Comparing and integrating artificial intelligence and similarity search detection techniques: application to seismic sequences in Southern Italy

Authors: Francesco Scotto di Uccio [1], Antonio Scala [1,2], Gaetano Festa [1,2], Matteo Picozzi [1], Gregory C. Beroza [3]

[1] Department of Physics "Ettore Pancini", University of Napoli Federico II, 80126 Napoli, Italy.

[2] Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Roma 1, Rome, Italy.

[3] Department of Geophysics, Stanford University, Stanford, CA 94305, USA.

Corresponding author: Francesco Scotto di Uccio; email : francesco.scottodiuccio@unina.it

Tel: +39 081679927; Fax : +39 0810093584

Antonio Scala; email : antonio.scala@unina.it

Gaetano Festa ; email : gaetano.festa@unina.it

Matteo Picozzi; email: matteo.picozzi@unina.it

Gregory C. Beroza; email : beroza@stanford.edu

Short title: Comparing advanced detection techniques

.

Summary

Understanding mechanical processes occurring on faults requires detailed information on the microseismicity that can be enhanced today by advanced techniques for earthquake detection. This problem is challenging when the seismicity rate is low and most of the earthquakes occur at depth. In this study, we compare three detection techniques, the autocorrelation FAST, the machine learning EQTransformer, and the template matching EQCorrScan, to assess their ability to improve catalogs associated with seismic sequences in the normal fault system of Southern Apennines (Italy) using data from the Irpinia Near Fault Observatory (INFO). We found that the integration of the machine learning and template matching detectors, the former providing templates for the cross-correlation, largely outperforms techniques based on autocorrelation and machine learning alone, featuring an enrichment of the automatic and manual catalogs of factors 21 and 7 respectively. Since output catalogs can be polluted by many false positives, we applied refined event selection based on the cumulative distribution of their similarity level. We can thus clean up the detection lists and analyze final subsets dominated by real events.

The magnitude of completeness decreases by more than one unit compared to the reference value for the network. We report b-values associated with sequences smaller than the average, likely corresponding to larger differential stresses than for the background seismicity of the area. For all the analyzed sequences, we found that main events are anticipated by foreshocks, indicating a possible preparation process for mainshocks at sub-kilometric scales.

Keywords: Machine learning, Time series analysis, Statistical seismology, Earthquake detection, Seismic sequences.

1 Introduction

Earthquake detection is a fundamental task in seismology since it represents the first step in the analysis and interpretation of Earth's crustal deformation processes. Detection consists of analyzing continuous records to identify discrete seismic events, but it becomes challenging as the size of the events decreases and the signal amplitude for earthquakes approaches the noise level. Furthermore, detection becomes even more challenging during seismic sequences, due to the occurrence of many events close in time. With the recent increase in the number and quality of seismic data, and the demand for fast, automatic processing of large data volumes, detection techniques need to be scalable and flexible, exploiting the multiscale coherency of earthquake signals.

Several detection techniques have been established through the years. They can be grouped into two main classes: energy-based and similarity-based detectors. The energy-based STA/LTA method (Allen, 1978) is one of the reference approaches in seismology and compares the waveform energy over a short time window (STA) with that of a long time window (LTA). This approach is highly flexible in detecting impulsive arrivals but may fail for events with amplitude near the noise level and closely spaced in time such that their arrivals overlap. Characteristic functions based on higher-order statistics of the waveforms, such as skewness and kurtosis, have been shown to improve on STA/LTA for low signal-to-noise ratio and intense seismic activity (Baillard *et al*. 2014; Poiata *et al.* 2016; Grigoli *et al.* 2018).

When focusing on seismic sources that repeat in time at nearby locations and with a similar mechanism, techniques based on waveform similarity measured by cross-correlation are very efficient for the detection of low amplitude events (Gibbons & Ringdal 2006; Shelly *et al.* 2007; Schaff 2008; Dodge & Walter 2015). These are usually referred to as template matching methods,

since a set of selected and known earthquake records are used as templates to scan the continuous data stream. Template matching techniques have been efficiently applied to dense networks surrounding tectonic (e.g., Nomura *et al.* 2014), volcanic (Lengliné *et al.* 2016) and induced seismicity (Huang & Beroza 2015) areas. Template matching requires an a priori set of reference earthquakes, and the final catalog is limited to events with waveforms similar to them. For this reason, template matching might be biased by the number and quality of the reference signals. Autocorrelation techniques mitigate the latter issue using each portion of continuous data as a potential template (Brown *et al.* 2008), under the assumption that the main similarities occur between seismic signals instead of noise. Despite their good performance, the computational demand of such detectors increases quadratically with the time extent of the dataset, such that it becomes infeasible for long-term monitoring. The introduction of an efficient set-based similarity, built on compact fingerprints (Yoon *et al.* 2015) allows a significant reduction of the runtime. Similarity search methods are suitable both for induced seismicity monitoring (Yoon *et al.* 2017; Scala *et al.* 2022) and for seismic sequences in tectonic areas (Yoon *et al.* 2019; Festa *et al.* 2021).

More recently, seismologists successfully adopted artificial intelligence (AI) and deep-learning (DL) models for event detection and phase picking (Zhu & Beroza 2018; Wang *et al.* 2019; Dokht *et al.* 2019; Mousavi *et al.* 2020; Mousavi & Beroza 2022). These techniques work by learning general earthquake waveform features from their high-level representations and aim to increase the dimension of manually generated catalogs with models easily exportable across different application areas (Mousavi *et al.* 2020; Münchmeyer *et al.* 2022). However, it is worth considering that the performance of machine learning detectors is always conditioned by magnitude and distance ranges of the events included in the training dataset (Mousavi *et al.* 2020).

Understanding which detection technique is more efficient in enhancing catalogs is a challenging

task, which requires comparison across methods and in-depth analysis of the differences between the resulting catalogs. In this study, we performed this comparison for seismic sequences occurring at the Southern Italy Irpinia fault system. Italy is one of the most seismically hazardous countries in the Mediterranean region, and although the worst historical earthquakes were all characterized by magnitude between 6 and 7, more than 120,000 people were killed by earthquakes during the last century (Valensise & Pantosti 2001). The area of this study was struck by the 1980, M 6.9 Irpinia earthquake occurred along NW-SE striking faults and characterized by three main episodes within a few tens of seconds, causing about 3000 fatalities and severe damage (Bernard & Zollo 1989).

To better characterize the seismic hazard of the Irpinia area, the Irpinia Near Fault Observatory (INFO) was established in 2007. The key idea of Near Fault Observatories in Europe is to install dense networks of multi-parametric sensors close to faults aiming at understanding the underlying Earth instability processes over broad time intervals (Chiaraluce *et al.* 2022). INFO includes the Irpinia Seismic Network (ISNet, http://isnet.unina.it) made up of 31 seismic stations, equipped with strong-motion accelerometers and weak-motion sensors to be sensitive to microseismic events. ISNET covers an area of 100×70 km^2, including the epicenter of the 1980 Irpinia earthquake (Iannaccone *et al.* 2010), with an interstation distance of 10-20 km. The manually revised catalog of seismic events for the past 15 years includes ∼3000 earthquakes mainly occurring at depths between 8 and 15 km. The events cover a local magnitude range between Ml -0.4 and Ml 3.7, with a completeness magnitude of Ml 1.1 (Vassallo *et al.* 2012). In Figure 1, we represent the location of stations of ISNet (yellow triangles), the distribution of the background seismicity (grey dots) and that of seismic sequences (red dots), as reported in the INFO catalog, and the epicenter of the Ms 6.9, 1980 Irpinia earthquake (green star).

In recent years, several studies concerning both source parameters (i.e., Zollo *et al.* 2014; Picozzi *et al.* 2019, 2022a) and medium properties (Vassallo *et al.* 2016) allowed significant improvement in the understanding of the Irpinia fault system, highlighting a relation between hydrological changes in the shallow karst system and microseismicity generation, with the former likely acting as an external forcing mechanism (D'Agostino *et al.* 2018; Picozzi *et al.* 2022b; De Landro *et al.* 2022). The background seismicity that occurred during recent years appears to be distributed within a volume bounded by the main faults of the 1980 event and sometimes clustered in sequences with events of maximum magnitude Ml ~3.0 (Stabile *et al.* 2012). The monitoring strategies currently adopted at INFO led to manually revised catalogs containing between 8 and 74 events for the seismic sequences. Nevertheless, Festa *et al.* (2021) showed that these seismic sequences might include hundreds of detectable events when analyzed with advanced tools.
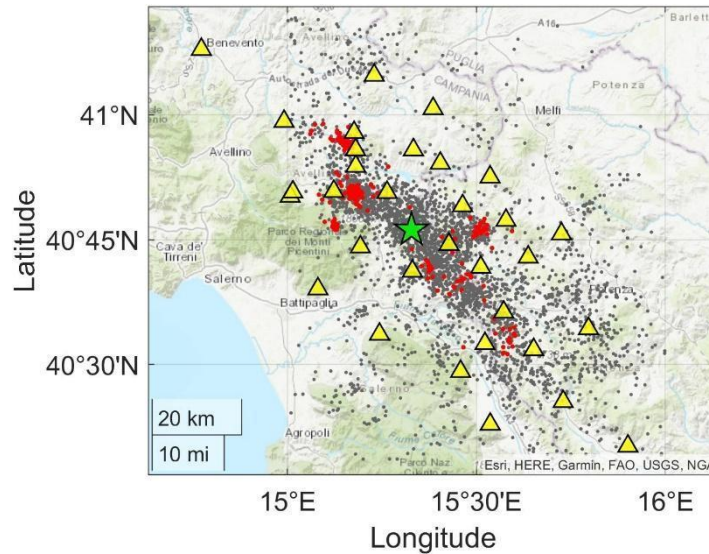

Figure (1). The location of the ISNet seismic stations (yellow triangles), the background seismicity (grey dots) and the events in the analyzed sequences (red dots), as detected by INFO from 2007. The green star marks the epicenter of the 1980, M 6.9 earthquake.

For this reason, we considered the continuous waveforms at ISNet stations for a set of ten seismic sequences (Figure 1) to test and compare the performance of three well-established automatic detection approaches: the autocorrelation method FAST (Yoon *et al.* 2015; Bergen & Beroza 2018), the machine learning technique EQTransformer (Mousavi *et al.* 2020) and the template matching technique EQCorrscan (Chamberlain *et al.* 2018). We assessed their performance in a particular context (i.e., a relatively small scale, dense network monitoring normal faults) and for a specific task (to provide augmented catalogs for seismic sequences).

