Earthquake Early Warning using 3 seconds of records on a single station

Pablo Lara[1], Quentin Bletery[1], Jean-Paul Ampuero[2], and Adolfo Inza[3]

[1]Université Côte d'Azur, IRD, CNRS, Observatoire de la Côte d'Azur, Geoazur
[2]Institut de Recherche pour le Développement
[3]Instituto Geofísico del Perú

February 27, 2023

Abstract

We introduce the Ensemble Earthquake Early Warning System (E3WS), a set of Machine Learning algorithms designed to detect, locate and estimate the magnitude of an earthquake using 3 seconds of P waves recorded by a single station. The system is made of 6 Ensemble Machine Learning algorithms trained on attributes computed from ground acceleration time series in the temporal, spectral and cepstral domains. The training set comprises datasets from Peru, Chile, Japan, and the STEAD global dataset. 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 [?] 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.

Earthquake Early Warning using 3 seconds of records on a single station

Pablo Lara[1,2], Quentin Bletery[1], Jean-Paul Ampuero[1], and Adolfo Inza[2]

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

Corresponding author: Pablo Lara, `pablo.elara@ieee.org`

Abstract

We introduce the Ensemble Earthquake Early Warning System (E3WS), a set of Machine Learning algorithms designed to detect, locate and estimate the magnitude of an earthquake using 3 seconds of P waves recorded by a single station. The system is made of 6 Ensemble Machine Learning algorithms trained on attributes computed from ground acceleration time series in the temporal, spectral and cepstral domains. The training set comprises datasets from Peru, Chile, Japan, and the STEAD global dataset. 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.

1 Introduction

Today millions of people live at risk from earthquakes. Real-time seismic monitoring near seismic sources opens the possibility of rapidly estimating earthquake parameters that control the potential impact of shaking, notably magnitude and hypocenter location. Taking advantage of such estimates and speed-of-light communications, Earthquake Early Warning Systems (EEWS) can generate an alarm before seismic shaking reaches the population, 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 warning time to take protective measures. In a common approach, the magnitude and location of the earthquake are rapidly estimated, then used as input in a Ground Motion Prediction Equation to forecast ground shaking intensity measures such as Peak Ground Acceleration, 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. 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 focus on single-station 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 and Global Navigation Satellite System (GNSS) stations (R. M. Allen & Melgar, 2019).

The τ_c-P_d Onsite algorithm (Böse et al., 2009), one of the three algorithms that contributed to the development of ShakeAlert, the EEWS 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 seconds 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, originally developed in Japan. 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).

Recently, Artificial Intelligence (AI) has been used in a number of applications in seismology that are relevant for EEW. In particular, studies not designed for EEW purposes used AI for specific targets, such as detection, picking and source characterization components. The Earthquake Transformer algorithm (EQTransformer) (Mousavi et al., 2020) uses 1 minute long seismograms to feed an architecture based on Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) to detect earthquakes, and estimate P and S phase arrivals. The model achieves an earthquake detection precision (true positives divided by total positives) of 1.0, and estimates the P and S phase arrivals with a mean of 0 seconds and standard deviation (STD) of 0.03 seconds for the P phase, and mean of 0.0 seconds and STD of 0.11 seconds for the S phase. Mousavi and Beroza (2020) estimate magnitudes with CNN and RNN trained with 30 seconds long seismograms ($M<5.7$). They obtain a mean error close to 0 magnitude units and standard deviation $\sim$0.2. Mousavi and Beroza (2019) estimate earthquake location based on Bayesian Deep Learning. The network is fed by 1 minute long seismograms in the case of training distances. For back-azimuth, the angle is represented as points of the unit circle (cosine, sine) and is trained using seismograms of 1.5 s long. The network achieves a mean localization error of 7.27 km with STD of 12.16 km. While these AI models are not designed for EEW (they use signal windows that are too long), 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 seconds of signal recorded by a single three-component sensor. E3WS is a system focused on early warning for populations living near seismic sources. 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. Its magnitude residuals are small enough to not generate false warnings, for both overestimates and underestimates. 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 Machine Learning, allowing (in contrast to Deep Learning approaches) some understanding of what controls the estimations. Thanks to its simplicity, E3WS can be installed in small 32-bits single board computers, such as a Raspberry Pi, as well as in more complex 64-bits processors.

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). 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 73,000 earthquake seismograms. Data statistics are shown in Fig. 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. We ignore sensors for which we did not find the instrument response.

The preprocessing also includes semi-automatic detection of the P-phase for the Chilean and Japanese datasets. Since the Chilean, KiK-net and K-NET datasets do not include P-phase picks, we used an STA/LTA algorithm (R. M. Allen, 2007) and theoretical travel times to estimate the P-phase and S-phase arrival times. Sometimes we find in the raw data of a given earthquake more than one earthquake. In these cases, the STA/LTA algorithm triggered in most of the earthquakes. To avoid selecting an earthquake that does not belong to the dataset catalog, we select the earthquake in which the difference in the estimated time between the STA/LTA algorithm and theoretical travel times is not greater than three seconds. Then, we chose the STA/LTA trigger time as the P-phase time.

For the Peru, Chile and STEAD datasets, we manually identified and removed signals that were saturated for M>5.0. We found no further saturated waveforms when inspecting events down to M4.5. We believe that for magnitudes smaller than 4.5, there are no more saturated waveforms. The Japanese data are recorded by strong-motion accelerometers, which do not saturate.

3 Proposed EEWS

E3WS consists of 6 ML algorithms : a detector, a P-phase picker and 4 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 seconds of P-wave signal, the 4 regression models run in parallel 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 6 models take as input a feature vector formed by concatenating 140 attributes extracted from the waveforms, their spectrum and their cepstrum.

We test several approaches to design the models, including Extreme Gradient Boosting (XGB), Support Vector Machine (SVM), Random Forest (RF) and Neural Networks (Multilayer Perceptron: MLP). We find that the approach yielding the best results is XGB (Chen & Guestrin, 2016), a Supervised Machine Learning 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 seconds of the P-wave. 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

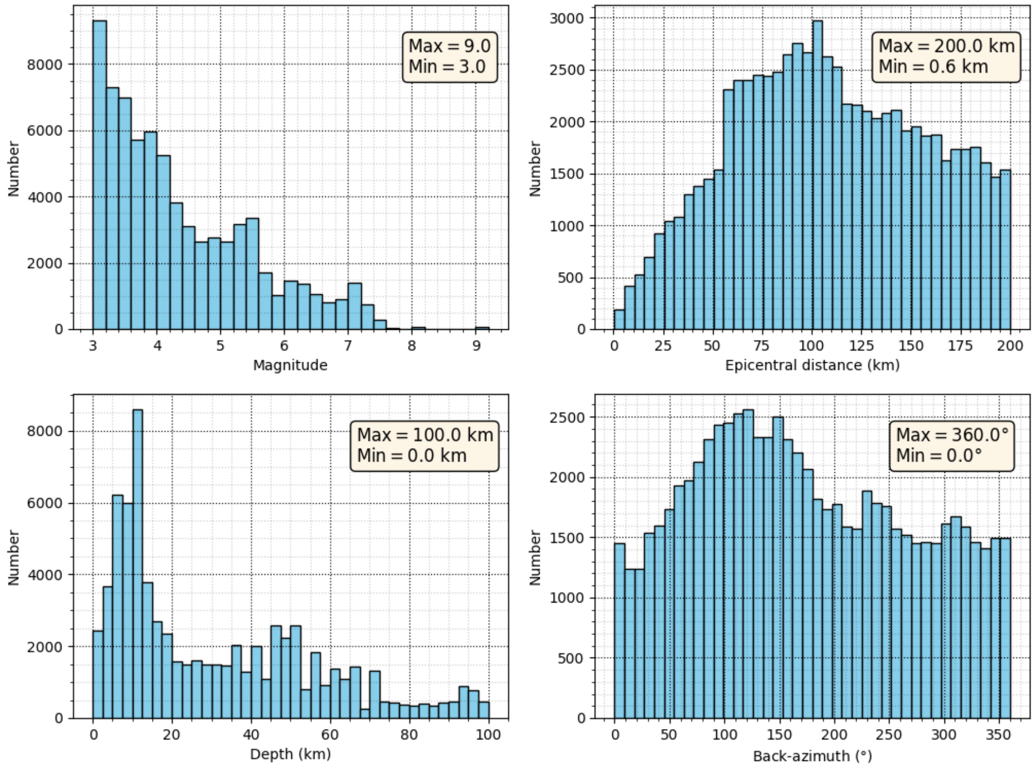

Figure 1. Magnitude, epicentral distance, depth and back-azimuth distributions of the earthquake waveform database compiled for this work.

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, on an Intel(R) Xeon(R) Silver 4114 processor. Even on a Single Board Computer like a Raspberry Pi 4, XGB takes 0.01 s on average. 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. Although XGB only supports single-label training, it performs better than inherently multi-output regression models such as RF.

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 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. Fig. S1 shows the general XGB scheme. For all models, we use the following hyperparameters for XGB training: depth = 4, number of trees = 6000, subset = 80% and learning rate = 0.1.

3.1 Detection

Since STEAD is a global dataset that also includes global noise samples, we extract 55,000 noise windows and add them to our database. We use waveforms filtered from 1 to 7 Hz and a fixed time window of 10 seconds including a minimum of 3 seconds of P-wave. We discard waveforms that do not contain 7 seconds of data before the P-

Table 1. Accuracy and errors using XGB, RF, SVM and MLP models for detection, P-phase picking and source characterization.

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

wave arrival. In our tests we obtained better accuracy using a 10-seconds-long window compared to shorter windows. For instance, we find false detections due to impulsive noise using shorter windows; 10 seconds-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 seconds of P or S-wave, respectively. 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.

We estimate the likelihood that a window contains a P wave, sliding the 10-second window by steps of 0.5 seconds. 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.

3.2 P-phase picking

Because the Japanese and Chilean datasets do not provide P-phase arrival times (t_p), we restrict the training set for phase-picking to the Peruvian and STEAD datasets. We use a fixed time window of 4 seconds filtered between 1 and 7 Hz.

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 (Fig. 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 $[t_p{-}3.5s, t_p{+}0.5s]$ 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 extracted every 0.01 s, the

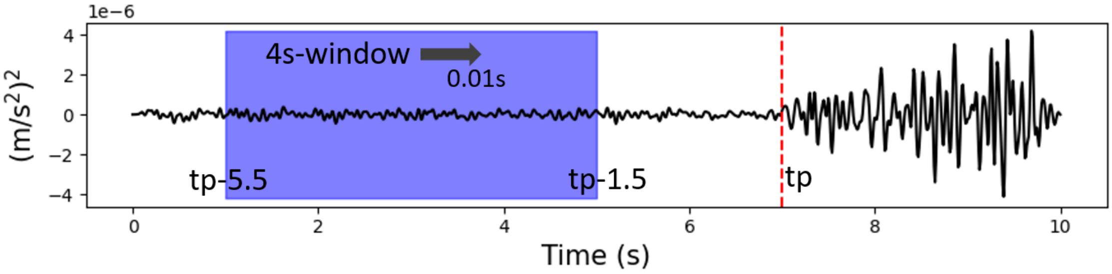

Figure 2. Labeling for the P-phase picking model. We extract attributes from a 4-seconds-long window, starting from $t_p - 5.5$ s as the blue box, for our entire database. We repeat the attribute extraction every 0.01 s until the blue box reaches $t_p + 2.5$ s.

input dataset for the P-phase picking model is made of approximately 36 million 140-dimensional samples.

3.3 Source characterization

For source characterization (esimation of earthquake magnitude, epicentral distance, back-azimuth and hypocentral depth), we use time windows that contain 7 seconds of noise and 3 seconds of P-wave signal extracted from our earthquake database. We apply a 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 datasets 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 (Fig. 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 dataset into K=10 groups and training each XGB base-model with nine out of the ten 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 ($\mathrm{OOF}_{\mathrm{pred}}$) to train the LASSO meta-model (Kukreja et al., 2006).

For the back-azimuth model, we divide the training into 2 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 domains. For the time domain, we extract attributes from the preprocessed signal s and from its envelope, defined as the absolute value of its analytic sig-

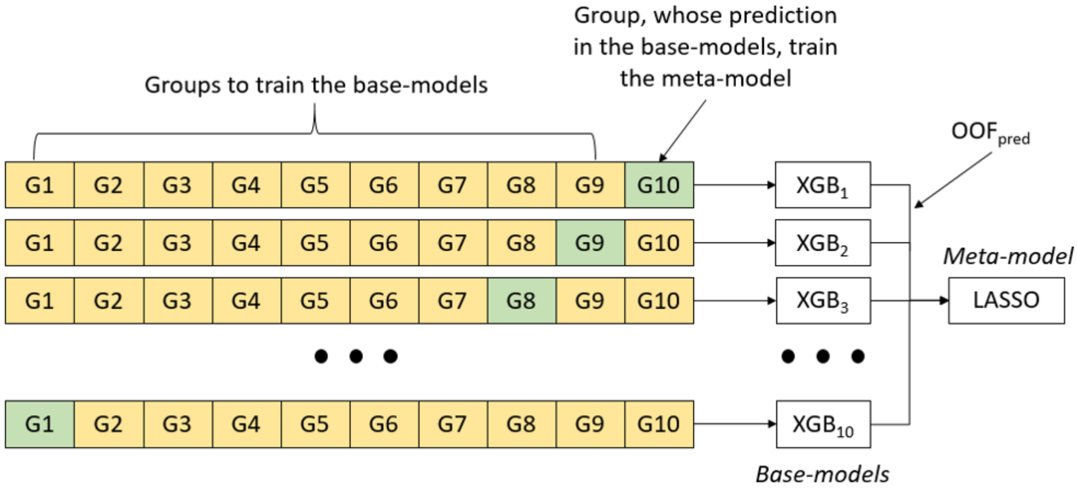

Figure 3. Source characterization model based on Stacking algorithm and K-Fold with K=10. For each K, nine groups train the XGB base-model. Estimates from the remaining group, using the corresponding trained XGB model, feed the LASSO meta-model.

nal $|s+iH\{s\}|$ where H is the Hilbert transform. For the spectral domain, we consider the Power Spectral Density (PSD) 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).

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. 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 seconds 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 seconds 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 (Fig. S2), 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 $M < 4.3$ (Fig. S3) and low signal-to-noise ratios (Fig. S4). Most of them have an epicentral distance greater than 100 km.

Table 2. Confusion matrix for the detection algorithm.

Overall (%):		**True class**	
99.9		Noise	Earthquake
Estimated	Noise	**11264**	10
class	Earthquake	0	**8778**
Accuracy (%):		100.0	99.9

4.1.2 P-phase picking

We evaluate the picker model on more than 10,000 seismograms of the test dataset compared to the "true" (manually picked) P-wave arrival times (Fig. S5). 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 (Figs. 5, S6, Table 3), given that our algorithm only uses 3 seconds of records on a single station.

Table 3. Performance of the source characterization algorithm with its mean error, STD error, MAE and coefficient of determination (R2).

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 (°)	-3.4	43.7	45.2	0.84

The magnitude estimates are very stable for earthquakes with magnitudes between 3.2 and 6.5, with magnitude average residuals ($|M_{pred} - M_{true}|$) of ∼0.2 for M<5.7 (Fig. 4a), and residuals between 0 and 0.4 for 5.7<M<6.5. We even observe magnitude residuals ∼0 for M6.2. This gives us confidence in estimating magnitudes for minor (M3.0-M3.9) to strong (M6.0-M6.9) earthquakes. For instance, for all M>6.0 earthquakes the average estimates are M¿6.0, so there would not be missed events in a EEW system that uses a threshold M>6.0 as a primary alert criterion. The small errors over the entire range of magnitudes are reflected in a high R2 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 (Fig. S7a), 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 seconds window duration.

We observe an average residuals at epicentral distances for distances very close to the seismic source (0-20 km) of $\sim$28 km (Fig. 4b). As the seismic gets farther away up to a distance of $\sim$120 km, the residuals decrease linearly even down to almost 0 error. From here, the errors grow linearly up to our training distance limit (200 km). If we keep our error tolerance at 28 km (errors at very close distances), we can estimate up to an epicentral distance of 165 km. Longer distances to this implies greater 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 smaller than 10 km (Fig. 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 back-azimuth, residuals exceed 35°. However, the STD of the estimates decreases significantly as the magnitude increases, achieving an STD of 21° for M>6.0 earthquakes (Fig. S8). The estimates have uniform performance throughout their range (Fig. S7d). The high R2 of 0.84 shows that the model contains relevant information in the whole back-azimuth range.

4.2 Performance of source characterization using longer signals

Meier et al. (2017) showed that the source time functions (STF) of large and very large 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 seconds of the P wave, i.e. the magnitude reached by the earthquake 3 seconds after its onset. This estimate can form the basis to generate a first warning and can be updated when longer records become available.

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 s to 46 s. Fig. 6 shows the evolution of two performance metrics, MAE and R2, as a function of the considered signal duration.

We observe a significant improvement in the estimations of magnitude and epicentral distance, with R2 scores increasing up to 0.94 and 0.93, respectively, and MAE dropping to about 0.25 and 9 km, respectively, at 46 s of signal (Fig. 7). 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 seems sufficient to converge to the best performance (Fig. 6a-d).

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

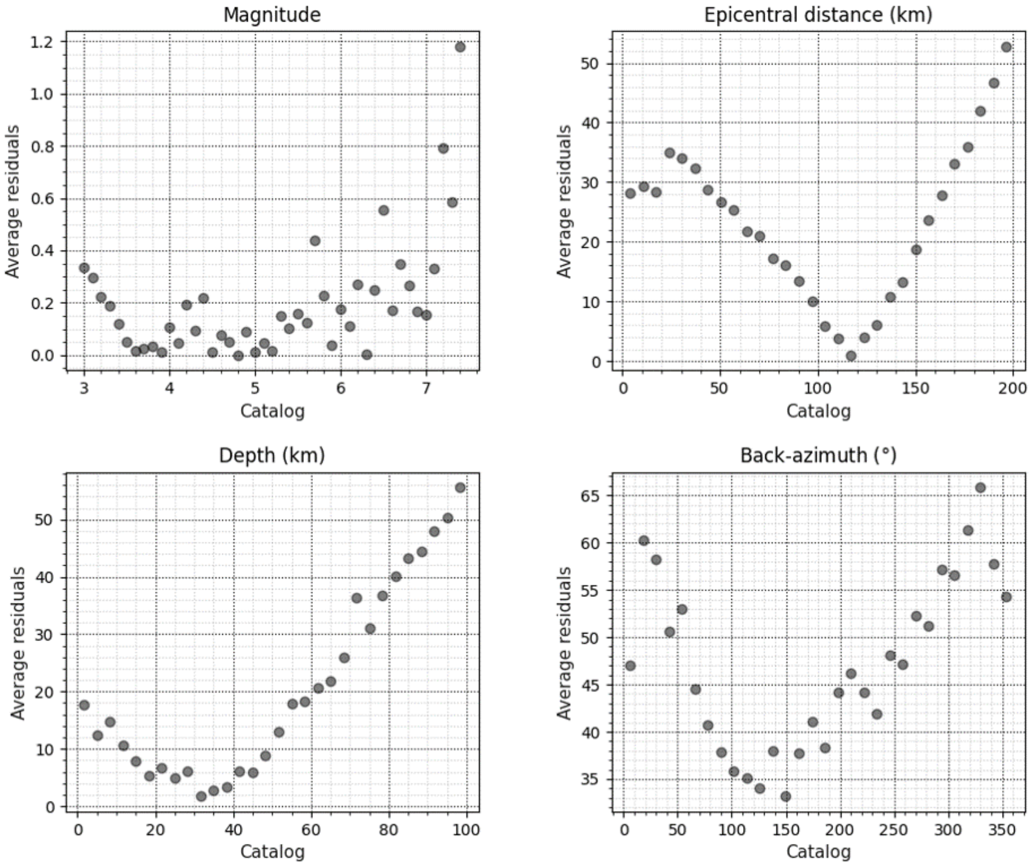

Figure 4. Average residuals ($|target_{pred} - target_{true}|$) for each target: Magnitude, epicentral distance, depth and back-azimuth, using the first 3 seconds of P-wave.

