Comparing and integrating artificial intelligence and similarity search detection techniques for seismic sequences in Southern Italy

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

[1]Università di Napoli Federico II
[2]University of Naples Federico II
[3]Stanford University

November 30, 2022

Abstract

Understanding mechanical processes occurring on faults requires detailed information on the microseismicity that can be today enhanced by advanced techniques for earthquake detection. This problem is more challenging when 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, for catalog improvement associated with seismic sequences in the normal fault system of Southern Apennines (Italy) using data from the Irpinia near fault observatory. We found that the integration of the machine learning and template matching detectors, the former providing templates for the cross-correlation, largely outperforms with respect to the techniques based on autocorrelation and machine learning alone, featuring an enrichment of the automatic and manual catalogs of factors 21 and 7 respectively. The output catalogs can be polluted by many false positives; so, we applied refined selection based on the cumulative distribution of the similarity level to clean up the detection lists and analyze final subsets dominated by real events. The magnitude of completeness decreases by more than one unit as compared to the reference value for the network. We report b-values 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.

Posted on 30 Nov 2022 — CC-BY-NC-ND 4 — https://doi.org/10.1002/essoar.10512172/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

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 today enhanced by advanced techniques for earthquake detection. This problem is more challenging when 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, for catalog improvement associated with seismic sequences in the normal fault system of Southern Apennines (Italy) using data from the Irpinia near fault observatory. We found that the integration of the machine learning and template matching detectors, the former providing templates for the cross-correlation, largely outperforms with respect to the techniques based on autocorrelation and machine learning alone, featuring an enrichment of the automatic and manual catalogs of factors 21 and 7 respectively. The output catalogs can be polluted by many false positives; so, we applied refined selection based on the cumulative distribution of the similarity level to clean up the detection lists and analyze final subsets dominated by real events.

The magnitude of completeness decreases by more than one unit as compared to the reference value for the network. We report b-values 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, Earthquake interaction, forecasting and prediction, Statistical seismology

1 Introduction

Earthquake detection is a fundamental task in seismology since it represents the first step for the analysis and the interpretation of Earth's crustal deformation processes. Detection consists in analyzing continuous records to identify discrete seismic events and becomes more difficult as the size of the events decreases and the amplitude of the earthquake signals approaches the noise level. With the increase of 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 techniques have become established through the years, and 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 (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 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 selected, 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, and hence might be biased by the number and quality of the reference signals.

Autocorrelation techniques mitigate this 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 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, using artificial intelligence, seismology has developed deep-learning models for event detection (Zhu & Beroza 2018; Wang *et al.* 2019; Dokht *et al.* 2019; Mousavi *et al.* 2020). 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).

Understanding which detection technique is more suitable to enhance catalogs is a challenging task that 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 within 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 these infrastructures 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 (red triangles), the distribution of the earthquakes in the INFO catalog (black dots), the location of the sequences analyzed in this study (yellow stars) and the epicenter of the M 6.9, 1980 Irpinia earthquake (red 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 of recent years appears to be spread 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 current monitoring strategies led to manually revised catalogs containing between 10 and 50 events for the seismic sequences. Nevertheless, Festa *et al.* (2021) showed that these seismic sequences may include hundreds of detectable events when analyzed with advanced tools.

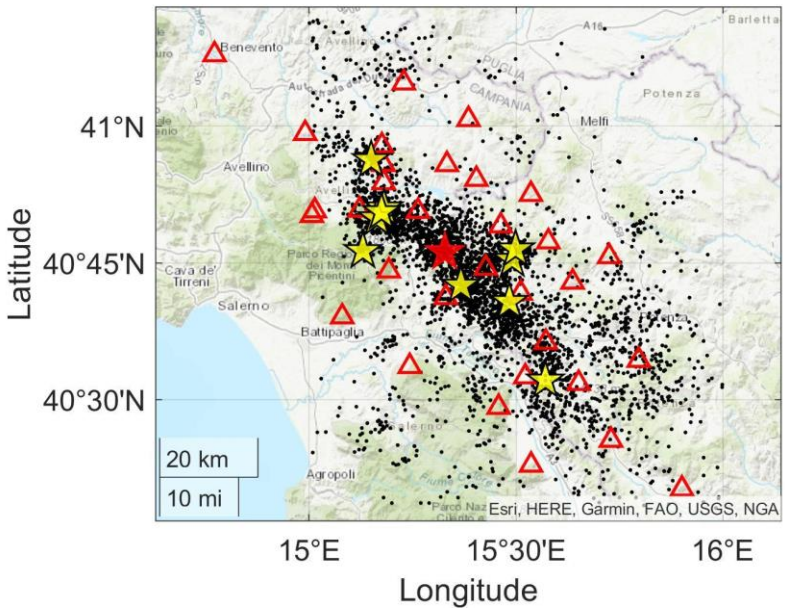

Figure (1). The location of the ISNet seismic stations (red triangles), the events detected by INFO from 2007 (black dots). The red star marks the epicenter of the 1980, M 6.9 earthquake; yellow stars indicate the locations of the main events of the 10 seismic sequences analyzed in this study.

