Geophysical Journal International

Geophys. J. Int. (2024) **236,** 675–687
Advance Access publication 2023 November 10
GJI Seismology

https://doi.org/10.1093/gji/ggad443

Real-time prediction of distance and PGA from *P*-wave features using Gradient Boosting Regressor for on-site earthquake early warning applications

Antonio Giovanni Iaccarino [1], Amalia Cristofaro,[1] Matteo Picozzi [1], Daniele Spallarossa[2] and Davide Scafidi[2]

[1] *Physics Department 'E. Pancini', University of Naples Federico II, Naples* 80126*, Italy. E-mail: antoniogiovanni.iaccarino@unina.it*
[2] *DISTAV, University of Genoa, Genoa* 16132*, Italy*

Accepted 2023 November 8. Received 2023 September 12; in original form 2022 October 19

SUMMARY

On-site earthquake early warning (EEW) systems represent an important way to reduce seismic hazard. Since these systems are fast in providing an alert and reliable in the prediction of the ground motion intensity at targets, they are particularly suitable in the areas where the seismogenic zones are close to cities and infrastructures, such as Central Italy.

In this work, we use Gradient Boosting Regressor (GBR) to predict peak ground acceleration (PGA), and hypocentral distance (D) starting from P-wave features. We use two data sets of waveforms from two seismic sequences in Central Italy: L'Aquila sequence (2009) and the Amatrice–Norcia–Visso sequence (2016–2017), for a total of about 80 000 three-component waveforms. We compute 60 different features related to the physics of the earthquake using three different time windows (1 s, 2 s and 3 s). We validate and train our models using the 2016–2017 data sets (the bigger one) and we test it on the 2009 data set.

We study the performances of GBR predicting D and PGA in terms of prediction scores, finding that the models can well predict both targets even using 1 s window, and that, as expected, the results improve using longer time windows. Moreover, we perform a residual analysis on the test set finding that the PGA can be predicted without any bias, while the D prediction presents a correlation with the moment magnitude.

In the end, we propose a prototype for a probabilistic on-site EEW system based on the prediction of D and PGA. The proposed system is a threshold-based approach and it releases an alert on four possible levels, from 0 (far and small event) to 3 (close and strong event). The system computes the probability related to each alert level. We test two different set of thresholds: the Felt Alert and the Damage Alert. Furthermore, we consider the lead time (LT) of the PGA to distinguish between useful alerts (positive LT) and Missed Alerts (MA). In the end, we analyse the performance of such a system considering four possible scenarios: Successful Alert (SA), Missed Alert (MA), Overestimated Alert (OA) and Underestimated Alert (UA). We find that the system obtains SA rate about 80 per cent at 1 s, and that it decreases to about 65 per cent due to the increase in MA. This result shows how the proposed system is already reliable at 1 s, which would be a huge advantage for seismic prone regions as Central Italy, an area characterized by moderate-to-large earthquakes ($M_w < 7$).

Key words: Europe; Machine learning; Earthquake early warning; Earthquake hazards.

1 INTRODUCTION

In the framework of natural phenomena, earthquakes are one of the most impressive. Large earthquakes have a strong impact on society: many human lives are loss due to these catastrophic events and modern society is severely affected by the enormous damages that they cause in the hit areas, where economic losses can easily reach the order of billions of euros.

Nowadays, real-time seismology is one of the most effective actions to mitigate seismic risk by decreasing the exposure of people and targets to the potentially destructive effects of an earthquake (Satriano *et al.* 2011; Picozzi et al., 2013). In particular, the pro-

gresses in real-time seismic data processing and fast telemetry have allowed the development of Earthquake Early Warning Systems (EEWSs; Kanamori 2005). These systems are aimed at detecting and characterizing in real-time the size of an earthquake and at providing a rapid alert notification of the ongoing event to vulnerable targets in the vicinity of the source region, only using information extracted from the first seconds of the available seismic waveforms (Wu & Kanamori 2005b, Wu & Zhao 2006; 2008).

The first idea of such a system dates back to 1868 by D.J. Cooper (Cooper 1868), which proposed the introduction of an alert system after the earthquake of magnitude 7 that strongly damaged the city of San Francisco. While Cooper's idea has never been implemented, nowadays there are few operating, or under testing, EEWSs around the world, such as in Japan, USA, Italy and Mexico (Espinosa-Aranda *et al.* 2009; Zollo *et al.* 2009; Allen & Melgar 2019; Allen *et al.* 2019; Cremen & Galasso 2020). EEWSs use independently or combine different strategies for extracting different pieces of information about the earthquake and the ground motion at different sites by exploiting extended seismic networks or single-station systems (Hoshiba *et al.* 2008; Zollo *et al.* 2014; Colombelli *et al.* 2015; Picozzi *et al.* 2015a, 2015b, 2015c; Caruso *et al.* 2017; Festa *et al.* 2018; Spallarossa *et al.* 2019; Song *et al.* 2022).

In this work, we focus on the on-site EEW approach, which consists of single seismic station, or a small array of sensors, placed directly or nearby a target site (Satriano *et al.* 2011). Typically, on-site EEWSs extract information from the analysis of *P* waves to predict the ground motion intensity at targets (Brondi *et al.* 2015; Caruso *et al.* 2017; Iaccarino *et al.* 2020). EEW on-site systems can provide faster alerts at target sites for small hypocentral distances with the respect to EEWS operating at regional scale (Satriano *et al.* 2011). Furthermore, predictions of regional EEW systems are hampered by the uncertainties of ground motion prediction equations, which is larger than the one of empirical models relating the *P* waves and *S* waves ground motion predictions of the ground motion intensity at the target (Bindi *et al.* 2011; Spallarossa *et al.* 2019). For this reason, on-site EEW systems are generally considered to provide more reliable predictions of the ground motion intensity at targets than regional ones.

A key issue of EEWS is the amount of time available for alerting a target site before it is hit by destructive seismic waves. Clearly, the greater the alerting time, the greater the effectiveness of the remedial measures aimed at protecting the target itself. We indicate as 'lead-time' the time interval between the arrival of the *S*-wave and the alert time, both measured at the target. Obviously, the lead-time depends on the distance between the target and the hypocenter (Satriano *et al.* 2011).

So far, EEW studies have been mostly based on features extracted from limited signal windows (e.g. 1, 2 or 3 seconds) of the *P* waves, assuming inner relationships between these features and the strong motion parameters. Many studies use the maximum amplitude of the *P*-wave displacement, P_d, to predict the Peak Ground Velocity, PGV, (Wu & Kanamori 2005b; Zollo *et al.* 2010; Caruso *et al.* 2017). Brondi *et al.* (2015) proposed the Integral of squared Velocity, IV^2, measured on a 3 s *P*-wave window as a proxy for predicting both PGV and the Housner intensity, I_H (Housner 1952), with this latter commonly used in seismic engineering for damage assessments. Caruso *et al.* (2017) proposed an EEWS that combines P_d and predominant period (τ_c) to predict PGV, magnitude and a distance classification.

Nowadays, the development of automatic learning algorithms, such as artificial intelligence and machine learning, can give us the chance to optimize the extraction of information from the data. This can further pave the way to new lines of investigation to refine well-established techniques or introduce innovative approaches. In the last decades, new techniques have been proposed by many authors in seismology using machine and deep learning to study earthquakes in real-time. Many authors proposed the use of deep-learning techniques on waveforms to provide phase picking, earthquake location, magnitude and earthquake characterization (Zhu & Beroza 2019; Mousavi & Beroza 2020; Mousavi *et al.* 2020; Kuang *et al.* 2021; Münchmeyer *et al.* 2021). Other authors prefer to use seismic features extracted from waveforms to directly insert physics principles within the models. Among others, Böse *et al.* (2012) used a neural network (NN) to predict magnitude, distance and PGV using 9 *P*-wave features measured from one station and considering progressive time windows. Ochoa *et al.* (2018) used Support Vector Machine Regressor on 25 real-time features to predict magnitude using 5 s, 10 s and 15 s. Furthermore, Münchmeyer *et al.* (2020) used gradient boosting on 6 features, automatically corrected by distance, to predict magnitude. Hsu & Huang (2021) used a Convolutional Neural Network on 3 s of *P* waves both in time and frequency domain to predict PGA in an on-site EEW framework using data from single station. Iaccarino *et al.* (2021) compared the results of different machine learning regressors on 9 features computed at different time windows from in-building data to predict structural Drift in real-time.

In this work, we explore the feasibility of using machine learning regression as a tool for the real time prediction of the hypocentral distance (hereinafter, *D*) and the peak ground motion acceleration (hereinafter, PGA or ground motion intensity) considering as input features extracted from *P* waves time windows. We use a big data set of waveforms from two seismic sequences happened in Central Italy, the 2009 L'Aquila sequence and the 2016–2017 Central Italy sequence. It is important to note that the Italian context is characterized by seismogenic zones placed close to cities and infrastructures. Therefore, despite the seismicity is characterized by moderate-to-large earthquakes ($M_w < 7$), it can cause huge human and economic losses. These characteristics make this area particularly suitable for an on-site EEW system.

Once we derive and discuss the performance of models connecting the *P* waves information to the distance and the ground motion intensity, we also test the on-site predictions within a framework of a prototypal, probabilistic alert system to be implemented in an on-site, threshold based EEWS.

2 DATA SET

Our analysis is based on a data set generated using the RSN-Picker2 (Spallarossa *et al.* 2014; Scafidi *et al.* 2018), which includes a chain of modular iterative algorithms (i.e. an automatic phase detector, a picker and a locator) aimed at improving recall, precision and accuracy of automatic picking and location procedures for a real-time monitoring of earthquakes and nontectonic events. To be consistent with the Italian official seismic catalog, the information about earthquake locations is extracted from the INGV (National Institute for Geophysics and Volcanology) bulletin (see Data and Resources).

The source parameters (i.e. seismic moment, moment magnitude, local magnitude, seismic radiated energy) are estimated using features extracted directly from waveforms (Picozzi *et al.* 2017; Bindi *et al.* 2018; Spallarossa *et al.* 2021a). They are provided together with the ground motion intensity by the service RAMONES—Rapid Assessment of MOmeNt and Energy Service—(Spallarossa

et al. 2021b), which exploits raw data retrieved from the European Integrated Data Archive (ORFEUS-EIDA, https://www.orfeus-eu.org/data/eida/) and the Italian Civil Protection, DPC (http://ran.protezionecivile.it/EN/index.php; Gorini *et al.* 2010), strong motion network. Most of the traces in the data set are recorded by accelerometers (≈80 per cent), while the remaining by velocimeters. The stations belong to the INGV national seismic network and to the RAN-Italian strong motion network (see Data and Resources).

