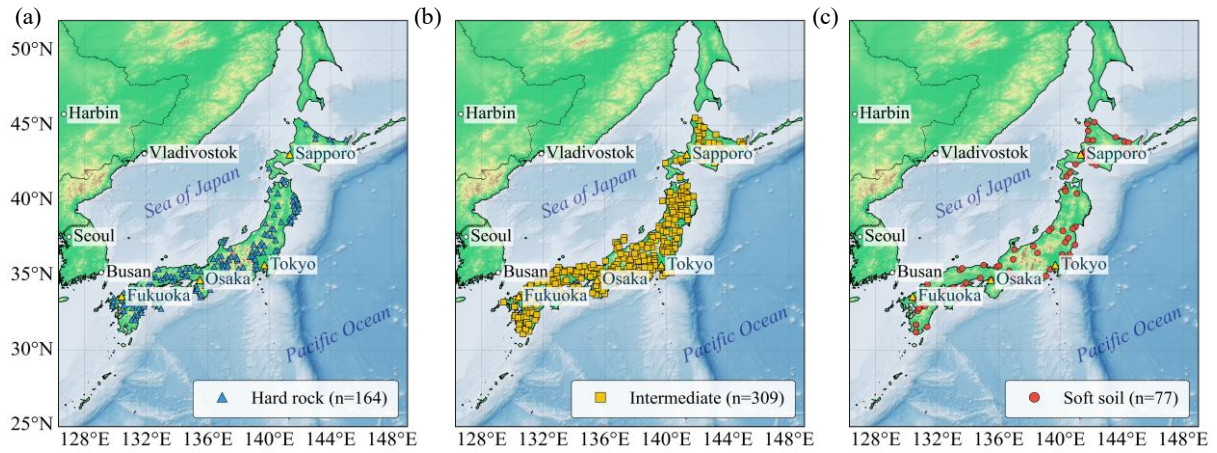


Fig. 3. Spatial distribution of the three site classes. (a) hard rock ($n = 164$), (b) intermediate ($n = 309$), and (c) soft soil ($n = 77$).

The classification identified 164 hard-rock stations, 309 intermediate stations, and 77 soft-soil stations, corresponding to 3,960, 7,453, and 1,832 ground-motion records, respectively. As shown in **Fig. 3**, hard-rock stations are mainly located in mountainous inland regions,

intermediate stations are widely distributed across the Japanese archipelago, and soft-soil stations are concentrated primarily in coastal lowlands and alluvial areas.

Table 1. Characteristics of the three site classes.

Parameter	Hard rock	Intermediate	Soft soil
Number of stations / records	164 / 3,960	309 / 7,453	77 / 1,832
M	4.96 ± 0.77	4.99 ± 0.77	4.96 ± 0.75
R (km)	109.8 ± 43.1	112.9 ± 44.4	115.4 ± 45.3
Vs30 (m/s)	708 ± 196	408 ± 90	245 ± 71
$f_{0,HV}$ (Hz)	9.84 ± 3.94	4.45 ± 2.70	1.38 ± 1.13
AF	0.887 ± 0.254	0.917 ± 0.206	1.559 ± 0.404

Table 1 summarizes the characteristics of the three site classes. The mean Vs30 decreases from 708 m/s for hard rock to 408 m/s for intermediate sites and 245 m/s for soft soil. A similar decreasing trend is observed for $f_{0,HV}$, whereas AF is substantially higher for soft-soil stations. Meanwhile, the magnitude and hypocentral-distance distributions are broadly comparable among the three classes.

3. VLPNet: A Unified Model for Variable-Length P-Wave Records

3.1 Variable-Length Input Construction and Prediction Target

For each station-event record, the borehole vertical acceleration component (UD2) was used as the input to VLPNet. Starting from the P-wave arrival identified in **Section 2.2**, the waveform was truncated at an arbitrary observation time between 0.5 and 6.0 s. For the i -th record, the input was expressed as:

$$\mathbf{x}_i(T_i) = \left[a_i^{\text{UD2}} \left(t_{P,i} + \frac{n}{f_s} \right) \right]_{n=0}^{\lfloor f_s T_i \rfloor - 1}, T_i \in [0.5, 6.0] \text{s}, f_s = 100 \text{Hz} \quad (4)$$

where a_i^{UD2} is the UD2 acceleration, $t_{P,i}$ is the identified P-wave arrival time, T_i is the effective observation duration after the P-wave arrival, and $\lfloor f_s T_i \rfloor$ is the corresponding number of valid samples. At a sampling frequency of 100 Hz, VLPNet can generate an estimate at a temporal

resolution of 0.01 s within the specified range, rather than being restricted to several predefined observation windows. Record-wise amplitude normalization was not applied because the absolute P-wave amplitude contains information relevant to earthquake magnitude.

Earthquake magnitude is conventionally related to the logarithm of ground-motion amplitude and source-to-station distance. The JMA magnitude procedure uses the logarithm of the combined peak horizontal displacement together with distance corrections (Katsumata, 2004). Similarly, Wang et al. (2023) predicted the logarithm of the complete-record peak displacement from the initial vertical P-wave and subsequently combined it with hypocentral distance to estimate magnitude. Accordingly, VLPNet predicts $\log_{10}A$ as a physically interpretable intermediate variable rather than estimating magnitude directly.

The prediction target was derived from the complete NS1 and EW1 surface records. The acceleration components were demeaned, linearly detrended, and filtered using a fourth-order zero-phase Butterworth band-pass filter between 0.1 and 25 Hz. They were then integrated twice using cumulative trapezoidal integration, with the mean removed after each integration:

$$d_i^h(t) = \mathcal{I}_c \left\{ \mathcal{I}_c \left[\mathcal{B}_{0.1-25} \left(\mathcal{D} \left[a_i^h(t) \right] \right) \right] \right\}, h \in \{\text{NS1}, \text{EW1}\} \quad (5)$$

where $\mathcal{D}$, $\mathcal{B}_{0.1-25}$, and $\mathcal{I}_c$ denote detrending, band-pass filtering, and mean-corrected integration, respectively. The two horizontal displacement components were combined vectorially at each time step, after which the maximum value over the complete record was determined. The regression target was defined as:

$$A_i = \max_t \sqrt{\left[d_i^{\text{NS1}}(t) \right]^2 + \left[d_i^{\text{EW1}}(t) \right]^2}, y_i = \log_{10} A_i \quad (6)$$

where A_i is the maximum vector-combined horizontal displacement of the complete surface record, expressed in centimeters, and y_i is the corresponding learning target. The complete NS1

and EW1 records were used only to construct the target and were not provided as model inputs. Thus, VLPNet predicts the eventual $\log_{10}A$ using only the P-wave information available at a given warning time.