For the back-azimuth estimation, the best model uses 5 seconds 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.

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 dataset 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 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) (Hoshiba & Ozaki, 2014) and PEGSNet (Licciardi et al., 2022), and GNSS stations such as BEFORES (Minson et al., 2014) and G-larmS (Grapenthin et al., 2014b,

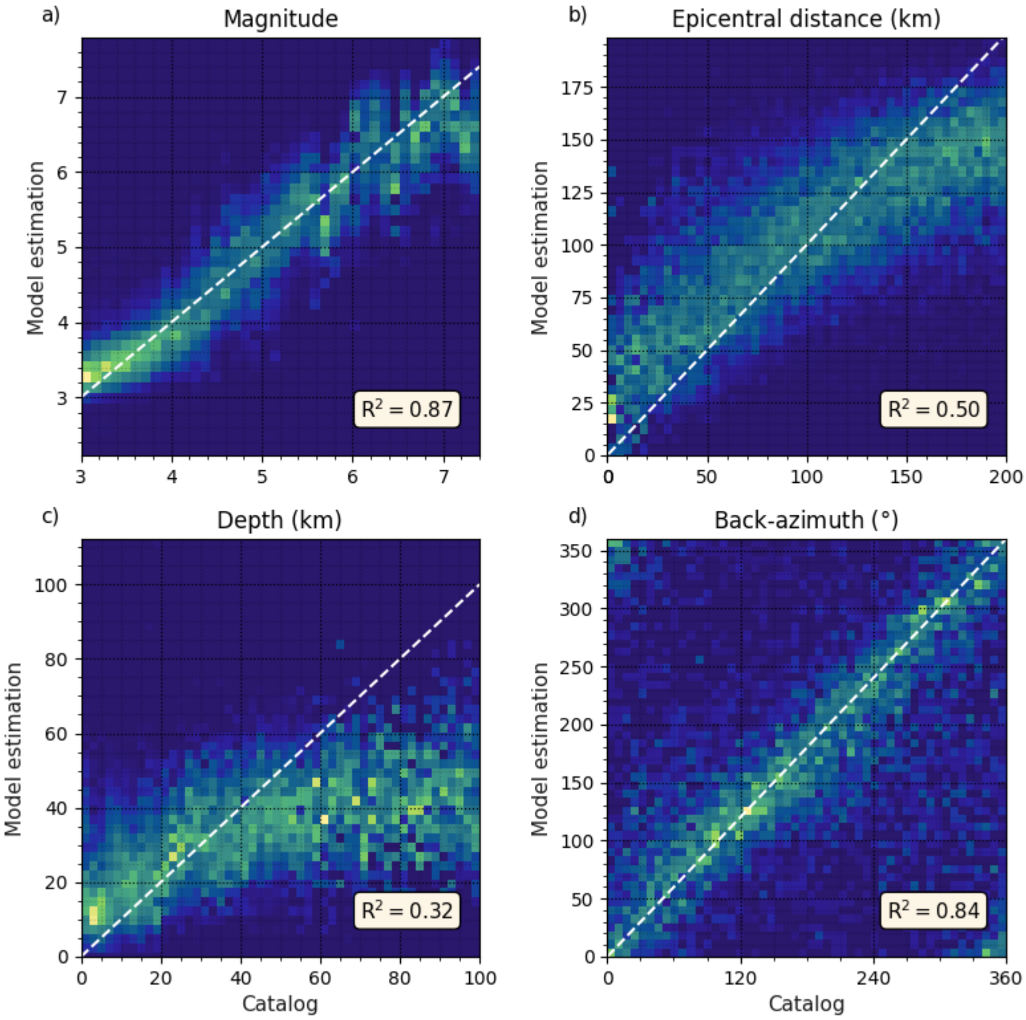

Figure 5. Estimated source parameters (magnitude, distance, depth, back-azimuth) using 3 s of records as a function of cataloged values.

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

Fig. 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). BE-

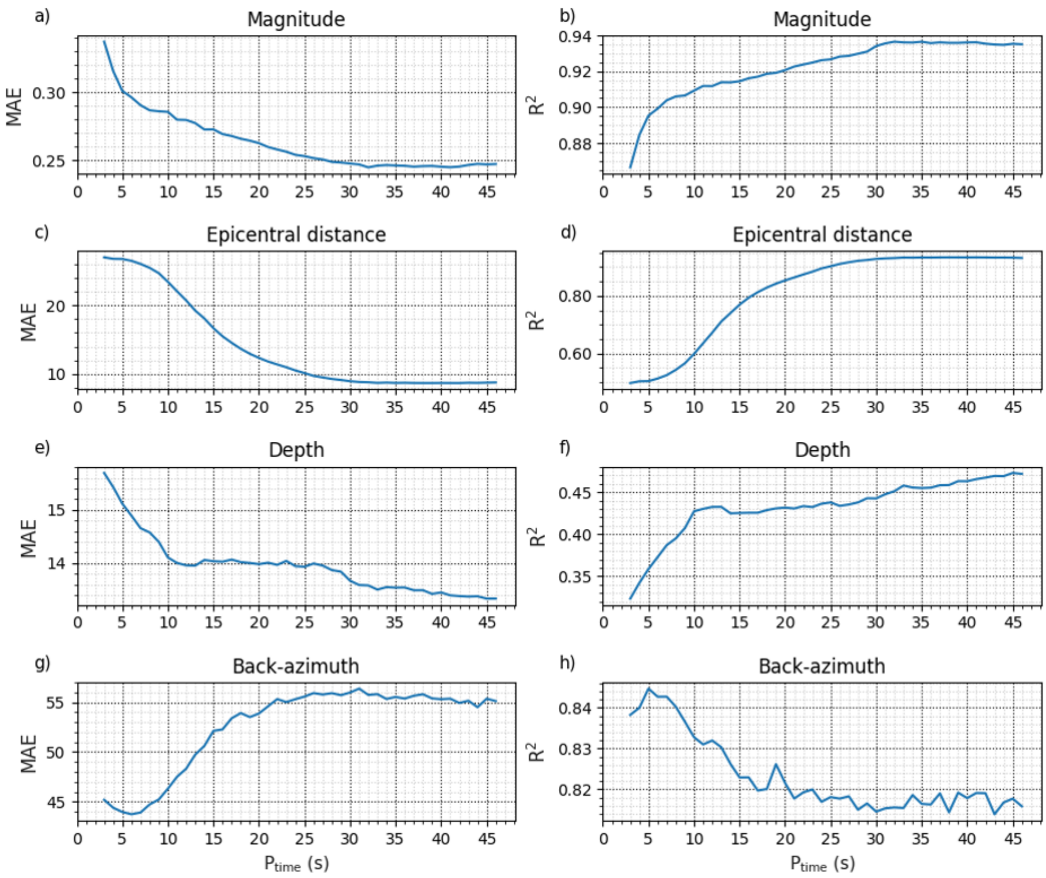

Figure 6. MAE and R2 results using 3 s to 46 s of P wave.

FORES makes its first estimate (M6.4) at 20 seconds after origin time (OT) when the true instantaneous magnitude is M7.3, outperforming the estimation of M5.7 by E3WS. However, one second 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 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 (Ptime) of each station, to compare them to the event's STF (Fig. 8b) given by the SCARDEC 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 dataset, 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

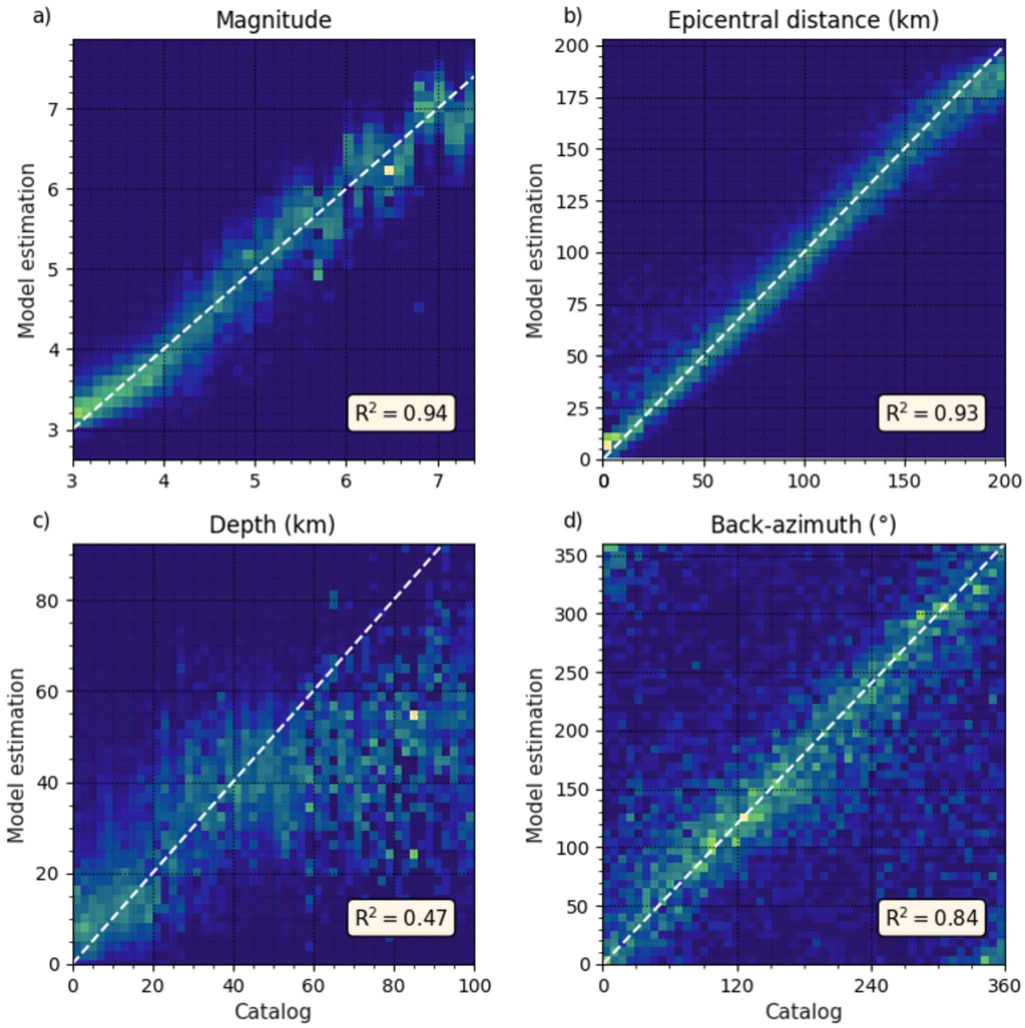

Figure 7. Same as Fig. 5 using 46 s of signal after P-wave arrival.

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 (Fig. 9), E3WS estimations track the magnitude evolution in agreement with the STF, with no systematic underestimation, 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). This 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.

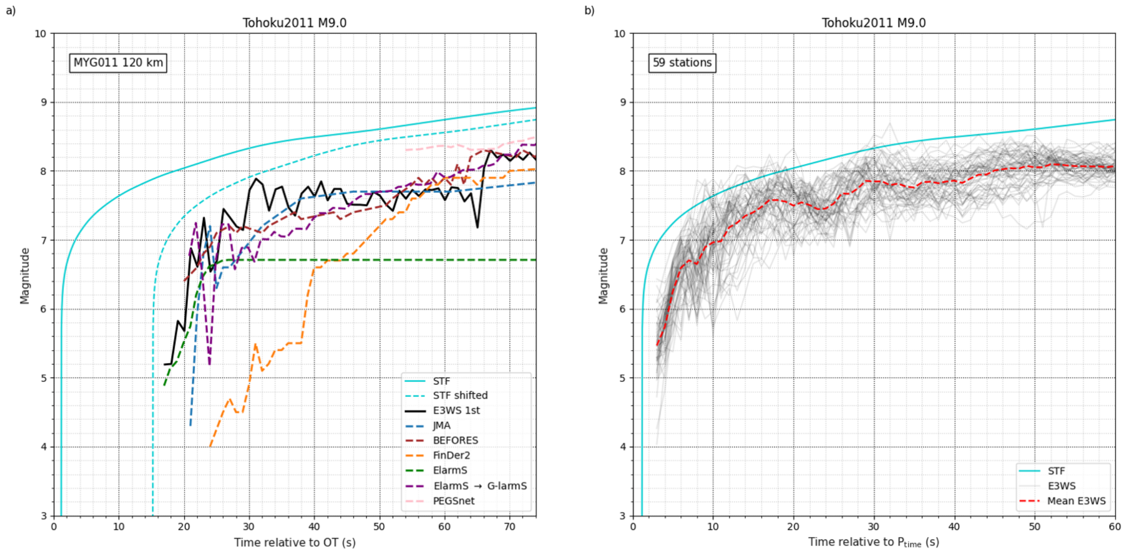

Figure 8. Real-time magnitude estimates for the 2011 M_W 9.0 Tohoku-Oki earthquake. (a) Magnitude evolution estimated by several EEW algorithms (see legend) as a function of time relative to the earthquake origin time. We also show the magnitude from the seismologically determined Source Time Function (STF) 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.

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.

E3WS does not trigger for 15 earthquakes (Table S1). The maximum magnitude of these missed events is 3.8, with a strong trade-off between magnitude and distance (Table S1). These magnitudes are below those that generate significant shaking in coastal Peru; they would not warrant an alert. E3WS triggers for 3 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 km, 396 km 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 > 1000 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.

Table 4. E3WS earthquake detections using 3 s of P wave in a continuous month (January 2022) at station SLN1.

M_{true}	M_{est}	Detection (s)	Warning time (s)
5.6	5.4	0.7	13.4
4.3	4.2	1.4	11.1
4.0	3.9	1.1	7.9
4.0	3.6	1.1	17.3
3.9	3.5	0.5	23.6
3.7	3.5	0.9	11.5
3.6	3.3	1.3	9.3
3.6	3.4	1.1	18.5
3.5	3.5	1.0	12.7
3.5	3.4	0.8	11.3
3.5	3.5	0.7	11.8
3.5	3.3	1.6	18.6
3.1	3.2	1.4	13.8
3.1	3.2	0.9	8.5

5 Discussion

5.1 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, i.e. it is a measure of the improvement in the estimates when using a particular attribute. 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.

The attributes that contribute the most to magnitude estimation, both using short and long portions of P wave, are the MFCC (Fig. 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), that share similarities with features that are widely used for magnitude estimation in other EEWS, such as P_d or τ_c. We hypothesize that the MFCC, by measuring energies on the Mel scale (a logarithmic frequency scale), 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, as is our approach, requires time integration, which is prone to amplify noise. Thus, it might preferable to not include them in E3WS. Indeed, our tests show better efficiency when using acceleration waveforms. Moreover, E3WS requires unclipped data for strong earthquakes as provided by accelerometers.

5.2 Comparative performance of E3WS and other EEWS

We compare the performance of E3WS with that of ElarmS (Brown et al., 2011), which estimates earthquake magnitude within the first 4 seconds of the 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 $\sim$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.

We also test the performance of E3WS compared to UrEDAS. Lockman and Allen (2005) report results applying UrEDAS using stations containing at least 5 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 seconds 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 seconds, E3WS achieved a MAE of magnitude of 0.22, significantly better than UrEDAS with 4 seconds. For location, E3WS yields results similar to UrEDAS, with MAE of 14 km for hypocentral distance and 20° for back-azimuth. Using 4 seconds 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 link 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°.

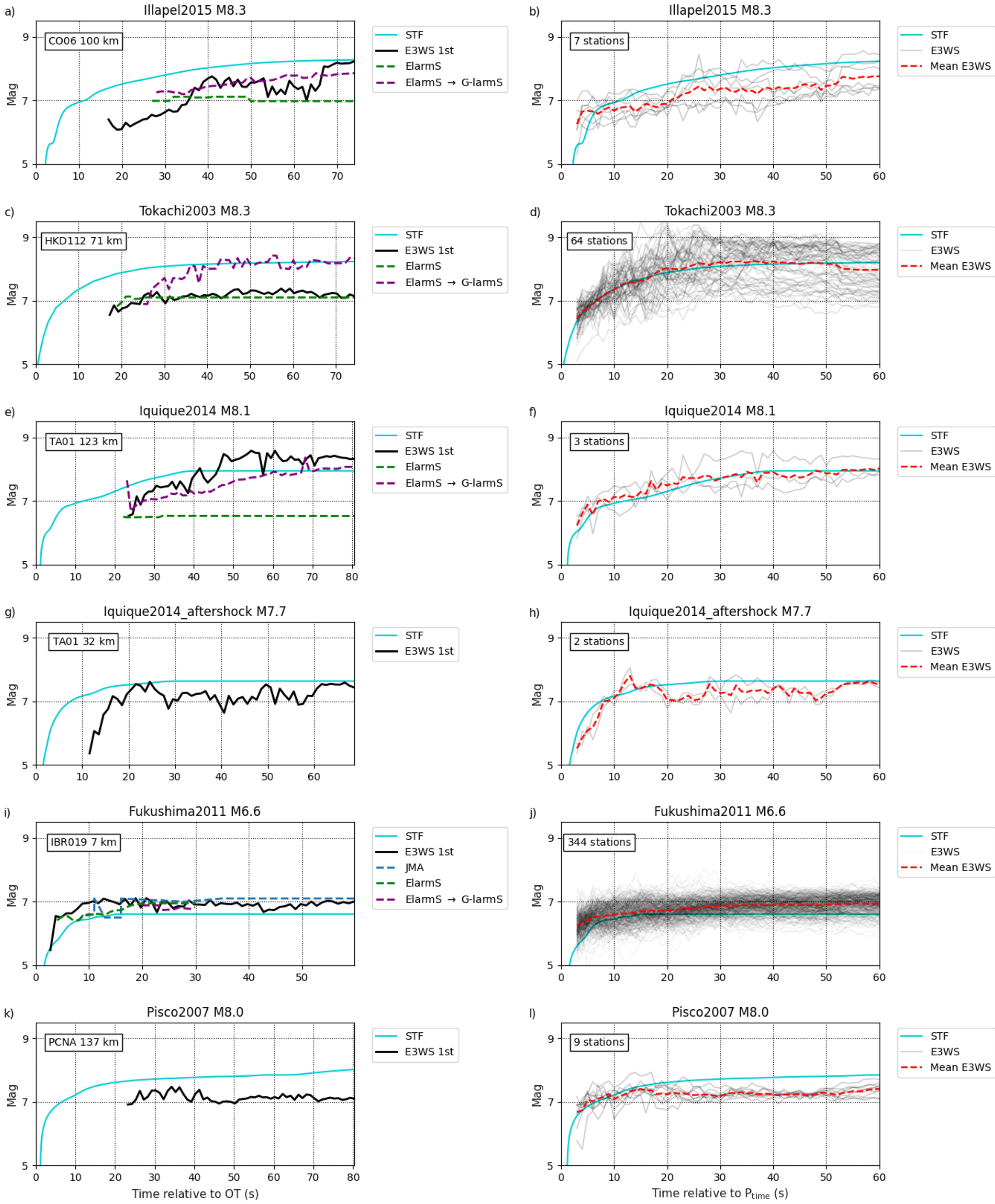

Figure 9. Magnitude estimates for the following earthquakes: 2015 M_W 8.3 Illapel, 2003 M_W 8.3 Tokachi-Oki, 2014 M_W 8.1 Iquique, 2014 M_W 7.7 Iquique aftershock, 2011 M_W 6.6 Fukushima aftershock, 2007 M_W 8.0 Pisco. Estimates are shown as a function of time relative to the earthquake's OT for the 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 with those obtained by other EEWS. On the right, we show all the estimates (gray), their mean (red), the moment function (the integral of the STF, light blue).