Our data set is in time made by two subsets: i) the first including 375 events (with M_w in the range [1.8; 6.3]) recorded at 217 different stations (more than 10 600 three-components waveforms) from 2009 (hereinafter, we will refer to this set as S09); the other including 840 events (with M_w in the range [1.8; 6.5]) recorded at 605 different stations (more than 69 000 waveforms) from the period between January 2016 and September 2019 (hereinafter, we will refer to this set as S1619). In the end, we select only data with hypocentral distance lower than 150 km because further data would be useless for EEW purposes considering the Italian seismicity.

The data set S1619 is dominated by earthquakes of the 2016–2017 Central Italy seismic sequence (Fig. 1a), which includes three main events: the M_w 6.0, 24 August 2016, Amatrice event (orange star); the M_w 5.9, 26 October 2016, Visso earthquake (yellow star); and M_w 6.5, 30 October 2016, Norcia earthquake (red star). The data set S09 is instead mainly formed by the seismic sequence related to the M_w 6.1, 4 April 2009, L'Aquila earthquake (Fig. 1b, red star).

Since S1619 is significantly larger than S09, we will use the former as training set, while S09 is used as testing set.

Although the two relative main sequences are geographically separated, with S09 that is slightly southernmost than S1619, the two regions belong to the same extensional sector of the Apennines, and therefore the earthquakes have similar characteristics (e.g. hypocentral depth, focal mechanism).

Furthermore, we divide S1619 in two subsets, $S1619_T$ (80 per cent) and $S1619_V$ (20 per cent). The first one is the proper training set, while the other is the validation set that we use for the tuning of the machine learning (ML) techniques. The splitting is done using the Scikit-Learn function 'train_test_split' in Python (Pedregosa *et al.* 2011) and stratifying data by the station magnitude (the magnitude computed using the ground motion amplitude measured at the seismic station). In Supporting Information Fig. S1, we show the distributions of station magnitude, distance and PGA_H for the three data sets. From the latter figure, we can see that the majority of data refer to events with magnitude ≤ 4 and $D \leq 50$ km.

3 METHOD

3.1 Features

We analyze the waveforms belonging to the two data sets to extract features (Table 1) that will then be used as input for the ML regressor. The features are calculated considering the signal in different domains: in displacement, $D(t)$, velocity, $V(t)$, and acceleration, $A(t)$. We estimate 15 features, some of them derived from the analysis of signals recorded on the vertical component, Z, others from the north and east components, N and E. Moreover, we also estimate features from the geometrical mean of the horizontal components (hereinafter indicated as H). Hence, we have in total 60 features that we compute at three different time windows 1 s, 2 s and 3 s.

We think each feature to be connected to the source physics in different ways. As example, the peak amplitude values from *P* waves (P_d, P_v and P_a) are directly linked to the peak amplitudes on *S* waves. The signal-to-noise ratios (SNR_a, SNR_v and SNR_d) are linked both to the peak amplitude and the quality of the signal. The latter are useful features because the ML regressor does not know a priory the data quality. The integral quantities (IA^2, IV^2, ID^2A_{rms}, V_{rms}, D_{rms} and CAV) are closely related to the energy radiated during the rupture processes, while τ_c and τ_p are related to the dominant frequencies within the earthquake signals, and thus provide indications on the events' size. These latter parameters are, in fact, often used in EEW applications to predict the magnitude of the event, sometimes in combination with P_d (Wu & Kanamori 2005b; Zollo *et al.* 2010; Caruso *et al.* 2017).

3.2 Targets

We explore the capability of the set of features listed in Table (1) to predict the target parameters *D* and PGA, having in mind to apply them for an on-site, threshold based EEWS.

For the hypocentral distance, *D*, we follow the classic approach used in Ground Motion Prediction Equation (GMPE) studies by predicting the logarithm of *D*, $\log_{10} D$ (Bindi *et al.* 2011). Following again Bindi *et al.* (2011), we compute the PGA as the geometrical mean of the horizontal components (PGA_H) as follows,

$$PGA_H = \sqrt{PGA_N \cdot PGA_E} \quad (1)$$

Where PGA_N and PGA_E are values for the North and East components, respectively. In the end, we aim to predict $\log_{10} PGA_H$.

Our choice of the target parameters (i.e. $\log_{10} D$ and $\log_{10} PGA_H$) is motivated by the idea, and past experiences in EEW (e.g. Caruso *et al.* 2017; Iaccarino *et al.* 2020; Mousavi & Beroza 2020), that features extracted from the *P* waves vertical component of ground motion carry information useful to rapidly estimate the distance and expected ground shaking.

It is important to note that both *D* and the PGA_H determination are not linear problems, despite being a reasonable assumption in many cases. As matter of fact, there are multiple factors that concur in the non-linearity: 1) The magnitude of the event influences not only the amplitude of the waveforms but also the frequency content; 2) The frequency content also changes if we use waveforms in acceleration, velocity, or displacement; 3) The source time function duration depends itself on magnitude, and that influences how much information we can retrieve from a chosen time window; 4) The anelastic attenuation acts reducing the more the high frequency content the further away will be the event; 5) Site effects corrupt the signals related to the seismic source; 6) For shorter distances, part of the *S*-wave might be included in the chosen window. All these effects influence in different ways each feature at the same time, making the problem of estimating *D* and the PGA_H from features clearly non-linear. Therefore, in this paper, we use a non-linear machine learning regressor as the Gradient Boosting Regressor.

3.3 Lead-time

To assess the effectiveness of the EEW alerts, after the ML models calibration, we will consider the lead-time. Therefore, we calculate the lead-times, LT_{PGA}, for all the waveforms in our data set considering the arrival time of the PGA, t_{PGA}, as

$$LT_{PGA} = t_{PGA} - t_P - P_{window} - t_{comp} \quad (2)$$

where P_{window} is the considered time window in seconds, t_p is the *P*-wave arrival and t_{comp} is the computation time that we set at 0.5 seconds. It is worth to note that often EEWS approaches consider

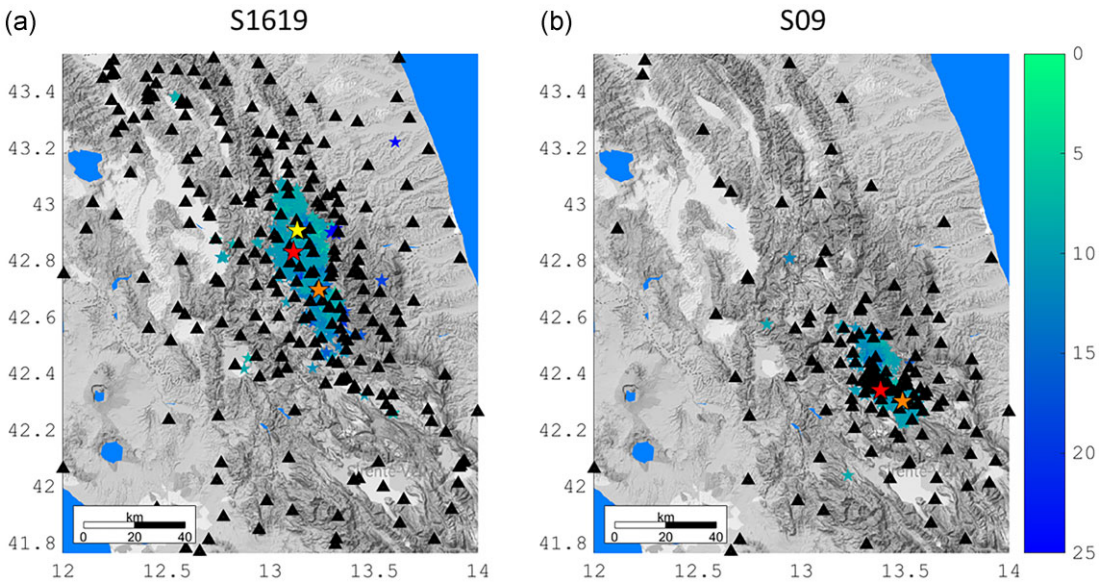

Figure 1. Map showing the spatial distribution of earthquakes and accelerometric/seismic stations. The events are displayed as stars coloured by depth while the stations are plotted as black triangles (a) S1619 set. The red star refers to the M_w 6.5 '161 030 064 017' Norcia event. The orange star refers to the M_w 6.0 '160 824 013 632' Amatrice event. Finally, the yellow star refers to the M_w 5.9 '161 026 191 805' Visso earthquake. (b) S09 data set. The red star refers to the M_w 6.0 '090 406 013 242' L'Aquila event. The orange star refers to the M_w 5.4 '090 407 174 739' Fossa event.

Table 1. List of the features used.