The paper is organized as follows. In section 2 we present the selected seismic sequences. Then, we discuss the different techniques and the associated processing used to generate the enhanced catalogs (Section 3). In section 4, the results are reported in terms of both performance of the techniques and differences in the catalog content across the three methods. In section 5, we discuss the time-magnitude features of the sequences (magnitude of completeness, b-value of the Gutenberg-Richter law and time evolution of the seismicity). Finally, we present in the Discussion how the results affect the strategies for catalog generation and for a statistical description of the seismicity.

2 Data

We analyzed 10 seismic sequences that occurred in the Irpinia region between 2011 and 2021 near the three fault segments of the 1980 Irpinia earthquake. For each sequence, we selected the continuous velocity records at the 5 to 7 closest stations to the sequence centroid, spanning a hypocentral distance range between 10 km and 30 km. Our choice was guided by the need to have

a good azimuthal coverage with respect to the sequence. At one station (VDS3), we selected a ± 0.25 g full scale accelerometer because no high-gain seismometer was available.

The seismic sequences were selected using the information in the INFO catalog, which is available at the webpage http://isnet.unina.it/. The catalog has two layers: the first one consists of a catalog automatically generated by the Earthworm software (Johnson *et al.* 1995), which runs on continuous data-streams. A second revised catalog is released after waveform inspection by network operators with the twofold aim of improving the phase picking and including detections missed by the automatic procedure. At this stage, an event is included in the manual catalog if at least 4 phases have been picked on the records, including one S phase, and the P and S pick residuals are smaller than 0.5 s and 1.0 s, respectively. The final catalog consists of located events (uncertainties in absolute locations are on average between 1 and 2 km) with a local magnitude estimated using a scaling law calibrated for the area (Bobbio *et al.* 2009) and a moment magnitude inferred by spectral fitting (Zollo *et al.* 2014). According to the INFO catalog, the selected seismic sequences last between 3 and 6 days. Each sequence consists of a number of earthquakes between 8 and 74 (Table 1). Since our aim is to compare the performance of different detection strategies under close to homogeneous conditions, we processed each sequence considering the continuous data stream from one day before to one day after the duration indicated in the catalog.

Table (1) List of analyzed sequences. The table contains the ID number of the sequence, the name, the date, the geographical information (latitude, longitude and depth) of the main event, its local magnitude and the number of events included in the INFO manual catalog.

ID	PLACE	DATE	LAT (deg)	LON (deg)	DEPTH (km)	Ml	# INFO EVENTS
1	Rocca San Felice (AV)	2020-07-03	40,938	15,150	9,6	3,0	74
2	Lioni (AV)	2011-08-02	40,850	15,181	11,4	2,7	48
3	San Gregorio Magno (SA)	2012-02-17	40,709	15,367	5,6	2,8	9
4	Lioni (AV)	2012-03-03	40,832	15,164	11,3	3,7	25
5	Laceno (AV)	2013-07-22	40,772	15,130	13,3	1,8	30
6	Ricigliano (SA)	2015-12-12	40,679	15,484	19,5	3,0	12
7	Sant'Angelo le Fratte (PZ)	2016-05-15	40,535	15,171	16,0	2,7	19
8	Lioni (AV)	2017-07-16	40,843	15,175	11,2	2,8	17
9	Capo di Giano (PZ)	2019-04-16	40,756	15,491	7,2	2,9	8
10	Bella (PZ)	2019-09-08	40,775	15,499	6,3	3,1	23

3 Methods

We applied three different detection techniques to the selected continuous data: the deep-learning based detector EQTransformer (Mousavi *et al.* 2020), the template matching technique EQCorrscan (Chamberlain *et al.* 2018) and the autocorrelation technique FAST (Yoon *et al.* 2015). We then compared the detections provided by each method against the automatic and manually revised INFO catalogs to evaluate their performance. Although all these techniques showed improved performance in comparison with standard detection approaches, it is well known that, depending on the tuning of working parameters, they can suffer from false detections, whose number becomes larger and larger as the acceptance threshold for event declaration is lowered.

Therefore, after applying the techniques, further analyses and detailed inspection are usually required to isolate real events. In this study, we explore the possibility of using adaptive thresholds, which refine the initial selections and automatically extract a final catalog, significantly reducing the rate of false positives.

3.1. EQTransformer

EQTransformer is an AI-based earthquake signal detector and phase (P and S) picker built on a deep neural network with an attention mechanism. It has a hierarchical architecture in which detection is performed on continuous time series and seismic phases identified along with the extracted declarations.

The detector is trained on the Stanford EArthquake Dataset (STEAD) (Mousavi *et al.* 2019), composed of worldwide labeled earthquakes and noise signals. During the training phase ~ 1 M earthquake and ~ 300 k noise waveforms recorded at stations within epicentral distance up to 300 km were used, representing ~ 450 k events. The majority of events within STEAD have magnitude lower than 2.5 and are recorded at epicentral distances within 100 km. The traces used in the training extended for 1 minute with a sampling rate of 100 Hz and were band-pass filtered in the range 1-45 Hz. In our work, we follow a data preprocessing similar to the one of STEAD. During the training phase, the data was augmented by adding secondary earthquake signals and a random level of Gaussian noise into the earthquake waveforms and shifting the event within the trace to vary its position. The first operation allowed the model to be sensitive, at station level, to multiple events occurring very close in time, as happens during a seismic sequence. We selected the trained model EqT_model_h5, optimized to minimize false positives.

EQTransformer output depends on five input parameters: the probability thresholds for detection (det_thresh), P (P_thresh) and S (S_thresh) picks, the overlap between consecutive time windows and the batch size. After a parametric study, we set the input values to det_thresh=0.3, P_thresh=S_thresh=0.1, overlap=30%, batch size=100 (see Supplementary material). For the association, we declared an event if it is detected at least at 2 stations within a time window of 10 s.

Since the association phase performs a simple count of detections declared inside a moving time window without checking any time and spatial coherence among the stations, EQTransformer may declare false positives if there are multiple triggers from the same station inside the considered time window. The choice of the trained model mitigates this issue; the declaration of non-seismic signals is further reduced by excluding from the association detections with an estimated duration below 1 s.

Because the EQTransformer detections will be used as templates for the analysis with EQCorrscan, we evaluated the quality of the EQTransformer picks for a set of events by comparing the picks to the manual ones. For this purpose, we selected 1097 phases (672 P and 425 S picks) from about 200 events in the local magnitude range -0.1 – 3.7, that were also included in the INFO catalog. We found that EQTransformer automatic picks are consistent with manual picks. The distributions of differences between picks are in fact centered at zero and feature a standard deviation of 0.08 s for P waves (Figure S1, central panel), and of 0.2 s for S waves (Figure S1, right panel), respectively. Integrated results including both P and S picks are shown in Figure S1, left panel.

3.2. EQCorrscan

We used EQCorrscan software (Chamberlain *et al.* 2018) for template matching analysis. We generated two different catalogs: the first one was built considering as templates the events declared by EQTransformer; the second one considering the events contained in the INFO manual catalog. For both cases, we extracted templates at stations with at least one pick (P or S). The threshold for event declaration is computed by considering for each master event the MAD of cross-correlation coefficients between the single-station template and one hour window of continuous data. The threshold is then defined as the sum of the MAD over all the picked stations for the master events. We selected portions of templates lasting 1.5 s and starting 0.15 s before the picks; traces were band-pass filtered in the range 2-9 Hz and decimated to 25 Hz.

For template matching, the number of detections critically depends on the selected threshold. For a lower threshold, we can retrieve more earthquakes, but with the risk of a dramatic increase in the number of false detections. On the other hand, higher thresholds would reduce the number of false positives, but at the expense of increasing missed real events. Thus, the threshold should be set to balance the number of false and real detections. We initially set the detection threshold to 8 times the sum of template cross-correlation MAD and investigated the resulting catalogs for some sequences. In Figure 2, we report, as an example, the results for the Laceno sequence (ID 5), which was characterized by a main event of magnitude Ml 1.8. In this case, the template matching analysis provides 233 detections. We inspected all the detections to isolate real events from false positives, checking the shape, the frequency content and the duration of the signals in different frequency bands, with a maximum explored range of 2-20 Hz In the left panel of Figure 2, we report the distribution of the ratio between the sum of the cross-correlation associated with the single declaration and the detection threshold (CCsum/thresh), distinguishing false events (red points) and real earthquakes (green points). The distributions of real and false detections appear

separated with an overlap in the range (1.1-1.4).

To investigate the distribution of the previous parameter further, in the right panel of Figure 2, we represent the cumulative number of real and false events (green and red curves, respectively) and the cumulative number of all detections in the catalog (blue curve) as a function of the CCsum/thresh parameter. For low values of this parameter, the cumulative number of detections is dominated by false events.
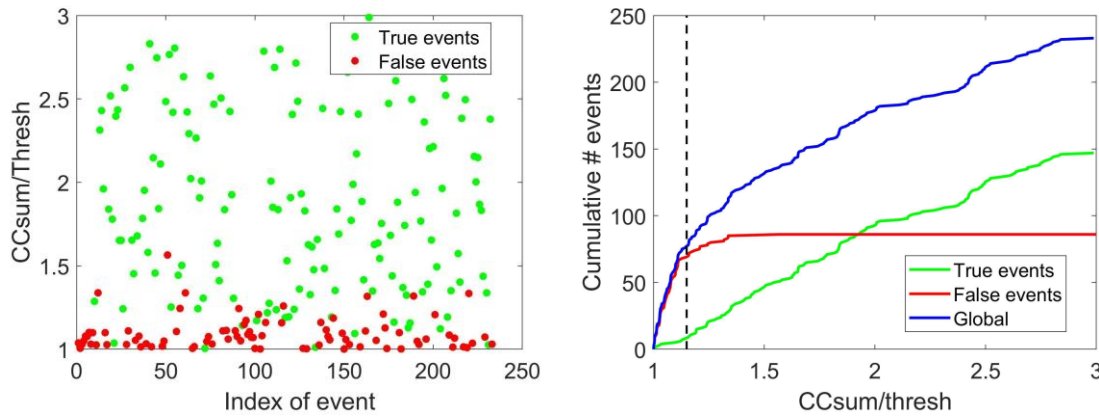

Figure (2) Left panel: Distribution of the ratio CCsum/threshold for the events in the initial catalog provided by EQCorrscan (green points correspond to real events, red points to false positives) for the ID 5 sequence. Right panel: Cumulative number of events as a function of the ratio CCsum/threshold for the subset of the real events (green line), false positives (red line) and for the whole initial catalog (blue line)

As the parameter increases, we retrieve fewer and fewer false events having such a high score, resulting in a flattening of the red curve. Thus, for large values of CCsum/thresh the increase of the blue curve is driven by the distribution of the real events. Since the real and false event distributions feature different behaviors, the cumulative distribution of all the detections exhibits

a change in the slope, which can be identified by fitting its initial and final trends and used as a refined criterion that allows us to significantly reduce the number of false detections in the catalog. In the right panel of Figure 2, the slope break corresponds to the value CCsum/thresh = 1.15, which results in a new threshold equal to 9.2 MAD, higher than the initial value.