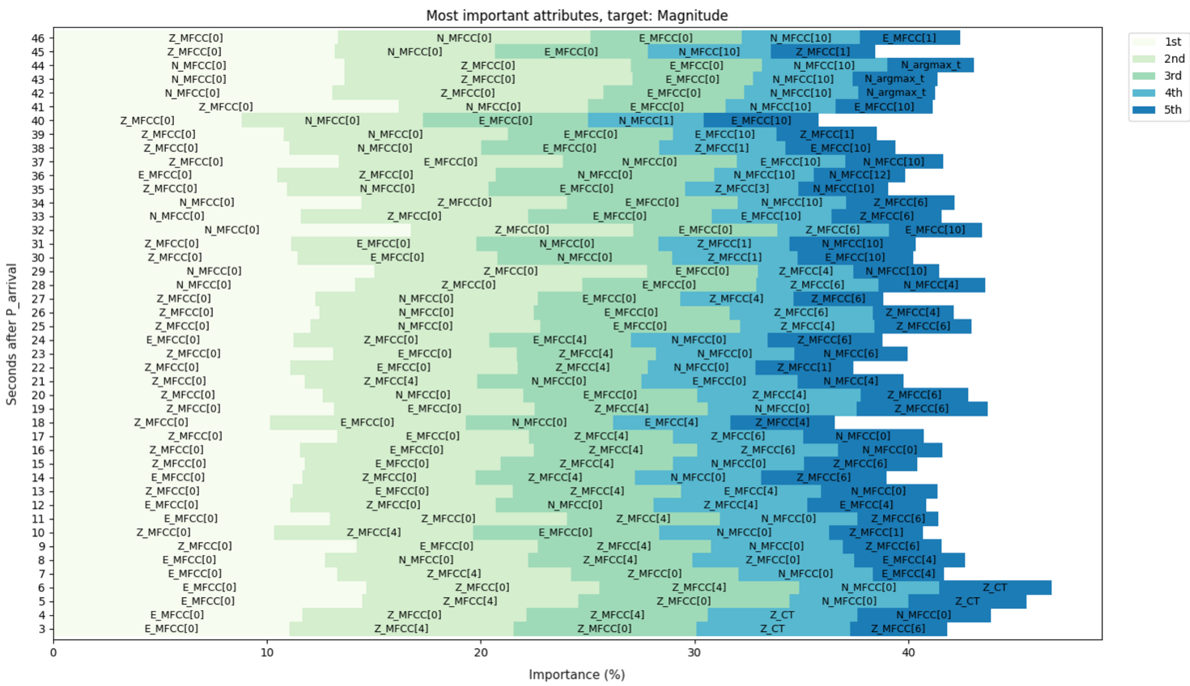

Figure 10. First (lightest color) to fifth (darkest color) most important features for magnitude estimation as a function of the P-wave window duration, from 3 to 46 s. For each time window, feature importance is based on the corresponding stacking model (see subsection 3.3), which consists of 10 XGB base models. Importance (%) shown is calculated as the gain mean plus STD 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.

6 Conclusion

We introduced E3WS, a set of Machine Learning algorithms using only 3 seconds of P-wave signal recorded by a single accelerometric station to detect, locate and estimate the magnitude of an earthquake. E3WS is made of 6 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 seconds of seismic recording, such as ElarmS and UrEDAS. 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

Waveforms and metadata used in this article were provided by the University of Chile downloaded by IRIS Web Services (`https://service.iris.edu/`), NIED K-NET, KiK-net, National Research Institute for Earth Science and Disaster Resilience, `doi:10.17598/NIED.0004` (`https://www.kyoshin.bosai.go.jp/`), the Stanford Earthquake Dataset (`https://github.com/smousavi05/STEAD`), and through petition to Instituto Geofísico del Perú (`https://www.gob.pe/igp`).

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.

References

Allen, R. M. (2007). The elarms earthquake early warning methodology and application across california. *Earthquake early warning systems*, 21–43.

Allen, R. M., & Kanamori, H. (2003). The potential for earthquake early warning in southern california. *Science*, *300*(5620), 786–789.

Allen, R. M., & Melgar, D. (2019). Earthquake early warning: Advances, scientific challenges, and societal needs. *Annual Review of Earth and Planetary Sciences*, *47*, 361–388.

Allen, R. V. (1978). Automatic earthquake recognition and timing from single traces. *Bulletin of the seismological society of America*, *68*(5), 1521–1532.

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.

Barrientos, S., & Team, N. S. C. C. (2018). The seismic network of chile. *Seismological Research Letters*, *89*(2A), 467–474.

Böse, M., Allen, R., Brown, H., Gua, G., Fischer, M., Hauksson, E., ... others (2014). Cisn 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.

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.

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.

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.

Cremen, G., Galasso, C., & Zuccolo, E. (2022). Investigating the potential effectiveness of earthquake early warning across europe. *Nature communications*, *13*(1), 1–10.

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.

Davis, S., & Mermelstein, P. (1980). Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *IEEE transactions on acoustics, speech, and signal processing*, *28*(4), 357–366.

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.

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.

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.

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.

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.

Lara, P. E. E., Fernandes, C. A. R., Inza, A., Mars, J. I., Métaxian, 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.

Licciardi, A., Bletery, Q., Rouet-Leduc, B., Ampuero, J.-P., & Juhel, K. (2022). Instantaneous tracking of earthquake growth with elastogravity signals. *Nature*, 1–6.

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.

Malfante, M., Dalla Mura, M., Métaxian, 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.

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.

Meier, M.-A., Ampuero, J., & Heaton, T. H. (2017). The hidden simplicity of subduction megathrust earthquakes. *Science*, *357*(6357), 1277–1281.

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.

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.

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.

Mousavi, S. M., Sheng, Y., Zhu, W., & Beroza, G. C. (2019). Stanford earthquake dataset (stead): A global data set of seismic signals for ai. *IEEE Access*, *7*, 179464–179476.

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.

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.

Nielsen, D. (2016). *Tree boosting with xgboost-why does xgboost win" every" machine learning competition?* (Unpublished master's thesis). NTNU.

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.

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.

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.

Yih-Min, W., Kanamori, H., Allen, R. M., & Hauksson, E. (2007). Determination of earthquake early warning parameters, τ c and pd, for southern california. *Geophysical Journal International*, *170*(2), 711–717.

Supporting Information for "Earthquake Early Warning using 3 seconds of records on a single station"

Pablo Lara[1,2], Quentin Bletery[1], Jean-Paul Ampuero[1], Adolfo Inza[2]

[1]Université Côte d'Azur, IRD, CNRS, Observatoire de la Côte d'Azur, Géoazur, France

[2]Instituto Geofísico del Perú, Lima, Perú

Contents of this file

1. Introduction

This supporting information includes the attributes used in this work, 9 supplementary figures and 1 supplementary table.

2. Attributes

We detail here the attributes computed to train the Machine Learning algorithms, with their corresponding domain and signal.

2.1. Attributes from 3-component seismograms

1. Maximum eigenvalue λ_1 of covariance matrix from the 3C seismogram.

2. Eigenvalue factor: ratio of the maximum eigenvalue to the sum of the remaining eigenvalues:

$$\lambda_{factor} = \lambda_1/(\lambda_2 + \lambda_3). \tag{1}$$

3. The 3 components of the eigenvector ν_1 associated with the maximum eigenvalue λ_1.

2.2. Attributes from each channel

N denotes the number of samples per channel within the time window. F_s denotes the sampling rate, in Hz. The envelope of the signal s is defined as $e = |s + iH\{s\}|$, where H is the Hilbert transform.

2.2.1. Time-domain attributes

4. Maximum energy of the preprocessed signal:

$$E_{max} = \max(s^2). \tag{2}$$

5. Time at which the maximum energy is reached:

$$t_{E_{max}} = argmax(s^2). \tag{3}$$

6. Total energy:

$$E_{ne} = \sum_{n=1}^{N} s^2[n]. \tag{4}$$

7. Energy centroid time:

$$CT = \frac{1}{E_{ne}} \sum_{n=1}^{N} n\ s^2[n]. \tag{5}$$

8. Bandwidth, characteristic duration, variance around the energy centroid:

$$BW_t = \sqrt{\frac{\sum_{n=1}^{N} (n - CT)^2 s^2[n]}{E_{ne}}}. \tag{6}$$

9. Skewness around bandwidth:

$$Skewness_{BW_t} = \begin{cases} \sqrt{SK_{pre}}, & \text{if } SK_{pre} \geq 0 \\ -\sqrt{-SK_{pre}}, & \text{otherwise.} \end{cases}$$

where

$$SK_{pre} = \frac{\sum_{n=1}^{N} (n - CT)^3 s^2[n]}{E_{ne}\ BW_t^3}. \tag{7}$$

10. Kurtosis around bandwidth:

$$Kurtosis_{BW_t} = \sqrt{\frac{\sum_{n=1}^{N} (n - CT)^4 s^2[n]}{E_{ne}\ BW_t^4}}. \tag{8}$$

11. Mean envelope:

$$< env >= \frac{1}{N} \sum_{k=1}^{N} e[k]. \tag{9}$$

12. Ratio of maximum amplitude envelope to its mean:

$$RMM_t = \frac{\max(e)}{< env >}. \tag{10}$$

13. Standard deviation of the envelope:

$$STD_{env} = \sqrt{\frac{\sum_{k=1}^{N} (e[k] - < env >)^2}{N}}. \tag{11}$$

14. Skewness of the envelope:

$$Skewness_{env} = \frac{1}{N} \sum_{k=1}^{N} \left(\frac{e[k] - < env >}{STD_{env}} \right)^3. \tag{12}$$

15. Kurtosis of the envelope:

$$Kurtosis_{env} = \frac{1}{N} \sum_{k=1}^{N} \left(\frac{e[k] - < env >}{STD_{env}} \right)^4. \tag{13}$$

16. Threshold-crossing rate of the envelope signal: how many times per second the signal envelope crosses the threshold of 80% of its maximum amplitude:

$$TCR_t = \frac{count(r[n]r[n-1] < 0)}{N/F_s}, \tag{14}$$

where:

$$r = e/\max(e) - 0.8. \tag{15}$$

17. Fraction of envelope samples that exceed a threshold of 80% of the envelope maximum:

$$fract(TCR_{env}) = count(e \geq 0.8 \max(e))/N. \tag{16}$$

18. Shannon entropy of the envelope, with $N_{bins} = 200$.

$$Shannon_{env} = -\sum_{i=1}^{N_{bins}} Prob_e[i] \log_2(Prob_e[i]), \tag{17}$$

where:

$$Prob_e[i] = Histogram(e, N_{bins}). \tag{18}$$

19. Renyi entropy of the envelope, with $\alpha = 2$.

$$Renyi_{env} = \frac{\log_2 \sum_{i=1}^{N_{bins}} Prob_e^{\alpha}[i]}{1-\alpha}, \tag{19}$$

20. Zero crossing rate, how many times per second the signal s changes sign:

$$ZCR_t = \frac{count(s[n]s[n-1] < 0)}{N/F_s} \tag{20}$$

2.2.2. Spectral-domain attributes

Attributes extracted from $p = PSD(s)$, the Welch's Power Spectral Density of the signal s. Here N denotes the number of frequency samples in the spectrum up to the Nyquist frequency $F_s/2$.

21. Mean PSD:

$$< PSD >= \frac{1}{N} \sum_{k=1}^{N} p[k]. \tag{21}$$

22. Maximum spectral energy:

$$PSD_{max} = \max(p). \tag{22}$$

23. Frequency index of maximum spectral energy:

$$f_{PSD_{max}} = argmax(p). \tag{23}$$

24. Centroid frequency of the spectrum:

$$CF = \frac{\sum_{k=1}^{N} k \; p[k]}{\sum_{k=1}^{N} p[k]}. \tag{24}$$

25. Frequency bandwidth, variance around the spectral centroid:

$$BW_f = \sqrt{\frac{\sum_{k=1}^{N} (k - CF)^2 \; p[k]}{\sum_{k=1}^{N} p[k]}}. \tag{25}$$

26. Skewness of the spectrum:

$$Skewness_{BW_f} = \begin{cases} \sqrt{SK_{pre}}, & \text{if } SK_{pre} \geq 0 \\ -\sqrt{-SK_{pre}}, & \text{otherwise.} \end{cases}$$

where

$$SK_{pre} = \frac{\sum_{k=1}^{N} (k - CF)^3 \; p[k]}{BW_f^3 \sum_{k=1}^{N} p[k]}, \tag{26}$$

27. Kurtosis of the spectrum:

$$Kurtosis_{BW_f} = \sqrt{\frac{\sum_{k=1}^{N} (k - CF)^4 \; p[k]}{BW_f^4 \sum_{k=1}^{N} p[k]}}. \tag{27}$$

28. Standard deviation of the PSD:

$$STD_{PSD} = \sqrt{\frac{\sum_{k=1}^{N} (p[k] - < PSD >)^2}{N}}. \tag{28}$$

29. Skewness of PSD:

$$Skewness_{PSD} = \frac{\sum_{k=1}^{Na} \left(\frac{p[k] - < PSD >}{STD_{PSD}} \right)^3}{N}. \tag{29}$$

30. Kurtosis of PSD:

$$Kurtosis_{PSD} = \frac{\sum_{k=1}^{N} \left(\frac{p[k] - < PSD >}{STD_{PSD}} \right)^4}{N}. \tag{30}$$

31. Shannon entropy, with $N_{bins} = 50$:

$$Shannon_{PSD} = - \sum_{i=1}^{N_{bins}} Prob_p[i] \; \log_2(Prob_p[i]), \tag{31}$$

where:

$$Prob_p[i] = Histogram(p[k], N_{bins}). \tag{32}$$

32. Renyi entropy, with $\alpha = 2$:

$$Renyi_{PSD} = \frac{\log_2 \sum_{i=1}^{N_{bins}} Prob_p^{\alpha}[i]}{1 - \alpha}. \tag{33}$$

33. Ratio of maximum PSD amplitude to its mean.

$$RMM_f = \frac{\max(p)}{< PSD >}. \tag{34}$$

34. Threshold-crossing rate of the PSD, how many times the PSD crosses a threshold of 40% of its maximum amplitude:

$$TCR_f = \frac{count(r[k]r[k-1] < 0)}{N/F_s}, \tag{35}$$

where:

$$r = PSD/\max(PSD) - 0.4 \tag{36}$$

35. Relative number of samples that exceed a threshold of 40% of its maximum.

$$fract(TCR_{PSD}) = count(p \geq 0.4 \max(p))/N. \tag{37}$$

2.2.3. Cepstral-domain attributes

36. The 13 first mel-frequency cepstrum coefficients (MFCC):

$$MFCC[m] = DCT\{log[\sum\{|F\{s\}|^2 \Lambda_m\}]\}, \tag{38}$$

where DCT is the Discrete Cosine Transform, $F\{.\}$ is the Discrete Fourier Transform, and Λ is a triangular filter bank function linearly spaced from 1 to 45 Hz in a Mel scale. In this work, we use $m = 26$ filter banks, and are compute as in (Kopparapu & Laxminarayana, 2010).

3. Figures

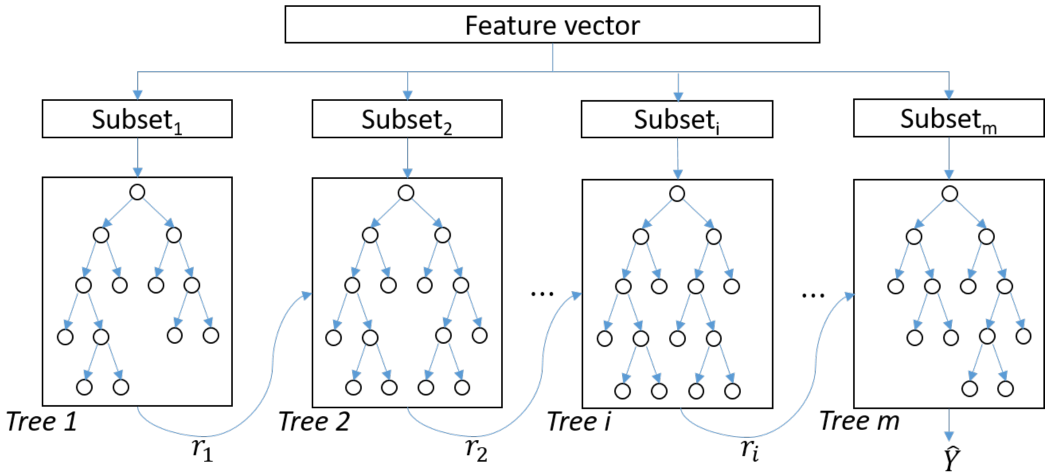

Figure S1. General architecture XGB.