Feature	Symbol	Equation	Reference
Peak of acceleration	Pa	$Pa = \max_{t_o<t<t_o+\Delta t} A(t)$	(Wu & Kanamori 2005a)
Peak of velocity	Pv	$Pv = \max_{t_o<t<t_o+\Delta t} V(t)$	(Wu & Kanamori 2005a)
Peak of displacement	Pd	$Pd = \max_{t_o<t<t_o+\Delta t} D(t)$	(Wu & Kanamori 2005a)
Integral of squared acceleration	IA^2	$IA^2 = \int_{t_o}^{t_o+\Delta t} A^2(t)\,dt$	This work
Integral of squared velocity	IV^2	$IV^2 = \int_{t_o}^{t_o+\Delta t} V^2(t)\,dt$	(Festa et al. 2008)
Integral of squared displacement	ID^2	$ID^2 = \int_{t_o}^{t_o+\Delta t} D^2(t)\,dt$	(Iaccarino et al. 2021)
Dominant period	τ_c	$\tau_c = 2\pi\sqrt{\frac{\int_{t_o}^{t_o+\Delta t} D^2(t)\,dt}{\int_{t_o}^{t_o+\Delta t} V^2(t)\,dt}}$	(Wu & Kanamori 2005a)
Predominant Period	τ_p	$\tau_p = 2\pi\sqrt{\frac{\text{mean}_{t_o<t<t_o+\Delta t} D^2(t)}{\text{mean}_{t_o<t<t_o+\Delta t} V^2(t)}}$	(Wolfe, 2006)
Cumulative absolute velocity	CAV	$CAV = \int_{t_o}^{t_o+\Delta t} \lvert A(t)\rvert\,dt$	(Fahjan et al. 2011)
Acceleration root mean square	Arms	$A_{rms} = \frac{\sqrt{\int_{t_o}^{t_o+\Delta t} A^2(t)\,dt}}{\Delta t}$	
Velocity root mean square	Vrms	$V_{rms} = \frac{\sqrt{\int_{t_o}^{t_o+\Delta t} V^2(t)\,dt}}{\Delta t}$	
Displacement root mean square	Drms	$D_{rms} = \frac{\sqrt{\int_{t_o}^{t_o+\Delta t} D^2(t)\,dt}}{\Delta t}$	
Acceleration signal-to-noise ratio	SNRa	$SNR_a = 20 * \log_{10} \frac{Pa}{Pa_{noise}}$	This work
Velocity signal-to-noise ratio	SNRv	$SNR_v = 20 * \log_{10} \frac{Pv}{Pv_{noise}}$	This work
Displacement signal-to-noise ratio	SNRd	$SNR_d = 20 * \log_{10} \frac{Pd}{Pd_{noise}}$	(Caruso et al. 2017)

the *S*-wave arrival time to compute the lead-time. However, many earthquakes, especially the largest ones, present peaks of amplitude that arrive well after the *S*-wave arrival, due to the finite source effect. For this reason, considering the arrival time of the PGA, which is the effective ground motion with respect to the EEWS should protect the users, provides a more realistic estimation of the effective lead-time (Parolai *et al.* 2015).

3.4 Gradient Boosting Regressor

One of the most powerful and applied techniques in ML is the ensemble of methods, whereas predictions from multiple, often weak, statistical models are combined to improve the predictive performance (Dietterich 2000). Among these ensembles, we recall the random forests (Breiman 2001) and boosted gradient trees (Friedman 2001).

Table 2. Scheme of the features used for the models.

	Z	N	E	H
D	✓	✓	✓	✓
PGA_H	✓	✗	✗	✓

It is worth noting that the ensemble predictions are not always better than the prediction from the individual models that constitute the ensemble. However, using large and completely random ensembles (for which the models are uncorrelated), it has been shown that the variance and, thus, the fluctuations due to finite-sample effects are cut down. On the other hand, when models in the ensemble are completely random, the bias of the aggregate predictor is the same of a single model. This is true for random forest method, but in general, the increase in bias can be neglected with respect to the reduction in variance. For this reason, these methods are well-suited for procedures where the error of the predictor is dominated by the variance and not by the bias (Mehta *et al.* 2019).

In this work, we use the Gradient Boosting Regressor (Friedman 2001), hereinafter called GBR. GBR combines the intuitions from boosting and gradient descent to construct ensembles of decision trees. Like in boosting, the ensembles are created by iteratively adding new decision trees to the structure. The central role is played by the cost function that measures the performance of our ensemble. At each step of the process, the gradient of the cost function is calculated with respect to the predicted value of the ensemble, and trees are added to shift the algorithm in the negative direction of the gradient. The GBR models are mainly controlled by three hyperparameters: the number of the trees, the maximum depth of these trees and the learning rate of the gradient descent. It is worth noting that, since GBR is an ensemble of decision trees, it cannot predict a target that is outside the range of targets which make up the training set. For example, the calibrated model in our application will never be able to predict a distance greater than 150 km. Moreover, as for all the machine learning methods, GBR has problem to make good predictions for events with different characteristic with respect to the training set (i.e. events from different regions, or with greater magnitude).

3.5 Training process

To improve the convergence of the gradient descent approach, we first standardize all the 60 features (Table 1) for each considered time window (i.e. 1 s, 2 s and 3 s). Then, we tune our GBR models. Specifically, we fix the number of trees to 300. For the GBR training, we consider the 80 per cent of S1619 (hereinafter, $S1619_T$), with data selection stratifying them by magnitude. The stratification in magnitude helps to avoid unwanted effects due to data imbalance (Supporting Information Fig. S1).

We use all the features presented in Table 1 for both the models, but, after a trial-and-error procedure, we decide to exclude features obtained from the N- and E- components for computing the PGA_H. In this way, the model for the prediction of D uses a total of 60 features, while, for the prediction of PGA_H, only 30 features are used. In Table 2, we present a simple scheme of the components used for both models.

During the validation process, we consider the remaining 20 per cent of S1619 (hereinafter, $S1619_V$). We let the maximum depth of the trees to vary from 3 to 20, and the learning rate of the gradient descent from 0.001 to 0.1. This choice for the hyperparameters ensures a robust predictive model (Raschka & Mirjalili 2017). In the tuning phase, we train the model using the $S1619_T$ set, and we evaluate the hyperparameters combination looking at the results on the $S1619_V$ set in terms of R^2 of the prediction. At the end of this process, we choose the best set of hyperparameters. Finally, we apply the best model on the S09 set. It is worth to note that we derive different models for each time window tuning the hyperparameters differently. This procedure is important to avoid or, at least mitigate, the effect of the overfitting. This is a natural effect of ML techniques, where the calibrated models tend to perfectly predict the training set, but at the cost of a drop in the prediction performance when models work on the testing set. The validation process helps to select models that reduces at maximum the overfitting.

It is important to note that the system presented in this paper is thought to be just a module of a complete EEW system. Therefore, we do not consider data with only seismic noise in the training process, because we assume that another module would distinguish, in real-time, seismic arrivals from noise.

4 RESULTS

4.1 GBR for hypocentral distance prediction

Following the GBR training on the $S1619_T$ data set, we show in Fig. 2 the results of the model validation on $S1619_V$ concerning the prediction of the logarithm of hypocentral distance ($\log_{10}$D). The results are presented in terms of R^2. We can see that the best hyperparameter configurations barely change with the time window length. The best maximum depth is always 10, while the learning rate is 0.063 for the 1 s and 3 s models, and 0.016 for the 2 s model. The R^2 slightly increases with the time window length from 0.810 to 0.828. Moreover, we see that R^2 has values that differ each other by less than 0.01 from the best solution when considering various configurations. This result tells us that it is not necessary to refine the search of the hyperparameters.

In Fig. 3, we report the prediction of the logarithm of hypocentral distance ($\log_{10}D$) for S09 at 1 s, 2 s and 3 s. The data are colored by magnitude to highlight the well-known trade-off between magnitude and distance that typically hampers the magnitude estimation in on-site EEW applications (Münchmeyer *et al.* 2020; Song *et al.* 2022). Furthermore, we report the predictions of $\log_{10}$D for $S1619_T$ (Supporting Information Fig. S2) and $S1619_V$ (Supporting Information Fig. S3) in the Supplemental Material. Overall, we can see that the predictions are reliable for all the time windows, having: i) $R^2 = 0.74$ and $\sigma = 0.15$ at 1 s; ii) $R^2 = 0.77$ and $\sigma = 0.14$ at 2 s; iii) $R^2 = 0.79$ and $\sigma = 0.14$ at 3 s. As expected, these values are lower than the validation results, but they still indicate a good prediction performance. Moreover, it is evident a slight improvement of results with the increase of the time window length. The residual panels (lower subplots in Fig. 3) well highlight the strong dependence of residuals with magnitude. Our results confirms that any on-site EEW approach aiming to predict magnitude from P waves measures would be prone to large uncertainty.

We can have a better look at the residual's distribution for all the data sets and time windows in Supporting Information Fig. S4. In the latter, we binned the residuals for D, and for each bin we show the residuals mean values and the standard deviation (as blue error bars). Looking at plot for the training set, $S1619_T$, we observe that the residuals are consistent with zero for each bin except for very small D ($D < \sim 7$ km) and large D ($D > \sim 125$ km). The latter result confirms us that the model is well trained. For very small

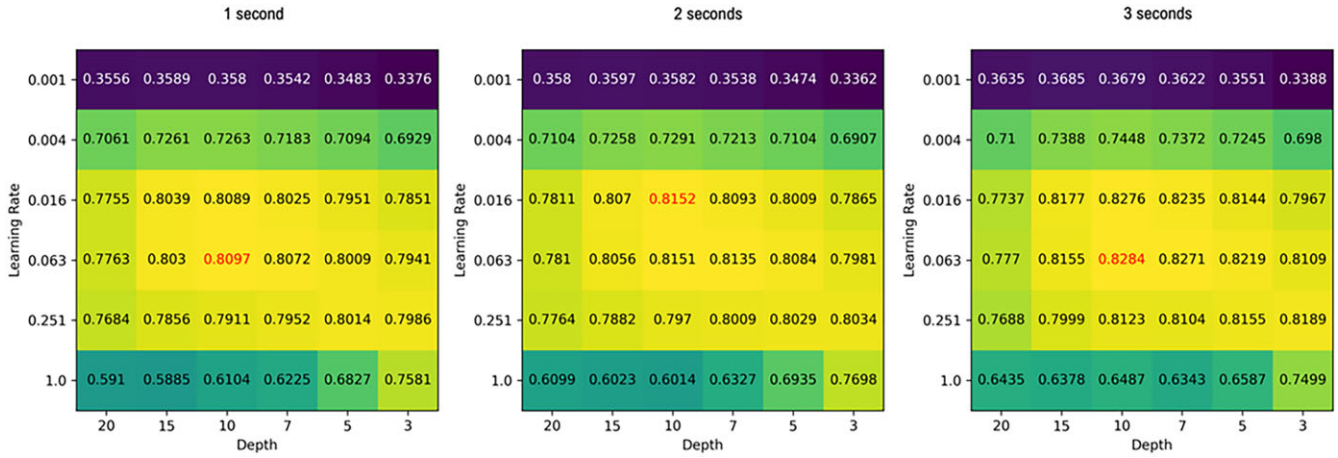

Figure 2. Results of the validation process performed for the three *P* time windows in terms of R^2. The heatmaps show the best couple of hyperparameters – depth and learning rate – (in red) to be used for the $\log_{10}D$ predictions.

