



UNIVERSITÀ DEGLI STUDI DI NAPOLI
FEDERICO II

Scuola Politecnica e
delle Scienze di Base



Università degli Studi di Napoli Federico II
Dottorato di Ricerca in
Ingegneria Strutturale, Geotecnica e Rischio Sismico

THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Detection and characterization of microseismicity using advanced techniques

by

Francesco Scotto di Uccio

Advisor: Prof. Gaetano Festa

Co-advisor: Prof. Matteo Picozzi



SCUOLA POLITECNICA E DELLE SCIENZE DI BASE
DIPARTIMENTO D STRUTTURE PER L'INGEGNERIA E L'ARCHITETTURA

Detection and characterization of microseismicity using advanced techniques

Ph.D. Thesis presented
for the fulfillment of the Degree of Doctor of Philosophy
in Ingegneria Strutturale, Geotecnica e Rischio Sismico
by
Francesco Scotto di Uccio

December 2024



Approved as to style and content by

Handwritten signature of Gaetano Festa in black ink.

Prof. Gaetano Festa

Handwritten signature of Matteo Picozzi in black ink.

Prof. Matteo Picozzi

Università degli Studi di Napoli Federico II

Ph.D. Program in Ingegneria Strutturale, Geotecnica e Rischio Sismico

XXXVII cycle - Chairman: Prof. Iunio Iervolino



www.dist.unina.it/dottorati-di-ricerca/dottorati

Candidate's declaration

I hereby declare that this thesis submitted to obtain the academic degree of Philosophiæ Doctor (Ph.D.) in Ingegneria Strutturale, Geotecnica e Rischio Sismico is my own unaided work, that have not used other than the sources indicated, and that all direct and indirect sources are acknowledged as references.

Parts of this dissertation have been published in international journals and/or conference proceedings (see list of the author's publications at the end of the thesis).

Napoli, December 4, 2024

A handwritten signature in black ink, reading "Francesco Scotto di Uccio". The signature is written in a cursive, slightly slanted style.

Francesco Scotto di Uccio

Abstract

Earthquakes are catastrophic phenomena which cause damage and raise concerns among population and institutions, with an increasing necessity of identifying actions for mitigating their impact and effects. Moreover, since the occurrence rate of high magnitude, damaging events is low, direct access to their observation is limited. Nevertheless, microseismicity continuously occurs within the same active seismogenic faults where a major earthquake might be generated, and thus, the analysis of small magnitude events can provide crucial insights into the mechanical and stress state of the faults and the preparatory phase of large earthquakes. However, the capability of identifying micro-earthquakes is strongly affected by the low amplitude level on the seismic records, which is typically comparable to the ambient noise. The growth of advanced monitoring systems and the development of improved strategies for earthquake identification and characterization are acting as a powerful lens to reveal the small-scale rupture processes. In this thesis, I apply advanced techniques for characterizing the microseismicity in terms of earthquake detection, accurate location and seismic source properties estimation. I generate enhanced seismic catalogs using machine learning and waveform similarity detection techniques, identifying one order of magnitude more earthquakes as compared to the existing catalogs, which I investigate for identifying seismogenic structures in different tectonic and volcanic environments. The achieved resolution on the hypocentral locations and on the properties characterizing the seismic sources led to the development of models for understanding the generation and the evolution of earthquakes.

Keywords: Microseismicity characterization, Seismic sequences, Deep learning techniques, Advanced detection strategies, Hypocenter determination, Seismic source.

Sintesi in lingua italiana

I terremoti sono fenomeni catastrofici che causano danni e diffondono preoccupazione tra la popolazione e le istituzioni, con un crescente aumento della necessità dell'individuazione di azioni per la mitigazione del loro impatto ed effetto. Inoltre, dato il basso tasso di occorrenza dei terremoti distruttivi di alta magnitudo, l'accesso alle loro osservazioni è limitato. Tuttavia, la microsismicità avviene con continuità all'interno della stessa area sismogenetica in cui un grande terremoto può essere generato e dunque l'analisi dei terremoti di bassa magnitudo può fornire informazioni cruciali sullo stato meccanico e di sforzo delle faglie e sulla fase preparatoria dei grandi terremoti. Tuttavia, la capacità di identificazione della microsismicità è fortemente influenzata dal basso livello di ampiezza sulle registrazioni sismiche, che è tipicamente confrontabile con il livello del rumore ambientale. La combinazione di sistemi di monitoraggio avanzati e lo sviluppo di strategie innovative per l'identificazione e la caratterizzazione dei terremoti agiscono come lenti per rivelare i processi di rotture su scale spaziali sempre più piccole. In questa tesi ho applicato tecniche avanzate per la caratterizzazione della microsismicità in termini di detezione dei terremoti, localizzazione accurata e stima dei parametri relativi alla sorgente sismica. Ho generato cataloghi sismici aumentati utilizzando tecniche basate sull'intelligenza artificiale e sulla similarità delle forme d'onda, identificando un numero di terremoti superiore di un ordine di grandezza a quanto riportato nei cataloghi esistenti, che ho investigato per l'identificazione delle strutture sismogenetiche in ambienti tettonici e vulcanici. La risoluzione ottenuta sulla localizzazione degli ipocentri e sulle proprietà della sorgente sismica ha permesso lo sviluppo di modelli per la comprensione della generazione ed evoluzione della sismicità.

Parole chiave: Caratterizzazione della microsismicità, Sequenze sismiche, Tecniche di intelligenza artificiale, Strategie di detezione avanzate, Localizzazione degli ipocentri, Sorgente sismica.

Contents

Abstract	i
Sintesi in lingua italiana	ii
Contents	iii
Acknowledgements	vi
List of Acronyms	viii
List of Figures	x
List of Tables	xxvii
List of Symbols	xxviii
Introduction	1
Advanced techniques for earthquake detection	5
2.1 Introduction.....	5
2.2 Methods	14
2.2.1 Machine learning models for earthquake detection.....	14
2.2.2 An efficient template matching strategy: EQCorrscan	19
2.2.3 An innovative technique for autocorrelation: FAST	22
2.3 Application of advanced detection techniques	27
2.3.1 Comparing and integrating advanced detection techniques for seismic sequences in Southern Italy	27

2.3.2	Seismic monitoring using temporary dense array deployments	41
2.3.3	An enhanced catalog of repeating earthquakes in Southern Italy	48
2.3.4	Monitoring microseismic activity in the gas storage area of Collalto (Northern Italy).....	51
2.3.5	An automatic detection strategy for continuous seismic monitoring.....	55
	Accurate determination of earthquake locations	65
3.1	Introduction.....	65
3.2	Methods	72
3.2.1	A global probabilistic approach for earthquake location: NLLoc	72
3.2.2	Relative location techniques for hypocenter location.....	75
3.2.2.1	A linearized approach for relative earthquake location: HYPODD.....	75
3.2.2.2	Global approach for relative earthquake location: NLLoc-SSST-Coherence	78
3.2.3	Accuracy of automatic phase arrival times and influence of the velocity model in the hypocenter determination	81
3.3	Unveiling seismogenic structures through accurate earthquake location.....	87
3.3.1	Depicting fault segments from the analysis of enhanced catalogs for seismic sequences	87
3.3.2	Using ultra-microseismic events for illuminating fault segments.....	95
3.3.3	Delineation of fault zones activated during the 2014–2024 unrest at the Campi Flegrei Caldera (Southern Italy)	103
	Microseismicity characterization	117

4.1	Introduction.....	117
4.2	Source characterization	118
4.2.1	Source parameters	118
4.2.2	Stress release modelling	126
4.2.3	Focal mechanisms estimation	131
4.2.4	Stress field inversion	137
4.3	Source characterization for microseismic events in the Southern Apennines	143
4.3.1	Source parameters estimation of events in enhanced catalogs for seismic sequences in Southern Apennines.....	143
4.3.2	Stress release modelling for seismic sequences in Southern Apennines.....	153
4.3.3	Focal mechanism estimation for earthquakes in Southern Apennines using dense array deployments	163
4.3.4	Stress field inversion from focal mechanisms using dense array deployments	184
	Conclusions	191
	Bibliography	199
	Author's publications	223

Acknowledgements

I would like to thank prof. Gaetano Festa and prof. Matteo Picozzi for guiding me during the PhD, giving me the opportunity to present and discuss the ongoing activities within the scientific community, building strong connections with different research groups.

I am grateful to Dr. Patricia Martinez-Garzon and prof. Greg Beroza for hosting me during the PhD. The continuous discussions with you and your research group have been truly inspiring, allowing me to significantly widen the research perspectives.

I would like to acknowledge the reviewers, prof. Margarita Segou and prof. George Kaviris for their precious and significant comments that improved the quality of the thesis.

Part of the activities that led to the results discussed in this PhD thesis received support by the PRIN-FRACTURES (supported under the National Recovery and Resilience Plan (NRRP), Mission 4, Component 2, Investment 1.1, program PRIN 2022, Project "Multiscale study of seismogenic processes in Campania-Lucania Apennines using machine learning algorithms and multiparametric observations - FRACTURES) and GEOInquire (Geo-INQUIRE is funded by the European Commission under project number 101058518 within the HORIZON-INFRA-2021-SERV-01 call) projects.

This PhD thesis used the Irpinia Near Fault Observatory (<https://isnet.unina.it>) data and products. Seismic data were available at the EIDA website (<https://eida.ingv.it/it>) and at the EPOS Data Portal (<https://www.epos-eu.org/dataportal>) - IRPINIA Seismic Velocity and Acceleration Waveforms (Continuous) provided by Università di Napoli Federico II and INGV, networks IX and IV. Seismic catalog was available at the Irpinia Near Fault Observatory website (<http://isnet-bulletin.fisica.unina.it/cgi-bin/isnet-events/isnet.cgi>) and at the EPOS Data Portal (<https://www.epos-eu.org/dataportal>)- IRPINIA Seismic Events provided by Università di Napoli Federico II.

The phase arrival times used in this study for the analysis of the seismicity in the Campi Flegrei caldera are available at the INGV-Osservatorio Vesuviano bulletin database (<https://terremoti.ov.ingv.it/gossip/index.html>). Information is available per event. Seismic waveforms can be accessed through EIDA portal (<https://eida.ingv.it/it/>), network code IV.



List of Acronyms

The following acronyms are used throughout the thesis.

CC	Cross Correlation
CNN	Convolutional Neural Network
DL	Deep Learning
EGF	Empirical Green's Function
EQT	EQTransformer
FF	Far Field
GPD	Generalized Phase Detection
IDX - ID	Index
IF	Intermediate Field
INFO	Irpinia Near Fault Observatory
INGV	Istituto Nazionale di Geofisica e Vulcanologia
INGV-OV	Istituto Nazionale di Geofisica e Vulcanologia – Osservatorio Vesuviano
ISNet	Irpinia Seismic Network
LSQR	Least Square inversion
MAD	Median Absolute Deviation
ML	Machine Learning
NF	Near Field
PDF	Probability Density Function
PSD	Power Spectral Density

RMS	Root Mean Square error
RSC	Rete Sismica di Collalto
STA/LTA	Short Time Average / Long Time Average
SNR	Signal to Noise Ratio
SSST	Source-Specific Station Travel time correction
SVD	Singular Value Decomposition
TM	Template matching

List of Figures

- Figure 2.1.** Panel (a): Temporal evolution of earthquake magnitudes during the L’Aquila seismic sequence. Panel (b): Number of events in each magnitude bin (white circles) and cumulative frequency-magnitude (black circles) distributions of events for all seismic events. Figure extracted from De Santis et al. (2011)..... 6
- Figure 2.2.** Application of the STA/LTA technique on a M2.9 earthquake occurred in Southern Apennines. Left panel: Application of the STA/LTA technique for the closest station (SCL3, hypocentral distance of ~5 km). The red and blue bars in the upper panels mark the onset and the offset of the trigger, respectively. The impulsive arrival at the station produces a sharp variation in the STA/LTA function, leading to the declaration of the event. Right panel: Application of the STA/LTA technique for a more distant station (AND3, hypocentral distance of ~40 km). The emergent signal is barely above the noise level, producing a maximum value of the STA/LTA window close to the detection threshold..... 8
- Figure 2.3.** Distribution of normalized cross-correlation values versus interevent separation distance for P (red) and S wave (black), extracted from Michele et al., (2020)..... 10
- Figure 2.4.** Autocorrelation runtimes (purple) as a function of the continuous data duration. For continuous data longer than 1 week, the runtime is extrapolated based on quadratic scaling (violet dashed line). Figure extracted from Yoon et al. (2015)..... 11
- Figure 2.5.** (a) Illustration of a convolutional neural network for generalized phase detection (GPD) and (b) application example. Probability time series for P (red) and S waves (blue) are used as characteristic functions for phase detection. The figure is extracted from Figure 1 in Ross et al., (2018)..... 12
- Figure 2.6.** Location – magnitude distribution of the STEAD dataset. The figure is extracted from Figure 3 of Mousavi et al., (2019)..... 15
- Figure 2.7.** Application of EQT for a seismic station in Southern Apennines. The three upper panels report the normalized and [1-45] Hz filtered waveforms for the Z, N and E component of velocity records, respectively. The bottom

panel reports the prediction scores for earthquake detection (green), P (red) and S (blue), respectively16

Figure 2.8. Network architecture of PhaseNet. The blue rectangles represent layers inside the neural network, while the numbers near them are the dimensions of each layer. The input seismic data go through four down-sampling stages and four up-sampling stages. The down-sampling is done by 1-D convolution and stride, while the up-sampling is done by deconvolution, which recovers the input length of the previous stage. Figure extracted from Figure 5 in Zhu & Beroza (2019)..... 17

Figure 2.9. Comparative analysis in logarithmic scale of the picks provided by EQT conservative model (blue bars), PhaseNet (orange bars) and EQT original models (red bars). For each model, an increase of one order of magnitude is typically observed, with the number of PhaseNet picks laying between the two version of EQT model..... 18

Figure 2.10. Example of a template event used the phase arrival times provided by EQT. Red waveforms indicate the 1.5s – long portions extracted around the P picks on the vertical components of the velocity records, while the cyan sectors indicate the waveforms extracted around the S picks on the horizontal components of the velocity records.....20

Figure 2.11. Left-hand panel: distribution of the ratio $CC_{sum}/threshold$ for the events in the initial catalog provided by EQCorrscan (green points correspond to real events, red points to false positives). Right-hand panel: cumulative number of events as a function of the ratio $CC_{sum}/threshold$ for the subset of the real events (green line), false positives (red line) and for the whole initial catalog (blue line). The black dashed line marks the similarity parameters associated with the slope change in the global $CC_{sum}/threshold$ distribution, which we use as refined threshold.....21

Figure 2.12. Conversion of time-domain waveforms into binary fingerprints containing only the discriminative features of earthquakes. Figure is redrawn after Yoon et al., (2015).....23

Figure 2.13. Left panel: The upper panel reports the raw waveform at LIO3 station, while the lower panel contains the [1-10] Hz filtered data, which feature a monochromatic behaviour, this latter causing an increase in the number of similar fingerprints for that station. Right panel: Daily power spectral density (PSD) for LIO3 station for the considered day. The PSD shows a spurious peak around the characteristic frequency of the monochromatic signal dominating the waveforms after the filtering.....26

- Figure 2.14.** Manual ISNet catalog from January 2007 to June 2022. Background seismicity is represented with shaded black circles, while the ten considered seismic sequences are colored with different colors. The red star marks the epicenter of the 1980 M 6.9 earthquake. Red triangles indicate the seismic stations of the ISNet network..... 29
- Figure 2.15.** 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..... 31
- Figure 2.16.** Number of detections for each sequence provided by FAST (blue), EQT (orange), EQT + TM (red), manual INFO catalog (pink) and INFO + TM (green). The results for sequence ID 1 (Rocca San Felice) is shown independently in the left panel, since it features the highest number of events..... 32
- Figure 2.17.** Foreshock/aftershock analysis. This figure shows the cumulative percentage of the events in the catalog as a function of the time since the main event 33
- Figure 2.18.** Performance of EQCorrscan and FAST on a one-minute-long waveform featuring near-overlapping events. Green boxes mark the disentangled detections from template matching catalog, while the wide red box marks the single long chain of similar fingerprints provided by FAST, which does not distinguish the overlapping events..... 34
- Figure 2.19.** Left panel: Application of the conservative model of EQT to a one-minute-long waveforms containing multiple events. EQT can correctly discriminate the events, featuring magnitude 1.0, 0.4 and 0.3 respectively. However, it misses an event reported in the template matching catalog, placed around sample 800..... 35
- Figure 2.20.** Comparison of the performance of EQT on events featuring different position of the waveforms within the one-minute-long considered windows. The window on the left panel starts at 2020-07-03 16:16:42 and contains the events in the central part of the window, while the right panel has been extracted from 2020-07-03 16:16:30 and contains the events closer to the right edge of the window. In the right panel, the peak probability scores obtained from the left panel are reported, with red and blue dashed lines marking the former P and S probability scores, respectively..... 36

Figure 2.21. Gutenberg–Richter distribution for FAST (Mc 0.0 and $b = 0.71 \pm 0.05$; upper-left-hand panel), EQT (Mc 0.2 and $b = 0.54 \pm 0.04$; upper-right-hand panel), EQT + TM (Mc -0.3 and $b = 0.71 \pm 0.03$; lower-left-hand panel) and INFO + TM (Mc -0.1 and $b = 0.72 \pm 0.04$; lower-right-hand panel) catalogs.	39
Figure 2.22. 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, coloured circles) for the sequences analysed here.	41
Figure 2.23. Distribution of the DETECT stations (red triangles), organized in the 20 arrays indicated with the corresponding code. Black triangles mark the location of the the Irpinia Near Fault Observatory stations.	42
Figure 2.24. Data availability during DETECT experiment. Drops of the data availability close to the edges are due to the progressive installation and uninstallation of the stations respectively.	43
Figure 2.25. The six subnetworks extracted from the DETECT network. Each subnetwork consists of six arrays, with an overlapping of three arrays with the adjacent subnetwork.	44
Figure 2.26. Median monthly number of earthquakes for each year from January 2009 to September 2021 (colored bars). The mean number of earthquakes for each month, which is marked with a dashed black line, is reported within the brackets, next to the median value.	46
Figure 2.27. Gutenberg- Richter law for the manual INFO catalog from 2089 (green) and for the DETECT catalog (blue). The dashed blue and green vertical lines mark the completeness magnitude of the DETECT and INFO catalogs, respectively.	47
Figure 2.28. Location of the cluster used in this analysis (Palo et al., 2023a), reported as blue dots. The red station marks the position of the SNR3 station, while the dashed black lines mark the fault segments activated during the 1980 M 6.9 earthquake.	49
Figure 2.29. Set of potential repeaters extracted from continuous data for cluster 14 of Palo et al. (2023a) at the SNR3 station. The red waveform marks the master event used for refining the detection list. CC coefficient with the master event is reported in parenthesis for each detection.	50

- Figure 2.30.** Location of the experimental seismic array (blue markers), along with the RSC seismic network (red markers). The black stars mark the M_l 0.1 and 0.0 Conegliano earthquakes, while the red star refers to the M_l 0.4 Vittorio Veneto earthquake. All these events were included in the RSC catalog..... 52
- Figure 2.31.** Vertical component of the velocity records, filtered in [1 -20] Hz, for the Conegliano M_l 0.1 (left panel), and the FAST detections M_l -0.4 and -0.2, in the central and right panels, respectively..... 54
- Figure 2.32.** Panel reporting a typical 7-minute-long window selected for visual inspection by network operators. Each waveform represents the vertical velocity (or acceleration) records for a specific station, which are ordered from North to South direction..... 56
- Figure 2.33.** [1 -20] Hz filtered waveforms for station VDS3, SCL3 and COL3. The event was not visually identified from the representation as in Figure 2.32, but was detected by the automatic detection strategy..... 58
- Figure 2.34.** Distribution of the time residuals between PhaseNet automatic picks and manual refined arrival times for the P (left panel), S (central panel) and combined phases (right panel). Both P and S phases declared by PhaseNet are compatible with the manual arrival times, with a larger dispersion observed for the S phases..... 59
- Figure 2.35.** Automatic and manual identifications of phase arrival times for a M_l 0.4 earthquake, reported as green and blue dashed lines, respectively. While the P arrival time is coherent among the two strategies, S arrival times feature a time residual of 0.375s, with the PhaseNet picks preceding the manual identification..... 60
- Figure 2.36.** Zoom of the bottom panel of Figure 2.35, around the arrival of S wave. In this representation the PhaseNet pick (green dashed line) appears to be more reliable as compared to the manual identification (blue dashed line), correctly identifying the change in amplitude and frequency characteristic for the arrival of the S phase..... 60
- Figure 2.37.** Gutenberg – Richter law extracted from events that occurred between 27th February and 27th August for each of the past years from 2008 to 2021 (shaded dots). The black dots mark the Gutenberg – Richter relation for the PhaseNet enhanced catalog. The magnitude of completeness decreases by almost 0.5 magnitude units..... 61
- Figure 2.38.** Examples of false detections raised by the declaration of arrival times for non-seismic transients 62

Figure 2.39. M_l 1.0 earthquake missed by the automatic detection strategy. Despite the higher magnitude, this event was observed at 4 stations only, with a single identification of the S arrival time, due to the low frequency content of the waveforms	63
Figure 2.40. M_l 0.4 earthquake missed by the automatic detection strategy, due to uncertainty in a pick declaration which was discarded within the association stage	64
Figure 3.1. Distribution of the cross-correlation coefficients resulting from different windows extracted around the P arrival time.....	77
Figure 3.2. Left panel: example of source-specific station P waves travel time corrections, extracted from Figure 1 in Lomax & Savvaidis (2022). SSST corrections are added to the current, P travel-time field for the station to produce updated, SSST corrected travel-times (right).....	80
Figure 3.3. Left - Top trace represents the target waveform while the reported coherence with the target is indicated to the upper right of each waveform. Right - Schematic representation of coherences between the target and each waveform, corresponding stacking weight after cosine taper mapping of coherence, location PDF's forming the stack (color intensity indicates stack weight), and final NLL-coherence location PDF for the target event.	81
Figure 3.4. Distribution of the time residual among the automatic and manual phase arrival times for the ten seismic sequences analysed in Section 2.3.1. Both P and S phase arrival times feature zero-mean residuals, with a slightly larger dispersion for the S phase.....	82
Figure 3.5. INSTANCE (left panel) and IQUIQUE (right panel) datasets, implemented in Seisbench. For estimating the phase arrival times, we used EQT models trained individually on one of the considered datasets.	83
Figure 3.6. Residual distributions for P and S phase arrival times predicted by EQT models trained on STEAD (upper panel), INSTANCE (central panel) and IQUIQUE (lower panel), respectively. The distribution observed for the INSTANCE training dataset is coherent with the STEAD dataset.....	84
Figure 3.7. Smooth velocity model obtained from the interpolation of the Matrullo et al., (2013) P wave velocity model. The blue line represents the model obtained interpolating the velocity at the top depth of the layers in Matrullo et al., (2013), while the green blue line represents the model obtained interpolating the velocity at the mean depth of the layers.....	86

Figure 3.8. Horizontal hypocenter uncertainty (first column), vertical location uncertainty (second column) and RMS (third column) for the obtained location using the 1D layer velocity model (red), the continuous model obtained by interpolating the top of each layer (blue) and the continuous model obtained by interpolating the mean depths of each layer (green). The text reports the median and the MAD values observed for each distribution	86
Figure 3.9. Absolute location for the 10 seismic sequences from the manual INFO catalog. In the left panel, the black segments mark the faults activated during the 1980 M 6.9 earthquake. In the right panel, the seismicity is projected along a vertical plane oriented N40E, (A-A' direction), orthogonal to the strike of the faults represented in the left panel.....	88
Figure 3.10. Location uncertainties from the events in the manual catalogs. The left panel reports the horizontal uncertainties, the central panel reports the vertical uncertainties, while the right panel reports the root-mean-square of the travel time residuals. In each panel, μ indicates the median value of each distribution.....	89
Figure 3.11. Absolute locations for the enhanced catalogs of seismic sequences obtained in Scotto di Uccio et al., (2023). 1130 events (60% of the starting detection catalogs) have been located	90
Figure 3.12. Location uncertainties from the events in the manual catalogs. The left panel reports the horizontal uncertainties, the central panel reports the vertical uncertainties, while the right panel reports the root-mean-square of the travel time residuals... ..	90
Figure 3.13. Distribution of the spatial location uncertainties for common located events in the manual and automatic catalogs. The location uncertainties for the events located using the automatic phase arrival times typically show lower values	91
Figure 3.14. Double difference relocation of the enhanced catalogs of Scotto di Uccio et al., (2023). The representation of the cross-section shows that the seismicity patterns feature clear alignments	92
Figure 3.15. Distribution of the spatial location uncertainties for the double difference relocation of the events in the enhanced catalogs. The spatial scale in this figure is much lower than the analogous representation of the kilometeric-size location uncertainties in Figures 3.12 and 2.13.....	93

- Figure 3.16.** Left panel: Spatio-temporal evolution of the epicenters for the Rocca San Felice seismic sequence (IDX 1), colored according to the occurrence time from the main event. The main event is represented by a black star. Right panel: Cross-section colored according to the occurrence time from the main event94
- Figure 3.17.** Distribution of the P (left panel) and S (right panel) phase arrival times for the 20 deployed arrays. Most of the phase arrival times are observed for the arrays 06, 09 and 11, which are deployed in the central region of the Southern Apennines. A red pixel indicates the presence of the phase arrival time for the considered station, otherwise a blue pixel indicates the absence of a pick.....96
- Figure 3.18.** Absolute location of the earthquakes in the manual catalog from September 2021 to August 2022, obtained by selecting the stations of the ordinary seismic network. Most of the seismicity occurred in the central and southern section of the Irpinia region. The cross-sections on the right panel represent the projection of the hypocenters along vertical planes oriented N40E. For each sector, only hypocenters within 15 km from the vertical planes are considered.....97
- Figure 3.19.** Upper panel: Representation of the epicenters from the absolute location of the earthquakes in the enhanced catalogs. The seismicity appears to occur mainly in the central and southern sector of the area. Lower panels: characteristic cross-section for the area. Absolute locations do not highlight seismicity alignments with activated structures.....98
- Figure 3.20.** Time residuals among the automatic and phase arrival times refined through hierarchical clustering and cross-correlation alignment. The three distributions suggest that the automatic identifications are coherent but slight corrections are necessary to appreciate the small-scale spatial distance among the events.....99
- Figure 3.21.** Double difference relocation of the enhanced catalog for the DETECT experiment, projected along the A-A' vertical plane. In the left panel, the epicentral representation confirms higher seismicity in the central and Southern sector of the area, while strongly marking the lack of seismicity in the southeastern edge, within the black dashed circle. The cross-sections in the right panel reports the presence of sparse shallow seismicity within 5 km of depth, while the deeper seismicity tends to appear more clustered in space and time. The cluster within the cyan circle represents a swarm-like sequence of 30 earthquakes, associated with the occurrence of a M_l 2.1 event in Dentecane (AV), located outside the array deployment.....100

- Figure 3.22.** Clustering analysis using DBSCAN for identifying seismicity clusters. We declared 22 seismic clusters populated by at least 10 earthquakes ... 101
- Figure 3.23.** Cross-sections along vertical planes oriented N40E for the Northern, Central and Southern region of the Irpinia, as in Figure 2.18 and Figure 2.19. The left panels refer to the period from September 1st 2021 to February 8th 2022, while the right panels contain the event from this latter date up to the end of the experiment..... 102
- Figure 3.24.** Schematic representation of the step fault identified from the analysis of the DETECT seismicity 103
- Figure 3.25.** Left panel: Ground deformation recorded at the RITE GNSS station from 2014 up to March 2024. The ground uplift reached the value of 100 cm. Vertical dashed lines mark the increase in the uplift rate, experienced in mid-2017, mid-2020 and end of 2022. Right panel: Temporal distribution of the number of earthquakes (black curve) and their maximum magnitude (red curve) in the same time interval. Both curves feature a marked increase, leading to seismic rates of several hundreds of events per month. Figure from Scotto di Uccio et al. (2024b).... 104
- Figure 3.26.** Shaded relief map of Campi Flegrei with simplified caldera boundaries (modified after Natale et al., 2022; Vitale & Isaia, 2014), showing the 2014–2024 seismicity recorded by INGV-Osservatorio Vesuviano seismic network, colored by hypocentral depth and magnitude-scaled. The upper-right inset shows the vertical deformation pattern..... 105
- Figure 3.27.** Left panel: Depth distribution for the located earthquakes in the INGV-OV catalog from 01/01/2014 to 14/03/2024. Right panel: Duration magnitude (M_d) distribution for the earthquakes that occurred within the same time interval 106
- Figure 3.28.** Cluster of similar earthquakes recorded at station COLB. The coherency is evaluated in terms of the cross-correlation coefficient with the master event, reported in the upper panel of the figure 108
- Figure 3.29.** Relocated NLL-SC seismicity 2014–2024 ($M_d \geq -0.8$, 8133 events) represented as circles color-coded and scaled according to M_d 109
- Figure 3.30.** Histograms of location uncertainties measured by vertical (top panel) and horizontal (middle panel) errors in km and RMS of travel time residuals in s (bottom panel). We found a dramatic reduction in the location errors due to the waveform coherence step of the location procedure..... 110

Figure 3.31. Histograms of location uncertainties from this study and from De Siena et al. (2024), for the earthquakes that occurred within 2014 and 2019, located using NNLoc and the 3D velocity model of Battaglia et al. (2008)111

Figure 3.32. Spatiotemporal evolution of the seismicity in periods 2014–2017, 2018–2019, 2020–2021 and 2022–2024. The epicenters, scaled according to M_d , are reported for earthquakes featuring $M_d \geq -0.8$ and ellipsoid major axis ≤ 2.0 km.....112

Figure 3.33. Simplified structural map showing the relationship between the epicentral distribution of relocated seismicity in the 2022–2024 period with the elliptical pattern and the main volcano-tectonic structures. Black lines map known geological features. Focal mechanisms solutions for selected 2023 $M_d > 3$ events are shown (ID according to Scotto di Uccio et al., 2024b) with their color coded by depth. The cross-sections for each vertical planes (panels b, c and d) report earthquakes occurred within the corresponding yellow box (s1, s2 and s3).113

Figure 4.1. Schematic representation of a fault, with area Σ and normal vector n . The point P is indicated, in a reference frame along the fault, by the coordinates (ξ_1, ξ_2) and the dislocation occurs when the displacement on the Σ^+ side differs from the displacement on the Σ^- side. Figure extracted from Zollo & Emolo (2011)...119

Figure 4.2. The real and complex rupture process can be simplified, under specific conditions, with an average dislocation model, which is analogous to the effect of a system of a double-couple of forces acting along and perpendicular to the fault plane.120

Figure 4.3. Theoretical displacement spectrum for an earthquake. At low frequency, the spectrum presents a plateau level, while at higher frequencies it starts decaying121

Figure 4.4. Seismic moment as a function of the fault length for different ranges of magnitudes. Continuous black lines mark the theoretical constant stress drop scaling of $\Delta\sigma=1$ MPa, 10 MPa. Yellow and black dots refer to the $M > 6$ earthquakes, while the grey cloud refers to the moderate magnitude seismicity analysed in Madariaga et al. (1991). Source parameters for most of the earthquakes fall within the continuous black lines, supporting the hypothesis of constant stress drop scaling (figure extracted from Zollo & Emolo 2011)125

Figure 4.5. Slip as a function of the distance, in the model of Burridge & Halliday (1971).....	127
Figure 4.6. Static shear stress from the slip model of Figure 4.5.....	129
Figure 4.7. Complete stress solution, computed for $\theta=0$ and $\theta=\pi/2$, compared to the approximate solution of Figure 4.6.....	130
Figure 4.8. Definition of the fault orientation parameters (strike Φ_s and dip δ angles) and the rake angle λ indicating the orientation of the slip vector along the fault.....	131
Figure 4.9. Schematic representations of different fault types: normal, strike-slip and inverse fault	132
Figure 4.10. Definition of the focal sphere and local polar coordinates in a reference frame centred in the earthquake hypocenter.....	133
Figure 4.11. Cartesian and spherical reference frames for the analysis of the radiation pattern associated with a shear dislocation occurring along the fault surface.....	134
Figure 4.12. Left panel: P wave radiation pattern for a vertical strike-slip fault. Grey arrows indicate the slip direction respect to the fault plane. Red arrows indicate the P wave amplitude for different θ angles. Right panel: First motion amplitude for the P wave respect to the fault and auxiliary planes.....	135
Figure 4.13. Beachball representation for the focal mechanisms associated with the main fault types. the colored quadrants indicate the volume of the medium around the source which undergoes to extension, whereas the white quadrants refer to the segments which undergo compression	137
Figure 4.14. Anderson's classification scheme of stress in the Earth's crust (left) and corresponding faulting regimes (right). The focal mechanisms with the P and T axes are shown in the lower-hemisphere equal-area projection	138
Figure 4.15. Mohr's circle diagram. Quantities σ_n and τ are the normal and shear stresses along a fault, σ_1 , σ_2 and σ_3 are the principal stresses. All permissible values of σ_n and τ acting on a fault must lie in the shaded area of the diagram	140
Figure 4.16. Mohr-Coulomb failure criterion. The red area shows all possible orientations of fault planes which satisfy the Mohr-Coulomb failure criterion. The blue dot with shear and normal stresses τ_c and σ_c denotes the principal fault plane which is optimally oriented with respect to stress, and C denotes the cohesion.....	141

- Figure 4.17.** Focal mechanisms associated with unstable fault planes. Randomly distributed fault planes inside the unstable area of the Mohr’s diagram (left), corresponding nodal lines (in the middle) and the P/T axes (right). The P (pressure) axes are marked by the red circles, the T (tension) axes by the blue crosses.....141
- Figure 4.18.** Left panel: Q_{EGF} estimation from linear fit (red line) of the logarithm of displacement event spectrum (blue dots) as a function of the linear frequency, for a $M_l = 0.41$ earthquake in the Rocca San Felice seismic sequence. Noise spectrum is reported as black dots. Right panel: histogram of the Q_{EGF} for the events $M_l < 1$ in the Rocca San Felice sequence (IDX 1) at NSC3 station.....145
- Figure 4.19.** Spectral inversion of a $M_l = 2.2$ earthquake at SFL3 station ($r \sim 40$ km from the main event, sequence IDX 1) using different attenuation factors for the station. Left panel: inversion using $Q = 230$. Right panel: inversion using $Q = 300$. In both panels, black vertical lines mark the frequency band in which the spectral fit is performed. The pink vertical lines indicate the estimated corner frequency. Right panel inset: average RMSE for the events at SFL3 for the sequence IDX 1, using different Q_s values in the inversion147
- Figure 4.20.** Spectral inversion of a $M_l = 2.8$ earthquake using different attenuation factors for the stations. Left panel: inversion using Q_{EGF} at NSC3 station. Central panel: inversion using Q_{LOC} at SCL3 station. Right panel: inversion using Q_{REG} at VDS3.....148
- Figure 4.21.** Synthetic tests for an event of magnitude $M_w 1.5$, $Q=230$ and $f_c = 15$ Hz - Panels a) and b) – and $f_c = 25$ Hz - Panels c) and d). We inverted the spectra assuming $Q=150$ – Panels a) and c) - and $Q=300$ – Panels b) and d). We found that the moment magnitude is well resolved, while the corner frequency estimate is strongly affected by the attenuation and limited bandwidth. In the table we report the retrieved values149
- Figure 4.22.** Synthetic tests for an event of magnitude $M_w 2.0$, $Q=230$ and $f_c=8$ Hz. We inverted the spectra assuming $Q=150$ – Panel a) $Q=300$ – Panel b). We found that both moment magnitude and corner frequency are well resolved. In the table we report the retrieved values150
- Figure 4.23.** $M_w - M_l$ distribution, with the 1:1 scaling (dashed red line). For $M_l < 2$ earthquakes, we observed $M_w = 0.89 (\pm 0.03) + 0.62 (\pm 0.02)M_l$ (green dashed line).151

- Figure 4.24.** Distribution of source parameters for $M_w > 2$, colored according to sequence IDX, with theoretical constant stress drops scaling of 0.1 MPa, 1 MPa and 10 MPa (red solid lines)..... 152
- Figure 4.25.** Left panel: Stress release model for the sequence IDX 1. Right panel: Histogram of the nearest neighbor distance for the relocated events of the sequence IDX 1. 154
- Figure 4.26.** Stress release model for the considered sequences. Small events are mainly concentrated in or around the area affected by stress changes due to the main events in the sequences, suggesting stress release as trigger mechanism..... 155
- Figure 4.27.** Comparison of the stress released by an earthquake with Brune radius and stress drop of 100 m and 1 MPa respectively, using Brune and Madariaga models. The solutions are represented along the directions parallel ($\theta=0$) and orthogonal ($\theta=\pi/2$) to the slip. The solution from the Brune model is amplified by a factor of 3. We found that the size of the region experiencing a stress increase after the earthquake is of the same order of magnitude for both models..... 157
- Figure 4.28.** Comparison of stress changes using Brune (left panel) and Madariaga models (right panel) with the same scale for the stress for the Rocca San Felice sequence. At the scale of the fault, the two images appear very similar, with the main differences around the hypocenters of the events, within the source radius..... 158
- Figure 4.29.** Zooming on the Figure 4.28, representing now the left cluster of the Rocca San Felice sequence. Left panel - Stress change with the Brune model; Central Panel - Stress change with the Madariaga model represented with the same scale as the Brune model; Right Panel - Stress change with the Madariaga model, represented with a wider scale that helps to identify the maximum amplitudes..... 158
- Figure 4.30.** Stress changes obtained assuming self-similarity (left panel) or the scaling model $M_0 \propto f_c^{-(3+\epsilon)}$, where the $\epsilon = -0.39$ according to Picozzi et al., (2022a) for the Rocca San Felice sequence. The two maps appear very similar, with differences around the hypocenters of small events..... 159
- Figure 4.31.** Schematic representation of seismicity lineations, oriented along the dip direction, which for a normal fault coincides with the slip direction... 160
- Figure 4.32.** Upper panel: map of the GPS stations near the Rocca San Felice seismic sequence and associated displacements. The vectors show the

horizontal displacements with 1-sigma error ellipses, from the static offset during the seismic sequence. Three different synthetic scenarios were evaluated assuming increasing slip (25, 50, 75 mm), with geometry and kinematics from Festa et al., 2021. The surface displacements produced by a slip of 50 mm on the fault drawn with purple lines is shown in the map (purple arrows). Lower panel. GPS position time series and average positions (white dashed lines) before and after the seismic sequence, with 1-sigma error shown with grey shading. Calculated displacements are shown (dashed lines) for the 25mm (green), 50mm (purple) and 75 mm (red) dislocation scenarios162

Figure 4.33. Focal mechanisms for the main earthquakes occurred in the Southern Apennines. The destructive 1980 M 6.9 and the 1996 M 5.1 Irpinia earthquakes feature pure normal-faulting mechanism, while the events of the 1990-1991 Potenza seismic sequence mainly report east-west oriented mechanisms.....164

Figure 4.34. Magnitude distribution of the events in the manual INFO catalog from September 2021 to August 2022.....165

Figure 4.35. Focal mechanisms for the events in the manual INFO catalog (6 events). Among the 6 earthquakes for which the focal mechanism was estimated, 4 events report a normal focal mechanism, coherent with the main fault segments in the area, while the remaining earthquakes reported a less reliable strike-slip and inverse faulting.....166

Figure 4.36. Instance of seismograms related to polarity U (upper-left panel), D (upper-right panel) and K (lower-left panel), respectively. Lower-right panel: distribution of the signal-to-noise ratio associated with each of the three polarities classes167

Figure 4.37. Precision (left column) and recall (right column) curves from the model training as a function of the SNR (upper panels), distance (central panels) and magnitude (bottom panels), respectively.....168

Figure 4.38. Examples of wrong polarity provided by the deep learning model. Despite the seismograms featuring a clear positive first motion amplitude for both stations, the model assigns polarity D for both instances.....170

Figure 4.39. Precision (left column) and recall (right column) curves from the model trained by randomly shifting the position of the P wave onset within the seismogram, as a function of the SNR (upper panels), distance (central panels) and magnitude (bottom panels), respectively.....171

- Figure 4.40.** Application of the model trained by fixing the position of the P wave onset in the centre of the seismogram (left panel) and the model trained allowing a random shift of the onset in $[-0.5, 0.5]$ s in the seismogram (right panel) to an instance featuring a delay in the identification of the P phase arrival time. While the former version of the model misclassified the polarity, the novel version of the model is able to handle the earlier onset of the P wave, correctly associating the class label with a high probability score. 172
- Figure 4.41.** Representation of the classification outputs for three stations reporting different impulsivity of the P wave onset (decreasing from the top to the bottom). The higher the impulsivity of the P wave onset, the higher is the probability score associated with the class..... 174
- Figure 4.42.** Histograms of the slopes (left panel) and intercepts (right panel) of the linear regression between the logarithm of the signal-to-noise ratio and the probability scores predicted by the classifier for the assigned polarities for each event. Coherence in the slope and the intercept suggests a correlation between the two parameters 175
- Figure 4.43.** Focal mechanism estimation of a M_l 1.7 earthquake, featuring 82 polarity estimates. Focal mechanism resulted in a pure normal faulting .. 176
- Figure 4.44.** Focal mechanism selecting only polarities featuring probability scores higher than 0.70, 0.80 and 0.90, respectively..... 177
- Figure 4.45.** Seismogram related to the station S0207, whose polarity is discarded when considering probabilities higher than 0.80. However, this station is expected to appear on the focal sphere nearby a nodal plane, better constraining the mentioned plane..... 178
- Figure 4.46.** Distribution of the number of polarities used within the inversion of focal mechanisms for earthquakes in three magnitude classes: $M_l \leq 0.5$, $0.5 < M_l \leq 1$ and $M_l > 1$. Within each subplot, the median number of polarities is reported 179
- Figure 4.47.** Focal mechanisms obtained from the inversion of the polarities from the ordinary seismic network and the dense array deployments. The size of the beach balls is proportional to the magnitude of the earthquakes. The 223 obtained fault planes mainly show pure normal faulting mechanisms, with a slight percentage of strike-slip component and inverse faulting..... 180
- Figure 4.48.** Rose diagram reporting the strike (left panel), dip (central panel) and rake (right panel) angles resulting from the 223 extracted focal mechanisms.