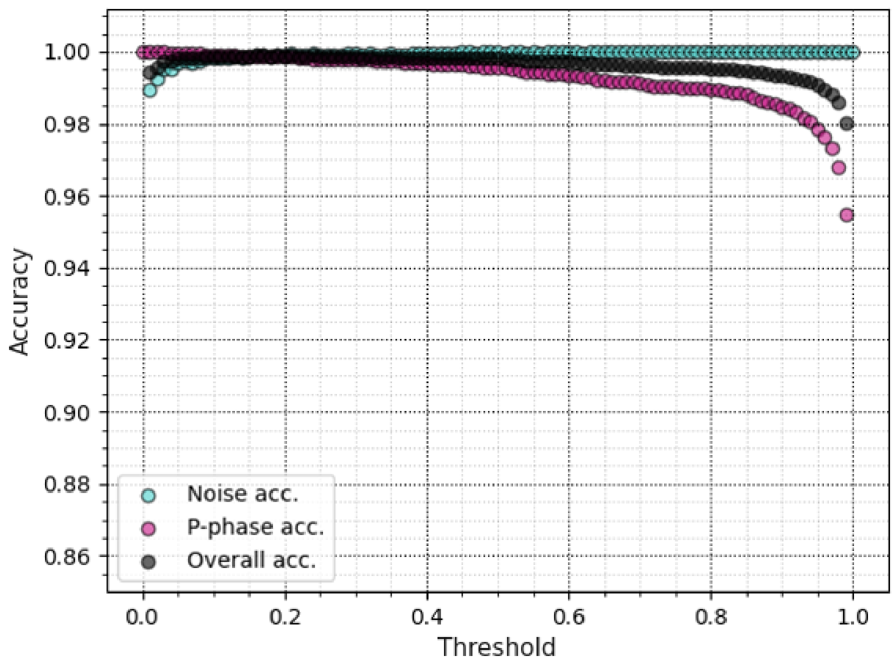

Figure S2. Accuracy of noise and earthquake classification, using different thresholds.

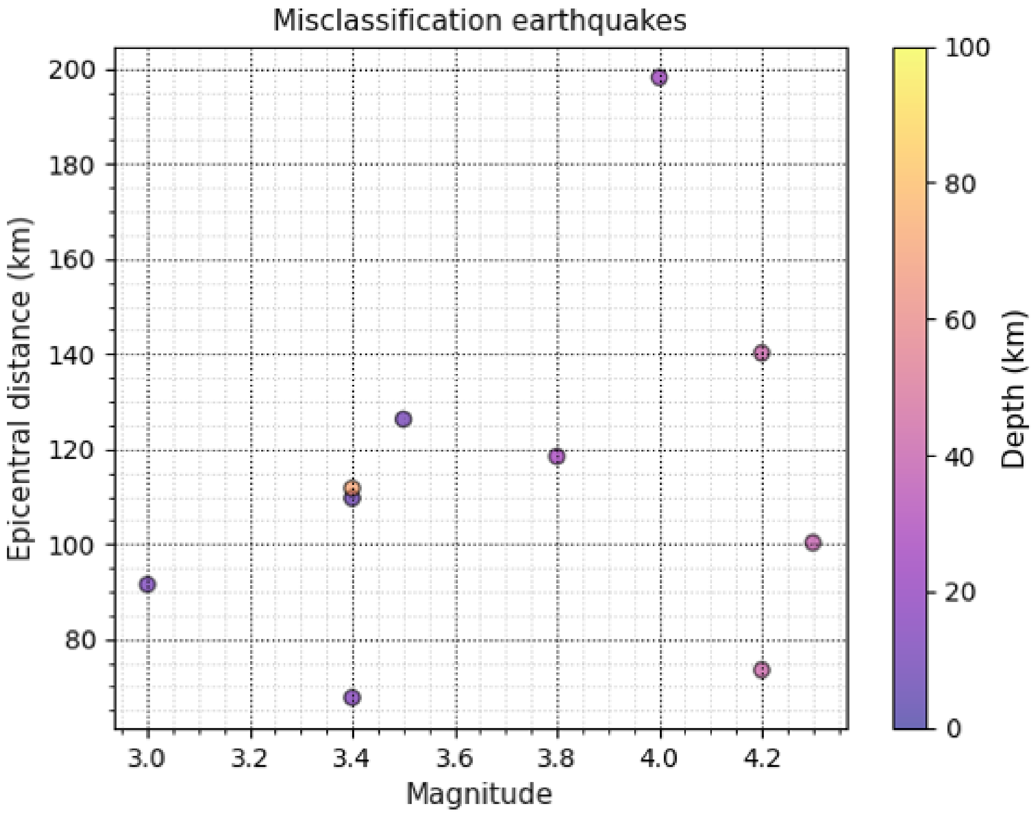

Figure S3. Magnitude, epicentral distance and depth of the misclassified signals shown in Fig. S4.

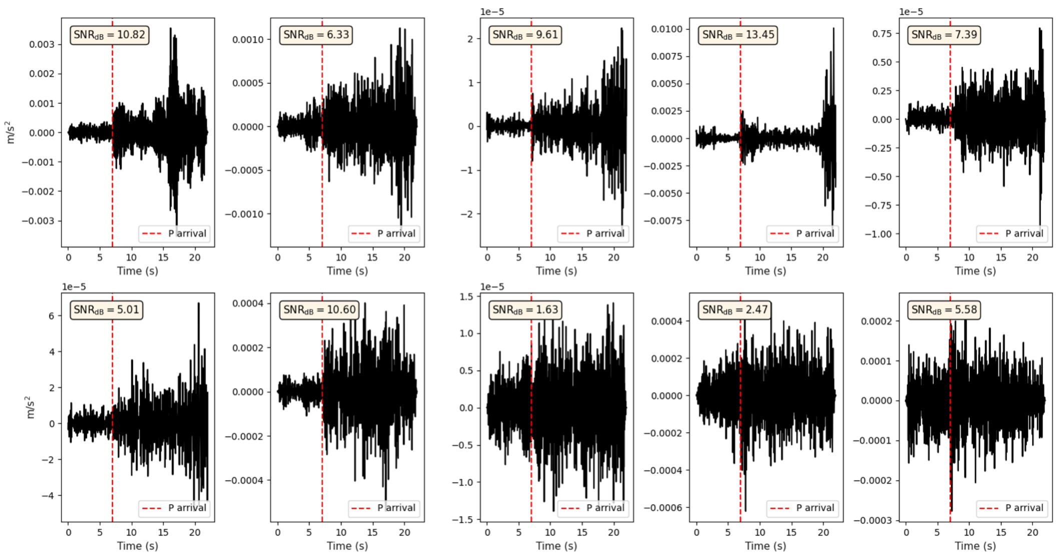

Figure S4. Earthquake signals misclassified as noise, and their signal-to-noise ratios (SNR).

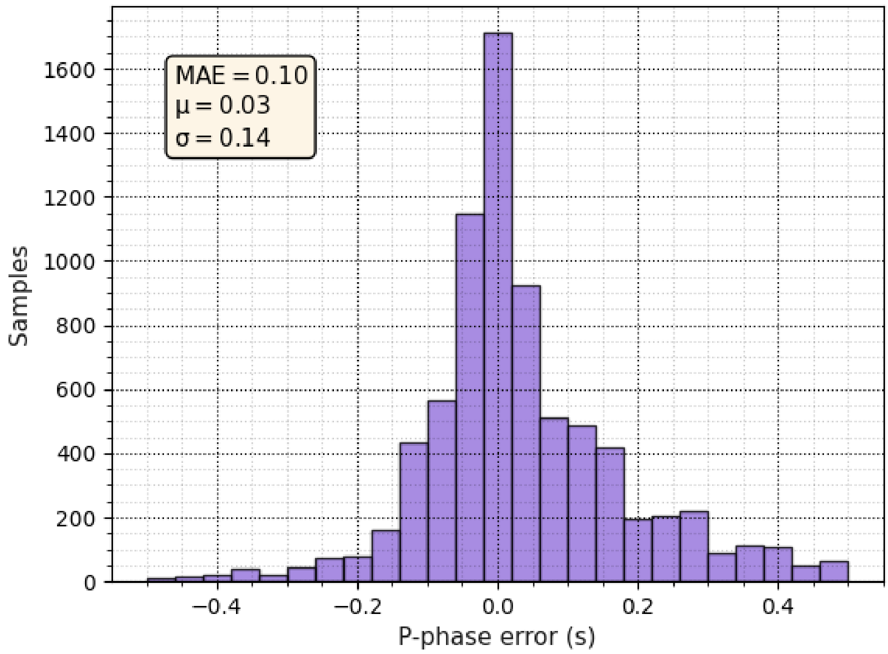

Figure S5. histogram of errors between the true and predicted P-phase arrival times.

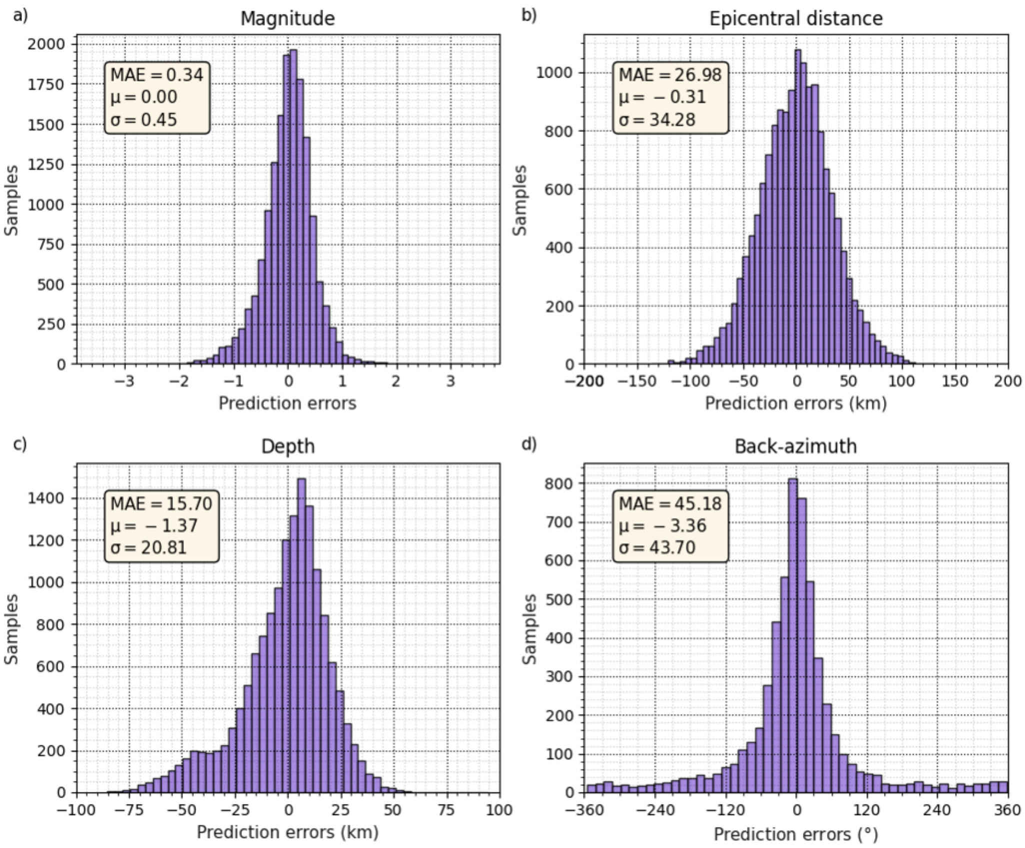

Figure S6. Histogram of the errors in the source characterization predictions using 3 s of P-wave.

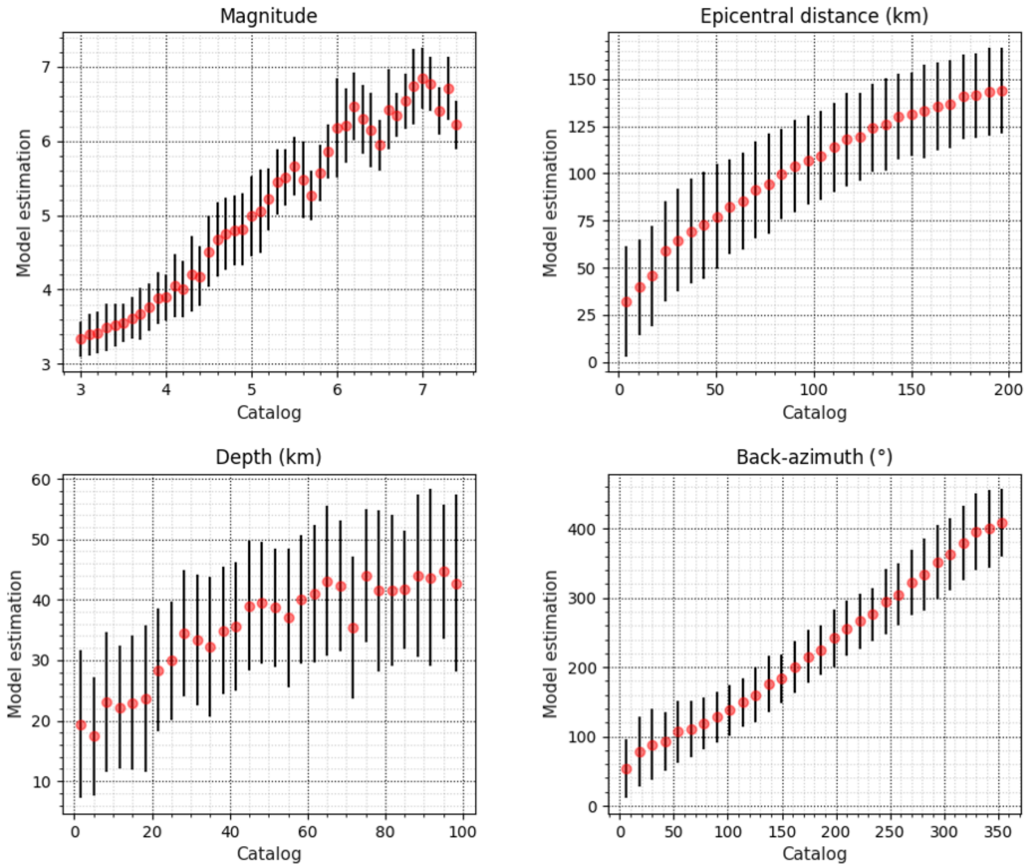

Figure S7. Mean (circle) and STD (bar) predictions per bin using 3 s of P-wave.

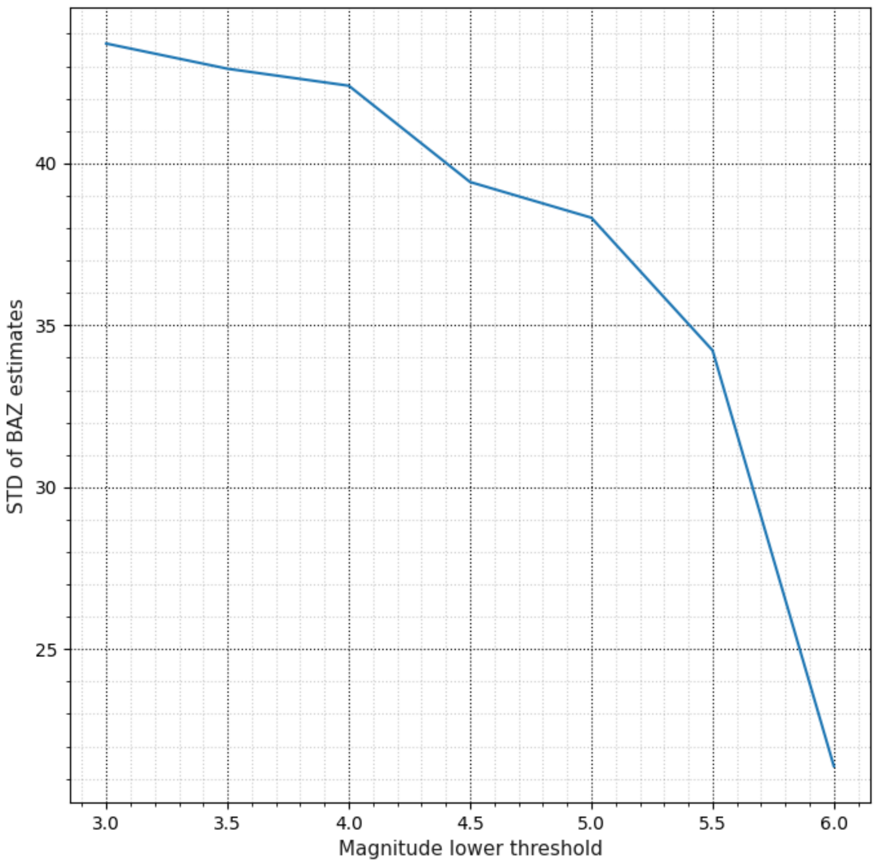

Figure S8. STD of the back-azimuth estimates, using different lower thresholds of magnitude.

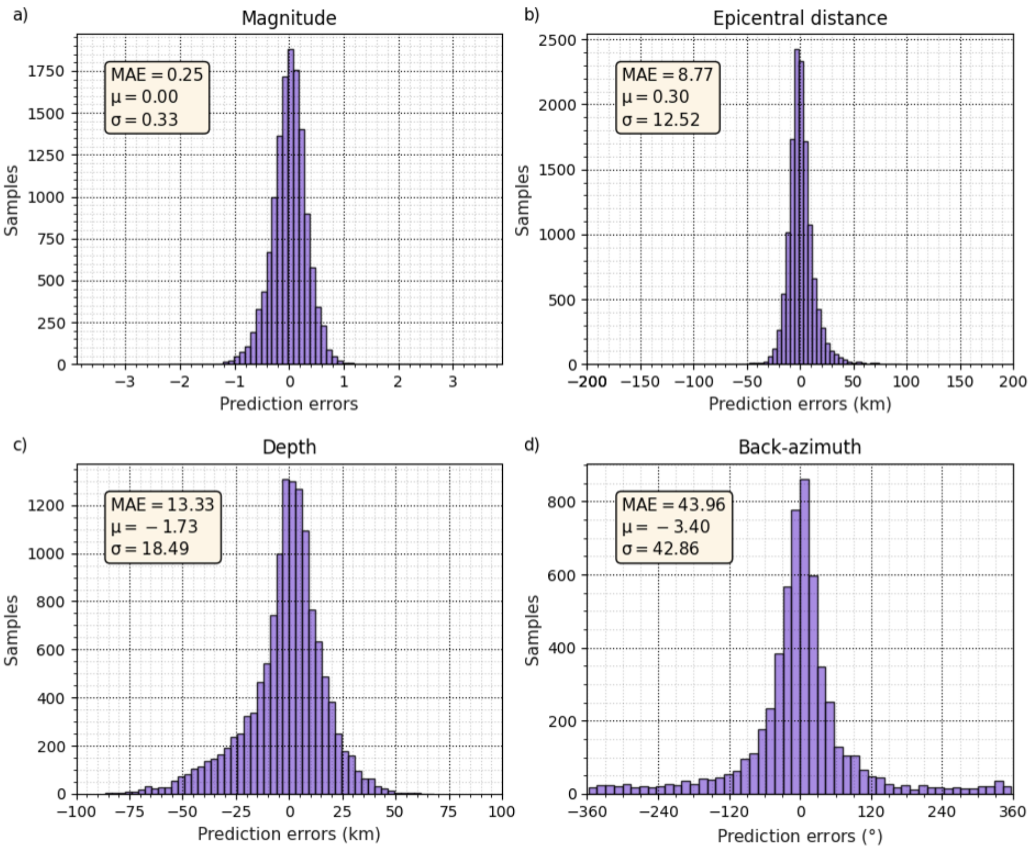

Figure S9. Histogram of the errors in the predictions using 46 s of P-wave.

Table S1. Real-time earthquake detection by E-EEWS using 3 s of P-wave in a continuous month.

