



HAL
open science

Earthquake Early Warning Starting From 3 s of Records on a Single Station With Machine Learning

Pablo Lara, Quentin Bletery, Jean-paul Ampuero, Adolfo Inza, Hernando Tavera

► To cite this version:

Pablo Lara, Quentin Bletery, Jean-paul Ampuero, Adolfo Inza, Hernando Tavera. Earthquake Early Warning Starting From 3 s of Records on a Single Station With Machine Learning. *Journal of Geophysical Research: Solid Earth*, 2023, 128 (11), 10.1029/2023JB026575 . hal-04280699

HAL Id: hal-04280699

<https://hal.science/hal-04280699>

Submitted on 17 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License


JGR Solid Earth



RESEARCH ARTICLE

10.1029/2023JB026575

Earthquake Early Warning Starting From 3 s of Records on a Single Station With Machine Learning

Pablo Lara^{1,2} , Quentin Bletery¹ , Jean-Paul Ampuero¹ , Adolfo Inza² , and Hernando Tavera² 

¹Université Côte d'Azur, IRD, CNRS, Observatoire de la Côte d'Azur, Géoazur, Valbonne, France, ²Instituto Geofísico del Perú, Lima, Perú

Key Points:

- Earthquake Early Warning System (EEWS) with faster and robust preliminary estimates of magnitude and location to generate a first alert
- First EEWS designed purely with Artificial Intelligence
- Earthquake Early Warning with continuous updates that follow the earthquake source time function

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

P. Lara,
pablo.elara@ieee.org

Citation:

Lara, P., Bletery, Q., Ampuero, J.-P., Inza, A., & Tavera, H. (2023). Earthquake Early Warning starting from 3 s of records on a single station with machine learning. *Journal of Geophysical Research: Solid Earth*, 128, e2023JB026575. <https://doi.org/10.1029/2023JB026575>

Received 21 FEB 2023
Accepted 22 OCT 2023

Author Contributions:

Investigation: Pablo Lara, Quentin Bletery, Jean-Paul Ampuero, Adolfo Inza, Hernando Tavera
Methodology: Pablo Lara
Resources: Hernando Tavera
Supervision: Quentin Bletery, Jean-Paul Ampuero, Adolfo Inza
Writing – original draft: Pablo Lara
Writing – review & editing: Quentin Bletery, Jean-Paul Ampuero

Abstract We introduce the Ensemble Earthquake Early Warning System (E3WS), a set of Machine Learning (ML) algorithms designed to detect, locate, and estimate the magnitude of an earthquake starting from 3 s of *P*-waves recorded by a single station. The system is made of six Ensemble ML algorithms trained on attributes computed from ground acceleration time series in the temporal, spectral, and cepstral domains. The training set comprises data sets from Peru, Chile, Japan, and the STEAD global data set. E3WS consists of three sequential stages: detection, P-phase picking, and source characterization. The latter involves magnitude, epicentral distance, depth, and back azimuth estimation. E3WS achieves an overall success rate in the discrimination between earthquakes and noise of 99.9%, with no false positive (noise mis-classified as earthquakes) and very few false negatives (earthquakes mis-classified as noise). All false negatives correspond to $M \leq 4.3$ earthquakes, which are unlikely to cause any damage. For P-phase picking, the Mean Absolute Error is 0.14 s, small enough for earthquake early warning purposes. For source characterization, the E3WS estimates are virtually unbiased, have better accuracy for magnitude estimation than existing single-station algorithms, and slightly better accuracy for earthquake location. By updating estimates every second, the approach gives time-dependent magnitude estimates that follow the earthquake source time function. E3WS gives faster estimates than present alert systems relying on multiple stations, providing additional valuable seconds for potential protective actions.

Plain Language Summary We designed E3WS, an Earthquake Early Warning System (EEWS) using the first 3 s of seismic waves recorded by a single station. E3WS is the first EEWS entirely made of Artificial Intelligence algorithms. Our results show an improvement in magnitude estimates compared to existing single-station based EEWS, such as Urgent Earthquake Detection and Alarm System and ElarmS, and slightly better localization. In order to track the magnitude of $M > 7.0$ earthquakes, whose source duration is typically longer than 6 s, we test E3WS on large earthquakes in real-time scenarios. E3WS provides time-dependent magnitude estimates that follow the earthquake source time functions. E3WS is highly exportable. Thanks to its simplicity, it can be installed in 32-bit Single Board Computers like a Raspberry Pi 4, and of course in more complex 64-bit processors. E3WS is not only theoretical, it is already running in alpha test mode as the core of the EEWS of Peru in construction. We believe that E3WS will provide valuable seconds to populations living at constant risk, giving the user extra seconds that may give them enough time to "drop, cover, and hold on" or to perform mitigation actions like stopping traffic, stopping elevators or evacuating the ground floor of buildings.

1. Introduction

Today millions of people live at risk from earthquakes and nearly 1 million deaths have been due to earthquakes since 1990 (Pal et al., 2023). Real-time seismic monitoring in strategic areas where large earthquakes are likely to occur opens the possibility of rapidly estimating earthquake parameters that control the potential impact of shaking, such as rupture geometry and dimension, magnitude and hypocenter location. Taking advantage of such estimates and speed-of-light communications, Earthquake Early Warning Systems (EEWS) can generate an alarm before the propagation of the hazardous seismic shaking completes, with the goal to mitigate human and material losses.

An ideal EEWS accurately estimates ground shaking and time of impact in a region soon after an earthquake is detected, then notifies the population or infrastructure at risk with sufficiently long lead time to take protective measures. In a common approach, the magnitude and location of the earthquake are rapidly estimated, then

© 2023. The Authors.

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

used as input in a Ground Motion Prediction Equation to forecast ground shaking intensity measures such as peak ground acceleration (PGA), whose values are used as a criterion to alert the population. However, there is a trade-off between achieving the earliest alert time and improving the accuracy of the estimates (Meier, 2017): waiting for additional data containing more information about the earthquake improves the magnitude and location estimates, but delays the issuance of alerts. More importantly, the earlier the estimates, the worse is the underestimation of the final magnitude (Melgar & Hayes, 2019). Therefore, for a single station, we must exploit a long enough window of earthquake signal to estimate whether the earthquake can be dangerous or not, even if we cannot predict the final magnitude of the earthquake. Three seconds of earthquake records is a good starting point because most of $M < 6$ earthquakes have half-durations shorter than 3 s (Meier et al., 2017). Thus, we can expect that, in most cases, at 3 s we can tell if an earthquake will be larger than M6, which is large enough to be damaging. Nowadays most EEWs are based on multi-station data, to improve accuracy by taking advantage of more information, at the expense of additional delays. Here we bring back the single-station based EEWs, as it has the potential to be faster since it does not require waiting for seismic wave arrivals at multiple stations.

The practice of EEWs dates back to 1988 with the deployment of the Urgent Earthquake Detection and Alarm System (UrEDAS) in Japan, the first operational system based on the analysis of a few seconds of P waves recorded by a single station to estimate earthquake source parameters (Nakamura, 1988; Nakamura et al., 2011). Since then, a number of EEW algorithms have been developed using records from broadband seismometers, strong-motion accelerometers, Global Navigation Satellite System (GNSS) stations (R. M. Allen & Melgar, 2019) and cellphones (Bossu et al., 2022).

The τ_c - P_d Onsite algorithm (Böse et al., 2009), one of the three algorithms that contributed to the development of ShakeAlert, the EEW of the US West coast (Böse et al., 2014), uses the period parameter τ_c and the peak initial-displacement amplitude P_d (Yih-Min et al., 2007) extracted from the first 3 s of the P-wave recorded by a single sensor. The algorithm estimates the P phase arrival based on a combination of the classic STA/LTA (R. V. Allen, 1978) with a P/S wave discriminator which uses the ratio of horizontal to vertical ground motions. It estimates the magnitude and the Modified Mercalli Intensity but not the earthquake location, thus it is intended for on-site warning instead of regional warning.

Most single-sensor-based algorithms only contain some components of an EEWs (detection, picking, magnitude or location), but not the whole package. The only exception is UrEDAS. However, it does not present the same performance when estimating the back azimuth for earthquakes outside Japan. In particular, when applied in California, UrEDAS estimates showed larger error and yielded several cases of magnitude overestimation for earthquakes with magnitudes between 3.0 and 5.0 (Nakamura & Saita, 2007).

Artificial Intelligence (AI) has been used in a number of applications in seismology that are relevant for EEW. Böse et al. (2012) rapidly estimates seismic source and ground motion parameters using a 3-component sensor with an artificial neural network, obtaining 60% smaller errors than the τ_c algorithm of Kanamori (2005). Meier et al. (2019) classifies between earthquakes and impulsive noises using fully connected neural network, Convolutional Neural Network (CNN), Recurrent Neural Network (RNN) and generative adversarial network with Random Forest (RF). Using 3-s-long records it achieves 99.5% of accuracy. Ochoa et al. (2018) uses Support Vector Machine (SVM) to estimate the local magnitude of earthquakes based on a single 3-component station and the first 5 s of the P-wave, reaching 0.19 of standard deviation (STD) error when estimating magnitudes between M3.0 and M4.5. Münchmeyer et al. (2021) estimates the source characterization based on a dynamic network of stations and the first seconds of P-waves with a transformer network. For the Chile region and using 0.5 s of P-waves, the model achieves a root mean square error for magnitude and location of ~ 0.3 magnitude units and 40 km, respectively. Moreover, there are relevant studies not designed for EEW purposes that use AI for specific targets, such as detection, picking and source characterization components. The Earthquake Transformer algorithm (EQTransformer) (Mousavi et al., 2020) uses 1 min long seismograms to feed an architecture based on CNN and RNN to detect earthquakes, and estimate P and S phase arrivals. The model achieves an earthquake detection precision (true positives divided by total predicted positives) of 1.0, and estimates the P and S phase arrivals with a mean \pm STD of 0.0 ± 0.03 s for the P phase, and 0.0 ± 0.11 s for the S phase. Mousavi and Beroza (2020) estimate magnitudes with CNN and RNN trained with 30 s long seismograms ($M < 5.7$). They obtain an error of 0.0 ± 0.2 magnitude units. Mousavi and Beroza (2019) estimate earthquake location based on Bayesian Deep Learning. The network is fed by 1 min long window seismograms in the case of training distances and 1.5 s long window seismograms in the case of back azimuth. The network achieves a localization error of

7.27 ± 12.16 km. While these AI models are not designed for EEW (they are trained to use many seconds of *P*-wave, signal windows that are too long from multiple stations, resulting in long processing times), they represent a useful reference to evaluate the performance of AI-based EEW approaches.

Here, we present E3WS, the first EEWS in which all components (detection, picking, and source parameter estimation) are based on AI. It uses only 3 s of signal recorded by a single three-component instrument. E3WS is a system focused on early warning for populations living near seismic sources (distances less than 200 km). Extra seconds of alert time can give the user enough time to “drop, cover and hold on” or to perform mitigation actions like stopping traffic, stopping elevators or evacuating the ground floor of buildings (Cremen et al., 2022). Compared to current single-station-based EEWS, E3WS estimates earthquake magnitudes with significantly better accuracy and locations with slightly better accuracy. It requires no additional software to estimate *P*-phase arrival, and estimates source characterization without applying signal-to-noise ratio constraints or acceleration thresholds. E3WS can be applied anywhere and is designed using simple Machine Learning (ML) models (fed by attributes given by humans), allowing in contrast to Deep Learning approaches (where the model itself is responsible for extracting its own attributes) some understanding of what controls the estimations. Thanks to its simplicity, E3WS can be installed in complex 64-bits processors or in small 32-bits single board computers, such as a Raspberry Pi, opening the possibility to process the data on site: the Raspberry Pi 4 computer of each station processes the three channels of that station. The advantage of the on-site-processing approach is that only warning information of small size needs to be transferred in real time to the warning center, not the whole waveform data, which makes communication lighter, faster and more robust.

