

A Machine-Learning Approach for the Reconstruction of Ground-Shaking Fields in Real Time

Simone Francesco Fornasari^{1*}, Veronica Pazzi¹, and Giovanni Costa¹

ABSTRACT

Real-time seismic monitoring is of primary importance for rapid and targeted emergency operations after potentially destructive earthquakes. A key aspect in determining the impact of an earthquake is the reconstruction of the ground-shaking field, usually expressed as the ground-motion parameter. Traditional algorithms compute the ground-shaking field from the punctual data at the stations relying on ground-motion prediction equations computed on estimates of the earthquake location and magnitude when the instrumental data are missing. The results of such algorithms are then subordinate to the evaluation of location and magnitude, which can take several minutes. To fill the temporal gap between the arrival of the data and the estimate of these parameters, a new data-driven algorithm that exploits the information from the station data only is introduced. This algorithm, consisting of an ensemble of convolutional neural networks (CNNs) trained on a database of ground-shaking maps produced with traditional algorithms, can provide estimates of the ground-shaking maps and their associated uncertainties in real time. Because CNNs cannot handle sparse data, a Voronoi tessellation of a selected peak ground parameter recorded at the stations is computed and used as the input to the CNNs; site effects and network geometry are accounted for using a (normalized) V_{S30} map and a station location map, respectively. The developed method is robust to noise, can handle network geometry changes over time without the need for retraining, and can resolve multiple simultaneous events. Although having a lower resolution, the results obtained are statistically compatible with the ones from traditional methods. A fully operational version of the algorithm is running on the servers at the Department of Mathematics and Geosciences of the University of Trieste, showing real-time capabilities in handling stations from multiple Italian strong-motion networks and outputting results with a resolution of $0.05^\circ \times 0.05^\circ$.

KEY POINTS




- Real-time seismic monitoring is of primary importance for rapid and targeted emergency operations.
- ShakeRec is robust to noise, handles network geometry changes, resolves multiple events, and operates in real time.
- ShakeRec is a useful tool for civil protection purposes during the early stages of an emergency.

[Supplemental Material](#)

INTRODUCTION

Ground-shaking maps serve a wide spectrum of purposes, from postearthquake emergency management and response (Wald, Lin, and Quitoriano, 2008; Wald *et al.*, 2020) to engineering (Allen *et al.*, 2009) and financial instruments (Wald, Lin, Porter, and Turner, 2008) to general scientific applications

(Moratto *et al.*, 2011). Multiple ground-motion parameters (GMPs) have been introduced in the literature to represent different characteristics of strong-motion recordings. GMPs can be classified into time-domain parameters and frequency-domain parameters: the former are derived directly or with minimal processing from the strong-motion recordings, and the latter are derived from Fourier or response spectra. Typical examples of time-domain parameters are peak ground acceleration (PGA), peak ground velocity (PGV), and peak ground displacement

1. Department of Mathematics and Geosciences, SeisRaM Working Group, University of Trieste, Trieste, Italy,  <https://orcid.org/0000-0001-8855-4478> (SFF);  <https://orcid.org/0000-0002-9191-0346> (VP);  <https://orcid.org/0000-0002-0656-2499> (GC)

*Corresponding author: simonefrancesco.fornasari@phd.units.it

Cite this article as Fornasari, S. F., V. Pazzi, and G. Costa (2022). A Machine-Learning Approach for the Reconstruction of Ground-Shaking Fields in Real Time, *Bull. Seismol. Soc. Am.* **112**, 2642–2652, doi: [10.1785/0120220034](https://doi.org/10.1785/0120220034)

© Seismological Society of America

TABLE 1

Ground-Motion Prediction Equations (GMPEs) for the Italian Territory

Region	Depth Range (km)	GMPEs
Shallow active crustal regions	<35	Bindi et al. (2011)
Volcanic areas	<5	Tusa and Langer (2016)
	>5	Bindi et al. (2011)
Deep events	>35	Bindi et al. (2014)
	5–35	Bindi et al. (2011)
Subduction zone	35–70	Bindi et al. (2014)
	>70	Abrahamson, Gregor, and Addo (2016)

(PGD). PGA is a measure of strong-motion amplitude, but it is generally related to the motion at shorter periods (<0.3 s), which is of interest only for the monitoring of very specific structures. PGV and PGD have gained increasing attention in the seismic engineering community due to their relation to motion at a period range of more interest for buildings. PGV is also considered to be related to the macroseismic intensity and thus a good indicator of the structural damaging potential of the ground motion ([Wald, Quitoriano, Heaton, and Kanamori, 1999](#); [Atkinson, 2003](#)). Spectral accelerations (SAs) are an example of frequency-domain parameters and are defined as the peak time-domain response to the ground motion of a single degree of freedom oscillator, with a given natural period and critical damping, usually 5%. De facto standard periods for the SAs are 0.3, 1.0, and 3.0 s, which are of interest for the engineering community. Because GMPs are retrieved from instrumental recordings only at a very limited number of points, the goal of this work is to best constrain each GMP at any location using a defined interpolation technique. ShakeMap ([Wald, Quitoriano, Heaton, Kanamori, Scrivner, and Worden, 1999](#); [Wald, Worden, Thompson, and Hearne, 2021](#)) is a multiperiod geospatial interpolation process constrained with limited (often imprecise) data; this interpolation is not only inherently uncertain but also fundamentally nonunique. ShakeMap outputs separate maps (hereinafter referred to as shakemaps) of the spatial distribution of peak ground motions (displacement, velocity, and acceleration), SAs, and instrumental derived seismic intensity. The interpolation model is constrained by site amplified ground-motion prediction equations (GMPEs) based on rapidly determined source parameters and conditioned by strong motion or macroseismic observations. ShakeMap uses an elaborate system of multiple-weighted GMPEs to appropriately model the ground motion, whereas the site amplifications are introduced using the time-averaged shear-wave velocity in the upper 30 m of the crust (V_{S30}). A fault model could be introduced in ShakeMap with benefits depending on several factors such as event size and location and station coverage, as well as the use of “Did You Feel It?” data. The macrointensity maps are produced using GMPEs and

ground motion to intensity conversion equations (GMICES). All ShakeMap estimates are probabilistic with several sources of uncertainty: the GMPEs used are nonunique, often with competing models for each tectonic region and event type; the use of V_{S30} map as a proxy for site amplification introduces uncertainties; and rapid estimate of earthquake parameters also introduces uncertainties, which trades accuracy for speed. ShakeMap also computes the uncertainty related to the estimates at each mapped point. Its greater utility is, therefore, in the early emergency stages, but there are intrinsic limitations to be aware of such as the approximations introduced using GMPEs and the probabilistic nature of the results ([Wald et al., 2021](#)), as well as the way site amplification is handled ([Cultrera et al., 2014](#)).

In Italy, the Italian Civil Protection (Dipartimento per la Protezione Civile, DPC) uses shakemaps, produced routinely by Istituto Nazionale di Geofisica e Vulcanologia (INGV, [Michelini et al., 2020](#)), together with other independent information to assess the areas affected by an earthquake. Four GMPEs are used for zones with different prevailing tectonic regimes and earthquake depths as reported in Table 1. The V_{S30} map of Italy, actually in use by INGV, was developed by [Michelini et al. \(2020\)](#). Using automated procedures ([Margheriti et al., 2021](#)), INGV communicates to DPC the preliminary parameters of every seismic event of magnitude $M \geq 3.0$ in the Italian territory. The rapid and final automatic locations are produced in 2 and 5 min after the earthquake occurrence, respectively; the verified location, revised by the seismologists on duty in the control room, are available, at most, 30 min after the event (12 min on average). Only once the origin and magnitude are manually evaluated, the automated ShakeMap procedure is triggered.

A different implementation of ShakeMap for the generation of (near-)real-time shakemaps in the southeastern Alps was proposed by [Moratto et al. \(2009\)](#) who calibrated the software for the specific geological settings of the area, which was subdivided into three classes based on the geological map by [DiSGAM \(2005\)](#), and used region-specific GMPEs, as reported in Table 2. The procedure, integrated into the BRTT Antelope system and automated, combines data from different national networks (northeast Italy, Austria, and Slovenia) and generates shakemaps within 5 min of the event.