Magnitude	Distance (km)	Depth (km)
3.1	175	13
3.5	179	56
3.3	195	85
3.2	145	78
3.2	173	54
3.3	89	50
3.3	98	86
3.2	163	17
3.7	162	19
3.3	159	84
3.8	155	10
3.0	47	35
3.4	189	88
3.5	138	47
3.5	97	49

References

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).

Earthquake Early Warning using 3 seconds of records on a single station

Pablo Lara[1,2], Quentin Bletery[1], Jean-Paul Ampuero[1], and Adolfo Inza[2]

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

Corresponding author: Pablo Lara, pablo.elara@ieee.org

Abstract

We introduce the Ensemble Earthquake Early Warning System (E3WS), a set of Machine Learning algorithms designed to detect, locate and estimate the magnitude of an earthquake using 3 seconds of P waves recorded by a single station. The system is made of 6 Ensemble Machine Learning algorithms trained on attributes computed from ground acceleration time series in the temporal, spectral and cepstral domains. The training set comprises datasets from Peru, Chile, Japan, and the STEAD global dataset. 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.

1 Introduction

Today millions of people live at risk from earthquakes. Real-time seismic monitoring near seismic sources opens the possibility of rapidly estimating earthquake parameters that control the potential impact of shaking, notably magnitude and hypocenter location. Taking advantage of such estimates and speed-of-light communications, Earthquake Early Warning Systems (EEWS) can generate an alarm before seismic shaking reaches the population, 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 warning time to take protective measures. In a common approach, the magnitude and location of the earthquake are rapidly estimated, then used as input in a Ground Motion Prediction Equation to forecast ground shaking intensity measures such as Peak Ground Acceleration, 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. 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 focus on single-station 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 and Global Navigation Satellite System (GNSS) stations (R. M. Allen & Melgar, 2019).

The τ_c-P_d Onsite algorithm (Böse et al., 2009), one of the three algorithms that contributed to the development of ShakeAlert, the EEWS 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 seconds 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, originally developed in Japan. 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).

Recently, Artificial Intelligence (AI) has been used in a number of applications in seismology that are relevant for EEW. In particular, studies not designed for EEW purposes used AI for specific targets, such as detection, picking and source characterization components. The Earthquake Transformer algorithm (EQTransformer) (Mousavi et al., 2020) uses 1 minute long seismograms to feed an architecture based on Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) to detect earthquakes, and estimate P and S phase arrivals. The model achieves an earthquake detection precision (true positives divided by total positives) of 1.0, and estimates the P and S phase arrivals with a mean of 0 seconds and standard deviation (STD) of 0.03 seconds for the P phase, and mean of 0.0 seconds and STD of 0.11 seconds for the S phase. Mousavi and Beroza (2020) estimate magnitudes with CNN and RNN trained with 30 seconds long seismograms ($M<5.7$). They obtain a mean error close to 0 magnitude units and standard deviation $\sim$0.2. Mousavi and Beroza (2019) estimate earthquake location based on Bayesian Deep Learning. The network is fed by 1 minute long seismograms in the case of training distances. For back-azimuth, the angle is represented as points of the unit circle (cosine, sine) and is trained using seismograms of 1.5 s long. The network achieves a mean localization error of 7.27 km with STD of 12.16 km. While these AI models are not designed for EEW (they use signal windows that are too long), 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 seconds of signal recorded by a single three-component sensor. E3WS is a system focused on early warning for populations living near seismic sources. 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. Its magnitude residuals are small enough to not generate false warnings, for both overestimates and underestimates. 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 Machine Learning, allowing (in contrast to Deep Learning approaches) some understanding of what controls the estimations. Thanks to its simplicity, E3WS can be installed in small 32-bits single board computers, such as a Raspberry Pi, as well as in more complex 64-bits processors.

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). 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 73,000 earthquake seismograms. Data statistics are shown in Fig. 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. We ignore sensors for which we did not find the instrument response.

The preprocessing also includes semi-automatic detection of the P-phase for the Chilean and Japanese datasets. Since the Chilean, KiK-net and K-NET datasets do not include P-phase picks, we used an STA/LTA algorithm (R. M. Allen, 2007) and theoretical travel times to estimate the P-phase and S-phase arrival times. Sometimes we find in the raw data of a given earthquake more than one earthquake. In these cases, the STA/LTA algorithm triggered in most of the earthquakes. To avoid selecting an earthquake that does not belong to the dataset catalog, we select the earthquake in which the difference in the estimated time between the STA/LTA algorithm and theoretical travel times is not greater than three seconds. Then, we chose the STA/LTA trigger time as the P-phase time.

For the Peru, Chile and STEAD datasets, we manually identified and removed signals that were saturated for M>5.0. We found no further saturated waveforms when inspecting events down to M4.5. We believe that for magnitudes smaller than 4.5, there are no more saturated waveforms. The Japanese data are recorded by strong-motion accelerometers, which do not saturate.

3 Proposed EEWS

E3WS consists of 6 ML algorithms : a detector, a P-phase picker and 4 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 seconds of P-wave signal, the 4 regression models run in parallel 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 6 models take as input a feature vector formed by concatenating 140 attributes extracted from the waveforms, their spectrum and their cepstrum.

We test several approaches to design the models, including Extreme Gradient Boosting (XGB), Support Vector Machine (SVM), Random Forest (RF) and Neural Networks (Multilayer Perceptron: MLP). We find that the approach yielding the best results is XGB (Chen & Guestrin, 2016), a Supervised Machine Learning 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 seconds of the P-wave. 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

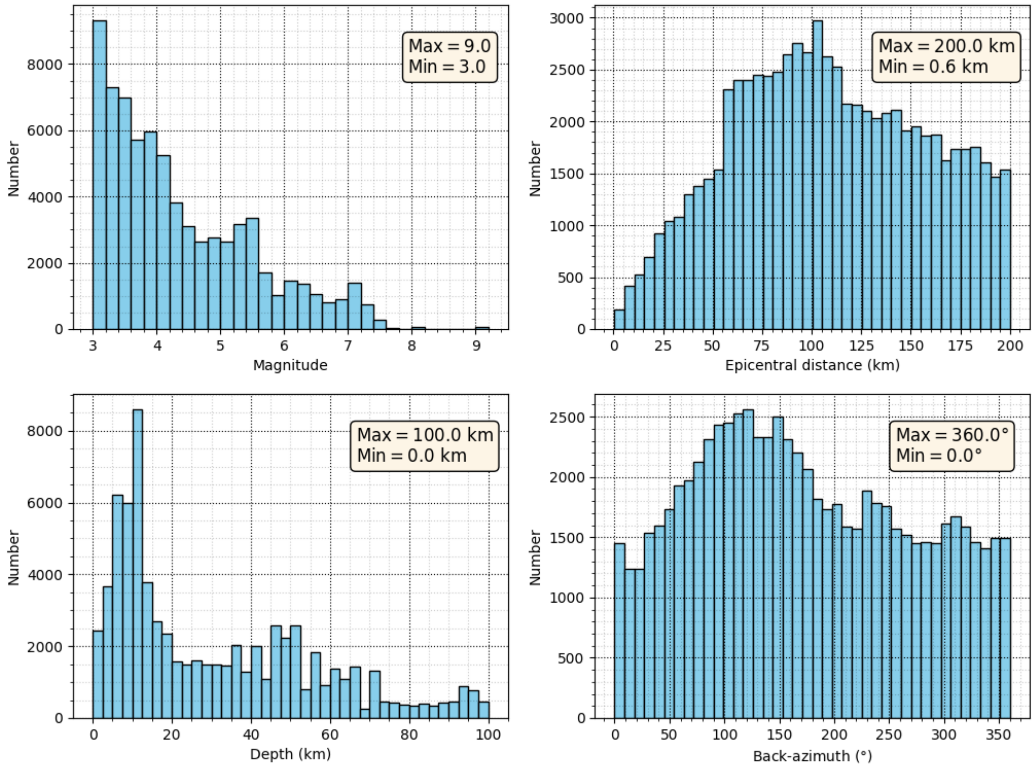

Figure 1. Magnitude, epicentral distance, depth and back-azimuth distributions of the earthquake waveform database compiled for this work.

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, on an Intel(R) Xeon(R) Silver 4114 processor. Even on a Single Board Computer like a Raspberry Pi 4, XGB takes 0.01 s on average. 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. Although XGB only supports single-label training, it performs better than inherently multi-output regression models such as RF.

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 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. Fig. S1 shows the general XGB scheme. For all models, we use the following hyperparameters for XGB training: depth = 4, number of trees = 6000, subset = 80% and learning rate = 0.1.

3.1 Detection

Since STEAD is a global dataset that also includes global noise samples, we extract 55,000 noise windows and add them to our database. We use waveforms filtered from 1 to 7 Hz and a fixed time window of 10 seconds including a minimum of 3 seconds of P-wave. We discard waveforms that do not contain 7 seconds of data before the P-

Table 1. Accuracy and errors using XGB, RF, SVM and MLP models for detection, P-phase picking and source characterization.

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

wave arrival. In our tests we obtained better accuracy using a 10-seconds-long window compared to shorter windows. For instance, we find false detections due to impulsive noise using shorter windows; 10 seconds-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 seconds of P or S-wave, respectively. 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.

We estimate the likelihood that a window contains a P wave, sliding the 10-second window by steps of 0.5 seconds. 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.

3.2 P-phase picking

Because the Japanese and Chilean datasets do not provide P-phase arrival times (t_p), we restrict the training set for phase-picking to the Peruvian and STEAD datasets. We use a fixed time window of 4 seconds filtered between 1 and 7 Hz.

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 (Fig. 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 $[t_p{-}3.5s, t_p{+}0.5s]$ 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 extracted every 0.01 s, the

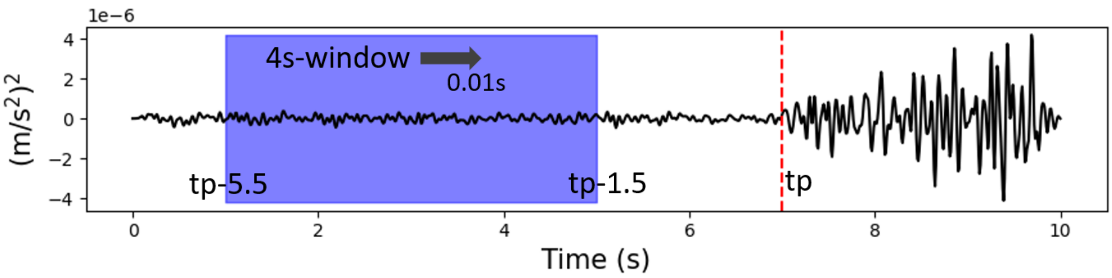

Figure 2. Labeling for the P-phase picking model. We extract attributes from a 4-seconds-long window, starting from $t_p - 5.5$ s as the blue box, for our entire database. We repeat the attribute extraction every 0.01 s until the blue box reaches $t_p + 2.5$ s.

input dataset for the P-phase picking model is made of approximately 36 million 140-dimensional samples.

3.3 Source characterization

For source characterization (esimation of earthquake magnitude, epicentral distance, back-azimuth and hypocentral depth), we use time windows that contain 7 seconds of noise and 3 seconds of P-wave signal extracted from our earthquake database. We apply a 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 datasets 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 (Fig. 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 dataset into K=10 groups and training each XGB base-model with nine out of the ten 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 ($\mathrm{OOF}_{\mathrm{pred}}$) to train the LASSO meta-model (Kukreja et al., 2006).

For the back-azimuth model, we divide the training into 2 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 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 sig-

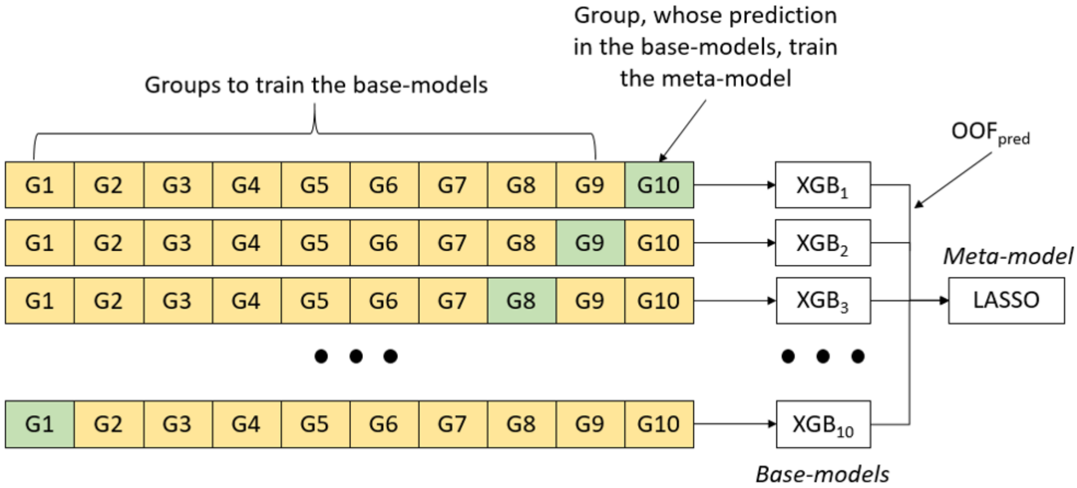

Figure 3. Source characterization model based on Stacking algorithm and K-Fold with K=10. For each K, nine groups train the XGB base-model. Estimates from the remaining group, using the corresponding trained XGB model, feed the LASSO meta-model.

nal $|s+iH\{s\}|$ where H is the Hilbert transform. For the spectral domain, we consider the Power Spectral Density (PSD) 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).

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. 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 seconds 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 seconds 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 (Fig. S2), 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 $M < 4.3$ (Fig. S3) and low signal-to-noise ratios (Fig. S4). Most of them have an epicentral distance greater than 100 km.

Table 2. Confusion matrix for the detection algorithm.

Overall (%):		**True class**	
99.9		Noise	Earthquake
Estimated	Noise	**11264**	10
class	Earthquake	0	**8778**
Accuracy (%):		100.0	99.9

4.1.2 P-phase picking

We evaluate the picker model on more than 10,000 seismograms of the test dataset compared to the "true" (manually picked) P-wave arrival times (Fig. S5). 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 (Figs. 5, S6, Table 3), given that our algorithm only uses 3 seconds of records on a single station.

Table 3. Performance of the source characterization algorithm with its mean error, STD error, MAE and coefficient of determination (R2).

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 (°)	-3.4	43.7	45.2	0.84

The magnitude estimates are very stable for earthquakes with magnitudes between 3.2 and 6.5, with magnitude average residuals ($|M_{pred} - M_{true}|$) of ∼0.2 for M<5.7 (Fig. 4a), and residuals between 0 and 0.4 for 5.7<M<6.5. We even observe magnitude residuals ∼0 for M6.2. This gives us confidence in estimating magnitudes for minor (M3.0-M3.9) to strong (M6.0-M6.9) earthquakes. For instance, for all M>6.0 earthquakes the average estimates are M¿6.0, so there would not be missed events in a EEW system that uses a threshold M>6.0 as a primary alert criterion. The small errors over the entire range of magnitudes are reflected in a high R2 of 0.87 (1.0 in the ideal case). For the smallest earthquakes of our database (M∼3), the magnitude estimates show a slight overestimation of 0.3 and STD of 0.2 (Fig. S7a), 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 seconds window duration.

We observe an average residuals at epicentral distances for distances very close to the seismic source (0-20 km) of ∼28 km (Fig. 4b). As the seismic gets farther away up to a distance of ∼120 km, the residuals decrease linearly even down to almost 0 error. From here, the errors grow linearly up to our training distance limit (200 km). If we keep our error tolerance at 28 km (errors at very close distances), we can estimate up to an epicentral distance of 165 km. Longer distances to this implies greater 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 ± STD (8-48 km depth), most events have average residuals smaller than 10 km (Fig. 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 back-azimuth, residuals exceed 35°. However, the STD of the estimates decreases significantly as the magnitude increases, achieving an STD of 21° for M>6.0 earthquakes (Fig. S8). The estimates have uniform performance throughout their range (Fig. S7d). The high R2 of 0.84 shows that the model contains relevant information in the whole back-azimuth range.

4.2 Performance of source characterization using longer signals

Meier et al. (2017) showed that the source time functions (STF) of large and very large 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 seconds of the P wave, i.e. the magnitude reached by the earthquake 3 seconds after its onset. This estimate can form the basis to generate a first warning and can be updated when longer records become available.

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 s to 46 s. Fig. 6 shows the evolution of two performance metrics, MAE and R2, as a function of the considered signal duration.

We observe a significant improvement in the estimations of magnitude and epicentral distance, with R2 scores increasing up to 0.94 and 0.93, respectively, and MAE dropping to about 0.25 and 9 km, respectively, at 46 s of signal (Fig. 7). After that time, most M≤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 seems sufficient to converge to the best performance (Fig. 6a-d).

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

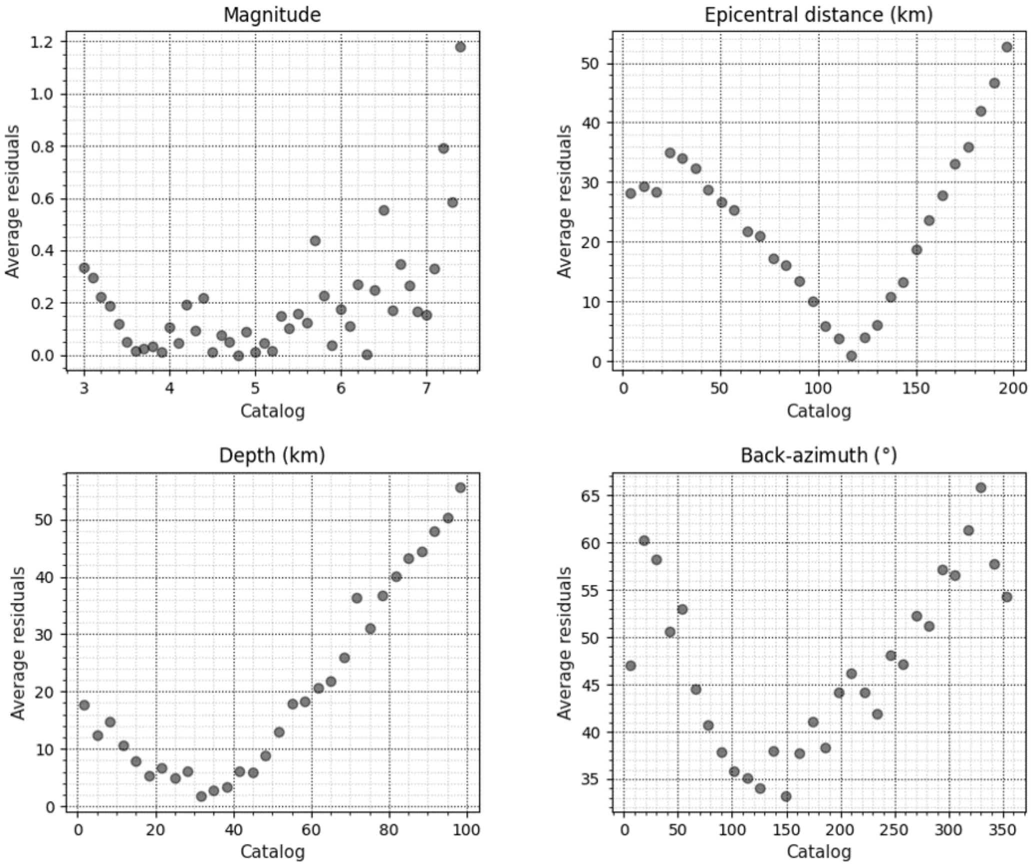

Figure 4. Average residuals ($|target_{pred} - target_{true}|$) for each target: Magnitude, epicentral distance, depth and back-azimuth, using the first 3 seconds of P-wave.