These histograms support the observations of a main normal faulting environment, with a slight component of strike-slip.181

Figure 4.49. Strike (left panel), dip (central panel) and rake (right panel) uncertainties for the 223 estimated focal mechanisms.182

Figure 4.50. Representation of focal mechanisms featuring an inverse faulting. Red stars mark the corresponding epicentres182

Figure 4.51. Focal mechanism for three clustered earthquakes that occurred within a span of 20 days in the Southern Sector of the region (black box in Figure 4.50). The estimated inverse focal mechanisms appear similar for the considered earthquakes.....183

Figure 4.52. Comparison of the estimated focal mechanisms for an earthquake featuring inverse faulting resulting from the fully automatic procedure (left panel) and manual identification of phase arrival times and P wave polarities (right panel).....184

Figure 4.53. Left panel: Stress field inversion from De Matteis et al. (2012). Right panel: Stress field retrieved from the inversion of the 223 focal mechanisms estimated in Section 4.3.3.....185

Figure 4.54. Distribution of the R value obtained from each bootstrap instance using the catalog of 223 focal mechanisms187

Figure 4.55. Dataset splitting for the inversion of the stress field in the Northern (blue), Central (green) and Southern sector.....187

Figure 4.56. Stress field inversion from focal mechanisms belonging to the Northern (left panel), Central (central panel) and Southern (right panel) sector.....188

Figure 4.57. Histogram of the R ratio from bootstrapping resample of the focal mechanisms for the Northern (left panel), Central (central panel) and Southern (right panel) sector.....189

Figure 4.58. Stress field inversion for the events occurred before (left panel) and after (right panel) February 1st 2022.189

List of Tables

Table 2.1. Number of detections in the manual catalogs and information regarding the main event for each of the analysed seismic sequences ...	30
Table 2.2. M_c and b-value estimation from the enhanced catalogs of each sequence.....	39
Table 2.3. Monthly detections for the manual INFO catalog, the EQT and EQT + TM catalogs for DETECT	45
Table 3.1. Weighting scheme proposed in De Landro et al. (2022), for assigning the importance of the phase arrival time within the location inversion.....	74
Table 3.2. Table proposed by Mousavi et al. (2020) for the converting the probability of EQT detections in weights for earthquake location.	74
Table 4.1. Summary of the classification results for the two versions of the trained model. The model trained randomizing the position of the P wave onset in the seismogram provided improved polarity classification performances as compared to the former version.....	172
Table 4.2. Proposed weighted scheme based on the conversion of class probability scores into discrete weights.....	175
Table 4.3. Plunge and azimuth angles for σ_1 , σ_2 , and σ_3 obtained from the inversion of focal mechanisms evaluated in Section 4.3.3. Red values refer to the results of De Matteis et al. (2012). Bracketed values refer to the 95% confidence interval from bootstrap resampling	186

List of Symbols

The following symbols are used within the thesis

α, v_p	P wave velocity
A_{ref}	Reference amplitude for amplitude ratio
β, v_s	S wave velocity
δ	Dip angle
$\Delta\sigma$	Stress drop
E	Seismic energy
f_c	Corner frequency
γ	v_p/v_s ratio
i_ξ	Incidence angle
λ	Rake angle
μ	Rigidity modulus / mean value
M_d	Duration magnitude
M_l	Local magnitude
$M_{l_{ref}}$	Reference local magnitude
M_w	Moment magnitude
M_0	Seismic moment
\mathbf{n}	Normal vector
Φ	Strike angle
Q	Quality factor

Q_{EGF}	Quality factor estimated using EGF
Q_{LOC}	Quality factor estimated minimizing inversion residuals
Q_{REG}	Regional value of the quality factor
s	Slip
σ	Standard deviation
σ	Stress tensor
R_P, R_S	Radiation pattern for the P/S wave
τ	Traction
T_0	Origin time of the earthquake
TT_P, TT_S	Travel time for the P and S waves



Introduction

Earthquake characterization is a crucial task for understanding the nucleation process leading to the occurrence of seismic events and inferring the mechanical state of the faults. The characterization of seismic events consists of: i) identification of the earthquakes within continuous records of ground motion quantities, that are typically dominated by ambient noise and non-seismic anthropogenic sources (earthquake detection); ii) determination of the nucleation point on the fault plane (hypocentral location); iii) estimation of the source properties accounting the rupture size, event magnitude and released stress (source characterization). Accurate analysis of major earthquakes has revealed the complexity of the seismogenic structures hosting the main events (Bernard & Zollo, 1989; Chiaraluce et al., 2017; Ross et al., 2019; Shelly, 2020). However, the global magnitude distribution reports a few high magnitude earthquakes in time, making difficult a direct access to their preparatory and nucleation phases, whose comprehension remains an open question in earthquake seismology.

On the other hand, lower magnitude events occur within the interseismic cycle between successive major earthquakes, typically within the same structures which can accommodate larger magnitude events. According to the global seismicity occurrence, we expect an increase of a factor of ten in the number of earthquakes when lowering the considered magnitude by one unit (Gutenberg & Richter, 1944). Therefore, a detailed analysis of low magnitude events could provide an enriched view of the small-scale rupture processes. However, microseismic events feature a low signal-to-noise ratio on the seismic records, being typically buried within the ambient noise level. Therefore, they show prominent amplitudes above the noise level just at the stations closest to the epicenter, strongly limiting the capability of being identified through visual inspection of the data or using standard detection techniques (Allen, 1978; Baillard et al., 2014). Seismic networks deployed close to seismogenic sources and temporary deployment of dense array constellations have been shown to record waveforms featuring improved

signal-to-noise ratio for low magnitude earthquakes (Ben-Zion et al., 2015; Chiaraluce et al., 2022a). However, standard algorithms are not able to handle the massive amount of data collected by dense monitoring systems, leading to the necessity of developing novel approaches (Yoon et al., 2015). Advanced techniques have been developed for improving the resolution of earthquake detection, using machine learning models (Ross et al., 2018; Zhu & Beroza, 2019; Mousavi et al., 2020) or exploiting similarity among seismic waveforms (Chamberlain et al., 2018; Vuan et al., 2018), leading to the generation of enhanced seismic catalogs with a ten-fold increase in the number of identified earthquakes (Tan et al., 2021). Effectively, deep-learning approaches have entered almost every subfield of seismology, for which they have shown the ability to outperform classical approaches, often dramatically, for seismological tasks such as denoising, earthquake detection, phase picking, seismic image processing, also strongly reducing the computational time (Mousavi & Beroza, 2022; Messuti et al., 2024). Coupling the improved catalogs with accurate techniques for the determination of the hypocentral coordinates (Waldhauser & Ellsworth, 2000; Lomax & Savvaidis, 2022) has revealed fine-scale seismic structures identified by the occurrence of low magnitude events, moving the observation size from the kilometric to decametric/hundred of meter scale (Chiaraluce et al., 2017, 2022b; Ross et al., 2019). In this direction the analysis of seismic sequences, characterized by a higher seismicity rate as compared to the background and identifying earthquakes occurring close in space and time, is a precious framework for investigating local seismic features (Ellsworth & Beroza, 1998; Festa et al., 2021; Sukan et al., 2023). Therefore, seismic sequences present as a tool for understanding the efficiency of advanced state-of-the-art detection techniques, which require comparison across methods and in-depth analysis of the differences between the resulting catalogs.

Within this thesis, we mainly focus on the characterization of the microseismic events occurred in the Southern Apennines, hosting the complex Irpinia normal system struck by the 1980 M 6.9 earthquake. This destructive event occurred along NW-SE-striking faults and was characterized by three main episodes within a few tens of seconds, causing about 3000 fatalities and severe damage (Bernard & Zollo, 1989). Moreover, this area is considered one of the regions with the highest seismic hazard in Italy (Stucchi et al., 2011). Since 2007, the area has been continuously monitored by the Irpinia Near Fault Observatory (INFO) including 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. Seismicity in Irpinia typically features a low seismic rate, consisting of few events per day, and moderate depth values, mainly ranging between 8 and 15 km, covering a local magnitude range between $M_l -0.4$ and $M_l 3.7$. The background seismicity that occurred during recent years appears to be distributed within a volume bounded by the main faults of the 1980 event without clearly identified fault segments at depth (De Landro et al., 2015), and sometimes occurring in sequences with events of maximum magnitude $M_l \sim 3.0$ (Stabile et al., 2012; Festa et al., 2021), lasting few days. Within this thesis, we also attempted to test these techniques to different environments, such as the Campi Flegrei area, where a seismic crisis associated with volcanic bradyseism is ongoing (Bianco et al., 2022), and Collalto region in Northern Italy, a gas storage site featuring induced seismicity (Priolo et al., 2015).

The thesis is structured as follows. In Chapter 2 we introduce earthquake detection, describing the challenges in the identification of seismic events at the noise level and the advanced state-of-the-art detection techniques which are used for generating enhanced catalogs. We apply the strategies in different tectonic contexts, presenting semi-automatic approaches for long-term seismic monitoring and statistically interpreting the obtained enhanced detection lists. In Chapter 3 we focus on the determination of the hypocentral coordinates using innovative location algorithms, discussing their impact in the resolution of small-scale seismogenic structures as depicted from the relocation of high-resolution catalogs. In Chapter 4 we tackle seismic source characterization, revealing rupture size and stress release model for earthquakes in enhanced catalogs, estimating focal mechanisms for microseismic events and inverting them for analyzing spatiotemporal characteristics of the acting stress field. In Chapter 5 we summarize and discuss the results obtained in the different analyses, identifying interesting and stimulating perspectives for the considered topics.

Chapter 2

Advanced techniques for earthquake detection

2.1 Introduction

Earthquake detection consists in the identification of seismic events within continuous ground motion records (seismograms), which are generally mostly populated by noise and non-seismic transients, as wind induced oscillations and anthropogenic sources. Earthquake rates vary according to the seismogenic settings of the areas, typically ranging from few events per day, as it is observed for the background seismicity in the Southern Apennines, (<http://isnet-bulletin.fisica.unina.it/cgi-bin/isnet-events/isnet.cgi>), to thousands of events per day, as reported from seismic catalogs produced for the Mw 6.0 L'Aquila (Tan et al., 2021) and Mw 7.8 Kahramanmaraş earthquakes (Becker et al., 2024). However, the short duration of earthquake signals on seismic waveforms makes detection a very challenging task, especially as the magnitude of the events decreases and the amplitude of the signal approaches the noise level.

Seismicity characterization in terms of determination of the source position, earthquake size and geometry allows to investigate and infer the mechanical properties of the faults where seismicity occurs, providing crucial information about earthquake nucleation and triggering processes (Abercrombie & Leary, 1993; Chiaraluce et al., 2017; Festa et al., 2021; Waldhauser et al., 2021). The higher the number of detected earthquakes, the higher the amount of information that can be extracted from their characterization. Locating low magnitude events, that typically follow the occurrence of a larger magnitude earthquake, can highlight the activated faults segments (Wiens et al., 1994; Ross et al., 2019; Michele et al., 2020). Information on earthquake size and geometry can be inferred from the determination of the source parameters (De Matteis et al., 2012; Martinez-Garzon et al., 2014; Supino et al., 2019). Following this direction, the identification of seismic events represents the first step for a comprehensive earthquake characterization, aiming to provide robust and “complete” catalogs of earthquakes to be investigated.

Moreover, seismic catalogs can be further inspected for extracting statistical properties. The most striking application concerns the determination of the regression parameters of the Gutenberg-Richter law (Gutenberg & Richter, 1944), relating the number and the magnitude of the events. It is a universal observation that, in a target area, the cumulative frequency-magnitude distribution of the earthquakes is not random, but appears to follow a specific trend, known as the Gutenberg-Richter law (G-R). Figure 2.1 reports the magnitude-time distribution (panel a) and the frequency-magnitude distribution (panel b) for the 2009 L'Aquila seismic sequence (De Santis et al., 2011).

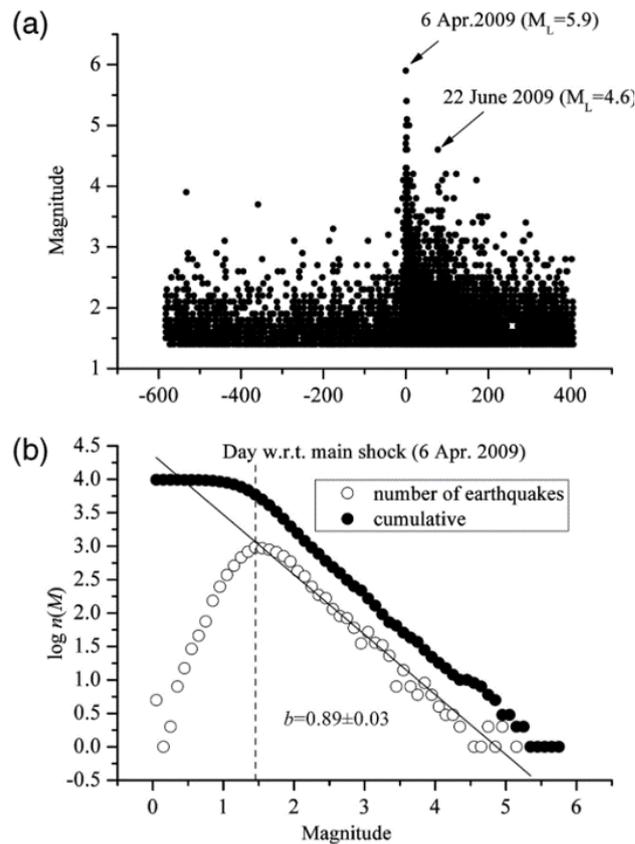


Figure 2.1) Panel (a): Temporal evolution of earthquake magnitudes during the L'Aquila seismic sequence. Panel (b): Number of events in each magnitude bin (white circles) and cumulative frequency-magnitude (black circles) distributions of events for all seismic events. Figure extracted from De Santis et al. (2011).

The Gutenberg-Richter law can be expressed as

$$\log_{10} N(M) = a - bM \quad (2.1)$$

where $N(M)$ is the number of earthquakes with magnitude greater than or equal to M in a given region and in a fixed time interval; a and b are positive parameters indicating the level and the type of seismicity in the region of concern. The slope of the Gutenberg – Richter law, often indicated as the *b-value*, has been claimed from several studies to be an indicator of the stress state of faults (Beroza et al., 2021; Picozzi et al., 2022a), and was analysed both in tectonic (Festa et al., 2021) and volcanic areas (Tramelli et al., 2021; 2024). As an example, Picozzi et al. (2022a) showed that the *b-value* distribution in Southern Apennines is not uniform and hypothesized that 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. Moreover, despite being still debated within the seismological community, monitoring the time variation of the *b-value* has been proposed to discriminate whether an ongoing sequence represents a decaying aftershock sequence or foreshocks to an upcoming large event (Gulia & Wiemer, 2019).

The reliability of the estimation of statistical parameters from a seismic catalog is strongly affected by their content, expressed in terms of the magnitude of completeness. This latter quantity is defined as the magnitude above which all the events are contained within the catalog. Some authors attribute biases in the estimation of the *b-value* to local incompleteness in the detection, driving an underestimation of the slope of the Gutenberg-Richter law (Herrmann & Marzocchi, 2021; Mancini et al., 2022). Thus, it appears to be crucial to deal with robust versions of seismic catalogs.

The content of the seismic catalogs features a twofold dependency from the density of the monitoring seismic networks and the adopted detection technique. Former versions of earthquake catalogs were grounded on visual inspection of seismic waveforms recorded by sparse seismic networks, leading to a high magnitude threshold in event identification. The deployment of increasingly dense networks (Iannaccone et al., 2010; Ben-Zion et al., 2015; Poiata et al., 2016), reaching even kilometeric or sub-kilometeric interstation distances, along with the integration of advanced automatic detection techniques (Yoon et al., 2015; Vuan et al., 2018; Zhu & Beroza, 2019) has moved the magnitude of the detected earthquakes to seismic signals at, or even below, the noise level of the seismic waveforms.

The main branches of the automatic detection techniques can be grouped in energy-based, similarity-based and deep-learning algorithms. 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), the latter window characterizing the average noise level. Earthquake detection is declared when the STA/LTA ration exceeds an imposed threshold. While this approach

is highly flexible in detecting impulsive arrivals, it might fail in identifying events with amplitude near the noise level or closely spaced in time, such that their arrivals overlap. In Figure 2.2 we report the application of the STA/LTA technique on the vertical component of the velocity records of a M 2.9 earthquake occurred in Southern Apennines, for two seismic stations (SCL3, left panel, and AND3, right panel), at hypocentral distances of ~ 5 km and ~ 40 km, respectively (<http://isnet-bulletin.fisica.unina.it/cgi-bin/isnet-events/event-card.cgi?id=17713r>). In this example we used a STA and LTA windows of 1 s and 10 s, respectively, filtering the seismic signals in the [1 – 20] Hz band.

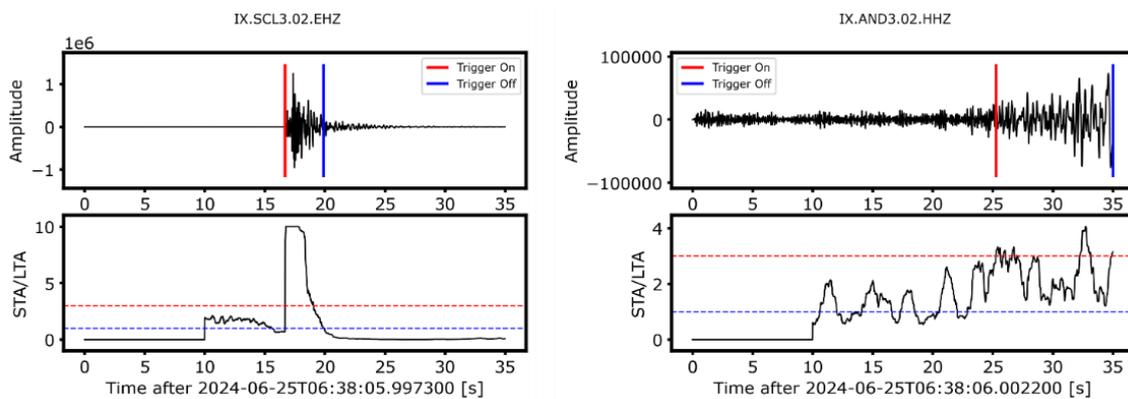


Figure 2.2) Application of the STA/LTA technique on a M2.9 earthquake occurred in Southern Apennines. Left panel: Application of the STA/LTA technique for the closest station (SCL3, hypocentral distance of ~ 5 km). The red and blue bars in the upper panels mark the onset and the offset of the trigger, respectively. The impulsive arrival at the station produces a sharp variation in the STA/LTA function, leading to the declaration of the event. Right panel: Application of the STA/LTA technique for a more distant station (AND3, hypocentral distance of ~ 40 km). The emergent signal is barely above the noise level, producing a maximum value of the STA/LTA window close to the detection threshold.

While for the closest station (SCL3) the arrival is impulsive, producing a sharp variation in the STA/LTA function and leading to the detection of the event, the application of the same strategy on a more distant station (AND3) provides a maximum value of the STA/LTA window close to the detection threshold, given the emergent characteristic of the seismic signal, close to the noise level. Moreover, since this technique only accounts for the energy content of the considered windows, it is not able to discriminate whether the energy change is driven by a real earthquake or a high-amplitude anthropogenic source, eventually leading to a high number of false detections when applied in noisy and harsh environments. 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; Grigoli et al., 2016).

On the other hand, similarity-based detectors exploit the expected similarity among the waveforms for earthquakes occurring close in space, when recorded at the same station. The similarity among the waveforms is often measured in terms of the normalized cross-correlation coefficient (CC), that can be expressed as:

$$CC(\vec{x}, \vec{y}) = \frac{\sum_{i=1}^m x_i y_i}{\sqrt{\sum_{i=1}^m x_i x_i} \sqrt{\sum_{i=1}^m y_i y_i}} \quad (2.2)$$

where \vec{x} and \vec{y} represent the time domain earthquake waveforms and m refers to the number of samples of the signals. According to this definition, the CC coefficient ranges from -1 (perfect anticorrelation) to 1 (perfect correlation). Similarity-based detectors can be further divided into two main branches, known as template-matching search and autocorrelation. The substantial difference among the two branches is that template matching search requires an a-priori known set of templates, that will be used for finding similar events in the continuous records. Thus, template matching is often referred to as matched-filter detection (many-vs-all). In this framework, the similarity of the continuous data is tested against each template, and a detection is declared whether the CC coefficient with at least one template event overcomes an imposed similarity threshold. Before performing the template matching search, pre-processing operations, such as downsampling and filtering in narrow frequency bands are typically applied to continuous records and template waveforms aiming to increase, and eventually to emerge, the waveform similarity for events featuring an amplitude close to the noise level (Vuan et al., 2018; Ross et al., 2019; Sukan et al., 2019; Martinez-Garzon et al., 2023). The success of the technique in identifying similar earthquakes in the continuous records strongly relates to the content of the starting template catalog. It is straightforward to relate the capability of recognizing similar events to the number of available templates. However, since the similarity of the earthquake records has been proved to quickly decrease as a function of the inter-event distance (Schaff & Waldhauser, 2005; Michele et al., 2020), for performing a comprehensive template-matching analysis of a seismic region, it is required to have a large set of templates, also distributed in space. As an example, in Figure 2.3 we report the distribution of the normalized CC coefficients as a function of the interevent distances, extracted from Michele et al. (2020).

Moreover, the frequency content of the seismic waveforms, and thus the shape in the time-amplitude domain in which the cross-correlation is evaluated, is a function of the magnitude. Higher-magnitude events feature lower frequency content, while lower magnitude events, which are typically uncataloged and represent the target for the template matching detection, are characterized by a substantially higher frequency content, although limited by anelastic attenuation due to the wave propagation through the media. Thus, it is more likely to find the similarity among earthquakes characterized

by similar magnitudes, beyond that having close focal parameters (hypo-center locations and focal mechanisms), further emphasizing the need of complete catalogs for an appropriate template-matching analysis. Despite these apparent limitations, template matching is a versatile and powerful technique that has found undetected events in a wide range of seismicity studies: uncataloged low-magnitude earthquakes (Schaff & Waldhauser, 2010), foreshocks (Kato & Nakagawa, 2014), aftershocks (Peng & Zhao, 2009), triggered earthquakes (Meng et al., 2013) and earthquake swarms (Shelly et al., 2013).

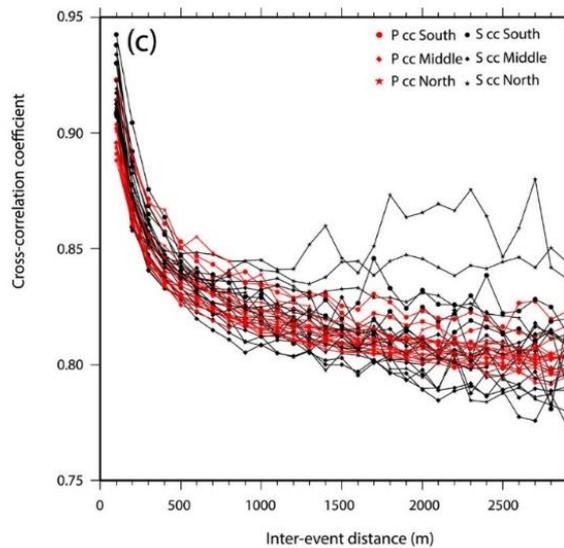


Figure 2.3) Distribution of normalized cross-correlation values versus interevent separation distance for P (red) and S wave (black), extracted from Michele et al. (2020).

Autocorrelation, on the other hand, does not require an a-priori known set of earthquakes. Rather, it assumes each portion of the continuous data as a candidate template and tests it against all the records (all-vs-all). Despite being particularly promising in frameworks featuring low seismicity rates, where the number of templates might be insufficient, autocorrelation has the major disadvantage of being computationally intensive and ultimately infeasible for detecting earthquakes in massive continuous data sets. The $O(N^2)$ algorithm complexity required for comparing all the possible couples from N windows extracted from the continuous data leads to a significant amount of redundant work, because most pairs of windows are uncorrelated and not of interest for detection. Indeed, as discussed in the previous paragraphs, seismic waveforms are typically dominated by noise and non-seismic transients. In Figure 2.4, we report the runtime scaling of autocorrelation as a function of the time extent of the analyzed data, reported as a violet line (figure extracted from Yoon et al., 2015). Autocorrelation suits in detecting

similar earthquakes in a few hours of continuous data (Brown et al., 2008, Yoon et al., 2015). The algorithm complexity limits the application on days, weeks, months, or even years of continuous seismic, especially for dense networks, requiring large-scale computational resources (Yoon et al., 2015). However, autocorrelation has been successfully applied to seismic monitoring, leading to the declaration of uncataloged low-magnitude earthquakes (Zhang et al., 2014; Yoon et al., 2015; Liu et al., 2017).

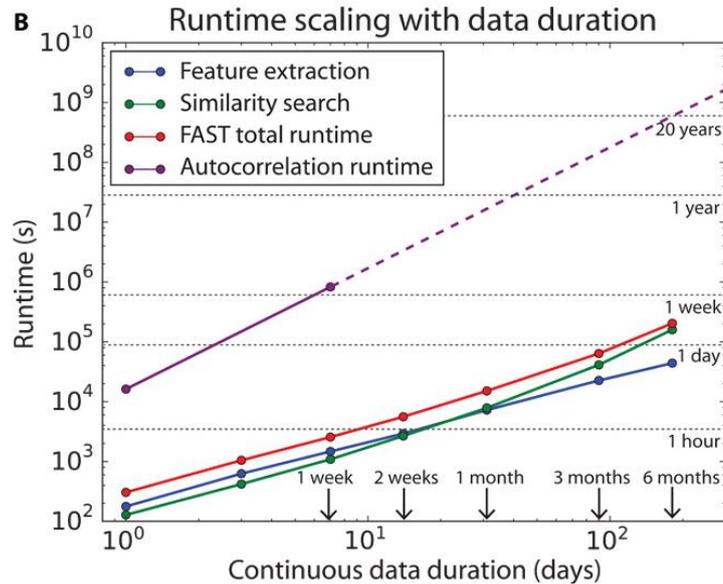


Figure 2.4) Autocorrelation runtimes (purple) as a function of the continuous data duration. For continuous data longer than 1 week, the runtime is extrapolated based on quadratic scaling (violet dashed line). Figure extracted from Yoon et al. (2015)

In the recent years, the explosion of the use of machine-learning (ML) models for solving complex scientific problems has led to the definition of deep learning techniques for the identification of seismic signals in continuous data (Perol et al., 2018; Ross et al., 2018; Zhu & Beroza, 2019; Mousavi et al., 2020). Deep-learning detection models work by learning general characteristics of earthquake waveforms and seismic phases from high-level representations, iteratively minimizing a loss function to achieve the best values for weights and biases in the framework of a classification problem. Despite differences in the network architecture, the main elements of a ML model for earthquake detection are the feature extraction layers, which effectively perform the classification problem minimizing a cost function, and a final layer which associates the probability of the input waveforms of being noise or earthquakes. More advanced networks are also able to detect the arrival of seismic phases (P and S waves) beyond the earthquake/noise classification. Figure 2.5 reports the architecture of ML models performing detection and phase picking on the input seismic waveforms (extracted from Figure 1 in Ross et al., 2018).

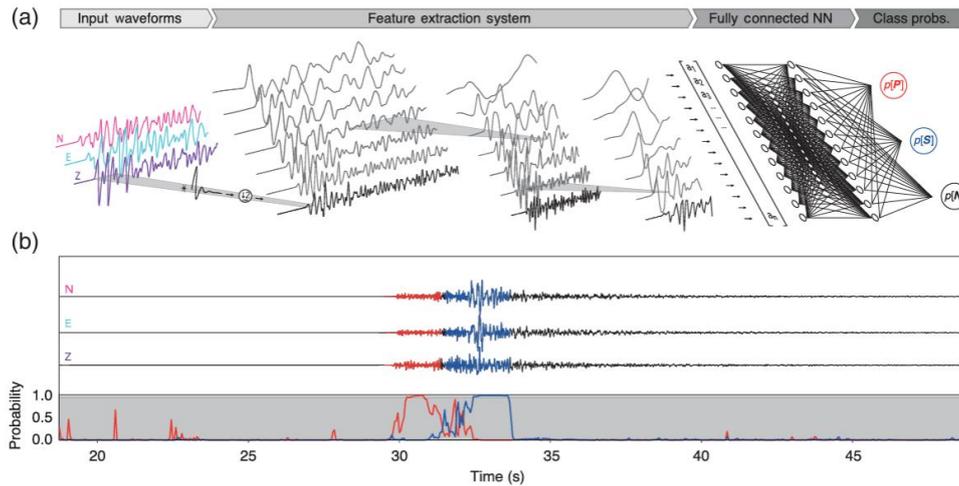


Figure 2.5) (a) Illustration of a convolutional neural network for generalized phase detection (GPD) and (b) application example. Probability time series for P (red) and S waves (blue) are used as characteristic functions for phase detection. The figure is extracted from Figure 1 in Ross et al. (2018)

Crucial for the training of a ML model is the selection of the training dataset. The deployment of dense and worldwide seismic monitoring networks has given rise to massive datasets, often containing millions of standardized labelled seismic waveforms, in terms of earthquake detection and phase arrival times, generated by expert seismologists (Mousavi et al., 2019; Woollam et al., 2019; Michelini et al., 2021; Cole et al., 2023). Moreover, the use of a massive dataset ensures a wide number of events to be used after the train/validation/test split of the training dataset. Using the labelled data, high-level characteristic features are automatically extracted from the seismic waveforms by applying a series of filtering and decimation operations and used to classify data windows. For this task, most of the state-of-the-art ML models ground on convolutional neural networks (CNN, Ross et al., 2018; Zhu & Beroza, 2019) and attention mechanisms (Mousavi et al., 2020) for identifying the most discriminative and characteristic properties of the seismic signals. These down-sampled features are transformed into high-level representations through a series of residual convolution blocks, and global attention section at the end of the encoder can help in aiming the attention of the network to the parts associated with the earthquake signal (Mousavi et al., 2020). These data-driven, high-level features are then directly mapped to a vector of probabilities representing the existence of an earthquake signal (detection) using decoder branches (Figure 2.5, panel b). Recent studies applied these techniques to different environments, reporting encouraging results as compared to the manual catalogs and to the application of ordinary detection strategies. In particular, the application of ML models in complex tectonic contexts provided an increased number of detected earthquakes of more than one order of magnitude (Zhu & Beroza, 2019; Tan et al., 2021; Chiaraluce et al., 2022b; Mancini et

al., 2022), even when applying the techniques on areas that were not included in the training datasets (Mousavi et al., 2020). These characteristics can, thus, support the exportability of the ML models in different contexts, overcoming the low generalization that typically affects the similarity-based models. Moreover, the application of ML models has been proved to provide reliable estimations of the P and S phase arrival times, as compared to the manual identification of expert seismologists (Mousavi et al., 2020; Munchmeyer et al., 2022), allowing an unbiased characterization of enhanced seismic catalogs in terms of earthquake location and source parameter estimations. Moreover, ML methods can detect events which are not correlated with other events in the catalog, and would be missed by a template matching procedure, as there is no matching template for such events.

In this chapter, we tested and compared different detection techniques in several frameworks, in order to generate robust and complete enhanced catalogs of microseismic events for statistical analysis and monitoring purposes. Moreover, these catalogs will represent the starting point for a comprehensive analysis of the microseismicity, discussed in the next chapters, improving the knowledge of the earthquake generation processes.

Specifically, we apply advanced detection techniques for:

- Generating enhanced detection catalogs for seismic sequences in Southern Apennines, testing and comparing the performance of similarity-based and ML techniques and interpreting, in a statistical framework, the information contained in the resulting catalogs (Section 2.3.1)
- Analysing the data collected by a pioneering survey, consisting of the deployment of 200 seismic stations in Southern Apennines, organized in dense arrays, for detecting the ultra-microseismicity ($M_I < 0$) and improving the knowledge about the seismogenic structures (Section 2.3.2)
- Investigating repeater and near-repeater earthquakes in Southern Apennines on the timescale of decades, to infer temporal variations in the properties of the propagation medium and to analyse the temporal evolution of the seismic cycle (Section 2.3.3)
- Monitoring the seismicity in a gas-storage area in Northern Italy, for tracking the induced earthquakes during the injection and extraction stages (Section 2.3.4)
- Defining a near-real-time automatic workflow for the continuous monitoring of seismicity in Southern Apennines, aiming to replace the routine manual surveillance by network operators (Section 2.3.5).