For this reason, in this study 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 small scale, dense network monitoring normal faults) and for a specific task (to provide augmented catalogs for seismic sequences). We found that the integration of the machine learning and template matching techniques outperforms the autocorrelation detector. Enhanced catalogs feature a smaller magnitude of completeness as compared to the INFO catalogs, down to one point in magnitude, and enable to characterize the seismicity rate discovering a foreshock activity before the main event.

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 performance of the techniques and the 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, in the Discussion we present how the results affect the strategies for catalog generation and for a statistical description of the seismicity.

2 Data

In this work, 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 stations closest to the sequence centroid, spanning a hypocentral distance range between 10 km and 30 km. This choice was guided by the need for having a good azimuthal coverage with respect to the sequence.

The seismic sequences were selected using the information of the INFO catalog, available at the webpage http://isnet.unina.it/. The catalog has two layers: the first one consists of a catalog

automatically generated by the software Earthworm (Johnson *et al.* 1995), which runs on continuous data-streams. Then, 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. The final catalog consists of located events 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 10 and 50 (Table 1). 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 with respect to standard detection approaches, it is well-known that, depending on the tuning of working parameters, they can suffer from the occurrence of false detections, whose number becomes larger and larger as the acceptance threshold for event declaration is lowered. Therefore, after applying the techniques, further analysis 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. In the training phase ~ 1 M earthquake and ~ 300 k noise waveforms recorded at stations within epicentral distance up to 300

km are used, representing ~ 450 k events. The traces used in the training extend for 1 minute with a sampling rate of 100 Hz and are band-pass filtered in the range 1-45 Hz. During the training phase, the data was augmented by adding secondary earthquake signals and a random level of Gaussian noise into the earthquake waveform, and shifting the event within the trace to vary its position. The first operation allows 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 10s.

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.

Since these detections will be used as templates for template matching analysis, we evaluated the quality of the EQTransformer picks by comparing them to the manual picks. Therefore, we selected 75 phases from 13 events in the local magnitude range 1.0 – 3.0 that were also included in the INFO catalog. We found that EQTransformer automatic picks tend to anticipate manual picks. The distributions of the differences in the picks feature a mean and standard deviation of -

0.1 s and 0.2 s for P waves (Figure S1, left panel), and -0.2 s and 0.5 s for S waves (Figure S1, right panel) respectively.

3.2. EQCorrscan

We used EQCorrscan software (Chamberlain *et al.* 2018) for template matching analysis. We considered as templates the events declared by EQTransformer 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 approaches, 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 (ID5), 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. 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 further investigate the distribution of the previous parameter, 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.

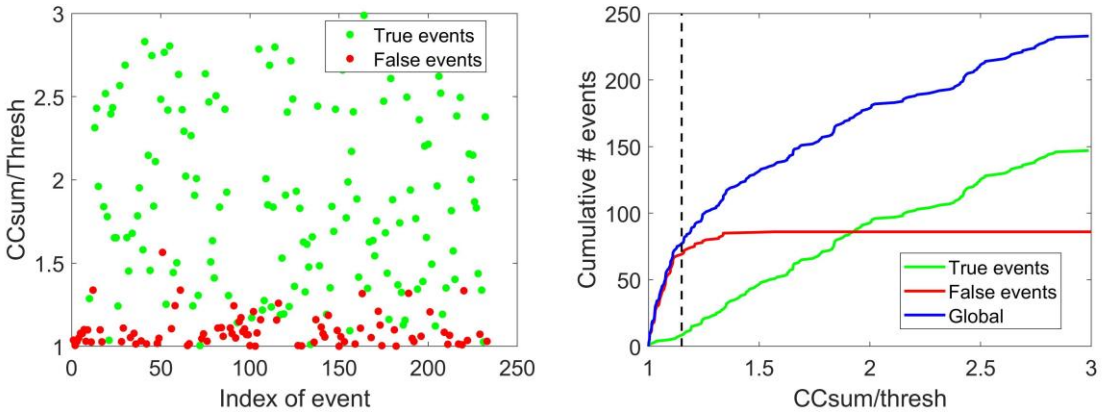

Figure (2) Left panel: Distribution of the ratio CCsum/threshold for the events in the initial catalog provided by EQCorrscan (green points correspond to real events, red points to false positives) for the ID5 sequence. Right panel: cumulative number of events as a function of the ratio CCsum/threshold 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 the separation between the distributions of false and real events can be identified when the number of samples for the two populations is comparable. 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, given the inter-station distance.

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 the 3% of the day length (~ 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 of which 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 having performed the detection analysis of the seismic sequences, we estimated the local magnitude Ml of seismic events. We first selected one event in each sequence with waveforms clearly emerging from the noise at all the stations. 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: $Ml = Ml_{ref} + \frac{A}{A_{ref}}$. 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 can be estimated to be 0.1 units of magnitude (Festa *et al.* 2021).

To investigate the frequency–magnitude distribution from the resulting catalogs, we evaluated the parameters of the Gutenberg-Richter relation ($logN = a - bM$, N number of events with local

magnitude > *M*) and the magnitude of completeness for each sequence, using the software ZMAP (Wiemer, 2001) and considering the local magnitude to characterize the event size.