To convert the predicted $\log_{10}A$ into magnitude, an empirical amplitude-magnitude-distance relationship was established using the training dataset:

$$\log_{10} A_i = \alpha M_i + \beta \log_{10} R_i + \gamma \Rightarrow \hat{M}_i = \frac{\hat{y}_i - \beta \log_{10} R_i - \gamma}{\alpha} \quad (7)$$

where M_i is the observed JMA magnitude, R_i is the hypocentral distance in kilometers, and α , β , and γ are regression coefficients estimated by ordinary least squares using only the training dataset. The VLPNet output is denoted by $\hat{y}_i$, and $\hat{M}_i$ is the resulting magnitude estimate. This formulation preserves the physical relationship among waveform amplitude, propagation distance, and earthquake magnitude while preventing information from the complete surface records from entering the model input.

3.2 VLPNet Architecture

Fig. 4 illustrates the architecture of VLPNet, which consists of a convolutional feature-extraction module, a valid-length packing module, a two-layer LSTM, and a fully connected regressor. The model receives zero-padded UD2 P-wave sequences together with their valid lengths and predicts the logarithm of the combined peak horizontal displacement amplitude. All input durations share the same network parameters, enabling a single model to process P-wave records observed at different warning times.

The convolutional module contains two one-dimensional convolutional blocks. The first block transforms the single-channel acceleration sequence into 32 feature channels, whereas the second expands the feature dimension from 32 to 64. Both convolutional layers use a kernel

size of 5, a stride of 1, and a padding size of 2, followed by batch normalization and ReLU activation. Because the padding preserves the temporal dimension, the effective output lengths of the two blocks remain L_i , with feature sizes of $L_i \times 32$ and $L_i \times 64$, respectively. No temporal pooling is applied, thereby preserving the correspondence between the extracted features and the original sampling instants.

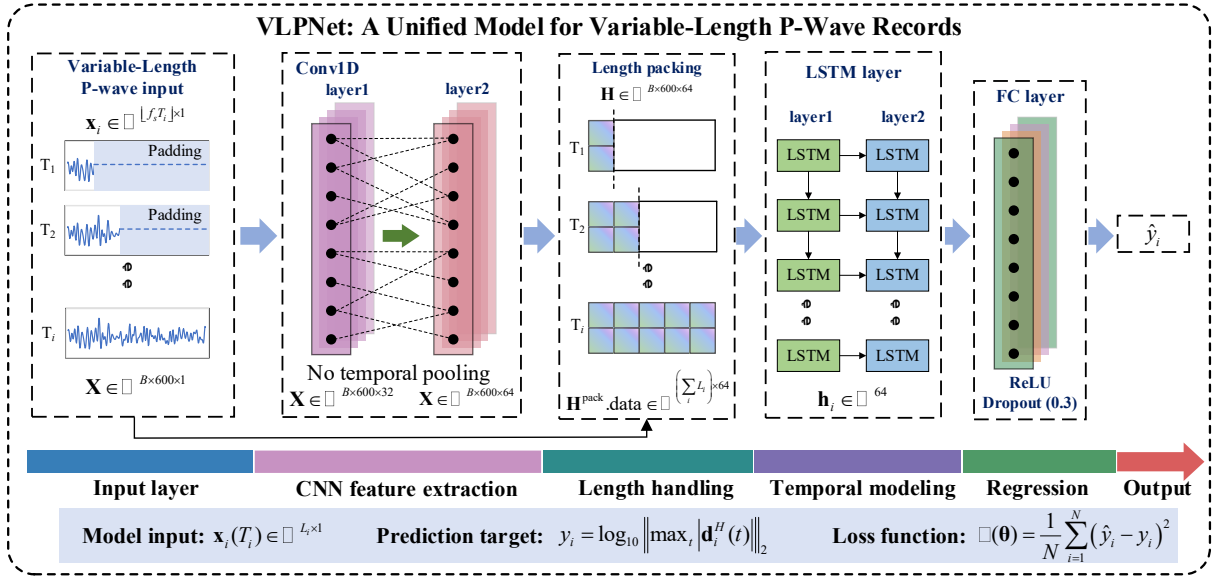


Fig. 4. Architecture of VLPNet for variable-length P-wave processing.

For mini-batch processing, the input sequences are zero-padded to 600 samples before convolution, while their valid lengths are retained. The resulting CNN feature tensor has a padded size of $B \times 600 \times 64$. Before being passed to the LSTM, the feature sequences are packed according to their valid lengths using “pack_padded_sequence”. Consequently, only the first L_i feature vectors of each sequence participate in the recurrent calculation, while the padded time steps are excluded. This operation allows samples with different observation durations to be processed within the same mini-batch.

The packed feature sequences are subsequently encoded using a two-layer LSTM with 64 hidden units in each layer, which enables the model to capture temporal dependencies at

multiple levels of abstraction. A dropout rate of 0.3 is applied between the two LSTM layers. For each record, the final hidden state of the second layer corresponds to its actual observation endpoint rather than the end of the padded sequence. The resulting 64-dimensional hidden state is passed to a fully connected regressor consisting of a 32-neuron hidden layer with ReLU activation and a dropout rate of 0.3, followed by a single-neuron output layer.

Together, these components enable VLPNet to combine local waveform feature extraction with variable-length temporal modeling. The convolutional blocks identify local patterns in the initial P-wave without reducing the temporal resolution, while the LSTM integrates the evolving CNN features over the actual observation duration. Valid-length packing prevents padded time steps from affecting the recurrent computation and ensures that the final hidden state corresponds to the true endpoint of each sequence. Most importantly, the same network parameters are shared across all input durations, eliminating the need to train a separate model for each warning time and allowing the magnitude estimate to be continuously updated as additional P-wave information becomes available.

3.3 Model Training Strategy

The data used for model development were obtained from the KiK-net dataset described in **Section 2**, comprising 13,245 station-event records from 1,916 earthquakes. The earthquakes were chronologically divided at the event level into training, validation, and testing sets containing 1,341, 287, and 288 events, respectively. During training, each record was randomly truncated to a valid length ranging from 50 samples to its maximum available length, capped at 600 samples. Samples beyond the selected endpoint were set to zero, while the valid length was retained for packed LSTM processing. Thus, the training inputs covered observation durations

from 0.5 to 6.0 s at a temporal resolution of 0.01 s. All truncated sequences derived from the same record shared the complete-record target defined in **Section 3.1**.

VLPNet was trained for 200 epochs by minimizing the mean squared error (MSE) between the predicted and target logarithmic displacement amplitudes, as illustrated in **Fig. 4**. The Adam optimizer was adopted with a mini-batch size of 32, an initial learning rate of 1.0×10^{-3} , a weight decay of 1.0×10^{-5} , and coefficients of $\beta_1=0.9$ and $\beta_2=0.99$. The learning rate was exponentially decayed by a factor of 0.99 after each epoch. For model selection, the validation set was evaluated at seven observation durations, $T_k \in \{0.5, 1, 2, 3, 4, 5, 6\}$ s. The overall validation loss was defined as the equally weighted mean of the duration-specific MSEs:

$$\bar{L}_{\text{val}} = \frac{1}{K} \sum_{k=1}^K \left\{ \frac{1}{N_k} \sum_{i \in I_k} \left[f_{\theta}(\mathbf{x}_i^{(T_k)}, T_k) - y_i \right]^2 \right\} \quad (8)$$

where I_k and N_k denote the eligible validation records and their number at duration T_k , respectively. The model weights yielding the lowest $\bar{L}_{\text{val}}$ were retained for evaluation on the independent testing set. VLPNet was implemented in PyTorch 2.3.1 with CUDA 11.8 and trained on a single NVIDIA GeForce RTX 4060 Ti GPU.