2. Database

We build a database of seismic waveforms combining data from the Instituto Geofísico del Perú (IGP) recorded between 2017 and 2019, the STEAD global database (Mousavi et al., 2019), the Seismic Network of Chile (Barrientos & Team, 2018), and the Japanese seismic networks K-Net and KiK-net (Aoi et al., 2004). The data are recorded from strong-motion accelerometers (Japanese database) and broadband velocimeters (Peru, Chile, and STEAD). We select events with magnitude greater than 3.0, depth shallower than 100 km and recordings at epicentral distance shorter than 200 km. We consider 3-component accelerograms oriented to the east, north and vertical directions, respectively. In total, we compile a database of more than 22,000 earthquake events associated with ~100,000 earthquake observations that contain at least 30 s of *P*-wave. Data statistics are shown in Figure 1.

As the data come from different sources and have different sampling frequencies, sensors and digitizer types, we preprocess them to standardize our database. Preprocessing steps consist of removal of the mean to avoid low-frequency artifacts, removal of a least-squares-fitted linear trend, multiplication by a cosine taper at each end over 2.5% of the total window duration (see Section 3 for the analysis window time setting), and resampling using the Fourier method at 100 Hz. We convert the preprocessed data from broadband seismometers and accelerometers to acceleration in m/s^2 by deconvolution. We ignore sensors for which we did not find the instrument response.

For the Peru and Chile data sets, we ignore seismograms that exceed 80% of the dynamic range of the digitizer, to avoid clip level saturation. Given that the STEAD data set contains preprocessed signals (detrended and resampled at 100 Hz), we manually identified and removed signals that were saturated for $M > 5.0$. We found no further saturated waveforms when inspecting events down to $M 4.5$. We believe that for magnitudes smaller than 4.5, there are no more saturated waveforms. The Japanese data are recorded by strong-motion accelerometers that did not show saturation.

3. Proposed EEWS

E3WS consists of 6 ML algorithms: a detector, a *P*-phase picker and four regression models estimating the magnitude, epicentral distance, back azimuth, and depth of the source. The detector model monitors the seismic activity. When it detects an earthquake, the *P*-phase picker is triggered. Then, using a minimum of 3 s of *P*-wave signal, the four regression models run independently and estimate the magnitude and location of the event. Estimations are updated at regular times thereafter, as the signal window lengthens. For each signal window, these six models take as input a feature vector formed by concatenating 140 attributes extracted from the waveforms, their spectrum and their cepstrum. We built E3WS with the goal that all components of the EEWS (detection, *P*-phase

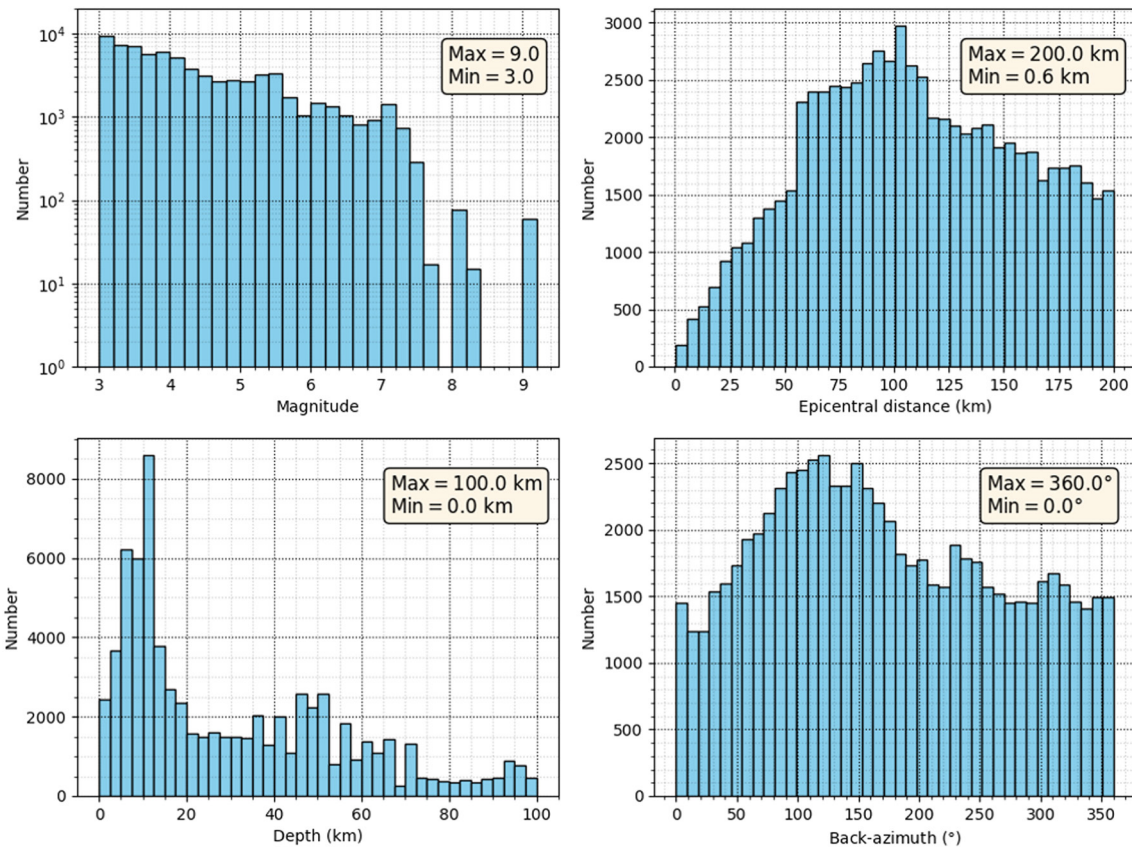


Figure 1. Magnitude, epicentral distance, depth, and back azimuth distributions of the earthquake waveform database compiled for this work.

picking, and source characterization) are based on the same definition of attributes, arranged in a feature vector of length 140, avoiding the use of other algorithms that can increase the processing latency. Applying simple ML models to the E3WS provides some understanding of how these attributes work in the EEWS.

We test several approaches to design the models, including Extreme Gradient Boosting (XGB), SVM, RF, and Neural Networks (Multilayer Perceptron [MLP]). We find that the approach yielding the best results is XGB (Chen & Guestrin, 2016), a Supervised ML model that has become popular for its leading performance in Kaggle competitions (Nielsen, 2016) and that has been recently applied to seismology (Shokouhi et al., 2021).

We train the models using 80% of the database and we test on the remaining 20% (more details in Section 4.1), based on the first 3 s of the *P*-wave, on an Intel(R) Xeon(R) Silver 4114 processor. In the testing stage, detection results show that XGB has an overall accuracy (correct classifications divided by total test samples) of 99.95%, slightly better than the other models (Table 1). For *P*-phase picking, XGB and RF achieve the best performances, with similar responses. However, XGB is far superior in terms of computational cost of real-time processing: the average time complexity per estimate is 0.01 s for XGB compared to 0.49 s for RF. Even on a Single Board Computer like a Raspberry Pi 4, XGB takes 0.01 s on average and it only increases power by 0.75 W. Furthermore, due to the large storage required by the RF model (1.6 GB) compared to XGB (7.6 MB), it was not possible to load the RF model on the Raspberry Pi 4. For source characterization, XGB gives a smaller Mean Absolute Error (MAE) than the other models.

Moreover, XGB can be accelerated by multi-threading, which we exploit here: we train our models on 80 CPU cores in parallel. XGB is based on the Ensemble Learning approach: it uses multiple sub-models (decision trees) to

Table 1
Accuracy and Errors Using XGB, Random Forest, Support Vector Machine, and MLP Models for Detection, P-Phase Picking and Source Characterization

| Performance | XGB | RF | SVM | MLP |
|-------------------------|-------|-------|-------|-------|
| DET _% | 99.95 | 99.94 | 99.92 | 99.90 |
| PICK _{MAE} (s) | 0.10 | 0.14 | 0.18 | 0.16 |
| MAG _{MAE} | 0.34 | 0.38 | 0.47 | 0.39 |
| DIS _{MAE} (km) | 27 | 29 | 33 | 30 |
| DEP _{MAE} (km) | 15.7 | 28.9 | 18.0 | 17.7 |
| BAZ _{MAE} (°) | 45.1 | 47.0 | 52.3 | 51.1 |

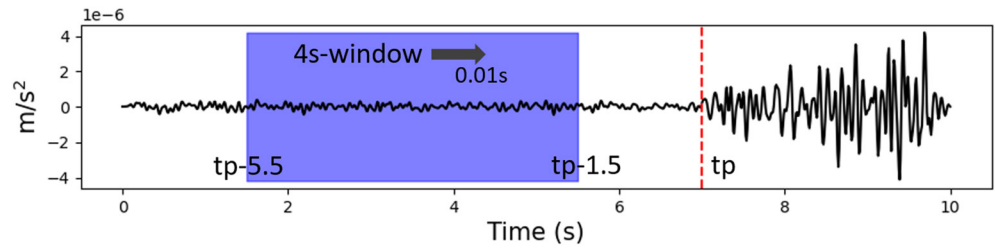


Figure 2. Labeling for the P-phase picking model. We extract attributes from a 4-s-long window, starting from $t_p - 5.5$ s as the blue box, for our entire database. We repeat the attribute extraction every 0.01 s until the blue box reaches $t_p + 2.5$ s.

improve the final estimation. It applies the ensemble technique boosting, which consists of sequentially decreasing the residuals along each tree, and a gradient descent algorithm to minimize the loss function. Figure S1 in Supporting Information S1 shows the general XGB scheme. For all models, we use the following hyperparameters for XGB training: depth = 4, number of trees = 6,000, subset = 80%, and learning rate = 0.1.

3.1. Detection

Since STEAD is a global data set that also includes global noise samples which includes environmental and cultural noise, we extract 55,000 noise windows and add them to our database. We estimate the likelihood that a window contains a *P*-wave, sliding a 10-s window by steps of 0.5 s. We use 10-s waveforms, which we preprocess and filter from 1 to 7 Hz applying a fourth-order Butterworth band-pass filter. To avoid triggers caused by impulsive noise, we consider the average over several sliding windows: if the average of the likelihood of containing a *P* phase of three consecutive windows is less than a threshold of 0.21, we classify it as noise; otherwise, we classify it as an earthquake. The choice of the threshold value is described in Section 4.1.1.

In our tests we obtained better accuracy using a 10-s-long window compared to shorter windows. For instance, we find false detections due to impulsive noise using shorter windows; 10 s-long windows limit false detections by lowering the weight of impulsive noise in the attributes.

We train the XGB model as a classifier between noise, *P*-waves and *S*-waves. We label a window as class 0 if it contains only noise, and class 1 or 2 if the window contains 0.5, 1.0, ..., 4.0 s of *P*- or *S*-wave, respectively. Therefore, each earthquake seismogram generates eight samples of class 1 and 8 samples of class 2. We discard waveforms that do not contain 10 s of data. For instance, a class 1 sample with 0.5 s of *P*-wave, requires 9.5 s of noise before the *P*-wave arrival. Leading to a database for the detection of ~60,000 earthquake seismograms. Although our focus is on the analysis of the *P*-wave signal, we add an *S*-phase class in the training so that our system does not trigger with *S*-waves.

3.2. P-Phase Picking

Because the Japanese and Chilean data sets do not provide *P*-phase arrival times (t_p), we restrict the training set for phase-picking to the Peruvian and STEAD data sets. We use a 4-s window, which we preprocess and filter between 1 and 7 Hz applying a fourth-order Butterworth band-pass filter, to detect 0.5 s of *P*-wave. We discard waveforms that do not contain 3.5 s of noise before the *P*-wave arrival.

We train the XGB model as a classifier between noise, 0.5 s of *P*-wave and 0.5 s of *S*-wave. This classifier works as a scan, where we label class 1 when the 4 s-long window contains 0.5 s of *P*-wave signal, class 2 when it contains 0.5 s of *S*-wave, and class 0 otherwise. We include the *S*-phase to minimize the error in *P*-phase picking when the 4 s-long window contains both the *P* and *S* phases.

We feed the model with attributes extracted from a 4 s-long window sliding with a step of 0.01 s covering the interval $t_p - 5.5$ s to $t_p + 2.5$ s (Figure 2). The estimated *P* arrival time is the ending time of the first 4 s-long window classified as Class 1 minus 0.5 s. We proceed similarly for the *S*-phase.