For the Rocca San Felice seismic sequence (ID1), 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 applying the autocorrelation FAST, the machine learning EQTransformer (hereinafter, EQT) and the template matching EQCorrscan (hereinafter, TM) detectors for ten seismic sequences in the Irpinia region. In Figure 3, we show the detection catalogs for the three techniques, integrated over the ten sequences and organized in a Venn diagram. When analyzing the performances of the different approaches, it is worth highlighting that for the considered sequences, the manually revised INFO catalog contains 265 events only, 82 of which were automatically declared by the STA/LTA-based detector operating on the network.

The catalog obtained by merging all the independent detections (i.e., hereinafter 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 techniques to the merged catalog in terms of number of detections changes among sequences. We

report one sequence (ID1) with 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 performances of the single techniques, we report that FAST declares 941 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. Finally, TM declares 1715 events, with a catalog content similar to the merged one.

As can be seen from Figure 3, most of the events (95%) declared by FAST are also retrieved by the TM, using earthquakes declared by EQT as templates. This is the reason why the EQT detections are always included in the TM detection list.

FAST is able to declare 73 events that are missed by the other two 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 nearby the sequence centroid. For the former class of events, they exhibit a smaller cross-correlation value with the used templates and can be still retrieved by the TM technique, by lowering the acceptance threshold, at the cost of a significant increase of the number of false events. As an example, for the San Gregorio Magno sequence (ID3), the 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.

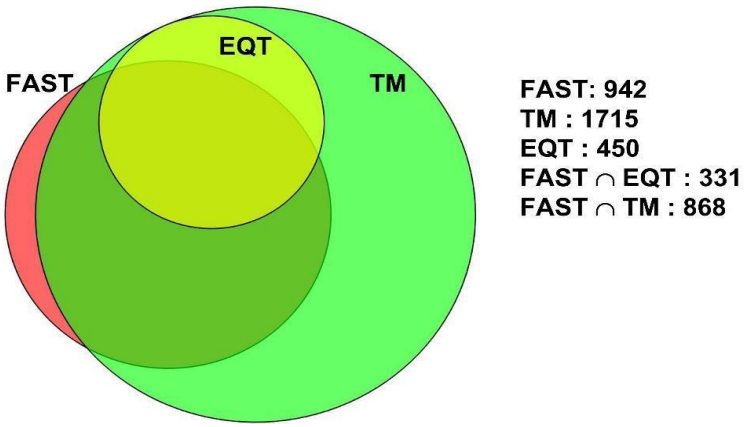

Figure (3) Venn diagram showing the performance of the different detectors: the autocorrelation (FAST, red), machine learning (EQT, yellow) and template matching (TM, green) techniques. The EQT detections are included in the TM ones because the EQT output is used to form templates.

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.

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 occur clustered in time, especially when the earthquakes occur in the coda of previous ones. For the other sequences, the performances of TM and FAST are similar. However, we observe that also when we analyze the catalog for single sequences, TM is always able to catch more than 85% of the merged catalog.

We performed some tests 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 (ID3) as an example, the decrease of 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 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 catalog). Indeed, when we limit the set of templates to those provided by the revised manual catalog, the number of the TM detections significantly decreases and it becomes comparable to or smaller than the FAST catalog.

Finally, the number of EQT detections is much smaller than the other two 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 TM improving its performance.

In Figure 4 we report the detection performances of each technique for the 10 analyzed seismic sequences, also comparing them with the INFO manual catalogs. The Rocca San Felice sequence (ID1, left panel of Figure 4) is separated to improve the visualization of the results of the other sequences (ID2-10, right panel of Figure 4). We note that when TM outperforms FAST we also have a wider set of templates provided by EQT (orange bars) with respect to the dimension of the INFO catalog (violet bars).

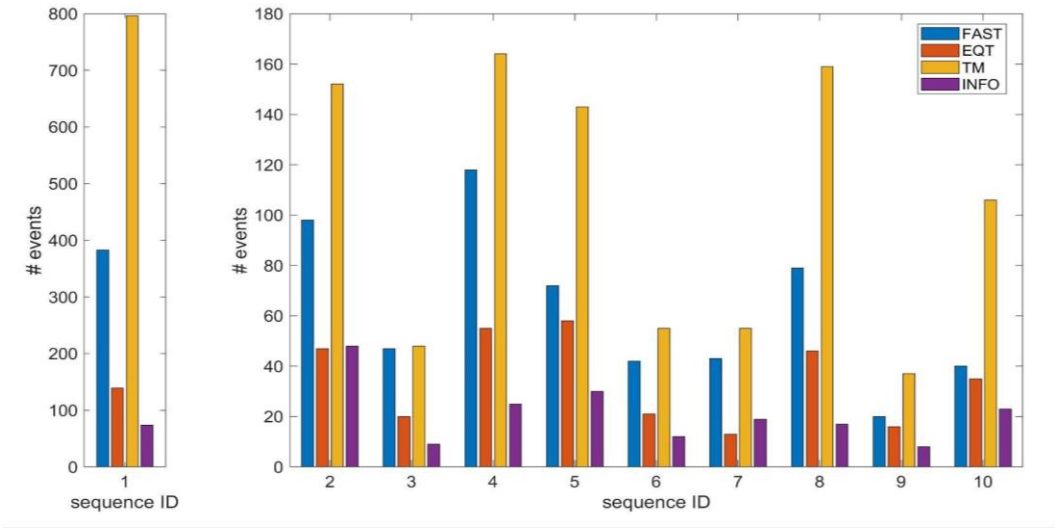

Figure (4) Detection performance of each technique for the analyzed sequences. The Rocca San Felice sequence (ID1) is shown in the left panel, all the other sequences in the right panel. The histograms report the detections of FAST (blue), EQT (orange), TM (yellow) and INFO (violet).