For the back-azimuth estimation, the best model uses 5 seconds 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.

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 dataset 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 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) (Hoshiba & Ozaki, 2014) and PEGSNet (Licciardi et al., 2022), and GNSS stations such as BEFORES (Minson et al., 2014) and G-larmS (Grapenthin et al., 2014b,

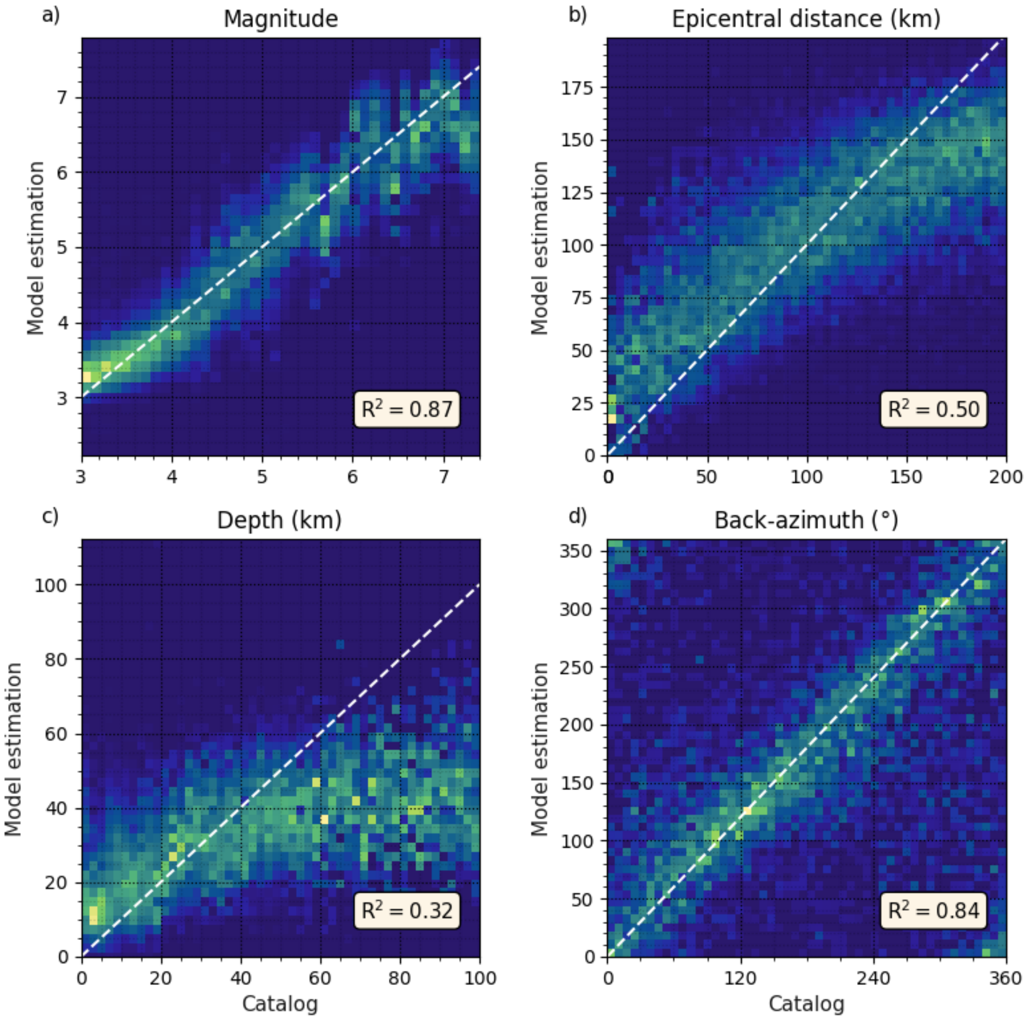

Figure 5. Estimated source parameters (magnitude, distance, depth, back-azimuth) using 3 s of records as a function of cataloged values.

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

Fig. 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). BE-

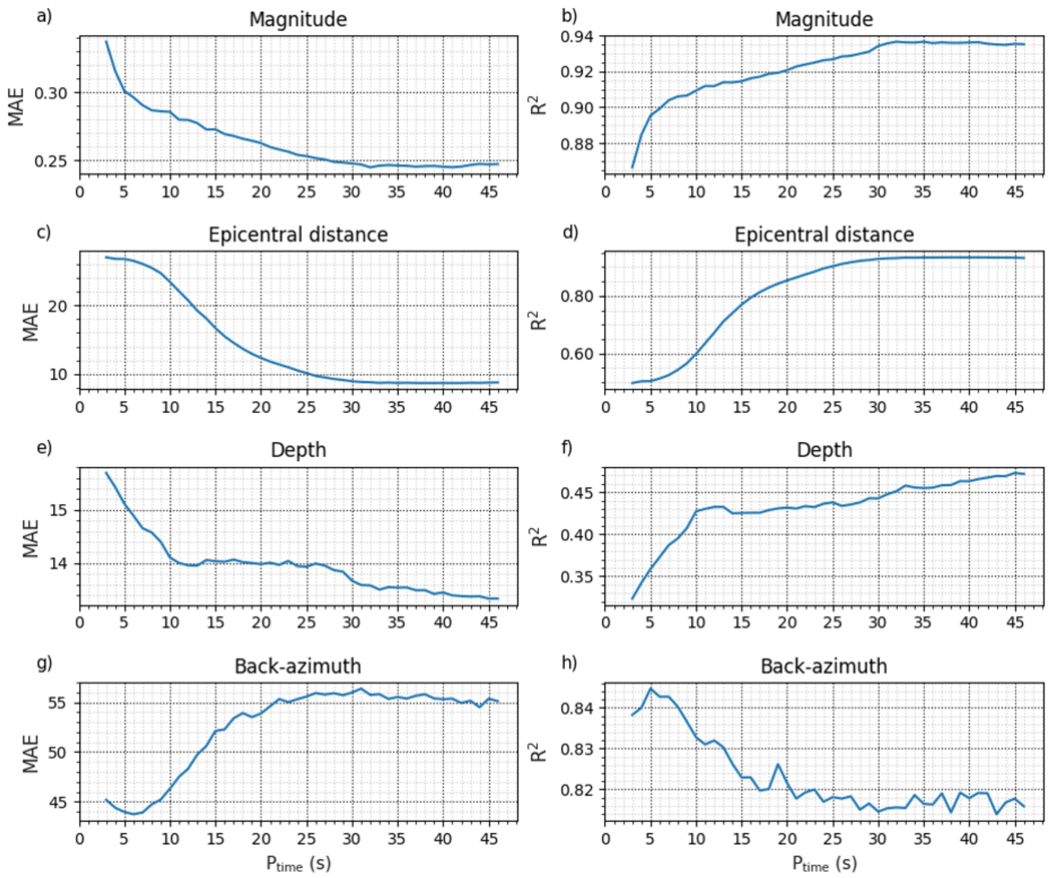

Figure 6. MAE and R2 results using 3 s to 46 s of P wave.

FORES makes its first estimate (M6.4) at 20 seconds after origin time (OT) when the true instantaneous magnitude is M7.3, outperforming the estimation of M5.7 by E3WS. However, one second 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 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 (Ptime) of each station, to compare them to the event's STF (Fig. 8b) given by the SCARDEC 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 dataset, 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

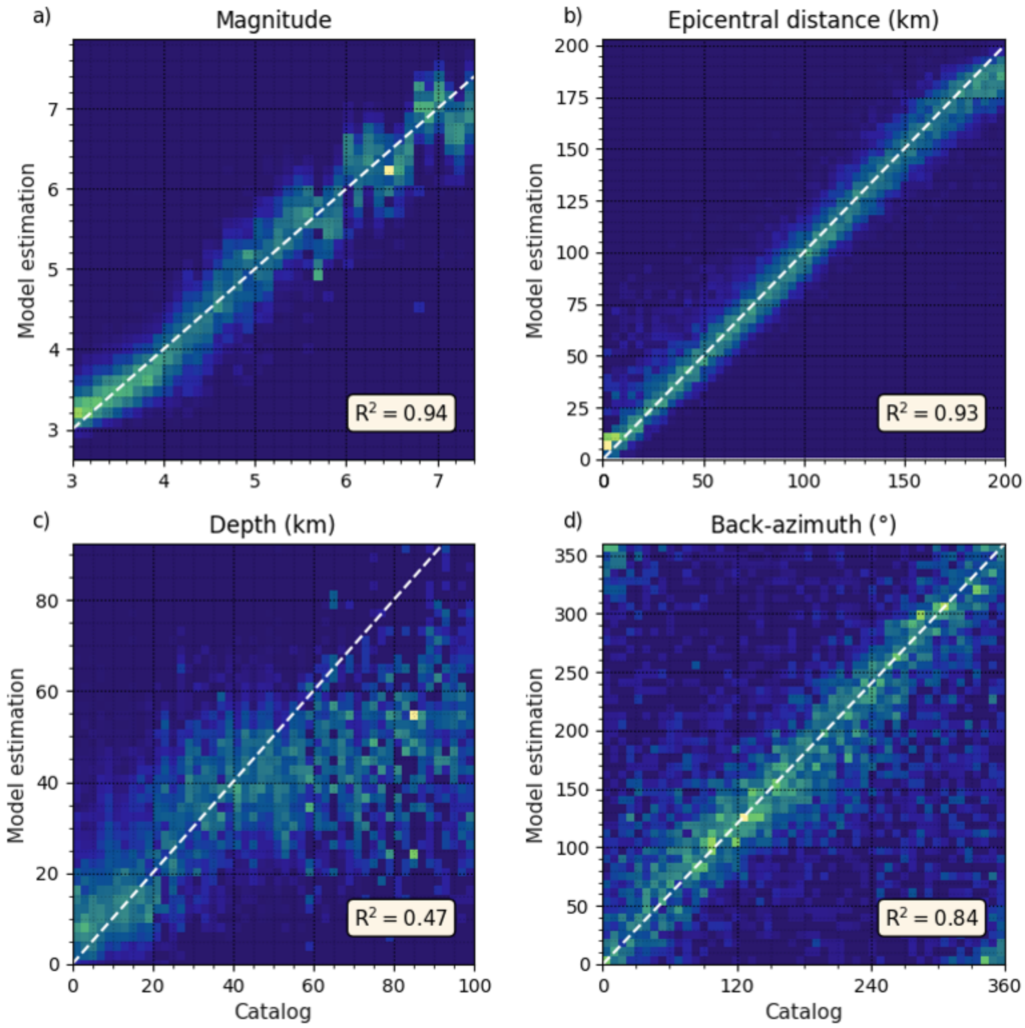

Figure 7. Same as Fig. 5 using 46 s of signal after P-wave arrival.

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 (Fig. 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). This 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.

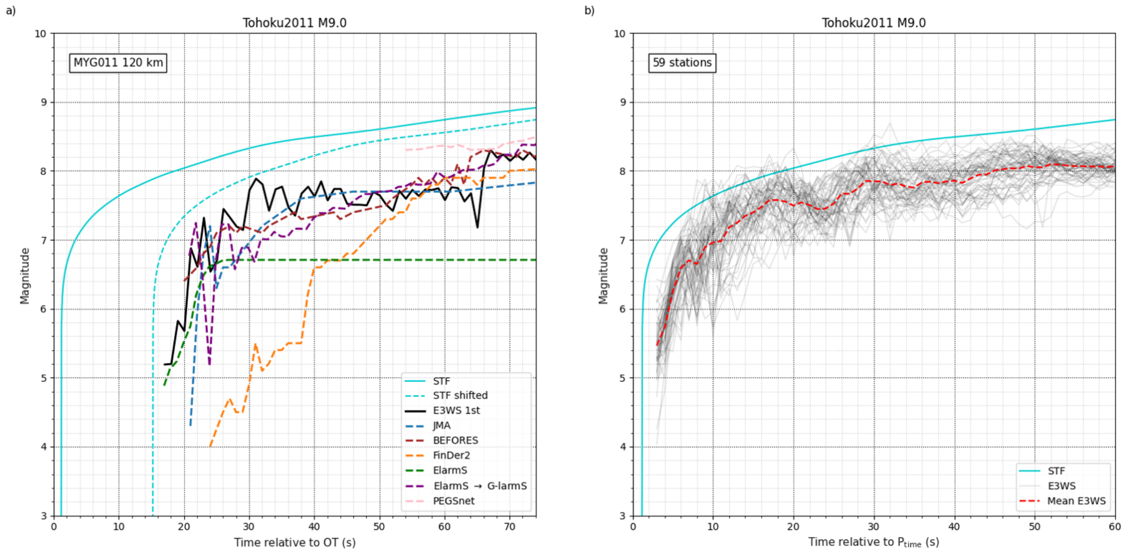

Figure 8. Real-time magnitude estimates for the 2011 M_W 9.0 Tohoku-Oki earthquake. (a) Magnitude evolution estimated by several EEW algorithms (see legend) as a function of time relative to the earthquake origin time. We also show the magnitude from the seismologically determined Source Time Function (STF) 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.

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.

E3WS does not trigger for 15 earthquakes (Table S1). The maximum magnitude of these missed events is 3.8, with a strong trade-off between magnitude and distance (Table S1). These magnitudes are below those that generate significant shaking in coastal Peru; they would not warrant an alert. E3WS triggers for 3 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 km, 396 km 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 > 1000 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.

Table 4. E3WS earthquake detections using 3 s of P wave in a continuous month (January 2022) at station SLN1.

M_{true}	M_{est}	Detection (s)	Warning time (s)
5.6	5.4	0.7	13.4
4.3	4.2	1.4	11.1
4.0	3.9	1.1	7.9
4.0	3.6	1.1	17.3
3.9	3.5	0.5	23.6
3.7	3.5	0.9	11.5
3.6	3.3	1.3	9.3
3.6	3.4	1.1	18.5
3.5	3.5	1.0	12.7
3.5	3.4	0.8	11.3
3.5	3.5	0.7	11.8
3.5	3.3	1.6	18.6
3.1	3.2	1.4	13.8
3.1	3.2	0.9	8.5

5 Discussion

5.1 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, i.e. it is a measure of the improvement in the estimates when using a particular attribute. 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.

The attributes that contribute the most to magnitude estimation, both using short and long portions of P wave, are the MFCC (Fig. 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), that share similarities with features that are widely used for magnitude estimation in other EEWS, such as P_d or τ_c. We hypothesize that the MFCC, by measuring energies on the Mel scale (a logarithmic frequency scale), 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, as is our approach, requires time integration, which is prone to amplify noise. Thus, it might preferable to not include them in E3WS. Indeed, our tests show better efficiency when using acceleration waveforms. Moreover, E3WS requires unclipped data for strong earthquakes as provided by accelerometers.

5.2 Comparative performance of E3WS and other EEWS

We compare the performance of E3WS with that of ElarmS (Brown et al., 2011), which estimates earthquake magnitude within the first 4 seconds of the 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.

We also test the performance of E3WS compared to UrEDAS. Lockman and Allen (2005) report results applying UrEDAS using stations containing at least 5 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 seconds 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 seconds, E3WS achieved a MAE of magnitude of 0.22, significantly better than UrEDAS with 4 seconds. For location, E3WS yields results similar to UrEDAS, with MAE of 14 km for hypocentral distance and 20° for back-azimuth. Using 4 seconds 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 link 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°.

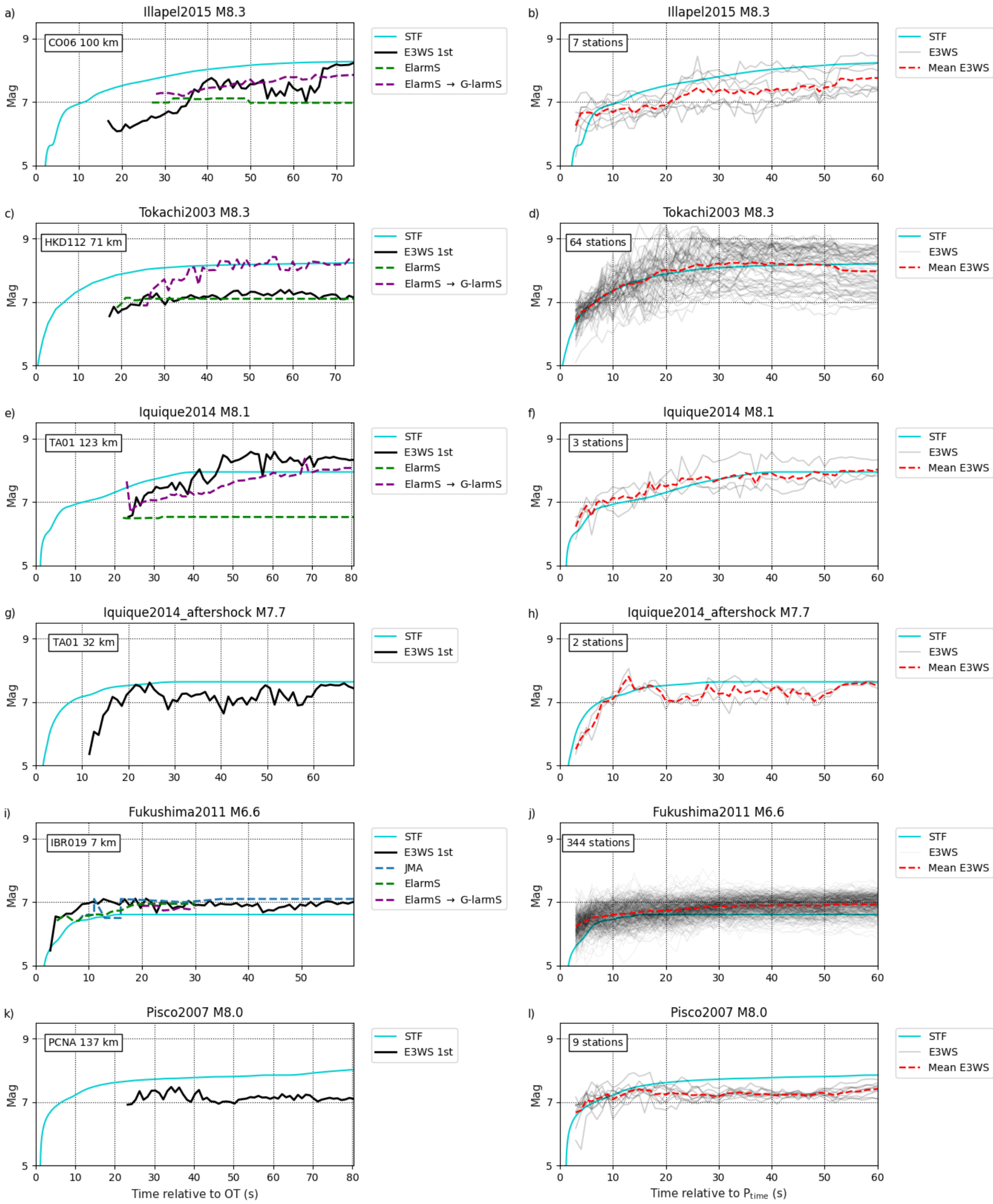

Figure 9. Magnitude estimates for the following earthquakes: 2015 M_W 8.3 Illapel, 2003 M_W 8.3 Tokachi-Oki, 2014 M_W 8.1 Iquique, 2014 M_W 7.7 Iquique aftershock, 2011 M_W 6.6 Fukushima aftershock, 2007 M_W 8.0 Pisco. Estimates are shown as a function of time relative to the earthquake's OT for the 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 with those obtained by other EEWS. On the right, we show all the estimates (gray), their mean (red), the moment function (the integral of the STF, light blue).