We use the time window $[t_p - 3.5$ s, $t_p + 0.5$ s] as label 1 because of the natural uncertainty in the catalog arrival times. We trust that the uncertainties in the *P*-phase arrival times of the catalogs are less than 0.5 s. With attributes

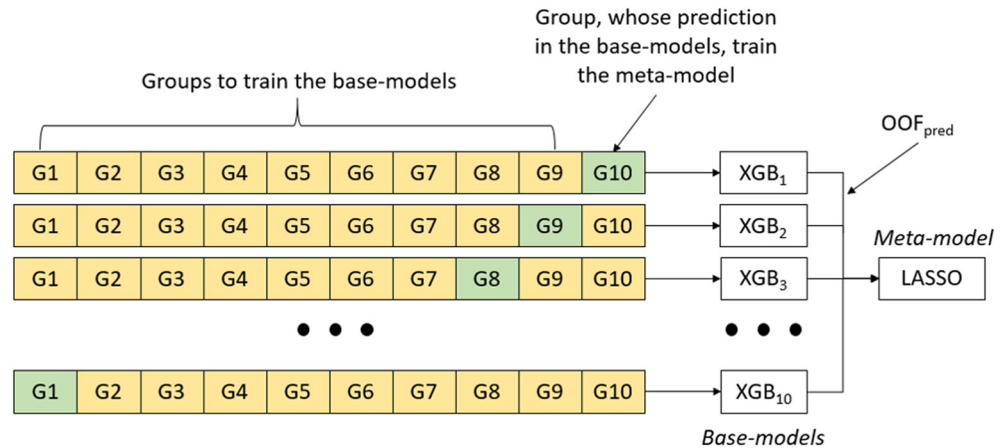


Figure 3. Source characterization model based on Stacking algorithm and K-Fold with $K = 10$. For each K , nine groups train the XGB base-model. Estimates from the remaining group, using the corresponding trained XGB model, feed the Least Absolute Shrinkage and Selection Operator meta-model.

extracted every 0.01 s, the input data set for the P-phase picking model is made of approximately 36 million 140-dimensional (140 attributes) samples.

3.3. Source Characterization

For source characterization (estimation of earthquake magnitude, epicentral distance, back azimuth, and hypocentral depth), we use time windows that contain 7 s of noise and 3 s of *P*-wave signal extracted from our earthquake database. We preprocess and apply a fourth-order Butterworth band-pass filter from 1 to 45 Hz. Because *P*-phase accuracy is crucial when estimating back azimuth using only one station, we select only data sets that have a *P*-phase catalog. To train the back azimuth estimation model, we only use STEAD and Peruvian samples and select only the stations that are properly oriented to the east (azimuth 90°) and north (azimuth 0°).

We train each model independently. These models are based on the Stacking algorithm (Cui et al., 2021), which uses a set of models per layer. The outputs of the models in the first layer, called base-models, feed a model in the second layer, called meta-model (Figure 3). The main idea of using Stacking is to reduce the error by increasing the heterogeneity of the data by using multiple subsets of the original database, and combining them with the meta-model to generate the estimates.

The base-models are obtained by an XGB regressor, with the same hyper-parameters as used for detection and *P*-phase picking. The meta-model is obtained by the Least Absolute Shrinkage and Selection Operator (LASSO).

For each model, we perform K -fold validation, splitting the data set into $K = 10$ groups and training each XGB base-model with nine out of the 10 groups. Then, the remaining Out-of-Fold group of validation is estimated by the trained XGB model. Finally, we combine all the estimates for each Out-of-Fold group (OOF_{pred}) to train the LASSO meta-model (Kukreja et al., 2006).

For the back azimuth model, we divide the training into two targets. Because the angle is represented in non-Euclidean space, we train two separate models to estimate its cosine and sine, respectively.

3.4. Feature Vector

For all of the algorithms, we compute the same set of 140 attributes, in the time, spectral, and cepstral (spectrum of the spectrum) domains. For the time domain, we extract attributes from the pre-processed signal s and from its envelope, defined as the absolute value of its analytic signal $|s + iH\{s}|$ where H is the Hilbert transform. For the spectral domain, we consider the Power Spectral Density of the signal estimated by Welch's method using an overlap of 75%, a Fourier Transform length of 512 samples and a Hanning taper function. For the cepstral domain, we use the first 13 Mel-frequency cepstral coefficients (MFCC) (Davis & Mermelstein, 1980), attributes that showed high relevance in seismic signals, even better than the temporal and spectral domain in volcano seismology (Lara et al., 2020).

Table 2
Confusion Matrix for the Detection Algorithm

| | | True class | |
|-------------------|------------|------------|------------|
| | | Noise | Earthquake |
| Overall (%): 99.9 | | | |
| Detected class | Noise | 11,264 | 10 |
| | Earthquake | 0 | 8,788 |
| Accuracy (%): | | 100.0 | 99.9 |

In total, we extract 45 attributes for each channel: 17 in the time domain, 15 in the spectral domain, and 13 in the cepstral domain. We add 5 attributes from the combination of the 3-component signal: the maximum eigenvalue, the eigenvector associated with the maximum eigenvalue, and the ratio of the maximum eigenvalue to the sum of the remaining eigenvalues. We then concatenate all the features in a single vector, generating a 140-dimensional feature vector. Generating the 140-attribute feature vector takes 0.1 s on the Raspberry Pi 4. We provide the complete list of attributes in the Supporting Information. Most of them were previously used in (volcano) seismology by Malfante et al. (2018) and Lara et al. (2020).

4. Results

Here, we evaluate the performance of E3WS. First, we analyze the models that compose E3WS using hold-out validation, with 3 s of *P*-wave signal. Next, we evaluate the behavior of the system when using longer signal time windows. Then, we apply E3WS to track the magnitude of a set of earthquakes with $M > 6.0$ in simulated real-time conditions and compare the performance with existing EEWS methods. Finally, we show an application of E3WS in a real-time scenario in Peru.

4.1. Hold-Out Validation Using 3 s of *P*-Wave Signal

We assess the behavior of the E3WS target models through Hold-out validation. We assign 80% of the seismic events and their associated observations to the training stage, and the remaining 20% to the testing stage. To avoid data leakage, we use Hold-out validation on seismic events and then we associate their observations, which prevents having events with observations in both the training stage and the testing stage.

4.1.1. Detection

The detector model achieves its best performance for a *P*-phase likelihood threshold of 0.21 (Figure S2 in Supporting Information S1), reaching an overall success rate of 99.9% in the discrimination between noise and earthquakes (Table 2). For a total of 11,264 noise observations, 100% of noise samples are correctly classified. From 8,788 earthquake observations, 10 are misclassifications, leading to a success rate of 99.9% for earthquake classification. All of these misclassifications belong to earthquakes with $M < 4.3$ (Figure S3 in Supporting Information S1) and low signal-to-noise ratios (Figure S4 in Supporting Information S1). Most of them have an epicentral distance greater than 100 km.

4.1.2. *P*-Phase Picking

We evaluate the picker model on more than 10,000 seismograms of the test data set compared to the “true” (manually picked) *P*-wave arrival times (Figure S5 in Supporting Information S1). The model achieves a *P*-phase arrival time error with mean of 0.03 s, STD of 0.14 s, and MAE of 0.10 s.

4.1.3. Source Characterization

The performance of the source characterization is remarkable (Figure 5, Figure S6 in Supporting Information S1, Table 3), given that our algorithm only uses 3 s of records on a single station.

The magnitude estimates are very stable for earthquakes with magnitudes smaller than $M7.2$, with magnitude average residuals ($M_{pred} - M_{true}$) between 0.2 and ~ 0.4 for $M < 6.0$ (Figure 4a), and residuals between ~ 0.3 and 0.6 for $M6.0 < M < M7.1$. This gives us confidence in estimating magnitudes for minor ($M3.0$ – $M3.9$) to strong ($M6.0$ – $M7.1$) earthquakes. The small errors over the entire range of magnitudes are reflected in a high R^2 of 0.87 (1.0 in the ideal case). For the smallest earthquakes of our database ($M \sim 3$), the magnitude estimates show a slight overestimation of 0.3 and STD of 0.2 (Figure S7a in Supporting Information S1), but that is not a problem for EEWS because such small earthquakes do not warrant alerts. For $M > 6.5$ the estimated magnitudes saturate and underestimate the real values. This magnitude saturation is expected: the half duration of $M > 6.5$ earthquakes is typically longer than the 3 s window duration (Meier et al., 2017; Melgar & Hayes, 2019).

We observe an average residuals at epicentral distances for distances very close to the seismic source (0–20 km) of ~ 33 km (Figure 4b). As the seismic source gets farther away up to a distance of ~ 100 km, the residuals

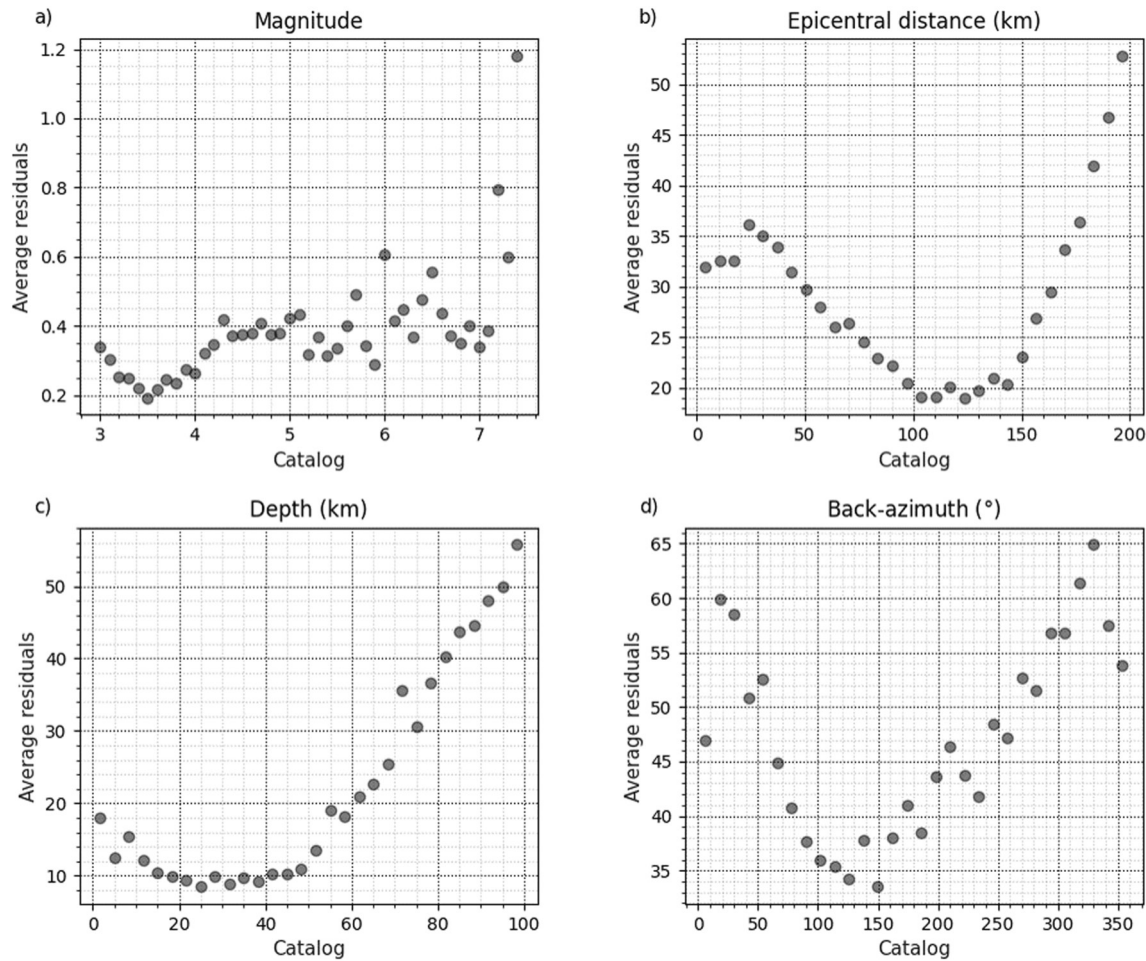


Figure 4. Average residuals ($target_{pred} - target_{true}$) for each target: Magnitude, epicentral distance, depth, and back azimuth, using the first 3 s of P -wave.

decrease linearly to almost 20 km of error, and continues with this error up to 145 km. From here, the errors grow linearly up to our training distance limit (200 km), which is why R2 is not high (0.50). If we keep our error tolerance at ~ 33 km (errors at very close distances), we can estimate up to an epicentral distance of 170 km. Longer distances are associated with larger errors. This behavior shows that the information contained within 3 s of P phase is not sufficient to resolve accurately such large epicentral distances.