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 (ID1), 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 three detectors.

When looking at the outcomes of the three techniques, FAST output consists of 1406 detections, reduced to 500 after the application of the automatic selection criterion. These events were then 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 latter as templates for TM, we obtain a list of 1394 declarations, 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 TM). The size of the final detection catalogs from the three techniques are presented in Figure 4.

As previously observed, comparing the three catalogs we note that most of the detections retrieved by TM and missed by FAST are 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 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. This parametrization quickly becomes prohibitive, when applied to longer data-streams, and cannot be used in standard processing.

The few cases (i.e., only 7 events) where TM is not able to detect events seen by FAST correspond to signals buried in the noise. In Figure S5, we show a set of 3 events declared by 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 TM.

The final combined catalog (i.e., considering the independent detections of FAST and 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, TM outperforms FAST due to the wide magnitude range and the large number of the templates provided by EQT. Indeed, when we limit the set of templates to the 74 manually detected events in INFO, we observe that TM yields a catalog of about 300 events (~ 40 % of the previous TM catalog), which is smaller than the FAST final event list.

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 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 TM coincides with those obtained for the combined catalogs. Also FAST provides Mc estimates similar to those of the combined catalogs, with the exception of sequences ID1 and ID5, 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 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 (ID1) in Figure 5.

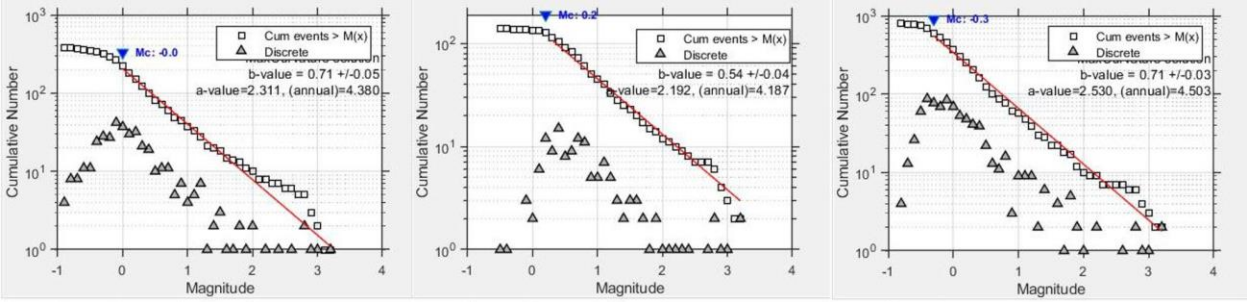

Figure (5) Gutenberg-Richter law for FAST (Mc 0.0 and b=0.71 ± 0.05; left panel), EQT (Mc 0.2 and b = 0.54 ± 0.04; central panel) and TM (Mc -0.3 and b = 0.71 ± 0.03; right panel) catalogs.

We find that the slope and the magnitude of completeness are Mc 1.2 and b_{M_l} = 0.83±0.12 for manual INFO catalog (not shown in Figure), Mc 0.0 and b_{M_l} = 0.71±0.05 for FAST, Mc 0.2 and b_{M_l} = 0.54±0.04 for EQT, Mc -0.3 and b_{M_l} = 0.71±0.03 for 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.

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 local magnitude estimates into seismic moment using the empirical relationship $\log M_0 = 1.31Ml + 10.55$ (standard deviation 0.12, Figure S6) which has been derived considering the ISNet 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 associating to a given point a number of events that occurred within a given distance range. Therefore, the b-values from Picozzi *et al.* (2022b) provide a spatial average of the stress level in a rather large crustal volume with respect to the b-values obtained for the sequences, which typically have limited spatial 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.