2.2 Methods

We hereby present the advanced detection techniques that we applied for the aforementioned analysis. We selected the ML models for earthquake detection and phase picking EQTransformer (Mousavi et al., 2020) and PhaseNet (Zhu & Beroza, 2019), the template-matching detector EQCorrscan (Chamberlain et al., 2018) and the autocorrelation technique FAST (Yoon et al., 2015).

2.2.1 Machine learning models for earthquake detection

The wide spread of artificial intelligence in Earth sciences has led to the development of machine learning models for identifying earthquakes and recognizing the arrival times of seismic phases (Perol et al., 2018; Ross et al., 2018). ML models learn discriminative data-driven features from labelled and massive training datasets, leading to robust and precise earthquake detection on continuous data, even in regions non included within the training dataset (Mousavi et al., 2020; Munchmeyer et al., 2022). These characteristics make the models particularly suitable for seismic monitoring purposes and catalog enhancement.

EQTransformer (EQT, Mousavi et al., 2020) is an AI-based earthquake signal detector and phase (P&S) picker employing a deep neural network with an attention mechanism, characterized by a hierarchical architecture specifically designed for detecting earthquake signals, inspired by human visual attention. Humans focus on a certain region of the waveforms for discriminating seismic signals from noise, i.e. sharp amplitude or frequency variations. EQT implements two levels of attention mechanisms, one at the global level for identifying earthquake signals within seismic waveforms, and one at the local level for identifying phase arrival times within that earthquake signal. The architecture of the model consists of 56 activation layers, which is the deepest network that has been ever trained for seismic signal processing, simultaneously performing the detection and phase picking, using separate loss functions. EQT is trained using the STEAD dataset (Mousavi et al., 2019), a large-scale global dataset of labelled earthquake and non-earthquake signals. The trained dataset consisted of ~1 M earthquake and ~300 K noise waveforms (including both ambient and cultural noise) recorded by ~ 2600 seismic stations at epicentral distances up to 300 km. Earthquake waveforms are associated with about 450 K earthquakes with a diverse geographical distribution around the world and were randomly split into training (85%), validation (5%), and test (10%) sets. Most of these earthquakes are smaller than M 2.5 and have been recorded within 100 km from the epicentre, these features being particularly promising for local and

microseismic monitoring. In Figure 2.6 we report the location-magnitude distribution of the STEAD dataset (Mousavi et al., 2019).

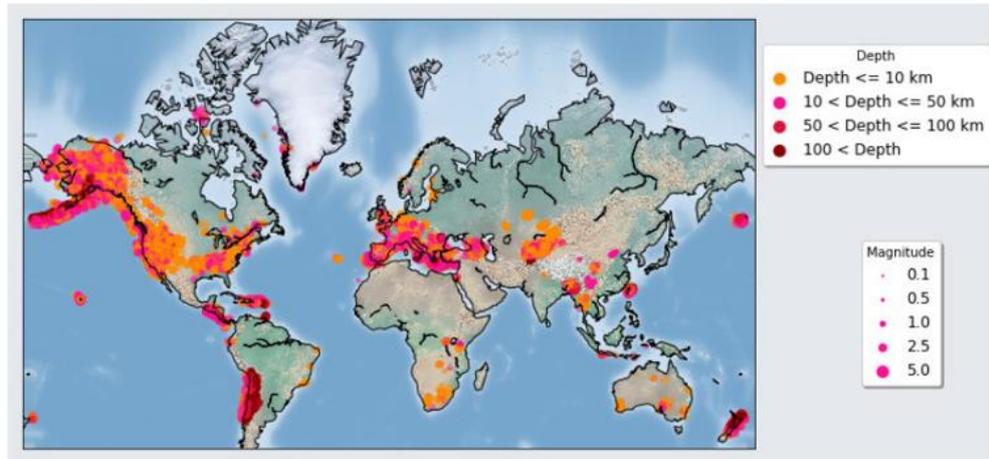


Figure 2.6) Location – magnitude distribution of the STEAD dataset. The figure is extracted from Figure 3 of Mousavi et al. (2019)

The 1-minute-long training waveforms were sampled at 100 Hz, extending for 1 minute and band-passed filtered in 1.0-45.0 Hz. When applying EQT on seismic records, we reproduced the same preprocessing operation, matching those of the training dataset. During the training stage, data were further augmented by adding a secondary earthquake signal into the empty part of the trace, in order to let the model be sensitive to multiple events and earthquakes occurring close in time and randomly shifting the event within the trace through array rotation, in order to avoid the model learning only the windowing scheme.

The output of EQT is a probability vector, each value associated with a sample of the analysed window, for earthquake detection, P and S phase arrival time. An earthquake is declared whether the detection probability overcomes an imposed score threshold, while phase arrival times are provided when the pick probability is higher than a second, independent threshold. In Figure 2.7 we report the results of the EQT for a 1-minute-long three-component normalized velocity record for a seismic station in Southern Apennines. The reported waveforms are downsampled from 125 Hz to 100 Hz and bandpass-filtered in [1 – 45] Hz. EQT allows to select two versions of the trained model: EqT_original_model.h5 (hereinafter “original model”) and EqT_model_conservative.h5 (hereinafter “conservative model”) (<https://github.com/smousavi05/EQTransformer/tree/master/ModelsAndSampleData>, last access 3rd July 2024).

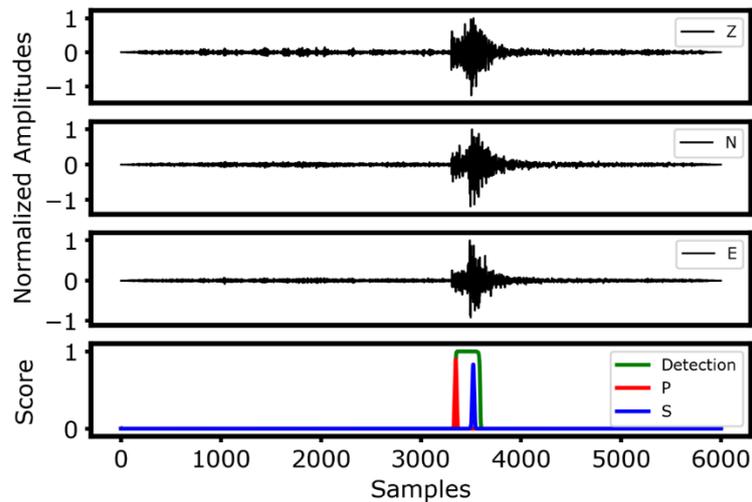


Figure 2.7) Application of EQT for a seismic station in Southern Apennines. The three upper panels report the normalized and [1-45] Hz filtered waveforms for the Z, N and E component of velocity records, respectively. The bottom panel reports the prediction scores for earthquake detection (green), P (red) and S (blue), respectively.

The former model was optimized to minimize the rate of false negatives, while the latter was trained to minimize the false positive rates. Both models are trained using the same dataset and the same network architecture but selecting and tuning differently the hyperparameters. Mousavi et al. (2020) applied the original model to the region that was struck by the 2000 Mw 6.6 Western Tottori earthquake, searching for aftershocks over more than one month of continuous data, using threshold values of 0.5, 0.3, and 0.3 for detection, P-picking, and S-picking respectively. This analysis led to the definition of a catalog composed of ~ 21.1 k earthquakes, along with reliable estimations of the phase arrival times, featuring zero-mean residuals and standard deviation lower than 0.1 s, when compared to manual identifications, for both P and S picks.

We tested the application of both models on seismic sequences in Southern Apennines, also analysing the reliability of the phase arrival times. For homogenizing the probability thresholds, we set the probability scores to 0.3 and 0.1 for declaring detections and phase arrival times, respectively. For binding single-station phase arrival times to seismic events, EQT is coupled with a phase associator, which simply counts the number of detections occurring in moving time windows, without inspecting for time and space coherency of the moveout of the declarations. In Figure 2.9, we report the number of phase arrival times declared by the conservative model (blue bars) and the original model (red bars) at the closest stations to a seismic sequence. When adopting the original model, which minimizes the rate of false negatives, we retrieved a list of ~ 3 k detections after the association, ~ 550 of which were real earthquakes (visual inspection of the declaration).

On the other hand, the application of the conservative model led to the declaration of 140 associated earthquakes, featuring very few false positives. However, since the binding criterion of EQT is too weak in the selection of real events, and even when using more advanced associators (Zhang et al., 2019) the resulting catalogs were dominated by false events and we preferred to maintain a safer approach for EQT. We thus adopted the conservative model for earthquake detection, and we further enhanced the catalogs using the enlarged set of events detected by EQT as master events (templates) for a template matching search, which requires a reliable and robust set of phase arrival times. Indeed, an event with low signal-to-noise ratio waveforms and only few P/S picks will not be an optimal template, eventually leading to a family of many false detections.

Similarly to EQT, PhaseNet (Zhu & Beroza, 2019) has become an established model for identifying earthquakes and phase arrival times on continuous data. The architecture of PhaseNet is modified from U-Net (Ronneberger et al., 2015) to deal with 1-D time-series data (Figure 2.8).

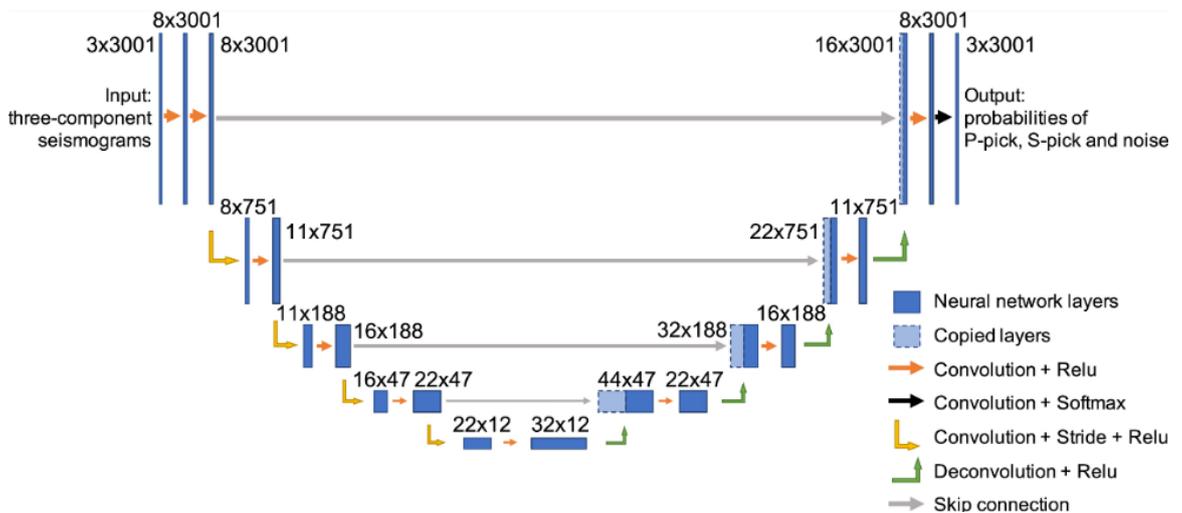


Figure 2.8) Network architecture of PhaseNet. The blue rectangles represent layers inside the neural network, while the numbers near them are the dimensions of each layer. The input seismic data go through four down-sampling stages and four up-sampling stages. The down-sampling is done by 1-D convolution and stride, while the up-sampling is done by deconvolution, which recovers the input length of the previous stage. Figure extracted from Figure 5 in Zhu & Beroza (2019).

The input seismic data are processed by four down-sampling and four up-sampling stages, involving 1-D convolutions and rectified linear unit (ReLU) activation functions. Down-sampling extracts useful information from raw seismic data, while the up-sampling process determines the probability distributions of P wave, S wave and noise for each time sample (Zhu & Beroza, 2019). During the training stage, minimal data pre-processing is applied to the training data, which consisted of a 30-s time window that

included the P and S arrival times. As for EQT, the position of the arrivals within the window is varied to ensure that the algorithm does not just learn the windowing scheme. All data are sampled at 100 Hz and normalized by removing its mean and dividing it by the standard deviation. However, in contrast to EQT, the training dataset PhaseNet was not filtered in the frequency domain, in order to also deal with event waveforms close to the noise level within the training stage. PhaseNet was trained on seismic waveform data based on the Northern California Earthquake Data Center Catalog (NCEDC, 2014), composed of ~780k recordings. As for other ML pickers, the output of the technique is a vector of probability scores. Zhu & Beroza (2019) applied the technique to different environments, retaining only the phase arrival times with probability scores higher than 0.5, for both P and S picks, showing similar distribution in terms of time residuals between automatic and manual picks as for EQT. Moreover, the technique appeared to be robust in detecting events in noisy data and even on clipped waveforms (Zhu & Beroza, 2019).

We tested PhaseNet on the same dataset used for evaluating the performances of two models of EQT, selecting the same probability score as for the same former tests. In Figure 2.9, we added the number of picks provided by PhaseNet as orange bars, along with the two versions of EQT.

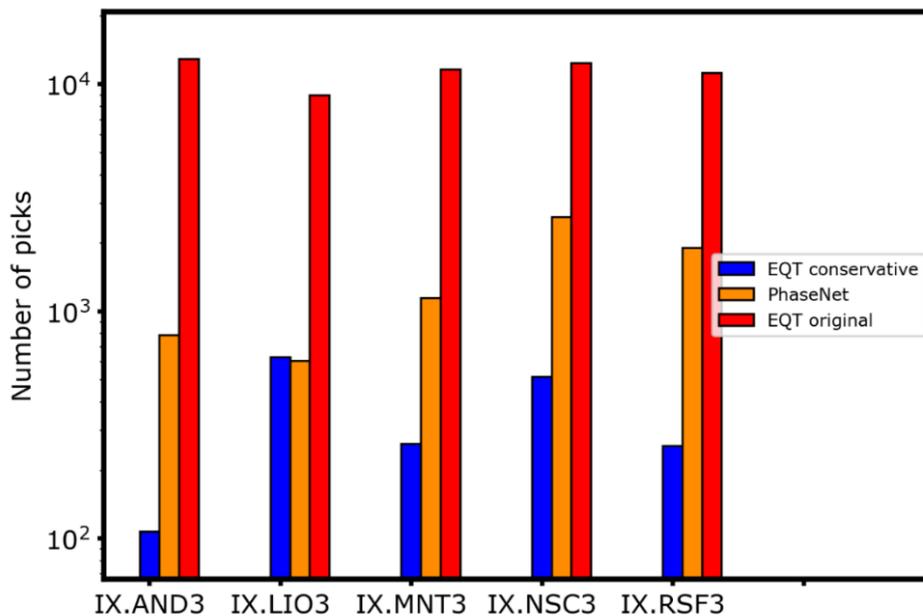


Figure 2.9) Comparative analysis in logarithmic scale of the picks provided by EQT conservative model (blue bars), PhaseNet (orange bars) and EQT original models (red bars). For each model, an increase of one order of magnitude is typically observed, with the number of PhaseNet picks laying between the two version of EQT model.

From PhaseNet, we obtained a number of detections that falls between the resulting number from the application of the conservative and original EQT model but, as for the EQT original model, this technique still requires a good associator to produce a seismic catalog that is robust enough to not be polluted by false declarations. When coupling the PhaseNet detections with REAL associator (Zhang et al., 2019), a grid-search-based binding algorithm, we obtained a detection list of 250 real earthquakes. This number is lower than the cleaned catalog provided by the original EQT model but almost doubles the number of detections from the EQT conservative version. However, since the effect of the phase association appears to be crucial in ruling out the false declaration and it might not be trivial to disentangle the role of the association stage from the effective detection capability, we preferred to not include this technique in the analysis of seismic sequences in Southern Italy (Section 2.3.1) and the data collected by the dense arrays (Section 2.3.2). On the other hand, since the number of detections provided by PhaseNet falls between those of the two EQT models, this technique looks particularly encouraging for seismic monitoring application as stand-alone ML model, approaching near-real-time application, which is a challenging framework for template-matching strategies. We thus adopted PhaseNet, coupled with the phase associator REAL, to implement an automatic detection workflow for daily monitoring of the seismicity in Southern Apennines, complementing, and in the near future aiming to substitute, the routine operations by human seismologists (Section 2.3.5).

2.2.2 An efficient template matching strategy: EQCorrscan

Template matching is an efficient similarity-based approach for earthquake detection, exploiting the waveform similarity among events occurring close in space, making this technique particularly suitable for the analysis of seismic sequences. In Section 2.3.1, 2.3.2 and 2.3.3, we applied the template matching technique EQCorrscan (Chamberlain et al., 2018) for the analysis of seismic sequences, continuous data recorded by a dense temporary survey and for identifying repeating events in Southern Apennines, respectively. As template matching techniques require the knowledge of the set of earthquakes to be used for finding similar events within the continuous data, we grounded our detection strategy using the earthquakes detected by EQT in a previous stage as master events, in both Section 2.3.1 and Section 2.3.2. On the other hand, since for the search of the repeating earthquakes we did not have the phase arrival times provided by the machine learning technique for 10 years of continuous data, we used the available manual identification of earthquakes and corresponding phase arrival times. However, to further understand the performance of the technique, we discussed the content of template matching enhanced catalogs obtained using either the manual or the EQT detections as template sets for the analysis of the seismic sequences in Southern Apennines (Section 2.3.1).

When building a template event, we extracted a portion of waveforms for stations with at least one pick (P or S). Several stations might contribute to the template, with the moveout of the event among the stations being indicative of the direction of propagation. We selected portions of templates lasting 1.5 s and starting 0.15s before the picks. For enhancing the similarity between the template and the continuous data and for reducing the computational efforts, both waveforms were bandpass-filtered in the range 2–9 Hz and decimated to 25Hz, similarly to other template matching applications (Vuan et al., 2018; Ross et al., 2019). In Figure 2.10, we report the portion of the waveforms composing a template event (magnitude M_l 1.2) extracted by selecting the phase arrival times provided by EQT. Red waveforms indicate the 1.5s – long portions extracted around the P picks on the vertical components of the velocity records, while the cyan sectors indicate the waveforms extracted around the S picks on the horizontal components of the velocity records.

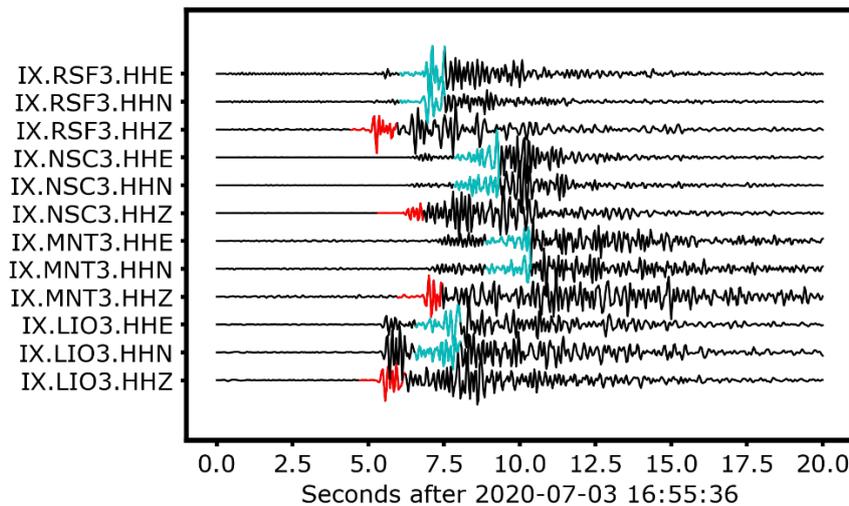


Figure 2.10) Example of a template event used the phase arrival times provided by EQT. Red waveforms indicate the 1.5s – long portions extracted around the P picks on the vertical components of the velocity records, while the cyan sectors indicate the waveforms extracted around the S picks on the horizontal components of the velocity records.

When performing the similarity-based detection using all the available templates, EQcorrscan implements a delay-and-stack method for the calculation of normalized cross-correlation coefficients, with correlation vectors aligned according to the relative delays between template channels. For each template, the cross-correlation with the continuous data is only evaluated, among the stations composing the template, on the channels providing the phase arrival times. For attributing a detection score to each template, we used the sum of the cross-correlation of the considered master events and the continuous data among the different channels. We then compared the similarity score

with an imposed threshold, defined from the median absolute deviation (MAD), of the cross-correlation coefficient evaluated for one hour of continuous data. In this framework, the MAD value will be indicative of the level of cross-correlation of the template with the noise, being the one hour of continuous data being typically dominated by noise signals. We declared an event when the sum of the similarity score of the analysed records, with at least one template, overcame 8 times the MAD value. This latter value is comparable to, and even slightly lower than, the similarity thresholds imposed in other template matching analyses (Skoumal et al., 2014; Li & Zhan, 2018; Sukan et al., 2023). 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, which might be due to the artificial similarity observed with noise waveforms enhanced by the narrow filtering operation. 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 thus explored the possibility of automatically identifying a refined threshold for ruling out false positives from the enhanced catalogs (Scotto di Uccio et al., 2023). For several test cases, we visually inspected all the detections with similarity scores higher than 8 times the MAD, to isolate real events from false positives accounting for the shape and the frequency content. In the left-hand panel of Figure 2.11, we report the distribution of the ratio between the sum of the cross-correlation associated with the single declaration and the detection threshold ($CC_{sum}/thresh$), distinguishing false events (red points) and real earthquakes (green points) for an enhanced catalog composed of 233 detections.

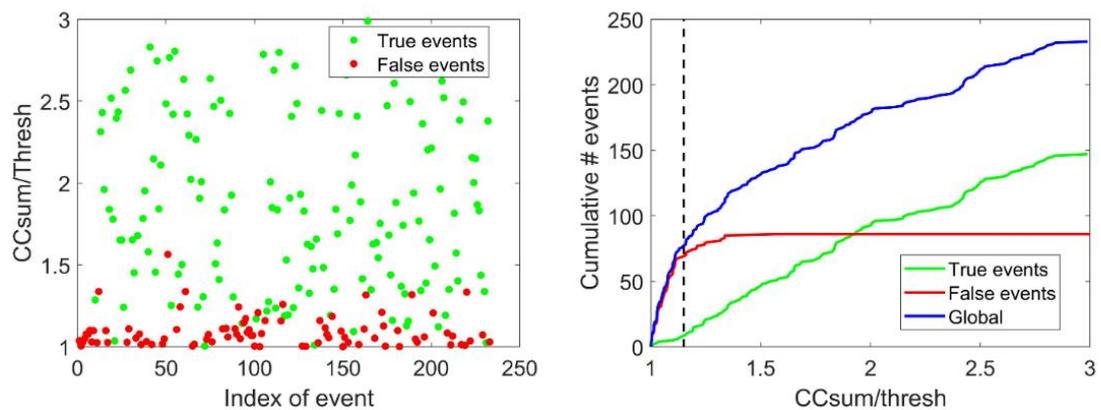


Figure 2.11) Left-hand panel: distribution of the ratio $CC_{sum}/threshold$ for the events in the initial catalog provided by EQCorrscan (green points correspond to real events, red points to false positives). Right-hand panel: cumulative number of events as a function of the ratio $CC_{sum}/threshold$ for the subset of the real events (green line), false positives (red line) and for the whole initial catalog (blue line). The black dashed line marks the similarity parameters associated with the slope change in the global $CC_{sum}/threshold$ distribution, which we use as refined threshold.

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 Fig.1.11, 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. 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 behaviours, the cumulative distribution of all the detections exhibits a change in the slope, which can be identified (black dashed line in the right panel of Figure 2.11) by fitting its initial and final trends and used as a refined threshold that allows us to significantly reduce the number of false detections in the catalog. In the right panel of Figure 2.11, the slope break corresponds to the value $CCsum/thresh=1.15$, which results in a new threshold of 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.

Finally, we also estimated the phase arrival times for detections, this time exploiting the individual cross-correlation coefficient. We imposed a minimum cross-correlation coefficient of 0.6 for accepting the phase arrival time, while selecting a maximum time lag of ± 0.2 s in the exploration. This approach allowed us to fine-tune the individual picks around the detection made with all channels simultaneously.

2.2.3 An innovative technique for autocorrelation: FAST

As introduced in Section 2.1, the autocorrelation is a similarity-based strategy that is grounded on the expected similarity among seismic signals, as compared to random noise records. The autocorrelation is an uninformed detector, since it does not require the knowledge of the desired seismic signals. Rather, it considers every time window extracted from the continuous data as a candidate waveform to be searched within the records. This approach appears to be particularly promising especially for seismic sequences, in which earthquakes occur closely in space and time with a seismic rate higher than background seismicity, and for monitoring induced seismicity, given the high similarity expected from the low inter-event distance (Schaff & Waldhauser, 2005). However, the computational cost required by the autocorrelation ($O(N^2)$, see Figure 2.4), makes this strategy unfeasible for long-term monitoring. To overcome this limitation, Yoon et al. (2015) introduced a new algorithm, known as FAST (Fingerprinting And Similarity Thresholding), claiming to feature the same resolution as standard autocorrelation techniques while strongly lowering the computational time. The

innovation introduced within FAST stands in the conversion of time-domain waveforms into binary objects, known as fingerprints, which only contain discriminative features of earthquakes. Fingerprints serve as a proxy for waveforms, thus two similar waveforms should produce similar fingerprints, while two dissimilar waveforms should be converted into dissimilar fingerprints. Figure 2.12 illustrates the main stages of the conversion of time-domain waveforms into binary fingerprints.

The core of FAST consists of three stages:

- Feature extraction: in which the time-domain data are effectively compressed and the features are extracted for generating the sparse and binary fingerprints containing the most discriminative characteristics.
- Similarity search: given a query fingerprint, an optimized Min-Hashing algorithm (Broder et al., 1998), maps the object in a local database and returns the most similar fingerprints, without needing to compare it with the entire set.
- Network pseudo-association: if multiple stations are used, single-station results are combined to declare events in the network, ruling out eventual false detections at the station level.

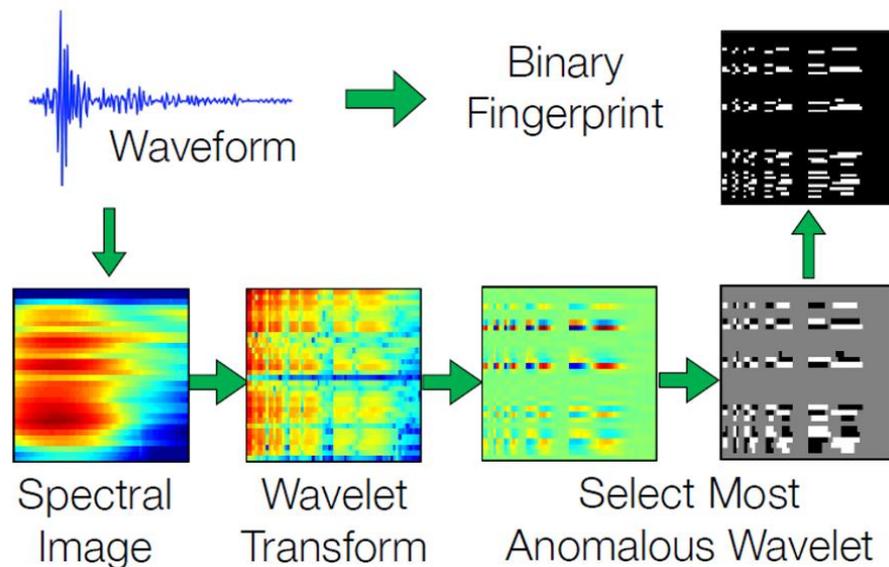


Figure 2.12) Conversion of time-domain waveforms into binary fingerprints containing only the discriminative features of earthquakes. Figure is redrawn after Yoon et al. (2015).

In the pre-processing phase, gaps and zeroes in the data stream are replaced with Gaussian uncorrelated noise, since these characteristics would have been misleadingly interpreted as similar. Moreover, with the twofold aim of further reducing the computational time and enhancing the signal-to-noise ratio of the waveforms, time-domain waveforms are downsampled and filtered in a narrow frequency band. When applying FAST to enhance catalogs for seismic sequences in Southern Apennines (Section 2.3.1) and to monitor the induced seismicity in the gas-storage area of Collalto (Section 2.3.4), we applied a downsampling factor of 5 to the raw continuous data, leading to a new sampling rate of 25 Hz and 50 Hz, respectively. The different sampling rates implied different frequency bands to be used for filtering, given the lower Nyquist frequency for data collected in the southern Apennines, as compared to the data from the gas-storage area. We thus filtered the former records in the [1-10] Hz frequency band, while the latter in the [1-20] Hz frequency band. However, both frequency bands agree with those selected for different applications of similarity-based detectors (Yoon et al., 2015; 2019).

In the feature extraction phase, single-component waveforms extracted from the pre-processed continuous data are converted into spectrograms, using overlapping windows and subdividing the frequency domain into equally spaced bins. Using this representation, earthquakes appear in the spectrogram as transient, high-energy events. The spectrogram is further split into overlapping windows in the time dimension, each window referred to as a "spectral image". Each spectral image extracted from the spectrogram is then compressed by applying the Haar wavelet transform and maintaining only a certain fraction of the coefficients, those that are most detached from their daily average. This compression is shown to preserve the seismic features in the fingerprints under the assumption that the daily average is representative of the ambient noise. The selected coefficients are then binarized, to generate final sparse and binary fingerprints. In both applications of FAST, we adopted 6.0 s of time-window signal with a shift of 0.2 s for evaluating an individual spectrogram, while the spectral images, resampled in 32 x 32 samples, contain information of 12.4 s of time-domain windows. The time extension of the spectral image is long enough to include the whole earthquake in its content and, on the other hand, sufficiently short for separating events occurring close in time, as it might occur during seismic sequences. When applying the Haar wavelet transform, we maintained only 200 out of the 1024 wavelet coefficients, selecting them among those featuring the highest Z-score. Z-score is a useful metric for quantifying the deviation of the i -th coefficient c_i from the mean μ_i weighted by the standard deviation σ_i of the same coefficient, being defined as:

$$Z_i = \frac{c_i - \mu_i}{\sigma_i} \quad (2.3)$$

Hypothesizing that μ_i is dominated by noise, the highest the Z_i , the more that coefficient detaches from non-seismic behaviour.

After feature extraction, the goal is to identify pairs of similar fingerprints to detect earthquakes. FAST creates a database in which similar fingerprints are grouped, using hashing functions, into the same hash bucket. Then the similarity search returns all similar fingerprints to a given search query object, measuring the similarity in terms of the Jaccard similarity. Given two binary fingerprints, the Jaccard similarity is defined as $J(A, B) = \frac{|A \cap B|}{|A \cup B|}$. The numerator contains the number of bits in both A and B that are equal to 1, whereas the denominator is the number of bits in either A, B, or both A and B that are equal to 1. Locality-sensitive hashing (LSH) efficiently compares fingerprints, returning a shorter list of “candidate pairs” that are similar with high probability. Each fingerprint is inserted into one hash bucket that is selected based on the output of a hash function. A hash table contains many hash buckets, and the hash function determines how items are distributed among the different hash buckets. LSH allows to search for pairs of similar items (seismic signals) within the same hash bucket—these pairs become candidate pairs, ignoring pairs of items that do not appear together in the same hash bucket, which comprise most pairs. 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). The number of hash functions, r , is the most sensitive parameter in the similarity search, also affecting the computational time: lower values lead to fewer missed detections, more false detections and longer runtime, while higher values produce more missed detections, fewer false detections, shorter runtime (Yoon et al., 2015). Thus, increasing the number of hash functions is, in a probabilistic framework, equivalent to requiring a higher similarity among the waveforms. Yoon et al. (2015) suggested values for the number of hash functions to be selected according to the time extension of the dataset. For days/weeks long datasets, the suggested value is 4. We thus set $b = 100$ and $r = 4$ (Yoon et al., 2015; Festa et al., 2021). The search is fast and scalable with increasing database size, with near-constant runtime for a single search query, so that the total runtime is near-linear (Yoon et al., 2015).

In the application of the similarity search for seismic sequences in Southern Apennines, we noted an anomalous and suspicious high number of similar fingerprints for a specific station, one order magnitude greater than for the other stations. Being a class of false positives in autocorrelation techniques represented by coherent noise occurring over time due to local ambient source, we investigated the time-domain waveforms for that station. In Figure 2.13, left panel, we show 30 seconds of raw data recorded at the considered station in the upper panel, while the lower panel shows the [1 – 10] Hz filtered waveform, the input for the fingerprint generation. This representation clearly shows that the [1-10] Hz bandpass filtering on this station strongly affects its frequency content, leading to a

nearly monochromatic time-domain signal. Consequently, since such a signal extended for several minutes and was randomly distributed within the daily records, the similarity search found a massive number of similar fingerprints, that should be considered as false positives given their non-seismic nature. Moreover, inspecting the power spectral density for the considered station (PSD, displayed in the right panel of Figure 2.13), we noted a spurious frequency peak around 1.5 Hz, which falls within the frequency band used for filtering the data. Thus, the selected frequency band significantly alters the shape of the time-domain waveforms, leading to an artificial similarity over time. A straightforward solution for overcoming this issue might have been represented by the selection of a different frequency band, not containing 1.5 Hz. However, this frequency range has particular interest in seismic signals; thus, to mitigate this effect, we discarded those fingerprints that are similar for more than 3 percent of the day length (~ 15 min). This choice effectively reduced the number of false detections due to the occurrence of coherent and repeated noise. Indeed, FAST might also be sensitive to non-seismic, repeated transients resulting from anthropogenic activities (e.g. quarry blasts, Yoon et al., 2017).

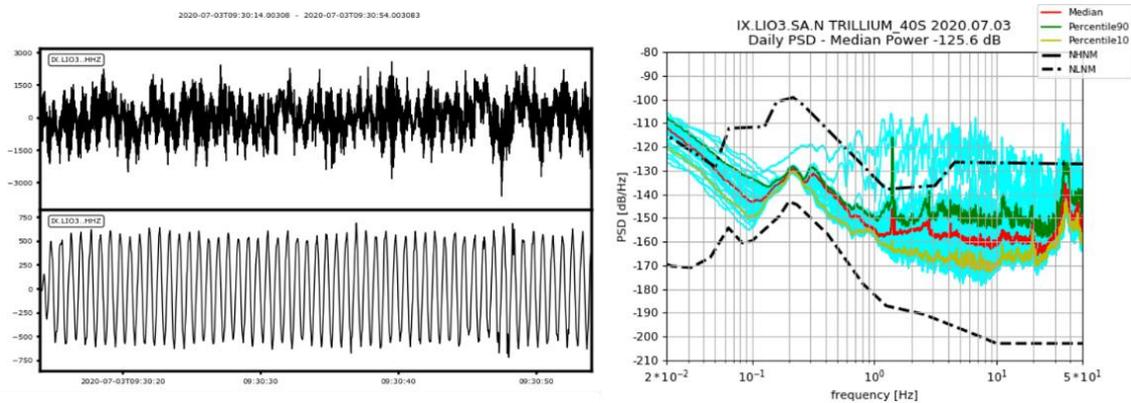


Figure 2.13) Left panel: The upper panel reports the raw waveform at LIO3 station, while the lower panel contains the [1-10] Hz filtered data, which feature a monochromatic behaviour, this latter causing an increase in the number of similar fingerprints for that station. Right panel: Daily power spectral density (PSD) for LIO3 station for the considered day. The PSD shows a spurious peak around the characteristic frequency of the monochromatic signal dominating the waveforms after the filtering.

In the final association stage, we merge the single-station declarations requiring similarity among fingerprints to occur in at least 2 stations for the application to seismic sequences in Southern Apennines and at 4 stations for the application to the gas storage area in Collalto. Moreover, in order to account for the propagation of waves among the different seismic stations, we imposed a maximum lag between detection times of 3 s. This latter value was selected considering the average inter-station distances.

Similarly to the approach introduced for the analysis of the template matching catalogs, for limiting the false detections within the FAST declarations we investigated the possibility of setting up a semi-automatic criterion, based on the analysis of the cumulative distribution of the similarity score parameter for all the detections, at a fixed number of stations. We typically observe a slope break in the cumulative distribution, and below this corner, the reported similarity scores are very close to each other. When inspecting the waveforms for events featuring similarity scores below this value, 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. 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 was large enough to allow statistical analysis in terms of the distribution of the total similarity scores (*peaksum*). To validate this criterion, we also performed a visual inspection of the declarations to discriminate 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 were found similar to those of events above the threshold (repechage phase).

2.3 Application of advanced detection techniques

In this section, we report the main results from the application of the aforementioned techniques in several contexts. We compared the detection performance of similarity-based detection techniques for catalog enhancement in Southern Apennines (Section 2.3.1). We integrated machine learning and template matching approaches for detection of the microseismic events recorded by a dense array deployment in Southern Apennines (Section 2.3.2) and using template matching we sought for repeated earthquakes in 10 years of continuous data from the same area (Section 2.3.3). In Section 2.3.4 we used autocorrelation for monitoring the microseismicity in the gas storage area of Collalto (Northern Italy), while in Section 2.3.5 we built up an automatic detection system for daily monitoring of the seismicity in the Southern Apennines.

2.3.1 Comparing and integrating advanced detection techniques for seismic sequences in Southern Italy

The analysis of seismic sequences, characterized by a higher seismicity rate as compared to the background seismicity and featuring earthquakes occurring close in locations and occurrence time, have been proven to provide precious information about earthquake generation processes, with the analysis of enhanced catalogs helping in extracting

statistical and geometrical information (Vuan et al., 2018; Beroza et al., 2021; Tan et al., 2021; Chiaraluce et al., 2022b).