From our database, the earthquakes that represent significant hazard ($M > 6.0$) have hypocentral depths of 28 km on average and STD of 20 km. Within the range of the average \pm STD (8–48 km depth), most events have average residuals of 10 km (Figure 4c). The residuals do not exceed 20 km down to depths less than 60 km. This means that if E3WS estimates an earthquake with $M > 6.0$, it is very likely that the error in depth is not greater than 10 km, and almost certainly the error is less than 20 km. For earthquakes deeper than 60 km, the residual errors grow linearly up to the maximum depth in our database (100 km), which explains the small R2 of 0.32. For $M > 6$ earthquakes that do not belong to our database, that is, earthquakes deeper than 48 km (28 km average plus 20 km STD), E3WS estimates around our maximum hypocentral depth of 48 km. A clear example is the 2023 Ecuador earthquake of magnitude 6.8 and hypocentral depth 68 km: E3WS estimates a magnitude of M6.4 and a depth of 43 km (Figure S11 in Supporting Information S1).

For back azimuth, residuals exceed 33° . However, the STD of the estimates decreases significantly as the magnitude increases, achieving an STD of $\sim 21^\circ$ for $M > 6.0$ earthquakes (Figure S8 in Supporting Information S1). The estimates have uniform performance throughout their range (Figure S7d in Supporting Information S1). The high R2 of 0.84 shows that the model contains relevant information in the whole back azimuth range.

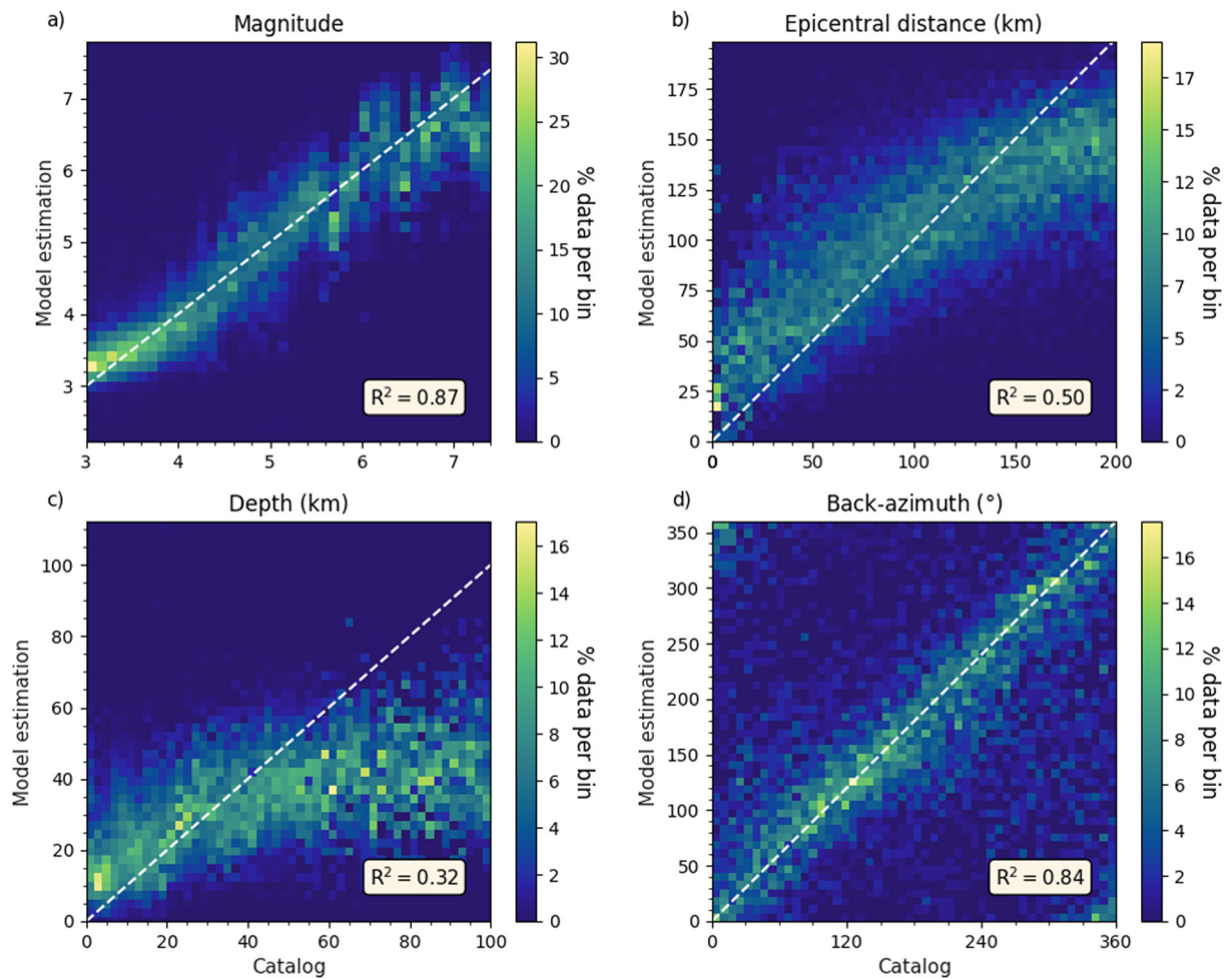


Figure 5. Estimated source parameters (magnitude, distance, depth, and back azimuth) using 3 s of records as a function of cataloged values. Color bar represents the percentage of data per bin.

4.2. Performance of Source Characterization Using Longer Signals

Meier et al. (2017) showed that the source time functions (STF) of $M \geq 7$ shallow subduction earthquakes have a similar evolution until the maximum moment rate is reached, suggesting that the beginning of the rupture does not contain enough information to estimate the final magnitude of the event. However, we can estimate the instantaneous magnitude using the first 3 s of the P -wave, that is, the magnitude reached by the earthquake 3 s after its onset. This estimate can form the basis to generate a first warning and can be updated when longer records become available.

Table 3
Performance of the Source Characterization Algorithm With Its Mean Error, Standard Deviation Error, Mean Absolute Error, and Coefficient of Determination (R^2)

| Performance | Mean | STD | MAE | R2 |
|-------------|------|------|------|------|
| MAG | 0.0 | 0.45 | 0.34 | 0.87 |
| DIS (km) | -0.3 | 34.3 | 27.1 | 0.50 |
| DEP (km) | -1.4 | 20.8 | 15.7 | 0.32 |
| BAZ (°) | 2.7 | 43.7 | 45.2 | 0.84 |

To evaluate how much information the ML algorithms can leverage with more time, we retrain our algorithms using longer seismic signals. We increment the P -phase window duration by steps of 1 s from 3 to 46 s of both training and test data sets. Each additional second requires reading 140 MB of source characterization model files, and uses 300 MB of RAM with 0.05 W on a Raspberry Pi 4 device. E3WS takes 0.4 s to estimate the source characterization (magnitude and location). Figure 6 shows the evolution of two performance metrics, MAE and R^2 , as a function of the considered signal duration.

We observe a significant improvement in the estimations of magnitude and epicentral distance, with R^2 scores increasing up to 0.94 and 0.93,

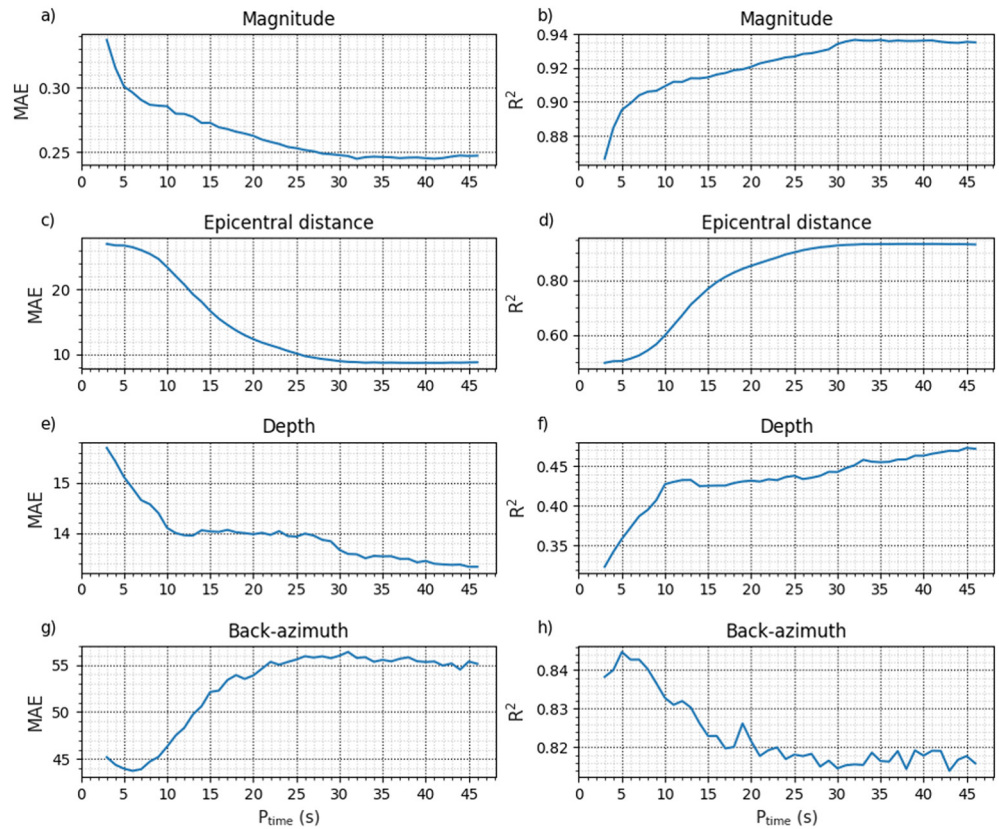


Figure 6. Mean Absolute Error and R2 results using 3–46 s of *P*-wave.

respectively, and MAE dropping to about 0.25 and 9 km, respectively, at 46 s of signal (Figure 7, Figure S9 in Supporting Information S1). After that time, most $M \leq 7$ earthquakes are indeed over, which allows the model to estimate the final magnitude, and the S phase has arrived, which allows the model to infer the epicentral distance from the arrival time difference between P and S phases. A signal duration of 30 s of training seems sufficient to converge to the best performance (Figures 6a–6d).

The depth estimates improve slightly over time (Figures 6e and 6f). From 10 to approximately 27 s, the estimates do not improve. After this time the model improves slightly.

For the back azimuth estimation, the best model uses 5 s of *P*-wave, because the relevant information (likely the polarization) is contained in the first few seconds of the signal. The two most important attributes for the cosine model are the eigenvectors in the north and vertical components associated with the maximum eigenvalue, and for the sine model the vertical and east components. The importance of attributes is based on the contribution of the attributes to the model, where an attribute is relevant if it improves the estimates (see Section 5.2 for details).