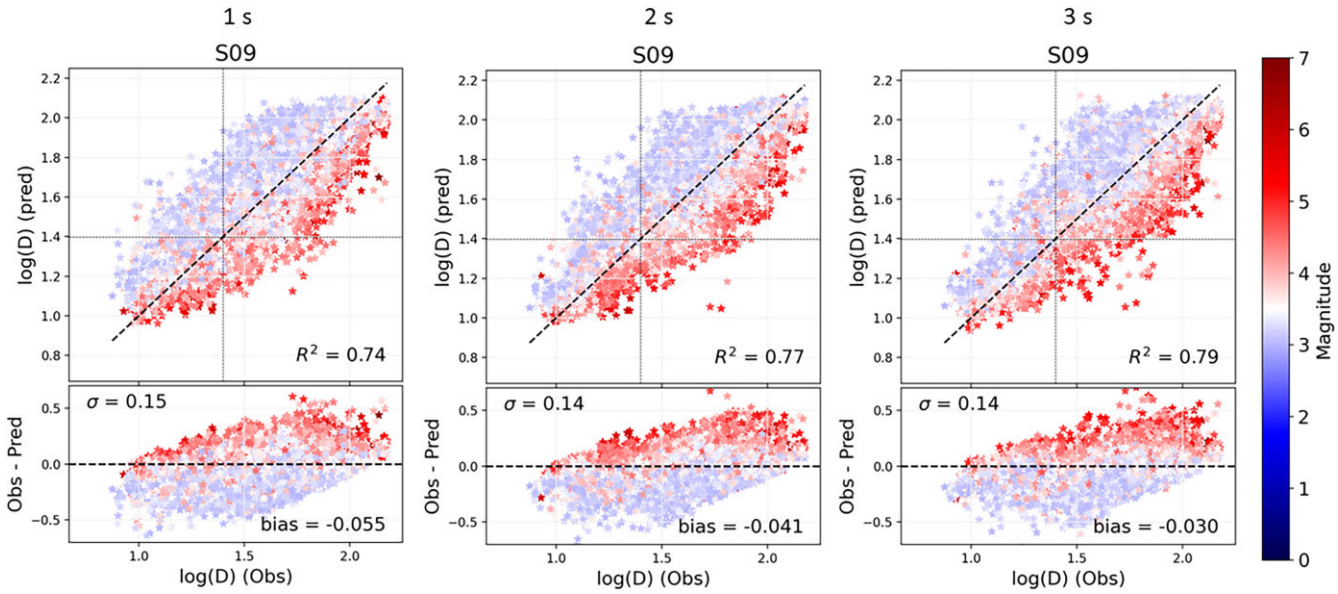

Figure 3: Predictions of $\log_{10}D$ versus the observed values for the three time windows for S09 data. The plots are colored by station magnitude. We report the time windows from left to right. Diagonal dashed lines report the 1vs1 relation, while horizontal and vertical ones denote the alert threshold used to assess the quality of prediction. The bottom panels show the residuals of the predictions versus $\log_{10}D$.

and large distances, on the contrary, data are too few for model convergence (Supporting Information Fig. S1). On the other hand, for $S1619_V$ and S09 we observe an increased dispersion of the data due to overfitting, but also in this case we note that the residuals remain consistent with zero for all the bins, with again the exception of $D < \sim 7$ km and $D > \sim 125$ km.

Interestingly, for the 2 s and 3 s windows, we observe that all the data sets present a notch of small overestimation (Supporting Information Fig. S4). The latter cases correspond to around $\log_{10}(D) = 1.2$ ($D \approx 16$ km) for the 2 s window and around $\log_{10}(D) = 1.4$ ($D \approx 25$ km) for the 3 s window. These values are consistent with the *S*-wave propagation, so we believe that this effect is mainly due to the fact that the model is adapting itself to whether or not the *S*-wave is already in the considered window (before or after the notch), while it has a small difficulty on the transition.

For a better evaluation of the models' performance, we analyze the ROC (Receiver Operating Characteristic) curve. We select this analysis because it allows to evaluate the prediction independently from choosing a specific alert threshold. This is particularly useful in EEW applications, since many studies pointed out the dependence of the EEW performances with the alert threshold (Meier 2017; Minson *et al.* 2019).

In Fig. 4, we show the ROC curves and the ROC-AUC (Area Under the ROC Curve) values for $\log_{10}D$ predictions, for $S1619_T$, $S1619_V$ and S09 sets. Looking at the ROC curves for the three data sets, we observe the best performance is obtained for the training set ($S1619_T$), while the results slightly worsen for the validation ($S1619_V$) and the testing (S09) sets.

Despite the strategies put in place to avoid it, we think that the difference in performances among the three data sets is likely due to a small overfitting in the model training. However, despite the overfitting, we observe that the AUC reaches values greater than 0.92 for S09 in all the time windows, and this represents a particularly good performance. Moreover, the performances between the validation and the testing data sets are similar, confirming that the two data sets are quite homogeneous. Finally, we can see a slight

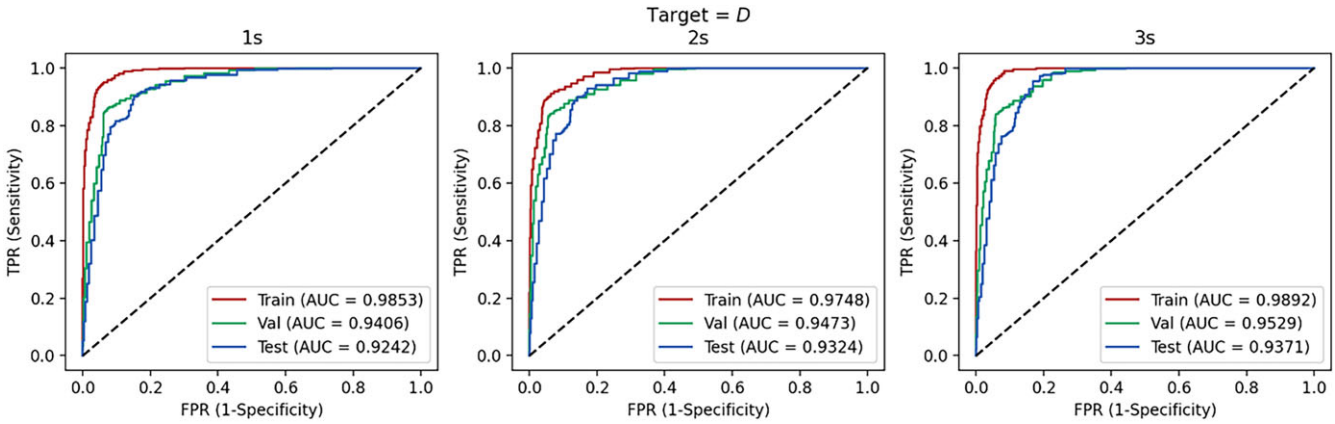

Figure 4: ROC-curves for $\log_{10}D$. Each panel shows the three curves for training, validation and testing sets. In parenthesis, we reported the AUC values for the predictions.

improvement in the AUC value for the testing prediction, switching from 1 s to 3 s window.

4.2 GBR for peak ground acceleration prediction

In Fig. 5, we report the results of the validation process on $S1619_V$ for the prediction of the $\log_{10}PGA_H$ (the figure is analogue to Fig. 2). In this case, we obtain the same best hyperparameters configuration for the three time window lengths: maximum depth equal to 10 and learning rate equal to 0.016. As in previous case, the validation R^2 increases with the time window length from 0.888 to 0.940. The performances for this target are clearly better than those for D, suggesting a better correlation of PGA with the used features.

In this case, we use the GBR regressor to retrieve the $\log_{10}PGA_H$, and we start, as before, from P waves features without any analytical hypothesis on the function form of the model. Here, we consider only features computed on the vertical, Z, and the horizontal, H, components. Moreover, we exclude features for the E and N components after a trial-and-error optimization.

In Fig. 6, we show the predictions of $\log_{10}PGA_H$ obtained by applying the GBR, along with the residual plots for the 1 s, 2 s and 3 s windows, for S09 set. In the Supplemental Material, we present the predictions of $\log_{10}PGA_H$ for $S1619_T$ (Supporting Information Fig. S5) and $S1619_V$ (Supporting Information Fig. S6).

The results are already very good for the 1 s window, reaching a value for the R^2score equal to 0.90. For larger time window lengths, we can observe a slight improvement in the predictions, which reaches $R^2 = 0.95$ for 3 s. The standard deviation ranges between 0.27 for the 1 s window to 0.19 for the 3 s case. For comparison, Bindi *et al.* (2011) provided a GMPE relation for $\log_{10}PGA_H$ using Italian data with a total standard deviation equal to 0.337.

In Supporting Information Fig. S7, we again report all the residuals together with the mean and the standard deviation computed on data binned in PGA. The latter shows us that the residuals for $S1619_T$ are always consistent with zero for each bin, except for very small and very large values of $\log_{10}PGA_H$ (i.e. for these latter bins likely data are too few to properly constrain the model, Supporting Information Fig. S1). No trend is visible in the residuals, which confirms us the goodness of GBR models. Moreover, we note an increased dispersion for $S1619_V$ and S09 due to overfitting. By the way, it is evident that the predictions present a high underestimation for large PGA for $S1619_V$. However, since this effect is not visible for S09, we believe that is due to a heterogeneity of the data sets in this range of PGA_H, which maybe is due to the presence of peculiar source, propagation or site effects.

As done in the previous paragraph, we report in Fig. 7 the ROC curves and AUC values for $\log_{10}PGA_H$ predictions, for $S1619_T$, $S1619_V$ and S09 sets. Also in this case, we can observe a small overfitting, which is identifiable as the small area between the red curve and the other ones, but also from the decrease of the AUC by just 0.01–0.02. Nevertheless, it is worth to say that the GBR models show a very good performance in the prediction of PGA_H.

4.3 Prototype of EEWS

To highlight the potential benefits and limitations of an on-site EEW system predicting the hypocentral distance and ground motion intensity using the calibrated models in this study, we present in the following a prototype of a single-station alert system. In this framework, the alert notification of an upcoming earthquake damage is issued based on the local measurements of P-wave ground motion, with no need for accurate estimation of source parameters (Zollo *et al.* 2010). Many authors (Kanamori 2005; Wu & Kanamori 2005a, 2005b; Zollo *et al.* 2010) have suggested the introduction of an alert system for both regional and on-site EEWSs, using the information extracted from the early P-wave signals, such as the P_d and/or P_a, P_v and τ_c parameters, as proxies for the estimation of the ground motion parameters, such as Peak Ground of Displacement (PGD), PGV and PGA. The setting of specific thresholds for these parameters allows the system to rapidly issue an earthquake alert based on a four-entry decision-table scheme (Zollo *et al.* 2010). Nevertheless, we must consider that a given level of uncertainty is always associated with the model prediction used in EEWS. Therefore, we establish an alert criterion in which the comparison between the predicted ground motion and the threshold selected for the ground motion parameter is done using a probabilistic approach (Minson *et al.* 2019; Iaccarino *et al.* 2020).