In this analysis (Scotto di Uccio et al., 2023), we applied the autocorrelation technique FAST, the machine learning detector EQT and the template matching approach EQCorrscan to 10 seismic sequences that occurred in Southern Apennines, between 2011 and 2020. Southern Apennines host the complex Irpinia normal system, which was struck by the 1980, M 6.9 earthquake that occurred along NW–SE-striking faults and was characterized by three main episodes within a few tens of seconds. It caused about 3000 fatalities and severe damage (Bernard & Zollo, 1989; event epicentre is reported as a red star in Figure 2.14). Moreover, this area is considered one of the regions with the highest seismic hazard in Italy (Stucchi et al., 2011). Since 2007, the area is continuously monitored by the Irpinia Near Fault Observatory (INFO). The key idea of Near Fault Observatories is to install dense networks of multiparametric sensors close to faults, aiming at understanding the underlying Earth instability processes over broad time intervals (Chiaraluce et al., 2022a). 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 $100\text{km} \times 70\text{km}$, including the epicentre of the 1980 Irpinia earthquake (Iannaccone et al., 2010), with an interstation distance of 10–20 km. ISNet daily releases a seismic catalog (<http://isnet-bulletin.fisica.unina.it/cgi-bin/isnet-events/isnet.cgi>), which consists of two layers: the first one is an automatic catalog 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 missed detections by the automatic procedure. At this stage, an event is included in the manual catalog if at least 4 phases have been identified on the records, including one S phase, and the P and S pick residuals are smaller than 0.5 and 1.0 s, respectively. The manually revised catalog of seismic events for the past 15 yr includes ~ 3000 earthquakes, which is reported in Figure 2.14 with shaded black circles, along with the stations of the ISNet network.

Seismicity in Irpinia typically features a low seismic rate, consisting of few events per day, and moderate depth values, mainly ranging between 8 and 15 km. Moreover, recent studies have shown a seasonal dependence of the seismic rate, caused by hydrological forcing due to the variation in the charge level of the karst aquifer (D’Agostino et al., 2018; Tarantino et al., 2024). The events cover a local magnitude range between $M_l -0.4$ and $M_l 3.7$, with a completeness magnitude of $M_l 1.1$ (Vassallo et al., 2012). 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 is sometimes clustered in sequences with events of maximum magnitude $M_l \sim 3.0$ (Stabile et al., 2012), lasting for a few days.

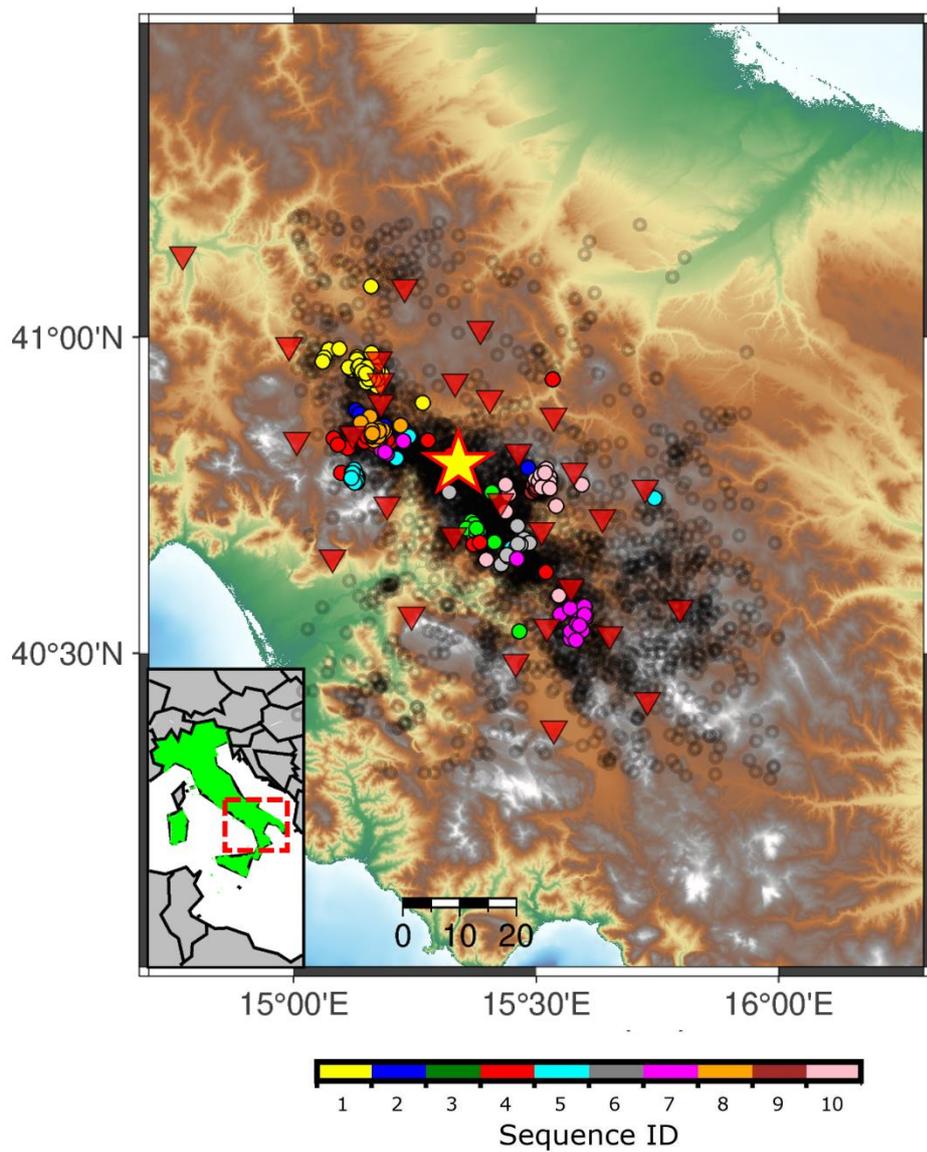


Figure 2.14) Manual ISNet catalog from January 2007 to June 2022. Background seismicity is represented with shaded black circles, while the ten considered seismic sequences are colored with different colors. The red star marks the epicenter of the 1980 M 6.9 earthquake. Red triangles indicate the seismic stations of the ISNet network.

The monitoring strategies currently adopted at INFO led to manually revised catalogs containing between 8 and 74 events for the seismic sequences. In Table 2.1, we report the

information regarding the main event and the number of detections in the manual catalog for each of the ten considered sequences.

Table 2.1) Number of detections in the manual catalogs and information regarding the main event for each of the analysed seismic sequences.

ID	PLACE	DATE	LAT	LON	DEP	M_l	# 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)	2019/08/08	40.775	15.499	6.3	3.1	23

Aiming at increasing the content of catalogs for seismic sequences including low-magnitude earthquakes, for the analysis of each sequence we selected the continuous velocity records (more sensitive to microseismic events) at the 5–7 closest stations to the sequence's centroid, spanning a hypocentral distance range between 10 and 30 km. Our choice was guided by the need to have optimal 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. When cumulating the number of detections for the sequences, the manually revised INFO catalog contains 265 events, 82 of which were automatically declared by the STA/LTA-based detector operating on the network.

To assess the overall performance of the adopted detection techniques, we analysed the results obtained from applying four strategies to the data of ten seismic sequences in the Irpinia region: (i) FAST, (ii) EQT, (iii) EQCorrscan using EQT's templates (hereinafter, EQT + TM) and (iv) EQCorrscan using as templates the manual detections from INFO (hereinafter, INFO + TM). In Figure 2.15 we show the number of events provided by each technique, after applying the quality selections described in Section 2.2, organized in a Venn diagram. To extract the most complete catalogs, we merged the lists provided by each technique according to the detection time. The merged catalog contains 1792 events, increasing by a factor of ~ 7 the revised manual catalog and by a factor of ~ 21 the automatic one. Looking at the overall performance of the single techniques, we report that FAST declares 942 events ($\sim 3.5\times$ the manual catalog, $\sim 11\times$ the automatic catalog), EQT detects 450 events, increasing by factors of 1.5 and 5.0 the revised and automatic

catalogs, respectively. EQT + TM declares 1715 events, with a catalog content similar to the merged one, while INFO + TM detects 1165 events ($\sim 68\%$ of the EQT + TM catalog). Most of the events (95%) declared by FAST are also retrieved by EQT + TM, while the detections common to FAST and INFO + TM decrease to 734, representing 78% of the FAST catalog.

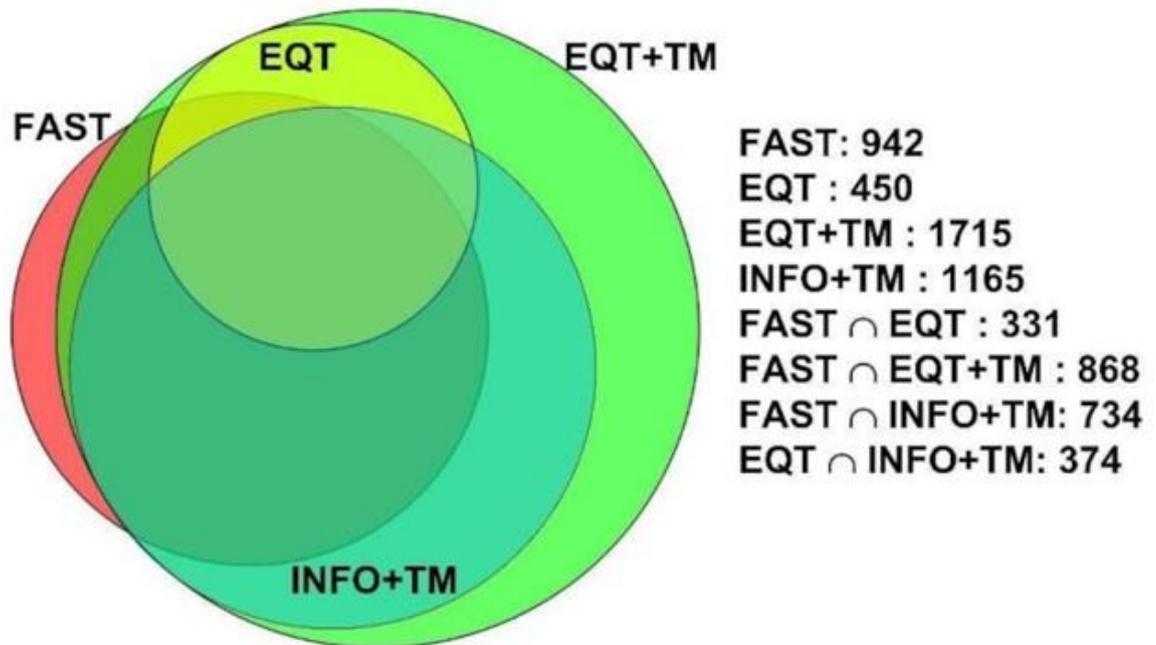


Figure 2.15) Venn diagram showing the performance of the different detectors: the autocorrelation (FAST, red), machine learning (EQT, yellow), template matching (EQT + TM, green; INFO + TM, cyan) techniques. The EQT detections are included in the EQT + TM ones because the EQT output is used to form templates.

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 (Rocca San Felice) has a significantly larger number of earthquakes (~ 300 events in one day), while for the other sequences the rate is lower, typically presenting between 50 and 200 events in 4 days. In Figure 2.16, we report the number of detections provided by the techniques for each sequence, with the Rocca San Felice sequence (ID 1) reported separately in the left panel, to improve the visualization of the results for the other sequences (ID 2–10), due to the different seismic rate. Inspecting the results for each sequence, we note consistent performances among the detection strategies: for each sequence, the most complete catalog is obtained using EQT events as template set for the template matching detection using EQCorrscan (red bars in Figure 2.16). In fact, the number of events detected by

EQT (orange bars in Figure 2.16) is typically twice larger than the existing manual catalog (pink bars in Figure 2.16). The use of manual catalogs as template sets (INFO + TM, green bars in Figure 2.16) produces lower matched detections, due to the fewer available templates, but the resulting catalogs are generally still richer than the autocorrelation catalogs. On the other hand, FAST provides a number of detections twice larger than EQT, but since it does not provide information about the phase arrival times along with the detections, necessary for extracting the template events, it is not feasible to apply a template matching based on its catalogs.

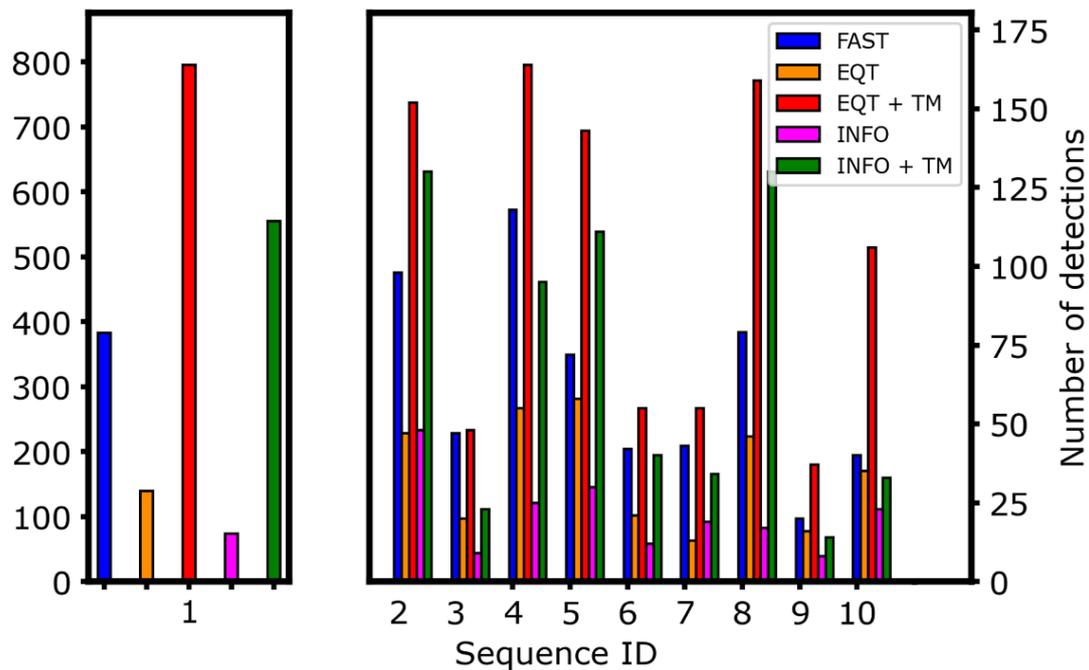


Figure 2.16) Number of detections for each sequence provided by FAST (blue), EQT (orange), EQT + TM (red), manual INFO catalog (pink) and INFO + TM (green). The results for sequence ID 1 (Rocca San Felice) is shown independently in the left panel, since it features the highest number of events.

Inspecting the temporal distribution of the events, as a function of the occurrence time since the main event, we highlight the presence of foreshocks anticipating the mainshocks. In Figure 2.17, we report the cumulative fraction of events occurring 24 hours before and after the main event, color-coded according to Figure 2.14 and Table 2.1. Looking at the time evolution of sequences, we can identify two main behaviours. Most of the sequences generate the majority of earthquakes (around 70 percent of the catalog) within 6 hours after the mainshock and are characterized by a similar temporal evolution of the aftershocks. Three sequences feature swarm-like behaviour with several magnitudes of the aftershocks being comparable to that of the largest event. For the most

populated sequence (ID 1, yellow curve in Figure 2.17), we observe 21 low-magnitude foreshocks within 14 hours before the main event. Conversely to the other cases, during this sequence, characterized by a main event of M_l 3.0, the seismic rate further accelerated two days after the mainshock, following the occurrence of a M_l 2.8 event, driving the sequence for the next two days.

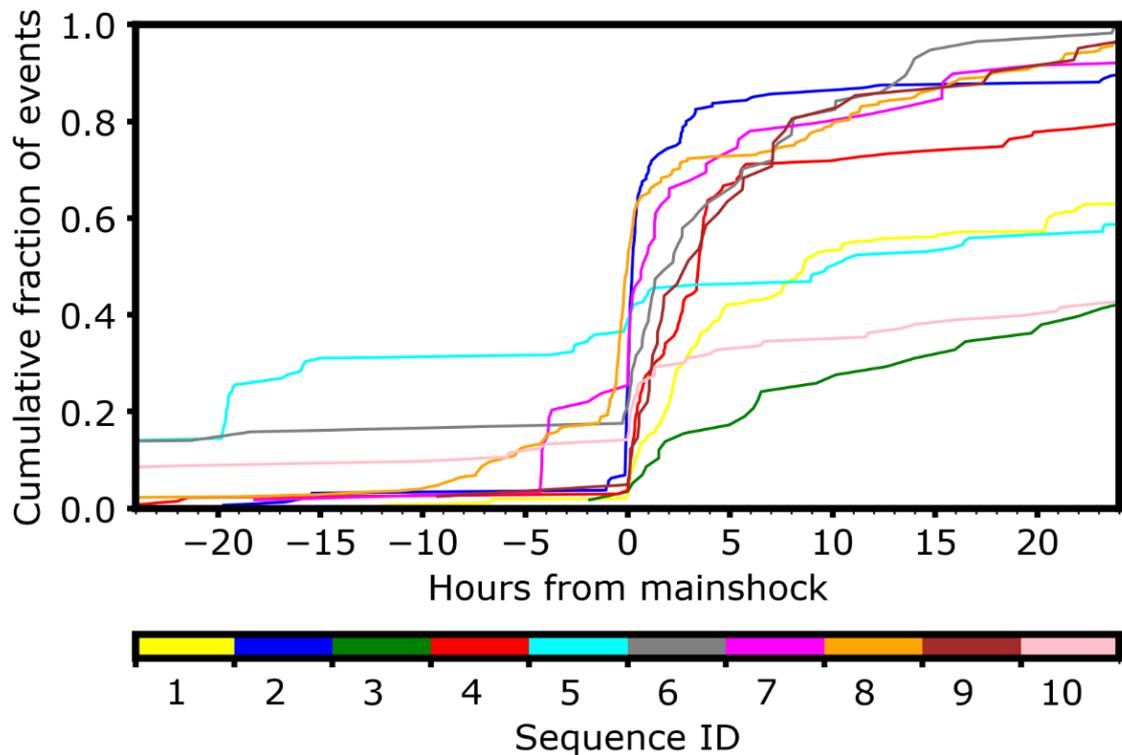


Figure 2.17) Foreshock/aftershock analysis. This figure shows the cumulative percentage of the events in the catalog as a function of the time since the main event.

FAST declares 73 events that are missed by the other techniques, which represents 5 % of the merged catalog. After checking these events, we observe that they typically feature a low, close to one, signal-to-noise ratio. For these events, we find that they exhibit a smaller cross-correlation value with the used templates, but they can still be retrieved by the EQT + TM technique by lowering the acceptance threshold, at the cost of a significant increase in the number of false events. As an example, for the San Gregorio Magno sequence (ID 3), EQT + TM technique is able to detect 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.

Most of the detections retrieved by EQT + TM and missed by FAST correspond to earthquakes occurring in the coda of the previous events. In Figure 2.18, we report a one-minute-long waveform extracted during the Rocca San Felice seismic sequence, which contains multiple and overlapping low-magnitude earthquakes. While the template matching can effectively discriminate nearly overlapping events due to the short extension of the templates (1.5 s), leading to the declarations of 5 earthquakes in the considered window (none of these reported in the manual catalog), FAST groups the events together into a single and extended chain of similar fingerprints. In Figure 2.18, the green boxes refer to the individual detections provided by EQT + TM, while the extended red box marks the long fingerprint chain that was reported by FAST.

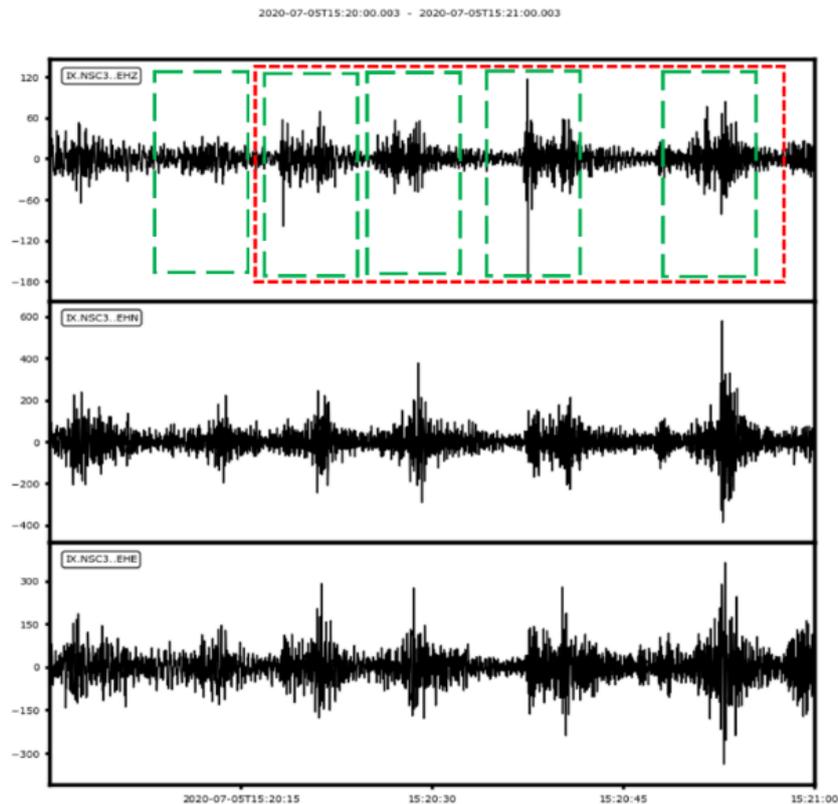


Figure 2.18) Performance of EQCorrscan and FAST on a one-minute-long waveform featuring near-overlapping events. Green boxes mark the disentangled detections from template matching catalog, while the wide red box marks the single long chain of similar fingerprints provided by FAST, which does not distinguish the overlapping events.

When shortening the time window for fingerprint generation from 6 to 3 s, FAST is able to separate most 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.

Analysing the content of the catalogs provided by the machine learning technique EQT, we note that it provides the least enlarged detection list, but it is still twice larger than the existing manual catalogs. We ascribe the lower performances of the technique both to the stressing test bed of this application, consisting of sequences of microearthquakes, which might not be fully representative of the training dataset (Mousavi et al., 2019), occurring close in time, and to the choice of the selected version of the EQT model.

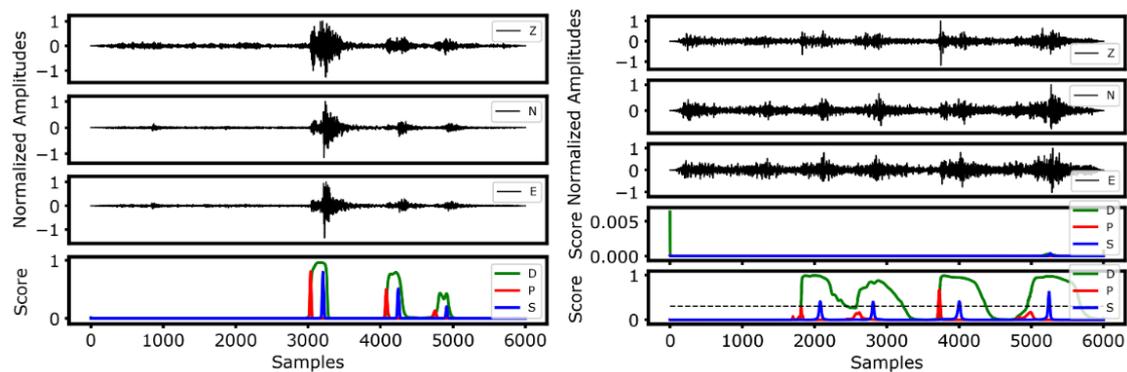


Figure 2.19 – Left panel: Application of the conservative model of EQT to a one-minute-long waveforms containing multiple events. EQT can correctly discriminate the events, featuring magnitude 1.0, 0.4 and 0.3 respectively. However, it misses an event reported in the template matching catalog, placed around sample 800.

In Figure 2.19 – left panel, we report the application of the conservative model of EQT for a one-minute-long waveform containing multiple earthquakes. In the left panel of Figure 2.19, we note that EQT can correctly identify and discriminate three events occurring in 20 seconds (from sample 3000 to sample 5000), which feature magnitudes 1.0, 0.4 and 0.3 respectively. However, it does not recognize a low-magnitude event placed around sample 800, contained in the template matching catalog, where EQT does not provide any variation in the probability distributions. We also tested EQT on the same windows analyzed in Figure 2.18 that, despite being in the same amplitude scale as the left panel of Figure 2.19 due to the normalization, features multiple $M < 0$ earthquakes. We report the results in the right panel of Figure 2.19. The application of the conservative model of EQT does not provide indications of the presence of the events, apart from a very faint variation in the flat trend of the probability distribution for the event close to the right edge of the considered window. On the other hand, the application of the original model of EQT, reported in the bottom panel, can help in discriminating the earthquakes, with probability values for the P and S phase arrival times close to the threshold values from Mousavi et al., (2020) (0.3, reported as a black dashed line in the bottom panel).

The original model provides the P arrival time for the event close to sample 4000, while it reports probability below the threshold for the other events. Conversely, the S phases are provided for all the detected events. However, since the original model is too sensitive in providing detections without resulting in a massive dataset dominated by false detections, which might be difficult to be automatically ruled out, we preferred to obtain a smaller but more reliable starting catalog of machine learning detections to be used for detecting lower magnitude events through a similarity search. We thus propose the use of the original model of EQT as a machine learning picker rather than a detector for scanning continuous waveforms. This approach would provide a large number of arrival times on a validated set of detection catalogs.

Recently, several studies have shown that the performance of machine learning-based detectors might be affected by the position of the event within the analyzed window, penalizing events close to the edge of the window, leading to biases in the completeness of the resulting manual catalogs (Park et al., 2023; Pita-Slim et al., 2023). They demonstrated that even a small perturbation in the data (e.g. varying the position of the events of a few samples within the analyzed windows) causes the neural network to output different answers, not only on the classification probability, which might be also lowered below the detection threshold, but also on the location of the peaks, and thus the corresponding phase arrival time (Park et al., 2023). In Figure 2.20 we tested the application of the conservative model of EQT on the same window as the left panel in Figure 2.19, also extracting the one-minute-long waveforms 12 seconds after the start of the former window, featuring the events in a more central position within the analyzed window.

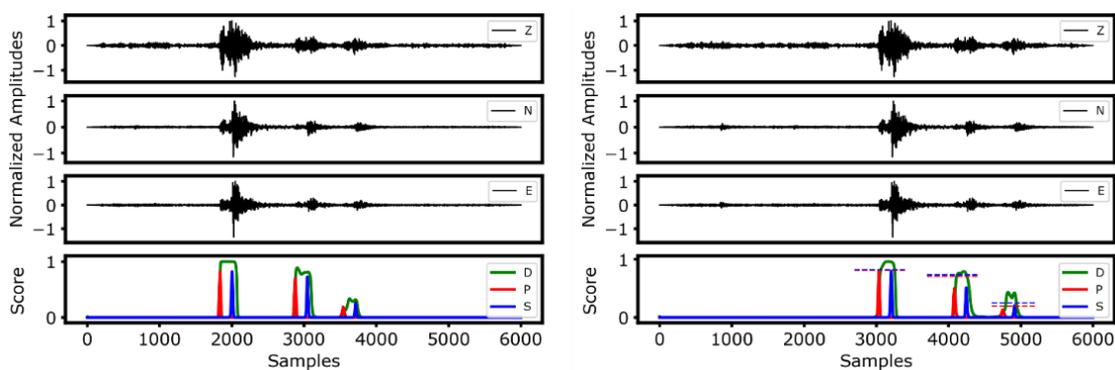


Figure 2.20) Comparison of the performance of EQT on events featuring different position of the waveforms within the one-minute-long considered windows. The window on the left panel starts at 2020-07-03 16:16:42 and contains the events in the central part of the window, while the right panel has been extracted from 2020-07-03 16:16:30 and contains the events closer to the right edge of the window. In the right panel, the peak probability scores obtained from the left panel are reported, with red and blue dashed lines marking the former P and S probability scores, respectively.

The comparison of the results for the two time-windows shows that, effectively, the probabilities provided by EQT for the events placed in a central position of the window (left panel of Figure 2.20) are higher than the those obtained for earthquakes close to the edge of the window (right panel of Figure 2.20, which represents the same waveforms discussed in Figure 2.19 – left panel). In the right panel of Figure 2.20, we also report the three peaks of the probability distribution predicted by EQT for the three events as for the left panel, marking with dashed red and blue lines the former scores for P and S waves, respectively. Although the P and S probabilities for the first, higher magnitude event do not differ significantly, due to the central position of the event in both windows, we observe a decrease of about 20 % in the peak probability scores for the following two lower magnitude events, close to the edge of the window in the right panel of Figure 2.20. The use of high overlap values among consecutive time windows to be fed to the neural networks has been proposed as a mitigation strategy for this issue, in order to ensure that events always feature an adequate position within the considered window (Park et al., 2023). In this analysis, we used an overlapping of 30 % among consecutive windows (Mousavi et al., 2020), lower than the values proposed by Park et al. (2023), which suggested overlapping even higher than 90 %. We can thus attribute some of the missed detections within our EQT catalog to this issue. However, despite being still affordable for the analysis of short datasets, the use of higher overlapping values can easily lead to higher computational times required for the analysis of massive datasets. Although we are aware of this issue, we recovered the missed detections through the template matching detection stage, confirming that the proposed strategy of the integration of machine learning and template matching can provide the most robust and complete seismic catalogs.

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, M_c and the b-value of the Gutenberg–Richter frequency magnitude distribution using the software ZMAP (Wiemer 2001) and considering the local magnitude to characterize the event size. The estimation of the statistical parameters using the INFO catalog is typically not possible for most of the considered sequences due to the small number of detected events (i.e. in some cases consisting of 10 earthquakes only). For performing a statistical analysis of the enhanced catalogs, we assigned a local magnitude value to each of the earthquakes detected within individual sequences, using amplitude ratio. We first selected one event in each sequence, referred to as the reference event: 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 ensuring a similar frequency content as for the lower magnitude events. 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 within the same sequence, we provided a magnitude estimation through the displacement amplitude ratio, assuming colocation:

$$M_l = M_{l_{ref}} + \log_{10} \left(\frac{A}{A_{ref}} \right) \quad (2.4)$$

where $M_{l_{ref}}$ is the magnitude of the reference event; A and A_{ref} are half of the maximum 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).

Considering the combined catalog for each sequence (obtained by combining EQT + TM and FAST detections), we obtain a magnitude of completeness ranging between $M_c -0.3$ and $M_c 0.4$ (in the local magnitude scale), with an average improvement with respect to the INFO manual catalog of 1.1 (Vassallo et al., 2012). Distinguishing the results from the different techniques, we find that the M_c for EQT + TM coincides with those obtained for the combined catalogs. Moreover, FAST and INFO + TM provide M_c 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 an M_c larger by 0.3 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 comparable values within uncertainties. Interestingly, we observe that, in general, EQT provides b-values systematically lower than the other two techniques, mainly driven by local incompleteness in the catalogs (Herrmann & Marzocchi, 2021). In other words, it seems that EQT progressively loses detections while approaching M_c , which leads to a biased population of magnitude bins and lower b-value estimates. This suggests the need to cover almost two to three orders of magnitude to estimate the relative rate of occurrence reliably for seismic events in sequences for the area. In Figure 2.21, we show the estimation of the statistical parameter of the Gutenberg-Richter law for the most populated seismic sequence (Rocca San Felice, ID 1), while in Table 2.2 we report the M_c and b-value for each sequence and enhanced catalog. For the considered seismic sequence in Figure 2.21, we observed a breakup from the linear trend at higher magnitudes. Moreover, this separation appears to be different according to the technique used for generating seismic catalogs. We remark that the magnitude of individual earthquakes in Figure 2.21 could vary among the techniques, since the stations used in the evaluation of the magnitude are those at which the event was detected. Moreover, the separation appears for less populated and less significant magnitude bins, whose extensions are comparable with the uncertainties on the magnitude estimates, and are not used to determine the parameters of the log-linear relationship.

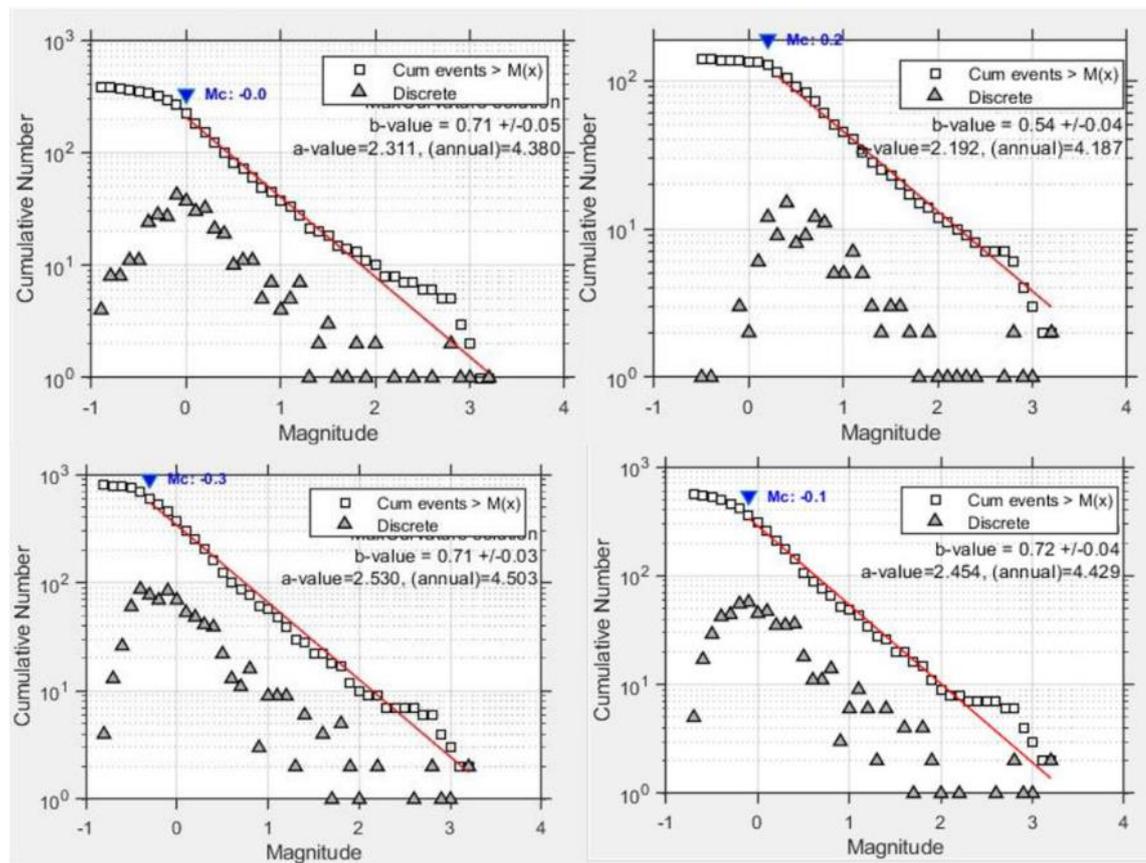


Figure 2.21) Gutenberg–Richter distribution for FAST ($M_c = 0.0$ and $b = 0.71 \pm 0.05$; upper-left-hand panel), EQT ($M_c = 0.2$ and $b = 0.54 \pm 0.04$; upper-right-hand panel), EQT + TM ($M_c = -0.3$ and $b = 0.71 \pm 0.03$; lower-left-hand panel) and INFO + TM ($M_c = -0.1$ and $b = 0.72 \pm 0.04$; lower-right-hand panel) catalogs.

Table 2.2) M_c and b -value estimation from the enhanced catalogs of each sequence

ID	M_l main	FAST	EQT	EQT + TM	INFO + TM
1	3.0	$M_c = 0.0$ $b = 0.71 \pm 0.05$	$M_c = 0.2$ $b = 0.54 \pm 0.04$	$M_c = -0.3$ $b = 0.71 \pm 0.03$	$M_c = -0.1$ $b = 0.72 \pm 0.04$
2	2.7	$M_c = -0.3$ $b = 0.60 \pm 0.05$	$M_c = -0.2$ $b = 0.51 \pm 0.07$	$M_c = -0.3$ $b = 0.68 \pm 0.06$	$M_c = -0.3$ $b = 0.62 \pm 0.06$
3	2.8	$M_c = 0.3$ $b = 0.62 \pm 0.11$	N.A	$M_c = 0.4$ $b = 0.76 \pm 0.13$	$M_c = 0.6$ $b = 0.65 \pm 0.16$
4	3.7	$M_c = -0.1$ $b = 0.76 \pm 0.08$	$M_c = 0.1$ $b = 0.54 \pm 0.08$	$M_c = -0.2$ $b = 0.65 \pm 0.07$	$M_c = -0.1$ $b = 0.70 \pm 0.09$
5	1.8	$M_c = 0.3$ $b = 0.83 \pm 0.10$	$M_c = 0.3$ $b = 0.75 \pm 0.10$	$M_c = -0.3$ $b = 0.73 \pm 0.05$	$M_c = 0.1$ $b = 0.83 \pm 0.08$

6	3.0	Mc = 0.3 $b=0.75\pm0.18$	Mc = 0.4 $b=0.53\pm0.13$	Mc = 0.2 $b=0.69\pm0.11$	Mc = 0.3 $b=0.8\pm0.2$
7	2.7	Mc = 0.4 $b=0.66\pm0.13$	N.A	Mc = 0.3 $b=0.78\pm0.14$	Mc = 0.5 $b=0.67\pm0.15$
8	2.8	Mc = 0.1 $b=1.16\pm0.11$	Mc = 0.9 $b=1.00\pm0.14$	Mc = 0.2 $b=1.26\pm0.13$	Mc = 0.2 $b=1.08\pm0.09$
9	2.9	N.A	N.A	Mc = -0.3 $b=0.60\pm0.09$	N.A
10	3.1	Mc = 0.2 $b=0.73\pm0.16$	Mc = 0.2 $b=0.53\pm0.09$	Mc = 0.1 $b=0.76\pm0.13$	Mc = 0.2 $b=0.71\pm0.16$

The considered seismic sequences occurred in different sectors of the Irpinia area. Picozzi et al. (2022a) showed that the b-value distribution in this area is not uniform and hypothesized that b-value 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 of Picozzi et al. (2022a). To this aim, we need to consider the moment magnitude of the detected events. We thus converted the local magnitude estimates into seismic moment ones using the empirical relationship $\log_{10} M_0 = 1.5 M_l + 10.55$ which has been derived considering the earthquakes in the INFO catalog (i.e. M_l and M_w values from Picozzi et al., 2022a). The M_0 can in turn be used to retrieve the moment magnitude M_w (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 to $b_{M_w} = \frac{1.5}{1.31} b_{M_l}$. As reported in Figure 2.22, which represents the b-value spatial distribution in Irpinia Picozzi et al. (2022a), the smallest b-values from Picozzi et al. (2022a) 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 3-D 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.