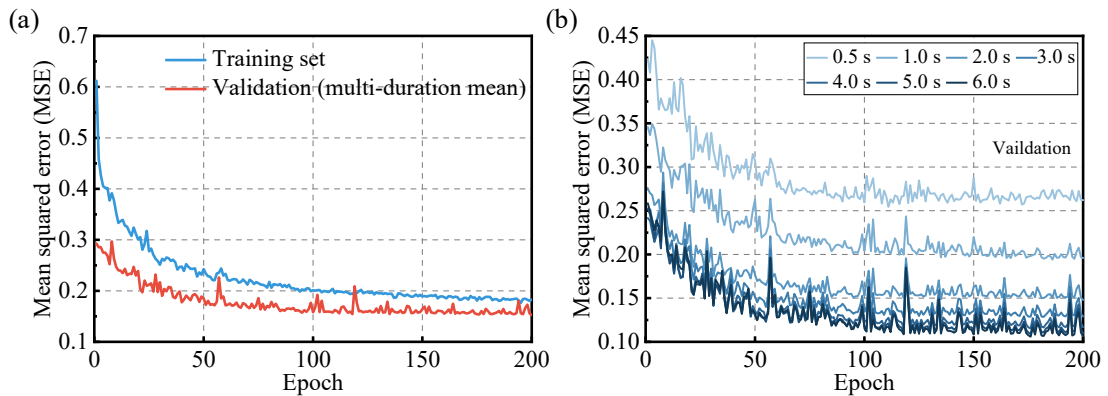


Fig. 5. Training and validation loss curves of VLPNet.

Fig. 5 presents the loss histories of VLPNet over 200 training epochs. As shown in **Fig. 5(a)**, both the training loss and the duration-averaged validation loss decreased rapidly during

the initial epochs and subsequently converged with only minor fluctuations, indicating stable model optimization without an evident increase in validation error. The training loss remained slightly higher than the validation loss, which likely reflects the random-length truncation and dropout applied during training, as well as the different aggregation of observation durations between training and validation. **Fig. 5(b)** shows that the duration-specific validation loss generally decreased as the P-wave observation duration increased, demonstrating that the additional waveform information improved the prediction of $\log_{10}A$. However, the rate of improvement gradually diminished, and the losses tended to stabilize after approximately 3–4 s, suggesting decreasing marginal benefits from further extension of the observation window.

4. Results and Discussion

4.1 Time-Dependent Prediction Performance

To evaluate the overall performance of VLPNet, the test records with complete 6-s P-wave segments were used at all observation durations, thereby ensuring that the observed performance variation was not affected by changes in sample composition. Five independently trained models with different random seeds were evaluated. The coefficient of determination (R^2), Pearson correlation coefficient ($corr$), mean absolute error (MAE), and root mean square error (RMSE) were adopted to assess the prediction of $\log_{10}A$. The estimated $\log_{10}A$ values were subsequently converted into magnitude estimates using the magnitude–distance–amplitude relationship established from the training dataset.

Fig. 6 presents the variation in prediction of $\log_{10}A$ performance with P-wave observation duration. As the observation duration increased from 0.5 to 6.0 s, both R^2 and $corr$ increased continuously, whereas MAE and RMSE decreased. This indicates that the predictive capability

of VLPNet improved progressively as additional P-wave information became available. The relatively narrow standard-deviation bands demonstrate that the performance trends were consistent across the five independent runs.

The most rapid improvement occurred during the initial observation period. As the window increased from 0.5 to approximately 3.0 s, R^2 and $corr$ increased markedly, accompanied by substantial reductions in MAE and RMSE. Beyond approximately 3.0 s, the improvements became more gradual. Nevertheless, the performance continued to increase up to 6.0 s, indicating that the later portion of the P-wave signal still provided useful information for estimating the subsequent peak displacement. This pattern reflects a trade-off between prediction accuracy and warning timeliness: short windows provide earlier estimates, whereas longer windows improve accuracy but reduce the available warning time.

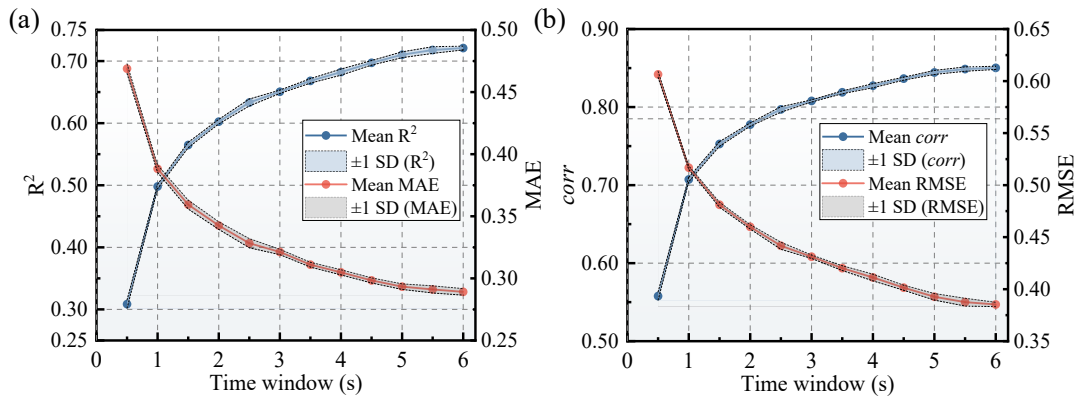


Fig. 6. Time-dependent prediction performance of VLPNet for $\log_{10}A$: (a) coefficient of determination (R^2) and MAE; (b) Pearson correlation coefficient ($corr$) and RMSE.

Fig. 7(a) further compares the predicted and actual $\log_{10}A$ values at representative observation durations of 0.5, 3.0, and 6.0 s. At 0.5 s, the predicted values exhibited relatively large dispersion around the 1:1 line, with $corr=0.558$, $MAE=0.466$, and $RMSE=0.606$. The predicted range was also narrower than the actual range, indicating that the very short P-wave segment provided insufficient information to distinguish records with substantially different

peak displacement amplitudes. When the observation duration increased to 3.0 s, the data points became more concentrated around the 1:1 line. The correlation coefficient increased to 0.808, while MAE and RMSE decreased to 0.320 and 0.432, respectively. At 6.0 s, the agreement between the predicted and actual values improved further, with $corr=0.851$, MAE =0.290, and RMSE =0.386. These results confirm that the additional temporal information contained in longer P-wave segments enabled the model to characterize the subsequent ground-motion intensity more accurately.