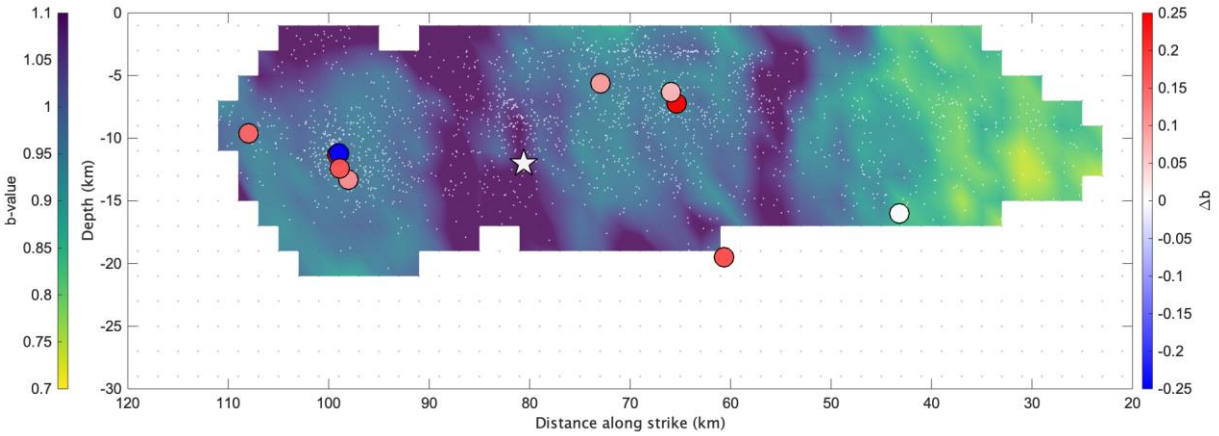

Figure (6) Map showing the spatial distribution of average b-values from bootstrap analysis (redrawn from Picozzi et al. 2022b). In the figure we show the nucleation point for the 1980 Irpinia earthquake (white star) and Δb-value (see text for the definition, colored circles) for the sequences analyzed here.

Finally, we analyzed the time evolution of 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. In particular, the larger value in seismic moment is associated with a main event of Ml 2.7 (ID2), 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 (ID5), 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 80%, of 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 % foreshocks.

In the time evolution of the sequences, we can distinguish two behaviors. Most of the sequences generate the majority of earthquakes (around 70 % of the catalog) within 6 hours after the mainshock and are 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 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.

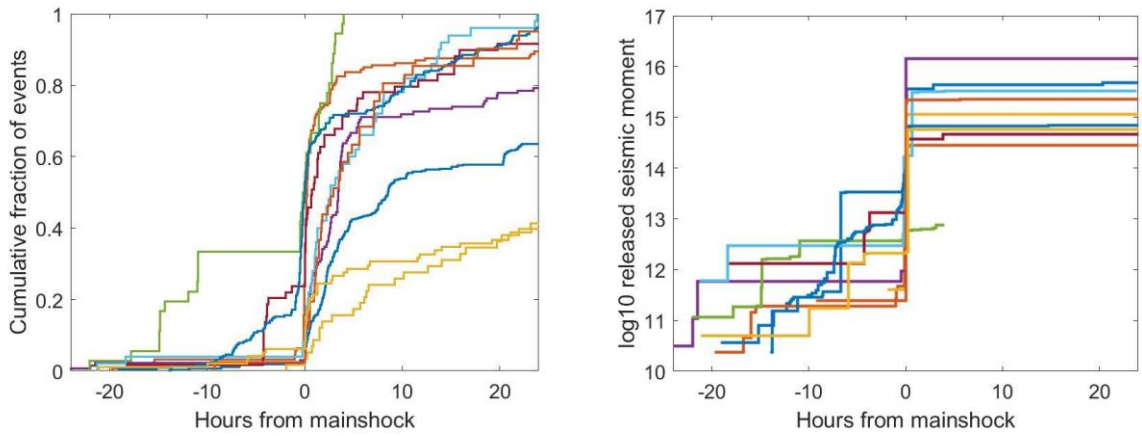

Figure (7) Foreshock/aftershock analysis: left panel shows the cumulative percentage of the events in the catalog as a function of the time from the main event. The right panel contains the cumulative seismic moment in a shorter time window before the main event.

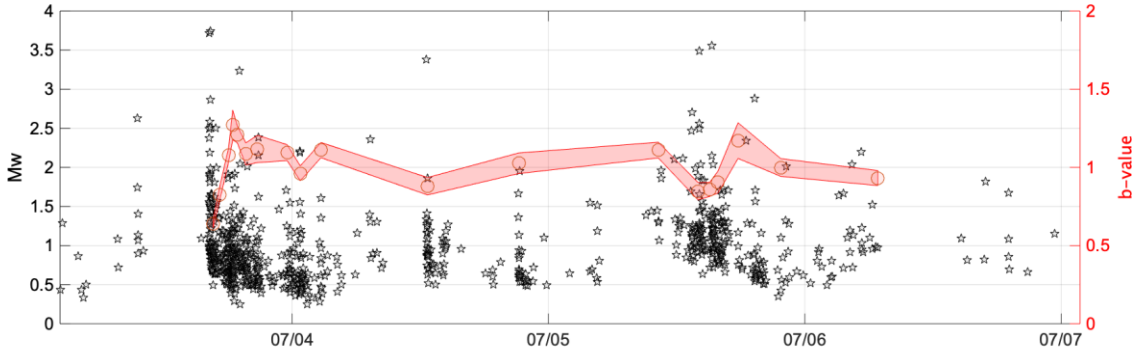

Figure (8): temporal evolution of b-value (red circles) computed for the Rocca San Felice sequence (ID1) with 95% probability uncertainty (red band) and event distribution as a function of time and moment magnitude (black stars).

6. Discussion

The use of new, advanced techniques for enhancing seismic catalogs is crucial: both to monitoring how seismicity evolves with time within a complex geological system, and to understanding how faults slip to generate moderate to large earthquakes, possibly anticipated by the occurrence of small seismic events. This latter scientific issue becomes more 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 (TM) outperforms the autocorrelation technique FAST.