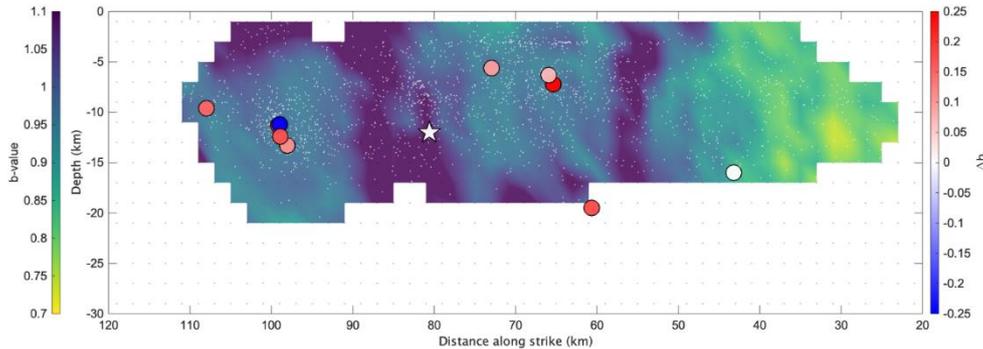


Figure 2.22) Map showing the spatial distribution of average b-values from bootstrap analysis (redrawn from Picozzi et al. 2022b). In the figure, we show the nucleation point for the 1980 Irpinia earthquake (white star) and Δb -value (see text for the definition, coloured circles) for the sequences analysed here.

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 analysing 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 the 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.

In the next sections, we will locate the earthquakes in the enhanced catalogs, using a relative location technique, for illuminating the fault segments where seismic sequences occurred (Section 3.2.1). Furthermore, we will investigate the source parameters for the seismic sequences assessing the seismic moment and the source size for the considered earthquakes (Section 4.2.1). Finally, we will define the stress release model of the sequences for understanding the triggering and the evolution of seismic events (Section 4.2.2).

2.3.2 Seismic monitoring using temporary dense array deployments

Near-fault observations can provide insights into the physical process interaction between fault slip activation, fluid presence/migration and seismicity production and processes

acting at different timescales that generate large earthquakes. As shown in Section 2.3.1, the integration of dense monitoring networks and advanced detection techniques can provide enhanced seismic catalogs of almost one order of magnitude larger than the revised manual catalogs, which can feature completeness magnitude closer or even lower than M_l 0 (Tan et al., 2021; Scotto di Uccio et al., 2023). With the aim of exploring the ultra-microseismicity in Southern Apennines, within the framework of an innovative experiment (DETECT, Bindi et al., 2021; Picozzi et al., 2022b), 200 seismic stations have been deployed in Irpinia for almost one year (September 2021 – August 2022), finely filling the gap of the less dense network of the Irpinia Near Fault Observatory. The 200 seismic stations have been organized in 20 arrays of 10 stations each, with an aperture of a few hundred meters on average, and featuring different geometrical configurations. Each array was equipped with one broadband sensor (station code 01), one 1-Hz sensor (station code 02) and eight 4.5 Hz short-period geophones (station codes 03 to 10). In Figure 2.23 we report the DETECT station distribution (red triangles), along with the stations of the Irpinia Near Fault Observatory (black triangles).

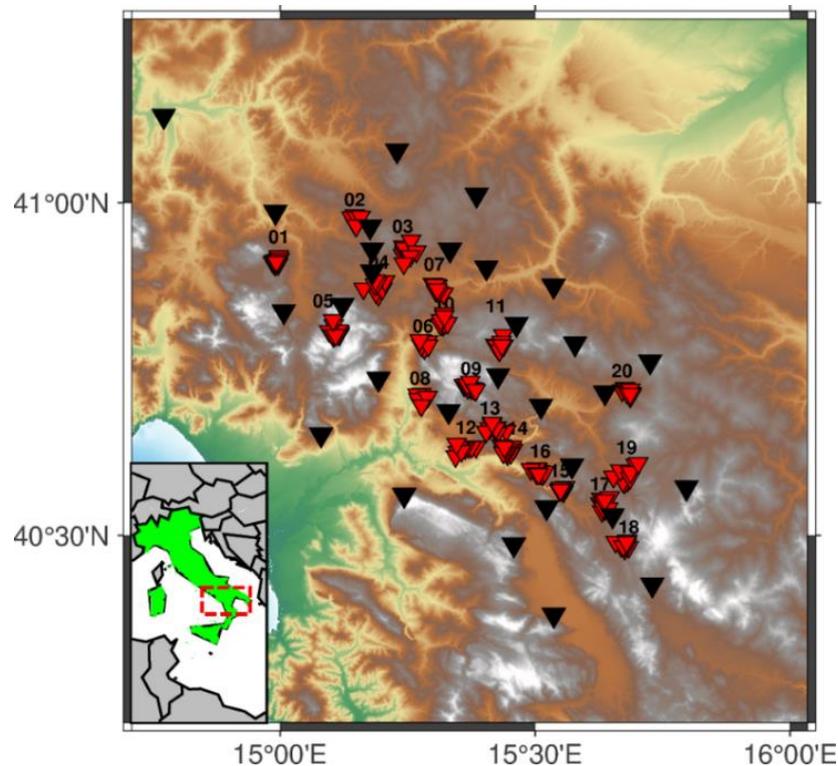


Figure 2.23 Distribution of the DETECT stations (red triangles), organized in the 20 arrays indicated with the corresponding code. Black triangles mark the location of the the Irpinia Near Fault Observatory stations.

The DETECT stations continuously recorded ground motion velocity with a sampling rate of 200 Hz. In Figure 2.24, we report the data availability during the whole DETECT experiment, reporting the fraction of recording stations as a function of time.

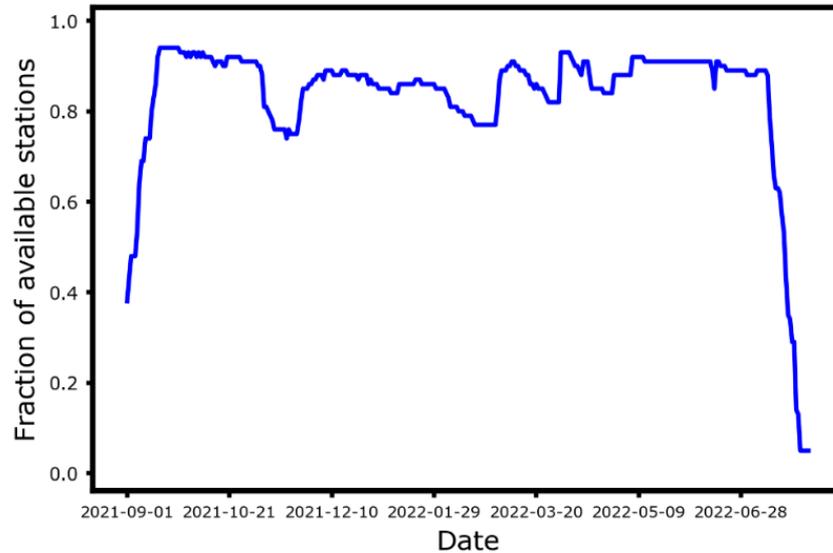


Figure 2.24) Data availability during DETECT experiment. Drops of the data availability close to the edges are due to the progressive installation and uninstallation of the stations respectively.

Despite the low data availability near the starting and the ending period of the experiment due to the progressive installation and uninstallation of the seismic stations, respectively, we achieved an average data availability higher than 80 % during the whole monitoring period. However, since the stations did not transmit waveforms in real-time, the records were stored locally in a system coupled with the stations and gradually downloaded directly on-field. This issue sometimes led to a lack of data due to the memory filling of the storage systems, which typically occurred with characteristic times of 3 months. This feature can be also noticed by local minima in the data availability distribution.

Earthquake detection followed the scheme proposed by Scotto di Uccio et al. (2023) for seismic sequences in the same region. The workflow was based on the use of the machine learning detector EQTransformer (Mousavi et al., 2020), which provides a wide set of templates for further similarity-based detection using the template matching technique EQCorrscan (Chamberlain et al., 2018). During the detection stage, the dense network was split into 6 subnetworks of 6 arrays each, with an overlapping of 3 arrays between consecutive subnetworks (Figure 2.25). Earthquake detection was performed independently for each subnetwork, integrating the declarations among the subnetworks according to the detection times.

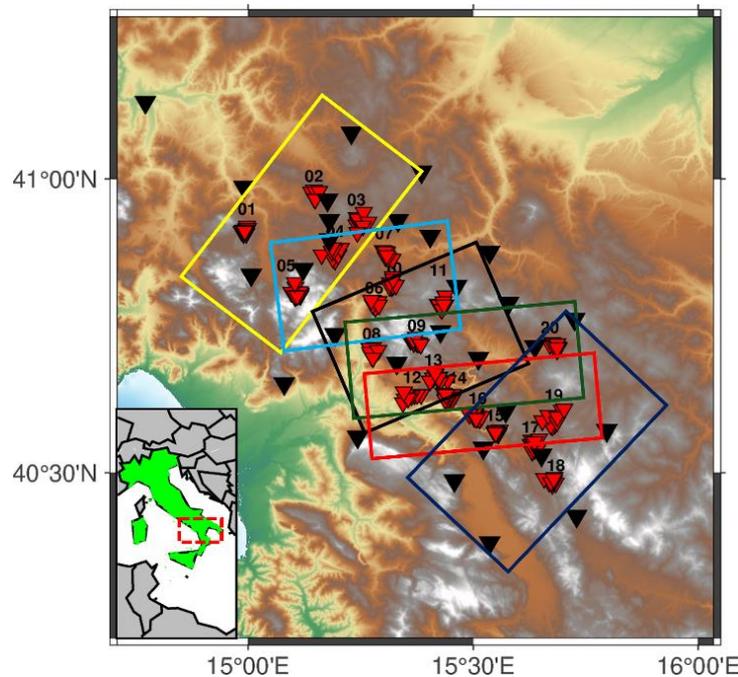


Figure 2.25) The six subnetworks extracted from the DETECT network. Each subnetwork consists of six arrays, with an overlapping of three arrays with the adjacent subnetwork.

This strategy allowed us to focus on the volume beneath the considered receivers, increasing the capability to detect local and very low magnitude events, and limiting the false declarations provided by coherent noise at distant stations. We used the same parameterizations for EQT and EQCorrscan as for Section 2.3.1 (Scotto di Uccio et al., 2023). We just modified the overlapping condition among adjacent time windows in EQT, increasing this value from 30% to 50%, following the results and the discussion of Section 2.3.1. For EQT, detections were declared when at least 5 picks were associated, within a single subnetwork, in time windows of 10 s, and then visually inspected.

The proposed detection strategy led to the declaration of 3563 real earthquakes. We remark that this value is comparable with the number of earthquakes in the manual catalog of INFO for more than 12 years of continuous monitoring (Figure 2.14). While the manual INFO catalog contains 443 earthquakes during the monitoring period of the DETECT experiment, the integration of ultra-dense monitoring and advanced detection techniques increased the existing catalog by a factor of ~ 8.2 , higher than the enhancing factors obtained with the same detection strategy on the ordinary seismic network (Scotto di Uccio et al., 2023). No events with a magnitude higher than M_l 3 occurred within the experiment, with just a M_l 2.8 occurring close, but outside the DETECT network. The high number of events in the DETECT catalog indicates that monitoring active

seismogenic sources with ultra-dense seismic arrays can strongly enhance the seismic catalog, even when the seismic activity is not intense. Considering the performance of EQT, the number of confirmed earthquakes was enhanced by a factor of ~ 4 compared to the manual catalog of the ordinary INFO network. Thus, the contribution of EQTransformer to the enhancement is almost twice as large as than what was observed by Scotto di Uccio et al. (2023) for seismic sequences. We owe this result both to the ability of the machine learning picker to identify very-low magnitude events at the array close to the earthquakes and to the use of higher overlapping values. However, we noted a slightly larger rate of false detections ($\sim 25\%$), mainly dominated by local transients observed at a single array, declared by the simple association algorithm of EQTransformer. On the other hand, the similarity-based detector provided more false positives compared to the former application, mainly driven by the high similarity of the noise at the nearby stations within the same array.

The obtained enhanced seismic catalog allows to track the time distribution of the seismicity during the monitoring period. In Table 2.3, we report the monthly number of earthquakes for the manual INFO catalog, based on the ordinary network, and the DETECT catalogs obtained applying EQT, and EQT + TM, respectively.

Table 2.3) Monthly detections for the manual INFO catalog, the EQT and EQT + TM catalogs for DETECT.

MONTH	INFO	EQT	EQT + TM
September 2021	43	221	349
October 2021	31	169	341
November 2021	24	123	223
December 2021	18	104	215
January 2022	28	123	215
February 2022	92	201	522
March 2022	43	131	216
April 2022	26	122	394
May 2022	39	192	374
June 2022	54	181	548
July 2022	45	99	256
TOTAL	443	1666	3653

Considering the monthly distribution, we note that the number of seismic events generally decreases from September 2021 to March 2022, while it increases from this latter month up to the end of the experiment. Due to the progressive uninstalling of the arrays during July 2022, this month is not fully representative of the observed trend. February 2022 appears to detach from the discussed behaviour. However, during this month, three

swarm-like seismic sequences (maximum M_l 1.9) occurred in Irpinia, with enhanced catalogs obtained from the integration of the detection techniques of hundreds of events each. The observed results appear to be consistent with the seasonal seismic trend observed in Irpinia, as extracted from the existing manual catalog. In Figure 2.26, we report the median and mean (in brackets) monthly number of earthquakes in the INFO manual catalog from January 2009 to September 2021, the starting time of the DETECT experiment.

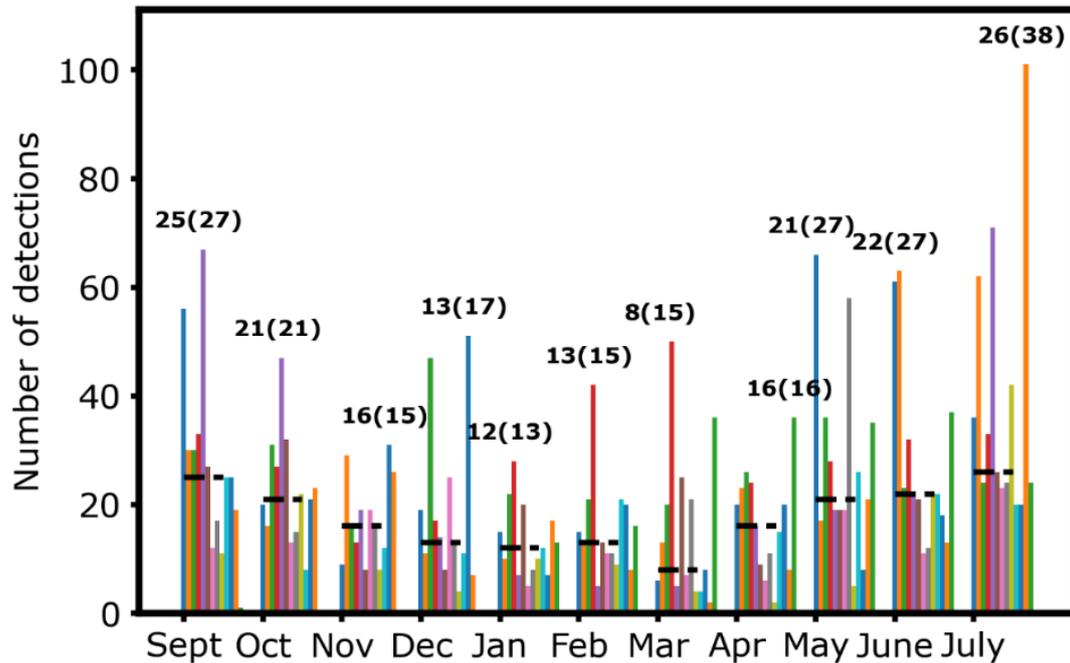


Figure 2.26) Median monthly number of earthquakes for each year from January 2009 to September 2021 (colored bars). The mean number of earthquakes for each month, which is marked with a dashed black line, is reported within the brackets, next to the median value.

Coherently with the trend observed for the DETECT experiment, we also identify a clear seasonal behaviour from the analysis of the manual catalog, with a decreasing and increasing seismic rate approaching and starting from the month of March, respectively. This feature agrees with the hypothesis of a seismic response to the hydrological stress perturbation in Irpinia, induced by the variations in the level of the karst aquifer (D'Agostino et al., 2018).

For assessing the statistical parameters of the enhanced catalog, we assigned a magnitude value to the detections. Conversely to the analysis of seismic sequences, in which we assumed colocation for events between each sequence due to the limited area where the

sequence evolved, in this framework the hypothesis of colocation cannot be held anymore. We therefore performed earthquake location (see Section 3.3.2), selecting the automatic phase arrival times from the detection stage, using NLLoc (Lomax et al., 2009) and the 1D-layer velocity model for the area proposed by Matrullo et al. (2013). We used the relation of Bobbio et al. (2009) for estimating the local magnitude of the earthquakes, considering the average peak-to-peak displacement amplitude on synthetic Wood-Anderson records. In Figure 2.27, we report the Gutenberg-Richter law for the INFO catalog from 2008 and the DETECT catalog from September 2021 to July 2022.

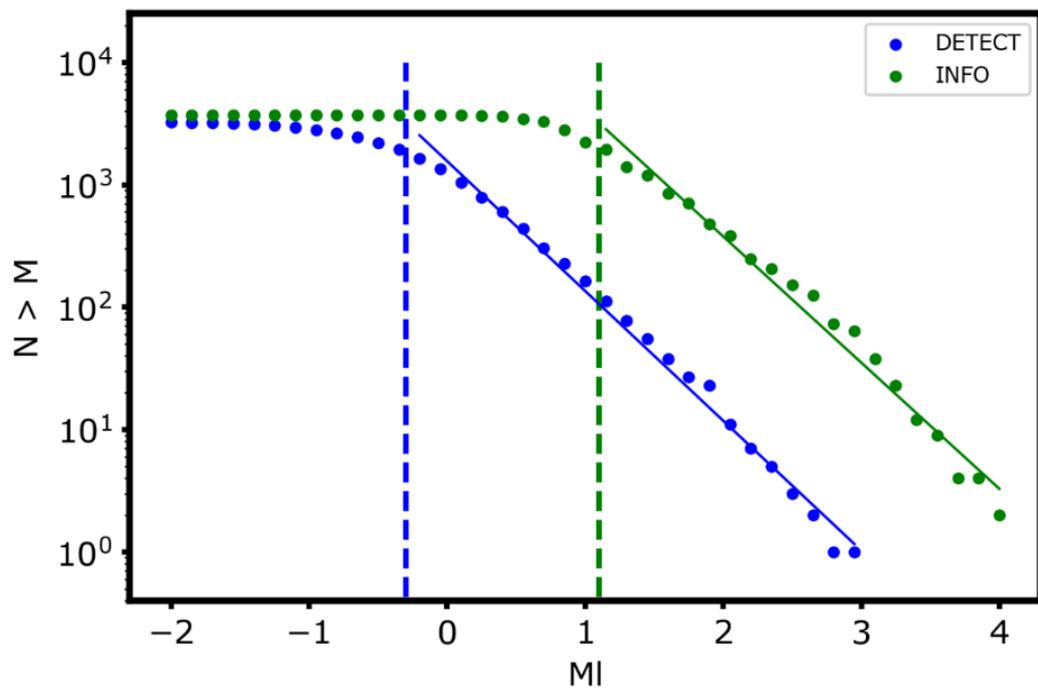


Figure 2.27) Gutenberg- Richter law for the manual INFO catalog from 2089 (green) and for the DETECT catalog (blue). The dashed blue and green vertical lines mark the completeness magnitude of the DETECT and INFO catalogs, respectively.

The analysis of the magnitude of completeness for the two catalogs shows that monitoring the seismogenic sources in Irpinia with temporary dense array deployment allows to decrease the completeness of the catalog by almost 1.5 orders of magnitude, moving from M_c 1.1 (Vassallo et al., 2011) to $M_c = -0.3$. Moreover, the comparison of the b -values for both catalogs indicates similar slope of the Gutenberg – Richter law, being $b_{INFO} = -1.04 \pm 0.08$ and $b_{DETECT} = -1.03 \pm 0.10$ for the INFO and DETECT catalogs respectively. On the other hand, the analysis of the slope of the Gutenberg-Richter

relationship for seismic sequences in the DETECT catalogs still reveals lower b-value for the seismic sequences as compared to the background seismicity, in agreement with the former results of Scotto di Uccio et al. (2023). In the next sections, we will perform accurate earthquake location for the DETECT catalog, interpreting the seismicity structures highlighted by the microseismicity (Section 3.3.2).

2.3.3 An enhanced catalog of repeating earthquakes in Southern Italy

Conversely to the seismic sequence, in which earthquakes occur with high seismic rate typically over co-located sources, repeaters are defined as earthquakes displaying very similar waveforms when recorded at the same sensor but repeating over a timescale longer than for seismic sequences (Uchida & Bürgmann, 2019).

Repeaters nucleate stably on a patch of the fault, so that path and site effects between source and station are common (Vidale et al., 1994; Nadeau et al., 1995). Due to their stability over time, the source of the repeaters is usually modeled as the repeated activation of an asperity surrounded by domains with different frictional properties. The size of the asperities producing the repeaters can be as small as a few hundred meters or even less, making the analysis of the repeaters a powerful and unique tool to map the local interface stress state at depth (Uchida et al., 2012). Repeaters usually show a temporal pattern with typical recurrence intervals of a few years or less and are often modeled as nucleating over locked patches distributed on an otherwise creeping fault zone.

The main goal of this work is the identification and characterization of repeating earthquakes in the volume of the 1980 Irpinia earthquake, eventually aiming at the identification of small-scale segments of the fault system generating repeater events (Palo et al., 2023b). In this framework, we selected the densest cluster of near-repeaters defined in Palo et al. (2023a), which grouped similar events contained in the manual INFO catalog from February 2008 to March 2020, for generating an enhanced catalog of repeating earthquakes using the events of this cluster as templates, scanning the continuous seismic signal with a similarity search approach (Chamberlain et al., 2018; Scotto di Uccio et al., 2023).

In this analysis, we thus selected the cluster ID 14 in Palo et al. (2023a), composed of 20 earthquakes (16 out of the 20 earthquakes were located in Palo et al. 2023a, represented with blue dots in Figure 2.28), and we focused on the INFO station where all events of the cluster were detected (labelled as SNR3 and marked with a red triangle in Figure 2.28). Black dashed rectangles mark the position of the main seismogenic source (DISS, 2021), responsible for the multiple ruptures of the 1980 M 6.9 earthquake. The daily velocimetric recordings at the reference station were scanned from 1 January 2012 to 22 April 2022, which is the time interval of the available continuous signals at this station.

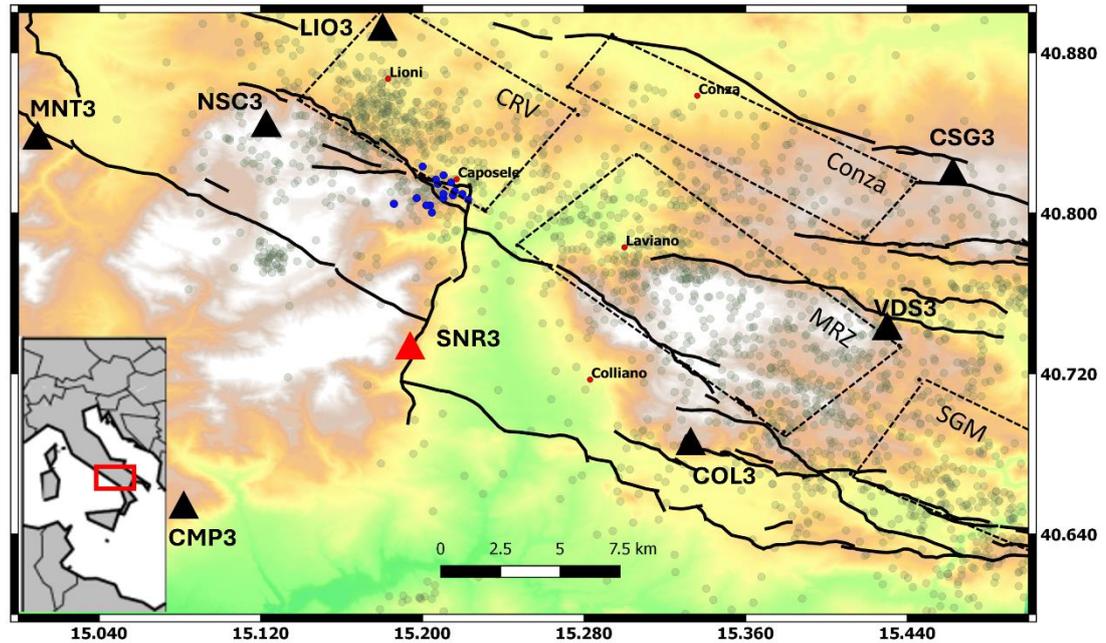


Figure 2.28) Location of the cluster used in this analysis (Palo et al.,2023a), reported as blue dots. The red station marks the position of the SNR3 station, while the dashed black lines mark the fault segments activated during the 1980 M 6.9 earthquake.

To homogenize the preprocessing operation to the initial parameterization of the clustering analysis in Palo et al. (2023a), we extracted the templates bracketing the vertical component of velocity records from 0.4 s before to 1 s after the manual identification of P arrival times, downsampling the traces to 50 Hz. As in Palo et al. (2023a), templates were filtered in the frequency band of 1–20 Hz using a 4th-order Butterworth bandpass filter. Continuous data were preprocessed matching the parameterization adopted for generating the templates. In both cases, preprocessing also included a linear detrend and a demeaning of the signals. An event was declared when the cross-correlation (CC) between the portion of the continuous window and a template overcame 0.8, which was the threshold fixed in Palo et al. (2023a). After refining the detection list, we obtained a catalog of 12 potential repeaters, increasing by more than 50% the content in the initial cluster of similar earthquakes. In Figure 2.29, we report the vertical component of the velocity records for the 12 detected repeaters, marking with red the waveform of the master event and reporting the CC coefficient for each slave event.

For evaluating the area activated by the occurrence of the repeaters, we compared the S-P travel times for the new events with the corresponding values of the master events. In particular, we can relate the differences in the propagation time to differences in the

earthquake location, since $\Delta t_{i,M} = \frac{\gamma-1}{V_P} \Delta r_{i,M}$, where $\Delta t_{i,M} = \Delta t_{P,i} - \Delta t_{S,i} = (t_{P,i} - t_{P,M}) - (t_{S,i} - t_{S,M})$ and $\Delta r_{i,M}$ represent the spatial distance, projected along the raypath to station SNR3, between the i -th event and the master event. Given the location results for the cluster in Palo et al. (2023a), which placed the events within [11-12] km of depth, we assumed $V_P = 6$ km/s and $V_P/V_S = \gamma = 1.95$ (Matrullo et al., 2013; Amoroso et al., 2014). $\Delta t_{P,i}$ was estimated as the time shift that maximizes the CC function between the master event and the i -th potential repeater, both windowed starting 0.1 s before the P onset, as explained above. Similarly, $\Delta t_{S,i}$ was estimated in the same way, but windowing, in this case, 0.2 s before S onset. All the solutions but one feature $\Delta t_{i,M}$ in the range of $(-0.05$ s, 0.05 s), which corresponded to a $\Delta r_{i,M}$ within 300 m from the hypocenter of the master event.

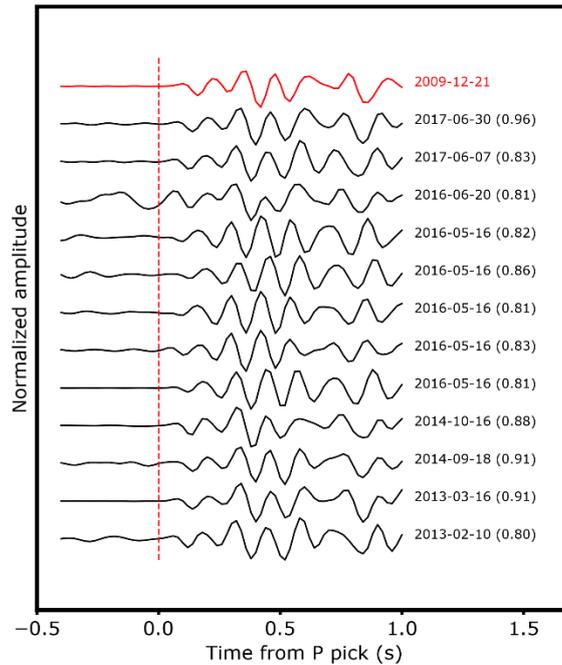


Figure 2.29 Set of potential repeaters extracted from continuous data for cluster 14 of Palo et al. (2023a) at the SNR3 station. The red waveform marks the master event used for refining the detection list. CC coefficient with the master event is reported in parenthesis for each detection.

Following the approach of Scotto di Uccio et al. (2023), we evaluated the local magnitude for the set of potential repeaters assuming colocation with a located event within the cluster, resulting in magnitude values ranging from M_l 0.4 and 1.7. Using the relation between moment magnitude and local magnitude proposed in Malagnini & Munafó (2018), $M_w = \frac{2}{3} M_l + 1.14$, we inferred the moment magnitude M_w (and then the seismic

moment M_0) for the set of repeaters. To associate a size to the source fault of each event from the moment magnitudes, we assumed an instantaneous and homogeneous stress release on a circular fault of size a , which is connected to the seismic moment and the stress drop $\Delta\sigma$ through $M_0 = \frac{16}{7} \frac{\Delta\sigma}{a^3}$. The knowledge of the source radius allows, finally, the estimation of the coseismic slip s , through:

$$\Delta\sigma = \frac{7\pi}{16} \mu \frac{s}{a} \quad (2.5)$$

where μ represent the rigidity coefficient of the medium. In this evaluation, we used the average stress drop inferred for the events in the cluster in Palo et al. (2023a), which corresponds to $\Delta\sigma = 4.0 \pm 3.3$ MPa. The cumulative slip from this the average stress drop value was about 35 mm, and the time distribution of the events show groups of repeaters occurred every 1.2–1.8 years.

This analysis opened the way to explore the possibility of constraining small-scale slip histories on the Irpinia fault by low-magnitude repeating earthquakes. In this direction, this approach can be extended to other clusters of near-repeaters to potentially identify other small-scale fault domains whose existence is suggested by the different fault sizes between the source volumes of the original set of near-repeaters used as a template and the inferred catalog of repeaters.

2.3.4 Monitoring microseismic activity in the gas storage area of Collalto (Northern Italy)

Seismic monitoring in areas where induced earthquakes could occur is a challenging topic for seismologists due to the generally very low signal-to-noise ratio. Seismic monitoring aims to characterize the spatio-temporal evolution of the seismicity in a sub-surface volume where industrial exploitation activities take place with the aim of discriminating the natural seismicity from the anthropogenic one (i.e., induced) and eventually intercepting variations in the background seismicity rate that, if needed, will guide the re-modulation, interruption, and restart of industrial activities. Indeed, tracking the microseismicity in time after fluid injection and accurately locating it can allow one to detect pore pressure changes and intercept migration fluid patterns (Goertz-Allmann et al., 2011). High-sensitivity monitoring networks have been proposed for detecting microseismic events, such as the one deployed around the Collalto underground gas storage in north-eastern Italy (Priolo et al., 2015). The Collalto seismic network (Rete Sismica di Collalto, RSC) aims to monitor the natural and induced seismicity potentially related to the industrial activity of the Collalto gas storage facility. It is composed of ten seismological stations equipped with borehole seismometers with periods varying between 10 s and 120 s operating at a sampling rate of 200 Hz, and accelerometric sensors at the surface of five sites, resulting in a completeness magnitude M_c for the whole area

of 0.6 in local magnitude (Romano et al., 2019). The Collalto gas field is a natural, depleted reservoir with a working gas storage capacity of approximately 600 million standard m^3 , covering an area of almost 89 km^2 (<https://www.edison.it/it/centrale-stoccaggio-di-collalto>). The field is equipped with 17 active wells through which the gas is seasonally injected into the reservoir during the April–October period and extracted during the November–March period.

In this analysis, we aim to test how and to what extent a small aperture, low-cost seismic array combined with the advanced microseismicity detection technique FAST can locally improve, e.g., in terms of the magnitude of completeness, the performance of a permanent, high-quality, less-dense seismic network, in enhancing the local earthquake catalog (Scala et al., 2022). For this purpose, we installed a seismic array, with a 2 km maximum aperture, composed of eight seismic stations located in the area near the gas storage site in Collalto, with two stations co-located with two seismometers of the RSC network. The temporary network, which is reported in Figure 2.30 along with the RSC network, was installed in wintertime and was operative for about one month (from 8th January 2021 to 31st January 2021), collecting continuous velocity records with a sampling rate of 250 Hz.

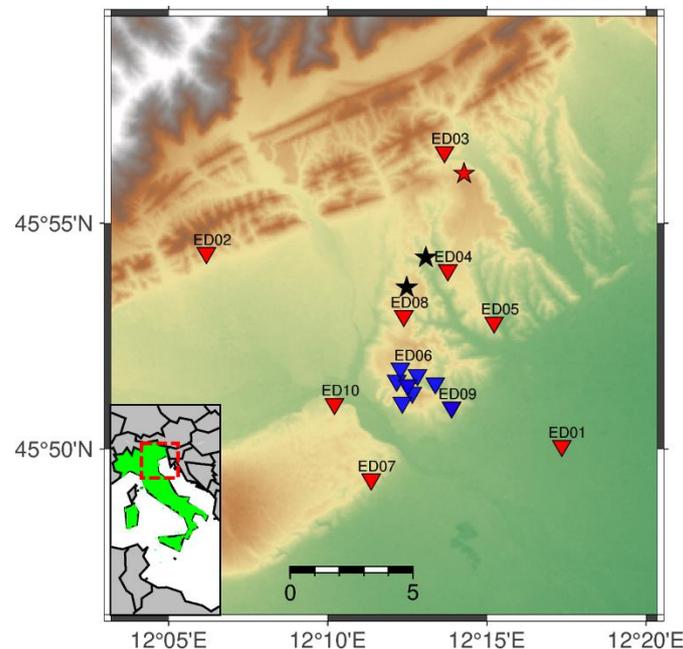


Figure 2.30 Location of the experimental seismic array (blue markers), along with the RSC seismic network (red markers). The black stars mark the M_l 0.1 and 0.0 Conegliano earthquakes, while the red star refers to the M_l 0.4 Vittorio Veneto earthquake. All these events were included in the RSC catalog.

In the preprocessing operation for the autocorrelation detection, we replaced gaps and zeroes in the data stream with Gaussian uncorrelated noise. Moreover, the waveforms were bandpass-filtered between 1 and 20 Hz using a 2-poles acausal Butterworth filter and decimated to 50 Hz to dump the high frequency noise and improve computational efficiency. We adopted the same parameterization for feature extraction in FAST as for Scotto di Uccio et al. (2023), selecting the detections only if similarity was found in least at 4 stations. The transient list was further investigated in terms of first arrival times and duration, for ruling out eventual false declarations. For the Collalto array, several detections were removed because of low apparent velocity across the array, long duration, and almost monochromatic character, mainly related to anthropogenic activity and/or weather effects. After this last step, the catalog was further enhanced, by including all the transients associated with fewer than 4 stations but being similar, at least at 2 stations, to accepted earthquakes.

This strategy led to the detection of 38 transients. Among them, 20 transients were discarded during the post-processing association phase due to unrealistic apparent velocity within the array. Indeed, they are characterized by signals having short durations and a large difference between the arrival times at the different array nodes, even higher than 10 s. A visual inspection of the detections confirmed that these waveforms were not seismic signals but were rather characterized by either monochromatic wave trains or the superposition of two monochromatic components at close frequencies. At the end of the procedure, we obtained a list of 18 detections.