We stress that this criterion is directly applied to the cumulative number of detections in the catalog output by the template matching technique and can be automated, without preliminarily identifying the two families. We observed the same behavior for all the sequences and thus we applied the same strategy to refine the threshold in all cases.

We remark that separating the two distributions is effective when false and real events are comparable in number. On the contrary, if the population of one distribution is significantly larger than the other, the global cumulative distribution almost reproduces the shape of the larger size distribution, and the criterion cannot be applied. Thus, the selection of the initial detection threshold is also important for the refinement. In Figure S2, we superimpose the cumulative distributions computed for thresholds of 6 and 8 MAD. When lowering the threshold value to 6 MAD, the detection list is composed of ~ 13k declarations and the shape of the cumulative distribution only represents the behavior of false events (red curve in Figure S2).

3.3. FAST

The FAST technique (i.e., Fingerprint And Similarity Thresholding, Yoon *et al.* 2015) is an uninformed similarity search technique that converts time-domain waveforms into binary fingerprints containing discriminative features of earthquakes. It performs an optimized search to identify couples of similar fingerprints associated with seismic events. FAST compresses the

single-component time-frequency spectrogram into fingerprints using the Haar wavelet transform, maintaining only the fraction of the Haar coefficients that most differ from the average noise-descriptive behavior.

The similarity among fingerprints is evaluated using a Min-Hash algorithm (Broder *et al.* 2000) and the Jaccard theorem (Leskovec *et al.* 2014). This latter ensures that the probability of having the same outcome from a Min-Hash function applied to two different fingerprints is equal to the Jaccard similarity, which provides a set-based estimate of the similarity between binary objects. Therefore, the use of a set of Min-Hash functions allows for an a priori selection of similar fingerprints without comparing all the couples, significantly reducing the computational time.

Information about similarity is integrated across components for fingerprints within the same time window; the station declaration list is thus built considering fingerprints overcoming a fixed station similarity threshold. After identification of couples of fingerprints at station level, a network association criterion requires time delays compatible with travel times at a minimum number of stations to provide the final list of detections. Following Bergen and Beroza (2018), the detections are ranked depending on the number of stations declaring the event and their similarity score, referred to as the peaksum, which measures the maximum similarity between a single detection and all the others.

We bandpass filtered traces in the range 1-10 Hz and downsampled to 25 Hz. A similar frequency band was considered in other applications (4-10 Hz, Yoon *et al.* 2015; 1-6 Hz, Yoon *et al.* 2019).

For fingerprint generation, we selected 6.0 s long time windows with a lag of 0.2 s between consecutive windows. We kept 200 out of the 1024 coefficients of the Haar transform based on the daily MAD of the coefficients. The Min-Hash algorithm application is grounded on two independent parameters: the number of tables b and the number of hash functions per table r (Yoon

et al. 2015). We set $b=100$ and $r=4$ (Festa *et al.* 2021; Scala *et al.* 2022). In the association phase, we required similarity occurring at least at 2 stations with a maximum lag between detection times of 3 s, the latter selected considering the available inter-station distances.

As with all the autocorrelation techniques, FAST suffers from a large number of false detections (Yoon *et al.* 2017). A class of false positives is represented by coherent noise occurring over time due to local ambient sources. To mitigate this effect, we discarded those fingerprints that are similar for more than 3% of the day length ($\sim$ 15 minutes). Starting from the threshold refinement established for the template matching technique, we investigated the possibility to set up a similar criterion, based on the cumulative distribution of the peaksum parameter for all the detections, at a fixed number of stations. In Figure S3, we present the cumulative distribution of the peaksum for events declared at 2 stations, where a change in slope occurs around the 90th percentile. Below this corner, the reported peaksum values are very close to each other and, when inspecting the waveforms, the vast majority corresponds to false positives. On the other hand, above this corner, the values follow a different distribution, and they appear to be indicative of real events. We found a similar behavior for other sequences and when changing the number of stations. Thus, the slope break criterion is adaptively applied for each sequence to refine the threshold for event selection. We applied this strategy only when the number of detections is large enough to allow a statistical analysis in terms of peaksum distribution. To validate this criterion, we also performed a visual inspection of the declarations discriminating real and false events according to the shape, the frequency content of the signal and the propagation throughout the considered stations.

Finally, events below the threshold were considered in a later stage if their fingerprints are found similar to those of events above the threshold. When building the final detection lists, we removed regional and teleseismic events.

3.4. Catalog merging, magnitude and b-value estimation

After performing the detection analysis of the seismic sequences and removing false positives, we estimated the local magnitude Ml of seismic events. We first selected one event in each sequence: the event is required to be located in the INFO catalog, with waveforms clearly emerging from the noise at all the stations and local magnitude between 1.0 and 2.0. For this event we computed the local magnitude from half of the maximum peak-to-peak Wood-Anderson displacement averaged on the horizontal components and on the stations, using the local relationship of Bobbio *et al.* (2009) and the INFO catalog location. For all other events (considered as 'offsprings'), we provided a magnitude estimation through the displacement amplitude ratio, assuming colocation:

$Ml = Ml_{ref} + \log_{10} \frac{A}{A_{ref}}$. Ml_{ref} is the magnitude of the reference event; A and A_{ref} are half of the maximum of peak-to-peak amplitudes for the considered and the reference events, respectively. The error in the magnitude estimation due to colocation is estimated to be 0.1 units of magnitude (Festa *et al.* 2021).

Catalogs are compared according to the detection times related to single events and are combined to obtain the merged catalog containing all the independent detections relevant to the catalogs derived by the different techniques.

To investigate the frequency–magnitude distribution from the resulting catalogs, we evaluated the parameters of the Gutenberg-Richter relation ($\log_{10} N = a - bM$, where N is the number of events with local magnitude $> M$), the magnitude of completeness for each sequence by using the software ZMAP (Wiemer, 2001) and considering the local magnitude to characterize the event size.

For the Rocca San Felice seismic sequence (ID 1), which is characterized by the highest number

of detections, we also investigated the temporal evolution of the b-value. For this purpose, we considered a sliding window of 80 events with 50% of overlap. We estimated the average b-value and the uncertainty by means of a bootstrap approach (Efron, 1979), using for each dataset, 200 realizations of random samples with replacement.

4 Results

To assess the overall performance of the adopted detection techniques, we analyzed the results obtained from applying four detectors to data from ten seismic sequences in the Irpinia region: i) the autocorrelation FAST, ii) the machine learning EQTransformer (hereinafter, EQT), iii) the EQCorrscan using EQT's templates (hereinafter, EQT+TM), iv) EQCorrscan using as templates the manual detections from INFO (hereinafter, INFO+TM). In Figure 3, we show the detection catalogs for the four techniques, integrated over the ten sequences and organized in a Venn diagram. When analyzing the performance of the different approaches, it is worth highlighting that for the considered sequences, the manually revised INFO catalog contains only 265 events, 82 of which were automatically declared by the STA/LTA-based detector operating on the network.

The merged catalog contains 1792 events, increasing by a factor 6.7 the revised catalog and by a factor 21 the automatic one. These results clearly highlight the superior performance of the advanced techniques in discovering smaller magnitude events relative to standard approaches. The contribution of the different techniques to the merged catalog in terms of number of detections changes among sequences. We observe that sequence ID 1 has a high seismicity rate (~500 events in 1 day), while for the other sequences the rate is lower, presenting between 50 and 200 events in 4 days.

Looking at the overall performance of the single techniques, we report that FAST declares 942 events (~ 3.5x the manual catalog, ~ 11x the automatic catalog), EQT detects 450 events, increasing by factors 1.5 and 5.0 the revised and automatic catalogs. EQT+TM declares 1715 events, with a catalog content similar to the merged one, while INFO+TM detects 1165 events (~ 68 % of the EQT+TM catalog).

As can be seen from Figure 3, most of the events (95%) declared by FAST are also retrieved by EQT+TM. The detections common to FAST and INFO+TM decrease to 734, representing 78% of the FAST catalog.

FAST is able to declare 73 events that are missed by the other techniques, which represents 5% of the merged catalog. After checking these events, we observe that they can be grouped in two classes: either they feature a low, close to one, signal-to-noise ratio, or their waveforms are different from those of the events occurring near the sequence centroid. For the former class of events, we find that they exhibit a smaller cross-correlation value with the used templates, and they can still be retrieved by the EQT+TM technique by lowering the acceptance threshold, but at the cost of a significant increase of the number of false events. As an example, for the San Gregorio Magno sequence (ID 3), the EQT+TM technique is able to catch all the events in this class when decreasing the declaration threshold to 6 MAD. These events appear within a set of more than 3.5k detections (8 MAD catalog is composed of 82 declarations), mostly corresponding to false positives.

The second class of events corresponds to events with observed S-P travel-times compatible with shallower epicenters, as compared to the other events in the sequence. These events often occur as isolated couples for which their mutual cross-correlation is very high, while they feature a much smaller cross-correlation value with the templates.
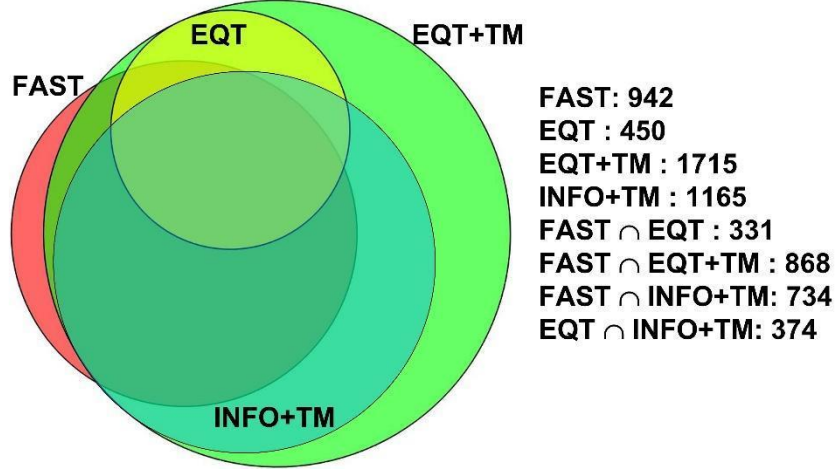
Figure (3) Venn diagram showing the performance of the different detectors: the autocorrelation (FAST, red), machine learning (EQT, yellow), template matching (EQT+TM, green; INFO+TM, cyan) techniques. The EQT detections are included in the EQT+TM ones because the EQT output is used to form templates.