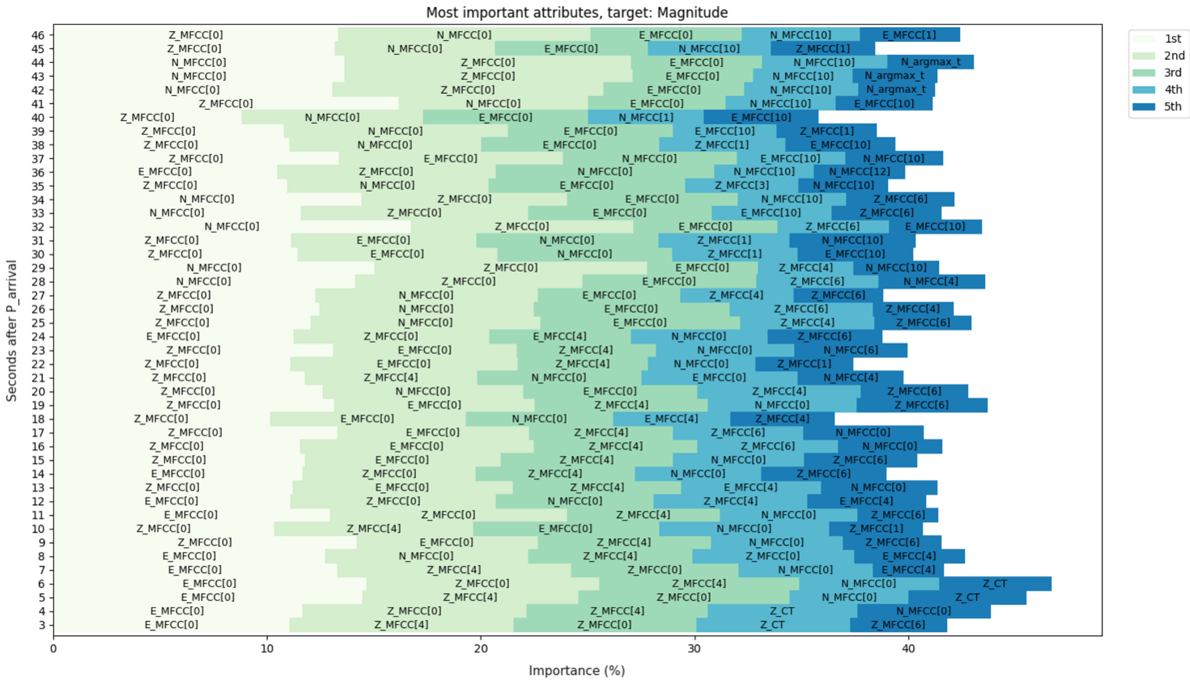

Figure 10. First (lightest color) to fifth (darkest color) most important features for magnitude estimation as a function of the P-wave window duration, from 3 to 46 s. For each time window, feature importance is based on the corresponding stacking model (see subsection 3.3), which consists of 10 XGB base models. Importance (%) shown is calculated as the gain mean plus STD 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.

6 Conclusion

We introduced E3WS, a set of Machine Learning algorithms using only 3 seconds of P-wave signal recorded by a single accelerometric station to detect, locate and estimate the magnitude of an earthquake. E3WS is made of 6 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 seconds of seismic recording, such as ElarmS and UrEDAS. 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

Waveforms and metadata used in this article were provided by the University of Chile downloaded by IRIS Web Services (`https://service.iris.edu/`), NIED K-NET, KiK-net, National Research Institute for Earth Science and Disaster Resilience, `doi:10.17598/NIED.0004` (`https://www.kyoshin.bosai.go.jp/`), the Stanford Earthquake Dataset (`https://github.com/smousavi05/STEAD`), and through petition to Instituto Geofísico del Perú (`https://www.gob.pe/igp`).

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.

References

Allen, R. M. (2007). The elarms earthquake early warning methodology and application across california. *Earthquake early warning systems*, 21–43.

Allen, R. M., & Kanamori, H. (2003). The potential for earthquake early warning in southern california. *Science*, *300*(5620), 786–789.

Allen, R. M., & Melgar, D. (2019). Earthquake early warning: Advances, scientific challenges, and societal needs. *Annual Review of Earth and Planetary Sciences*, *47*, 361–388.

Allen, R. V. (1978). Automatic earthquake recognition and timing from single traces. *Bulletin of the seismological society of America*, *68*(5), 1521–1532.

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.

Barrientos, S., & Team, N. S. C. C. (2018). The seismic network of chile. *Seismological Research Letters*, *89*(2A), 467–474.

Böse, M., Allen, R., Brown, H., Gua, G., Fischer, M., Hauksson, E., ... others (2014). Cisn 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.

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.

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.

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.

Cremen, G., Galasso, C., & Zuccolo, E. (2022). Investigating the potential effectiveness of earthquake early warning across europe. *Nature communications*, *13*(1), 1–10.

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.

Davis, S., & Mermelstein, P. (1980). Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *IEEE transactions on acoustics, speech, and signal processing*, *28*(4), 357–366.

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.

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.

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.

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.

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.

Lara, P. E. E., Fernandes, C. A. R., Inza, A., Mars, J. I., Métaxian, 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.

Licciardi, A., Bletery, Q., Rouet-Leduc, B., Ampuero, J.-P., & Juhel, K. (2022). Instantaneous tracking of earthquake growth with elastogravity signals. *Nature*, 1–6.

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.

Malfante, M., Dalla Mura, M., Métaxian, 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.

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.

Meier, M.-A., Ampuero, J., & Heaton, T. H. (2017). The hidden simplicity of subduction megathrust earthquakes. *Science*, *357*(6357), 1277–1281.

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.

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.

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.

Mousavi, S. M., Sheng, Y., Zhu, W., & Beroza, G. C. (2019). Stanford earthquake dataset (stead): A global data set of seismic signals for ai. *IEEE Access*, *7*, 179464–179476.

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.

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.

Nielsen, D. (2016). *Tree boosting with xgboost-why does xgboost win" every" machine learning competition?* (Unpublished master's thesis). NTNU.

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.

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.

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.

Yih-Min, W., Kanamori, H., Allen, R. M., & Hauksson, E. (2007). Determination of earthquake early warning parameters, τ c and pd, for southern california. *Geophysical Journal International*, *170*(2), 711–717.

Supporting Information for "Earthquake Early Warning using 3 seconds of records on a single station"

Pablo Lara[1,2], Quentin Bletery[1], Jean-Paul Ampuero[1], Adolfo Inza[2]

[1]Université Côte d'Azur, IRD, CNRS, Observatoire de la Côte d'Azur, Géoazur, France

[2]Instituto Geofísico del Perú, Lima, Perú

Contents of this file

1. Introduction

This supporting information includes the attributes used in this work, 9 supplementary figures and 1 supplementary table.

2. Attributes

We detail here the attributes computed to train the Machine Learning algorithms, with their corresponding domain and signal.

2.1. Attributes from 3-component seismograms

1. Maximum eigenvalue λ_1 of covariance matrix from the 3C seismogram.

2. Eigenvalue factor: ratio of the maximum eigenvalue to the sum of the remaining eigenvalues:

$$\lambda_{factor} = \lambda_1/(\lambda_2 + \lambda_3). \tag{1}$$

3. The 3 components of the eigenvector ν_1 associated with the maximum eigenvalue λ_1.

2.2. Attributes from each channel

N denotes the number of samples per channel within the time window. F_s denotes the sampling rate, in Hz. The envelope of the signal s is defined as $e = |s + iH\{s\}|$, where H is the Hilbert transform.

2.2.1. Time-domain attributes

4. Maximum energy of the preprocessed signal:

$$E_{max} = \max(s^2). \tag{2}$$

5. Time at which the maximum energy is reached:

$$t_{E_{max}} = argmax(s^2). \tag{3}$$

6. Total energy:

$$E_{ne} = \sum_{n=1}^{N} s^2[n]. \tag{4}$$

7. Energy centroid time:

$$CT = \frac{1}{E_{ne}} \sum_{n=1}^{N} n \ s^2[n]. \tag{5}$$

8. Bandwidth, characteristic duration, variance around the energy centroid:

$$BW_t = \sqrt{\frac{\sum_{n=1}^{N} (n - CT)^2 s^2[n]}{E_{ne}}}. \tag{6}$$

9. Skewness around bandwidth:

$$Skewness_{BW_t} = \begin{cases} \sqrt{SK_{pre}}, & \text{if } SK_{pre} \geq 0 \\ -\sqrt{-SK_{pre}}, & \text{otherwise.} \end{cases}$$

where

$$SK_{pre} = \frac{\sum_{n=1}^{N} (n - CT)^3 s^2[n]}{E_{ne}\ BW_t^3}. \tag{7}$$

10. Kurtosis around bandwidth:

$$Kurtosis_{BW_t} = \sqrt{\frac{\sum_{n=1}^{N} (n - CT)^4 s^2[n]}{E_{ne}\ BW_t^4}}. \tag{8}$$

11. Mean envelope:

$$< env >= \frac{1}{N} \sum_{k=1}^{N} e[k]. \tag{9}$$

12. Ratio of maximum amplitude envelope to its mean:

$$RMM_t = \frac{\max(e)}{< env >}. \tag{10}$$

13. Standard deviation of the envelope:

$$STD_{env} = \sqrt{\frac{\sum_{k=1}^{N} (e[k] - < env >)^2}{N}}. \tag{11}$$

14. Skewness of the envelope:

$$Skewness_{env} = \frac{1}{N} \sum_{k=1}^{N} \left(\frac{e[k] - < env >}{STD_{env}} \right)^3. \tag{12}$$

15. Kurtosis of the envelope:

$$Kurtosis_{env} = \frac{1}{N} \sum_{k=1}^{N} \left(\frac{e[k] - < env >}{STD_{env}} \right)^4. \tag{13}$$

16. Threshold-crossing rate of the envelope signal: how many times per second the signal envelope crosses the threshold of 80% of its maximum amplitude:

$$TCR_t = \frac{count(r[n]r[n-1] < 0)}{N/F_s}, \tag{14}$$

where:

$$r = e/\max(e) - 0.8. \tag{15}$$

17. Fraction of envelope samples that exceed a threshold of 80% of the envelope maximum:

$$fract(TCR_{env}) = count(e \geq 0.8 \max(e))/N. \tag{16}$$

18. Shannon entropy of the envelope, with $N_{bins} = 200$.

$$Shannon_{env} = -\sum_{i=1}^{N_{bins}} Prob_e[i] \log_2(Prob_e[i]), \tag{17}$$

where:

$$Prob_e[i] = Histogram(e, N_{bins}). \tag{18}$$

19. Renyi entropy of the envelope, with $\alpha = 2$.

$$Renyi_{env} = \frac{\log_2 \sum_{i=1}^{N_{bins}} Prob_e^{\alpha}[i]}{1-\alpha}, \tag{19}$$

20. Zero crossing rate, how many times per second the signal s changes sign:

$$ZCR_t = \frac{count(s[n]s[n-1] < 0)}{N/F_s} \tag{20}$$

2.2.2. Spectral-domain attributes

Attributes extracted from $p = PSD(s)$, the Welch's Power Spectral Density of the signal s. Here N denotes the number of frequency samples in the spectrum up to the Nyquist frequency $F_s/2$.

21. Mean PSD:

$$< PSD >= \frac{1}{N} \sum_{k=1}^{N} p[k]. \tag{21}$$

22. Maximum spectral energy:

$$PSD_{max} = \max(p). \tag{22}$$

23. Frequency index of maximum spectral energy:

$$f_{PSD_{max}} = argmax(p). \tag{23}$$

24. Centroid frequency of the spectrum:

$$CF = \frac{\sum_{k=1}^{N} k \; p[k]}{\sum_{k=1}^{N} p[k]}. \tag{24}$$

25. Frequency bandwidth, variance around the spectral centroid:

$$BW_f = \sqrt{\frac{\sum_{k=1}^{N} (k - CF)^2 \; p[k]}{\sum_{k=1}^{N} p[k]}}. \tag{25}$$

26. Skewness of the spectrum:

$$Skewness_{BW_f} = \begin{cases} \sqrt{SK_{pre}}, & \text{if } SK_{pre} \geq 0 \\ -\sqrt{-SK_{pre}}, & \text{otherwise.} \end{cases}$$

where

$$SK_{pre} = \frac{\sum_{k=1}^{N} (k - CF)^3 \; p[k]}{BW_f^3 \sum_{k=1}^{N} p[k]}, \tag{26}$$

27. Kurtosis of the spectrum:

$$Kurtosis_{BW_f} = \sqrt{\frac{\sum_{k=1}^{N}(k - CF)^4 \; p[k]}{BW_f^4 \sum_{k=1}^{N} p[k]}}. \tag{27}$$

28. Standard deviation of the PSD:

$$STD_{PSD} = \sqrt{\frac{\sum_{k=1}^{N}(p[k]- < PSD >)^2}{N}}. \tag{28}$$

29. Skewness of PSD:

$$Skewness_{PSD} = \frac{\sum_{k=1}^{Na}\left(\frac{p[k]- < PSD >}{STD_{PSD}}\right)^3}{N}. \tag{29}$$

30. Kurtosis of PSD:

$$Kurtosis_{PSD} = \frac{\sum_{k=1}^{N}\left(\frac{p[k]- < PSD >}{STD_{PSD}}\right)^4}{N}. \tag{30}$$

31. Shannon entropy, with $N_{bins} = 50$:

$$Shannon_{PSD} = -\sum_{i=1}^{N_{bins}} Prob_p[i] \; \log_2(Prob_p[i]), \tag{31}$$

where:

$$Prob_p[i] = Histogram(p[k], N_{bins}). \tag{32}$$

32. Renyi entropy, with $\alpha = 2$:

$$Renyi_{PSD} = \frac{\log_2 \sum_{i=1}^{N_{bins}} Prob_p^{\alpha}[i]}{1 - \alpha}. \tag{33}$$

33. Ratio of maximum PSD amplitude to its mean.

$$RMM_f = \frac{\max(p)}{< PSD >}. \tag{34}$$

34. Threshold-crossing rate of the PSD, how many times the PSD crosses a threshold of 40% of its maximum amplitude:

$$TCR_f = \frac{count(r[k]r[k-1] < 0)}{N/F_s}, \tag{35}$$

where:

$$r = PSD/\max(PSD) - 0.4 \tag{36}$$

35. Relative number of samples that exceed a threshold of 40% of its maximum.

$$fract(TCR_{PSD}) = count(p \geq 0.4\max(p))/N. \tag{37}$$

2.2.3. Cepstral-domain attributes

36. The 13 first mel-frequency cepstrum coefficients (MFCC):

$$MFCC[m] = DCT\{log[\sum\{|F\{s\}|^2\Lambda_m\}]\}, \tag{38}$$

where DCT is the Discrete Cosine Transform, $F\{.\}$ is the Discrete Fourier Transform, and Λ is a triangular filter bank function linearly spaced from 1 to 45 Hz in a Mel scale. In this work, we use $m = 26$ filter banks, and are compute as in (Kopparapu & Laxminarayana, 2010).

3. Figures

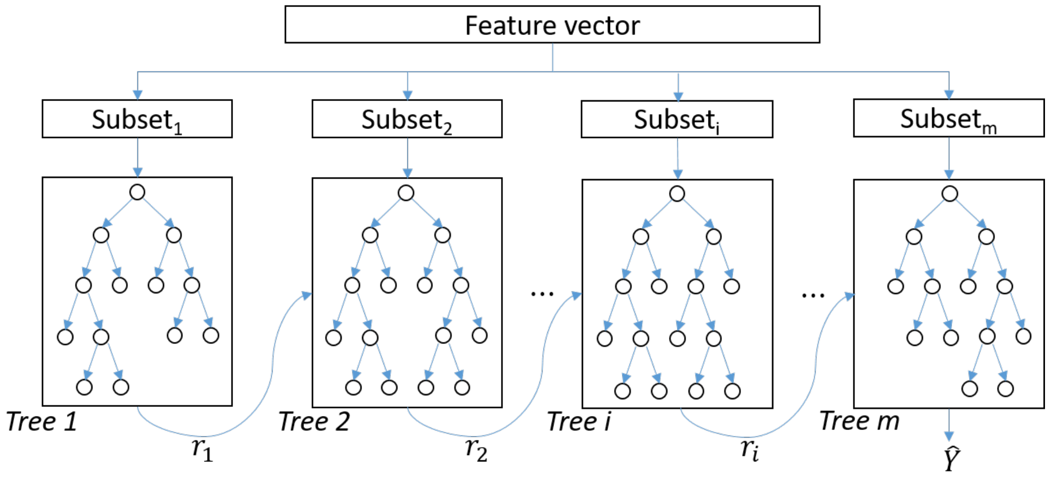

Figure S1. General architecture XGB.

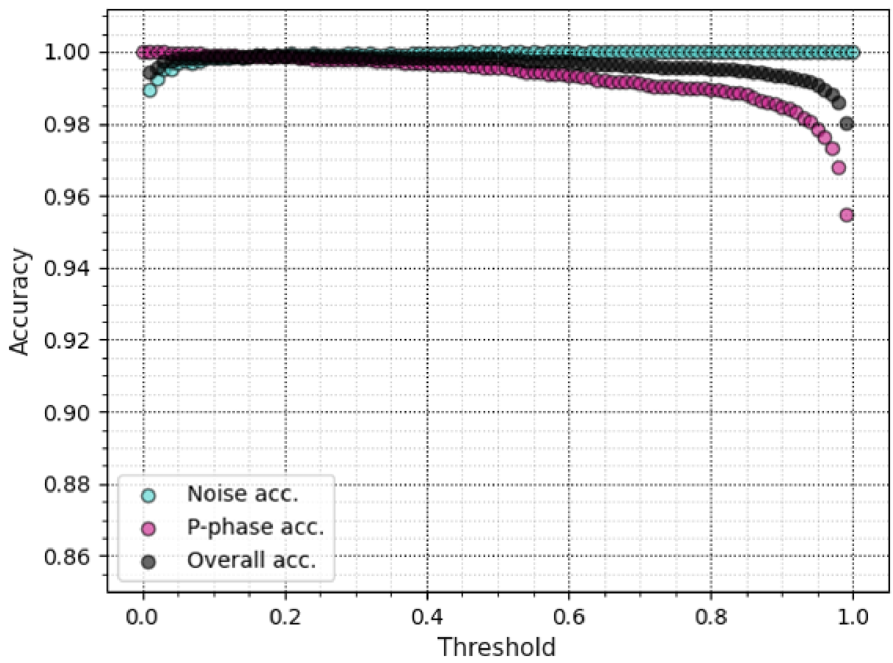

Figure S2. Accuracy of noise and earthquake classification, using different thresholds.