During the experiment, a seismic sequence occurred in the district of Ravenna, at an epicentral distance of about 230 km from the array centre, with the largest event having magnitude M_w 4.3. Inspecting the detected events, FAST declared five different transients that can be associated with three events of this sequence. The number of transients (5) is larger than the number of associated events (3) since two events were detected twice (FAST separately declared body-waves and later surface waves), as the difference in phase travel times was higher than the time window defined for the generation of fingerprints.

The final catalog is composed of 16 seismic events, consisting of 7 local earthquakes that occurred around Collalto, while the remaining ones are regional or teleseismic events. Focusing on the local events, we found two detections located close to Conegliano (black stars in Figure 2.30), included in the RCS catalog (the closest earthquakes to the array), featuring magnitudes M_l 0.1 and 0.0, respectively. Two more events were detected by FAST that were not contained in the RCS catalog. These events were declared because of their reciprocal similarity and their similarity with the M_l 0.1 Conegliano event. They both feature a very low signal-to-noise ratio on raw waveforms ($SNR \sim 1.5$) and cannot be located using the array records. However, considering the waveform similarity with the

Conegliano events, we assumed these events are colocated and provided an estimate of the magnitude (M_l -0.4 and -0.2 respectively). In Figure 2.31, we report the vertical component of velocity records, filtered in $[1 - 20]$ Hz, for the Conegliano M_l 0.1 (left panel), and the FAST detections M_l -0.4 and -0.2 , in the central and right panels, respectively.

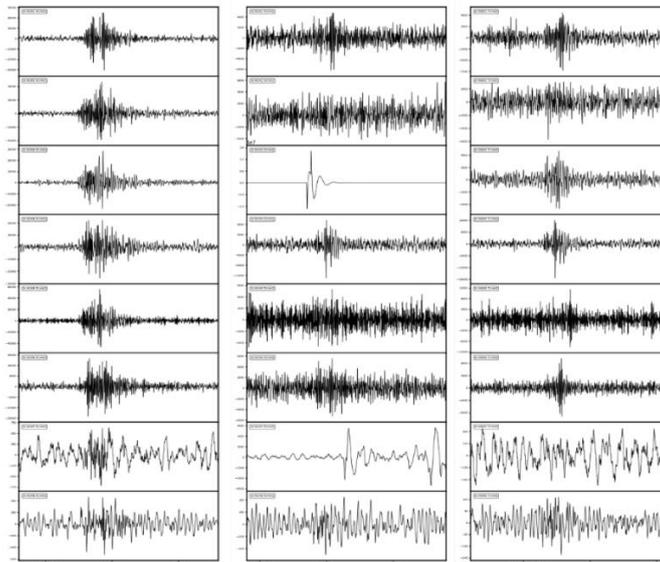


Figure 2.31) Vertical component of the velocity records, filtered in $[1 - 20]$ Hz, for the Conegliano M_l 0.1 (left panel), and the FAST detections M_l -0.4 and -0.2 , in the central and right panels, respectively.

In addition, FAST also correctly identified the M_l 0.4 event located at Vittorio Veneto (red star in Figure 2.30), which was also included in the RCS catalog, and two further events similar to it, whose magnitudes were estimated to be -0.4 for both earthquakes. It is worth noting that the negative magnitude events of Conegliano and Vittorio Veneto occurred before the corresponding larger-magnitude events. In particular, the Conegliano events occurred ~ 15.5 and ~ 14 h before the M_l 0.1 event, while the Vittorio Veneto doublet occurred ~ 6.5 h and ~ 13 min before the origin time of event M_l 0.4. Thus, despite the small number of these new events, we can interpret them as foreshocks of the two larger-magnitude reference events.

The limited duration of the experiment did not allow us to fully estimate the capability of the permanent and temporary deployment in monitoring the microseismicity with respect to the storage area target. However, starting from our results, in the following, we propose a strategy to infer the smallest magnitude that could have been detected in the study area. We investigated the noise distribution in the waveforms processed through the similarity search of FAST. For this analysis, the data stream at all stations of the array was

subdivided into one-minute-long windows, filtered in the band 1–20 Hz, and the noise level was estimated as the RMS amplitude of the whole signal. For local events detected by FAST, we computed the signal-to-noise ratio (SNR) on the horizontal components of the stations associated by FAST. Considering 7 s-long windows before and after the origin time as the noise $N(t)$ and the signal $S(t)$ contributions, we computed the Fourier transforms of both time series $N(f)$ and $S(f)$. The SNR is defined as S/N , where:

$$S = \int_{f1=1 \text{ Hz}}^{f2=20 \text{ Hz}} S(f) df \quad (2.6)$$

and an analogous representation stands for the noise. For the three largest events of the two clusters (M_l 0.4, 0.1 and 0.0), the SNR ranges between 4 and 6 on both horizontal components within the array, while for the smallest ones ($M_l = 0.4, -0.4$ and -0.3) the SNR ranges between 1.01 and 1.67 with a mean value $\langle SNR \rangle = 1.34$, which is assumed to be limit for the detection capability of the array system processed by FAST. Around this level, we have detected events of minimum magnitude $M = -0.4$ at a maximum hypocentral distance $R = 13$ km. Fixing this SNR level and reducing the distance in the range between 2 km and 3 km, to account for events potentially occurring near the gas-storage area, we can estimate the minimum detectable magnitude using the local magnitude scaling law proposed by Bragato & Tonto (2005), valid for north-eastern Italy and considering the overall level of noise and the estimated quantity $\langle SNR \rangle$. Considering that the Collalto gas-storage has an average depth of 1.5–1.6 km, and taking into account a range of reference distances between 2 km and 3 km from the array stations, we have retrieved an average limit magnitude $M_{l_{min}}$ between -0.8 and -0.6 , when we select a noise level equal to the 90th percentile of the noise distribution. Thus, we can state, with a 90% level of confidence, that no events with local magnitude larger than -0.6 have occurred within the study area during the experiment.

2.3.5 An automatic detection strategy for continuous seismic monitoring

Earthquake identification in continuous data for monitoring the seismic activity is typically a routine task performed by expert human operators. Despite the expertise of the operators, earthquake detection, but especially the identification of phase arrival times, is strongly affected by the subjective ability of the operator to discriminate the event within waveforms dominated by noise records, eventually leading to inhomogeneous catalogs. Moreover, in case of seismic sequences and higher seismic rate epochs, the monitoring effort requested for the operators might easily become overwhelming. Since catalogs generated using advanced detection techniques have been shown to be robust and even richer than the manual catalogs, also in case of seismic sequences, we hereby present a

novel detection strategy for the daily monitoring of the seismic activity in Southern Italy, complementing, but soon aiming to substitute, the work of human analysts.

In Southern Italy, the Irpinia Near Fault Observatory (INFO) manages the Irpinia Seismic Network (ISNet, <http://isnet.unina.it>) made up of 31 seismic stations (red triangles in Figure 2.14), equipped with strong-motion accelerometers and weak-motion sensors to be sensitive to microseismic events. While an automatic energy-based detector analyses ground motion records in real-time, typically detecting the higher magnitude events, each day a network operator visually scans the continuous data searching for missed earthquakes in the previous day. An event is included in the manual catalog if at least 4 phases have been picked on the records, including one S phase, and at least one station presents both P and S picks. After location, phase residuals are required to be smaller than 0.5 and 1.0 s, respectively. However, several factors, such as the extension of the time window for visually inspecting the continuous data (often 7 to 9 minutes), highly noisy data on windy days or even just the level of attention of the operators during the scanning stage might bias the detection capability. In Figure 2.32 we report a 7-minute-long window as it appears during the visual inspection of the continuous waveforms. This figure shows the vertical component of the velocity records (or acceleration records for stations equipped by strong-motion sensors only), spatially ordering the stations from North to South direction. The selected window, at first glance, does not appear to contain any clear and coherent seismic signal across multiple stations.

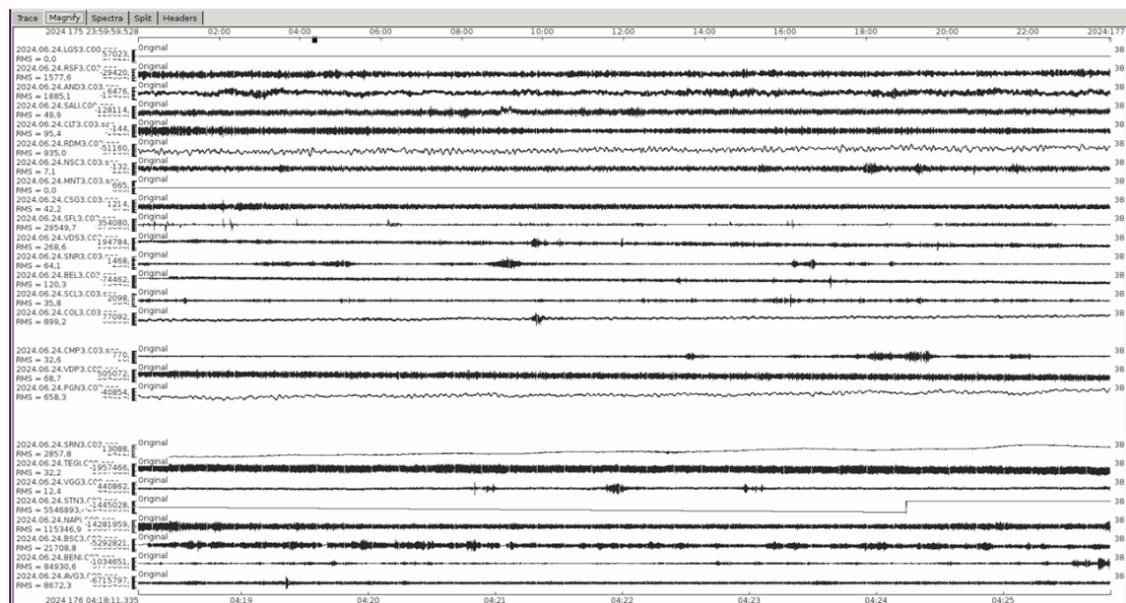


Figure 2.32) Panel reporting a typical 7-minute-long window selected for visual inspection by network operators. Each waveform represents the vertical velocity (or acceleration) records for a specific station, which are ordered from North to South direction.

To homogenize the detection performances and to limit the subjective bias, we built up a daily detection strategy based on PhaseNet (Zhu & Beroza, 2019), guided by the results of the comparison of machine learning-based detectors (Figure 2.9). As observed in the analysis of Section 2.2.1, PhaseNet appears to be the most promising technique to be implemented stand-alone, providing an encouraging trade-off between the number of detected events and the number of false positives. For associating the seismic phases provided by PhaseNet on the extended ISNet network, we used REAL (Zhang et al., 2018), a grid-search associator for relating individual phase arrival times to seismic sources, using theoretical travel times generated from a selected velocity model. For this application, we selected the velocity model proposed for the Irpinia region by Matrullo et al. (2013). In this workflow, we declared an event when at least 5 picks were associated, with at least 3 P and 2 S arrival times, and 2 stations provided both phases. It is worth noting that these criteria are stricter than those required for the visual identification of the earthquakes, eventually leading to some missed detections. However, on the other hand, this parameterization was chosen after some tests for balancing the number of missed and false declarations, being more numerous when using the same criteria as for visual identification.

The role of the automatic detection strategy is, up to now, to complement the visual inspection of the continuous waveforms, providing, before the start of the manual research, a list of potential earthquakes that occurred in the previous day. Moreover, the predictions of the phase arrival times are already set on the raw waveforms, extracted around the detection time for each declaration in the list. The network operator still inspects the continuous waveforms for missed events, eventually adding them to the detection list. Finally, before proceeding with earthquake location, the operator manages the provided phase arrival times by adding, refining or eventually discarding automatic picks. The complementing detection strategy has been operating since 27th February 2024, in a period in which the seismic rate is often observed to increase as a function of time (Figure 2.26). In this section, we summarize the results obtained up to 27th August 2024 (167 days).

In the considered period, the automatic detection strategy declared 685 events, 535 of which were real earthquakes (78%), while 150 detections did not represent seismic events (22%). 391 earthquakes were effectively located (74% of the real earthquakes), whereas, for the remaining percentage either the location uncertainties did not match the imposed quality threshold or there were not enough phase arrival times available after the manual revision of the picks. Considering the number of missed detections, the automatic workflow did not declare 19 earthquakes that were manually identified within the continuous data. However, the related percentage for the missed event is around 5% as compared to the number of detected and located earthquakes, respectively.

To highlight the main improvement resulting from the application of the automatic detection strategy, we can recall the time window shown in Figure 2.32. Despite no clear events appearing from the visual inspection of the continuous waveforms, an automatic detection was declared at 04:21:17, reporting phase arrival times at the stations VDS3, SCL3 and COL3. The filtered waveforms are shown in Figure 2.33 for the mentioned stations.

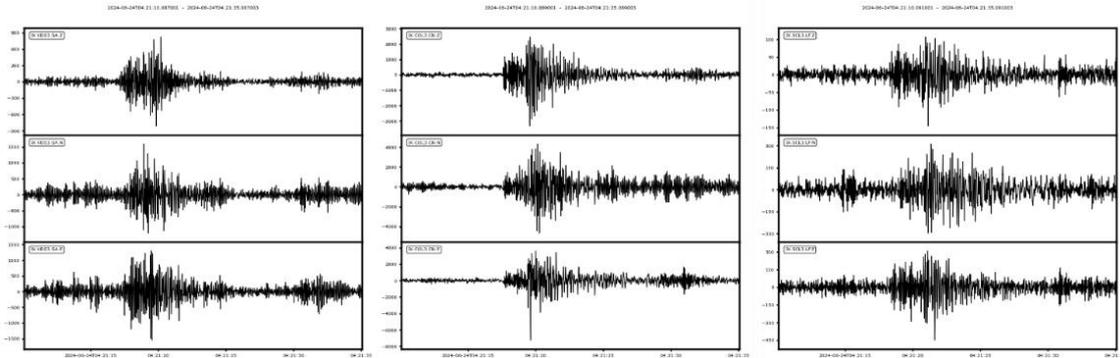


Figure 2.33 [1 -20] Hz filtered waveforms for station VDS3, SCL3 and COL3. The event was not visually identified from the representation as in Figure 2.32 but was detected by the automatic detection strategy.

When inspecting the waveforms for the detected event provided by the automatic strategy, the filtering operation increases the visibility of the event, allowing the operator to refine the phase arrival times. The event was finally located by selecting 3 P and 2 S phase arrival times, leading to the insertion in the catalog of a M_l 0.3 earthquake located at 3.0 km depth that would have been lost without the implementation of the automatic detection strategy.

To assess the reliability of the arrival times identified by the automatic detection strategy, we compared the phase arrival times before (automatic declarations) and after the manual refinement. In the following Figure 2.34, we report the time residuals between the automatic PhaseNet picks and the refined manual picks, the latter being used for the final earthquake location (around 2500 values). We observe consistency between automatic and manual phase arrival times. The distributions for P and S pick time residuals both feature mean values consistent with zero, suggesting that the automatic phases are accurate enough to obtain reliable earthquake locations even without the refinement step, leading to a complete automatization of the workflow. While we observe a narrower distribution for the P picks, we note wider dispersion for the (more complex) S phases. This feature is not surprising, since identifying the correct onset for the S picks is typically more challenging than for the primary waves, due to the lower signal-to-noise ratio and the complexity of the medium that might generate secondary phases. Thus, while we can assume the manual P pick as the ground truth for the arrival time of the primary wave,

for the larger time residuals in the S phases it is not trivial to distinguish whether the higher value results from an incorrect estimation of the automatic or the manual pick. As an example, we show in Figure 2.35 the results for the phase picking for a M_l 0.4 earthquakes, reporting with a green dashed line the automatic Phasenet pick and the corresponding manually refined arrival time, using a blue dashed line.

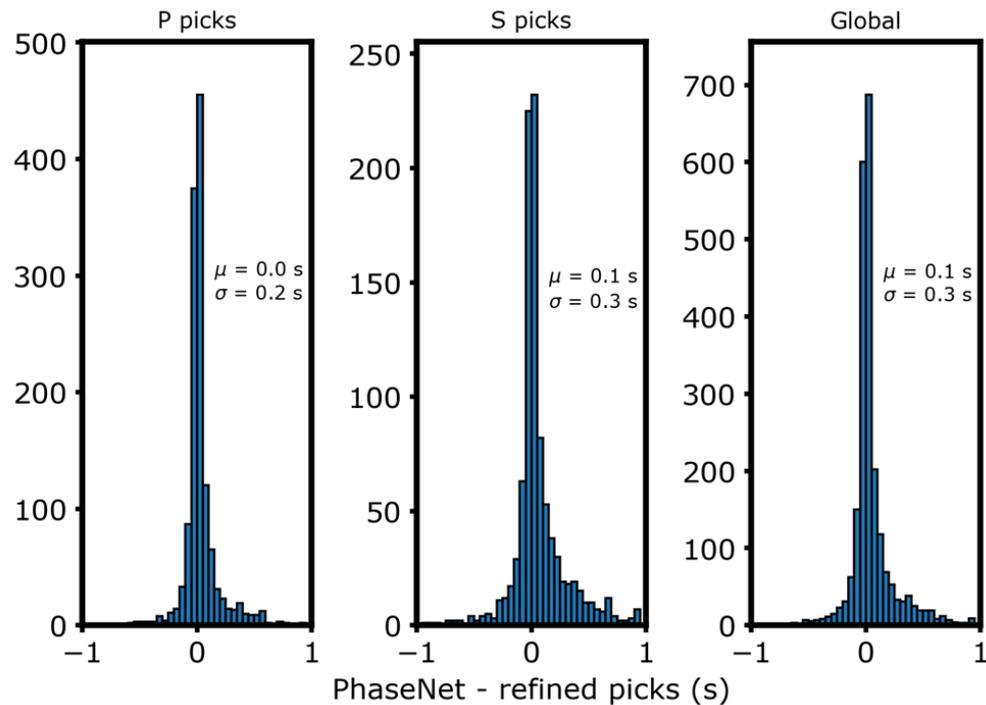


Figure 2.34) Distribution of the time residuals between PhaseNet automatic picks and manual refined arrival times for the P (left panel), S (central panel) and combined phases (right panel). Both P and S phases declared by PhaseNet are compatible with the manual arrival times, with a larger dispersion observed for the S phases.

While we observed coherent arrival times for the P phase between the automatic and the manual picks, we noted a larger residual for the estimation of the S arrival time. In particular, the manual arrival time is placed 0.375 s after the declaration of PhaseNet, this value being close to the standard deviation of the S residual distribution. However, looking at the bottom panel of Figure 2.35, we note that the PhaseNet estimation is closer to the correct arrival time of the S phase, as compared to the manual identification, matching the change in amplitude and frequency characteristic for the arrival of the S phase. This feature is even more clear in the following Figure 2.36, which is a close-up of the East component around the S arrival time.

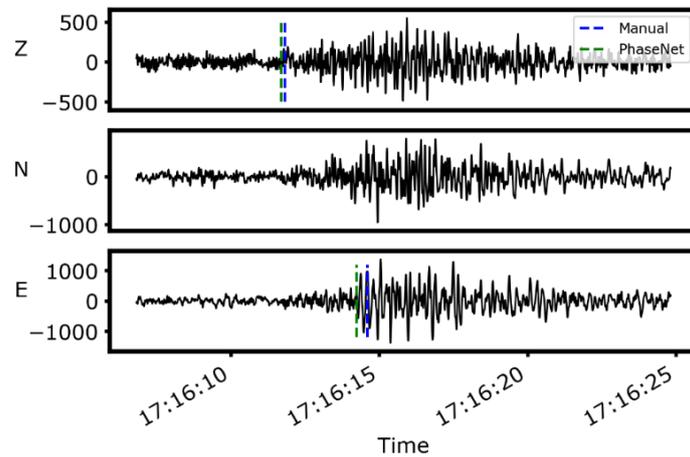


Figure 2.35) Automatic and manual identifications of phase arrival times for a M_l 0.4 earthquake, reported as green and blue dashed lines, respectively. While the P arrival time is coherent among the two strategies, S arrival times feature a time residual of 0.375s, with the PhaseNet picks preceding the manual identification.

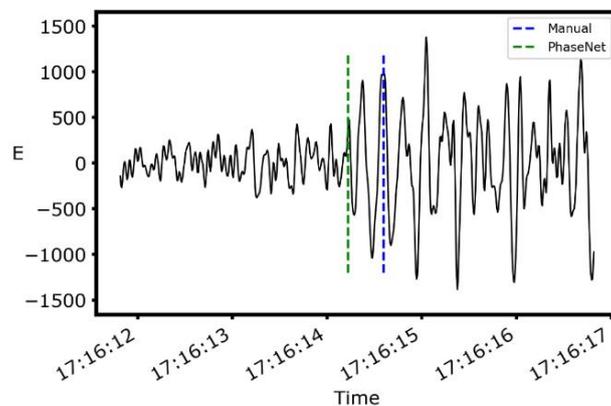


Figure 2.36) Zoom of the bottom panel of Figure 2.35, around the arrival of S wave. In this representation the PhaseNet pick (green dashed line) appears to be more reliable as compared to the manual identification (blue dashed line), correctly identifying the change in amplitude and frequency characteristic for the arrival of the S phase.

Thus, despite the higher residual, the most reliable S arrival time is provided by the automatic detection strategy, warning about the assumption of the manual identification as ground truth given the complexity of the waveforms, which typically contain converted and refracted phases. Moreover, we can compare the resulting catalog with those extracted from the past years within the same time extent, which is solely based on manual detections. While in the considered period, the automatic detection strategy led to the

location of 391 events, we observed almost 150 earthquakes in the past years, on average. Thus, the application of PhaseNet provides an almost three times richer catalog than the manual one, even for the background seismicity. Moreover, investigating the Gutenberg-Richter law for the considered period from different years (shaded dots for the past years black dots for the automatic catalog, Figure 2.37), we observed a reduction in the magnitude of completeness, by almost 0.5. Since we did not observe any seismic sequence during the operation of the automatic detection strategy and we note a similar slope in the Gutenberg-Richter relation, suggesting that we are observing the same seismicity as for the past years, we can attribute the higher number of retrieved events to the higher sensitivity of the novel approach.

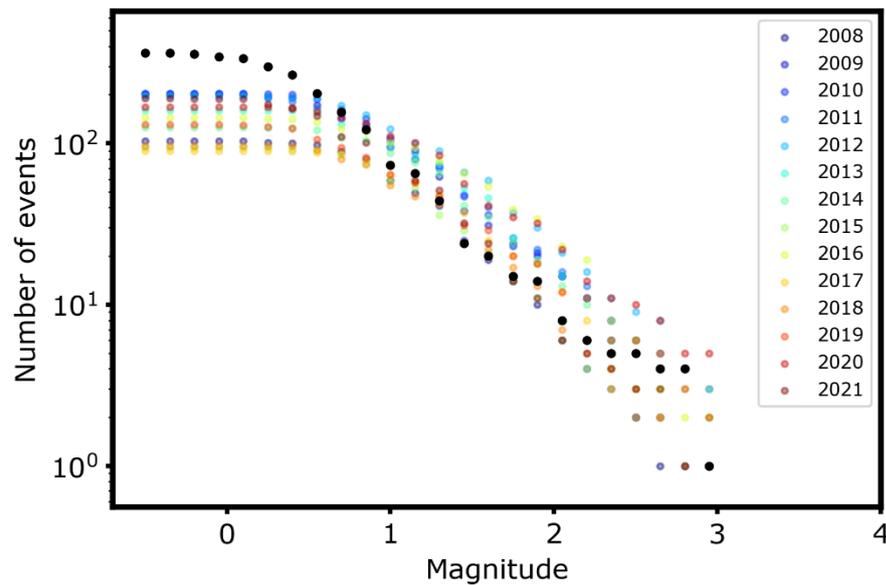


Figure 2.37) Gutenberg – Richter law extracted from events that occurred between 27th February and 27th August for each of the past years from 2008 to 2021 (shaded dots). The black dots mark the Gutenberg – Richter relation for the PhaseNet enhanced catalog. The magnitude of completeness decreases by almost 0.5 magnitude units.

However, the automatic detection strategy produces both false positives and false negatives, which represent a minor fraction of the detection list. False positives are typically raised from detections that match the minimum requirements in the identification stage, showing either phase arrival times declared due to ambient noise or characteristic anthropogenic sources. In Figure 2.38, we report the waveform associated with false picks declared for non-seismic transients.

Most of the false events are typically observed at specific stations that tend to contribute more to the false detections than to the identification of real earthquakes, being more

affected by anthropogenic and cultural noise. However, when averaged among the monitoring period, the number of daily false detection is still lower than the unit, supporting the robustness of the automatic approach. To limit the number of false detections, we tested the possibility of imposing a higher probability score for considering the phase arrival times at the 6 most critical stations, which typically populate the class of false detections. For these latter receivers, we therefore require a minimum probability score of 0.6, twice larger as compared to the usual threshold adopted for the other stations. We apply the novel approach for 20 days of continuous monitoring, strongly reducing of 60 % the number of false positives within the daily detection list, from 41 to 16 false detections. Promisingly, we did not observe any missed real earthquake within the testing window. Following these results, we introduced the novel thresholds in the operational automatic detection workflow.

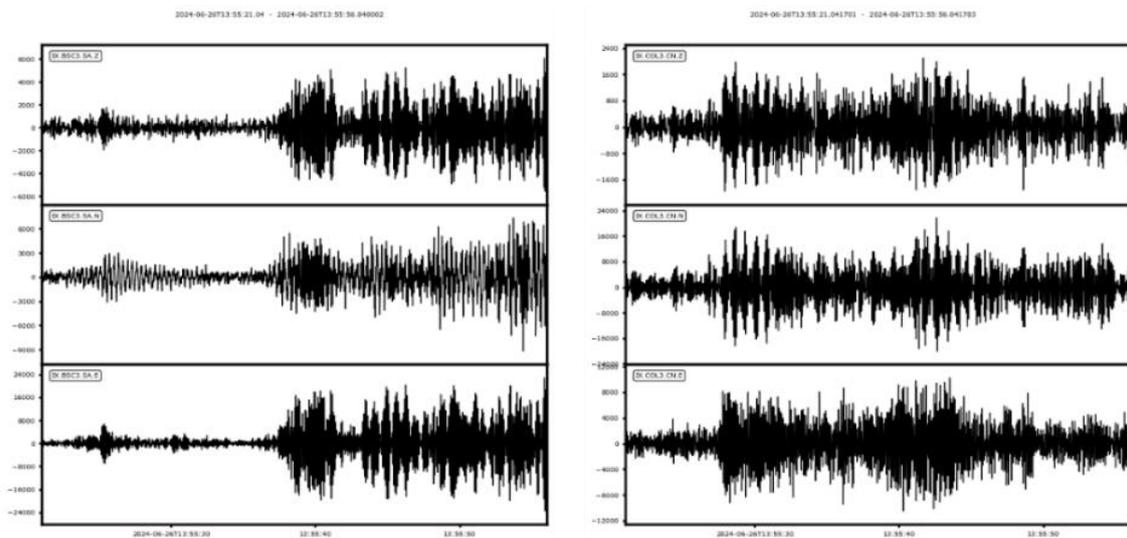


Figure 2.38) Examples of false detections raised by the declaration of arrival times for non-seismic transients.

On the other hand, the automatic detection strategy missed 18 earthquakes which were identified after the visual inspection of the traces. Typically, these events feature low magnitude, with an average magnitude of M_l 0.4, a value lower than the completeness magnitude of both manual (Vassallo et al., 2012) and automatic catalogs (see Figure 2.37). The highest magnitude event which was missed by the automatic procedure is a M_l 1.0 earthquake, which featured 4 P and 1 S manual phase arrival times (<http://isnet-bulletin.fisica.unina.it/cgi-bin/isnet-events/event-card.cgi?id=17711d>). However, the low number of picked stations as compared to similar magnitude events, and the

uncertainties in the earthquake location (of the order of a few kilometres), might have led to a higher magnitude value than the real one. For the considered earthquake, the raw waveforms appear to have a dominant low frequency content and a faint S arrival at all stations (Figure 2.39).

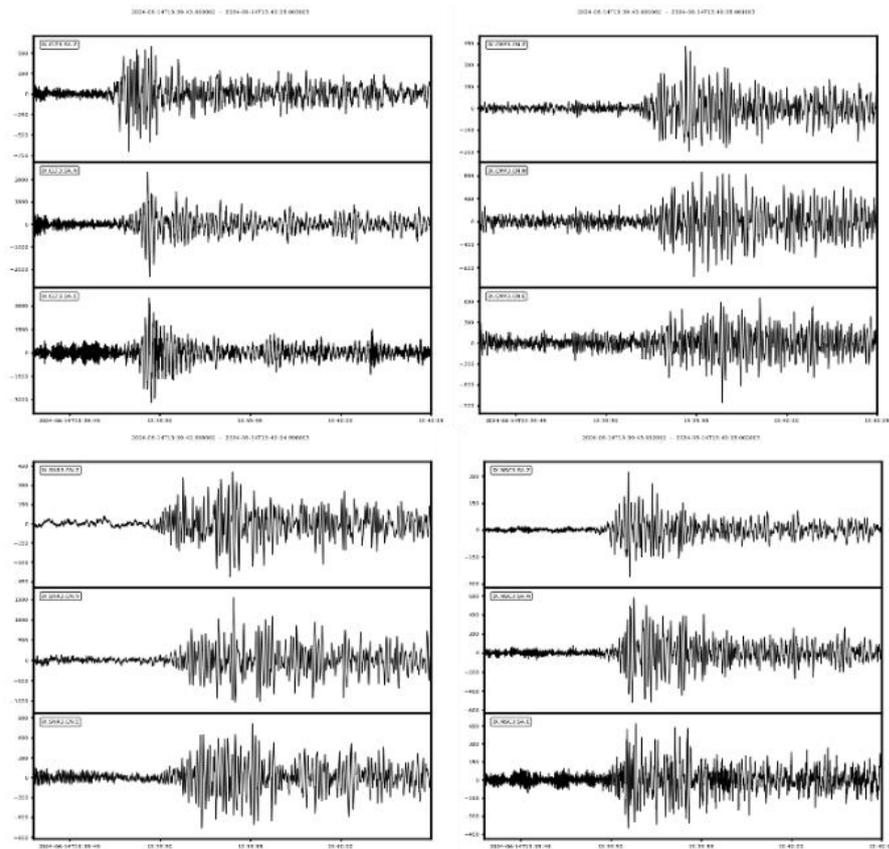


Figure 2.39) M_l 1.0 earthquake missed by the automatic detection strategy. Despite the higher magnitude, this event was observed at 4 stations only, with a single identification of the S arrival time, due to the low frequency content of the waveforms.

For this event, PhaseNet correctly provided the P arrival times at all four stations which were also picked manually, but no probabilities for the S phase overcame the threshold. Thus, the detections did not match the minimum number of associated phases (3 P and 2 S), and it was discarded during the association stage. A second family of missed events is characterized by low-magnitude earthquakes that, at individual stations, feature a sufficient number of picks provided by PhaseNet to be associated, but some of them are discarded, eventually due to uncertainties in the phase arrival times. As an example, we analysed a M_l 0.4 event, manually detected at 3 stations, featuring 3 P and 2 S manual picks (<http://isnet-bulletin.fisica.unina.it/cgi-bin/isnet-events/event-card.cgi?id=17646e>,

Figure 2.40). As the network operator, PhaseNet provided 3 P and 2 S picks among the three stations but, due to a higher residual on the P pick at station COL3 (right panel of Figure 2.40), the declared phase arrival time was discarded in the association, leading to an insufficient number of phases to declare the event.

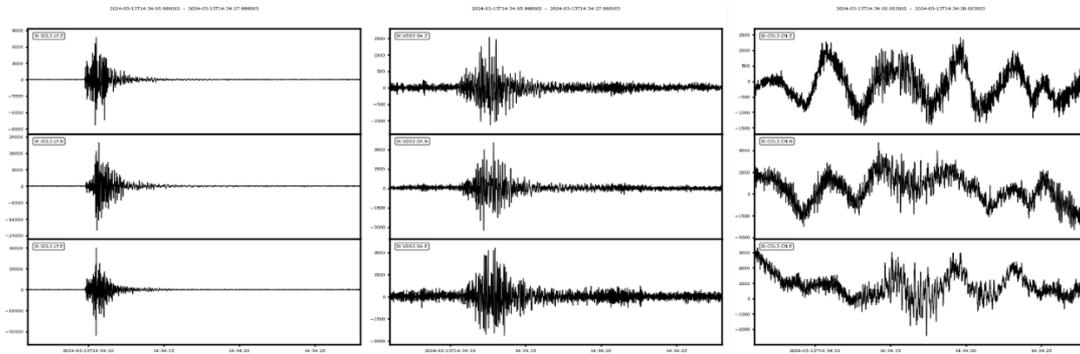


Figure 2.40) M_l 0.4 earthquake missed by the automatic detection strategy, due to uncertainty in a pick declaration which was discarded within the association stage.

Despite the discussed limitations, the proposed detection strategy has widely increased the content of the manual catalog, reporting consistent phase arrival times as compared to the identifications of the network operators. Moreover, this strategy can limit the subjective ability of the operators to identify earthquakes within continuous noisy records, leading to a more homogenous and complete catalog. These results support to feasibility of further automating the workflow, aiming for a fully automatic location workflow and approaching a near-real-time application.

Chapter 3

Accurate determination of earthquake locations

3.1 Introduction

Earthquake location is an inverse problem consisting in the determination of four parameters: the three spatial coordinates of the nucleation point (hypocenter) and the origin time of the event, using the P and S arrival times at different stations. Determining the position of the hypocenters provides powerful information for investigating the geometry of the faults that may generate large-magnitude earthquakes. In the case of a major event, accurate location of foreshocks and aftershocks can provide information on the rupture process from the preparation phase to the arrest by illuminating the structural complexity of the causative fault (e.g., Lomax, 2020; Waldhauser et al., 2021). Moreover, the analysis of the spatio-temporal evolution of the hypocentral locations can reveal the triggering mechanism of seismic sequences and swarms, being driven by fluid diffusion (Antonioli et al., 2005; Chiarabba et al., 2009) or concentrated in small-size asperities on the fault plane (Stabile et al., 2012). Accurate earthquake location of enhanced catalogs has shown complex faults segments activated during major events (Waldhauser & Ellsworth, 2000; Chiaraluce et al., 2017; Michele et al., 2020; Spallarossa et al., 2021; Ding et al., 2023; Kwiatek et al., 2023), identifying the fault planes responsible for the occurrence of the main events. The analysis of these cross sections resulting from accurate earthquake location for the 2016-2017 Centro Italia seismic sequences revealed the strongly complex geometry of the whole fault system composed of relatively younger normal segments emplaced in an environment modelled by the previous compressional tectonics, with additional evidence of the seismicity compartmentalization within specific layers deduced from the distribution of hypocenters at depth (Michele et al., 2020). On the other hand, in the Parkfield region, which is often struck by M6 earthquakes, location analyses have shown a vertical distribution of seismicity, whose shape suggested that the fault has spiral geometry, dipping NE in the northern region, nearly vertical in the central region, and SW in the southern region (Kim et al., 2016). The rapid twisting of the fault plane occurs in a short distance of approximately 50 km. The seismic velocity anomalies and fault geometry suggest location-dependent piecewise faulting, which may cause the

quasi-periodic high-magnitude earthquakes in the Parkfield region (Kim et al., 2016). The uncertainties in the estimation of source positions can be even lower than the source size, typically being of the order of one hundred meters (Ross et al., 2019; Michele et al., 2020). However, hypocenter determination of low-magnitude earthquakes in enhanced catalogs is challenging because they typically emerge from the noise only at the few closest stations with uncertain arrival times. A typical percentage of template-matched events that can be relatively located from enhanced catalogs is $< 50\%$ (Cabrera et al., 2022). Nevertheless, accurate locations from deep catalogs can provide a high-resolution image of fault structures, help to discern their interaction (e.g., Ross et al., 2019; Park et al., 2022; Sukan et al., 2023), illuminating paths for possible fluid migration (Ross et al., 2020; Vuan et al., 2020).

The estimation of the hypocenter coordinates is a non-linear inverse problem, given the relationship between linking the phase arrival times (data) and the unknown location of the source (parameters). Even in the assumption of homogeneous medium, characterized by a constant velocity value for the P and S waves, the phase arrival time at a considered station can be expressed as a function of the hypocentral coordinates as:

$$T_{P,S} = T_0 + TT_{P,S} = T_0 + \frac{\sqrt{(x^{sta} - x_H)^2 + (y^{sta} - y_H)^2 + (z^{sta} - z_H)^2}}{v_{P,S}}$$

where (T_0, x_H, y_H, z_H) represent the origin time and the hypocentral coordinates, $(x^{sta}, y^{sta}, z^{sta})$ refers to the position of the station where the phase arrival time is identified and $V_{P,S}$ indicates the velocity of the considered phase. In a more general and realistic framework, in which the phase velocity is not uniform but varies within the Earth, the former relation can be generalized as

$$T_{P,S} = T_0 + \int_{r_0(s)} u(\mathbf{r}_0) ds \quad (3.1)$$

where $\mathbf{r}_0(s)$ denotes a point at distance s along ray path \mathbf{r}_0 between source and receiver locations and u represents the inverse of the velocity field, also referred to as slowness.