The integration of the two techniques (i.e., EQT and TM) is beneficial because EQT allows building a rich catalog of templates for TM. Using only as templates the events automatically detected by the INFO system or revised by network operators, the final catalog of TM significantly decreases, at the same level or below the performances of FAST. This latter technique shows similar performances 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 here analyzed.

Both TM 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 TM 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 corresponds 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 (ID1) 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 thus prefer to maintain a more conservative approach for EQT and to enhance the catalog integrating it with the TM 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 TM and FAST provide consistent estimates, while the b-value from EQT is systematically lower. This is attributable 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 near 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.* 2022).

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 decrease of the b-value for the sequences indicates that they usually occur 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 (ID8) 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 7). Here we found a rapid increase of the b-value around the main event of the sequence, followed by an almost constant level. Monitoring the b-value with time has been claimed to be a discriminator of foreshocks and aftershocks. In some cases, both for real earthquakes (Gulia *et al.* 2016; Gulia & Wiemer 2019) and laboratory experiments (Amitrano, 2003), it has been observed to increase for aftershocks and decrease for foreshocks. Hence, these enhanced catalogs allow to study the variation of b-value in time and to confirm the possibility to be used as a risk mitigation 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 125Hz, 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 too large uncertainties and biases (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 and 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 picking with FAST would require additional post-processing on the detected events.

7 Conclusions

In this study, we compared the performance of three detectors: the autocorrelation method FAST, the machine learning technique EQTransformer (EQT) and the template matching technique EQCorrscan (TM). We selected seismic sequences occurring in the Irpinia region, Southern Italy as case study. 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 with respect to the reference catalogs for the area. We found that the integration of EQT and TM, that is the use of TM from a set of templates provided by EQT significantly outperforms FAST, almost doubling its 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 TM detections) individuated by FAST and missed by TM occur at low signal-to-noise ratio or they appear as isolated couples exhibiting shallower location. As expected, the performance of EQT alone, with the adopted trained model, is lower than the other two techniques. However, it still improves the INFO catalog by a factor 1.5. Finally, the merged catalog improves the INFO catalogs of factors ~7 (manual catalog) and ~21 (automatic catalog).

Both FAST and TM 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 TM) 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 integrating EQT and 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 is compatible with that of TM in most cases, while EQT provides larger Mc estimates. The *b*-values from TM and FAST are compatible between 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 biases in the Gutenberg-Richter determination, when covering a small range in magnitude above the completeness threshold.

We also recognize 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 to better understanding 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).

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.

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

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.

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

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., Maria Adinolfi, G., 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

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

Supplementary material for :

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.

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

Parametric test for EQTransformer

To better understand how the input parameters influence the detection performance of EQTransformer, we performed a parametric study on a specific sequence (Rocca San Felice sequence, ID1) starting from the following values: det_thresh=0.3, P_thresh=S_thresh=0.1, overlap=30%, batch size =500.

We systematically explored the batch size value, fixing the threshold parameters to the default values and reported an almost constant number of detections within the range 100-600. Reducing the detection threshold parameter from 0.3 to 0.15, to be more permissive in the selection of transients, we did not find significant changes in the number of events. When we lower the threshold for all the parameters (det_thresh=0.1, P_thresh=S_thresh=0.08), the final catalog increases by only about 10%. This latter increase, however, is not considered significant with respect to the catalog obtained with the initial parameterization, while it increases the risk of false detections. Therefore, we decided to keep the initial values of the parameters as reference for this study. We only changed the batch size to 100, to reduce memory demand**.**

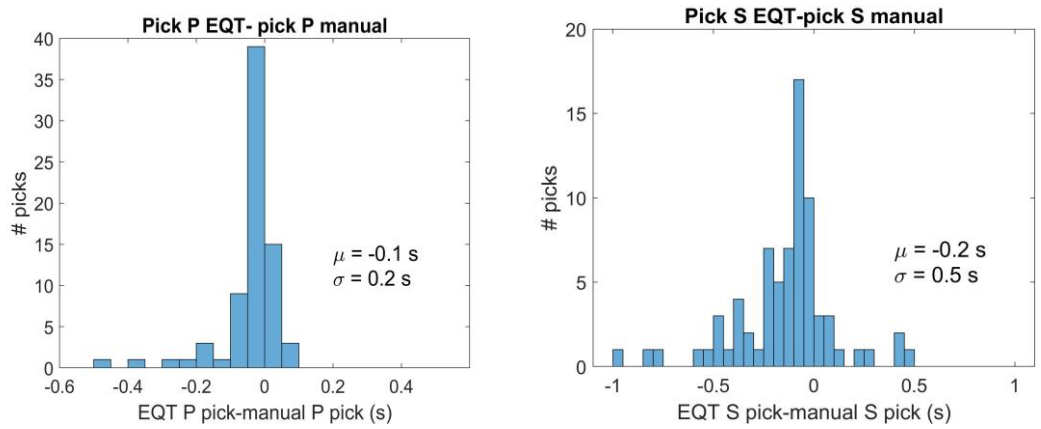

Figure S1. Comparison between the manual and EQTransformer automatic picks for 13 events (left panel: P picks; right panel: S picks). Automatic picks tend to anticipate manual ones, with a larger lag for S waves**.**