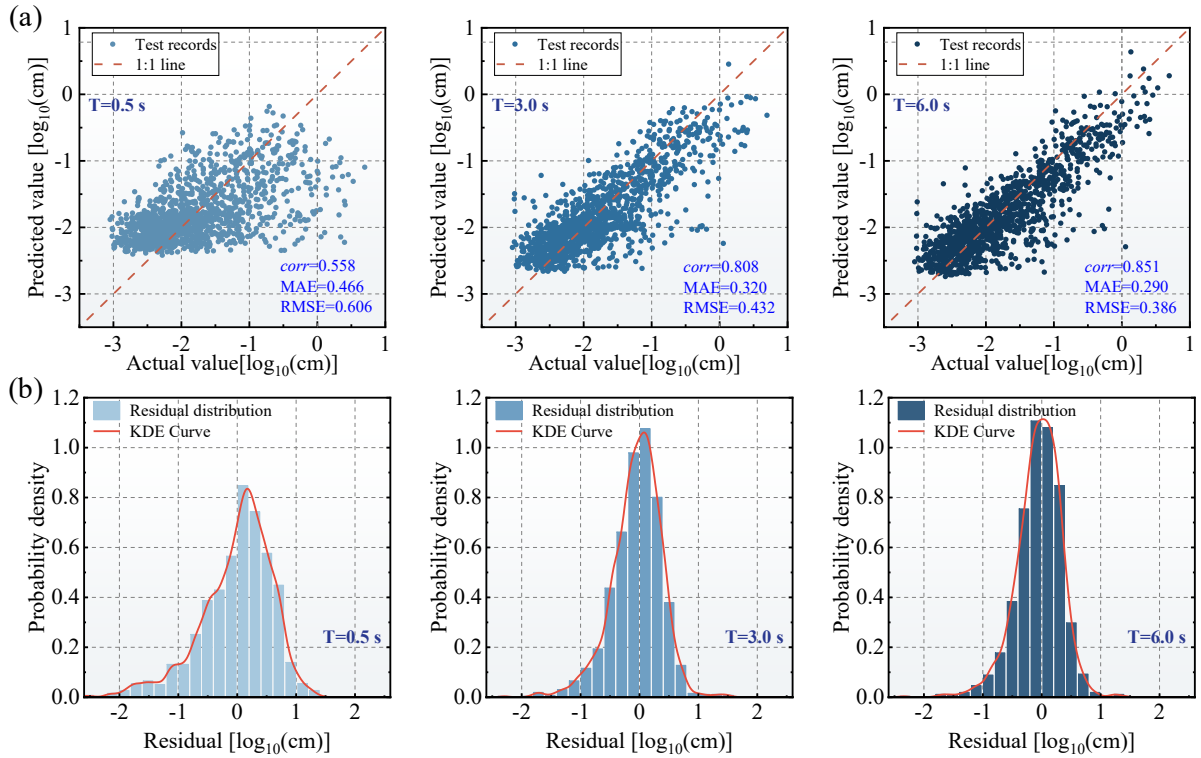


Fig. 7. Actual-predicted comparisons and residual distributions of $\log_{10}A$ for the representative run at P-wave observation durations of 0.5, 3.0, and 6.0 s.

The residual distributions shown in **Fig. 7(b)** provide further evidence of this improvement. At 0.5 s, the residual distribution was relatively broad and asymmetric, indicating large prediction uncertainty during the earliest observation stage. With increasing observation duration, the distribution became progressively narrower and more concentrated around zero.

In particular, the 3.0- and 6.0-s residual distributions were considerably sharper than that for 0.5 s. The reduction in both the distribution width and tail probability indicates that longer observation windows reduced not only the average prediction error but also the occurrence of relatively large errors.

Because the final objective is earthquake-magnitude estimation, the predicted $\log_{10}A$ values were converted into magnitude estimates, and the resulting magnitude residuals were analyzed. **Fig. 8** presents the relationships between magnitude residual and target magnitude, as well as between magnitude residual and epicentral distance, at observation durations of 0.5, 3.0, and 6.0 s. At 0.5 s, the magnitude estimates yielded an MAE of 0.458 and an RMSE of 0.632. Approximately 66.0% and 88.4% of the estimates were within ± 0.5 and ± 1.0 magnitude units, respectively. Increasing the observation duration to 3.0 s reduced the MAE and RMSE to 0.345 and 0.481, while the corresponding proportions increased to 77.9% and 94.9%. At 6.0 s, the MAE decreased further to 0.328 and the RMSE to 0.447, with 78.3% and 95.7% of the estimates falling within ± 0.5 and ± 1.0 magnitude units, respectively. Thus, a substantial improvement was achieved within the first 3 s, whereas the gain from 3.0 to 6.0 s was comparatively modest, particularly for the proportion of estimates within ± 0.5 magnitude units.

The residual-magnitude relationships in **Fig. 8(a)** reveal a magnitude-dependent error pattern. For relatively small events, the residuals were generally distributed around zero, whereas increasingly negative residuals were observed for larger events. This tendency was particularly pronounced at the 0.5-s observation duration and was reduced as the observation window increased. The systematic underestimation of large events can be attributed partly to the limited information contained in the initial P-wave segment and the lower representation of

large-magnitude earthquakes in the training dataset. In addition, the use of error-based regression objectives tends to favor the more densely populated moderate-magnitude range, thereby producing predictions that regress toward the center of the training distribution.

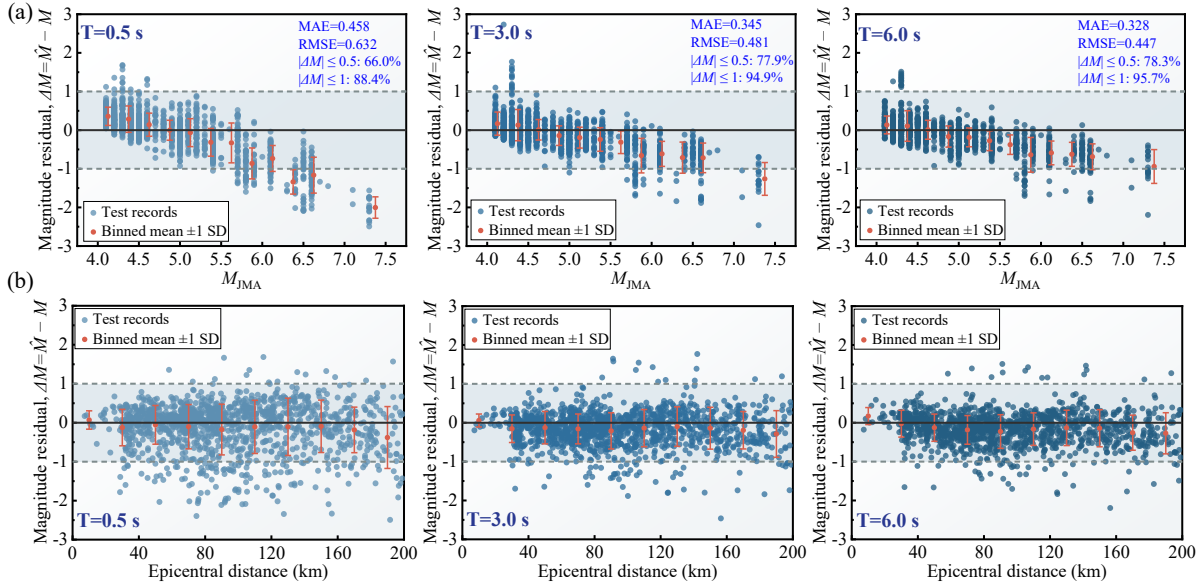


Fig. 8. Magnitude residuals with respect to JMA magnitude and epicentral distance.