EQT+TM contains almost all events retrieved by FAST and doubles the size of the FAST catalog. The significant outperformance of template matching is mainly driven by sequences where the seismicity rate is high and when the events are clustered in time, especially when the earthquakes occur in the coda of previous events. For the other sequences, the performance of EQT+TM and FAST are similar. However, we also observe that when we analyze the catalog for a single sequence, EQT+TM can always catch more than 85% of the merged catalog.

We performed some tests by changing the parametrization of FAST, either shortening the time window for fingerprint generation or decreasing the number of hash functions to be more permissive in the similarity search. The former action helps in retrieving consecutive events, the latter in reducing the similarity threshold for declaration (Yoon *et al.* 2015). In these cases, we retrieve the vast majority of events missed by the previous parametrization. However, if we take the San Gregorio Magno sequence (ID 3) as an example, decreasing the number of hash functions

from 4 to 3 results in a catalog of 3.5k detections, with a huge increase in the run time (from ~0.5 hours to almost 1 day of analysis) and in the memory requirements for the storage of similar fingerprint information (from ~ 100MB to 5 GB).

The better performance of EQT+TM is also related to the large set of available templates that are provided by EQT (i.e., for some sequences the number of templates is doubled with respect to the revised INFO catalog). Indeed, when we limit the set of templates to those provided by the revised manual INFO catalog (INFO+TM), the number of the detections significantly decreases, becoming intermediate between FAST and EQT+TM for half of the sequences and comparable to or smaller than the FAST catalog for the other half.

Finally, the number of EQT detections is much smaller than the other catalogs (it represents about 25% of the merged catalog). However, there is a fraction of events detected by EQT that are not seen by FAST, still related to consecutive seismic events. Indeed, the main advantage in using EQT in this framework is to provide a richer set of templates to template matching improving its performance.

In Figure 4, we report the detection performance for each technique and for the 10 analyzed seismic sequences, also comparing them with the INFO manual catalogs. The Rocca San Felice sequence (ID 1, left panel of Figure 4) is separated from the others to improve the visualization of the results for the other sequences (ID 2-10, right panel of Figure 4). We note that the cases in which EQT+TM outperforms FAST correspond to situations where EQT provides a wider set of templates (orange bars) with respect to the dimension of the INFO catalog (violet bars). The performance of INFO+TM is always lower than EQT+TM and becomes either comparable to or smaller than FAST for half of the sequences.
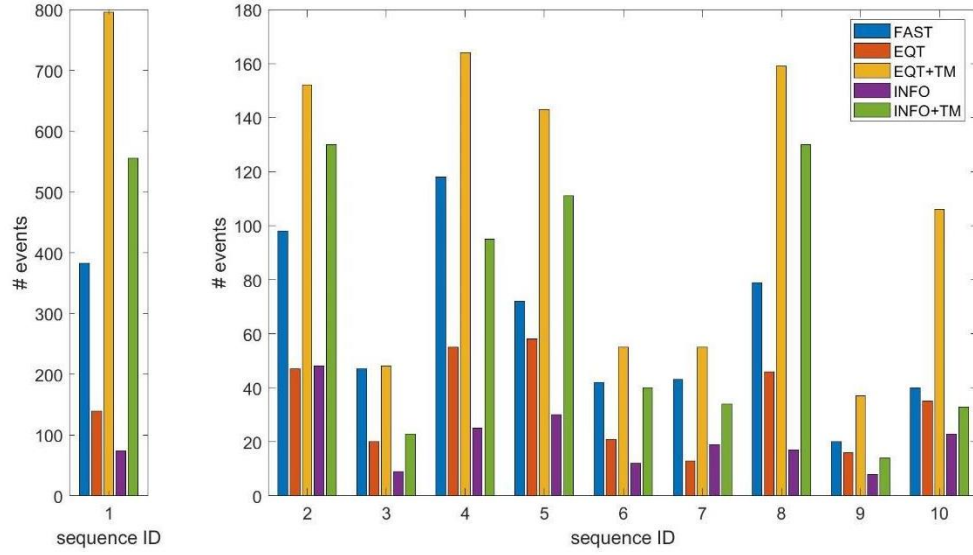
Figure (4) Detection performance of each technique for the analyzed sequences. The Rocca San Felice sequence (ID 1) is shown in the left panel, all the other sequences in the right panel. The histograms report the detections of FAST (blue), EQT (orange), EQT+TM (yellow), INFO (violet) and INFO+TM (green).

Note that the analysis was carried out on velocity records, with the exception of the station VDS3, which is equipped with a lower full-scale accelerometer. We found that about 30% of events in the merged catalog have been declared at this station, pointing out that the use of accelerometers with lower gain can contribute to microseismicity detection and characterization.

4.1. Rocca San Felice sequence

To analyze the differences in the performance of the techniques in detail, we inspected the detections of the largest sequence, the Rocca San Felice sequence (ID 1), which occurred from July 3 to 7, 2020. The Rocca San Felice area is a natural laboratory for studying the interaction between fluids and seismicity, since it hosts one of the largest sources of natural non-volcanic CO_2

gas emissions ever measured (Mefite d'Ansanto, Chiodini *et al.* 2010). The CO_2 degassing in central and southern Italy is well documented (Chiodini *et al.* 2004), and its accumulation within the Apennines' crust is considered a possible triggering mechanism for large earthquakes in Italy (Savage 2010). This sequence was previously studied by Festa *et al.* (2021), in terms of event detection, location and source parameters. Here, we reanalyzed the same waveform dataset to test and compare the detectors.

When looking at the outcomes of the detection techniques, FAST output consists of 1406 detections, which is reduced to 500 after the application of the automatic selection criterion. These events were visually inspected, limiting the final catalog to 383 earthquakes, similar to the results of Festa *et al.* (2021). EQT declares 155 detections, 139 of which are local earthquakes. Using these as templates for EQT+TM, we obtain a list of 1394 detections, reduced to 969 after threshold refinement, and finally to 796 events after visual inspection. Thus, the application of the adaptive thresholds for catalog refinement, allows significant reduction in the number of false events (from 73% to 25% for FAST and from 55% to 18% for EQT+TM). Finally, the INFO+TM catalog is built on 74 manually detected templates, resulting in 555 detections, all in common with the EQT+TM catalog.

The size of the final detection catalogs from the techniques are presented in Figure 4.

As previously observed, most of the detections retrieved by EQT+TM and missed by FAST correspond to earthquakes occurring in the coda of the previous events. An example of this situation is shown in Figure S4, where 4 different events, correctly and separately detected by EQT+TM appear as a single, longer detection in FAST. When shortening the time window for fingerprint generation from 6 s to 3 s, FAST is able to separate most of grouped detections, at the cost of increasing the computational time and the number of false declarations. When applied to

longer data-streams, this parametrization seems to quickly bring to prohibitive computational time that can hardly be imagined to be adopted in standard processing.

The few cases (i.e., only 7 events) where EQT+TM cannot detect events seen by FAST correspond to signals buried in the noise. In Figure S5, we show a set of 3 events declared by EQT+TM (Ml 0.08, 0.15, 0.18, red box) and the last event (blue box), featuring Ml -0.24, detected FAST but missed by EQT+TM.

The final combined catalog (i.e., considering the independent detections of FAST and EQT+TM) is composed of 803 events, enhancing the manual and automatic INFO catalogs of factors 11 and 20, respectively. It is worth stressing that also for this sequence, EQT+TM outperforms FAST due to the wide magnitude range and the large number of the templates provided by EQT. When we limit the set of templates to the 74 manually detected events in INFO, we observe that INFO+TM yields a catalog of about 555 events ($\sim$ 70 % of the previous TM catalog).

5. Statistical analysis of the catalogs

In this section, we highlight the importance of the enhanced catalogs to improve the characterization of the seismic sequences in terms of well-established statistical parameters. We used the seismic catalogs for estimating the magnitude of completeness, Mc, and the b-value of the Gutenberg-Richter frequency magnitude distribution. The estimation of the statistical parameters using the INFO catalog is not possible for some of the considered sequences due to the small number of detected events (i.e., in some cases consisting of 10 earthquakes only).

If we consider the combined catalog for each sequence (obtained combining EQT+TM and FAST detections), we obtain a magnitude of completeness ranging between Mc -0.3 and Mc 0.4 (in local

magnitude units), with an average improvement with respect to the INFO manual catalog of 1.1 units.

Considering the overall performance of the single techniques, we find that Mc for EQT+TM coincides with those obtained for the combined catalogs. Also FAST and INFO+TM provide Mc estimates similar to those of the combined catalogs, with the exception of sequences ID 1 and ID 5, where it is larger. On the contrary, EQT features a Mc larger by 0.3 units on average, with a large variability (from 0.1 to 0.7), but still smaller than the one from INFO.

Focusing on the b-value, we find that EQT+TM, INFO+TM and FAST provide compatible values within uncertainties. Interestingly, we observe that in general EQT provides b-values systematically lower than the other two techniques. We return to this issue in the Discussion.

In Table S1, we report the estimated values of the magnitude of completeness Mc and the b-values in local magnitude (here referred to as b_{M_l}) for the 10 analyzed sequences and the different techniques.

As an example, we show the frequency-magnitude distributions for the different detection techniques for the Rocca San Felice sequence (ID 1) in Figure 5.

We find that the slope and the magnitude of completeness are Mc 1.2 and $b_{M_l} = 0.83\pm0.12$ for manual INFO catalog (not shown in Figure), Mc 0.0 and $b_{M_l} = 0.71\pm0.05$ for FAST, Mc 0.2 and $b_{M_l} = 0.54\pm0.04$ for EQT, Mc -0.3 and $b_{M_l} = 0.71\pm0.03$ for EQT+TM, Mc -0.1 and $b_{M_l} = 0.72\pm0.04$ for INFO+TM .