This work describes the development of a new tool for the reconstruction in real time of ground-shaking maps, hereinafter referred to as ShakeRec. The lack of information such as earthquake location and magnitude in real-time analysis implies that GMPEs cannot be applied and a machine-learning (ML) approach is used instead. ML is a form of data analysis that allows for the development of computational models for which parameters are extracted directly from data using well-defined optimization rules. Deep learning (DL) is a subset of ML focused on extracting useful features and patterns from data using neural networks. DL has gone into a resurgence phase since the early 2000s due to the improved and more easily accessible high-performance computational resources

TABLE 2
GMPEs for the Southeastern Alps

Magnitude Range	PGA and PGV	SA
2.5–3.5	Bragato and Slejko (2005)	–
3.5–5.5	Massa <i>et al.</i> (2008)	Massa <i>et al.</i> (2008)
>5.5	Sabetta and Pugliese (1996)	Akkar and Bommer (2007)

PGA, peak ground acceleration; PGV, peak ground velocity; SA, spectral acceleration.

(especially graphic processing units), the development of a variety of open-source software that simplifies the design and use of DL models, and the availability of big datasets. LeCun *et al.* (2015) give a concise but comprehensive review of the development of DL. Seismology itself is undergoing rapid changes in the volume and variety of the data and the velocity of data analysis. As a consequence, there is a widespread application of ML models to seismological tasks; Kong *et al.* (2018) reviewed the main field of ML applications in seismology. Ensembles are well-established methods that are introduced to limit some of the shortcomings of traditional ML methods (Dietterich, 2000). Viewing a learning algorithm as searching a space of hypotheses, ensemble techniques can limit the problems arising from the sparse sampling of such space of hypotheses that are due to limited availability of training data, the learning algorithm getting stuck in local minima during the search, and the target function being outside the searching space. Furthermore, the ensemble spread, defined as the (non-biased) standard deviation of the outputs of the ensemble members, provides a measure of the level of (epistemic) uncertainty of the output of the ensemble (World Meteorological Organization [WMO], 2012). ShakeRec is based on the work of Fukami *et al.* (2021) who proposed a data-driven spatial field recovery technique that makes use of Voronoi tessellation (Voronoi, 1908; Aurenhammer, 1991) to obtain a structured grid representation to enable computational, tractable use of convolutional neural networks (CNNs). This architecture, first introduced in its modern conception by LeCun *et al.* (1989), was developed as one of the main DL algorithms to handle matrix-like data due to its low number of learnable parameters (with respect to fully connected neural networks) and its position-invariant pattern recognition ability. Each convolutional layer is defined on the base of its filters, and each one is a collection of weights learned during the training phase that addresses a different feature of the input data that are convolved; the output of the convolution (added to a bias term) is then passed to an activation function.

The model architecture is presented in the [Model architecture](#) section. In the [Model training](#) section, the preparation of the datasets and the details of the model training are explained. The results of the tests performed with ShakeRec are detailed in the [Results and discussion](#) section.

MODEL ARCHITECTURE

The objective of this work is to reconstruct a 2D ground-shaking field from measurements at the seismic stations. The conceptual workflow of ShakeRec for a single GMP is shown in Figure 1. The reconstruction for a single GMP is performed using an ensemble of five CNN models, thereafter called ensemble members, depicted by the central panel in Figure 1. The architecture of each ensemble member was adapted from Fukami *et al.* (2021): each ensemble member is composed of four convolutional layers, with zero-padding, with 12 filters of size 5×5 and a final convolutional layer with a single filter of size 5×5 . This architecture was developed considering both its performance in reconstructing the ground-shaking field and its run time. Each layer uses a rectified linear activation function (ReLU) activation function (Nair and Hinton, 2010) to introduce nonlinearity in the model. The activation function threshold is set to 0.01, so it is high enough to remove small computational artifacts and improve the overall loss function values but small enough not to remove any useful information.

The ensemble takes the following as input (left side of Fig. 1):

1. A Voronoi tessellation of the station peak ground values over the last T seconds.
2. A map of the active station locations.
3. A normalized V_{S30} map, from which a mask of the sea is computed and applied to the output.

The Voronoi tessellation and the map of the active stations provide a full description of the value at the stations, in terms of GMP amplitude and location of the recording, in a format that is manageable by a CNN, whereas the V_{S30} map is introduced as a proxy for the site effects. All ensemble inputs are tensors of the shape $(1, N, M, 1)$ with N, M being the number of grid points along the latitudinal and longitudinal directions, respectively.

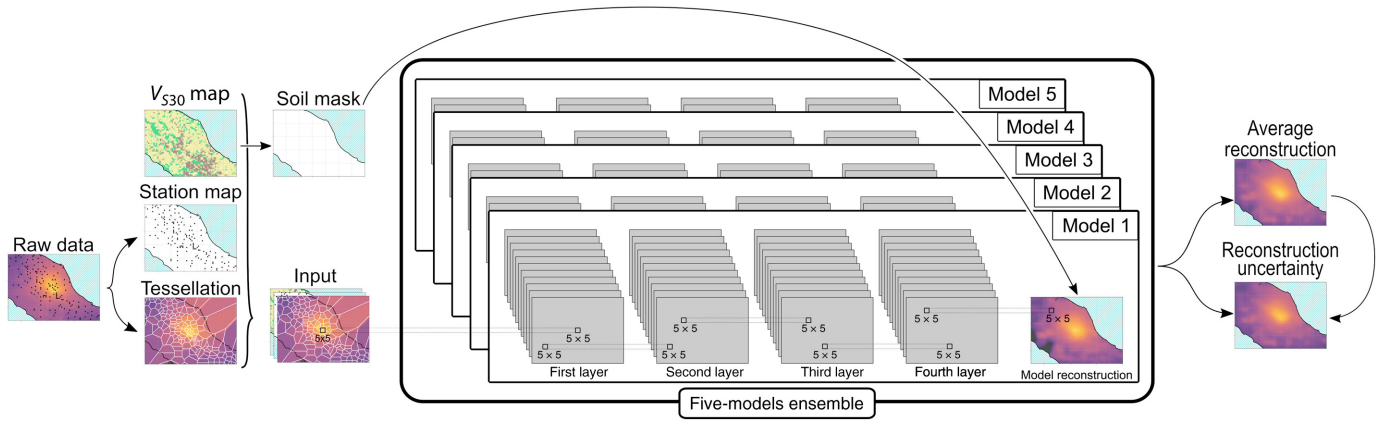
The outputs of the ensemble are the mean ground-shaking field $\tilde{\mu}$ and its uncertainty $\tilde{\sigma}$ (right side of Fig. 1). The former is computed as the mean of the ensemble members' reconstructions and is comparable to the GMPE expected value at each point; the latter is computed as the combination of an aleatoric σ_a and an epistemic σ_e term as

$$\tilde{\sigma} = \sqrt{\sigma_a^2 + \sigma_e^2}. \quad (1)$$

Because the GMPEs are usually expressed as log-normal distributions, the aleatoric term σ_a of the uncertainty is computed as

$$\sigma_a = \tilde{\mu} \sqrt{e^{\ln(10)^2 \sigma_G^2} - 1}, \quad (2)$$

with σ_G being the mean standard deviation of the GMPEs used in Michelini *et al.* (2020) (expressed in log-10 units). The



epistemic component σ_e of uncertainty is computed as the standard deviation of the reconstructions of the ensemble members:

$$\sigma_e = \sqrt{\frac{1}{m} \sum_{i=0}^m \mu_i^2 - \left(\frac{1}{m} \sum_{i=0}^m \mu_i \right)^2}, \quad (3)$$

with m being the number of members in the ensemble and μ_i being the prediction of a single ensemble member.