As shown in **Fig. 8(b)**, the magnitude residuals displayed a weaker dependence on epicentral distance than on target magnitude. The binned mean residuals remained relatively stable across most of the distance range, although a modest negative shift was observed at larger distances. This behavior may be related to distance-dependent attenuation, variations in signal-to-noise ratio, and residual uncertainty in the magnitude-distance-amplitude relationship. Importantly, extending the observation duration narrowed the residual distributions over nearly the entire distance range, indicating that the benefit of incorporating additional P-wave information was not restricted to a particular distance interval.

4.2 Comparison with Baseline Models

To assess the effectiveness of VLPNet, its magnitude-estimation performance was compared with three conventional P-wave parameter-based methods: peak displacement P_d ,

maximum predominant period τ_p^{max} , and their combination $P_d + \tau_p^{max}$. For each observation duration, the empirical relationships of the baseline methods were calibrated using only the training dataset and subsequently applied to the common test dataset. The same data partition and test records were used for all methods to ensure a consistent comparison.

Fig. 9 compares the performance of the four methods over observation durations ranging from 0.5 to 6.0 s. Overall, VLPNet achieved the lowest MAE and the highest proportion of estimates within ± 1 magnitude unit throughout the investigated duration range. At 0.5 s, VLPNet produced an MAE of 0.458, while 88.4% of its estimates fell within ± 1 magnitude unit. Despite the limited information available in such a short P-wave segment, its performance was already superior to that of the three conventional parameter-based methods. As the observation duration increased, the performance of all methods generally improved, although the degree of improvement varied considerably. The P_d -based method exhibited a marked reduction in MAE, indicating that peak displacement became increasingly informative as the P-wave window expanded. However, its error remained higher than that of VLPNet at all observation durations.

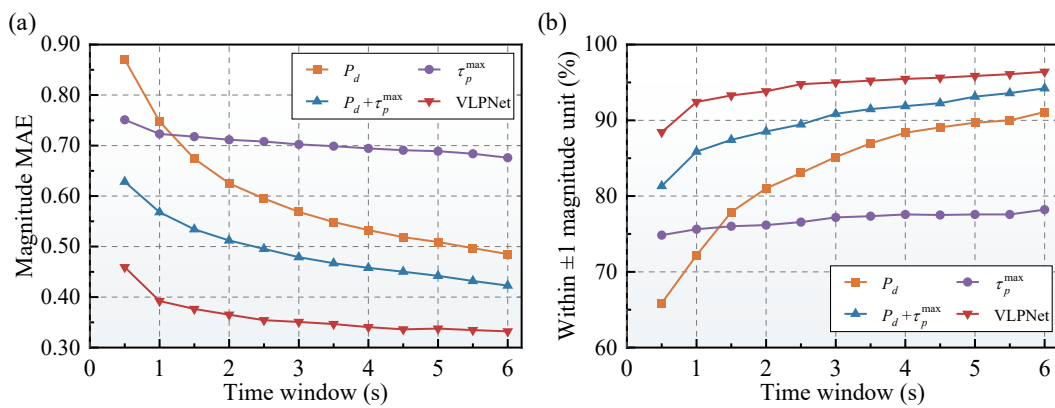


Fig. 9. Performance comparison of VLPNet and baseline methods: (a) MAE and (b) estimates within ± 1 magnitude unit.

Combining P_d and τ_p^{max} consistently improved the results compared with either parameter alone, particularly with respect to the percentage of estimates within ± 1 magnitude

unit. This improvement indicates that amplitude- and frequency-related parameters contain complementary information. Nevertheless, the combined empirical method remained less accurate than VLPNet. Unlike the baseline methods, which reduce each P-wave segment to one or two predefined scalar parameters, VLPNet learns amplitude, frequency, and temporal-evolution characteristics directly from the waveform. Its consistently better performance therefore suggests that the original waveform contains useful predictive information that is not fully retained by the conventional parameters.

Fig. 10 presents the magnitude residuals of the three baseline methods at an observation duration of 3.0 s. The P_d -based method yielded an MAE of 0.569 and an RMSE of 0.751, with 52.6% and 85.2% of the estimates falling within ± 0.5 and ± 1 magnitude units, respectively. Its residuals showed a relatively wide distribution and tended to become more negative with increasing magnitude. The τ_p^{max} -based method exhibited the largest errors, with an MAE of 0.702 and an RMSE of 0.983, while the corresponding proportions were only 47.6% and 77.2%. Its broad residual distribution suggests that τ_p^{max} alone provided relatively limited magnitude-discrimination capability in the present dataset. Combining P_d and τ_p^{max} improved the MAE and RMSE to 0.479 and 0.658 and increased the proportions within ± 0.5 and ± 1 magnitude units to 62.8% and 90.9%, respectively. This improvement indicates that the two parameters provide complementary amplitude- and frequency-related information. Nevertheless, some magnitude-dependent variation remained in the binned mean residuals, and the overall dispersion was still greater than that obtained using VLPNet.

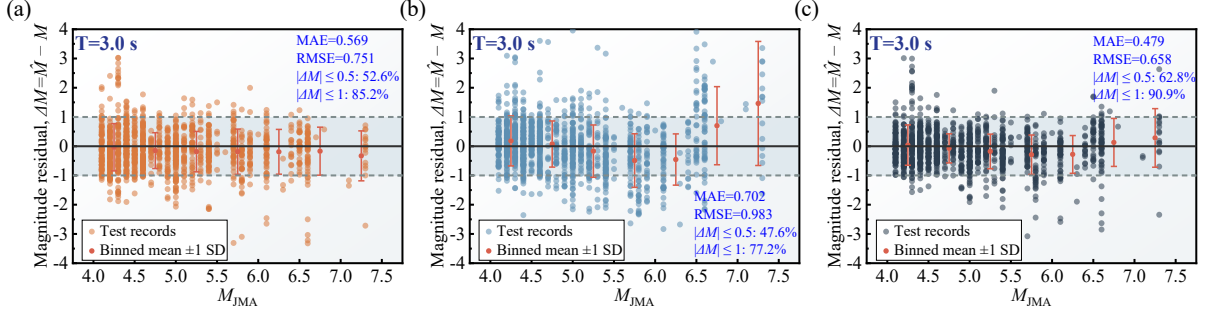


Fig. 10. Magnitude residuals of the baseline methods at a 3-s P-wave observation duration. (a) P_d , (b) τ_p^{max} , (c) $P_d + \tau_p^{max}$.