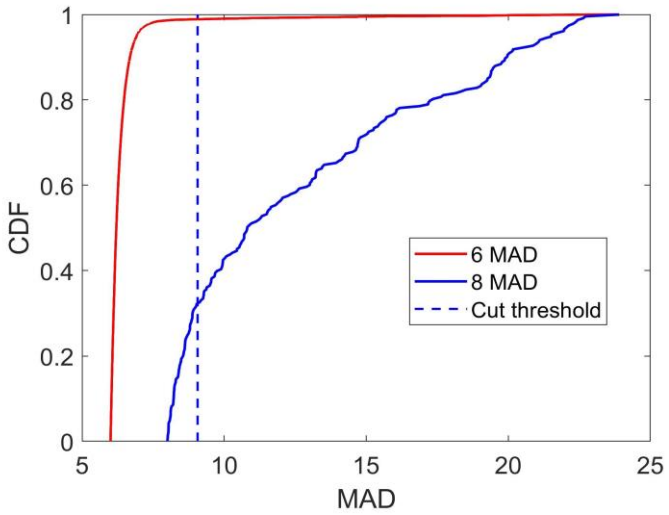

Figure S2. Cumulative distribution of the ratio *CCsum*/MAD for the catalog obtained with 6 MAD (red curve) and 8 MAD (blue curve) as thresholds.

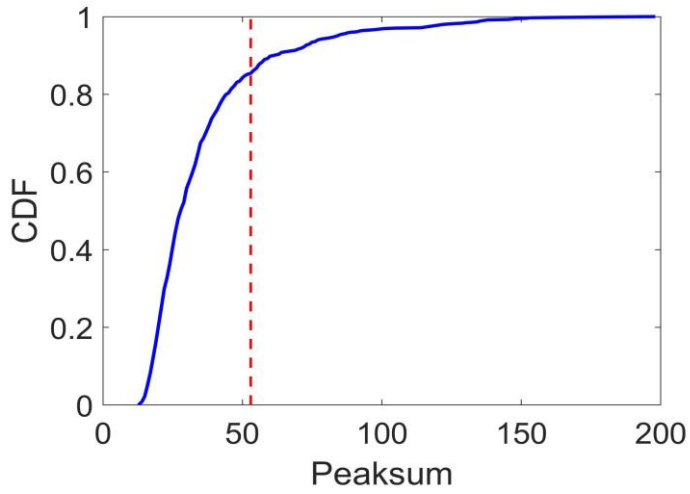

Figure S3. An example of cumulative distribution of the *peaksum* parameter, for events declared at 2 stations with FAST. We clearly identify a change in the slope of the distribution, around the 90% percentile, which is considered as the threshold to discriminate real events from false positives.

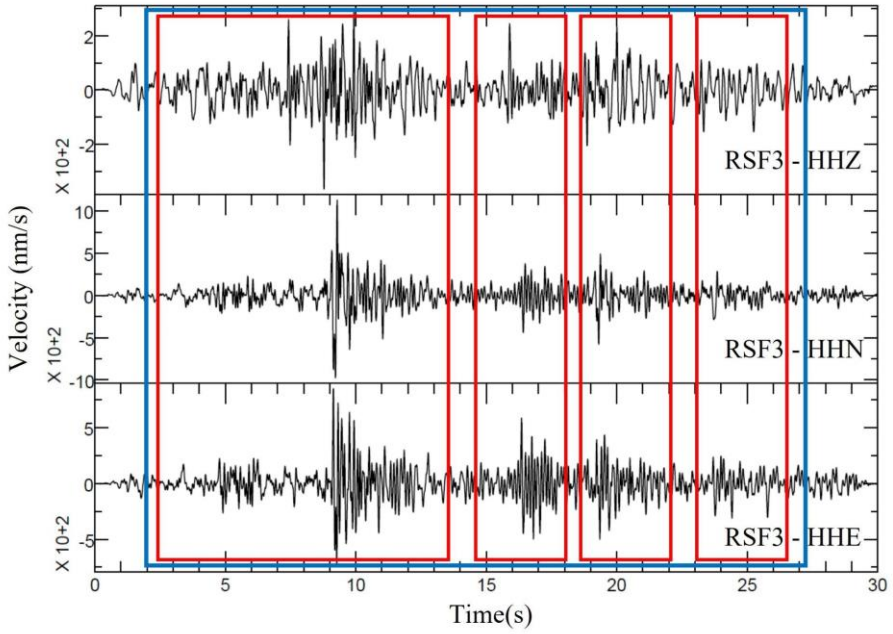

Figure S4. Events detected from TM (red boxes) and FAST (blue box): FAST is not able to discriminate events occurring closely in time.