MODEL TRAINING

The training dataset for each GMP (namely, PGA, PGV, SA03, SA10, and SA30) is generated from the corresponding INGV shakemap database. The location, magnitude, and depth of the events in this database are reported in Fig. 2. Each shakemap of the original database has been interpolated to the area 6.5° – 19.0° E and 35.5° – 47.1° N and resampled with a resolution of $0.05^\circ \times 0.05^\circ$. The area corresponds to the overall area for which the INGV shakemaps are evaluated. The resolution is chosen to limit the size of the network inputs and to reflect the real capabilities in reconstructing the ground-shaking field of the model. The processed data are matrices of shape 258×238 considered to be ground truth. The following seismic networks have been considered: the Italian strong-motion network (RAN), Friuli Venezia Giulia accelerometric network (RAF), Irpinia seismic network (ISNet), northeast Italy broadband network (NI), and Reti accelerometriche centri storici italiani (RR); see [Data and Resources](#). The GMP values at the stations are drawn from the processed maps considering only the stations that were active during each specific event. Taking the value from the processed maps (as opposed to taking the real value recorded at the station) ensures that the input data from the stations and the corresponding target shaking maps have a one-to-one correspondence. The position of the station and the peak values are used to generate the Voronoi tessellation. The shakemaps with extensions that do not involve any station of the selected networks, with maximum values that differ by more than an order of magnitude from the maximum value at the stations, or with maximum values that are below the 0.1%g

Figure 1. ShakeRec flowchart: real-time and static data are used as the input for the five-models ensemble, the outputs of which are the average reconstructed ground-motion parameters (GMP) field and its associated uncertainty. The color version of this figure is available only in the electronic edition.

(or 0.1 cm s^{-1} for PGV) threshold are discarded to stabilize the CNN training process. Many of the discarded events happened at sea, were outside the Italian territory, or were small events in regions with poor instrumental distribution. The processed dataset is split into a training set and a test set with a ratio of 4:1.

The V_{S30} map developed by [Michelini et al. \(2020\)](#) is normalized using min–max scaling and used as input for the convolutional neural network. Within the neural network, the normalized map is also used to produce a sea mask used to zero all offshore grid points; this step is useful for improving the results of the network in the area of interest.

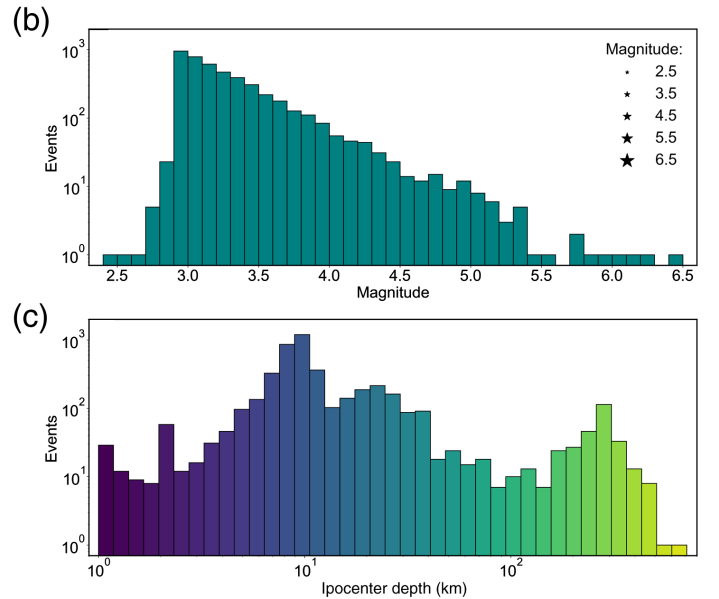
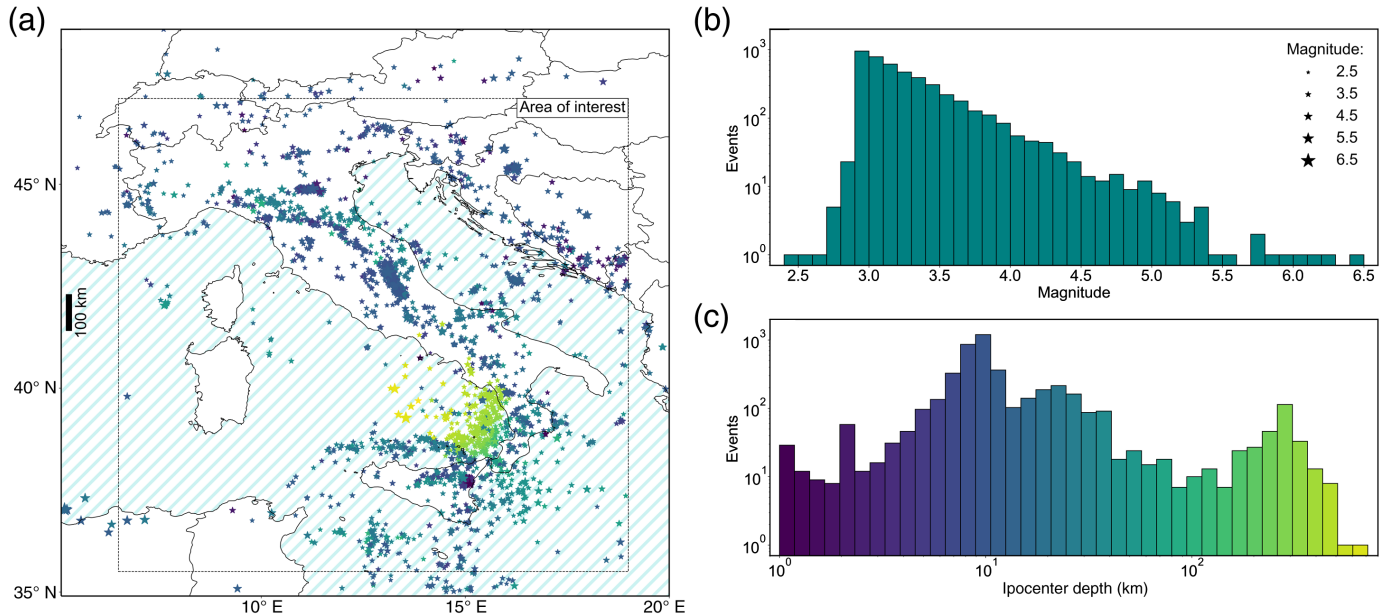
The loss function \mathcal{L} has the form

$$\mathcal{L} = \frac{\|\mu - \tilde{\mu}\|_2}{\|\mu\|_2}, \quad (4)$$

with μ and $\tilde{\mu}$ being the expected and reconstructed ground-shaking fields, respectively. The use of a normalized loss function promotes small relative differences between the expected and reconstructed values.

The optimizer chosen is Adam ([Kingma and Ba, 2017](#)) using a learning rate of 0.001, as suggested by the parametric studies performed.

The 10-fold cross validation is adopted to train the models composing the ensemble. The training dataset is divided in 10 chunks, with one of each routinely used for validation and the others used for training. Ensembles constructed with all models trained in this way are called cross-validated committees ([Parmanto et al., 1996](#)); a further step is performed selecting only five of them to limit the computational costs. All possible combinations of the 10 trained models are tested using the test



dataset, and the five-model ensemble configuration giving the best results, in terms of loss function value, is selected. A schematic representation of this process is depicted in Figure 3.

The training is done using sample batches (of 48 samples) to limit the memory requirement and to speed up the training process.

Each model is trained for a maximum of 500 epochs with early stopping in the case of the validation error not improving for 50 consecutive epochs, which restores the model weights from the epoch with the lowest validation error.

The whole process is performed only once for each GMP, resulting in a multiple ensemble. The model obtained is then saved and can be exported to other systems (providing that they operate on the same area). As discussed in Results and discussion section, the model retraining is necessary only for major changes in the network geometry or, obviously, if the area where it would be applied is changed.

RESULTS AND DISCUSSION

ShakeRec provides reliable results in reconstructing the ground-shaking field with average losses over the test datasets reported in Table 3. The reconstructions with the highest loss values are generally relative to very small events, badly constrained offshore events, or localized intense events (e.g., linked to eruptions). Being the results of the upcoming tests similar for all GMPs considered, they are reported only for the PGA for the sake of brevity.