The considered seismic sequences occurred in different sectors of the Irpinia area. Picozzi *et al.* (2022b) showed that the b-value distribution in this area is not uniform and hypothesized that b-values differences are related to different stress levels (i.e., the lower the b-value, the higher the

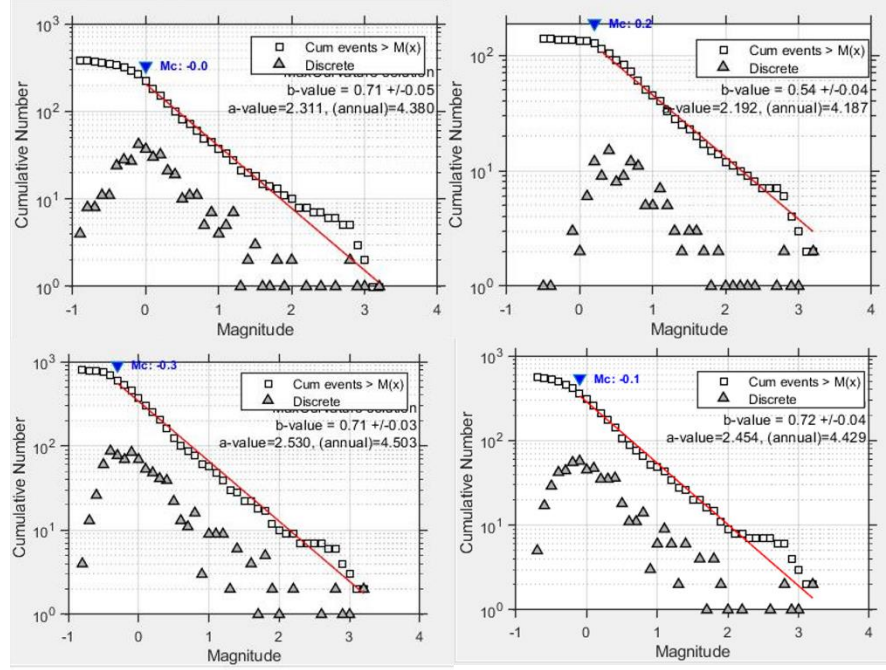stress, Scholz 2015) associated with the different seismogenic zones in this region.

Figure (5) Gutenberg-Richter distribution for FAST (Mc 0.0 and b=0.71 ± 0.05; upper-left panel), EQT (Mc 0.2 and b = 0.54 ±0.04; upper-right panel), EQT+TM (Mc -0.3 and b = 0.71 ±0.03; lower-left panel) and INFO+TM (Mc -0.1 and b = 0.72 ±0.04; lower-right panel) catalogs.

We therefore compare the b-values obtained for the seismic sequences with those obtained by Picozzi *et al.* (2022b). To this aim, we need to consider the moment magnitude of the detected events. We converted the local magnitude estimates into seismic moment ones using the empirical relationship $\log_{10} M_0 = 1.31 Ml + 10.55$ (standard deviation 0.12, Figure S6) which has been derived considering the INFO catalog earthquakes (i.e., Ml and Mw values from Picozzi *et al.* 2022a). The M_0 can in turn be used to retrieve the moment magnitude Mw (Hanks & Kanamori 1979). This magnitude scaling relation allows b-value estimates in terms of moment magnitude

(b_{M_w}) from the b-value based on the local magnitude (b_{M_l}), yielding $b_{M_w} = 1.5 \, b_{M_l} / 1.31$.

Figure 6 shows that the b-values for the whole area are on average rather small (i.e., mostly below 1). The smallest b-values from Picozzi *et al.* (2022b) are observed for the Southern sector of the Irpinia region (~0.7). The central and northern sectors are associated with slightly larger b-values (i.e., between 0.8 and 0.9).

We computed the difference between the b-values from Picozzi *et al.* (2022b) and those obtained for the sequences using the combined catalog (Δb-value). It is worth clarifying that the former values are obtained by considering a 3D grid of the area and associating to each node the closest events that occurred between 2007 and 2020. Therefore, the b-values from Picozzi *et al.* (2022b) provide a spatial average of the stress level in a rather large crustal volume and over a long time period with respect to the b-values obtained for the sequences, which have limited spatial and temporal extent. Despite this methodological issue, we observe that the Δb-values are small, within a ± 0.25 range, and larger than zero for most cases, indicating an overall decrease of the b-value during the sequences. We do not believe that b-value estimates for the sequences are biased by the duration of the selected time windows (< 1 week) or by the presence of events in the catalog not belonging to the sequences. Indeed, when inspecting the merged catalog, we note that the seismic activity strongly decreases near the end of the selected time window, with a negligible percentage of events in the last hours of analysis. For example, we report in Figure S7 the number of detected events as a function of time, showing that, for the merged and INFO catalogs and in the case of two sequences (ID 1 and ID 9), the number of events in the last six hours is smaller than 3% of the whole catalog. Moreover, when analyzing the distribution of the differences between the S and P arrival times at the stations with the largest number of detections, we found that the events are contained in a volume with characteristic size in the range 2-8 km.
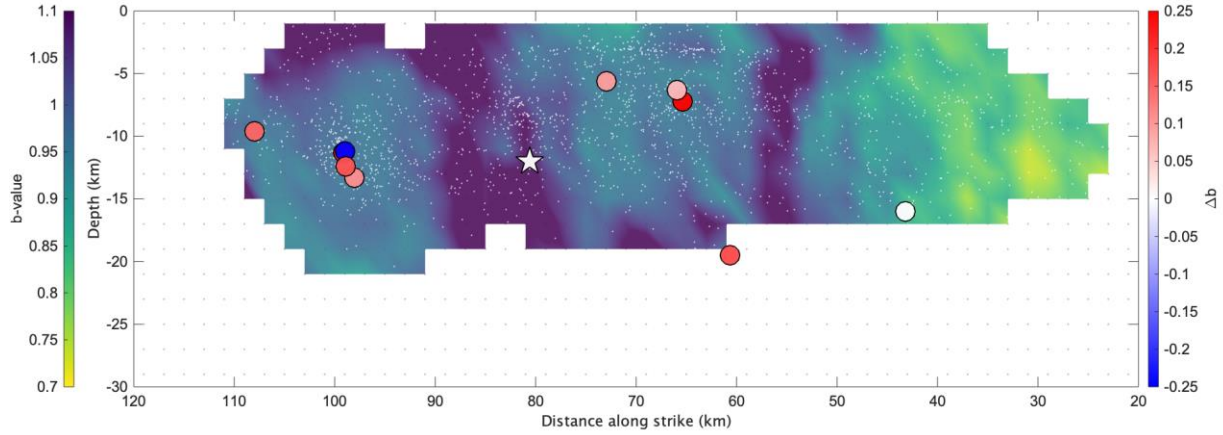
Figure (6) Map showing the spatial distribution of average b-values from bootstrap analysis (redrawn from Picozzi et al. 2022b). In the figure we show the nucleation point for the 1980 Irpinia earthquake (white star) and Δb-value (see text for the definition, colored circles) for the sequences analyzed here.

Finally, we analyzed the time evolution for the event occurrence and the seismic moment during the sequences. Considering the occurrence time of the main event in each sequence as a time reference, we observe that the enhanced catalogs allow discovery of foreshocks. In Figure 7, we report the cumulative number of events and the cumulative released seismic moment, as a function of occurrence time from the mainshock.

The foreshocks represent a fraction of the single-sequence catalog ranging between the 5% and 20% of the detected events, with a cumulative seismic moment between $2 \cdot 10^{11}$ and $6 \cdot 10^{13}$ Nm, representing up to 10% of the seismic moment of the main event, except for two sequences where this value increases to 60% (ID 5) and 90% (ID 2).

In particular, the larger value in seismic moment is associated with a main event of Ml 2.7 (ID 2), anticipated by two earthquakes of Ml 2.0 and a Ml 2.4 within 10 minutes before the mainshock.

The green curve of Figure 7 represents a particular sequence (ID 5), with a main event of magnitude Ml 1.8 that was anticipated by several foreshocks of maximum magnitude Ml 1.2. For this sequence, we find that the catalog is composed of more than its 80% by foreshocks. A similar percentage is observed for a Ml 2.8 main event sequence anticipated by several Ml >1.5 events that features around 50 % of the events being foreshocks.

Looking at the time evolution of sequences, we can identify two main behaviors. Most of the sequences generate the majority of earthquakes (around 70 % of the catalog) within 6 hours after the mainshock and result being characterized by a similar time evolution of the aftershocks (Figure 7, left panel). Three sequences feature a swarm-like behavior with magnitudes of the aftershocks approaching the size of the largest event.
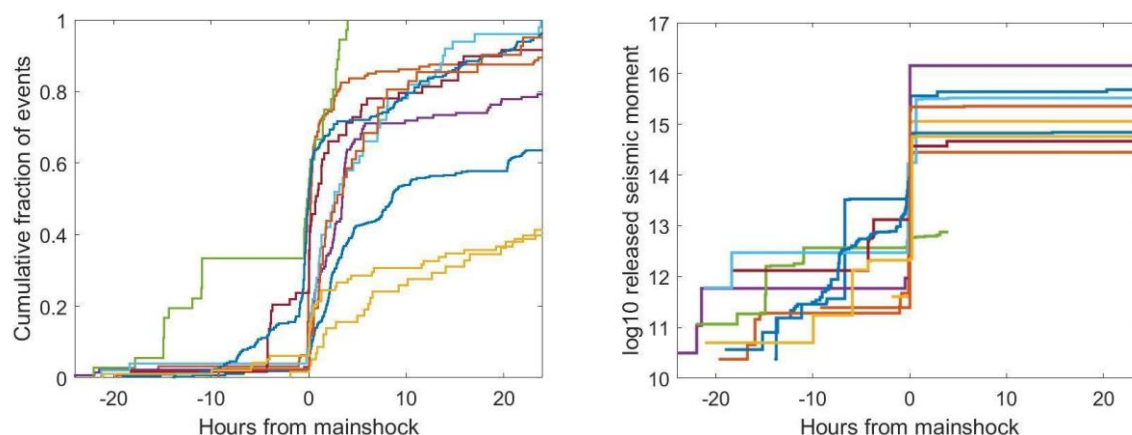

Figure (7) Foreshock/aftershock analysis: left panel shows the cumulative percentage of the events in the catalog as a function of the time from the main event. The right panel contains the cumulative seismic moment in a shorter time window before the main event.

Figure 8 shows the b-value temporal evolution for the Rocca San Felice sequence. We see that the b_{M_w} has an average value close to 1. However, during the sequence evolution, the b-value shows

significant changes, with variations larger than the associated uncertainties. The interpretation of these changes is beyond the scope of this work. We report these results to highlight the importance of having enhanced catalogs for monitoring the evolution of seismic sequences. Future studies will investigate the possible mechanism related to the b-value evolution during seismic sequences in Irpinia.
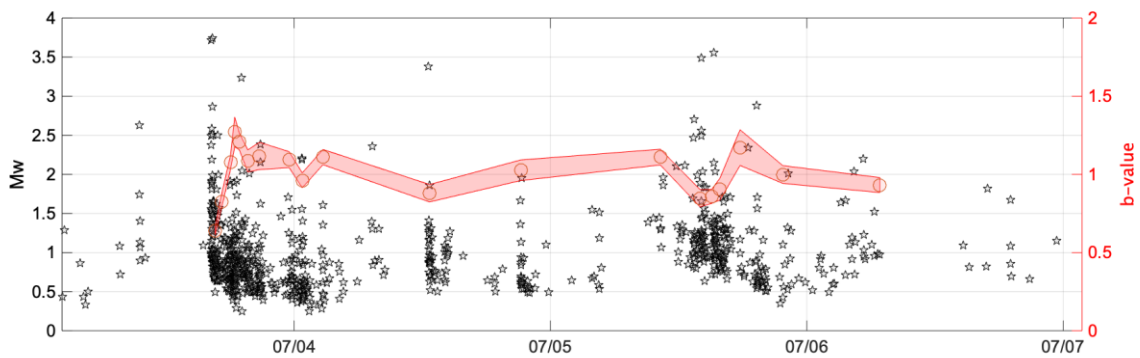

Figure (8): Temporal evolution of b-value (red circles) computed for the Rocca San Felice sequence (ID 1) with 95% probability uncertainty (red band) and event distribution as a function of time and moment magnitude (black stars).