4.3. Performance of E3WS on Selected Large Earthquakes

We test the performance of E3WS to estimate the magnitude over time for different large earthquakes ($M > 6$) using strong-motion accelerometers located in Japan, Chile, and Peru. We apply the Leave-one-out method: in each example, the selected event and all its observations are put in the test data set and the remaining observations in the training set. We convert the data from these earthquakes into Earthworm Tankplayer format to simulate real-time data processing, with a transmission of data packets every second and neglecting transmission delay, and we estimate the magnitude using a minimum of 3 s and a maximum of 60 s after the P-phase arrival. We compare E3WS estimations to those obtained by other EEW algorithms based on multiple stations, using broadband or strong-motion sensors such as ElarmS-3 (Chung et al., 2019), Finder2 (Böse et al., 2018), Japan Meteorological Agency (JMA) (Hoshiya & Ozaki, 2014) and PEGSNet (Licciardi et al., 2022), and GNSS stations such as

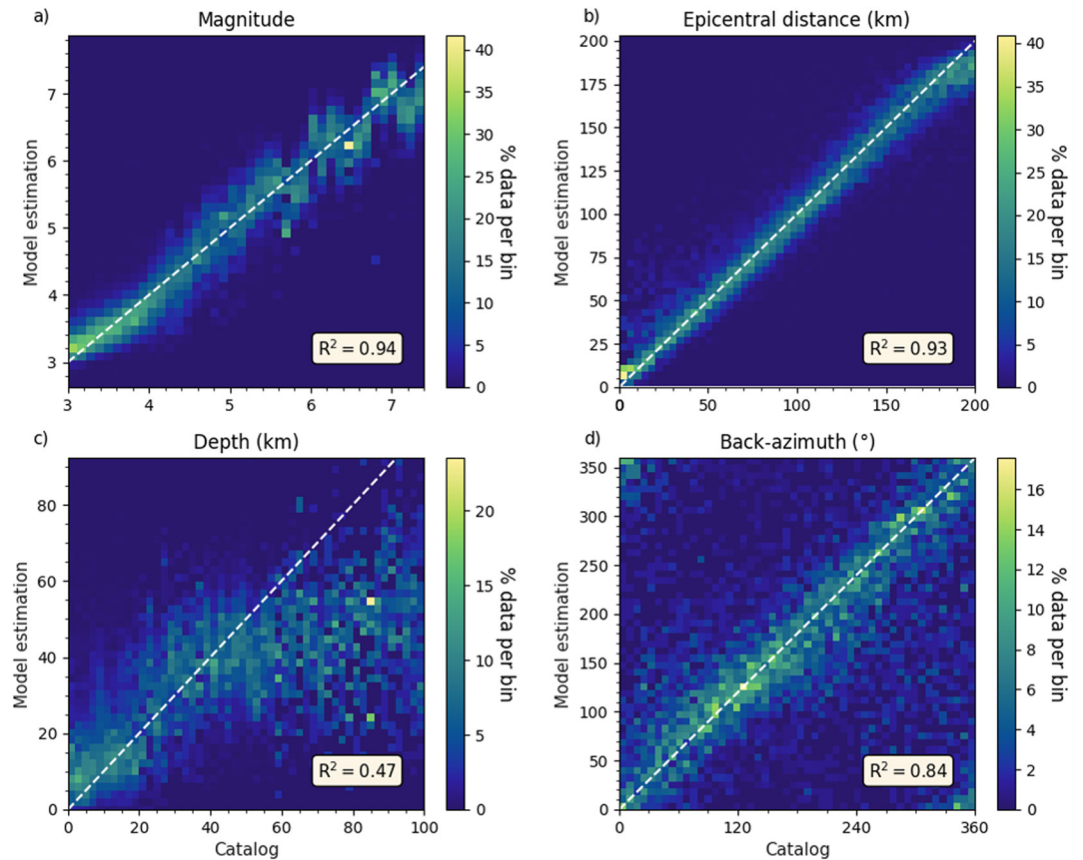


Figure 7. Same as Figure 5 using 46 s of signal after *P*-wave arrival.

BEFORES (Minson et al., 2014) and G-larmS (Grapenthin et al., 2014a, 2014b). For a true real-time comparison, we use the G-larmS triggered by ElarmS (ElarmS → G-larmS), as mentioned in Ruhl et al. (2019).

Figure 8a shows the results for the 2011 M_w 9.0 Tohoku, Japan earthquake. For reference, we show also the STF (the “true” instantaneous magnitude) and the STF shifted by the *P* arrival time at station MYG011, to compare both timeliness and accuracy. The first E3WS estimate uses 3 s of records after the first arrival at the station closest to the epicenter (MYG011, 120 km from the epicenter) and is obtained approximately 17 s after origin time (OT). ElarmS-3 uses at least 0.2 s of *P*-waves recorded by 3 stations (Ruhl et al., 2019). Owing to the high density of seismic stations in Japan and to the shortness of its first data window, ElarmS-3 issues its first estimation almost at the same time as E3WS.

E3WS outperforms in timeliness and accuracy the first estimates of the other EEWs based on broadband or strong-motion sensors. At the time of the first E3WS estimate, the true instantaneous magnitude (shifted by *P*-wave arrival time) is $M6.9$, while E3WS estimates $M5.2$, ElarmS $M4.9$, JMA $M4.3$ (4 s later) and Finder2 $M4.0$ (7 s later). BEFORES makes its first estimate ($M6.4$) at 20 s after OT when the true instantaneous magnitude is $M7.4$, outperforming the estimation of $M5.7$ by E3WS. However, 1 s later, E3WS outperforms the GNSS station-based systems in accuracy, giving $M6.9$ compared to $M6.5$ by BEFORES and $M6.8$ by G-larmS, when the true magnitude is $\sim M7.4$. E3WS magnitude estimates increase until 31 s after OT (17 s of *P*-wave) with estimates that are very close to the true instantaneous magnitude, then remains similar to the JMA estimate up to 62 s after OT. At the end of our analysis window, at 74 s after OT, E3WS and BEFORES achieve similar performance, 0.2 points of magnitude below PEGSnet. We take only $M_w \geq 8.3$ estimates for PEGSnet, because estimates are not reliable below this magnitude (Licciardi et al., 2022).

We also generate instantaneous magnitude estimates using all the strong-motion recordings available within a distance of 200 km from the epicenter. We show these estimates as a function of time relative to the *P*-wave arrival time (P_{time}) of each station, to compare them to the event's STF (Figure 8b) given by the SCARDEC

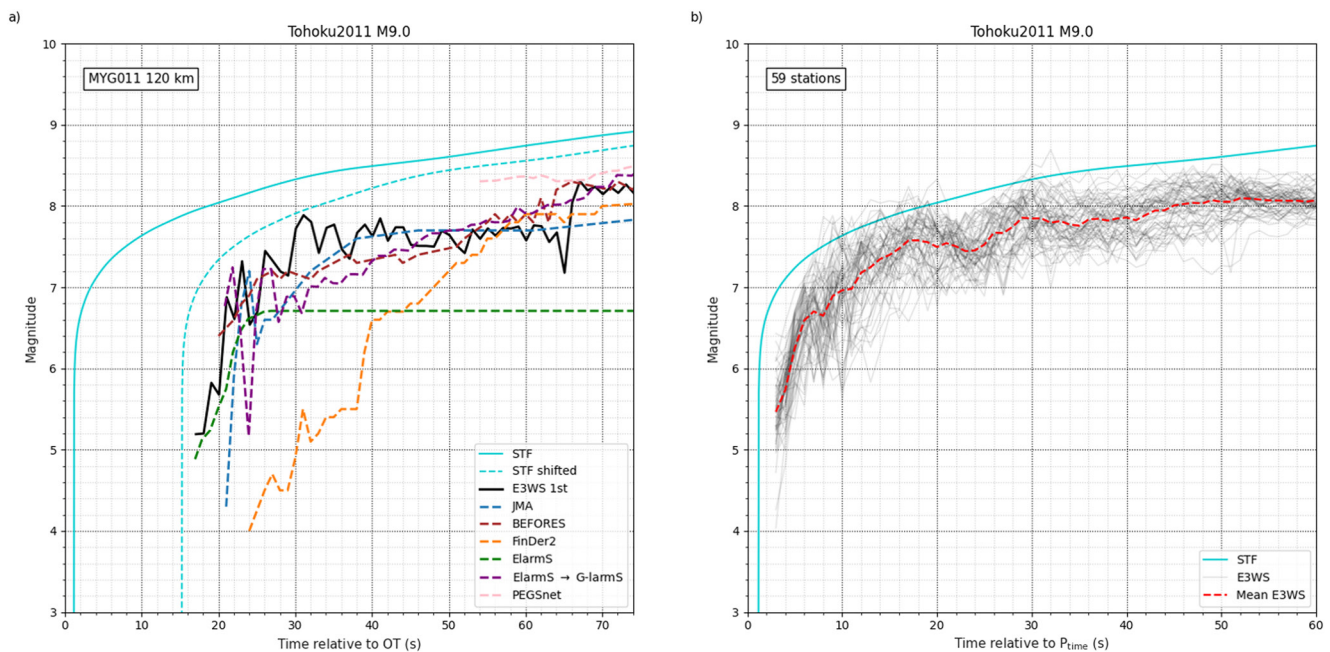


Figure 8. Real-time magnitude estimates for the 2011 M_w 9.0 Tohoku-Oki earthquake. (a) Magnitude evolution estimated by several EEW algorithms (see legend) as a function of time relative to the earthquake origin time. We also show the magnitude from the seismologically determined Source Time Function and after shifting it by the P -wave arrival time at the closest station to the source used by E3WS (name and epicentral distance shown in the top-left corner). (b) Magnitude evolution estimated by E3WS at several stations, as a function of time relative to the P -wave arrival times at each station. We indicate the number of available stations at a maximum of 200 km from the source in the top-left corner.

catalog (Vallée & Douet, 2016). We observe that all the magnitude estimates as a function of time follow the magnitude evolution given by the STF, but with significant underestimation. These underestimations are most likely due to the scarcity of $M_w \geq 8.3$ earthquakes in the training data set, which the system tries to compensate by extrapolating from the magnitudes closest to 9.0 found in our database.

Extrapolation is not required for the Illapel (2015, M_w 8.3), Tokachi (2003, M_w 8.3), Iquique (2014, M_w 8.1), Iquique aftershock (2014, M_w 7.7), Fukushima (2016, M_w 6.6), and Pisco (2007, M_w 8.0) earthquakes, for which the M_w 9.0 Tohoku earthquake observations are part of the training data. For these cases (Figure 9), E3WS estimations track the magnitude evolution in agreement with the STF, with no systematic under-estimation, some even overestimate the STF.

4.4. E3WS in a Real-Time Scenario

We install and test E3WS during one continuous month, with a transmission of data packets every second, at the San Lorenzo (SLN1) station, located in an island offshore Lima, Peru. This station is located at about 130 km from the trench, close to potential seismic sources. The performance of the detector model improves by retraining it with 10 days of noise recorded by the station (overlapping windows sliding by 1 s). We consider noise all seismic data that is not an earthquake with $M > 3.0$. The improvement of the detector is reflected in the decrease of the estimated likelihood that noise traces contain a P phase. The likelihood decreases from a mean of 0.15 with STD of 0.14, to a mean of 0.00017 with an STD of 0.0078, demonstrating the importance of including station-specific noise in the model.

We get 0 false detections and detect 14 earthquakes (Table 4), with mean and STD magnitude errors between the estimated magnitude (M_{est}) and the ground-truth (M_{true} taken from the IGP catalog) of -0.2 and 0.2 , respectively. We compute the detection time as the time at which the system triggers with respect to the P arrival time. E3WS detects earthquakes in less than 1.5 s, on average in 1.0 s. We define the warning time as the difference between the time in which the system computes the source characterization parameters, and the S -arrival time. The system generates an average warning time of 13.5 s with an STD of 4.3 s.