To give an example of the capabilities of ShakeRec, the PGA field for the 30 October 2016 M_w 6.5 Norcia earthquake (hereinafter referred to as the M_w 6.5 Norcia earthquake) was reconstructed. Figure 4a,b shows the reconstruction and uncertainty map obtained from ShakeRec and (resampled) ShakeMap, respectively. The overall agreement of the results obtained with the two methods can be noticed, as also shown by the absolute

Figure 2. (a) Location, (b) magnitude, and (c) depth of the events in the Istituto Nazionale di Geofisica e Vulcanologia (INGV) shakemap archive used to develop the training and test datasets. The size of the markers in panel (a) is proportional to the event magnitude and their color indicates the event depth, using the same chromatic scale as in panel (c). The color version of this figure is available only in the electronic edition.

difference of the reconstructions reported in Figure 4c. On a side note, the relatively high value of the uncertainty with respect to the reconstruction hints to using shaking maps mainly as a monitoring tool, as already addressed in the Introduction section. Similar results are obtained for the other GMPs considered and are reported in the supplemental material available to this article.

The M_w 6.5 Norcia earthquake is considered the most suitable event in the test dataset because the high number of stations involved reduces the effects of gaps in the network but at the same time involves an area sufficiently broad to be constrained differently by different network geometries (considering also the resolution of the model). Because both ShakeMap and ShakeRec provide estimates for the uncertainties of the ground-shaking field, the agreement of the results is tested through a T -test for each point on the map. On average by fixing the confidence level to 0.05, less than 1 grid point value per map would be rejected. This is also a consequence of, except for the area proximal to the epicenter, the small values and relatively high uncertainties that characterize the grid points. To assess the robustness of the ensemble reconstructions to noise, different levels of Gaussian noise are added to the local sensor measurements. Considering the M_w 6.5 Norcia earthquake, the station data are perturbed with Gaussian noise with an amplitude of a given percentage of the initial value. To obtain stable results, the process is repeated

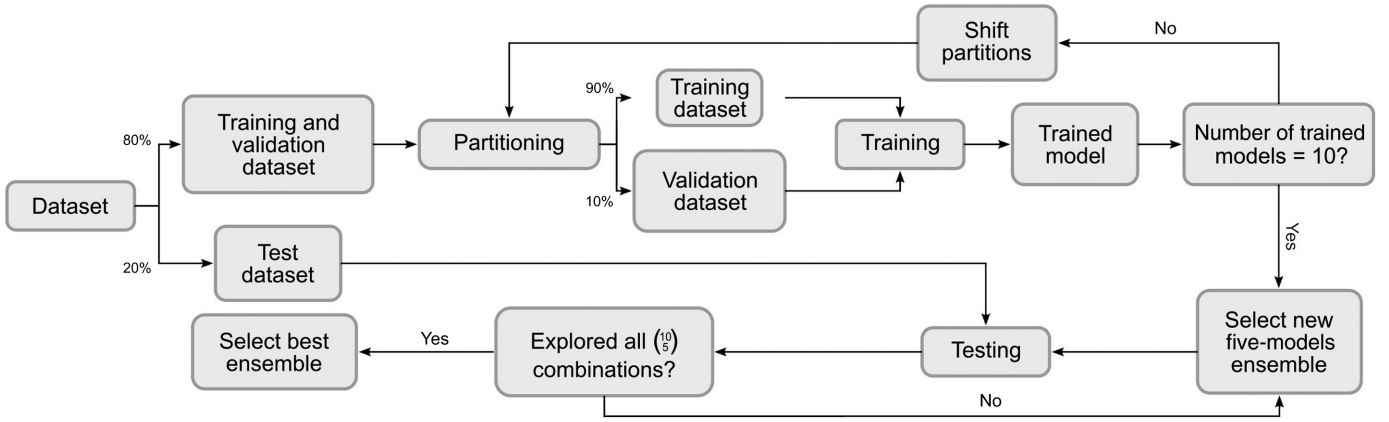


Figure 3. Workflow of ShakeRec training: the process is performed once for each GMP ensemble and defines the best five-model ensemble, in terms of loss over the test dataset, by testing all possible combinations of models, each trained using different training and validation datasets.

TABLE 3
Average Ensemble Loss Over the Test Dataset

Peak Ground Parameter	Average Loss
PGA	0.160
PGV	0.175
SA03	0.170
SA10	0.229
SA30	0.222

multiple times, and the mean value and standard deviation of the loss are taken. The results, reported in Table 4, show good robustness against noise, with great increase in the loss values only for low signal-to-noise ratios.

Some interesting features arise from the use of the CNN in the model architecture. In the case of simultaneous events, ShakeRec can reconstruct the resulting joint ground-shaking field, even if the areas affected by the events spatially overlap, because the reconstruction is a local process. This feature is tested by simulating the occurrence of earthquakes with target areas that partially overlap, using data from the 2012 M_w 4.9 Brescello, M_L 4.8 Mirandola, and M_w 5.8 Finale Emilia earthquakes. In the cases of events of different magnitudes (case I in Fig. 5) and similar magnitudes (case II in Fig. 5), the proposed method can reconstruct reasonably well the ground-shaking field and separate the areas involving the two events. This feature can be particularly useful during actual real-time monitoring situations when the analysis is performed on moving windows with lengths that could be involved with more than one event, especially during seismic sequences. In this situation, because the analysis is performed on the peak values recorded at the stations, multiple events occurring in the same time window are treated as simultaneous, and the ability to reconstruct their joint ground-shaking field is essential to having an accurate representation of their effects.

Another interesting feature of the model is that it is not necessary to specify a GMPEs (or multiple ones for different tectonic and regional conditions) as this selection is implicit in the training data used. Because INGV uses different GMPEs for different regional and tectonic settings, ShakeRec will implicitly learn a mixture of the main features of such models. ShakeRec can deal with network geometry changes over time. This feature is particularly useful when dealing with real data because changes in the seismic network, both temporary (e.g., offline stations and temporary stations) and permanent (e.g., addition or removal of stations), can be handled with low loss increase and no need for model retraining. The effects of the network geometry, both in terms of the number of stations and station distribution, are tested considering the M_w 6.5 Norcia

earthquake. The test is performed considering trained and untrained station positions. Fixing the number of stations N_s , the test is then repeated 100 times choosing different random configurations to study the effect of different station distributions. The results in Figure 6 show that the mean loss follows an hyperbolic growth with respect to the number of affected stations $\mathcal{L} \propto \frac{1}{N_s}$; the standard deviation values increase as the number of the affected station N_s decreases as an effect of the lower station coverage and the increasing importance of the station geometry. The importance of the network geometry also is reflected in the better performance obtained using untrained network positions, which densely populate the affected area in the test.

One of the downsides of the use of CNNs is that the values recorded at the stations are not maintained in the reconstruction. Even though this can be solved by reintroducing such values in the final output, because there is an overall agreement of the values at the stations, the smoother CNN reconstruction was preferred.