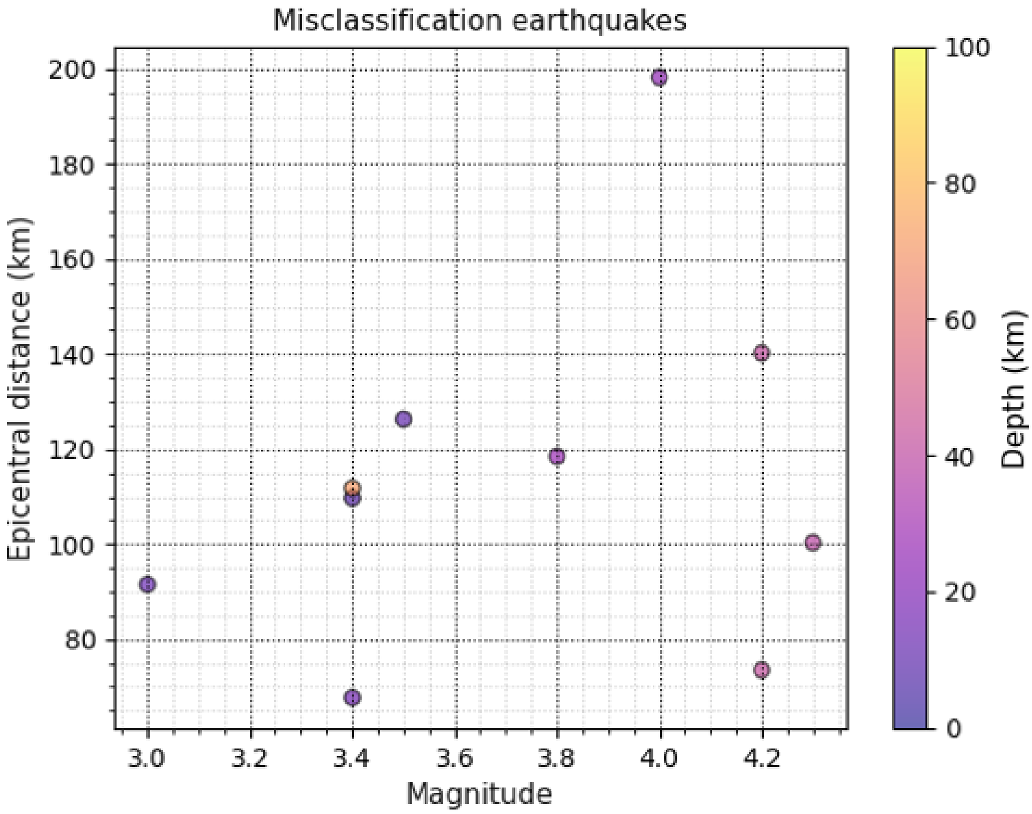

Figure S3. Magnitude, epicentral distance and depth of the misclassified signals shown in Fig. S4.

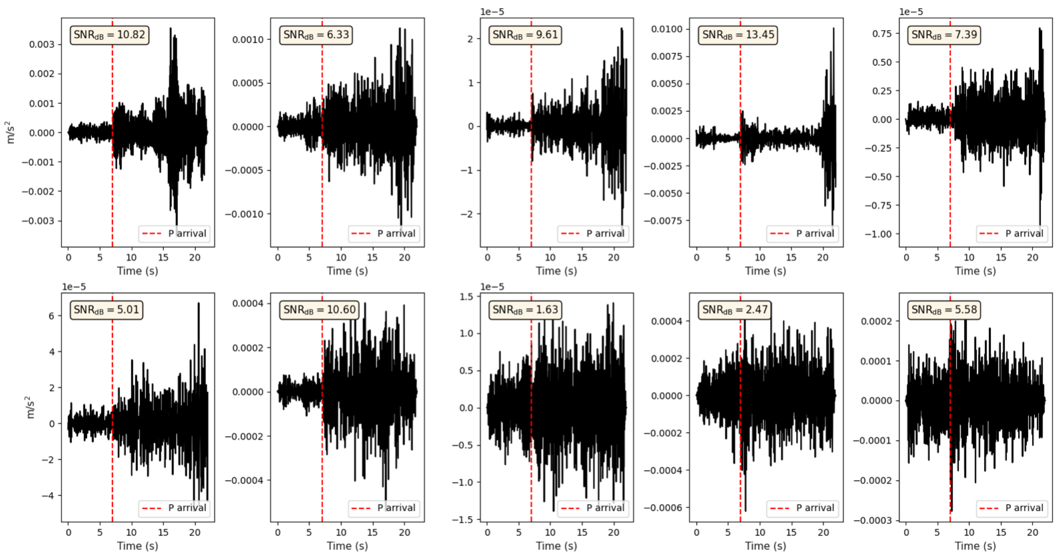

Figure S4. Earthquake signals misclassified as noise, and their signal-to-noise ratios (SNR).

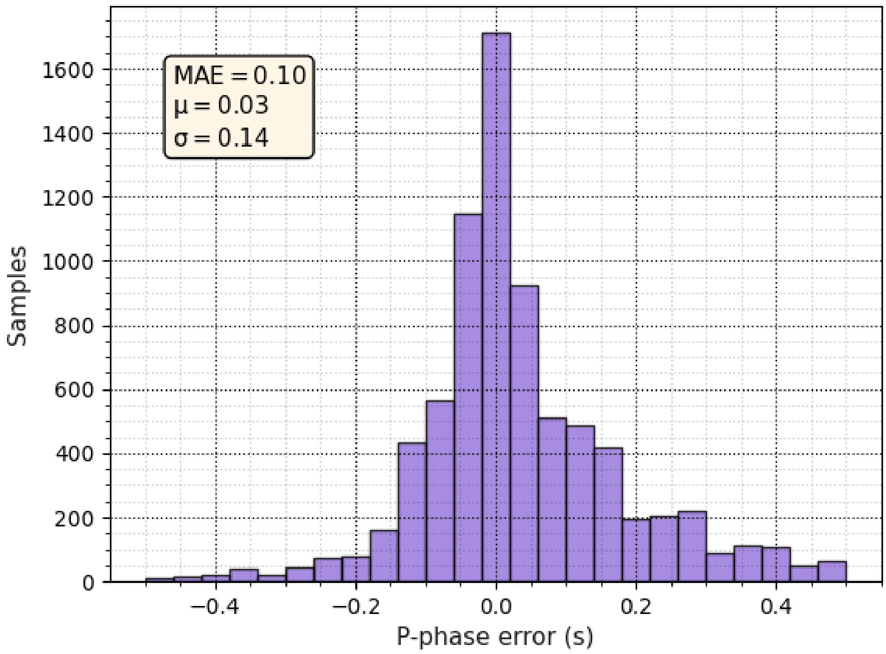

Figure S5. histogram of errors between the true and predicted P-phase arrival times.

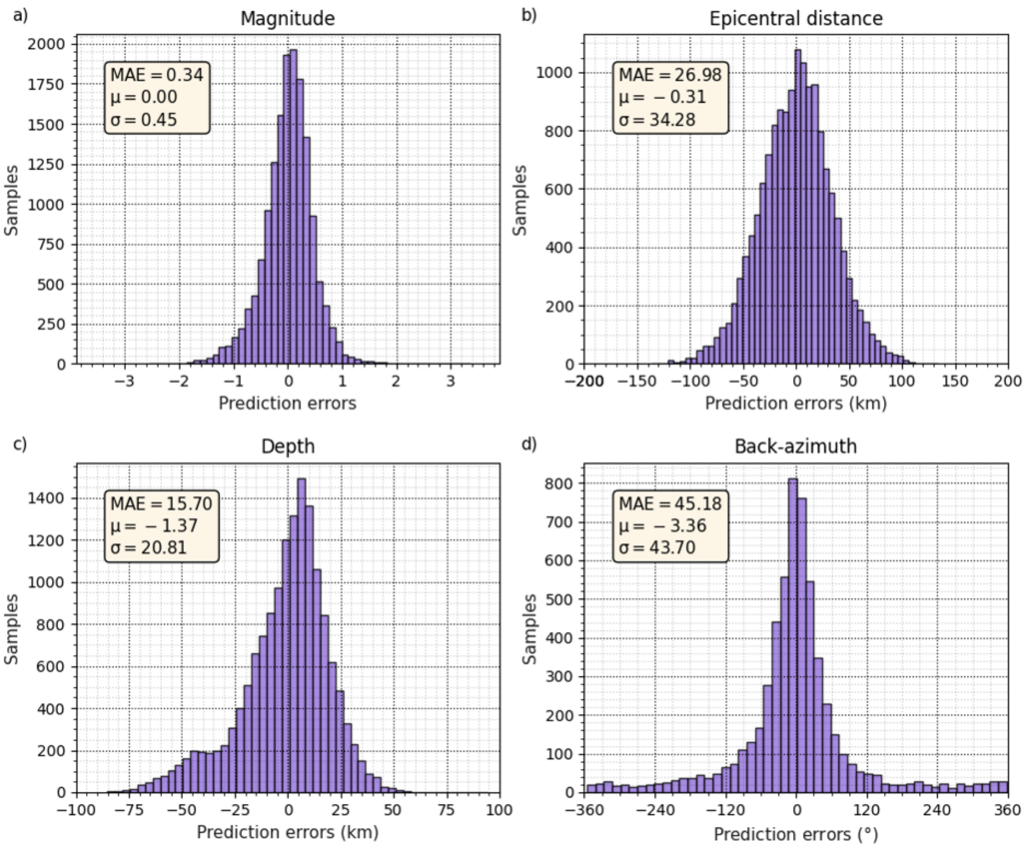

Figure S6. Histogram of the errors in the source characterization predictions using 3 s of P-wave.

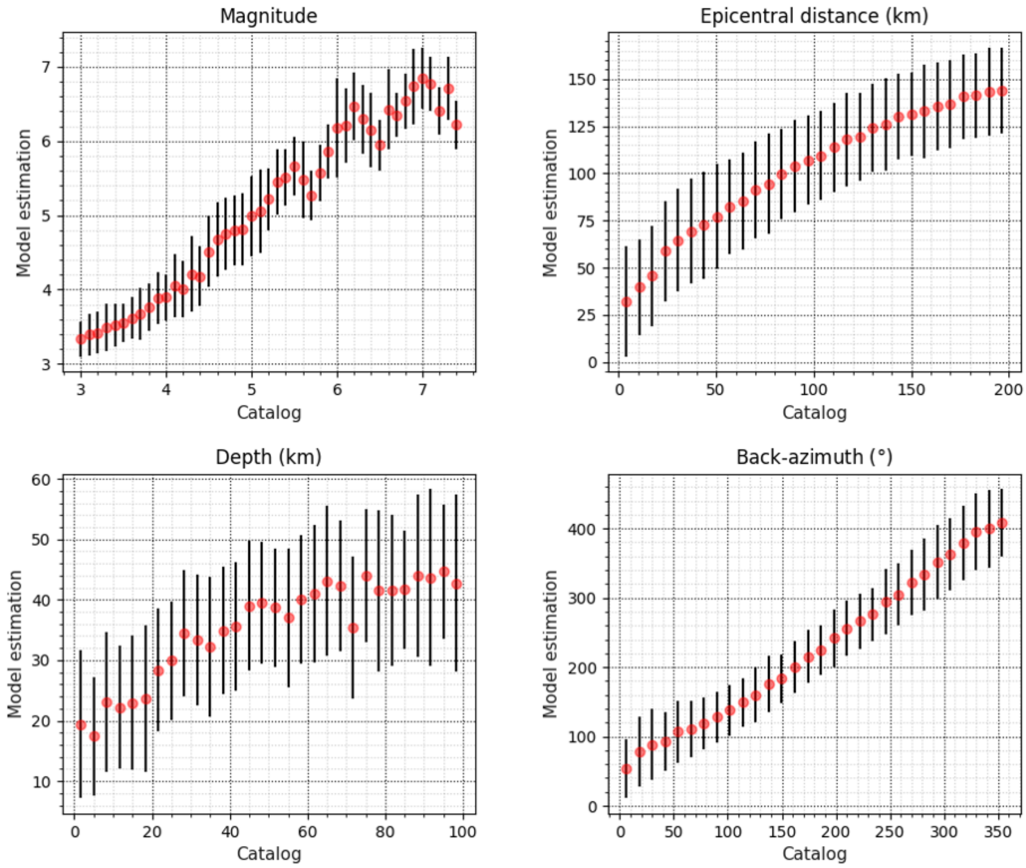

Figure S7. Mean (circle) and STD (bar) predictions per bin using 3 s of P-wave.

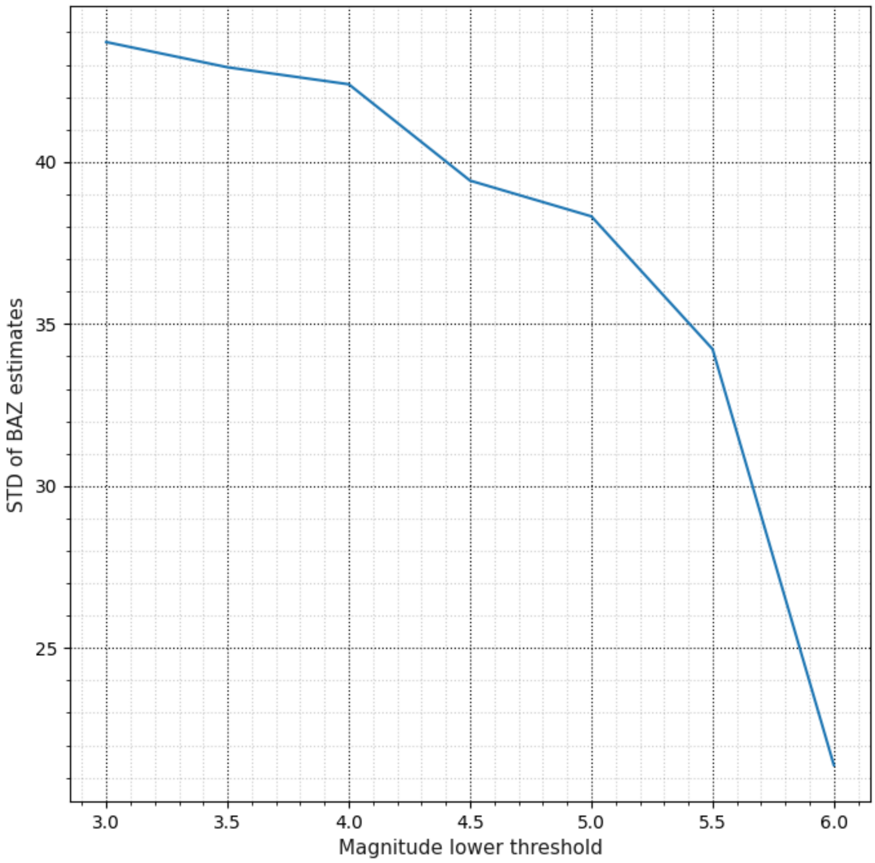

Figure S8. STD of the back-azimuth estimates, using different lower thresholds of magnitude.

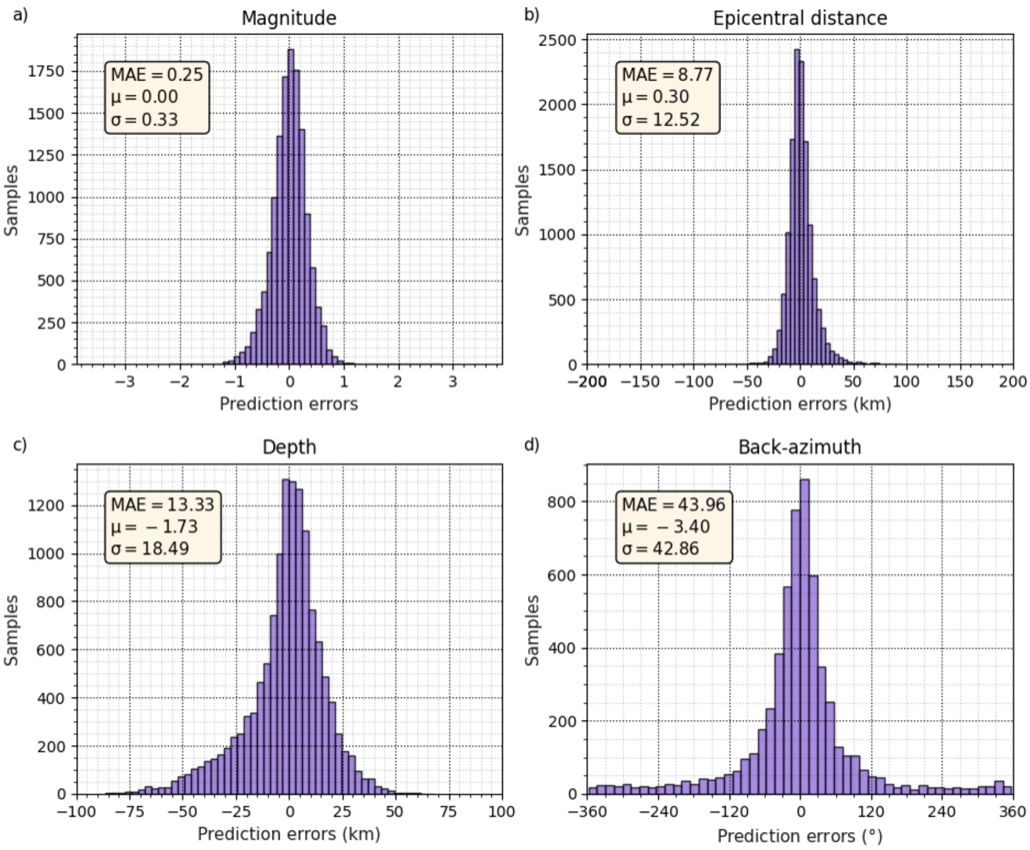

Figure S9. Histogram of the errors in the predictions using 46 s of P-wave.

Table S1. Real-time earthquake detection by E-EEWS using 3 s of P-wave in a continuous month.

Magnitude	Distance (km)	Depth (km)
3.1	175	13
3.5	179	56
3.3	195	85
3.2	145	78
3.2	173	54
3.3	89	50
3.3	98	86
3.2	163	17
3.7	162	19
3.3	159	84
3.8	155	10
3.0	47	35
3.4	189	88
3.5	138	47
3.5	97	49

References

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).