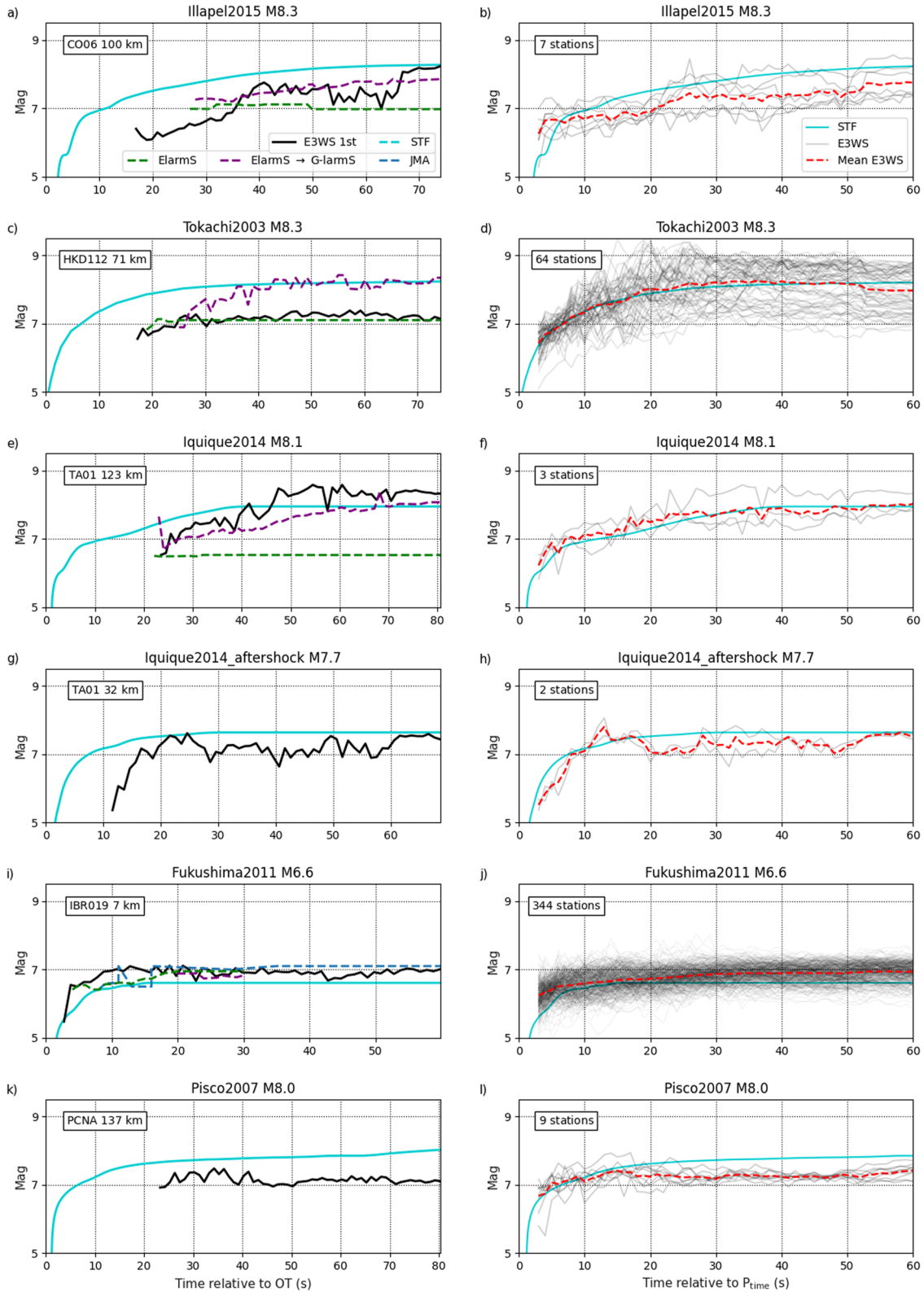


Figure 9.

Table 4
E3WS Earthquake Detections Using 3 s of P-Wave in a Continuous Month (January 2022) at Station SLN1

| M_{true} | Distance (km) | Depth (km) | M_{est} | Detection (s) | Warning time (s) |
|-------------------|---------------|------------|------------------|---------------|------------------|
| 5.6 | 43 | 105 | 5.4 | 0.7 | 13.4 |
| 4.3 | 81 | 47 | 4.2 | 1.4 | 11.1 |
| 4.0 | 61 | 22 | 3.9 | 1.1 | 7.9 |
| 4.0 | 122 | 108 | 3.6 | 1.1 | 17.3 |
| 3.9 | 167 | 128 | 3.5 | 0.5 | 23.6 |
| 3.7 | 85 | 46 | 3.5 | 0.9 | 11.5 |
| 3.6 | 54 | 54 | 3.3 | 1.3 | 9.3 |
| 3.6 | 154 | 56 | 3.4 | 1.1 | 18.5 |
| 3.5 | 150 | 102 | 3.5 | 1.0 | 12.7 |
| 3.5 | 85 | 44 | 3.4 | 0.8 | 11.3 |
| 3.5 | 87 | 22 | 3.5 | 0.7 | 11.8 |
| 3.5 | 148 | 18 | 3.3 | 1.6 | 18.6 |
| 3.1 | 101 | 64 | 3.2 | 1.4 | 13.8 |
| 3.1 | 44 | 55 | 3.2 | 0.9 | 8.5 |

E3WS does not trigger for 15 earthquakes (Table S1 in Supporting Information S1). The maximum magnitude of these missed events is 3.8, with a strong trade-off between magnitude and distance (Table S1 in Supporting Information S1). These magnitudes are below those that generate significant shaking in coastal Peru; they would not warrant an alert. E3WS triggers for three regional earthquakes outside the geographical target area (distances >200 km). The magnitudes of these earthquakes are 4.8, 4.5, and 4.2, with epicentral distances of 321, 396, and 357 km, respectively. As the signals contain low energy level at station SLN1, the magnitude estimates are ~M3.5. We have no false positives corresponding to teleseismic earthquakes (distances >1,000 km), which contain high energy at very low frequencies. This is one of the reasons why we filter between 1 and 7 Hz in our detector.

5. Discussion

5.1. Interpretation of Results

Our results using 3 s of earthquake signals show independence of the epicentral distance and depth estimate errors (STD) with respect to magnitude (Figure S8 in Supporting Information S1). The errors in depth practically remain constant at different magnitudes, while the errors in distance grow from $M > 3$ to $M > 6$ earthquakes but only slightly (by only 4 km). This growth is more related to the lack of $M > 6$ earthquakes compared to $M > 3$ earthquakes, than to magnitude. This means that the shape of the earthquake

seismograms in the 3 s of records is the main feature for estimating distance and depth. This shows consistency with the results of Odaka et al. (2003), who estimate the epicentral distance based on the slope parameter B . This attribute is extracted from the envelope of the acceleration waveform in the first 3 s of the P -wave, through fitting the function $Bt \cdot \exp(-At)$. Moreover, feature importance analysis shows that the most important attributes in the estimation of epicentral distance are: the Mel_0 coefficient (associated to the energy), the centroid in time, and the time in which the seismogram registers the highest amplitude. These attributes describe the signal envelope. Feature importance is described in Section 5.2.

For the vast majority of $M > 6.0$ earthquakes, the average E3WS estimates are above M6 (Figure S7a in Supporting Information S1). This gives us confidence that for most cases where the earthquake is larger than M6, E3WS can estimate within 3 s that the earthquake is larger than M6 (Figure 9). For instance, in the case of the Kahramanmaraş earthquake doublet (Turkey 2023) and using the first 3 s of records from the nearest station from the source, E3WS estimates are M6.6 for the M7.8 mainshock (Figure S12 in Supporting Information S1) and M6.0 for the M7.5 aftershock (Figure S13 in Supporting Information S1), respectively. The purpose of the estimation with 3 s is to provide a first alert, with limited information and a conservative estimate of the shaking intensity, which will be eventually extended to more distant areas if the magnitude estimate increases in later estimations.

For large earthquakes ($M > 6.5$), the errors decrease strongly from 3 s of P -wave (MAE of 0.59) to 7 s (MAE of 0.37), then decrease very smoothly to a MAE of 0.34 in 60 s (Figure S10 in Supporting Information S1). The inflection point is 7 s of P -wave for earthquakes with $M > 6.5$, which corresponds to the characteristic half duration of M6.5 earthquakes (Meier et al., 2017).

5.2. Importance of Different Waveform Attributes in E3WS

We estimate the importance of attributes based on their gain. The gain is the relative contribution of the attribute in each tree in XGB, that is, it is a measure of the improvement in the estimates when using a particular attribute.

Figure 9. Magnitude estimates for the following earthquakes: 2015 M_w 8.3 Illapel, 2003 M_w 8.3 Tokachi-Oki, 2014 M_w 8.1 Iquique, 2014 M_w 7.7 Iquique aftershock, 2011 M_w 6.6 Fukushima aftershock, and 2007 M_w 8.0 Pisco. Estimates are shown as a function of time relative to the earthquake's origin time for the closest station (left, name of station and epicentral distance indicated in the top-left corner) and as a function of time relative to the P -wave arrival time at each station for all seismic stations available (right, number of stations indicated in the top-left corner). On the left, we compare E3WS results (using the closest station to the source) with those obtained by other Earthquake Early Warning Systems. On the right, we show all the estimates (gray), their mean (red), the moment function (the integral of the source time functions, light blue).

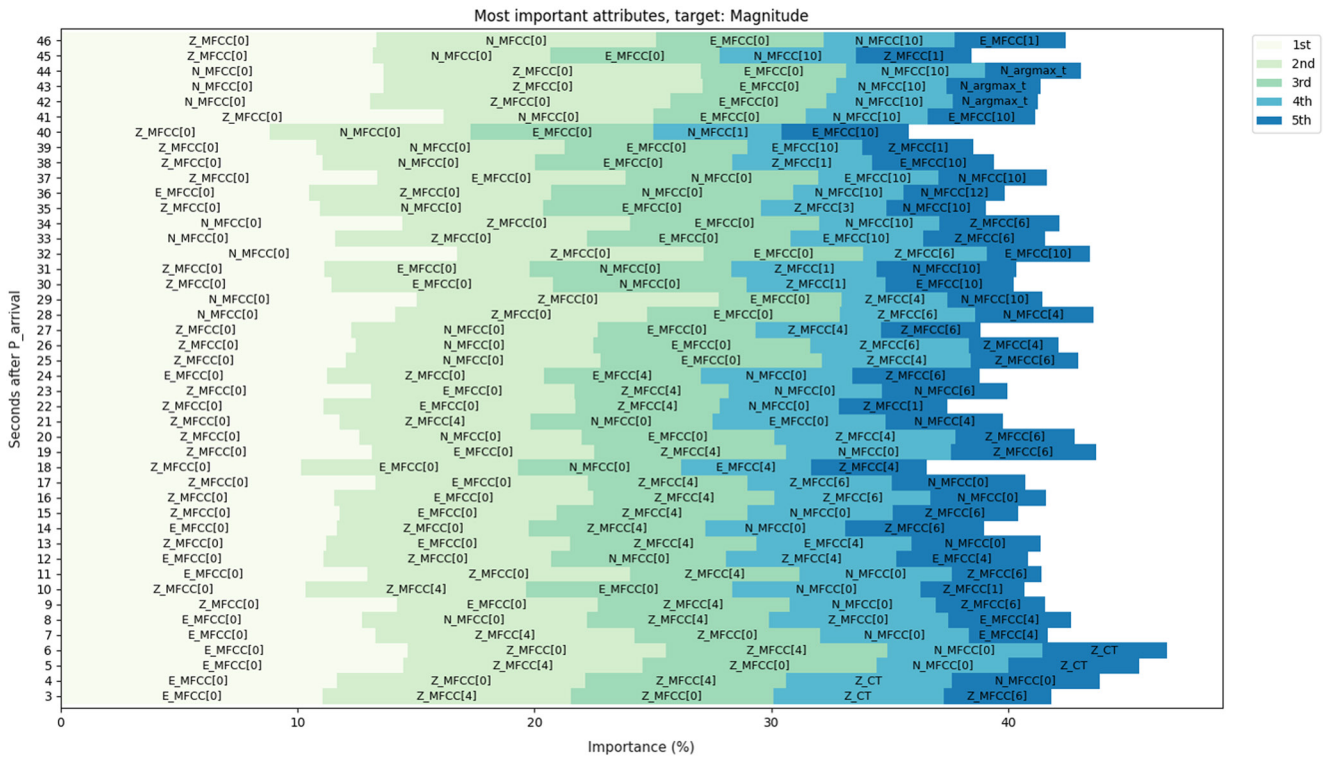


Figure 10. First (lightest color) to fifth (darkest color) most important features for magnitude estimation as a function of the P -wave window duration, from 3 to 46 s. For each time window, feature importance is based on the corresponding stacking model (see Section 3.3), which consists of 10 XGB base models. Importance (%) shown is calculated as the gain mean plus standard deviation of each base model, multiplied by 100 and divided by the total sum. The horizontal axis shows the gain, a measure of attribute importance when making estimates, defined as the relative contribution of the attribute in each tree in XGB. The vertical axis represents the duration of P -wave signal used to train the model. Z, N, and E represent attributes extracted from the vertical, north, and east channel, respectively.