Finally, the reconstruction ability of ShakeRec based on data available in real time is tested by reconstructing the PGA field for the M_w 6.5 Norcia event from the values extracted from the recorded waveforms. In Figure 7, the reconstructions based on

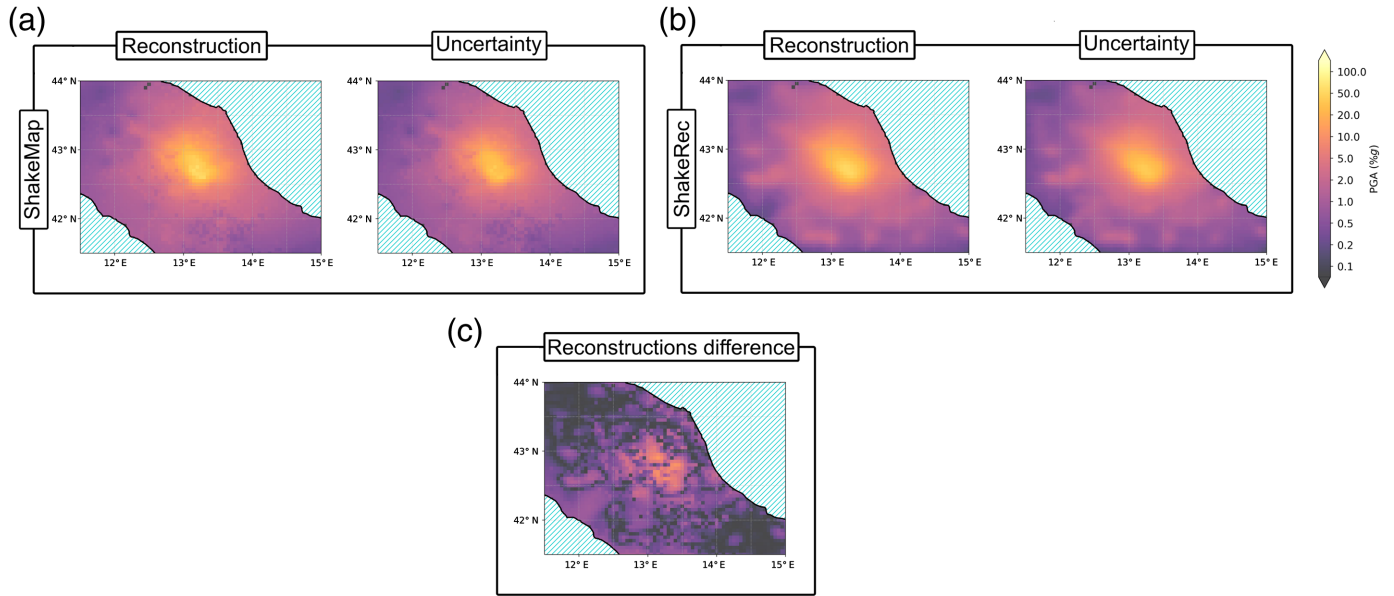


Figure 4. Reconstruction of the peak ground acceleration (PGA) field for the M_w 6.5 Norcia earthquake: (a) the (resampled) reconstruction and uncertainty from ShakeMap, (b) the reconstruction and uncertainty obtained

with ShakeRec, and (c) the absolute difference of the two reconstructions. All figures share the same colormap. The color version of this figure is available only in the electronic edition.

the data available at different times ΔT after the event are shown. It is noticeable how the reconstruction changes only at further epicentral distances as more time elapses since the event. This simulation represents the optimal scenario

in which no delays due to data transmission or computation are introduced. In practice, the same reconstructions would be obtained with a few seconds delay with respect to the time shown.

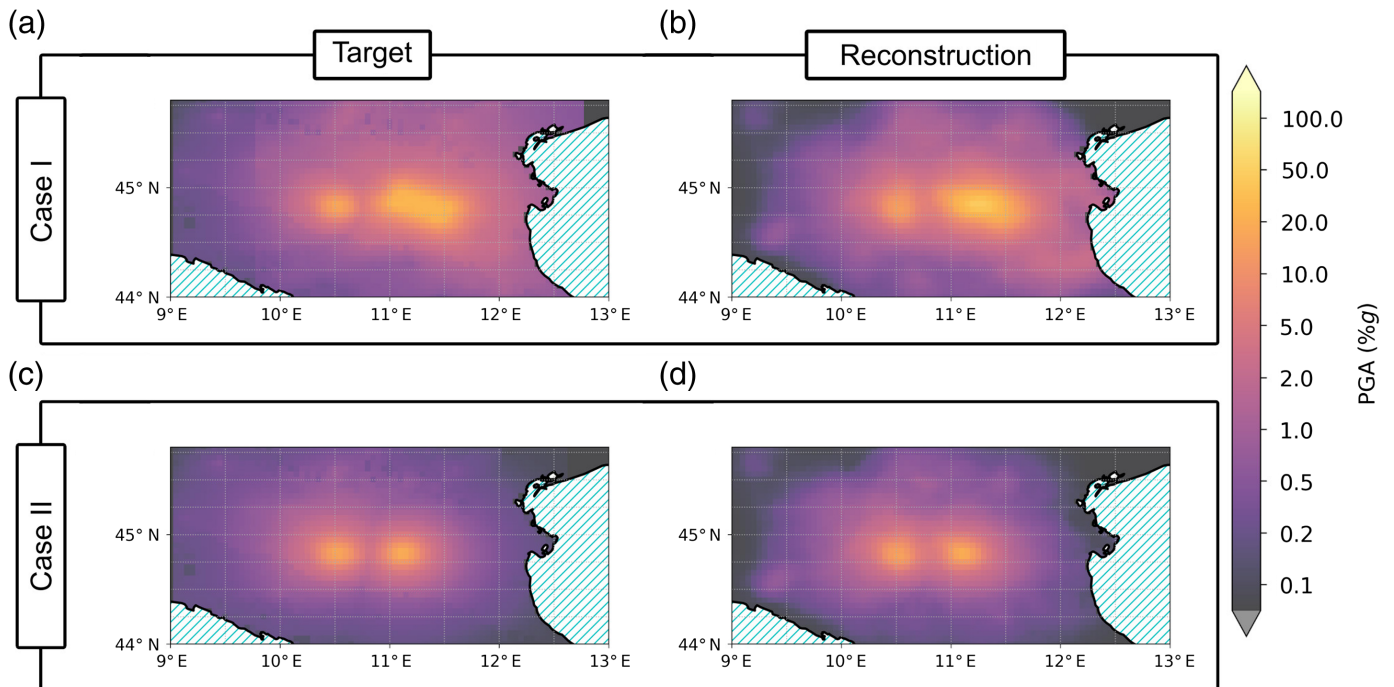


Figure 5. Reconstruction of the PGA field for multievent scenarios: M_w 4.9 Brescello and M_w 5.8 Finale Emilia earthquakes (case I), and M_w 4.9 Brescello and M_L 4.8 Mirandola (case II) are assumed simultaneous. (a,c) The

superposition of the shakemaps for the two events computed independently. (b,d) The ShakeRec reconstructions of the multievent scenarios. The color version of this figure is available only in the electronic edition.

TABLE 4

Noise Robustness Test Results*

Noise Level (%)	PGA	PGV	SA03	SA10	SA30
0	0.160	0.175	0.170	0.229	0.222
5	0.163 ± 0.007	0.177 ± 0.005	0.172 ± 0.005	0.230 ± 0.007	0.224 ± 0.004
10	0.169 ± 0.015	0.183 ± 0.009	0.178 ± 0.010	0.237 ± 0.015	0.230 ± 0.007
25	0.214 ± 0.032	0.229 ± 0.032	0.217 ± 0.037	0.270 ± 0.032	0.272 ± 0.022
50	0.355 ± 0.088	0.367 ± 0.091	0.359 ± 0.084	0.375 ± 0.083	0.377 ± 0.046
100	0.669 ± 0.148	0.605 ± 0.144	0.681 ± 0.119	0.560 ± 0.133	0.585 ± 0.079

*Mean and standard deviation of the loss value over 100 trials for the M_w 6.5 Norcia earthquake.

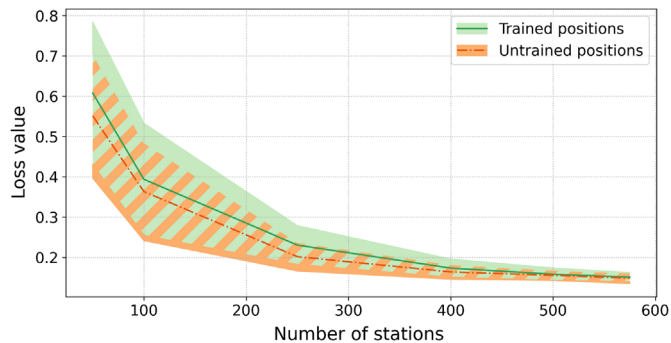


Figure 6. Mean and standard deviation of the loss as a function of the total number of seismic stations for the M_w 6.5 Norcia earthquake reconstruction for both trained and untrained station positions. The color version of this figure is available only in the electronic edition.