6. Discussion

The use of new, advanced techniques for enhancing seismic catalogs is crucial both for monitoring how seismicity evolves with time within a complex geological system, and for understanding how faults slip to generate moderate to large earthquakes, possibly anticipated by the occurrence of small seismic events. This latter scientific issue becomes challenging in areas where seismicity occurs deep in the crust, as happens for seismic events in Southern Apennines, with earthquakes located at depths between 8 and 15 km. Moreover, in this area, the background seismic rate is

pretty low (the INFO revised catalog contains about 3000 events in 14 years), despite the high seismic hazard of the region (Meletti *et al.* 2021). When applying automatic detection techniques during seismic sequences, we report that the combination of the machine learning detector EQTrasformer (EQT) and the template matching EQCorrscan (EQT+TM) outperforms the autocorrelation technique FAST.

Integrating the two techniques (EQT+TM) is in our opinion beneficial because EQT allows us to build a richer, automated catalog of templates for template matching. Using as templates only the events manually detected by the INFO system by network operators, the final catalog INFO+TM significantly decreases (70% of EQT+TM; for half of the sequences at the same level or below the performances of FAST). FAST shows similar performance with respect to the combination of the other two methods when the seismicity rate in the sequence is low and events are well separated in time. However, the computational cost and the memory requirements of FAST make this detector challenging to apply to seismic sequences. Furthermore, the computational resources significantly increase as we extend the time window of the analysis beyond the few days analyzed here.

Both template matching and FAST detectors suffer from a high rate of false events with standard thresholds. We suggest the use of a refinement procedure based on the cumulative distribution of parameters that measure the waveform similarity (the sum of cross-correlation for template matching and the *peaksum* for FAST). Several additional strategies have been proposed to rule out false positives from the detection lists (e.g., Yoon *et al.* 2017; Scala *et al.* 2022). However, these are based on specific features of false events (frequency content, energy distribution in time and/or frequency) and require further, often visual, inspection of the events. The strategy proposed here, instead, can be automatically applied on the shape of the cumulative distribution, without knowing

specific characteristics related to false positives.

The performance of EQT alone has been shown to be lower than the other two techniques. This can be ascribed to the selected trained model which minimizes false events. Consequently, most of the EQT declarations are real events within the sequences and the majority of the discarded detections still correspond to local or regional earthquakes occurring at different locations. When adopting the trained model EqT_model2.h5, which minimizes the rate of false negatives, in the case of the Rocca San Felice sequence (ID 1) we retrieved a list of ~3k detections, ~550 of which are effective events, almost all contained in the previous EQT+TM catalog. Since the binding criterion of EQT is weak in the selection of events, we prefer to maintain a more conservative approach for EQT and to enhance the catalog integrating it with the template matching technique. The use of a more robust associator (Ross *et al.* 2019; Zhang *et al.* 2019), and/or a model trained on local data (e.g., INSTANCE, Michelini *et al.* 2021) could improve the detection performance of EQT, at least to the level of FAST.

While standard or even revised catalogs are in general too sparse to provide reliable statistics for seismic sequences in the Southern Apennines, the rapid, automatic generation of enhanced catalogs allows monitoring of seismicity evolution in near real time, in terms of frequency-magnitude event occurrence (Gutenberg-Richter law) and event occurrence with time. For several sequences, no reliable estimate of the b-value can be provided using only events included in the INFO catalog. Here, we report that all the detectors but EQT only for a few cases provided enhanced catalogs over which it is possible to estimate the parameters of the Gutenberg-Richter law. Analyzing the b-values for the sequences, we report that EQT+TM, INFO+TM and FAST provide consistent estimates, while the b-value from EQT is systematically lower. This is ascribable to the larger magnitude of completeness of this technique and to the lack of some events

(~10%) above its completeness level. In other words, it seems that EQT progressively loses detections approaching Mc, which leads to a biased population of magnitude bins and lower b-value estimates. This suggests the need to cover almost two-three units in magnitude to estimate the relative rate of occurrence reliably for seismic events in sequences for the area.

We found that for most of the sequences the b-value is smaller (within a range of 0.25 units) than the average value computed from background seismicity. This systematically happens both in the northern and central sectors of the fault systems. Differences in b-value between these two Irpinia sectors have been ascribed to the presence of fluid saturated volumes, eventually filled with different fluid content, CO_2 and water respectively (Picozzi *et al.* 2022a).

Worldwide observations indicate that the b-value is directly connected to differential stress (Schorlemmer *et al.* 2005; Scholz 2015) and it increases as the differential stress decreases. The systematic lower b-values for the sequences might indicate that they occurred in regions where the stress is higher than in the surrounding areas, and they likely rupture compact, sub-kilometric size asperities (Festa *et al.* 2021). When analyzing in detail two sequences in the area, Stabile *et al.* (2012) and Festa *et al.* (2021) retrieved large stress drops and focal mechanisms compatible with main orientation of the large faults that generated the 1980 Irpinia earthquake. These sequences either occurred on subparallel, smaller scale faults, or they ruptured some patches on the main faults that were unruptured during previous events, or they map small scale, geometrical discontinuities, which impede rupture growth into a large earthquake. In all cases, the sequences did not evolve into a large event, indicating that both static and dynamic stresses were not able to sustain a rupture over larger spatial scales.

We only report one sequence featuring a b-value larger than the average. This sequence (ID 8) occurred in the northern sector and is characterized by a main event of Ml 2.8 and the second

largest one in magnitude of Ml 1.4. We retrieved more than 80 foreshocks (~50% of the catalog) within 1 hour before the main event and an acceleration of the seismicity preceding the mainshock. This specific behavior could be ascribed to different stress mechanisms generating and controlling the sequence, as compared to the nearby ones.

In most cases, the retrieved b-value is representative of the behavior of the seismicity after the main event in the sequence. However, when the rate of the seismicity is high, we can also monitor the evolution of the b-value as a function of time, as in the case of the Rocca San Felice sequence (Figure 8). Here we found a rapid increase of the b-value around the main event of the sequence, followed by an almost constant level. b-value monitoring with time as a discriminator of foreshock and aftershock activity is a debated topic in literature (Gulia & Wiemer 2019; Marzocchi *et al.* 2020; Lombardi 2022). b-value has been claimed to increase for aftershocks and decrease for foreshocks, both for real earthquakes (Gulia *et al.* 2016; Gulia & Wiemer 2019) and laboratory experiments (Amitrano 2003). Hence, these enhanced catalogs allow us to study the variation of b-value in time and to investigate the possibility of using it as a forecasting tool as suggested by Gulia and Wiemer (2019).

For each of the analyzed sequences, we found small magnitude earthquakes anticipating the main event. For two cases, we just report very few foreshocks; in other cases, their number (>10) is an indication of a preparation process of the main event in the area. Most of the foreshocks that appear in these catalogs have a local magnitude M ~ 0 and below. If we extend the stress drop self-similarity to such events, they ruptured a decametric area, enlightening a new space scale for seismic ruptures, never investigated before in this area. The extension and validity of statistical analysis and parameters at these scales might provide additional clues on earthquake self-similarity and upscaling at larger size events, also for monitoring purposes. Nevertheless, deeper analysis of

the sources of such events is limited by current instrumental deployment, sampling at 125 Hz, and by the low signal to noise ratio at high frequency, preventing a robust determination of the event size. Installation of instruments at depth or organized in arrays may help in improving our understanding of such small-scale ruptures and in filling the gap between laboratory-controlled experiments and moderate-to-large size events.

Further insight on the sequence generation and evolution can come from accurate location of events, to discern structures and faults, and to discriminate whether the sequence is driven by fluid diffusion or associated with sequential stress release. Automatic picks provided by EQT exhibit large uncertainties, especially for S phases (Figure S1), to be usable for locations with decametric precision. Large uncertainties arise from the complexity of the structure across which waves propagate, often showing converted phases preceding the S-wave or emergent P signals (De Landro *et al.* 2015). Future direction for picking improvement could be grounded on transfer learning to refine picking criteria based on local data and analyst measurements (e.g., Chai *et al.* 2020), on arrival time consistency across multiple stations or including more observables, such as wave polarization (Zollo *et al.* 2021). Template matching provides accurate relative arrival times that can be further improved by narrowing the time window around the main phases (Schaff & Waldhauser 2005). FAST, instead, does not provide any arrival time information, since the time windows for similarity search are too large for accurate waveform alignment. Automatic arrival time individuation with FAST would require the integration of external advanced picking tools, such as kurtosis or machine learning techniques.

7   Conclusions

In this study, we compared the performance of the autocorrelation method FAST, the machine

learning technique EQTransformer (EQT) and the template matching technique EQCorrscan using

EQT (EQT+TM) and INFO (INFO+TM) catalogs as template sets. As a case study, we selected

seismic sequences occurring in the Irpinia region, Southern Italy. In this region, the seismicity

occurs at depths between 8 and 15 km, and the complexity in the geological structure results in

waveforms enriched by several secondary phases, often featuring low SNR values and emergent

P and S arrivals.

We report that all the techniques provide enhanced catalogs for the area if compared to the

reference INFO catalog. We found that the integration of EQT and TM (EQT+TM), that is the use

of template matching from a set of templates provided by EQT, significantly outperforms FAST,

almost doubling the number of detections. The better performance is mainly driven by the

sequences where the seismicity rate is high and when the events occur clustered in time, often in

the coda of previous earthquakes. The few events (5% of the merged catalog, i.e., the catalog

obtained combining all the FAST and EQT+TM detections) detected by FAST and missed by

EQT+TM occur at low signal-to-noise ratio or they appear as isolated pairs exhibiting shallower

location. EQT+TM also outperforms INFO+TM, which is grounded on a manual catalog, by a

factor 1.5 due to the large set of available templates provided by EQT. As expected, the

performance of EQT alone, with the adopted trained model, is lower than the other two techniques.

However, it contains 85% of detections in the INFO manual catalog and improves it by a factor

1.5. Finally, the merged catalog improves the INFO catalogs by factors ~7 (manual catalog) and

~21 (automatic catalog).