For our purpose, we consider that the estimates of $\log_{10}PGA_H$ and $\log_{10}D$ are the mean values of two gaussian functions with standard deviations equal to σ_{PGA_H} and σ_D, respectively. We thus combined the two functions by computing their joint probability distribution and we obtain a probability ellipsoid in the space of $\log_{10}PGA_H$ and $\log_{10}R$ (Figs 8a and c). This allows us to build an alert system based on the exceedance probability with respect to any threshold selected for the two parameters.

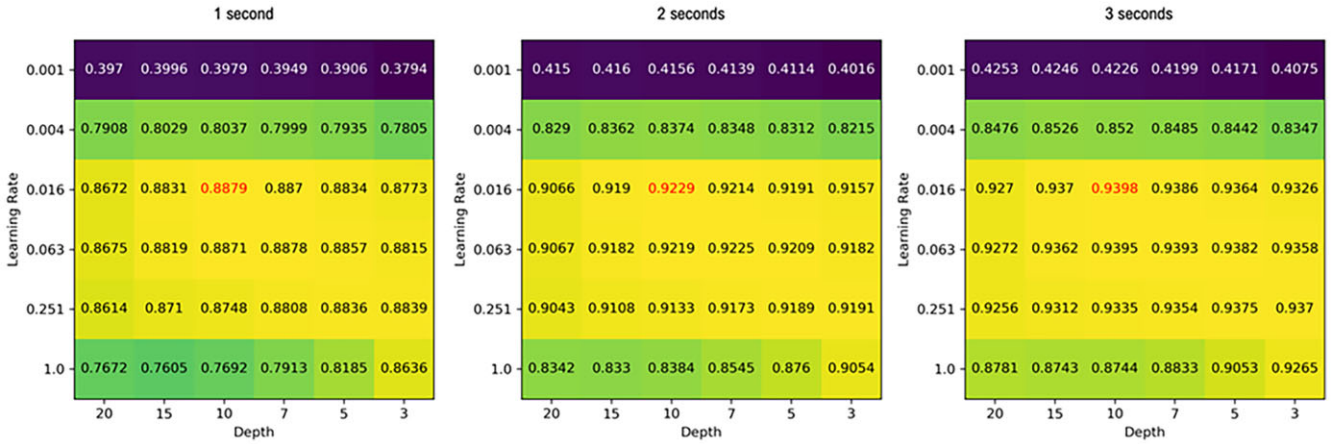

Figure 5. Results of the validation process performed for the three *P* time windows in terms of R^2. The heatmaps show the best couple of hyperparameters – depth and learning rate – (in red) to be used for the $\log_{10}PGA_H$ predictions.

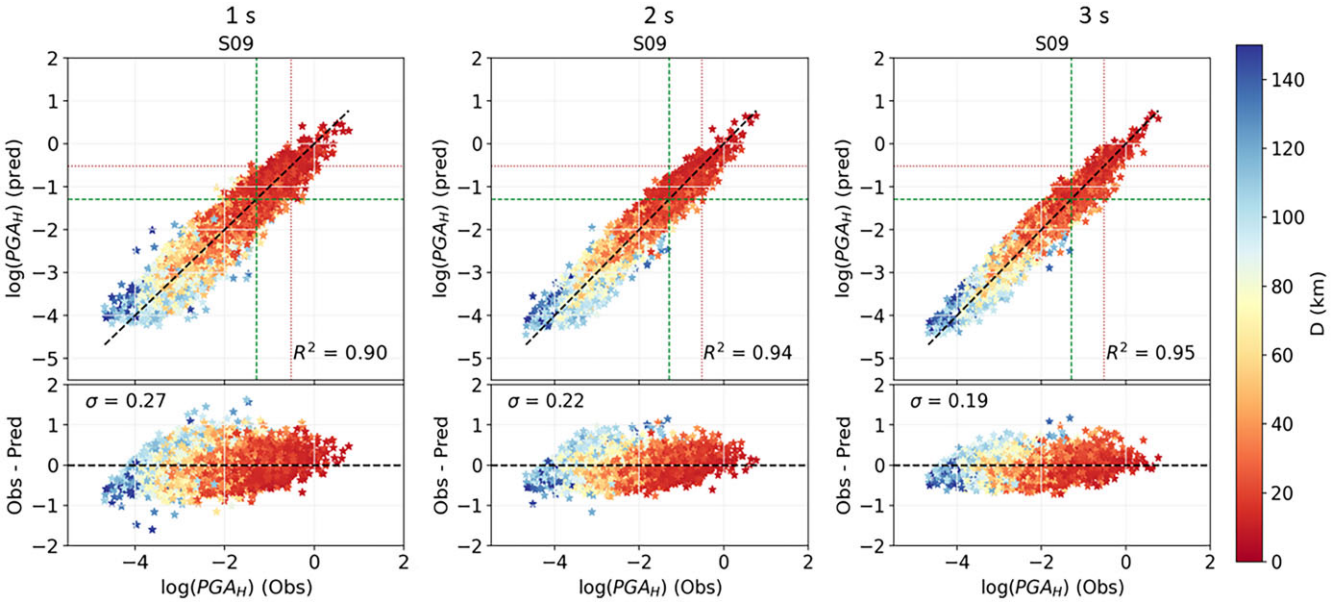

Figure 6. Predictions of $\log_{10}PGA_H$ versus the observed values for the three time windows for S09 data. The plots are coloured by observed *D*. Diagonal dashed lines report the perfect predictions, while horizontal and vertical ones denote the alert thresholds (see Table 3). The bottom panels show the residuals of the predictions.

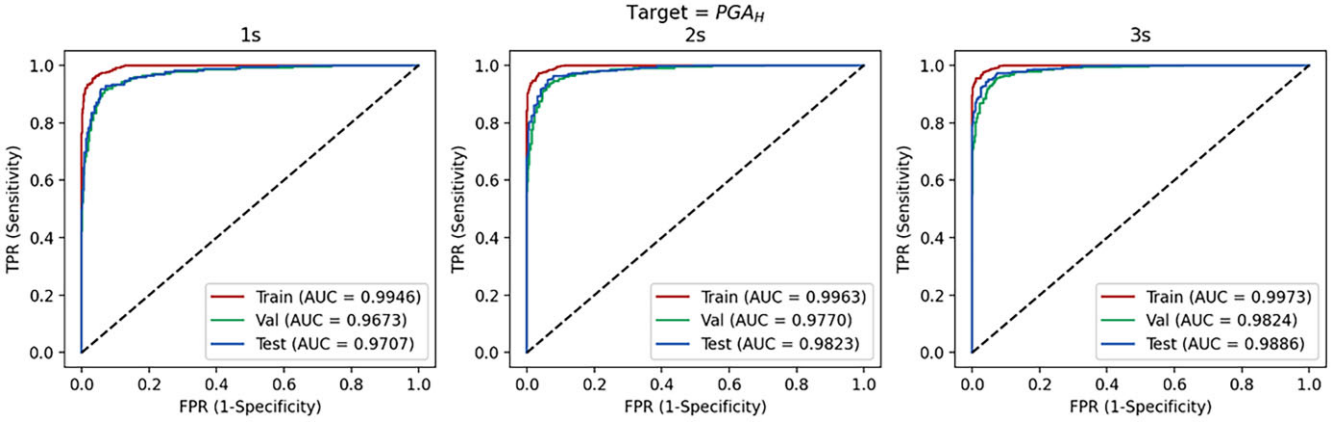

Figure 7. ROC-curves for $\log_{10}PGA_H$. Each panel shows the three curves for training, validation and testing sets. In parentheses, we reported the AUC values for the predictions.