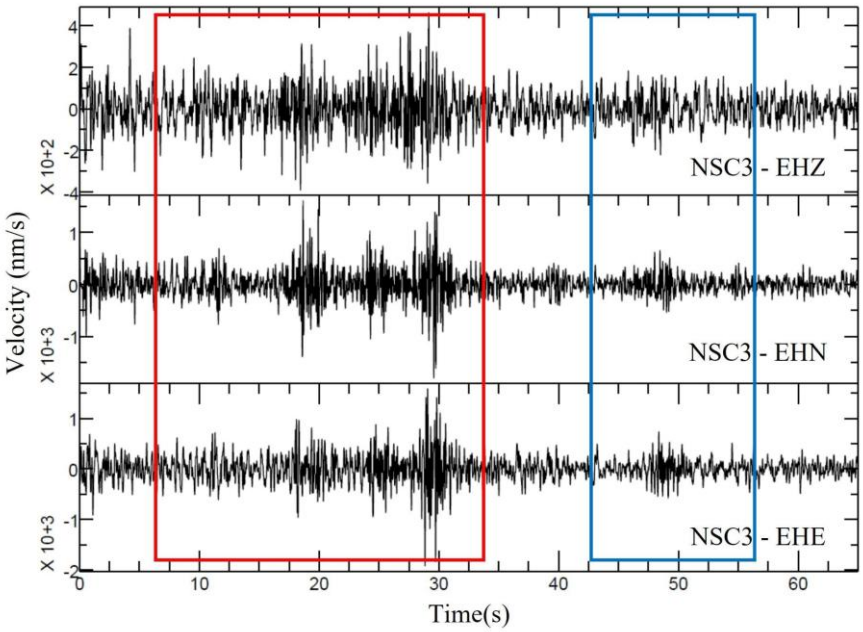

Figure S5. An example of a low SNR event declared by FAST and missed by TM. The three events close in time in the red box are correctly separately detected by TM.

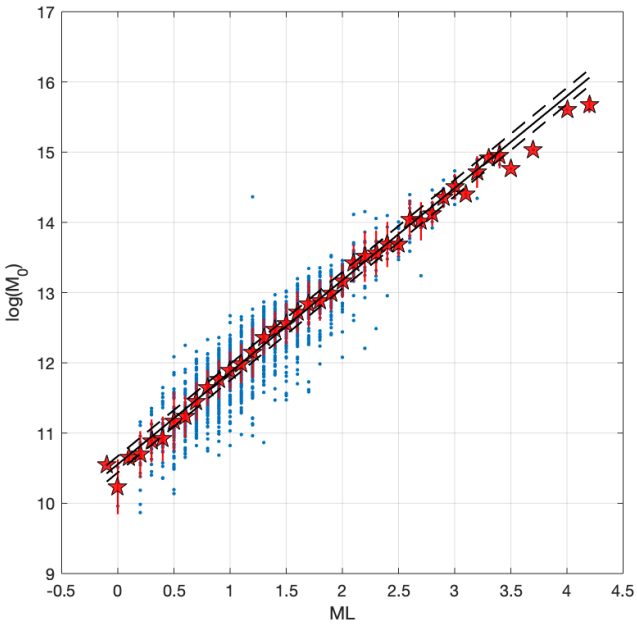

Figure S6. Linear regression between $\log M_0$ and Ml using the historical earthquakes recorded in the Irpinia region. We obtain $\log M_0 \ = \ 1.31 \ Ml \ + \ 10.55$, with $\sigma = 0.12$.

ID	Ml main event	FAST	EQT	TM
1	3.0	Mc 0.0 $b = 0.71 \pm 0.05$	Mc 0.2 $b = 0.54 \pm 0.04$	Mc -0.3 $b = 0.71 \pm 0.03$
2	2.7	Mc -0.3 $b = 0.60 \pm 0.05$	Mc -0.2 $b = 0.51 \pm 0.07$	Mc -0.3 $b = 0.68 \pm 0.06$
3	2.8	Mc 0.3 $b = 0.62 \pm 0.11$	N.A	Mc 0.4 $b = 0.76 \pm 0.13$
4	3.7	Mc -0.1 $b = 0.76 \pm 0.08$	Mc 0.1 $b = 0.54 \pm 0.08$	Mc -0.2 $b = 0.65 \pm 0.07$
5	1.8	Mc 0.3 $b = 0.83 \pm 0.10$	Mc 0.3 $b = 0.75 \pm 0.10$	Mc -0.3 $b = 0.73 \pm 0.05$
6	3.0	Mc 0.3 $b = 0.75 \pm 0.18$	Mc 0.4 $b = 0.53 \pm 0.13$	Mc 0.2 $b = 0.69 \pm 0.11$
7	2.7	Mc 0.4 $b = 0.66 \pm 0.13$	N.A	Mc 0.3 $b = 0.78 \pm 0.14$
8	2.8	Mc 0.1 $b = 1.16 \pm 0.11$	Mc 0.9 $b = 1.00 \pm 0.14$	Mc 0.2 $b = 1.26 \pm 0.13$
9	2.9	N.A	N.A	Mc -0.3 $b = 0.60 \pm 0.09$
10	3.1	Mc 0.2 $b = 0.73 \pm 0.16$	Mc 0.2 $b = 0.53 \pm 0.09$	Mc 0.1 $b = 0.76 \pm 0.13$

Table S1. Magnitude of completeness, b-values and magnitude of the main event in the sequence, using the different detection techniques for the 10 analyzed sequences