Given the non-linearity of the inverse problem, two main approaches have been proposed for obtaining the hypocentral coordinates from the phase arrival times, which can be classified as linearized or global approaches. In the linearized approach, the hypocentral solution is iteratively obtained by updating a starting trial source location, assuming that the real position is close to the chosen starting coordinates (Geiger, 1912, Lee & Lahr, 1975). In this framework, we can linearly expand the model relation at the first order using the Taylor series around the starting solution \mathbf{m}^0 as

$$T_i \approx T_i^0 + \sum_{j=1}^4 \frac{\partial T_i}{\partial m_j} \Big|_{\mathbf{m}^0} \Delta m_j^0 \quad (3.2)$$

In the former relationship, T_i^0 represents the theoretical arrival time at the i -th station obtained from the starting solution, while Δm_j^0 refers to the residual between the actual and the trial parameters. Reformulating the expression in terms of the residual between the observed and the theoretical arrival times, the relation can be expressed as a linear combination of the residual between the actual and the trial parameters, as follows:

$$\mathbf{d}_i \approx T_i - T_i^0 = \sum_{j=1}^4 \frac{\partial T_i}{\partial m_j} \Big|_{\mathbf{m}^0} \Delta m_j^0 = \sum_{j=1}^4 G_{ij} m_j \quad (3.3)$$

Thus, the relation can be expressed as a linear problem $\mathbf{d} = \mathbf{G} \Delta \mathbf{m}$, in which \mathbf{d} is the vector containing the residuals between the observed and theoretical arrival times (dimension $n \times 1$, with n number of P and S arrival times), $\Delta \mathbf{m}$ provides information about the residual between the actual and the trial parameters (dimension 4×1). In the case of a uniform velocity value, we can analytically express the elements of the matrix \mathbf{G} (dimension $n \times 4$) as:

$$\begin{aligned} G_{i,1} &= \frac{\partial T_i}{\partial x_H} \Big|_{\mathbf{m}^0} = -\frac{1}{v_{P,S}} \frac{x^{sta} - x_H^0}{\sqrt{(x^{sta} - x_H^0)^2 + (y^{sta} - y_H^0)^2 + (z^{sta} - z_H^0)^2}} \\ G_{i,2} &= \frac{\partial T_i}{\partial y_H} \Big|_{\mathbf{m}^0} = -\frac{1}{v_{P,S}} \frac{y^{sta} - y_H^0}{\sqrt{(x^{sta} - x_H^0)^2 + (y^{sta} - y_H^0)^2 + (z^{sta} - z_H^0)^2}} \\ G_{i,3} &= \frac{\partial T_i}{\partial z_H} \Big|_{\mathbf{m}^0} = -\frac{1}{v_{P,S}} \frac{z^{sta} - z_H^0}{\sqrt{(x^{sta} - x_H^0)^2 + (y^{sta} - y_H^0)^2 + (z^{sta} - z_H^0)^2}} \\ G_{i,4} &= \frac{\partial T_i}{\partial T_0} \Big|_{\mathbf{m}^0} = 1 \end{aligned} \quad (3.4)$$

Since \mathbf{G} is a non-square matrix, the least square solution to the inverse problem is given by

$$\Delta \mathbf{m} = (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \mathbf{d} \quad (3.5)$$

Equation 3.5 provides the best solution in terms of minimizing the quadratic deviation between the observed and theoretical data. Using a starting trial solution $\Delta \mathbf{m}^0$, it is thus

possible to determine the perturbation $\Delta\mathbf{m}$ to the starting model that provides the best data fit. The new model will be given by $\mathbf{m}^1 = \mathbf{m}^0 + \Delta\mathbf{m}^0$. This procedure can be iterated using \mathbf{m}^1 as new starting solution, obtaining the perturbation $\Delta\mathbf{m}^1$ and so a new model \mathbf{m}^2 . This iterative loop can be stopped when the prediction error becomes comparable with the mean data error (the error on the individuation of the seismic phase on the seismogram) or when the perturbations with respect to the previous model are not so significant. The linearized approach has the basic advantage of being relatively fast in computational time, but suffers from the presence of multiple local minima, thus location results are strongly connected with the chosen trial solution, eventually leading to a convergence to a local minimum rather than to the global one. To limit this issue, it is possible to include a multi-start approach of the starting solution, that can be chosen randomly, for ensuring a more complete exploration of the domain.

On the other hand, a global approach can systematically explore the whole parameter space, providing a complete view of the solution. An initial grid with a fixed size, number of nodes and location defines the full search region. Subsequently, nested grids are centered automatically on the optimal hypocentral node of the containing grid in one or more directions. The nested grids are typically smaller in size but may have more nodes than the containing grid. For every node of each location grid, the grid-search algorithm must obtain travel-times for every observation from the corresponding travel-time grid. Most of the global location techniques are based on the definition of a probability density function for associating the most probable source among all the possible combinations of parameters, allowing the use of 3D models and providing comprehensive uncertainty information (Lomax et al., 2000). Conversely to the linearized model, these methods are less affected by multiple local minima in the probability, given the complete investigation of the parameter domains using efficient grid-search approaches (Lomax et al., 2000; Lomax, 2008). The maximum likelihood (or minimum misfit) point of the complete, non-linear location PDF is selected as the "optimal" hypocentre. Considering a vector \mathbf{d} of arrival times, let $p(\mathbf{d})$ be the probability density function (PDF) describing the uncertainty of \mathbf{d} due to measurement uncertainties. Similarly, let \mathbf{m} denote the vector of source location parameters and $p(\mathbf{m})$ be the prior PDF representing the a priori available information about the location (i.e. the knowledge of the active faults zones in the area). A solution to the earthquake location model is found by combining the information in the observed data, $p(\mathbf{d})$, the prior PDF, $p(\mathbf{m})$, and the ability of the forward problem to predict the observed data $F(\mathbf{d}, \mathbf{m})$. This is achieved in a probabilistic framework by constructing a PDF, Q , describing the state of posterior information by (Lomax et al., 2009):

$$Q(\mathbf{d}, \mathbf{m}) = k \frac{p(\mathbf{d})F(\mathbf{d}, \mathbf{m}) p(\mathbf{m})}{\mu(\mathbf{d}, \mathbf{m})} \quad (3.6)$$

In which k is a normalization factor and $\mu(\mathbf{d}, \mathbf{m})$ is the homogeneous distribution over data and parameters. The final, posterior state of information about location parameters \mathbf{m} is given by integrating over the data \mathbf{d} to obtain the marginal posterior PDF:

$$Q(\mathbf{m}) = k p(\mathbf{m}) \int_{\mathbf{d}} \frac{p(\mathbf{d})F(\mathbf{d}, \mathbf{m})}{\mu(\mathbf{d}, \mathbf{m})} d\mathbf{d} \quad (3.7)$$

Equation 3.7 is the general, probabilistic solution to the inverse problem of event location from the available data, since it describes the uncertainty in event location \mathbf{m} given all available information. This quantity is often referred to as the likelihood function $L(\mathbf{m})$, which gives a (non-normalized) measure of how good any model \mathbf{m} is in explaining the observed data $p(\mathbf{d})$. $p(\mathbf{d})$ is often approximated by a Gaussian distribution, described by mean \mathbf{d}_0 and covariance matrix \mathbf{C}_d . It is also usually assumed that \mathbf{d} and \mathbf{m} are independent and hence that $\mu(\mathbf{d}, \mathbf{m})$ can be written as $\mu(\mathbf{d})\mu(\mathbf{m})$, with $\mu(\mathbf{d})$ usually taken to be constant. Following these assumptions, the likelihood function can be expressed as:

$$L(\mathbf{m}) = \exp\left\{-\frac{1}{2} [\mathbf{d}_0 - \mathbf{f}(\mathbf{m})]^T \mathbf{C}_d^{-1} [\mathbf{d}_0 - \mathbf{f}(\mathbf{m})]\right\} \quad (3.8)$$

With the mentioned simplifications a maximum likelihood estimation of the origin time, t_0 , can be determined analytically from weighted means of the observed arrival times and the predicted travel times (Tarantola & Valette, 1982). If the observed and predicted arrival times are uncorrelated, the likelihood function can further be shortened as

$$L(\mathbf{x}) = \exp\left\{-\frac{1}{2} \sum_i \frac{[T_i^0 - T_i^c(\mathbf{x})]^2}{\sigma_i^2}\right\} \quad (3.9)$$

In Equation 3.9, \mathbf{x} is the spatial part of \mathbf{m} , T_i^0 are the observed travel times, T_i^c are the calculated travel times for observation i and σ_i summarizes the associated standard deviation of uncertainty in the observed and calculated travel times. Though not normalized, $L(\mathbf{x})$ is sufficient to provide the relative probability of any location \mathbf{m} being the best estimate of the event location given the available data measurements. The definition of the probability density function allows to statistically estimate the uncertainty of the earthquake location. Following this approach, Lomax et al. (2000) provide Gaussian error uncertainty estimates, such as the expectation hypocentre location and the 68% confidence ellipsoid. Large differences between the expected hypocentre

location and the maximum likelihood hypocentre result from an ill-conditioned location. (Lomax et al., 2000).

The accuracy of hypocenter locations depends on the network geometry, accuracy of phase arrival times, and knowledge of the velocity model (Pavlis, 1986; Gomberg et al., 1990; Waldhauser et al., 1999). Often, for computational convenience or due to lack of information, the velocity model is parameterized with velocity varying only with depth. This is commonly called a laterally homogeneous or 1-dimensional (1D) model. Such a model may consist of one or more layers of constant or vertical-gradient wave-speeds which, for local or near-regional scale, may be horizontal and flat. However, the use of a one-dimensional reference velocity model to locate earthquakes might reduce the location accuracy, since the three-dimensional velocity variations which are not considered in less complex models can introduce biases into the estimated travel times. A strategy for partially accounting for the velocity variations consists in the inclusion of station and/or source correction terms in the location procedure (Shearer, 1997; Lin & Shearer, 2005; Lomax, 2008; Lomax & Savvaidis, 2022).

Location uncertainties can also be minimized by using relative earthquake location methods (Poupinet et al., 1984; Fremont & Malone, 1987; Got et al., 1994; Waldhauser & Ellsworth, 2000; Trugman & Shearer, 2017). If the spatial separation between two seismic events is small compared to the hypocentral distance and the scale-size of the velocity heterogeneities, the ray paths between the sources and a certain station are similar along almost the entire ray path. Therefore, the difference in travel times for two events observed at the station can be related to the spatial offset between the events, increasing location accuracy. This is because the absolute errors are of common origin except in the small region where the ray paths differ at the sources. Relative location techniques are based on the use of differential travel times of couples of earthquakes at the same station (also referred to as double difference) rather than absolute phase arrival times at the receiver. Double differences can be expressed as the residual between observed and calculated differential travel times between the two events:

$$dt_k^{i,j} = (t_k^i - t_k^j)^{obs} - (t_k^i - t_k^j)^{theo} \quad (3.10)$$

In the former relation, the indexes i and j refer to a couple of earthquakes, while k indicates the station that recorded the phase arrival time. Using this quantity, Waldhauser & Ellsworth (2000) showed that a generally valid equation for the change in hypocentral distance between a whole set of earthquakes can be written as a linear problem as:

$$\mathbf{WGm} = \mathbf{Wd} \quad (3.11)$$

Where \mathbf{G} defines a matrix of size $M \times 4N$ (M number of double-difference observations, N number of events) containing the partial derivatives, \mathbf{d} is the data vector containing the double-differences, \mathbf{m} is a vector of length $4N$, $(\Delta x, \Delta y, \Delta z, \Delta T)$, containing the changes in hypocentral parameters we are interested in, and \mathbf{W} is a diagonal matrix to weight each equation.

Moreover, it is also possible to improve the accuracy of the relative arrival-time readings using waveform cross-correlation methods. In fact, two earthquakes produce similar waveforms at a common station if their source mechanisms are virtually identical and their sources are colocated so that signal scattering due to velocity heterogeneities along the ray paths is small. Differences in the cross-correlation differential travel time can help in finely quantifying small variations in the earthquake locations.

A standard approach for solving the former equation, in a weighted least-squares sense (i.e. minimizing the L2-norm of the residual vector) is the use of normal equations:

$$\hat{\mathbf{m}} = (\mathbf{G}^T \mathbf{W}^{-1} \mathbf{G})^{-1} \mathbf{G}^T \mathbf{W}^{-1} \mathbf{d} \quad (3.12)$$

With \mathbf{W} containing a priori quality weights expressing the normalized quality of the data, that is the precision of the cross-correlation measurements. For small clusters, and for well-conditioned systems, we can solve the former equation by the method of singular value decomposition (SVD):

$$\hat{\mathbf{m}} = \mathbf{V} \mathbf{\Lambda} \mathbf{U}^T \mathbf{d} \quad (3.13)$$

Where \mathbf{U} and \mathbf{V} are two matrices of orthonormal singular vector of the weighted matrix \mathbf{G} and $\mathbf{\Lambda}$ is a diagonal matrix of the singular values of \mathbf{G} . \mathbf{U} , \mathbf{V} and $\mathbf{\Lambda}$ store information on the resolvability of the parameters \mathbf{m} and the amount of information supplied by the data \mathbf{d} . From these matrices, is it possible to estimate the least square errors, e_i , for each model parameter by

$$e_i^2 = C_{ii} * var \quad (3.14)$$

where C_{ii} are the diagonal elements of the covariance matrix $\mathbf{C} = \mathbf{V} \mathbf{\Lambda}^{-2} \mathbf{V}^T$ and var is the variance of the weighted residuals considering the mean of the residual vector and the residual of the i -th observation. As the system to be solved becomes larger, SVD is inefficient. In these cases, the solution $\hat{\mathbf{m}}$ can be found by using the conjugate gradient algorithm LSQR (Paige & Saunders, 1982), which takes advantage of the typical sparseness of the matrix \mathbf{G} . LSQR solves the damped least-squares problem:

$$\left\| \mathbf{W} \begin{bmatrix} \mathbf{G} \\ \lambda \mathbf{I} \end{bmatrix} \mathbf{m} - \mathbf{W} \begin{bmatrix} \mathbf{d} \\ \mathbf{0} \end{bmatrix} \right\|_2 = 0 \quad (3.15)$$

with λ being the damping parameter.

However, the reliability of the LSQR uncertainties is not guaranteed, since the diagonal elements of the covariance matrix are critically dependent on the convergence during internal iterations (Paige & Saunders, 1982; Waldhauser & Ellsworth, 2000). The reliability of the errors reported by LSQR, can be verified using SVD or by implementing a statistical resampling.

In this chapter, we use advanced techniques for determining the hypocentral coordinates in different seismogenic areas. Specifically, we:

- obtained accurate earthquake locations for the enhanced catalogs of seismic sequences in Southern Italy (generated in Section 2.3.1), integrating absolute and relative location techniques using automatic phase arrival times, for understanding the evolution of the seismic sequences (Section 3.3.1)
- located the enhanced catalog (obtained in Section 2.3.2) resulting from the application of the advanced detection techniques on the data collected by the dense array deployment in Southern Italy (Section 3.3.2)
- performed earthquake location for the seismicity observed within the Campi Flegrei caldera from 2014 to 2024, selecting an innovative technique for absolute and relative location of the earthquake grounding on the definition of a source-specific station travel time correction field (Section 3.3.3)

3.2 Methods

3.2.1 A global probabilistic approach for earthquake location: NLLoc

NLLoc is a probabilistic earthquake location technique adopting a non-linear, global sampling algorithm to obtain the hypocenter coordinates (Lomax et al., 2000). This technique has been established as the reference method for obtaining absolute earthquake locations, being applied in different contexts as tectonic (Chiaraluce et al., 2017; Festa et al., 2021), volcanic (Tramelli et al., 2022; De Siena et al., 2024), induced seismicity (Coccia et al., 2013; Lim et al., 2024) and even for earthquake early warning purposes (Satriano et al., 2008). NLLoc produces an estimate of the posterior density function (PDF) for the spatial hypocentre location, $Q(\mathbf{m})$, using either a systematic grid-search or a stochastic, Metropolis-Gibbs sampling approach. PDF values obtained by grid-search

samples represent the complete, probabilistic spatial solution to the earthquake location problem. This solution indicates the uncertainty in the spatial location due to Gaussian picking and travel-time calculation errors, the network-event geometry, and the incompatibility of the picks (Lomax et al., 2000). The travel times between a station and all nodes of a 3D grid are calculated using the Eikonal finite-difference scheme of Podvin & Lecomte (1991), while the grid-search algorithm in NLLoc performs successively finer, nested grid-searches within a spatial volume to obtain an estimate of the location PDF. The grid-search performs a systematic, exhaustive coverage of the search region and thus can identify multiple optimal solutions and highly irregular confidence volumes. The estimated travel times are evaluated to obtain a non-normalized location PDF value, which is stored at the appropriate node. When the grid-search for the final, fine grid is complete, the gridded PDF values are normalized by assuming that the integral of PDF over the search volume is unity.

For estimating the earthquake location, NLLoc evaluates the probability density function on each node of a chosen spatial grid, accounting for the weighted time residuals among the observed and the theoretical travel times. Therefore, crucial elements in the hypocentral inversion are the grid size, the weighting factors associated with the phase arrival times and the adopted velocity model. The step and the dimension of the grid are guided by the extension of the seismogenic region, balancing the location resolution and the computational time. For the location analysis in the Southern Apennines (Sections 3.3.1 and 3.3.2), which extends about $100\text{ km} \times 70\text{ km}$, we adopted the grid parameterization used in De Landro et al. (2015). We discretized the space domain by deploying 205 nodes along the latitude and longitude directions, while we selected 73 points along the depth direction. Specifically, to account for the station elevation, we extended the nodes also to a negative depth direction up to -2 km . The spatial step among consecutive nodes along each direction was set to 0.5 km . On the other hand, for the location analysis at the Campi Flegrei, whose seismicity typically interests an area of $8\text{ km} \times 8\text{ km}$, we adopted a finer step size of 0.2 km , discretizing the space using 61, 81, and 56 points along the latitude, longitude and depth direction, respectively. Also in this latter case, we extended the grid along the negative direction of the depth up to -1.0 km , to account for the station elevation.

To determine the hypocentral coordinates, NLLoc weights the time residual among the observed and the theoretical arrival times. Therefore, the weight refers to the importance of the individual phase arrival time within the hypocentral inversion stage. Classically, the weight is directly connected to the uncertainty in the manual estimation of the phase arrival times, being dependent on the signal-to-noise ratio of the phase onset. The assigned weights typically inversely refer to the importance within the inversion: the lower the weight, the higher the importance. As an example, De Landro et al. (2022)

proposed a weighting scheme for the phase arrival times as indicated in the following Table 3.1:

Table 3.1) Weighting scheme proposed in De Landro et al. (2022), for assigning the importance of the phase arrival time within the location inversion.

TIME UNCERTAINTY (s)	WEIGHT
< 0.05	0
[0.05, 0.1[1
[0.1, 0.2 [2
[0.2, 0.5[3
> 0.5	4

For the determination of the earthquakes within the Campi Flegrei caldera (Section 3.3.3), we used the manual phase arrival times provided by the INGV – Osservatorio Vesuviano (available at <https://terremoti.ov.ingv.it/gossip/flegrei/index.html>), along with the manual estimation of the uncertainty on the phase arrival times. On the other hand, in the location analysis in the Southern Apennines, we used fully automatic phase arrival times. The importance of the arrival times in the location stage was estimated considering the associated probability for machine learning picks and the CC values for template matching phases, respectively. We converted the probability values (ranging between 0.1 and 1.0) into discrete weights for location (from 4 to 0, decreasing numbers corresponding to larger importance in the location stage) according to the table proposed by Mousavi et al. (2020) for the EQT detections:

Table 3.2) Table proposed by Mousavi et al. (2020) for the converting the probability of EQT detections in weights for earthquake location.

PROBABILITY	WEIGHT
>0.7	0
]0.5, 0.7]	1
]0.2, 0.5]	2
]0.1, 0.2]	3
< 0.1	4

For the template matching picks, we imposed at least the same level of accuracy of the machine learning picks used for the declaration, eventually increasing the discrete weights for low cross-correlation values. We raised the discrete weights by one point for every decimal of the CC coefficient detaching from 1.0. However, since machine learning models result from complex and non-linear operations, it is not trivial to relate the quality and reliability of the identification of the phase arrival time to a single quantity, as the

signal-to-noise ratio in the manual picking. The analysis performed in Section 2.3.1 also showed that the probability scores associated with the declaration of a phase arrival time might depend on the position of the event within the considered window (Pita-Slim et al., 2023). We acknowledge that, at this stage, there is not a continuous mapping between probability values and phase arrival time quality evaluated through discrete weights. However, this conversion into weights mitigates the variability in the probability values, possibly associated with the specific position of the event in the analysed window. However, the probability value is the only metric that is provided along with the phase arrival time, and the conversion of the probability scores in weights to be used within earthquake location is a common approach when dealing with machine learning-based catalogs and/or automatic workflows for earthquake characterization (Mousavi et al., 2020; Tan et al., 2021; Zhang et al., 2022).

3.2.2 Relative location techniques for hypocenter location

Relative location techniques associate the residual between the observed and predicted phase travel-time difference for pairs of earthquakes observed at common stations to changes in the relative position of the hypocenters, through the partial derivatives of the travel times for each event with respect to the hypocenter parameters (Waldhauser & Ellsworth, 2000; Trugman & Shearer, 2017; Lomax & Savvaidis, 2022). This approach cancels common errors when the distribution of seismicity is sufficiently dense, i.e., where distances between neighbouring events are small relative to station distances (typically a few kilometers or less). By linking hundreds or thousands of earthquakes together through a chain of near neighbours, it is possible to obtain high-resolution relative hypocenter locations over a large area (Waldhauser & Schaff, 2008). Relative location techniques are widely used for providing fine-scale and accurate location catalogs, with decametric relative location uncertainties (Waldhauser & Schaff, 2008; Chiaraluce et al., 2017; Michele et al., 2020). As for the absolute location techniques, the main families for location algorithms can be classified as linear (HYPODD, Waldhauser & Ellsworth, 2000) or global (NLLoc-SSST-coherence, Lomax & Savvaidis, 2022)

3.2.2.1 A linearized approach for relative earthquake location: HYPODD

HYPODD (Waldhauser & Ellsworth, 2000) minimizes the residuals between observed and theoretical travel-time differences (double-differences), accounting for absolute travel-time and cross-correlation P- and S-wave differential travel-time measurements. HYPODD follows the approach described in Section 3.1, for the linearized relative location techniques. HYPODD implements a least-squares solution by iteratively adjusting the vector difference between hypocentral pairs. The use of cross-correlation

differential travel times has been shown to increase the spatial resolution of the relative hypocenter location, being indicative of the small-scale distance of the event pairs (Waldhauser & Schaff, 2008; Michele et al., 2021). Since catalog and cross-correlation data are jointly inverted, the interevent distances are determined to the accuracy of the cross-correlation data, while the relative locations between multiplets and uncorrelated events are determined to the accuracy of the absolute travel-time data.

For the location analysis in the Southern Apennines, we used absolute locations (estimated using NLLoc) as the starting point for relative re-locations of events in each sequence using HYPODD (Waldhauser & Ellsworth, 2000). In the analysis of the enhanced catalogs for seismic sequences, we evaluated the catalog delay times in each sequence using picks for event pairs separated by less than 10 km in absolute location at all the available stations. For CC differential travel times, we evaluated the delay times for events that were separated by less than 10 km, on seismograms decimated to 100 Hz and filtered in the frequency band [1.5 – 15] Hz (Schaff & Waldhauser, 2005; Michele et al., 2020). We assessed the length of the time windows for extracting the waveforms around the P and S arrival times by performing parametric tests. The typical extension of the time window for evaluating the cross-correlation among a couple of events is of the orders of seconds, with a slightly longer extension for the S phase and including different portions of waveform preceding the phase arrival time (Hauksson & Shearer, 2005; Waldhauser & Ellsworth, 2008; Stabile et al., 2012; Michele et al., 2020; Waldhauser et al., 2021). We here report the parametric test performed for the analysis of the P waves differential travel time for the EQT catalog of the Rocca San Felice seismic sequence. A similar analysis was conducted for the S waves. We compared the cross-correlation coefficients resulting from different window extensions, extracted [-0.3, 0.4] s, [-0.4, 1.0] s, [-0.5, 0.5] s and [-0.4, 0.7] s around the P arrival time. In Figure 3.1, we report the distribution of the obtained cross-correlation coefficients.

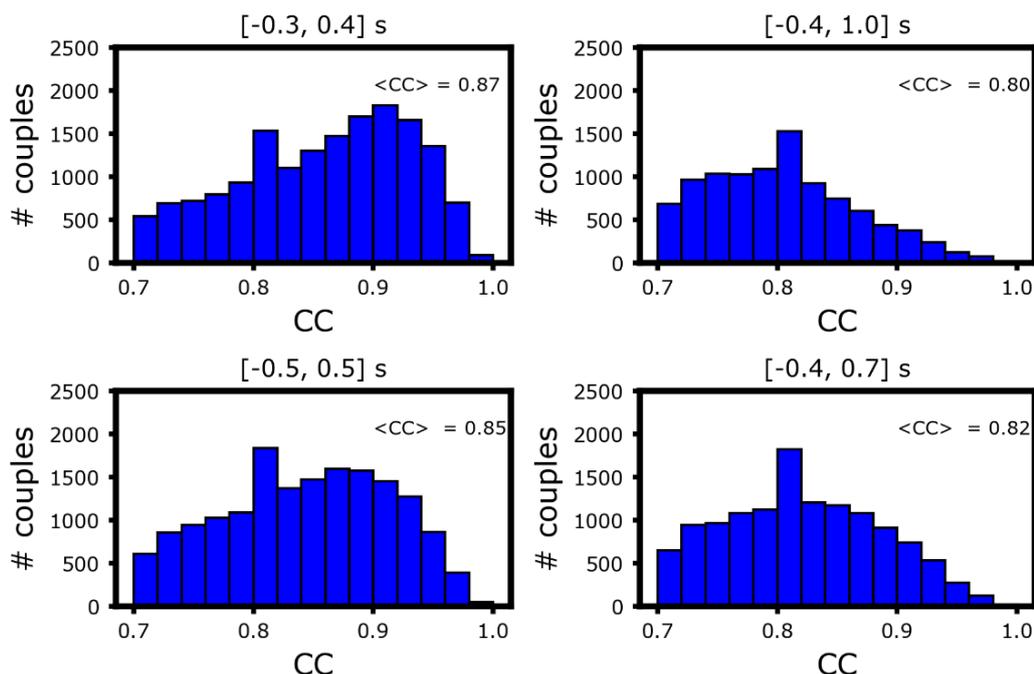


Figure 3.1) Distribution of the cross-correlation coefficients resulting from different windows extracted around the P arrival time.

Too short windows (i.e., $[-0.3, 0.4]$ s, $[-0.5, 0.5]$) resulted in too high values of CC coefficients such that the reliability of the lag measurement was overestimated. In particular, the shape of the distribution appeared to conflict with the analogous distribution observed in Schaff & Waldhauser (2005). Moreover, given the expected drop of the cross-correlation coefficient as a function of the interevent distance (Figure 2.3), which might lead to a decrease of several decimals in waveform similarity among events occurring a few kilometers apart (Michele et al., 2020), we should expect more populated bins in the central part of the distribution. We thus selected a 1.1 s ($[-0.4, 0.7]$ s) long window around the P phase arrival time for calculating the CC coefficients, imposing a maximum lag of 1s. Similarly, for the S phase we used a 1.4-s-long window. Finally, we only retained delay times for events with a CC coefficient higher than 0.7 (Chiaraluce et al., 2017; Michele et al., 2020; Waldhauser et al., 2021)

We estimated relative locations with HYPODD using an iterative least square procedure (LSQR) that minimized the differential time residuals for pairs of earthquakes at common stations by adjusting the vector connecting their hypocentres (Waldhauser & Ellsworth, 2000). We used 4 steps of 4 iterations (a total of 16 iterations) of damped and dynamically weighted least square inversions. In the initial settings, we assigned higher weights to catalog delay times, for better constraining the location of the clusters, and we increased

the contribution of the CC differential travel times in the following settings, to consider the different positions of the events within the cluster. The damping factor was selected to stabilize the problem (Waldhauser, 2001). To avoid inconsistency with ray patterns used in the absolute locations, we extracted a 1-D model composed of 20 thin layers resampling the velocity model used for absolute locations.

LSQR only approximates some aspects of the uncertainty (Waldhauser & Ellsworth, 2000), so we applied the Singular Value Decomposition (SVD) method for a more complete assessment of location errors. The SVD option can only solve a significantly lower number of earthquakes than LSQR. Nevertheless, we were able to apply the SVD technique for all the sequences apart from the Rocca San Felice sequence. For this latter sequence, we estimated location uncertainties using a bootstrap strategy, realizing 200 independent double differences relocation runs on subsets of events within the sequence. Each subset was obtained by randomly extracting 150 events, 60 % of which belong to the machine learning catalog. This selection ensures a more robust linkage to the cluster, since the number of picks associated with templates is generally larger than for template-matched events. We evaluated the location uncertainties from a statistical analysis based on the distance of each event from the cluster centroid in all the runs where it was located. This procedure allows the quantification of dependency of the results on the single subset. For the i -th event, we estimated the uncertainty along the j -th direction as $err_j^i = \text{median}_{p,m} |(x_{j,p}^i - x_{j,p}^c) - (x_{j,m}^i - x_{j,m}^c)|$, where p and m indicate two independent runs in which the i -th event was located, and the superscript c refers to the cluster's centroid of the considered run. The robustness of these estimates was verified by observing agreement with uncertainties from a SVD inversion for the subset of template events.

3.2.2.2 Global approach for relative earthquake location: NLLoc-SSST-Coherence

Relative location techniques are based on the use of differential travel times for resolving the small-scale hypocentral distance among couple of earthquakes. These techniques also leverage the use of waveform similarity and precise, cross-correlation, differential timing between events at individual stations to determine fine-scale, inter-event spatial relations. Despite these methods have been shown to highlight small-scale features of the seismicity, showing narrow streaks and sets of faulting structures (Got et al., 1994; Michele et al., 2020; Waldhauser et al., 2021), these techniques require dense monitoring coverage, with stations providing a massive amount of differential travel times, that is not always guaranteed. This is the case of the seismicity in the Campi Flegrei caldera, where despite a dense multiparametric network monitors the area (Bianco et al., 2022), not all the waveforms for the stations are publicly available. Following the strategy of

HYPODD, while we could have extracted the phase arrival times at all the stations from the released bulletin (<https://terremoti.ov.ingv.it/gossip/flegrei/index.html>) and computed the catalog delay times, we could only evaluate the waveform cross-correlation just for a limited fraction of the stations. Therefore, to avoid introducing biases in the location procedure, we adopted the strategy proposed by Lomax & Savvaidis (2022), NLLoc-SSST-coherence (NLLoc-SC), which is intended to achieve the same resolution of cross-correlation-based techniques, without explicitly evaluating the waveform similarity, and thus being particularly suitable for the analysis of the seismicity at the Campi Flegrei caldera (Section 3.2.3).

NLLoc-SC is a standard, arrival-time location procedure, modified to improve relative location accuracy through the use of spatially varying, source-specific station travel-time corrections (SSST) and a new, waveform coherence-based, multi-event location procedure. In a first relocation stage, NLLoc-SC iteratively develops SSST corrections on decreasing length scales (Richards-Dinger & Shearer, 2000; Lomax & Savvaidis, 2022), improving relative location accuracy and clustering of events. In contrast to station static corrections, which give a unique time correction for each station and phase type, SSST corrections vary smoothly throughout a 3D volume to specify a source-position-dependent correction for each station and phase type (Lomax & Savvaidis, 2022). These corrections account for 3D variations in velocity structure and corresponding distortion in source-receiver ray paths. Spatial-varying, SSST corrections are most effective for improving relative locations on all scales when the ray paths between stations and events differ greatly across the studied seismicity. SSST corrections can improve multi-scale precision when epistemic error in the velocity model is large, such as when a 1D, laterally homogeneous model or a large-wavelength, smooth model is used in an area with sharp, lateral velocity contrasts or smaller-scale, 3D heterogeneities. The iteration uses Gaussian smoothing kernels of decreasing size to produce final, NLL-SSST locations. Residuals from P and S arrivals and relocated events meeting minimum quality criteria are used for updates at each iteration. Figure 3.2 summarizes the obtained spatial distribution source-specific station corrections, extracted from Figure 1 in Lomax & Savvaidis (2022), along with the spatial travel time corrections.

In a second relocation stage, NLL-SSST-coherence reduces aleatoric location error by consolidating information across event locations based on waveform coherency between the events (Lomax & Savvaidis, 2022). NLL-coherence uses waveform similarity directly to improve relative location accuracy without the need for differential time measurements or many stations with waveform data. The method assumes that high coherency between waveforms for two events implies the events are nearly colocated, and also that all of the information in the event location, when corrected for true origin time shifts, should be nearly identical in the absence of noise.

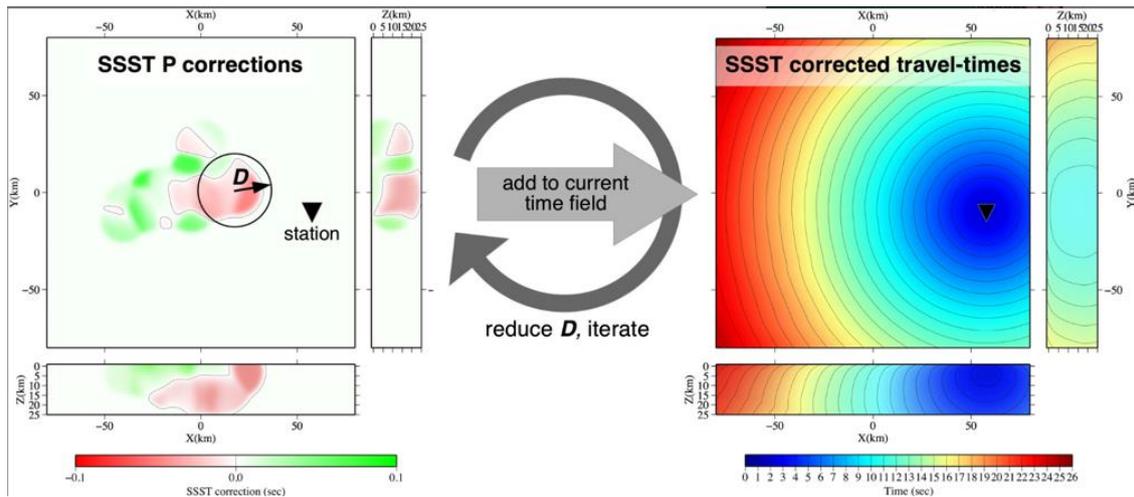


Figure 3.2) Left panel: example of source-specific station P wave travel time corrections, extracted from Figure 1 in Lomax & Savvaidis (2022). SSST corrections are added to the current, P travel-time field for the station to produce updated, SSST corrected travel-times (right).

Then, stacking over probabilistic locations for individual nearby events can be used to reduce the noise in this information and improve the location precision for individual, target events. We measured coherency as the maximum, normalized cross-correlation between vertical component velocity records from one or more stations for pairs of events within 2 km in the analysis of the seismicity at the Campi Flegrei caldera. We take the maximum station coherency between the target event and each other event as a proxy for true inter-event distances, and thus as stacking weights, to combine NLL-SSST location probability density functions over the events. Specifically, the procedure only uses high coherency waveforms, with a cross-correlation threshold of 0.5, evaluated on waveforms filtered in the [2 – 10] Hz band and downsampled to 50 Hz. The contribution of individual traces in the stack is then weighed by a cosine taper function to provide smoothly changing stacking weights between 0 and 1. Following this approach, the weight is near 1 only for coherence values > 0.8 . In effect, this stack directly improves the hypocenter location for each target event by combining and completing arrival-time data over nearby events and reducing aleatoric errors in the data, such as noise, outliers and missing arrivals. Figure 3.3, extracted from Lomax & Savvaidis (2022), illustrates the steps for obtaining the stacked PDF for a cluster of similar events.

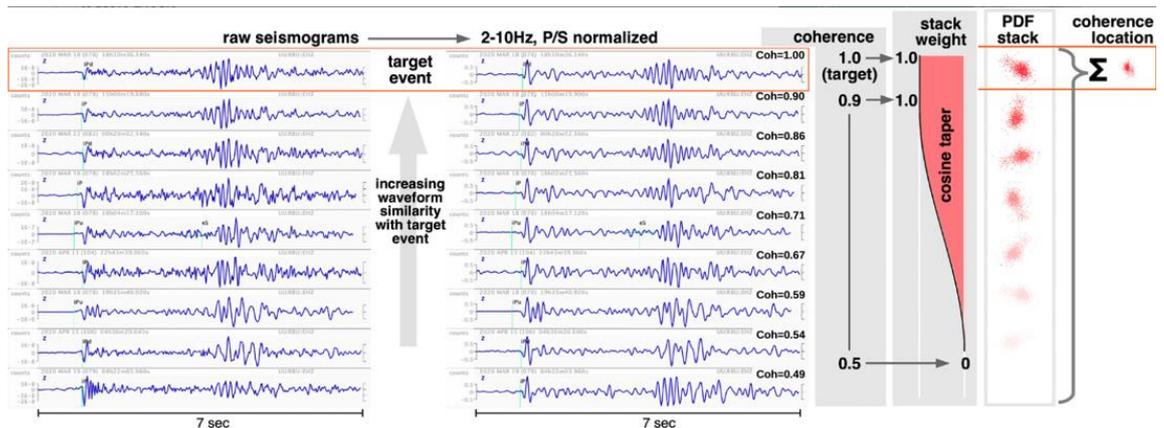


Figure 3.3) Left - Top trace represents the target waveform while the reported coherence with the target is indicated to the upper right of each waveform. Right - Schematic representation of coherences between the target and each waveform, corresponding stacking weight after cosine taper mapping of coherence, location PDF's forming the stack (color intensity indicates stack weight), and final NLL-coherence location PDF for the target event.