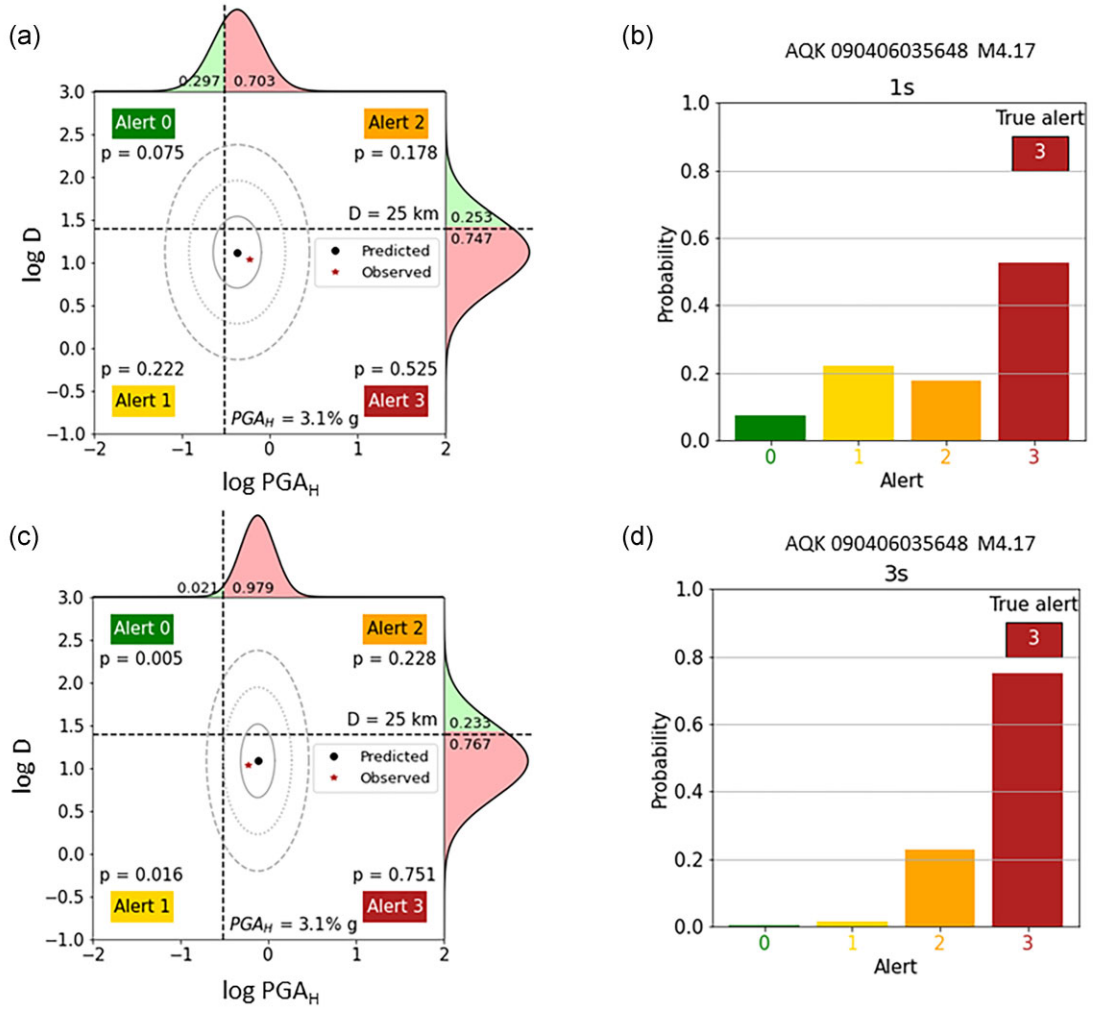

Figure 8. Probabilistic threshold-based alert system example for event M4.17 '090 406 035 648' recorded at the station AQK (observed hypocentral distance $D = 11.0$ km and $PGA_H = 0.580$ m s^{-2}). (a) Probabilistic scheme for the alert declaration using 1 s model. The black dot refers to the predicted values of distance and PGA for this datum. The red star indicates the observed value for the same data point. Vertical and horizontal dashed lines indicate the thresholds chosen for distance and PGA for the Damage Alert system. The confidence ellipse for the estimation of Distance and PGA is shown for 1σ, 2σ and 3σ. On the top and on the right side of the plot, we report the single target normal distributions, and the probabilities values corresponding to the area above or below the thresholds. (b) Bar plot of the joint probabilities for each possible alert level at 1 s. (c) Same as (a) at 3 s. (d) Same as (b) for 3 s.

Table 3. Felt (FA) and Damage (DA) Alerts. The PGA information is taken from Faenza & Michelini (2010).

	FA	DA
Distance (km)	50	25
PGA (%g)	0.52	3.1
Perceived shaking	Light	Strong
Intensity (Low bounds)	IV	VI

Therefore, we define two levels of alert: a Damage Alert (DA) and a Felt Alert (FA) (see Table 3). We define the thresholds in PGA_H associated to these alert levels considering the two MCS Intensity classes from (Faenza & Michelini (2010): intensity IV for FA, and VI for DA. We define the thresholds in D equal to 50 km for FA and 25 km for DA.

More specifically, the distance threshold for both DA and FA is considered as an upper bound, meaning that the alert is triggered if the prediction of D is below that threshold (i.e. if the event occurs at close distances to the target). On the contrary, we consider for the $\log_{10}PGA_H$ a lower bound threshold, for which the alert is triggered by predictions of $\log_{10}PGA_H$ surpassing the threshold itself (Figs 8a and c).

Computing the area under the normal distribution above or below the given threshold for a set of PGA and D predictions allows to determine the probability associated to the FA and DA alerts at a given target (here we consider the target co-located with, or nearby, the seismic station). Then, multiplying these probabilities, we obtain a joint probability (Figs 8b and d) that the predicted targets overcome either both thresholds, one threshold, or no threshold at all. This gives us four distinct levels of alert, given by:

(i) Alert 0: Distance $>$ Threshold, PGA $<$ Threshold (i.e. no danger for the target).

(ii) Alert 1: Distance $<$ Threshold, PGA $<$ Threshold (i.e. close event but negligible danger for the target).

(iii) Alert 2: Distance $>$ Threshold, PGA $>$ Threshold (i.e. far-away event but danger for the target).

(iv) Alert 3: Distance $<$ Threshold, PGA $>$ Threshold (i.e. close event and danger for the target).

In Fig. 8, we report the case of an event (ID 090406035648[1]) with magnitude $M_w = 4.17$ that has been recorded at the station AQK, and we consider the DA thresholds. The observed hypocentral distance D and PGA_H are 11.0 km and 0.580 m s^{-2} (5.92 per cent g), respectively. Therefore, according to our decision framework, the correct output of an EEWS should be 'Alert 3″. Already from the information related to the 1 s time window, the GBR models predict D equal to 13.2 km and PGA_H equal to 0.426 m s^{-2} (4.35 per cent g) (Fig. 8a). These predictions are within 1σ from the real value (Fig. 8a). These estimates of D and PGA_H lead our EEW prototype to correctly release the alert level to be '3' with a probability of 0.53 (Fig. 8b). In Figs 8(c) and (d), we present the outputs of the EEW prototype for the same event but using the 3 s features. For this time window length, the models predict $D = 12.2$ km and $PGA_H = 0.756$ m s^{-2} (7.71 per cent g). Therefore, also in this case the prediction of the targets is within the 1σ from the real value, even considering that the standard deviation is lower in this case (Fig. 8c). Again, the EEW prototype provides the right alert level, this time with a probability of 0.75 (Fig. 8d).

Furthermore, we present two more applications of the prototype in Supporting Information Figs S8 and S9. Here, we show the results for the M_w 6.0 L'Aquila main event ('090 406 013 242'), for two different stations, AQK at a distance of 8.43 km (Alert 3) and LSS at a distance of 42.23 km (Alert 0). In both cases, we have very good results and successful alert declarations.

The prediction of the hypocentral distance is fundamental to provide the EEW users with an estimation of the lead-time, which is the time available to mitigate their exposure to the incoming seismic risk.

Using our EEW prototype, we obtain a probability value for each of the four alert levels, and then we select for the hypothetical warning message the alert level corresponding to the maximum probability. It is worth noting that our choices are made for the sake of illustrating how an On-Site EEWS would work. Of course, for the operations in real conditions, the final users of the EEWS could decide to follow our scheme, but to tailor the decision alert criteria according to other needs (e.g. by a cost-benefit analysis).

In Fig. 9, we report the EEWS performances for the whole S09 data set considering each time window length and the threshold settings defined in Table 3. The performance is assessed comparing the GBR predictions for D and PGA_H with the observed measured values. Since we have 4 possible outputs for the EEWS (i.e. four alert levels), we have 12 different kinds of misclassification being that each alert level can be misclassified with the other three. Six of these misclassifications represent overestimations of the alert, '1|0″, '2|0″, '3|0″, '2|1″, '3|1″ and '3|2″ (released alert| real alert), the opposite ones are underestimations. Moreover, as noted by many authors (Satriano *et al.* 2011; Minson *et al.* 2019), for assessing the performance of the EEW system, we must also consider the events that happen too close to the station/target to release a useful alert (i.e. the sites that are within the so-called blind-zone, where there is no lead-time available to mitigate the users' exposure to the seismic risk). In the end, to simplify the problem, we define 4 types of performances: the rate of Successful Alert, SA, Missed Alert, MA, Underestimated Alert, UA and Overestimated Alert, OA.

These latter are defined as:

SA: Released Alert = True Alert.

MA: Lead Time $\Leftarrow$ 0 s.

UA: Released Alert < True Alert.

OA: Released Alert > True Alert.

Moreover, we recall that in this analysis false triggers are not considered, because here we only aim at assessing the performance of the calibrated modules. We can imagine them as modules of a future EEW system, and the issue of false trigger on noise will be assessed during the future testing of such system in the real word.

In Fig. 9, we report the total performances of our EEW prototype. Here, we see that using the GBR predictions of D and PGA_H we always obtain a SA rate over 65 per cent, with a maximum of 81.9 per cent for DA at 1 s. As matter of fact, we see that SA decreases with the increasing of the time window, but this is due to the increase of MA (i.e. large windows increase the blind-zone size), with this latter that ranges from 2.6 per cent at 1 s to 25.1 per cent at 3 s. Similarly, also UA and OA are lower for larger time windows, also in this case for the increase of MAs. These results show that the GBR model for the 1 s window can be extremely useful in EEW applications for regions as Central Italy being prone to moderate-to-large earthquakes ($M_w < 7$). For such kind of earthquakes, the 1 s window allows for an excellent compromise between a high performance in the prediction of target parameters and a small blind zone. Of course, the 1 s window would saturate with events for which the P-wave source time function is far greater than the considered window. Increasing the time window to 2 s or 3 s would slightly improve the precision of the target's prediction, but at the price of eroding the lead-time. Furthermore, we do not see many differences in performances between FA and DA thresholds. This confirms that the provided models are robust and reliable.

5 CONCLUSIONS

In this work, we explored the feasibility of the GBR for obtaining models suitable for the real-time prediction of hypocentral distance and ground motion intensity for single-station EEW purposes. The hypocentral distance, and Peak Ground Acceleration, PGA predictions are derived exploiting 60 features extracted from three P-wave time windows of length 1,2 and 3 seconds. To this purpose, we have used a large and high-quality data set consisting of about 80 000 earthquake waveforms and covering a magnitude range between $M_w = 1.8$ and $M_w = 6.5$ and a distance range between 5 km and 150 km. We divided the data set in three parts, using the bigger subset relevant to the 2016–2017 Central Italy sequence for training and validation of the models, while the smaller one relevant to the 2009 L'Aquila sequence as testing set.

The results of our analyses show that the proposed methodology can provide reliable predictions of both hypocentral distance and peak ground acceleration. As matter of facts, D is predicted on the S09 set with standard deviations equal to 0.15, 0.14 and 0.14 for the 1 s, 2 s and 3 s time windows, respectively. From the analysis of the residuals, we observed a trade-off with magnitude that, however, does not affect the performances of the estimations. The results have been evaluated considering the R^2 and the AUC-ROC values, which improve from 0.92 for the 1 s window to 0.94 for 3 s window on the S09. As said, the residuals present a remaining correlation with the magnitude. We observed that this trade-off causes an underestimation of D for event with $M_w > 4$ and large hypocentral distances (and clearly, vice versa, an overestimation for lower magnitude events). Despite this effect is not ideal, the underestimation of D for high magnitude events and large hypocentral distances would tend to produce safer alerts, which is a good side-effect.