A high gain of an attribute implies that the use of this feature improves the estimates. Our magnitude model is based on the Stacking algorithm, with 10 base-models. For each attribute, we generate the gain for each of the 10-base models trained for 3 s of P -wave signal and calculate the average of the gains and their STD. We order the results of all attributes from highest to lowest value. We repeat the process for longer time windows.

We include MFCCs as part of our attributes because of the relevance compared to time and spectral attributes in our previous work related to classification of volcano-seismic signals (Lara et al., 2020). Once again, MFCCs are shown to be powerful attributes in seismology, because the attributes that contribute the most to magnitude estimation, both using short and long portions of P -wave, are the MFCC (Figure 10). It is striking that cepstral attributes are more relevant than temporal or spectral attributes, such as peak signal energy, frequency centroid and dominant frequency (features 4, 23, and 24 in Section S2 in Supporting Information S1), that share similarities with features that are traditionally used for magnitude estimation in other EEWs, such as P_d or τ_c . MFCC is the discrete cosine transform of the logarithm of the spectrum. Then, the first coefficient, traditionally called the zero coefficient (MFCC[0]), is the logarithmic sum of the energies in the spectrum, which is correlated with the amplitude of the signal. The remaining coefficients give us information about the shape of the spectrum. We hypothesize that the MFCC, manages to capture properties of both signal amplitude and frequency content that are analogous to the traditional attributes P_d and τ_c , which are computed from displacement and velocity waveforms. Their computation from acceleration data requires time integration, which is prone to amplify noise. Since we use a database from strong-motion accelerograms and broadband velocigrams, time integration would lead to amplify the instrumental noise, which is larger in accelerograms and would produce a bias in the estimates.

5.3. Performance of E3WS Compared to Other EEWs

We compare the E3WS with the PreSEIS (Pre-SEISmic) On-site algorithm (Böse et al., 2012), which detects and estimates the magnitude and epicentral distance using the first few seconds of the P -wave recorded by a single

station using neural networks. For a fair comparison, we selected the same number of test samples, the same range of magnitudes (M3.1–M7.6) and epicentral distances (<115 km). E3WS misclassifies events in 0.1% of the test samples, while PreSEIS On-site in 2% using the first 3 s of the earthquake. We compute the uncertainties as the STD of the errors of the observed and estimated values. For the same time window, the E3WS uncertainty in magnitude is 0.41 magnitude units, outperforming PreSEIS On-site with an uncertainty of 0.49. The uncertainties are computed as the STD of the observed and estimated errors. For the epicentral distance, PreSEIS On-site has 18.4 km of uncertainty, while E3WS has 27 km. However, continuous updates do not seem to improve the PreSEIS On-site estimates, while E3WS does improve substantially with longer signals (Figure 6).

We also test the performance of E3WS compared to ElarmS (Brown et al., 2011), which estimates earthquake magnitude within the first 4 s of *P*-wave. To make a fair comparison, we select the same number of earthquake records associated with the same magnitudes within 100 km, as used by R. M. Allen and Kanamori (2003). ElarmS has a MAE of 0.70 magnitude units, while E3WS outperforms it in timeliness and accuracy, with MAE of 0.09 using 3 s of *P*-wave and 0.08 using 4 s. We also compare ElarmS with E3WS on data from the Japanese network. Similarly to R. M. Allen (2007), we select from our database Japanese earthquakes in the magnitude range from 3.8 to 7.4. ElarmS yields a MAE of ~0.75, while E3WS outperforms it again in timeliness and accuracy, with MAE of 0.23 using 3 s of *P*-wave and 0.17 using 4 s.

Finally, we compare the performance of E3WS with UrEDAS. Lockman and Allen (2005) report results applying UrEDAS using stations containing at least five earthquake records, with at least one of the records providing a magnitude estimate of at least M5.0, for earthquakes in southern California. For the best-performing quarter of the stations, with epicentral distances less than 150 km, and using the first 4 s of the earthquake record, UrEDAS achieves a MAE for magnitude estimation of 0.3 magnitude units. For source location, UrEDAS achieves MAEs of 15 km for hypocentral distances and of 20° for back azimuth. We select from our database stations with the same conditions. For the best-performing quarter of the stations and using 3 s, E3WS achieved a MAE of magnitude of 0.22, significantly better than UrEDAS with 4 s. For location, E3WS yields results similar to UrEDAS, with MAE of 14 km for hypocentral distance and 20° for back azimuth. Using 4 s of recording, E3WS achieves MAEs for magnitude, hypocentral distance and back azimuth of 0.20 magnitude units, 13.6 km and 19.1°, respectively.

The back azimuth error is currently the weakest part in E3WS. However, there are opportunities to improve the back azimuth estimates by including new attributes. For instance, Eisermann et al. (2015) combined three methods to estimate back azimuth and obtained an STD of 13°.

5.4. Limitations and Uncertainties

E3WS can operate from 3-component broadband velocimeters or strong-motion accelerometers, with a vertical component and two horizontal components oriented to the east and north, that have no saturation due to clip level. The system monitors $M > 3$ earthquakes with depths down to 100 km and has an application range of 200 km around the station where the E3WS monitors.

Our tests show that the detector model is highly dependent on the intrinsic noise of the station to be monitored, which requires retraining. For the other targets (*P*-phase picking and source characterization) no retraining is necessary. Real-time analysis indicates that 0.21 of trigger threshold for the detector model seems to be sufficient for stations without high noise levels. However, when installed in some locations with high environmental or anthropogenic noise levels, it is necessary to increase the threshold (up to 0.75 or 0.80) to avoid false positives.

Uncertainties in magnitude are between 0.2 and 0.6 magnitude units for the entire magnitude range of our test database (Figure S7a in Supporting Information S1) using the first 3 s of *P*-wave, and 0.35 for the maximum magnitude of our test database (M7.4). For very large earthquakes ($M > 8$), our tests show that all estimates using 3 s of *P*-wave are greater than M6.4, that is, above the alert threshold (M6.0). The only exception in our tests is the Tohoku M9.0 earthquake, for which the first estimate using 3 s of *P*-wave is M5.2, and the estimates exceed M6 (M6.9) using 7 s *P*-wave. This suggests that even with a high accuracy to generate a first alert (e.g., $M > 6$ estimates), it is necessary to complement the estimates with continuous updates using longer windows.

For $M > 6$ earthquakes that do not belong to our training database (earthquakes deeper than 48 km [mean + STD of our depth distribution]), magnitudes are underestimated by E3WS based on 3 s of the *P*-wave. For example, for

the 2023 M6.8 Ecuador earthquake, with hypocentral depth of 68 km, E3WS estimates M6.4 using the first 3 s (Figure S11 in Supporting Information S1). In an operational EEWS, this underestimation might be manageable. For instance, consider a system in which users are alerted within a radius around the epicenter defined by PGA exceeding 2%g (g: gravitational acceleration), where PGA is estimated using ground motion prediction equations based on the estimated source parameters. For the Ecuador earthquake, the E3WS would have set an alert radius of about 211 km, whereas the radius based on the true source parameters should have been 292 km. A tolerance or safety factor of 40%g in the alert radius would be an adequate compromise, given missed alerts are more costly than false alerts. More generally, tolerance factors in the warning system should be evaluated in order to compensate the uncertainties in the E3WS predictions based on the first 3 s of the earthquake.

E3WS has no limitations on the location of use, since we did not find a bias by region: we observed similar behavior in the estimates when testing separately on the Peru, Chile, and global STEAD data sets (Figure S14 in Supporting Information S1). However, we do find a slight bias in the magnitude estimates depending on the type of instrument. For the Peru, Chile, and global STEAD data sets, which contain only broadband velocimeter data, magnitudes are slightly underestimated for $M > 4$ earthquakes. For the Japan database, which contains only strong-motion accelerometer data, there is a slight overestimation in magnitude for $M < 5.5$ earthquakes and a better estimation for $M > 5.5$ earthquakes. We believe that this behavior is due to the fact that the Japanese database contains the largest number of large earthquake observations in our training stage.

5.5. Broad Implications and Future Directions

E3WS is a system based on a single station in order to optimize the warning time to be sent to the population. In the case of using E3WS in a network of stations, it is suggested to use the nearest station, or the first station whose estimates exceed a pre-established threshold, for example, a threshold of M6 to issue the warnings. If the network of stations is highly dense, the average of the estimates of a number of stations can be used to improve the estimates, as in Figure 8b or the images on the right side of Figure 9.

In the case of single-component stations, it is necessary to replicate that component twice. However, back azimuth estimates will not be consistent due to the limitations of single-component stations. E3WS works with broadband velocimeters or strong-motion accelerometers. If it is desired to incorporate E3WS with other types of data, such as GNSS stations or Distributed Acoustic Sensing, it will be necessary to incorporate such data into the database and retrain the entire system.

6. Conclusion

We introduced E3WS, a set of ML algorithms that analyze accelerometric data starting from the first 3 s of *P*-wave signal recorded by a single station to detect, locate and estimate the magnitude of an earthquake. E3WS is made of six independent algorithms performing detection, P-phase picking and estimation of magnitude, epicentral distance, depth, and back azimuth. The proposed system generates faster estimates than existing EEWS. E3WS could provide valuable additional seconds for warning. Although the final magnitude of $M_w \geq 7$ earthquakes cannot be estimated using only 3 s of signal, because their source duration is typically longer than 6 s, the system provides robust detection and preliminary estimations of the instantaneous magnitude and location of an ongoing event, which is valuable to send a first alert. E3WS provides better accuracy than other EEWS that can use one station and 3 s of seismic recording, such as the first version of ElarmS and the pioneering UREDAS. Indeed, our results warrant a revival of single-station methods. Continuous updates of the magnitude and location estimations can be made to update the alert radius as the earthquake grows to larger magnitude. The proposed system is not only theoretical: it is already running in alpha test mode for the EEWS of Peru. It has been installed on low-cost Raspberry Pi 4 devices connected to strong-motion sensors along the Peruvian coast. E3WS is easy to install, flexible to change, can be applied anywhere, and designed using free and open source software (Python3 with the Scikit-learn package) under the Linux operating system.

Data Availability Statement

Waveforms and metadata used in this article were provided by the University of Chile (Barrientos & Team, 2018) downloaded by IRIS Web Services (<https://service.iris.edu/>), NIED K-NET, KiK-net (NIED & KiK-net, 2019), National Research Institute for Earth Science and Disaster Resilience (<https://doi.org/10.17598/NIED.0004>), the

Stanford Earthquake Data set (Mousavi et al., 2019) available at <https://github.com/smousavi05/STEAD>, and the Peruvian data set through petition to the Instituto Geofísico del Perú (<https://www.gob.pe/igp>).

E3WS is public and available at <https://github.com/PabloELara/E3WS>. The release version v1.0 used in this article is available at <https://doi.org/10.5281/zenodo.7996838> (Lara, 2023). This system can be used as an EEWS that monitors seismic activity of $M > 3$ earthquakes around 200 km from the station to be installed.

Acknowledgments

This work has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Grant Agreement 949221). This work was granted access to the HPC resources of IDRIS under the allocations 2020-AD011012142, 2021-AP011012536, and 2021-A0101012314 made by GENCI. This work has been supported by the French government, through the UCAJEDI Investments in the Future project managed by the National Research Agency (ANR) ANR-15-IDEX-01. The Institut de recherche pour le développement (IRD) and the Instituto Geofísico del Perú (IGP).