Fig. 11 provides a direct quantitative comparison of the four methods at the 3.0-s observation duration. VLPNet achieved an MAE of 0.345 and an RMSE of 0.481, substantially lower than those of the baseline methods. Relative to the P_d -based method, VLPNet reduced MAE and RMSE by 39.4% and 36.0%, respectively. Compared with τ_p^{max} , the corresponding reductions were 50.9% and 51.1%. Even compared with the combined $P_d + \tau_p^{max}$ method, VLPNet reduced MAE by 28.0% and RMSE by 26.9%.

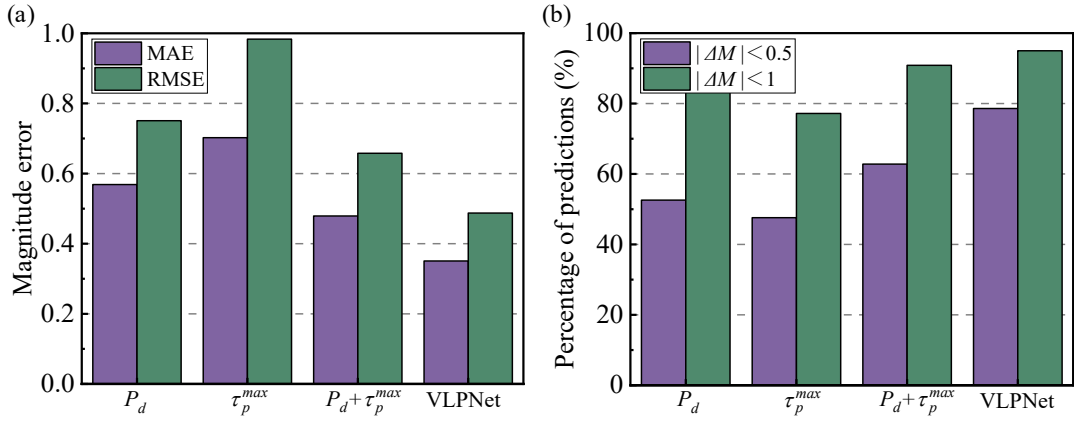


Fig. 11. Performance comparison at a 3-s P-wave duration: (a) errors and (b) estimates within specified error ranges.

The advantage of VLPNet was also evident under the two magnitude-error thresholds. At 3.0 s, 77.9% of the VLPNet estimates were within ± 0.5 magnitude units, exceeding the results of P_d , τ_p^{max} , and $P_d + \tau_p^{max}$ by 25.3, 30.3, and 15.1 percentage points, respectively. Similarly, 94.9% of the VLPNet estimates were within ± 1 magnitude unit, representing improvements of

9.7, 17.7, and 4.0 percentage points over the three baseline methods, respectively. Although the improvement over the combined method was smaller under the ± 1 criterion, VLPNet exhibited a clearer advantage under the stricter ± 0.5 criterion and in terms of both MAE and RMSE.

Overall, the comparison demonstrates that VLPNet provides more accurate and stable magnitude estimates than the conventional P_d and τ_p^{max} based methods. The combined empirical method benefits from incorporating both amplitude- and frequency-related information but still relies on a limited representation of the P-wave signal. By directly learning multiscale local features and their temporal dependencies from variable-length waveforms, VLPNet preserves more of the information carried by the evolving P-wave segment. This advantage is particularly relevant during the earliest observation stage, when robust estimation from limited waveform information is essential for earthquake early warning.

4.3 Performance under Different Site Conditions

To evaluate the generalization capability of VLPNet under different site conditions, the test dataset was divided into hard-rock, intermediate, and soft-soil groups according to the site classification described in **Section 2.3**. The same trained model was applied to all three site classes without site-specific retraining or parameter adjustment. To avoid interference caused by changes in sample composition, the same test records containing complete 6-s P-wave segments were used for all observation durations. Model performance was evaluated using the mean absolute error (MAE), the proportion of estimates within ± 1 magnitude unit, and the distribution of magnitude residuals.

Fig. 12 presents the variation in model performance with P-wave observation duration for the three site classes. Overall, as the observation duration increased from 0.5 to 6.0 s, the MAE

decreased and the proportion of estimates within ± 1 magnitude unit increased for all site classes. These results demonstrate that the magnitude-estimation error was progressively reduced as additional P-wave information became available. The performance improvement was more pronounced during the early observation period, whereas the changes in both metrics gradually diminished at longer observation durations. This trend indicates a decreasing marginal contribution of additional P-wave information to magnitude estimation. The performance curves for the three site classes exhibited generally consistent trends, indicating that the model extracted stable temporal features from records obtained under different site conditions. The improvement with increasing observation duration was therefore not limited to a particular site class. In addition, the error bars obtained from different random seeds remained relatively narrow and did not expand noticeably with observation duration, demonstrating satisfactory repeatability and training stability. Although some differences were observed among the three site classes, these differences primarily concerned the overall error level rather than the temporal pattern of performance improvement.

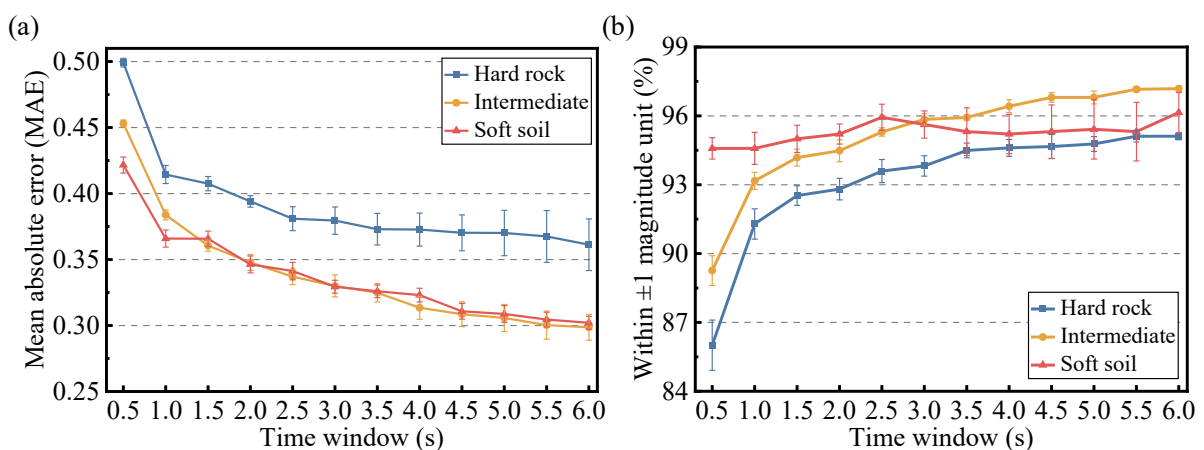


Fig. 12. Magnitude-estimation performance of the VLPNet model under different site conditions. (a) MAE, (b) proportion of predictions within ± 1 magnitude unit.