[1]The ID refers to the origin time of the event in the format 'yymmddHHMMSS'.

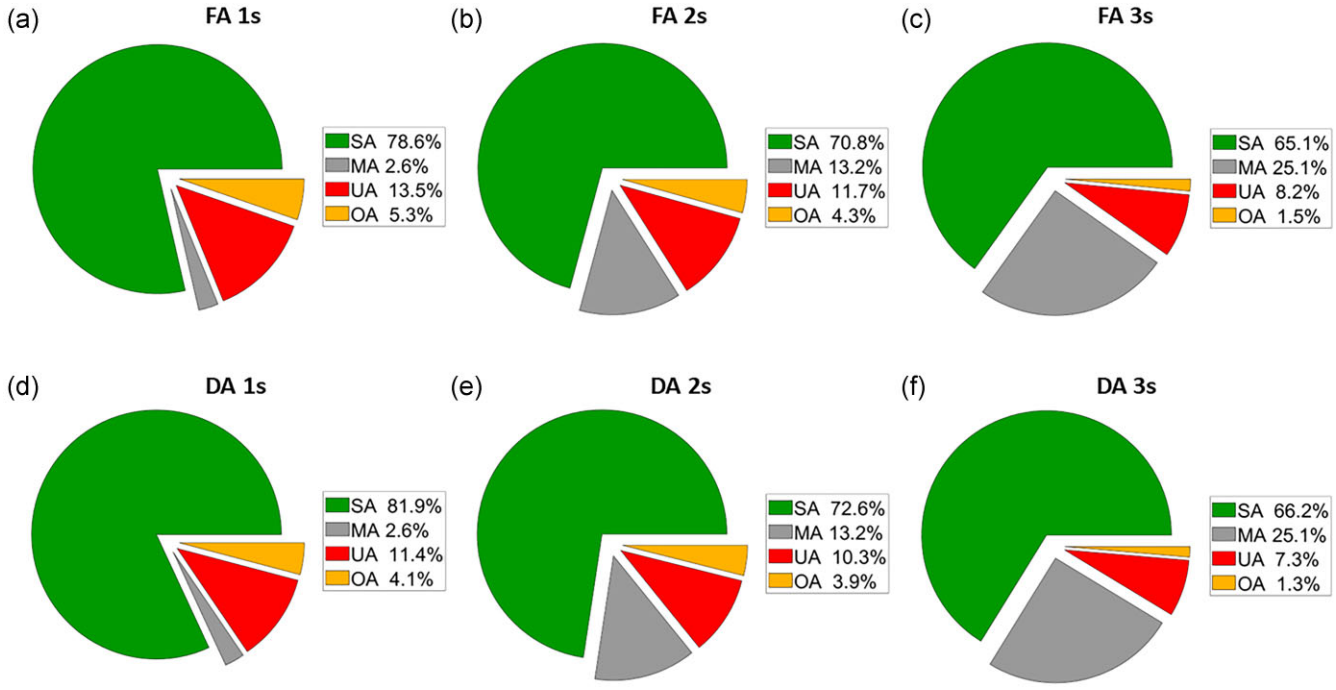

Figure 9. Pie charts of the EEWS performances on the S09 data set. (a) Results for 1 s model and FA thresholds. (b) Results for 2 s model and FA thresholds. (c) Results for 3 s model and FA thresholds. (d) Results for 1 s model and DA thresholds. (e) Results for 2 s model and DA thresholds. (f) Results for 3 s model and DA thresholds.

The PGA predictions exhibit good accuracy and reliability already using *P*-wave information extracted from a window length equal to 1 s. The analysis of the residuals does not present any bias or correlations. This means that PGA can be estimated already with a good level of accuracy from the first seconds of the *P*-wave for the range of magnitude explored. This is particularly important considering the perspective of including this methodology in a real-time framework for EEW on-site applications. It is important to note that not all the features are equally important for the retrieved models. To address this question, we perform a features importance analysis for the two models and show the results in Figs S10 (*D*) and S11 (PGA_H). From these figures we observe that a limited number of features outperforms the others since the ten most importance features accounts for 82–85 per cent of the performances for *D* prediction (10 out of 60 features) and for 94–97 per cent for PGA_H prediction (10 out of 30 features). Interestingly, we see in both Figs S10 and S11 that among the best ten features, some are recurring (i.e. Pa, Pv, SNR_a, ID^2 and τ_c), even if with different components. This set of features seems trying extracting from the event waveforms complementary information concerning the seismic source size and distance, as the radiated energy (ID^2), the peak amplitude (Pa, Pv, SNR_a), the frequency (τ_c) and the data quality (SNR_a). These pieces of information are evidently used by the GBR models to optimize the prediction. Despite the existence of a set of important features, their variability among the different targets and signal window duration testifies the absence of data leakage. In fact, in the case of data leakage, the GBR would have chosen the same 'leaked' feature for all the models. Our results let us to believe that, with further studies, it should be possible to develop a simpler model with less features and similar overall performances.

We have then assessed the performance of the calibrated predictive models by considering a prototype of an EEW alert system based on PGA and distance thresholds chosen for two macroseismic intensity scale levels, that we indicated as the FA and DA (Table 3). We adopted a probabilistic approach to describe the possibility to give an alert. In particular, we computed the joint probabilities of the two predictors, and we established a decision probability framework based on four levels from 0 to 3. Moreover, we considered the lead-time of the PGA to distinguish the useful alerts from the late ones that we counted as Missed Alert. In the end, we established 4 kinds of performances: successful, missed, underestimated, and overestimated alarms (Fig. 9). We showed that our prototype is particularly reliable already considering only 1 s of *P*-wave signals, having SA a rate greater than 78 per cent for both damage and felt thresholds. With the increasing of the time window length, the missed alerts increase significantly. This latter result well illustrates a problem in EEW applications for country like Italy (i.e. those with short distances between seismogenic zones and targets; Picozzi *et al.* 2015a; 2015c). Indeed, an increase of the time window length to pursue better predictions of EEW parameters (e.g. hypocentral distance, ground motion intensity, magnitude), would inevitably erode the lead-time and cause the occurrence of missing alerts for targets prone to the largest ground motion intensities.

In conclusion, we also wish to stress the concept that the tested models can be applied to the seismicity on which they have been trained, that is the seismicity with range of depths and magnitude of central Italy. Further studies and new data sets are necessary to safely transfer the developed models to other seismogenic areas. Despite that, we believe that the calibrated GBR models for predicting distance and peak ground acceleration by 1 s *P*-wave windows are particularly suitable for those regions where the target areas are close to the seismogenic threats.

SUPPORTING INFORMATION

Supplementary data and figures are available at *GJI* online.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.

ACKNOWLEDGMENTS

We want to thank the editor, Dr Sidao Ni, the reviewer, Dr Yuan Wang and the anonymous reviewer for their valuable comments, which allowed us to improve this paper.

This research was partly carried out in the frame of Program STAR PLUS (project DRAGON), financially supported by UNINA and Compagnia di San Paolo.

DATA AVAILABILITY

We used data and information retrieved from ORFEUS-EIDA (https://www.orfeus-eu.org/data/eida/), IRIS (https://www.iris.edu/hq/) and DPC (http://ran.protezionecivile.it/EN/index.php). The INGV bulletin is used to guide the data download (webservices.rm.ingv.it/fdsnws/event/1/) and to extract the earthquake locations. The International Federation of Digital Seismograph Networks (FDSN) specifications are available at http://www.fdsn.org/ and the Standard for the Exchange of Earthquake Data (SEED) manual is available at http://www.fdsn.org/pdf/SEEDManual_V2.4.pdf. For the calibration of the attenuation models, we used data from networks IV (DOI: 10.13127/SD/X0FXnH7QfY) and IT (DOI: 10.7914/SN/IT).

REFERENCES

Allen, R., Kong, Q. & Martin-Short, R., 2019, The MyShake platform: a global vision for earthquake early warning. *Pure appl. Geophys.,* **177,** 1699–1712.

Allen, R.M. & Melgar, D., 2019, Earthquake early warning: advances, scientific challenges, and societal needs. *Annu. Rev. Inc,* **47,** 361–388.

Bindi, D., Pacor, F., Luzi, L., Puglia, R., Massa, M., Ameri, G. & Paolucci, R., 2011, Ground motion prediction equations derived from the Italian strong motion database. *Bull. Earthq. Eng.,* **9**(6), 1899–1920.

Bindi, D., Spallarossa, D., Picozzi, M., Scafidi, D. & Cotton, F., 2018, Impact of Magnitude Selection on Aleatory Variability Associated with Ground-Motion Prediction Equations: part I—Local, Energy, and Moment Magnitude Calibration and Stress-Drop Variability in Central Italy. *Bull. Seismol. Soc. Am.,* **108**(3A), 1427–1442.

Böse, M., Heaton, T. & Hauksson, E., 2012, Rapid estimation of earthquake source and ground-motion parameters for earthquake early warning using data from a single three- component broadband or strong-motion sensor. *Bull. Seismol. Soc. Am.,* **102**(2), 738–750.

Breiman, L., 2001, Random forests. *Mach Learn,***45,** 5–32.

Brondi, P., Picozzi, M., Emolo, A., Zollo, A. & Mucciarelli, M., 2015, Predicting the macroseismic intensity from early radiated P wave energy for on-site earthquake early warning in Italy. *J. Geophys. Res. Solid Earth,* **120**(10), 7174–7189.

Caruso, A., Colombelli, S., Elia, L., Picozzi, M. & Zollo, A., 2017, An on-site alert level early warning system for Italy. *J. Geophys. Res. Solid Earth,* **122**(3), 2106–2118.

Colombelli, S., Caruso, A., Zollo, A., Festa, G. & Kanamori, H., 2015, A P wave-based, on-site method for earthquake early warning. *Geophys. Res. Lett.,* **42**(5), 1390–1398.

Cooper, J.D., 1868, Earthquake indicator, Letter to Editor. *San Francisco Daily Evening Bulletin,* Nov. 3.

Cremen, G. & Galasso, C., 2020, Earthquake early warning: recent advances and perspectives. *Earth Sci. Rev.,* **205,** 103184. doi: 10.1016/j.earscirev.2020.103184.

Dietterich, T.G., 2000, Ensemble methods in machine learning. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics),* **1857 LNCS,** 1–15.