References

- Allen, R. M. (2007). The ElarmS earthquake early warning methodology and application across California. *Earthquake Early Warning Systems*, 21–43. https://doi.org/10.1007/978-3-540-72241-0_3
- Allen, R. M., & Kanamori, H. (2003). The potential for earthquake early warning in southern California. *Science*, 300(5620), 786–789. <https://doi.org/10.1126/science.1080912>
- Allen, R. M., & Melgar, D. (2019). Earthquake early warning: Advances, scientific challenges, and societal needs. *Annual Review of Earth and Planetary Sciences*, 47(1), 361–388. <https://doi.org/10.1146/annurev-earth-053018-060457>
- Allen, R. V. (1978). Automatic earthquake recognition and timing from single traces. *Bulletin of the Seismological Society of America*, 68(5), 1521–1532. <https://doi.org/10.1785/bssa0680051521>
- Aoi, S., Kunugi, T., & Fujiwara, H. (2004). Strong-motion seismograph network operated by NIED: K-net and KiK-net. *Journal of Japan Association for Earthquake Engineering*, 4(3), 65–74. https://doi.org/10.5610/jae.4.3_65
- Barrientos, S., & Team, N. S. C. C. (2018). The seismic network of Chile. *Seismological Research Letters*, 89(2A), 467–474. <https://doi.org/10.1785/0220160195>
- Böse, M., Allen, R., Brown, H., Gua, G., Fischer, M., Hauksson, E., et al. (2014). Cism Shakealert: An earthquake early warning demonstration system for California. In *Early warning for geological disasters* (pp. 49–69). Springer.
- Böse, M., Hauksson, E., Solanki, K., Kanamori, H., Wu, Y.-M., & Heaton, T. (2009). A new trigger criterion for improved real-time performance of onsite earthquake early warning in southern California. *Bulletin of the Seismological Society of America*, 99(2A), 897–905. <https://doi.org/10.1785/0120080034>
- Böse, M., Heaton, T., & Hauksson, E. (2012). Rapid estimation of earthquake source and ground-motion parameters for earthquake early warning using data from a single three-component broadband or strong-motion sensor. *Bulletin of the Seismological Society of America*, 102(2), 738–750. <https://doi.org/10.1785/0120110152>
- Böse, M., Smith, D. E., Felizardo, C., Meier, M.-A., Heaton, T. H., & Clinton, J. F. (2018). Finder v. 2: Improved real-time ground-motion predictions for M2–M9 with seismic finite-source characterization. *Geophysical Journal International*, 212(1), 725–742. <https://doi.org/10.1093/gji/ggx430>
- Bossu, R., Finazzi, F., Steed, R., Fallou, L., & Bondár, I. (2022). “shaking in 5 seconds!”—Performance and user appreciation assessment of the earthquake network smartphone-based public earthquake early warning system. *Seismological Research Letters*, 93(1), 137–148. <https://doi.org/10.1785/0220210180>
- Brown, H. M., Allen, R. M., Hellweg, M., Khainovski, O., Neuhauser, D., & Souf, A. (2011). Development of the ElarmS methodology for earthquake early warning: Realtime application in California and offline testing in Japan. *Soil Dynamics and Earthquake Engineering*, 31(2), 188–200. <https://doi.org/10.1016/j.soildyn.2010.03.008>
- Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785–794).
- Chung, A. I., Henson, I., & Allen, R. M. (2019). Optimizing earthquake early warning performance: ElarmS-3. *Seismological Research Letters*, 90(2A), 727–743. <https://doi.org/10.1785/0220180192>
- Cremon, G., Galasso, C., & Zuccolo, E. (2022). Investigating the potential effectiveness of earthquake early warning across Europe. *Nature Communications*, 13(1), 1–10. <https://doi.org/10.1038/s41467-021-27807-2>
- Cui, S., Yin, Y., Wang, D., Li, Z., & Wang, Y. (2021). A stacking-based ensemble learning method for earthquake casualty prediction. *Applied Soft Computing*, 101, 107038. <https://doi.org/10.1016/j.asoc.2020.107038>
- Davis, S., & Mermelstein, P. (1980). Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *IEEE Transactions on Acoustics, Speech, & Signal Processing*, 28(4), 357–366. <https://doi.org/10.1109/tassp.1980.1163420>
- Eisermann, A. S., Ziv, A., & Wust-Bloch, G. H. (2015). Real-time back azimuth for earthquake early warning. *Bulletin of the Seismological Society of America*, 105(4), 2274–2285. <https://doi.org/10.1785/0120140298>
- Grapenthin, R., Johanson, I., & Allen, R. M. (2014a). The 2014 mw 6.0 Napa earthquake, California: Observations from real-time GPS-enhanced earthquake early warning. *Geophysical Research Letters*, 41(23), 8269–8276. <https://doi.org/10.1002/2014gl061923>
- Grapenthin, R., Johanson, I. A., & Allen, R. M. (2014b). Operational real-time GPS-enhanced earthquake early warning. *Journal of Geophysical Research: Solid Earth*, 119(10), 7944–7965. <https://doi.org/10.1002/2014jb011400>
- Hoshiba, M., & Ozaki, T. (2014). Earthquake early warning and tsunami warning of the Japan meteorological agency, and their performance in the 2011 off the Pacific coast of Tohoku earthquake (Mw 9.0). In *Early warning for geological disasters* (pp. 1–28). Springer.
- Kanamori, H. (2005). Real-time seismology and earthquake damage mitigation. *Annual Review of Earth and Planetary Sciences*, 33(1), 195–214. <https://doi.org/10.1146/annurev.earth.33.092203.122626>
- Kukreja, S. L., Löfberg, J., & Brenner, M. J. (2006). A least absolute shrinkage and selection operator (LASSO) for nonlinear system identification. *IFAC Proceedings Volumes*, 39(1), 814–819. <https://doi.org/10.3182/20060329-3-au-2901.00128>
- Lara, P. (2023). Pabloelara/e3ws: E3ws [Software]. Zenodo. <https://doi.org/10.5281/zenodo.7996838>
- Lara, P., Fernandes, C. A. R., Inza, A., Mars, J. I., Métaixian, J.-P., Dalla Mura, M., & Malfante, M. (2020). Automatic multichannel volcano-seismic classification using machine learning and EMD. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13, 1322–1331. <https://doi.org/10.1109/jstars.2020.2982714>
- Licciardi, A., Bletery, Q., Rouet-Leduc, B., Ampuero, J.-P., & Juhel, K. (2022). Instantaneous tracking of earthquake growth with elastogravity signals. *Nature*, 606(7913), 1–6. <https://doi.org/10.1038/s41586-022-04672-7>
- Lockman, A. B., & Allen, R. M. (2005). Single-station earthquake characterization for early warning. *Bulletin of the Seismological Society of America*, 95(6), 2029–2039. <https://doi.org/10.1785/0120040241>
- Malfante, M., Dalla Mura, M., Métaixian, J.-P., Mars, J. I., Macedo, O., & Inza, A. (2018). Machine learning for volcano-seismic signals: Challenges and perspectives. *IEEE Signal Processing Magazine*, 35(2), 20–30. <https://doi.org/10.1109/msp.2017.2779166>

- Meier, M.-A. (2017). How “good” are real-time ground motion predictions from earthquake early warning systems? *Journal of Geophysical Research: Solid Earth*, 122(7), 5561–5577. <https://doi.org/10.1002/2017jb014025>
- Meier, M.-A., Ampuero, J., & Heaton, T. H. (2017). The hidden simplicity of subduction megathrust earthquakes. *Science*, 357(6357), 1277–1281. <https://doi.org/10.1126/science.aan5643>
- Meier, M.-A., Ross, Z. E., Ramachandran, A., Balakrishna, A., Nair, S., Kundzic, P., et al. (2019). Reliable real-time seismic signal/noise discrimination with machine learning. *Journal of Geophysical Research: Solid Earth*, 124(1), 788–800. <https://doi.org/10.1029/2018jb016661>
- Melgar, D., & Hayes, G. (2019). Characterizing large earthquakes before rupture is complete. *Science Advances*, 5(5), eaav2032. <https://doi.org/10.1126/sciadv.aav2032>
- Minson, S. E., Murray, J. R., Langbein, J. O., & Gomberg, J. S. (2014). Real-time inversions for finite fault slip models and rupture geometry based on high-rate GPS data. *Journal of Geophysical Research: Solid Earth*, 119(4), 3201–3231. <https://doi.org/10.1002/2013jb010622>
- Mousavi, S. M., & Beroza, G. C. (2019). Bayesian-deep-learning estimation of earthquake location from single-station observations. arXiv preprint arXiv:1912.01144.
- Mousavi, S. M., & Beroza, G. C. (2020). A machine-learning approach for earthquake magnitude estimation. *Geophysical Research Letters*, 47(1), e2019GL085976. <https://doi.org/10.1029/2019gl085976>
- 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. *Nature Communications*, 11(1), 1–12. <https://doi.org/10.1038/s41467-020-17591-w>
- Mousavi, S. M., Sheng, Y., Zhu, W., & Beroza, G. C. (2019). Stanford earthquake data set (STEAD): A global data set of seismic signals for ai. *IEEE Access*, 7, 179464–179476. <https://doi.org/10.1109/access.2019.2947848>
- Münchmeyer, J., Bindl, D., Leser, U., & Tilmann, F. (2021). Earthquake magnitude and location estimation from real time seismic waveforms with a transformer network. *Geophysical Journal International*, 226(2), 1086–1104. <https://doi.org/10.1093/gji/ggab139>
- Nakamura, Y. (1988). On the urgent earthquake detection and alarm system (UREDAS). In *Proc. of the 9th world conference on earthquake engineering* (Vol. 7, pp. 673–678).
- Nakamura, Y., & Saita, J. (2007). UREDAS, the earthquake warning system: Today and tomorrow. *Earthquake Early Warning Systems*, 249–281. https://doi.org/10.1007/978-3-540-72241-0_13
- Nakamura, Y., Saita, J., & Sato, T. (2011). On an earthquake early warning system (EEW) and its applications. *Soil Dynamics and Earthquake Engineering*, 31(2), 127–136. <https://doi.org/10.1016/j.soildyn.2010.04.012>
- Nied, K., & KiK-net, N. (2019). National research institute for earth science and disaster resilience [Dataset]. National Research Institute of Earth Science And Disaster Resilience. <https://doi.org/10.17598/NIED.0004>
- Nielsen, D. (2016). *Tree boosting with xgboost-why does xgboost win “every” machine learning competition?* (Unpublished master's thesis). NTNU.
- Ochoa, L. H., Niño, L. F., & Vargas, C. A. (2018). Fast magnitude determination using a single seismological station record implementing machine learning techniques. *Geodesy and Geodynamics*, 9(1), 34–41. <https://doi.org/10.1016/j.geog.2017.03.010>
- Odaka, T., Ashiya, K., Tsukada, S., Sato, S., Ohtake, K., & Nozaka, D. (2003). A new method of quickly estimating epicentral distance and magnitude from a single seismic record. *Bulletin of the Seismological Society of America*, 93(1), 526–532. <https://doi.org/10.1785/0120020008>
- Pal, S. C., Saha, A., Chowdhuri, I., Ruidas, D., Chakraborty, R., Roy, P., & Shit, M. (2023). Earthquake hotspot and coldspot: Where, why and how? *Geosystems and Geoenvironment*, 2(1), 100130. <https://doi.org/10.1016/j.geogeo.2022.100130>
- Ruhl, C., Melgar, D., Chung, A., Grapenthin, R., & Allen, R. (2019). Quantifying the value of real-time geodetic constraints for earthquake early warning using a global seismic and geodetic data set. *Journal of Geophysical Research: Solid Earth*, 124(4), 3819–3837. <https://doi.org/10.1029/2018jb016935>
- Shokouhi, P., Girkar, V., Rivière, J., Shreedharan, S., Marone, C., Giles, C. L., & Kifer, D. (2021). Deep learning can predict laboratory quakes from active source seismic data. *Geophysical Research Letters*, 48(12), e2021GL093187. <https://doi.org/10.1029/2021gl093187>
- Vallée, M., & Douet, V. (2016). A new database of source time functions (STFS) extracted from the SCARDEC method. *Physics of the Earth and Planetary Interiors*, 257, 149–157. <https://doi.org/10.1016/j.pepi.2016.05.012>
- Yih-Min, W., Kanamori, H., Allen, R. M., & Hauksson, E. (2007). Determination of earthquake early warning parameters, τ_c and p_d , for southern California. *Geophysical Journal International*, 170(2), 711–717. <https://doi.org/10.1111/j.1365-246x.2007.03430.x>

References From the Supporting Information

- Grigoli, F., Cesca, S., Amoroso, O., Emolo, A., Zollo, A., & Dahm, T. (2014). Automated seismic event location by waveform coherence analysis. *Geophysical Journal International*, 196(3), 1742–1753. <https://doi.org/10.1093/gji/ggt477>
- Kopparapu, S. K., & Laxminarayana, M. (2010). Choice of Mel filter bank in computing MFCC of a resampled speech. In *10th international conference on information science, signal processing and their applications (ISSPA 2010)* (pp. 121–124).