A fully operational real-time implementation of this method is being tested at the Department of Mathematics and Geosciences (DMG) of the University of Trieste, where it has been integrated with Antelope. The raw data from multiple accelerometric networks (i.e., RAN, RAF, ISNet, RR, and Hareia, see [Data and Resources](#)), for a total of about 400 active stations, are retrieved from the Antelope buffer memory and processed; the GMP at each station over a predefined time window, alongside the information of the active stations, is used to produce the Voronoi tessellation and the station map. Because preprocessing involves standard procedures for the real-time analysis of the data, it could be shared with other running processes to reduce computational costs. Such maps are passed to ShakeRec, which updates its outputs in under 2 s, considering a single GMP. Figure 8 shows the workflow of ShakeRec implementation at DMG.

CONCLUSIONS

Rapid ground-shaking field reconstruction is an essential task for postearthquake emergency management and response. A data-driven approach, adapted from the literature, is used to develop an algorithm with real-time reconstruction capabilities that can fill the information gap between the arrival of the data at the data center and the generation of shaking maps using traditional algorithms (e.g., ShakeMap).

Multiple tests have been conducted to show the features and limits of the approach. The model is capable of retrieving the ground-shaking field from sparse GMP data with results comparable to those obtained using traditional methods; it presents good robustness to noise, can reconstruct simultaneous events, can handle network geometry changes over time, and finally does not need a GMPE to be specified because it is learned from the training data.

The applicability of the proposed method to other geographical areas is feasible after retraining the ensemble members. In areas where only limited data are available, transfer learning could be an effective approach, although in-depth analysis must be conducted to assess the quality of the reconstructions before implementing them in a monitoring workflow.

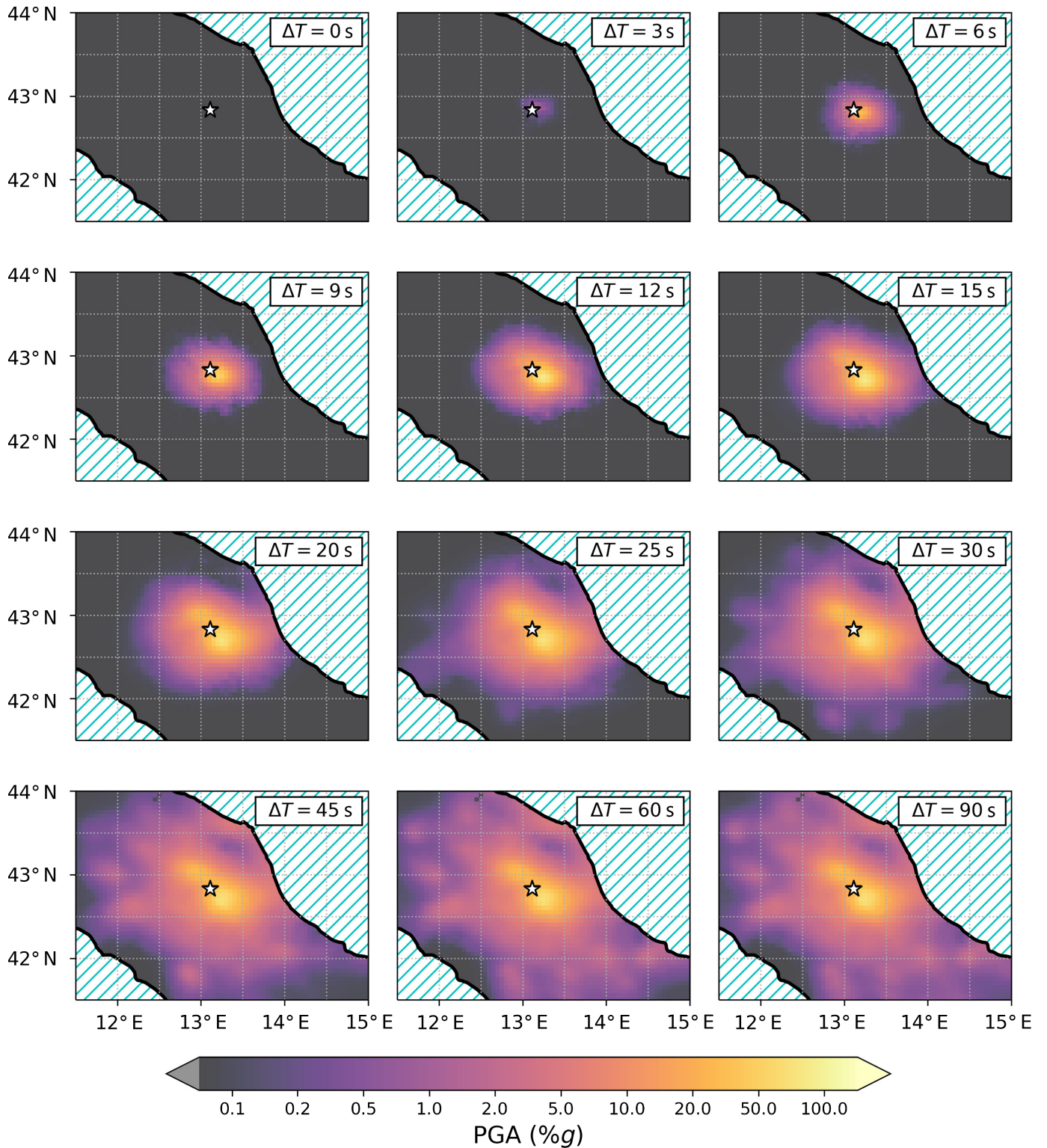
The two main limitations of ShakeRec are its inability to properly reconstruct small and outside of the network events and the overall smoothness of the outputs, compared with the ones from ShakeMap. The first is a consequence of the choices made in processing the training data and is related to the sole dependence of the model on the station data. The latter is more closely related to model training and its architecture. Both of these shortcomings depend ultimately on the choices made in the development phase and are considered to not undermine the purpose of the algorithm, which is linked to potentially destructive events. More sophisticated architecture could be tested to improve the reconstructions, but careful considerations should be made about the trade-off between the improvement obtained and the increase in complexity and computational cost, other than considering the real need in a monitoring environment.

ShakeRec outputs can be further used as input to GMICES to compute seismic macrointensity maps.

The integration of the algorithm in a real-time monitoring environment is tested at the Department of Mathematics and Geosciences of the University of Trieste.

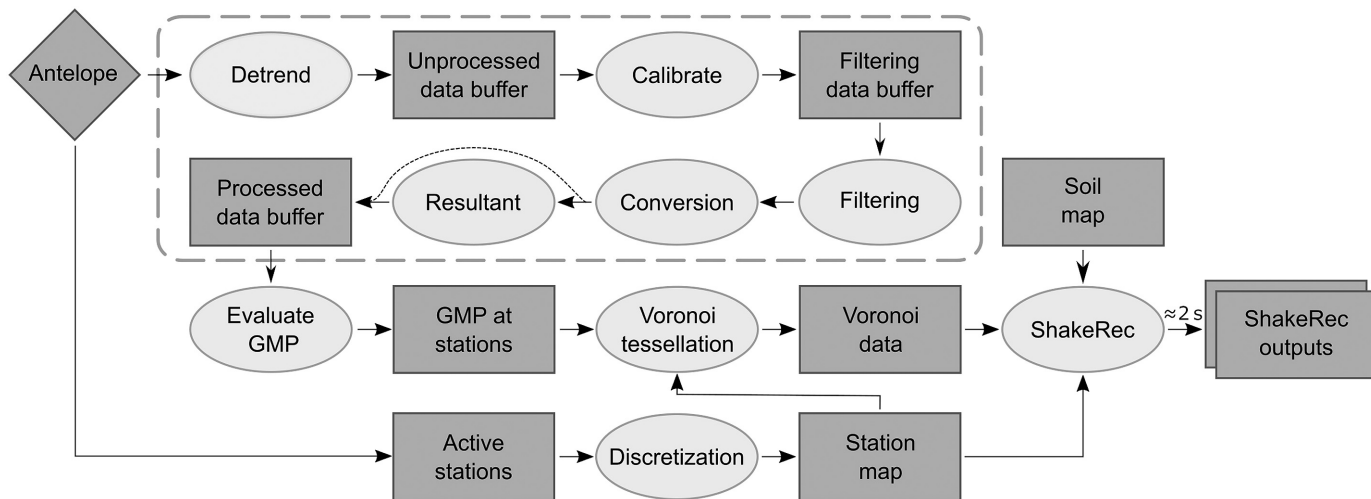
DATA AND RESOURCES

The development and testing of the proposed model were performed considering the data and metadata from multiple strong-motion networks in the Italian territory: the Italian strong-motion network (RAN, [Gorini et al., 2009](#)), Friuli Venezia Giulia accelerometric network (RAF,