Considering the balance between warning timeliness and prediction accuracy, a 3-s P-wave

observation duration was selected for a more detailed comparison of the residual distributions among the three site classes. **Fig. 13** presents the box-and-scatter plots of the magnitude residuals, while the corresponding quantitative results are summarized in **Table 2**. At the 3-s observation duration, the hard-rock, intermediate, and soft-soil subsets contained 356, 732, and 192 test records, respectively. The residual standard deviations ranged from 0.443 to 0.453, with a maximum difference of only 0.010. This close agreement indicates that the dispersion of the prediction errors was largely consistent across the three site classes. The intermediate sites achieved the lowest MAE of 0.333 and the highest proportion of estimates within ± 0.5 magnitude units, at 81.28%. The soft-soil sites produced a comparable MAE of 0.334 and the lowest RMSE of 0.456. In comparison, the hard-rock sites exhibited relatively larger errors, with an MAE of 0.396 and an RMSE of 0.523.

Table 2. Magnitude-estimation performance of the VLPNet model for different site classes at a 3-s observation duration.

Site class	Residual SD	MAE	RMSE	Within ± 0.5 (%)	Within ± 1.0 (%)
Hard rock	0.443	0.396	0.523	73.03	93.26
Intermediate	0.453	0.333	0.477	81.28	95.77
Soft soil	0.448	0.334	0.456	78.65	95.31

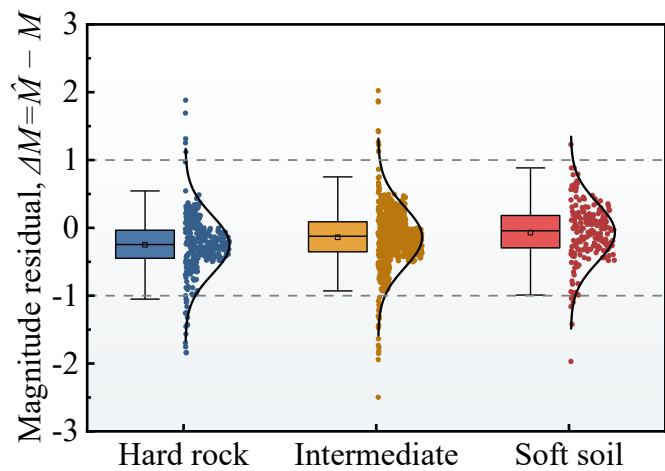


Fig. 13. Magnitude-residual distributions of the VLPNet model under different site conditions at a 3-s observation duration.

In terms of the practically acceptable error range, 93.26%, 95.77%, and 95.31% of the estimates for the hard-rock, intermediate, and soft-soil sites, respectively, were within ± 1 magnitude unit. All three proportions exceeded 93%, indicating that the model maintained satisfactory magnitude-estimation accuracy under different site conditions, even when only a relatively short 3-s P-wave segment was available. The proportions within ± 0.5 magnitude units ranged from 73.03% to 81.28%, with greater differences among the site classes than those observed for the ± 1 magnitude threshold. This result suggests that site conditions exert a more noticeable influence when a stricter error criterion is applied.

The residual distributions further show that most estimates were concentrated around zero, although a small number of records in each site class fell outside the ± 1 magnitude-unit range. The residual distribution for the hard-rock sites was more strongly shifted toward negative values, indicating a more pronounced tendency toward magnitude underestimation. By comparison, the residual centers for the intermediate and soft-soil sites were closer to zero. Meanwhile, the widths of the three residual distributions were similar, consistent with their nearly identical residual standard deviations. Therefore, the influence of site conditions was primarily reflected in changes in the central tendency of the residuals rather than in a substantial increase in prediction dispersion.

The observed differences may be associated with the effects of local site conditions on ground-motion amplitudes and frequency characteristics. VLPNet first predicts the logarithmic peak horizontal displacement from the borehole vertical P-wave record, after which the magnitude is estimated using a unified magnitude–distance–amplitude relationship. Hard-rock sites generally exhibit weaker site amplification and may therefore produce smaller ground-

motion amplitudes than softer sites for events with similar magnitudes and hypocentral distances. Consequently, applying a unified magnitude-inversion relationship without an explicit site term may result in a greater degree of magnitude underestimation for hard-rock sites. Nevertheless, this behavior may also be affected by differences in P-wave characteristics, the empirical inversion relationship, and the unequal sample sizes of the three site classes. The observed performance differences therefore cannot be attributed to a single factor based on the current analysis.

Overall, VLPNet exhibited consistent performance improvements with increasing P-wave observation duration across all three site classes. The residual standard deviations remained close to 0.45, and more than 93% of the 3-s estimates were within ± 1 magnitude unit for each site class, demonstrating satisfactory robustness to varying site conditions. However, the hard-rock group showed relatively higher MAE and RMSE values and a more pronounced underestimation tendency. Future improvements could incorporate a site correction term into the magnitude-inversion relationship or introduce site class and related site parameters as auxiliary model inputs to reduce systematic performance differences among site conditions.

5. Conclusions and Limitations

This study proposed VLPNet, a unified deep learning model for single-station magnitude estimation from variable-length P-wave records. The model predicts the logarithm of the complete-record peak horizontal displacement and subsequently combines it with hypocentral distance to estimate magnitude. Its performance was evaluated at different observation durations, against conventional parameter-based methods, and under different site conditions. The main conclusions are as follows:

(1) VLPNet processes P-wave records of any duration between 0.5 and 6.0 s using a single set of network parameters, enabling continuous magnitude updates without separate models for predefined observation windows. Valid-length packing excludes zero-padded samples from temporal modeling.

(2) Prediction accuracy improved with increasing observation duration. The magnitude MAE decreased from 0.461 ± 0.002 at 0.5 s to 0.317 ± 0.010 at 6.0 s, while the proportion of estimates within ± 1 magnitude unit increased from $89.2\% \pm 0.7\%$ to $96.5\% \pm 0.2\%$.

(3) Compared with conventional parameter-based models that rely on a limited number of manually defined P-wave features, VLPNet directly extracts magnitude-related information from the evolving waveform. It generally outperformed the P_d , τ_p^{max} , and $P_d + \tau_p^{max}$ based methods and maintained broadly consistent performance across the three site classes.

This study has several limitations that warrant further investigation. Although the present results demonstrate the effectiveness of VLPNet on the KiK-net dataset, its transferability across different seismic regions and magnitude ranges requires further evaluation. Event-aware and multi-station learning strategies could also be explored to make fuller use of records obtained at multiple stations during the same earthquake. In addition, the robustness of the framework to real-time uncertainties in P-wave picking and hypocentral-distance estimates should be systematically assessed.

574

Statements & Declarations

575 Author contributions

576 All authors contributed to the study conception and design. Material preparation, data
577 collection, and formal analysis were performed by Chunde Lu and Aihemaitijiang Tuerhong.
578 The first draft of the manuscript was written by Chunde Lu and Xiaopeng Gu, Yongtao Ma and
579 Xiaohong Long provided critical revisions and comments on previous versions of the
580 manuscript. All authors read and approved the final manuscript.