Both FAST and template matching techniques feature a large number of false events. Automatic

criteria, based on the cumulative distribution of the score of the similarity (the *peaksum* value for

FAST and the sum of cross-correlation for template matching), significantly reduce the number of

false events and can be systematically applied without identifying specific discriminative characteristics of false positives as compared to real events.

When inspecting the catalogs for evaluating the statistical parameters, we report that for the single-sequence catalog obtained by EQT+TM, the magnitude of completeness ranges between Ml -0.3 to 0.4, with an average improvement of 1.1 units in magnitude with respect to the INFO manual catalog. The magnitude of completeness for FAST and INFO+TM is compatible with that of EQT+TM in most cases, while EQT provides larger Mc estimates. The *b*-values from EQT+TM, INFO+TM and FAST are compatible with each other within uncertainties while the *b*-values obtained with EQT catalogs are systematically smaller, both due to the larger magnitude of completeness and the lack of events (~10%) above the completeness. This result cautions about the possibility of introducing some bias in the Gutenberg-Richter determination, when covering a small range in magnitude above the completeness threshold.

We also find that for all the seismic sequences, main events are anticipated by foreshocks (in most cases more than 10 events), indicating a possible preparation process for mainshocks at sub-kilometric scales. The presence of foreshocks, illuminated by advanced catalogs is crucial for a better understanding of crustal processes, as shown in recent studies of large Apennine earthquakes, where it helped in interpreting and understanding the nucleation processes leading to large magnitude events (Chiarabba *et al.* 2020; Sugan *et al.* 2022; Picozzi *et al.* 2022c).

Data availability statement

Data and products related to the Irpinia Near-Fault Observatory (INFO) are available at the Irpinia seismic network infrastructure portal (ISNet: http://isnet.unina.it); continuous waveforms are available at the EIDA platform (http://www.orfeus-eu.org/data/eida/, virtual network

_NFOIRPINA, network code IX) and at the EPOS portal (https://www.epos-eu.org/dataportal).

Software FAST is published on GitHub https://github.com/stanford-futuredata/FAST (Yoon *et al.*

2015). Software EQCorrScan is available at https://eqcorrscan.readthedocs.io/en/latest/

(Chamberlain *et al.* 2018). Software EQTransformer is published on GitHub

https://github.com/smousavi05/EQTransformer (Mousavi *et al.* 2020).

References

Allen, R. V., 1978. Automatic earthquake recognition and timing from single traces. *Bull. Seism. Soc. Am.*, 68(5), 1521-1532. https://doi.org/10.1785/BSSA0680051521

Amitrano, D., 2003. Brittle-ductile transition and associated seismicity: Experimental and numerical studies and relationship with the b-value. *J. Geophys. Res.*, 108. https://doi.org/10.1029/2001JB000680.

Baillard, C., Crawford, W. C., Ballu, V., Hibert, C. & Mangeney, A., 2014. An automatic kurtosis-based P-and S-phase picker designed for local seismic networks. *Bull. Seism. Soc. Am*, 104(1), 394-409. https://doi.org/10.1785/0120120347

Bergen, K. J. & Beroza, G. C., 2018. Detecting earthquakes over a seismic network using single-station similarity measures. *Geophys. J. Int.*, 213(3), 1984-1998. https://doi.org/10.1093/gji/ggy100

Bernard, P. & Zollo, A., 1989. The Irpinia (Italy) 1980 earthquake: detailed analysis of a complex normal faulting. *J. Geophys. Res.*: *Solid Earth*, 94(B2), 1631-1647. https://doi.org/10.1029/JB094iB02p01631

Broder, A. Z., Charikar, M., Frieze, A. M. & Mitzenmacher, M., 2000. Min-wise independent permutations. *J. Comput. Syst. Sci.,* 60(3), 630-659. https://doi.org/10.1006/jcss.1999.1690

Bobbio, A., Vassallo, M. & Festa, G., 2009. A local magnitude scale for southern Italy. *Bull. Seism. Soc. Am.*, 99(4), 2461-2470. https://doi.org/10.1785/0120080364

Brown, J. R., Beroza, G. C. & Shelly, D. R., 2008. An autocorrelation method to detect low frequency earthquakes within tremor. *Geophys. Res. Lett.*, 35(16). https://doi.org/10.1029/2008GL034560

Chai, C., Maceira, M., Santos-Villalobos, H. J., Venkatakrishnan, S. V., Schoenball, M., Zhu, W., Beroza, G. C., Turber, C. & EGS Collab Team, 2020. Using a deep neural network and transfer learning to bridge scales for seismic phase picking. *Geophys. Res. Lett.*, 47(16).e2020GL088651, https://doi.org/10.1029/2020GL088651

Chamberlain, C. J., Hopp, C. J., Boese, C. M., Warren-Smith, E., Chambers, D., Chu, S. X., Michailos, K. & Townend, J., 2018. EQcorrscan: Repeating and near-repeating earthquake detection and analysis in Python. *Seism. Res. Lett.*, 89(1), 173-181. https://doi.org/10.1785/0220170151

Chiarabba, C., Buttinelli, M., Cattaneo, M. & De Gori, P., 2020. Large earthquakes driven by fluid overpressure: The Apennines normal faulting system case. *Tectonics*, 39(4), e2019TC006014. https://doi.org/10.1029/2019TC006014

Chiaraluce, L., Festa, G., Bernard, P., Caracausi, A., Carluccio, I., Clinton, J., Di Stefano, R.,

Elia, L., Evangelidis, C., Ergintav, S., Jianu, O., Kaviris, G., Marmureanu, A., Sebela, S. & Sokos, E., 2022. The Near Fault Observatory community in Europe: a new resource for faulting and hazard studies. *Ann. Geophys.*, 65(3), DM316. https://doi.org/10.4401/ag-8778

Chiodini, G., Cardellini, C., Amato, A., Boschi, E., Caliro, S., Frondini, F. & Ventura, G., 2004. Carbon dioxide Earth degassing and seismogenesis in central and southern Italy. *Geophys. Res. Lett.*, 31(7). https://doi.org/10.1029/2004GL019480

Chiodini, G., Granieri, D., Avino, R., Caliro, S., Costa, A., Minopoli, C. & Vilardo, G., 2010. Non-volcanic CO2 Earth degassing: Case of Mefite d'Ansanto (southern Apennines), Italy. *Geophys. Res. Lett.*, 37(11). https://doi.org/10.1029/2010GL042858

D'Agostino, N., Silverii, F., Amoroso, O., Convertito, V., Fiorillo, F., Ventafridda, G. & Zollo, A., 2018. Crustal deformation and seismicity modulated by groundwater recharge of karst aquifers. *Geophys. Res. Lett.*, 45(22), 12-253. https://doi.org/10.1029/2018GL079794

De Landro, G., Amoroso, O., Stabile, T. A., Matrullo, E., Lomax, A. & Zollo, A., 2015. High-precision differential earthquake location in 3-D models: evidence for a rheological barrier controlling the microseismicity at the Irpinia fault zone in southern Apennines. *Geophysical Supplements to Mon. Notices Royal Astron. Soc*., 203(3), 1821-1831. https://doi.org/10.1093/gji/ggv397

De Landro, G., Amoroso, O., Russo, G., D'Agostino, N., Esposito, R., Emolo, A. & Zollo, A., 2022. Decade-long monitoring of seismic velocity changes at the Irpinia fault system (southern Italy) reveals pore pressure pulsations. *Sci. Rep.,* 12(1), 1-9. https://doi.org/10.1038/s41598-022-05365-x

Dodge, D. A. & Walter, W. R., 2015. Initial global seismic cross-correlation results: Implications

for empirical signal detectors. *Bull. Seism. Soc. Am.*, 105(1), 240-256.

https://doi.org/10.1785/0120140166

Dokht, R. M., Kao, H., Visser, R. & Smith, B., 2019. Seismic event and phase detection using time–frequency representation and convolutional neural networks. *Seism. Res. Lett.*, 90(2A), 481-490. https://doi.org/10.1785/0220180308

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Ann. Stat.*, 7, 1–26. https://doi.org/10.1214/aos/1176344552

Festa, G., Adinolfi, G. M., Caruso, A., Colombelli, S., De Landro, G., Elia, L., Emolo, A., Picozzi, M., Scala, A., Carotenuto, F., Gammaldi, S., Iaccarino, A. G., Nazeri, S., Riccio, R., Russo, G., Tarantino, S. & Zollo, A., 2021. Insights into mechanical properties of the 1980 Irpinia fault system from the analysis of a seismic sequence. *Geosci.*, 11(1), 28. https://doi.org/10.3390/geosciences11010028

Gibbons, S. J. & Ringdal, F., 2006. The detection of low magnitude seismic events using array-based waveform correlation. *Geophys. J. Int.*, 165(1), 149-166. https://doi.org/10.1111/j.1365-246X.2006.02865.x

Grigoli, F., Cesca, S., Krieger, L., Kriegerowski, M., Gammaldi, S., Horalek, J., Priolo, E. & Dahm, T., 2016. Automated microseismic event location using master-event waveform stacking. *Sci. Rep.,* 6(1), 1-13. https://doi.org/10.1038/srep25744.

Gulia, L., Tormann, T., Wiemer, S., Herrmann, M. & Seif, S., 2016. Short-term probabilistic earthquake risk assessment considering time-dependent b values. *Geophys. Res. Lett.*, 43(3), 1100-1108. https://doi.org/10.1002/2015GL066686

Gulia, L. & Wiemer, S., 2019. Real-time discrimination of earthquake foreshocks and

aftershocks. *Nature*, 574(7777), 193-199. https://doi.org/10.1038/s41586-019-1606-4

Hanks, T. C. & Kanamori, H., 1979. A moment magnitude scale. *J. Geophys. Res.*: *Solid Earth*, 84(B5), 2348-2350. https://doi.org/10.1029/JB084iB05p02348

Huang, Y. & Beroza, G. C., 2015. Temporal variation in the magnitude-frequency distribution during the Guy-Greenbrier earthquake sequence. *Geophys. Res. Lett.*, 42(16), 6639-6646. https://doi.org/10.1002/2015GL065170

Iannaccone, G., Zollo, A., Elia, L., Convertito, V., Satriano, C., Martino, C., Festa, G., Lancieri, M., Bobbio, A., Stabile, T. A., Vassallo, M. & Emolo, A., 2010. A prototype system for earthquake early-warning and alert management in southern Italy. *Bull. Earthq. Eng.,* 8(5), 1105-1129. https://doi.org/10.1007/s10518-009-9131-8

Johnson, C. E., Bittenbinder, A., Bogaert, B., Dietz, L. & Kohler, W., 1995. Earthworm: A flexible approach to seismic network processing. *Iris newsletter*, 14(2), 1-4.