Costa *et al.*, 2009), Irpinia seismic network (ISNet, Iannaccone *et al.*, 2007), northeast Italy broadband network (NI, OGS [Istituto Nazionale di Oceanografia e di Geofisica Sperimentale] and University of Trieste, 2002), Reti accelerometriche centri storici italiani (RR, Presidenza del Consiglio dei Ministri—Dipartimento della protezione civile, 2017), and Hareia (Hammerl *et al.*, 2012). The Istituto Nazionale di Geofisica e Vulcanologia (INGV) shakemaps archive that was used

Figure 7. Simulation of the real-time reconstruction of the PGA field for M_w 6.5 Norcia earthquake using ShakeRec. The PGA values at the stations are extracted from the recorded waveforms considering the data available at the time ΔT since the event. No delays due to data transmission or computation are considered. The epicenter, marked as a black star on the map, is shown as an indication of the earthquake location but would not be available in real time. The color version of this figure is available only in the electronic edition.



to produce the training and testing datasets is available at <http://shakemap.ingv.it> (last accessed February 2022). The proposed model was developed with Tensorow (Abadi et al., 2015) and deployed into production by integrating it with BRIT Antelope (<https://brtt.com>, last accessed June 2022). The figures were generated using Matplotlib (Hunter, 2007) and Cartopy (Met Office, 2010–2015). In the supplemental material, a comparison between the reconstruction of the ground-shaking field for different GMPs for the M_w 6.5 Norcia earthquake obtained from ShakeRec and ShakeMap for peak ground velocity (PGV) and spectral acceleration—SA03, SA10, and SA30 is reported. ShakeRec code is available upon request.

DECLARATION OF COMPETING INTERESTS

The authors acknowledge that there are no conflicts of interest recorded.

ACKNOWLEDGMENTS

This research received financial support from the Italian Civil Protection Department—Presidency of the Council of Ministers (DPC) under the agreement between DPC and UniTS for the accelerometric monitoring of the Friuli Venezia Giulia region and the advice on the analysis of the Italian strong-motion network (RAN) data (Agreement Number RAN2020-2022 CUP J91F20000110001). The authors thank the two anonymous reviewers for their constructive suggestions, which greatly helped to improve this article.

REFERENCES

Abadi, M., A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, et al. (2015). TensorFlow: Large-scale machine learning on heterogeneous systems, available at <https://www.tensorflow.org/> (last accessed in June 2022).

Abrahamson, N., N. Gregor, and K. Addo (2016). BC hydro ground motion prediction equations for subduction earthquakes, *Earthq. Spectra* **32**, no. 1, 23–44, doi: [10.1193/051712eqs188mr](https://doi.org/10.1193/051712eqs188mr).

Akkar, S., and J. J. Bommer (2007). Empirical prediction equations for peak ground velocity derived from strong-motion records from Europe and the Middle East, *Bull. Seismol. Soc. Am.* **97**, no. 2, 511–530, doi: [10.1785/0120060141](https://doi.org/10.1785/0120060141).

Allen, T. I., D. J. Wald, P. S. Earle, K. D. Marano, A. J. Hotovec, K. Lin, and M. G. Hearne (2009). An atlas of ShakeMaps and population

Figure 8. Workflow of ShakeRec implementation at Department of Mathematics and Geosciences (DMG). The preprocessing (in the dashed box), which contains standard procedures for real-time data analysis and could be shared with other algorithms, corresponds roughly to the left side of Figure 1.

exposure catalog for earthquake loss modeling, *Bull. Earthq. Eng.* **7**, no. 3, 701–718, doi: [10.1007/s10518-009-9120-y](https://doi.org/10.1007/s10518-009-9120-y).

Atkinson, G. M. (2003). Empirical ground-motion relations for subduction-zone earthquakes and their application to Cascadia and other regions, *Bull. Seismol. Soc. Am.* **93**, no. 4, 1703–1729, doi: [10.1785/0120020156](https://doi.org/10.1785/0120020156).

Aurenhammer, F. (1991). Voronoi diagrams—A survey of a fundamental geometric data structure, *ACM Comput. Surv.* **23**, no. 3, 345–405, doi: [10.1145/116873.116880](https://doi.org/10.1145/116873.116880).

Bindi, D., M. Massa, L. Luzi, G. Ameri, F. Pacor, R. Puglia, and P. Augliera (2014). Pan-European ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5 %-damped PSA at spectral periods up to 3.0 s using the RESORCE dataset, *Bull. Earthq. Eng.* **12**, no. 1, 391–430, doi: [10.1007/s10518-013-9525-5](https://doi.org/10.1007/s10518-013-9525-5).

Bindi, D., F. Pacor, L. Luzi, R. Puglia, M. Massa, G. Ameri, and R. Paolucci (2011). Ground motion prediction equations derived from the Italian strong motion database, *Bull. Earthq. Eng.* **9**, no. 6, 1899–1920, doi: [10.1007/s10518-011-9313-z](https://doi.org/10.1007/s10518-011-9313-z).

Bragato, P. L., and D. Slejko (2005). Empirical ground-motion attenuation relations for the eastern Alps in the magnitude range 2.5–6.3, *Bull. Seismol. Soc. Am.* **95**, no. 1, 252–276, doi: [10.1785/0120030231](https://doi.org/10.1785/0120030231).

Costa, G., L. Moratto, and P. Suhadolc (2009). Teneziueneziazia giulia accelerometric network: RAF, *Bull. Earthq. Eng.* **8**, no. 5, 1141–1157, doi: [10.1007/s10518-009-9157-y](https://doi.org/10.1007/s10518-009-9157-y) (in Italian).

Cultrera, G., L. Faenza, C. Meletti, V. D’Amico, A. Michelini, and A. Amato (2014). Shakemaps uncertainties and their effects in the post-seismic actions for the 2012 Emilia (Italy) earthquakes, *Bull. Earthq. Eng.* **12**, no. 5, 2147–2164, doi: [10.1007/s10518-013-9577-6](https://doi.org/10.1007/s10518-013-9577-6).

Dietterich, T. G. (2000). Ensemble methods in machine learning, in *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Vol. 1857 LNCS, 1–15, doi: [10.1007/3-540-45014-9_1](https://doi.org/10.1007/3-540-45014-9_1).