581 Funding

582 The authors would like to extend their thanks to the financial support of the National
583 Natural Science Foundation of China (Grant No. 51978306, 52408421) and the China
584 Postdoctoral Science Foundation - Hubei Joint Support Program (Grant No. 2025T024HB).

585 Competing Interests

586 The author has no relevant financial or non-financial interests to disclose.

587 Data availability

588 The KiK-net ground-motion records used in this study were obtained from the National
589 Research Institute for Earth Science and Disaster Resilience (NIED) and are available from the
590 NIED K-NET and KiK-net database at <https://www.kyoshin.bosai.go.jp/en/>.

591 References

- 592 Allen, R.M., Kanamori, H., 2003. The potential for earthquake early warning in southern California. *Science* 300
593 (5620), 786–789.
- 594 Allen, R.M., Gasparini, P., Kamigaichi, O., Bose, M., 2009. The status of earthquake early warning around the
595 world: An introductory overview. *Seismol. Res. Lett.* 80 (5), 682–693.
- 596 Allen, R.M., Melgar, D., 2019. Earthquake early warning: Advances, scientific challenges, and societal needs.
597 *Annu. Rev. Earth Planet. Sci.* 47 (1), 361–388.
- 598 Brown, H.M., Allen, R.M., Hellweg, M., Khainovski, O., Neuhauser, D., Souf, A., 2011. Development of the
599 ElarmS methodology for earthquake early warning: Realtime application in California and offline testing in
600 Japan. *Soil Dyn. Earthq. Eng.* 31 (2), 188–200.
- 601 Chakraborty, M., Fenner, D., Li, W., Faber, J., Zhou, K., Rümper, G., et al., 2022. CREIME—A convolutional
602 recurrent model for earthquake identification and magnitude estimation. *J. Geophys. Res. Solid Earth* 127 (7),
603 e2022JB024595.
- 604 Dai, H., Zhou, Y., Liu, H., Li, S., Wei, Y., Song, J., 2024. XGBoost-based prediction of on-site acceleration
605 response spectra with multi-feature inputs from P-wave arrivals. *Soil Dyn. Earthq. Eng.* 178, 108503.

- Hou, B., Zhou, Y., Li, S., Wei, Y., Song, J., 2024. Real-time earthquake magnitude estimation via a deep learning network based on waveform and text mixed modal. *Earth Planets Space* 76 (1), 1–18.
- Kanamori, H., 2005. Real-time seismology and earthquake damage mitigation. *Annu. Rev. Earth Planet. Sci.* 33 (1), 195–214.
- Katsumata, A., 2004. Revision of the JMA displacement magnitude. *Q. J. Seismol.* 67, 1–10. (In Japanese)
- Mousavi, S.M., Zhu, W., Sheng, Y., Beroza, G.C., 2019. CRED: A deep residual network of convolutional and recurrent units for earthquake signal detection. *Sci. Rep.* 9 (1), 10267.
- Mousavi, S.M., Beroza, G.C., 2020. A machine-learning approach for earthquake magnitude estimation. *Geophys. Res. Lett.* 47 (1), e2019GL085976.
- Mousavi, S.M., Ellsworth, W.L., Zhu, W., Chuang, L.Y., Beroza, G.C., 2020. Earthquake Transformer—An attentive deep-learning model for simultaneous earthquake detection and phase picking. *Nat. Commun.* 11 (1), 3952.
- Münchmeyer, J., Bindi, D., Leser, U., Tilmann, F., 2021. Earthquake magnitude and location estimation from real-time seismic waveforms with a transformer network. *Geophys. J. Int.* 226 (2), 1086–1104.
- Nakamura, Y., 1988. On the urgent earthquake detection and alarm system (UrEDAS). In: *Proceedings of the 9th World Conference on Earthquake Engineering*, pp. 673–678.
- Olivieri, M., Allen, R.M., Wurman, G., 2008. The potential for earthquake early warning in Italy using ElarmS. *Bull. Seismol. Soc. Am.* 98 (1), 495–503.
- Rafiei, M.H., Adeli, H., 2017. NEEWS: A novel earthquake early warning model using neural dynamic classification and neural dynamic optimization. *Soil Dyn. Earthq. Eng.* 100, 417–427.
- Saad, O.M., Chen, Y., Savvaidis, A., Fomel, S., Chen, Y., 2022. Real-time earthquake detection and magnitude estimation using vision transformer. *J. Geophys. Res. Solid Earth* 127 (5), e2021JB023657.
- Sadeh, M., Ziv, A., Wust-Bloch, H., 2014. Real-time magnitude proxies for earthquake early warning in Israel. *Geophys. J. Int.* 196 (2), 939–95
- Satriano, C., Wu, Y., Zollo, A., Kanamori, H., 2011. Earthquake early warning: Concepts, methods and physical grounds. *Soil Dyn. Earthq. Eng.* 31 (2), 106–118.
- Shieh, J., Wu, Y., Allen, R.M., 2008. A comparison of τ_c and τ_{pmax} for magnitude estimation in earthquake early warning. *Geophys. Res. Lett.* 35 (20).
- Spallarossa, D., Kotha, S.R., Picozzi, M., Barani, S., Bindi, D., 2019. On-site earthquake early warning: A partially non-ergodic perspective from the site effects point of view. *Geophys. J. Int.* 216 (2), 919–934.
- Wang, Y., Li, X., Wang, Z., Liu, J., 2023. Deep learning for magnitude prediction in earthquake early warning. *Gondwana Res.* 123, 164–173.
- Wolfe, C.J., 2006. On the properties of predominant-period estimators for earthquake early warning. *Bull. Seismol. Soc. Am.* 96 (5), 1961–1965.
- Wu, Y., Kanamori, H., 2005. Experiment on an onsite early warning method for the Taiwan early warning system. *Bull. Seismol. Soc. Am.* 95 (1), 347–353.
- Wu, Y., Zhao, L., 2006. Magnitude estimation using the first three seconds P-wave amplitude in earthquake early warning. *Geophys. Res. Lett.* 33 (16).
- Wu, Y.M., Kanamori, H., Allen, R.M., Hauksson, E., 2007. Determination of earthquake early warning parameters, τ_c and p_d , for Southern California. *Geophys. J. Int.* 170 (2), 711–717.
- Yamada, M., Mori, J., 2009. Using τ_c to estimate magnitude for earthquake early warning and effects of near-field terms. *J. Geophys. Res. Solid Earth* 114 (B5).
- Zhu, W., Beroza, G.C., 2019. PhaseNet: A deep-neural-network-based seismic arrival-time picking method. *Geophys. J. Int.* 216 (1), 261–273.

650 Zhu, J., Li, S., Ma, Q., Song, J., 2025. Onsite intensity prediction for earthquake early warning with multimodal
651 deep learning. *Soil Dyn. Earthq. Eng.* 195, 109430.
652 Zollo, A., Amoroso, O., Lancieri, M., Wu, Y., Kanamori, H., 2010. A threshold-based earthquake early warning
653 using dense accelerometer networks. *Geophys. J. Int.* 183 (2), 963–974.