Espinosa-Aranda, J.M., Cuellar, A., Garcia, A., Ibarrola, G., Islas, R., Maldonado, S. & Rodriguez, F.H., 2009, Evolution of the Mexican seismic alert system (SASMEX). *Seismol. Res. Lett.,* **80**(5), 694–706.

Faenza, L. & Michelini, A., 2010, Regression analysis of MCS intensity and ground motion parameters in Italy and its application in ShakeMap. *Geophys. J. Int.,* **180**(3), 1138–1152.

Fahjan, Y.M., Alcik, H. & Sari, A., 2011, Applications of cumulative absolute velocity to urban earthquake early warning systems. *J Seismol,* **15**(2), 355–373.

Festa, G. *et al.*, 2018, Performance of earthquake early warning systems during the 2016-2017 Mw5-6.5 central Italy sequence. *Seismol. Res. Lett.,* **89**(1), 1–12.

Festa, G., Zollo, A. & Lancieri, M., 2008, Earthquake magnitude estimation from early radiated energy. *Geophys. Res. Lett.,***35,** L22307. doi: 10.1029/2008GL035576.

Friedman, J.H., 2001, Greedy function approximation: a gradient boosting machine. *Ann Stat,* **Vol. 29, No. 5 (Oct., 2001),** 1189–1232.

Gorini, A. *et al.* 2010, The Italian strong motion network. *Bull. Earthq. Eng.,* **8**(5), 1075–1090.

Hoshiba, M., Kamigaichi, O., Saito, M., Tsukada, S. & Hamada, N., 2008, Earthquake early warning starts nationwide in Japan, Eos (Washington DC)., **89**(8), 73–74.

Housner, G.W., 1952, *Intensity of ground motion during strong earthquakes,* California Institute of Technology. Unpublished.

Hsu, T.Y. & Huang, C.W., 2021, Onsite Early Prediction of PGA Using CNN With Multi-Scale and Multi-Domain P-Waves as Input. *Front Earth Sci (Lausanne),* **9**. doi: 10.3389/feart.2021.626908.

Iaccarino, A.G., Gueguen, P., Picozzi, M. & Ghimire, S., 2021, Earthquake Early Warning System for Structural Drift Prediction Using Machine Learning and Linear Regressors. *Front Earth Sci (Lausanne),* **9**. doi: 10.3389/FEART.2021.666444.

Iaccarino, A.G., Picozzi, M., Bindi, D. & Spallarossa, D., 2020, Onsite earthquake early warning: predictive models for acceleration response spectra considering site effects. *Bull. Seismol. Soc. Am.,* **110**(3), 1289–1304.

Kanamori, H., 2005, Real-time seismology and earthquake damage mitigation. *Annu. Rev. Earth Planet. Sci.,* **33,** 195–214.

Kuang, W., Yuan, C. & Zhang, J., 2021, Real-time determination of earthquake focal mechanism via deep learning. *Nat. Commun.,* **12,**1432. doi: 10.1038/s41467-021-21670-x.

Mehta, P., Bukov, M., Wang, C.H., Day, A.G.R., Richardson, C., Fisher, C.K. & Schwab, D.J., 2019, A high-bias, low-variance introduction to Machine Learning for physicists. *Phys. Rep.,* **810,** 1–124.

Meier, M.A., 2017, How "good" are real-time ground motion predictions from Earthquake Early Warning systems?, *J. Geophys. Res. Solid Earth,* **122**(7), 5561–5577.

Minson, S.E., Baltay, A.S., Cochran, E.S., Hanks, T.C., Page, M.T., McBride, S.K., Milner, K.R. & Meier, M.A., 2019, The Limits of Earthquake Early Warning Accuracy and Best Alerting Strategy. *Sci. Rep.,* **9,** 2478. doi: 10.1038/s41598-019-39384-y.

Mousavi, S.M. & Beroza, G.C., 2020, Bayesian-Deep-Learning Estimation of Earthquake Location from Single-Station Observations. *IEEE Trans. Geosci. Remote Sens.,***58**(11), 8211–8224.

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

Münchmeyer, J., Bindi, D., Leser, U. & Tilmann, F., 2021, The transformer earthquake alerting model: a new versatile approach to earthquake early warning. *Geophys. J. Int.,* **225**(1), 646–656.

Münchmeyer, J., Bindi, D., Sippl, C., Leser, U. & Tilmann, F., 2020, Low uncertainty multifeature magnitude estimation with 3-D corrections and boosting tree regression: application to North Chile. *Geophys. J. Int.,* **220**(1), 142–159.

Ochoa, L., Niño, L. & Vargas, C., 2018, Fast magnitude determination using a single seismological station record implementing machine learning techniques. *Geod Geodyn,* **9**(1),34–41.

Parolai, S., Bindi, D., Boxberger, T., Milkereit, C., Fleming, K. & Pittore, M., 2015, On-Site Early Warning and Rapid Damage Forecasting Using Single Stations: outcomes from the REAKT Project. *Seismol. Res. Lett.,* **86**(5), 1393–1404.

Pedregosa, F. *et al.*, 2011, Scikit-learn: machine Learning in Python, Machine Learning in Python. *J. Mach. Learn Res.,* **12,** 2825–2830.

Picozzi, M., Bindi, D., Brondi, P., di Giacomo, D., Parolai, S. & Zollo, A., 2017, Rapid determination of P wave-based energy magnitude: insights on source parameter scaling of the 2016 Central Italy earthquake sequence. *Geophys. Res. Lett.,* **44**(9), 4036–4045.

Picozzi, M., Bindi, D., Pittore, M., Kieling, K. & Parolai, S. 2013. Real-time risk assessment in seismic early warning and rapid response: A feasibility study in Bishkek (Kyrgyzstan), *Journal of Seismology,* **17**(2), 485–505.

Picozzi, M., Elia, L., Pesaresi, D., Zollo, A., Mucciarelli, M., Gosar, A., Lenhardt, W. & Živcíc, M., 2015a, Trans-national earthquake early warning (EEW) in north-eastern Italy, Slovenia and Austria: first experience with PRESTo at the CE3RN network. *Advances in Geosciences,* **40,** 51–61.

Picozzi, M., Emolo, A. *et al.*, 2015b, Earthquake Early Warning System for Schools: a Feasibility Study in Southern Italy. *Seismol. Res. Lett.,* **86**(2A), 398–412.

Picozzi, M., Zollo, A., Brondi, P., Colombelli, S., Elia, L. & Martino, C., 2015c, Exploring the feasibility of a nationwide earthquake early warning system in Italy. *J. Geophys. Res. Solid Earth,* **120**(4), 2446–2465.

Raschka, S. & Mirjalili, V., 2017, Python Machine Learning: machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow.

Satriano, C., Wu, Y.M., Zollo, A. & Kanamori, H., 2011, Earthquake early warning: concepts, methods and physical grounds. *Soil Dyn. Earthq. Eng.,* **31**(2), 106–118.

Scafidi, D., Viganò, A., Ferretti, G. & Spallarossa, D., 2018, Robust Picking and Accurate Location with RSNI-Picker2: real-Time Automatic Monitoring of Earthquakes and Nontectonic Events. *Seismol. Res. Lett.,* **89**(4), 1478–1487.

Song, J., Zhu, J., Wang, Y. & Li, S., 2022, On-site alert-level earthquake early warning using machine-learning-based prediction equations. *Geophys. J. Int.,* **231**(2), 786–800.

Spallarossa, D., Cattaneo, M., Scafidi, D., Michele, M., Chiaraluce, L., Segou, M. & Main, I.G., 2021a, An automatically generated high-resolution earthquake catalogue for the 2016–2017 Central Italy seismic sequence, including P and S phase arrival times. *Geophys. J. Int.,* **225**(1), 555–571.

Spallarossa, D., Ferretti, G., Scafidi, D., Turino, C. & Pasta, M., 2014, Performance of the RSNI-Picker. *Seismol. Res. Lett.,* **85**(6), 1243–1254.

Spallarossa, D., Kotha, S.R., Picozzi, M., Barani, S. & Bindi, D., 2019, On-site earthquake early warning: a partially non-ergodic perspective from the site effects point of view. *Geophys. J. Int.,* **216**(2), 919–934.

Spallarossa, D., Picozzi, M., Scafidi, D., Morasca, P., Turino, C. & Bindi, D., 2021b, The RAMONES Service for Rapid Assessment of Seismic Moment and Radiated Energy in Central Italy: concepts, Capabilities, and Future Perspectives. *Seismol. Res. Lett.,* **92**(3), 1759–1772.

Wolfe, C.J., 2006, On the Properties of Predominant-Period Estimators for Earthquake Early Warning. *Bull. Seismol. Soc. Am.,* **96**(5), 1961–1965.

Wu, Y.M. & Kanamori, H., 2005a, Experiment on an onsite early warning method for the Taiwan early warning system. *Bull. Seismol. Soc. Am.,* **95**(1), 347–353.

Wu, Y.M. & Kanamori, H., 2005b, Rapid assessment of damage potential of earthquakes in Taiwan from the Beginning of P waves. *Bull. Seismol. Soc. Am.,* **95**(3), 1181–1185.

Wu, Y.M. & Kanamori, H., 2008, Development of an earthquake early warning system using real-time strong motion signals. *Sensors,* **8**(1), 1–9.

Wu, Y.M. & Zhao, L., 2006, Magnitude estimation using the first three seconds P-wave amplitude in earthquake early warning. *Geophys. Res. Lett.,* **33**(16), L16312. doi: 10.1029/2006GL026871.

Zhu, W. & Beroza, G.C., 2019, PhaseNet: a deep-neural-network-based seismic arrival-time picking method. *Geophys. J. Int.,***216**(1), 261–273.

Zollo, A. *et al.*, 2009, Earthquake early warning system in southern Italy: methodologies and performance evaluation. *Geophys. Res. Lett.,* **36**(5),L00B07. doi: 10.1029/2008GL036689.

Zollo, A., Amoroso, O., Lancieri, M., Wu, Y.M. & Kanamori, H., 2010, A threshold-based earthquake early warning using dense accelerometer networks. *Geophys. J. Int.,* **183**(2), 963–974.

Zollo, A., Colombelli, S., Elia, L., Emolo, A., Festa, G., Iannaccone, G., Martino, C. & Gasparini, P., 2014, An integrated regional and on-site Earthquake Early Warning system for southern Italy: concepts, methodologies and performances. 117–137. doi: 10.1007/978-3-642-12233-0_7.