- DiSGAM (2005). Preliminary Geological Map of the Regional Soils Classification. convenzione per la riclassificazione sismica del territorio della regione friuli venezia giulia, allegato 1, *Protezione Civile della regione Friuli Venezia Giulia* (in Italian).
- Fukami, K., R. Maulik, N. Ramachandra, K. Fukagata, and K. Taira (2021). Global field reconstruction from sparse sensors with Voronoi tessellation-assisted deep learning, *Nat. Mach. Intell.* **3**, no. 11, 945–951, doi: [10.1038/s42256-021-00402-2](https://doi.org/10.1038/s42256-021-00402-2).
- Gorini, A., M. Nicoletti, P. Marsan, R. Bianconi, R. D. Nardis, L. Filippi, S. Marcucci, F. Palma, and E. Zambonelli (2009). The Italian strong motion network, *Bull. Earthq. Eng.* **8**, no. 5, 1075–1090, doi: [10.1007/s10518-009-9141-6](https://doi.org/10.1007/s10518-009-9141-6).
- Hammerl, C., W. A. Lenhardt, and M. Innerkofler (2012). Historische erdbeben in tirol. forschungen im rahmen des interreg iv projektes hareia (historical and recent earthquakes in Italy and Austria), in *Forum hall in tirol, neues zur geschichte der stadt*, A. Zanesco (Editor), Vol. 3, Ablinger and Garber, Innsbruck, Austria, 174–205 (in German).
- Hunter, J. D. (2007). Matplotlib: A 2D graphics environment, *Comput. Sci. Eng.* **9**, no. 3, 90–95, doi: [10.1109/MCSE.2007.55](https://doi.org/10.1109/MCSE.2007.55).
- Iannaccone, G., A. Zollo, A. Bobbio, L. Cantore, V. Convertito, L. Elia, G. Festa, M. Lancieri, C. Martino, A. Romeo, *et al.* (2007). The Irpinia seismic network: An advanced monitoring infrastructure for earthquake early warning in the Campania region (southern Italy), *AGU Fall Meeting Abstracts*.
- Kingma, D. P., and J. Ba (2017). Adam: A method for stochastic optimization, doi: [10.48550/arXiv.1412.6980](https://doi.org/10.48550/arXiv.1412.6980).
- Kong, Q., D. T. Trugman, Z. E. Ross, M. J. Bianco, B. J. Meade, and P. Gerstoft (2018). Machine learning in seismology: Turning data into insights, *Seismol. Res. Lett.* **90**, no. 1, 3–14, doi: [10.1785/0220180259](https://doi.org/10.1785/0220180259).
- LeCun, Y., Y. Bengio, and G. Hinton (2015). Deep learning, *Nature* **521**, no. 7553, 436–444, doi: [10.1038/nature14539](https://doi.org/10.1038/nature14539).
- LeCun, Y., B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel (1989). Backpropagation applied to handwritten zip code recognition, *Neural Comput.* **1**, no. 4, 541–551, doi: [10.1162/neco.1989.1.4.541](https://doi.org/10.1162/neco.1989.1.4.541).
- Margheriti, L., C. Nostro, O. Cocina, M. Castellano, M. Moretti, V. Lauciani, M. Quintiliani, A. Bono, F. M. Mele, S. Pintore, *et al.* (2021). Seismic surveillance and earthquake monitoring in Italy, *Seismol. Res. Lett.* **92**, no. 3, 1659–1671, doi: [10.1785/0220200380](https://doi.org/10.1785/0220200380).
- Massa, M., P. Morasca, L. Moratto, S. Marzorati, G. Costa, and D. Spallarossa (2008). Empirical ground-motion prediction equations for northern Italy using weak- and strong-motion amplitudes, frequency content, and duration parameters, *Bull. Seismol. Soc. Am.* **98**, no. 3, 1319–1342, doi: [10.1785/0120070164](https://doi.org/10.1785/0120070164).
- Met Office (2010–2015). Cartopy: A cartographic python library with a matplotlib interface [Computer software manual], Exeter, Devon, available at <http://scitools.org.uk/cartopy> (last accessed in June 2022).
- Michellini, A., L. Faenza, G. Lanzano, V. Lauciani, D. Jozinović, R. Puglia, and L. Luzi (2020). The new ShakeMap in Italy: Progress and advances in the last 10 yr, *Seismol. Res. Lett.* **91**, no. 1, 317–333, doi: [10.1785/0220190130](https://doi.org/10.1785/0220190130).
- Moratto, L., G. Costa, and P. Suhadolc (2009). Real-time generation of ShakeMaps in the southeastern Alps, *Bull. Seismol. Soc. Am.* **99**, no. 4, 2489–2501, doi: [10.1785/0120080283](https://doi.org/10.1785/0120080283).
- Moratto, L., P. Suhadolc, and G. Costa (2011). ShakeMaps for three relevant earthquakes in the southeastern Alps: Comparison between instrumental and observed intensities, *Tectonophysics* **509**, no 1/2, 93–106, doi: [10.1016/j.tecto.2011.06.004](https://doi.org/10.1016/j.tecto.2011.06.004).
- Nair, V., and G. E. Hinton (2010). Rectified linear units improve restricted Boltzmann machines, J. Fürnkranz and T. Joachims (Editors), *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, June 21–24, 2010, Haifa, Israel, Omnipress, 807–814.
- OGS (Istituto Nazionale di Oceanografia e di Geofisica Sperimentale) and University of Trieste (2002). *North-East Italy Broadband Network*, International Federation of Digital Seismograph Networks, available at <https://www.fdsn.org/networks/detail/NI/>, doi: [10.7914/SN/NI](https://doi.org/10.7914/SN/NI).
- Parmanto, B., P. W. Munro, and H. D. Doyle (1996). Reducing variance of committee prediction with resampling techniques, *Connect. Sci.* **8**, no 3/4, 405–426, doi: [10.1080/095400996116848](https://doi.org/10.1080/095400996116848).
- Presidenza del Consiglio dei Ministri—Dipartimento della protezione civile (2017). *Reti accelerometriche centri storici italiani*, International Federation of Digital Seismograph Networks, available at <https://www.fdsn.org/networks/detail/RR/>, doi: [10.7914/SN/RR](https://doi.org/10.7914/SN/RR) (in Italian).
- Sabetta, F., and A. Pugliese (1996). Estimation of response spectra and simulation of nonstationary earthquake ground motions, *Bull. Seismol. Soc. Am.* **86**, no. 2, 337–352, doi: [10.1785/BSSA0860020337](https://doi.org/10.1785/BSSA0860020337).
- Tusa, G., and H. Langer (2016). Prediction of ground motion parameters for the volcanic area of mount Etna, *J. Seismol.* **20**, no. 1, 1–42, doi: [10.1007/s10950-015-9508-x](https://doi.org/10.1007/s10950-015-9508-x).
- Voronoi, G. (1908). New applications of continuous parameters to the theory of quadratic forms, *Z. Reine Angew. Math.* **134**, 198.
- Wald, D. J., K.-W. Lin, and V. Quitoriano (2008). Quantifying and qualifying USGS Shakemap uncertainty, *U.S. Geol. Surv.*
- Wald, D. J., K.-W. Lin, K. Porter, and L. Turner (2008). ShakeCast: Automating and improving the use of ShakeMap for post-earthquake decision-making and response, *Earthq. Spectra* **24**, no. 2, 533–553, doi: [10.1193/1.2923924](https://doi.org/10.1193/1.2923924).
- Wald, D. J., V. Quitoriano, T. H. Heaton, and H. Kanamori (1999). Relationships between peak ground acceleration, peak ground velocity, and modified Mercalli intensity in California, *Earthq. Spectra* **15**, no. 3, 557–564, doi: [10.1193/1.1586058](https://doi.org/10.1193/1.1586058).
- Wald, D. J., V. Quitoriano, T. H. Heaton, H. Kanamori, C. W. Scrivner, and C. B. Worden (1999). TriNet “ShakeMaps”: Rapid generation of peak ground motion and intensity maps for earthquakes in southern California, *Earthq. Spectra* **15**, no. 3, 537–555, doi: [10.1193/1.1586057](https://doi.org/10.1193/1.1586057).
- Wald, D. J., H. A. Seligson, J. Rozelle, J. Burns, K. Marano, K. S. Jaiswal, M. Hearne, and D. Bausch (2020). A domestic earthquake impact alert protocol based on the combined USGS PAGER and FEMA Hazus loss estimation systems, *Earthq. Spectra* **36**, no. 1, 164–182, doi: [10.1177/8755293019878187](https://doi.org/10.1177/8755293019878187).
- Wald, D. J., C. B. Worden, E. M. Thompson, and M. Hearne (2021). ShakeMap operations, policies, and procedures, *Earthq. Spectra*, doi: [10.1177/87552930211030298](https://doi.org/10.1177/87552930211030298).
- WMO (2012). *Guidelines on Ensemble Prediction Systems and Forecasting*, Vol. WMO-No. 1091, World Meteorological Organization, Geneva, Switzerland.