Lengliné, O., Duputel, Z. & Ferrazzini, V., 2016. Uncovering the hidden signature of a magmatic recharge at Piton de la Fournaise volcano using small earthquakes. *Geophys. Res. Lett.*, 43(9), 4255-4262. https://doi.org/10.1002/2016GL068383

Leskovec, J., Rajaraman, A. & Ullman, J. D., 2020. Mining of massive data sets. Cambridge university press.

Lombardi, A. M., 2022. Anomalies and transient variations of b-value in Italy during the major earthquake sequences: what truth is there to this?. *Geophys. J. Int*., ggac403. https://doi.org/10.1093/gji/ggac403

Marzocchi, W., Spassiani, I., Stallone, A. & Taroni, M., 2020. How to be fooled searching for significant variations of the b-value. *Geophys. J. Int*, 220(3), 1845-1856.

https://doi.org/10.1093/gji/ggz541

Meletti, C., Marzocchi, W., D'amico, V., Lanzano, G., Luzi, L., Martinelli, F., Pace, B., Rovida, A., Taroni, M., Visini, F. & Group MW., 2021. The new Italian seismic hazard model (MPS19). *Ann. Geophys.*, 64(1). https://doi.org/10.4401/ag-8579

Michelini, A., Cianetti, S., Gaviano, S., Giunchi, C., Jozinović, D. & Lauciani, V., 2021. Instance–the Italian seismic dataset for machine learning. *Earth Syst. Sci. Data*, 13(12), 5509-5544. https://doi.org/10.5194/essd-13-5509-2021

Mousavi, S. M., Sheng, Y., Zhu, W. & Beroza, G. C., 2019. STanford EArthquake Dataset (STEAD): A global data set of seismic signals for AI. IEEE Access, 7, 179464-179476. https://doi.org/10.1109/ACCESS.2019.2947848

Mousavi, S. M., Ellsworth, W. L., Zhu, W., Chuang, L. Y. & Beroza, G. C., 2020. Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. *Nature Com.*, 11(1), 1-12. https://doi.org/10.1038/s41467-020-17591-w

Mousavi, S. M. & Beroza, G. C., 2022. Deep-learning seismology. *Science*, 377(6607), eabm4470. https://doi.org/10.1126/science.abm4470

Münchmeyer, J., Woollam, J., Rietbrock, A., Tilmann, F., Lange, D., Bornstein, T., Diehl, T., Giunchi, C., Haslinger, F,.Jozinović, D. & Michelini, A., 2022. Which picker fits my data? A quantitative evaluation of deep learning based seismic pickers. *J. Geophys. Res.: Solid Earth*,;127(1):e2021JB023499. https://doi.org/10.1029/2021JB023499

Nomura, S., Ogata, Y. & Nadeau, R. M., 2014. Space-time model for repeating earthquakes and analysis of recurrence intervals on the San Andreas Fault near Parkfield, California. *J. Geophys. Res.*: *Solid Earth*, 119(9), 7092-7122. https://doi.org/10.1002/2013JB010558

Omori, F., 1894. On aftershocks. *Seism. J Japan*, (19), 71-80.

Picozzi, M., Bindi, D., Spallarossa, D., Oth, A., Di Giacomo, D. & Zollo, A., 2019. Moment and energy magnitudes: diversity of views on earthquake shaking potential and earthquake statistics. *Geophys. J. Int.*, 216(2), 1245-1259. https://doi.org/10.1093/gji/ggy488

Picozzi, M., Bindi, D., Festa, G., Cotton, F., Scala, A. & D'Agostino, N., 2022. Spatiotemporal Evolution of Microseismicity Seismic Source Properties at the Irpinia Near-Fault Observatory, Southern Italy. *Bull. Seism. Soc. Am.*, 112(1), 226-242. https://doi.org/10.1785/0120210064

Picozzi, M., Cotton, F., Bindi, D., Emolo, A., Adinolfi, G. M., Spallarossa, D. & Zollo, A., 2022. Spatiotemporal Evolution of Ground-Motion Intensity at the Irpinia Near-Fault Observatory, Southern Italy. *Bull. Seism. Soc. Am.*, 112(1), 243-261.. https://doi.org/10.1785/0120210153

Picozzi, M., Spallarossa, D., Iaccarino, A. G., & Bindi, D., 2022. Temporal evolution of radiated energy to seismic moment scaling during the preparatory phase of the Mw 6.1, 2009 L'aquila earthquake (Italy). *Geophys. Res. Lett*, 49, e2021GL097382. https://doi.org/10.1029/2021GL097382.

Poiata, N., Satriano, C., Vilotte, J. P., Bernard, P. & Obara, K., 2016. Multiband array detection and location of seismic sources recorded by dense seismic networks. *Geophys. J. Int.*, 205(3), 1548-1573. https://doi.org/10.1093/gji/ggw071

Ross, Z. E., Yue, Y., Meier, M. A., Hauksson, E. & Heaton, T. H., 2019. PhaseLink: A deep learning approach to seismic phase association. *J. Geophys. Res.*: *Solid Earth*, 124(1), 856-869. https://doi.org/10.1029/2018JB016674

Savage, M. K., 2010. The role of fluids in earthquake generation in the 2009 Mw 6.3 L'Aquila, Italy, earthquake and its foreshocks. *Geology*, 38(11), 1055-1056.

https://doi.org/10.1130/focus112010.1

Scala, A., Adinolfi, G. M., Picozzi, M., Scotto di Uccio, F., Festa, G., De Landro, G., Priolo, E., Parolai, S., Riccio, R. & Romanelli, M., 2022. Monitoring the Microseismicity through a Dense Seismic Array and a Similarity Search Detection Technique: Application to the Seismic Monitoring of Collalto Gas-Storage, North Italy. *Energies*, 15(10), 3504. https://doi.org/10.3390/en15103504

Schaff, D. P. & Waldhauser, F., 2005. Waveform cross-correlation-based differential travel-time measurements at the Northern California Seismic Network. *Bull. Seism. Soc. Am.*, 95(6), 2446-2461. https://doi.org/10.1785/0120040221

Schaff, D. P., 2008. Semiempirical statistics of correlation-detector performance. *Bull. Seism. Soc. Am.*, 98(3), 1495-1507. https://doi.org/10.1785/0120060263

Scholz, C. H., 2015. On the stress dependence of the earthquake b value. *Geophys. Res. Lett.*, 42(5), 1399-1402. https://doi.org/10.1002/2014GL062863

Schorlemmer, D., Wiemer, S. & Wyss, M., 2005. Variations in earthquake-size distribution across different stress regimes. *Nature*, 437(7058), 539-542. https://doi.org/10.1038/nature04094

Shelly, D. R., Beroza, G. C. & Ide, S., 2007. Non-volcanic tremor and low-frequency earthquake swarms. *Nature*, 446(7133), 305-307. https://doi.org/10.1038/nature05666

Stabile, T. A., Satriano, C., Orefice, A., Festa, G. & Zollo, A., 2012. Anatomy of a microearthquake sequence on an active normal fault. *Sci. Rep.*, 2(1), 1-7. https://doi.org/10.1038/srep00410

Sugan, M., Campanella, S., Michele, M., Chiaraluce, L. & Vuan, A., 2022. The unlocking process of the 2016 Central Italy seismic sequence. https://doi.org/10.1002/essoar.10511493.1

Valensise, G. & Pantosti, D., 2001. The investigation of potential earthquake sources in peninsular Italy: a review. *J. Seismol.*, 5(3), 287-306. https://doi.org/10.1023/A:1011463223440

Vassallo, M., Festa, G. & Bobbio, A., 2012. Seismic ambient noise analysis in southern Italy. *Bull. Seism. Soc. Am.*, 102(2), 574-586. https://doi.org/10.1785/0120110018

Vassallo, M., Festa, G., Bobbio, A. & Serra, M., 2016. Low shear velocity in a normal fault system imaged by ambient noise cross correlation: The case of the Irpinia fault zone, Southern Italy. *J. Geophys. Res.*: *Solid Earth*, 121(6), 4290-4305. https://doi.org/10.1002/2015JB012410

Vuan, A., Brondi, P., Sugan, M., Chiaraluce, L., Di Stefano, R. & Michele, M., 2020. Intermittent Slip Along the Alto Tiberina Low-Angle Normal Fault in Central Italy. *Geophys. Res. Lett.*, 47(17), e2020GL089039. https://doi.org/10.1029/2020GL089039

Wang, J., Xiao, Z., Liu, C., Zhao, D. & Yao, Z., 2019. Deep learning for picking seismic arrival times. *J. Geophys. Res.*: *Solid Earth*, 124(7), 6612-6624. https://doi.org/10.1029/2019JB017536

Wiemer, S., 2001. A software package to analyze seismicity: ZMAP. *Seism. Res. Lett.*, 72(3), 373-382. https://doi.org/10.1785/gssrl.72.3.373

Yoon, C. E., O'Reilly, O., Bergen, K. J. & Beroza, G. C., 2015. Earthquake detection through computationally efficient similarity search. *Sci. Adv.*, 1(11), e1501057. https://doi.org/10.1126/sciadv.1501057

Yoon, C. E., Huang, Y., Ellsworth, W. L. & Beroza, G. C., 2017. Seismicity during the initial stages of the Guy-Greenbrier, Arkansas, earthquake sequence. *J. Geophys. Res.*: *Solid Earth*, 122(11), 9253-9274. https://doi.org/10.1002/2017JB014946

Yoon, C. E., Yoshimitsu, N., Ellsworth, W. L. & Beroza, G. C., 2019. Foreshocks and

mainshock nucleation of the 1999 M w 7.1 Hector Mine, California, Earthquake. *J. Geophys. Res*: *Solid Earth*, 124(2), 1569-1582. https://doi.org/10.1029/2018JB016383.

Zhang, M., Ellsworth, W. L. & Beroza, G. C., 2019. Rapid earthquake association and location. *Seism. Res. Lett.*, 90(6), 2276-2284. https://doi.org/10.1785/0220190052

Zhu, W. & Beroza, G. C., 2019. PhaseNet: a deep-neural-network-based seismic arrival-time picking method. *Geophys. J. Int.*, 216(1), 261-273. https://doi.org/10.1093/gji/ggy423

Zollo, A., Orefice, A. & Convertito, V., 2014. Source parameter scaling and radiation efficiency of microearthquakes along the Irpinia fault zone in southern Apennines, Italy. *J. Geophys. Res.*: *Solid Earth*, 119(4), 3256-3275. https://doi.org/10.1002/2013JB010116

Zollo, A., Caruso, A., De Landro, G., Colombelli, S. & Elia, L., 2021. A Bayesian Method for Real-Time Earthquake Location Using Multiparameter Data. *J. Geophys. Res.*: *Solid Earth*, 126(3), e2020JB020359. https://doi.org/10.1029/2020JB020359