3.2.3 Accuracy of automatic phase arrival times and influence of the velocity model in the hypocenter determination

In the analyses of Sections 3.3.1 and 3.3.2, we use automatic picks provided by advanced detection techniques. In this section, we compare the accuracy of phase arrival times obtained through the integration of machine learning and similarity-based techniques (EQT + TM) with the corresponding manual identification. We here report the distributions of the time residuals between the automatic and manual phase arrival times for the enhanced catalogs obtained for the 10 seismic sequences in Irpinia (Section 2.3.1), extracted from the 265 common events among the enhanced and existing manual catalogs. Figure 3.4 reports the distribution of the time residuals with the P, S and both phases, respectively, analogously to Figure 2.34. Both P and S arrival times feature residual distributions centred around zero (Scotto di Uccio et al., 2023). In particular, the distribution of the P residuals features a standard deviation lower than 0.1 s, this value being half of that observed from the automatic detection strategy for the daily monitoring of the seismicity. However, while the latter analysis includes the whole seismic network, and hence also stations further from the hypocenter, in the analysis of the seismic sequence we selected the closest stations to each sequence, which typically feature a higher signal-to-noise ratio and thus are favoured in the identification of phase arrival times.

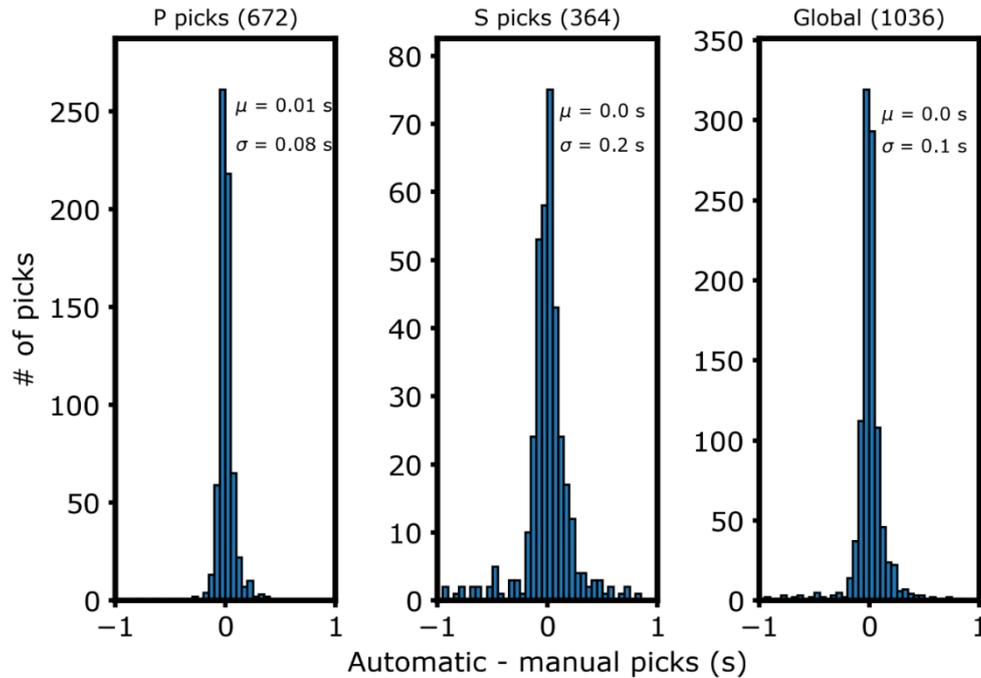


Figure 3.4) Distribution of the time residual among the automatic and manual phase arrival times for the ten seismic sequences analysed in Section 2.3.1. Both P and S phase arrival times feature zero-mean residuals, with a slightly larger dispersion for the S phase.

As observed in Figure 3.4, the distribution of the S residuals features a higher dispersion as compared to the corresponding distribution of the P residuals, being the former phase is more affected by the complexity of the media, which might generate converted phases. However, these results suggest that the automatic identifications are accurate enough to be used within the location stage, without leading to significant biases as compared to the manual locations.

We further assessed the accuracy of the automatic phase arrival times by investigating the dependency of the predictions to the training datasets used for generating the deep learning models. In Sections 2.3.1 and 2.3.2, we used the ML technique EQTransformer (Mousavi et al., 2020), originally trained on the worldwide STEAD dataset (Mousavi et al., 2019, see Figure 2.6), containing ~ 1 M earthquake and ~ 300 K noise waveforms recorded by ~ 2600 seismic stations at epicentral distances up to 300 km. For the most populated sequence analysed in Section 2.3.1, (Rocca San Felice, ID 1), we evaluated again the residual distributions for P and S phases using different training datasets for EQT. Aiming to create standardized formats for seismological analysis, Woollam et al. (2022) proposed a toolbox, known as Seisbench, collecting homogenized benchmark datasets and trained machine learning models. Seisbench provides a unified environment

for applying deep learning models to seismic waveforms, and for accessing and training machine learning algorithms on seismic datasets. SeisBench has been built to alleviate traditional bottlenecks when applying machine learning techniques to seismic data, in particular the steps of data preparation, collection and labelling. In this analysis, we selected the EQT models trained on INSTANCE (Michelini et al., 2021) and IQUIQUE (Woollam et al., 2019), reported in the left and right panels of Figure 3.5, respectively.

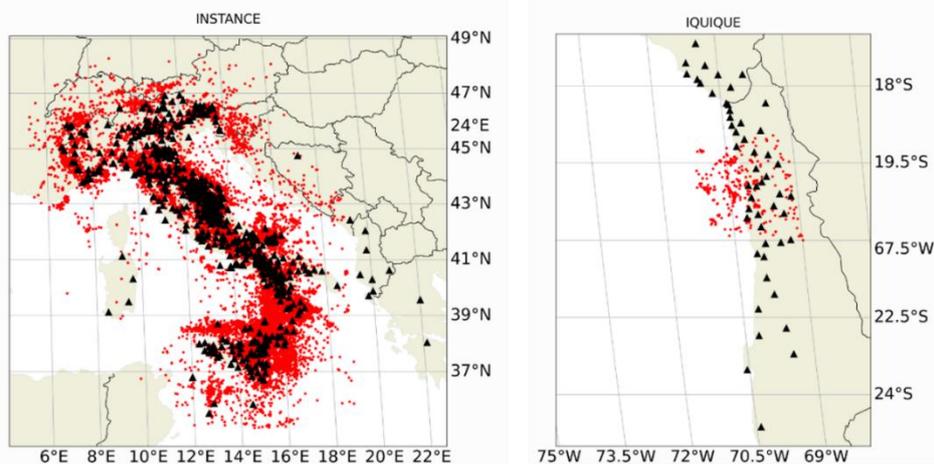


Figure 3.5) INSTANCE (left panel) and IQUIQUE (right panel) datasets, implemented in Seisbench. For estimating the phase arrival times, we used EQT models trained individually on one of the considered datasets.

The INSTANCE benchmark dataset is a dataset of signals compiled by the Istituto Nazionale di Geofisica e Vulcanologia (INGV), containing $\sim 130,000$ noise traces and ~ 1.2 million three-component waveforms, associated with $\sim 50,000$ earthquakes, whose magnitudes range from 0 - 6.5. The IQUIQUE dataset contains 13,400 instances of picked arrivals from the aftershock sequence following the $M_w=8.1$ Iquique earthquake occurring in Northern Chile in 2014. Conversely to the INSTANCE model, which also included noise waveforms in the dataset, the IQUIQUE waveforms only contain examples of earthquakes. We used Seisbench to apply these two versions of EQT on the same stations used in Section 2.3.1 for the detection analysis of the Rocca San Felice seismic sequence. In Figure 3.6, we report the residual distributions, as compared to the manual identifications, for the P and S phases of both models, along with the results from the original STEAD model used in Scotto di Uccio et al. (2023). The results suggest that machine learning models trained on massive datasets (STEAD and INSTANCE) feature similar performance, both in terms of accuracy and number of declared phase arrival times. Phase arrival times provided by both models feature zero-mean residuals as compared to the manual identifications for both phases, with a slightly larger dispersion for the S phase, coherently with what is observed in Figure 3.4.

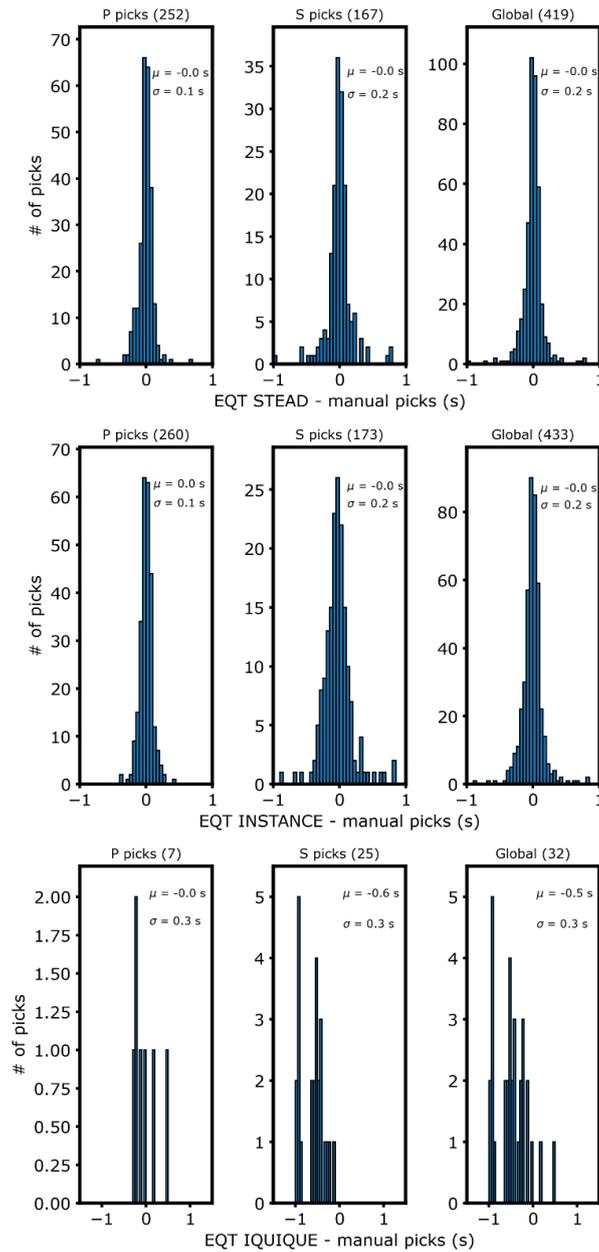


Figure 3.6) Residual distributions for P and S phase arrival times predicted by EQT models trained on STEAD (upper panel), INSTANCE (central panel) and IQUIQUE (lower panel), respectively. The distribution observed for the INSTANCE training dataset is coherent with the STEAD dataset.

These characteristics suggest that machine learning models trained on massive datasets can well generalize even on areas that are not fully representative of the training dataset,

implying that during the training stage the model has correctly learnt the global features characterizing the phase arrival times. Conversely, we observe degraded results when analyzing the automatic phase arrival times provided by the EQT model trained on IQUIQUE dataset. We note a strongly diminished number of phase arrival times declared by the model, being almost two orders of magnitude lower than the number of picks provided by EQT trained on both STEAD and INSTANCE datasets. Moreover, the accuracy of the picks declared by the model trained on IQUIQUE dataset is not compatible with the former results. While the number of declared P arrival times (7) is too low to be discussed statistically, the distribution of the S residuals features a strong non-zero mean, being centred in -0.6 s, implying that automatic phase arrival times tend to precede the manual identifications. We attribute this result to the differences in the seismogenic sources between the local IQUIQUE dataset and the target seismicity and to the low number of events included in the training catalog. While for the STEAD and INSTANCE catalogs, the number of earthquakes was typically in the order of tens/hundreds thousands (Mousavi et al., 2019; Michelini et al., 2021), the IQUIQUE catalog contained 411 events only (Woollam et al., 2019). Moreover, the typical depth of the IQUIQUE earthquakes varies between 40 and 150 km, being significantly deeper than the crustal earthquake which are we targeting. On the other hand, STEAD and INSTANCE catalogs are more populated by crustal earthquakes and thus learnt better the characteristics of the propagating waves.

Finally, to assess the dependency of the quality of the hypocentral coordinates on the adopted velocity model, we compared the location results from different velocity models. We focus here on the Rocca San Felice seismic sequence (ID 1 in Section 2.3.1) and, starting from a 1-D layered velocity model tailored for the Irpinia area (Matrullo et al., 2013, red line in Figure 3.7), we derived two gradient models, which smooth the discontinuities in the wave velocity across layer boundaries, by linearly interpolating values between either at the top (blue line) or at the middle points of the layers (green line). We note that the velocity model obtained by fixing the velocity value at the top of the layers systematically overestimates the velocity in each layer, while the continuous model obtaining by fixing the velocity of the layer at its mean point alternates depths for which the velocity is either underestimated or overestimated as compared to the model of Matrullo et al. (2013). We compared the location accuracy in terms of the corresponding uncertainties (horizontal and vertical hypocenter uncertainties and root-mean-square of the travel time residuals, RMS). Figure 3.8 summarizes the distribution of the corresponding location uncertainties for each model, color-coded as in Figure 3.7. Figure 3.7 contains the velocity models for the P wave, but we applied a similar interpolation also for the S wave, using a constant V_p/V_s ratio (1.85, Matrullo et al., 2013).

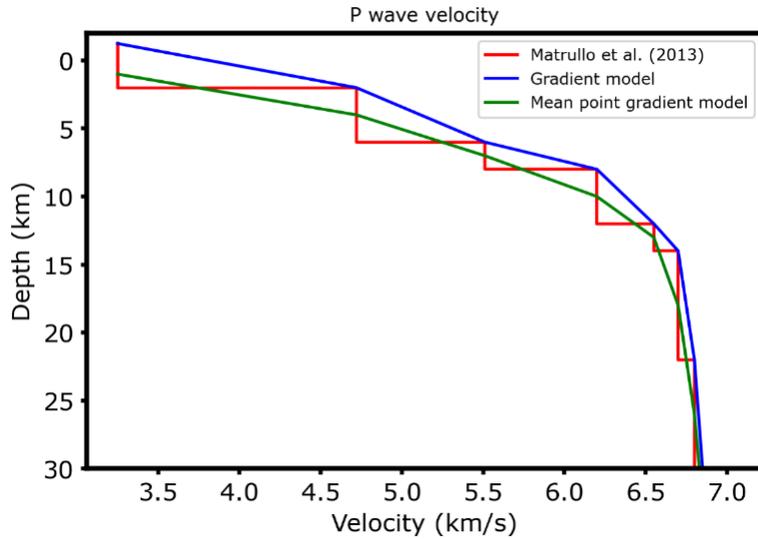


Figure 3.7) Smooth velocity model obtained from the interpolation of the Matrullo et al., (2013) P wave velocity model. The blue line represents the model obtained interpolating the velocity at the top depth of the layers in Matrullo et al. (2013), while the green blue line represents the model obtained interpolating the velocity at the mean depth of the layers

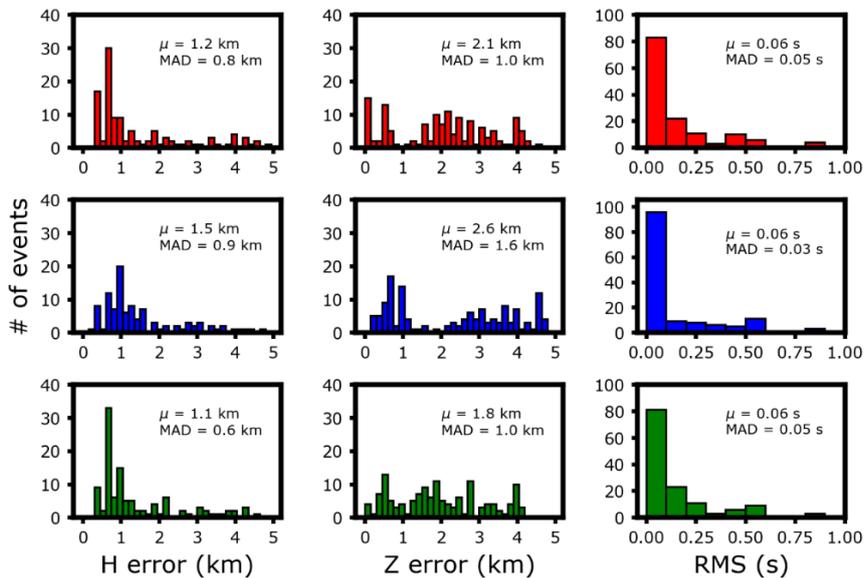


Figure 3.8) Horizontal hypocenter uncertainty (first column), vertical location uncertainty (second column) and RMS (third column) for the obtained location using the 1D layer velocity model (red), the continuous model obtained by interpolating the top of each layer (blue) and the continuous model obtained by interpolating the mean depths of each layer (green). The text reports the median and the MAD values observed for each distribution.

Despite observing comparable median RMS among the results obtained from the three models, the interpolated model which fixes the velocity at the mean point of the layers features the lowest spatial location uncertainties, as compared to the other two candidate models. The retrieved spatial uncertainties are reduced by $\sim 10\%$ and $\sim 15\%$ for the median horizontal and vertical hypocenter errors, respectively. Moreover, since the velocity model obtained by interpolating the velocity from the top of the layers (blue line in Figure 3.7), constantly overestimates the velocity of the waves at each depth, it results in greater depths, as compared to the other two models. Following these results, we adopted the continuous velocity model obtained by fixing the values at the mean point of the layers for the absolute locations in this chapter.

3.3 Unveiling seismogenic structures through accurate earthquake location

In this section, we report the location results for the seismicity occurring in the Southern Apennines and within the Campi Flegrei caldera. In Section 3.3.1, we locate the enhanced catalogs for seismic sequences in Southern Apennines, generated in Section 2.3.1, integrating absolute and relative location techniques. In Section 3.3.2, we perform an analogous location analysis using the enhanced catalogs generated from the data recorded by the dense arrays in Southern Apennines during its 1-year-long deployment. In Section 3.3.3, we investigate the seismicity of the last 10 years that occurred within the Campi Flegrei caldera to illuminate the seismogenic structures activated during the ongoing unrest phase.

3.3.1 Depicting fault segments from the analysis of enhanced catalogs for seismic sequences

In Section 2.3.1, we generated enhanced catalogs for 10 seismic sequences in Southern Apennines, integrating machine learning and template matching detection techniques (Scotto di Uccio et al., 2023). The obtained seismic catalogs improved the content of existing manual detection lists by a factor of ~ 7 , leading to the declaration of $\sim 1.8\text{k}$ earthquakes. The number of located events in the manual catalogs, obtained using a linearized approach for determining the absolute position of the hypocenters, is 265. We recall that an event is included in the manual INFO catalog if at least 4 phases have been picked on the records, including one S phase, and at least one station contains both P and S picks. Moreover, individual phase residuals are required to be smaller than 0.5 and 1.0 s for the P and S waves, respectively. In Figure 3.9, we report the absolute locations of the events from the manual INFO catalog for all the sequences in the left panel, along with the seismicity projected on a vertical plane, oriented according to a characteristic

direction for the Irpinia region (N40E) in the right panel. This angle marks the orthogonal direction to the strike of the major faults responsible for the 1980 M 6.9 earthquake (Bernard & Zollo, 1989).

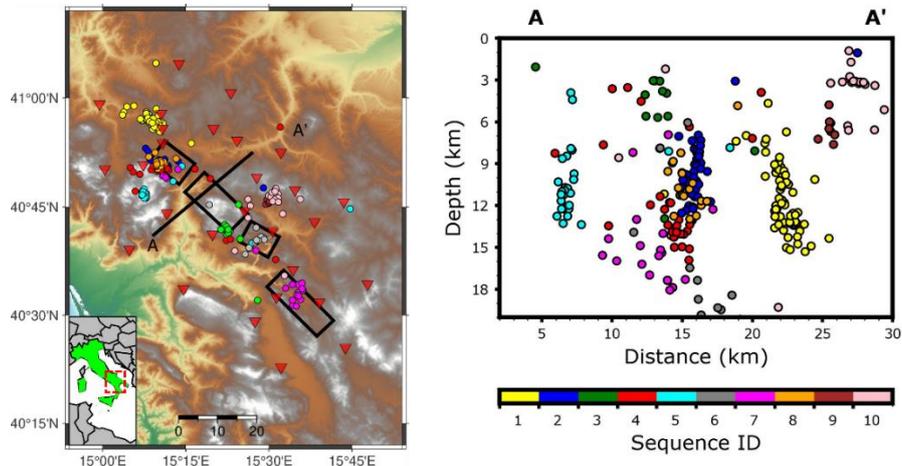


Figure 3.9 Absolute location for the 10 seismic sequences from the manual INFO catalog. In the left panel, the black segments mark the faults activated during the 1980 M 6.9 earthquake. In the right panel, the seismicity is projected along a vertical plane oriented N40E, (A-A' direction), orthogonal to the strike of the faults represented in the left panel.

As results from the manual INFO catalog, hypocenters mainly appear to be scattered and sparse within the region affected by the seismic sequences, without clearly mapping seismic structures in the vertical cross-section of Figure 3.9. The spatial location uncertainties are as large as a few kilometers, enough to obscure the fault segments or patches on which the seismicity takes place. The median (mean) values of spatial location uncertainties for the events in the manual catalog are 0.9 km (1.4 km) and 1.3 km (2.6 km) for the horizontal and vertical directions, respectively. Figure 3.10 summarizes the location uncertainties for the events in the manual catalogs.

We performed accurate earthquake location for the enhanced seismic catalogs generated in Section 2.3.1, integrating the absolute location technique NLLoc (Lomax et al., 2000, Section 3.2.1) and the relative location approach HYPODD (Waldhauser & Ellsworth, 2000, Section 3.2.2.1), using the parameterization described within the corresponding section (Scotto di Uccio et al., 2024a). For the absolute location with NLLoc, we used the available P and S arrival times from the enhanced catalogs of Scotto di Uccio et al. (2023). However, while within the detection analysis we limited the selected stations to the 5-7 closest ones to the sequence's centroid, to detect the lower-magnitude events, we here extended the starting phase-pick dataset to all the stations not included in the detection step of Scotto di Uccio et al. (2023), to ensure an appropriate azimuthal coverage. For the new stations, we selected the velocity records when available, and the acceleration data

as a second choice, following the same detection strategy used in Scotto di Uccio et al. (2023).

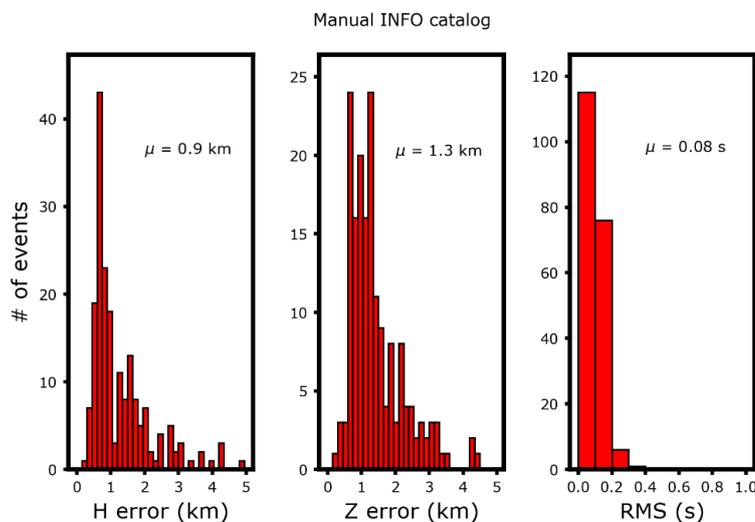


Figure 3.10) Location uncertainties from the events in the manual catalogs. The left panel reports the horizontal uncertainties, the central panel reports the vertical uncertainties, while the right panel reports the root-mean-square of the travel time residuals. In each panel, μ indicates the median value of each distribution.

Guided by the results presented in Section 3.2.1, we adopted a continuous velocity model obtained from Matrullo et al. (2013), fixing the velocity of the layers at the mean point of each layer (Figure 3.7), which resulted in the lowest median location uncertainties (Figure 3.8). The described strategy led to the estimation of the absolute locations for 1130 events ($\sim 60\%$ of the detection catalog of Scotto di Uccio et al., 2023). The number of absolute locations from the enhanced catalogs is 5 times larger than in the manual catalog, improving the number of located earthquakes by a factor 2.5 to 8 for each sequence, and providing a wide set of catalog and cross-correlation delay times for earthquake relocation. The cross-section in Figure 3.11, which was obtained analogously to the seismicity projection in Figure 3.9, reports a sparser hypocentral distribution, as compared to the locations in the manual catalog. However, it is worth recalling that the locations in the enhanced catalogs mainly refer to very low-magnitude events, which are typically less constrained than the higher-magnitude events included in the existing manual catalogs (Festa et al., 2021). In Figure 3.12, we report the hypocenter uncertainties for the absolute locations of the earthquakes in the enhanced catalogs. The median spatial uncertainties are estimated to be 1.6 km and 1.9 km for the horizontal and vertical locations, respectively, while the median RMS of the travel time residuals is 0.12 s.

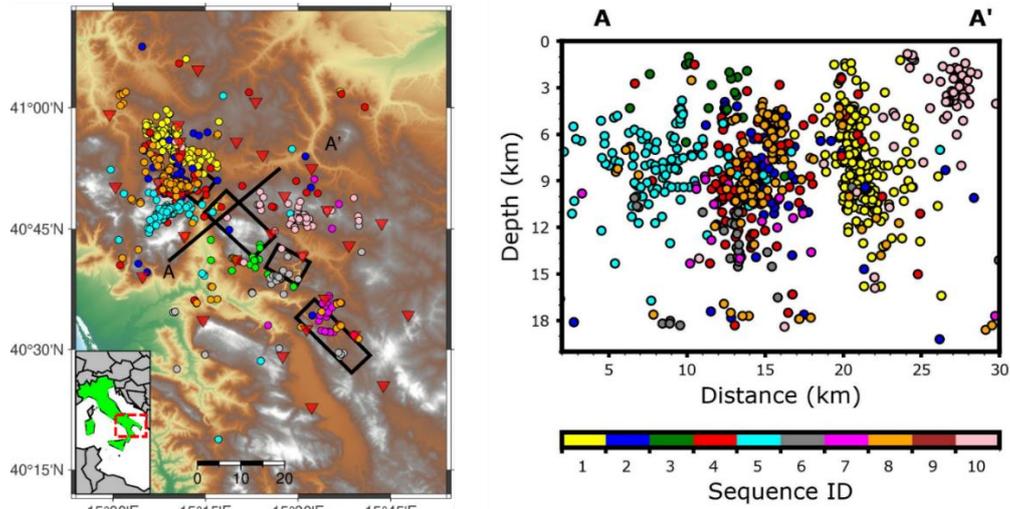


Figure 3.11) Absolute locations for the enhanced catalogs of seismic sequences obtained in Scotto di Uccio et al. (2023). 1130 events (60% of the starting detection catalogs) have been located.

Absolute location of enhanced catalogs (Scotto di Uccio et al. 2023)

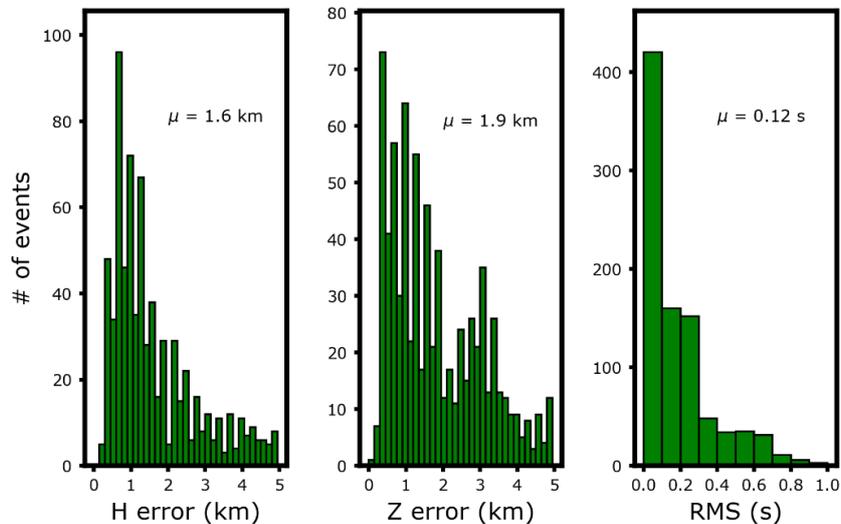


Figure 3.12) Location uncertainties from the events in the manual catalogs. Left panel reports the horizontal uncertainties, central panel reports the vertical uncertainties, while the right panel reports the root-mean-square of the travel time residuals

Despite these values being larger than the corresponding median uncertainties for the locations of the manual catalog, when extracting the spatial location errors for the common events among the two detection lists, we observe that the automatic location

errors are typically lower than those of the manual catalog, especially for the uncertainties on the hypocentral depth. In Figure 3.13, we report the uncertainties for the shared events in the manual and enhanced catalogs.

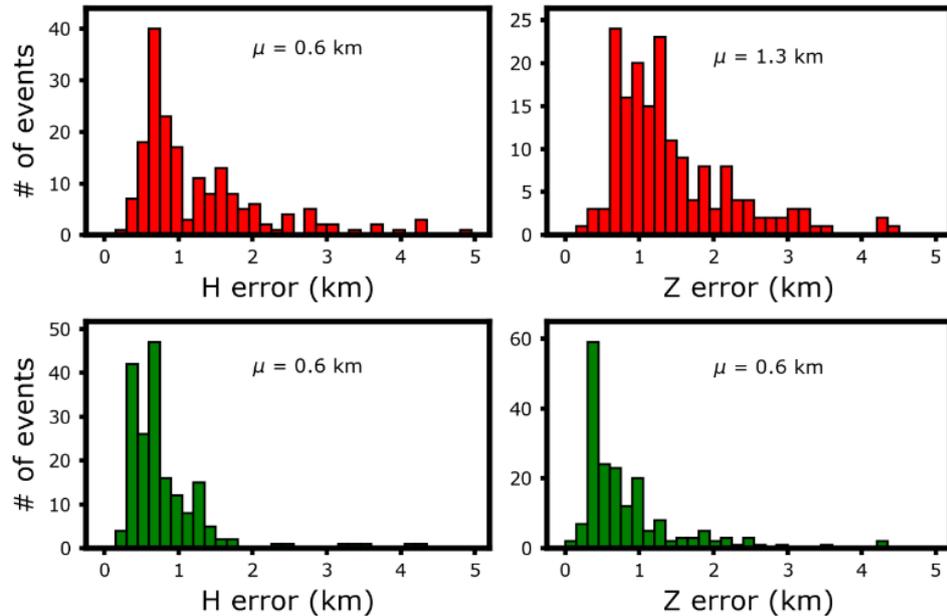


Figure 3.13) Distribution of the spatial location uncertainties for common located events in the manual and automatic catalogs. The location uncertainties for the events located using the automatic phase arrival times typically show lower values.

We attribute these features both to the accuracy of the automatic picks and to the choice of a continuous velocity model in the automatic location, which better reproduces the velocity variation than the 1D layer velocity model adopted in locating the manual catalog. Starting from the absolute positions of earthquakes estimated with NLLoc, we achieved double difference relocation with HYPODD (Waldhasuer & Ellsworth, 2000) using both catalog and cross-correlation differential travel times. We adopted a 1D velocity model of 20 thin layers, extracted from the continuous velocity model used in the absolute location, to minimize the inconsistency between the models. Using an iterative and dynamic weighting of the catalog and cross-correlation differential travel times, we achieved hypocentral relocations for 550 events, from 8 out of the 10 seismic sequences analyzed in Scotto di Uccio et al. (2023). The two sequences for which we did not get relocations (IDX 7 and IDX 9 in Scotto di Uccio et al., 2023) feature the lowest number of detections (about 40 events). The total number of relocated events represents $\sim 30\%$ of the enhanced catalog. A similar fraction is observed for each sequence and appears coherent with earthquake relocation in other template-matching-derived catalogs

(Cabrera et al., 2022; Ross et al., 2019). The reduction in the number of well-located events is driven by the different impact of the waveform similarity during the detection and location phases. Template matching detection algorithms leverage stacked cross-correlations across the entire network (Chamberlain et al., 2018; Vuan et al., 2018), resulting in a global similarity value, to which stations with both high and low cross-correlation values contribute. High-quality thresholds on the similarity coefficient for cross-correlation differential travel times requested by accurate double difference locations (Michele et al., 2020; Waldhauser et al., 2021), limit the number of available stations, especially for low-magnitude events. Figure 3.14 shows the double difference relocation of the earthquakes in the enhanced catalogs (sequence indexing with respect to Scotto di Uccio et al., 2023). In the left panel, we show the position of epicenters with respect to the seismic network. In the right panel, the hypocenters are projected along the vertical plane A-A' oriented perpendicular to the trend of the Apennines (N40°E). This plane represents the direction orthogonal to the main structures of the area, generating the 1980 Irpinia earthquake. For all the sequences, apart from the Rocca San Felice sequence (ID 1), where it was unfeasible to adopt the SVD approach due to the higher number of earthquakes, we estimated the statistically robust relative location uncertainties from the covariance matrix (Waldhauser & Ellsworth, 2000). From the inversion of the covariance matrix, we obtained median spatial location uncertainties of 70 m, 76 m and 119 m along the East, North and vertical direction, respectively. In Figure 3.15 we report the distribution of the spatial location uncertainties estimated with the SVD.

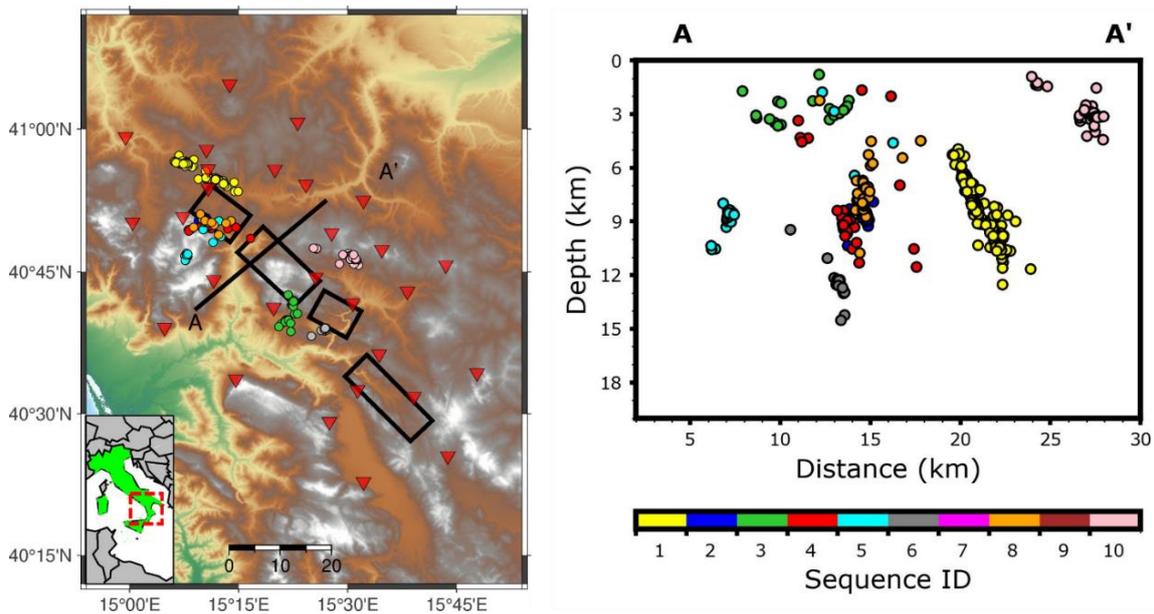


Figure 3.14) Double difference relocation of the enhanced catalogs of Scotto di Uccio et al. (2023). The representation of the cross-section shows that the seismicity patterns feature clear alignments

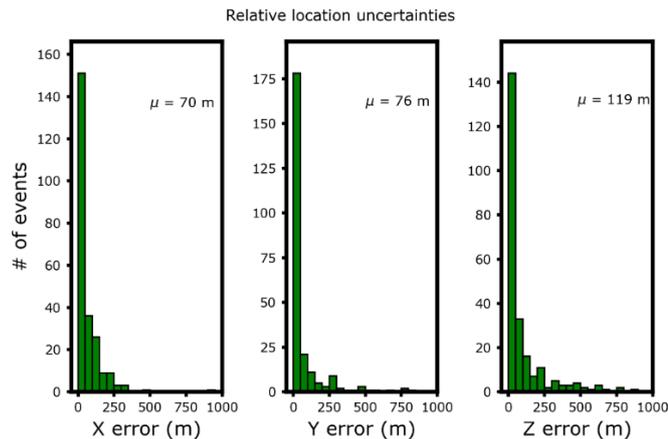


Figure 3.15) Distribution of the spatial location uncertainties for the double difference relocation of the events in the enhanced catalogs. The spatial scale in this figure is much lower than the analogous representation of the kilometeric-size location uncertainties in Figures 2.11 and 2.12

For the Rocca San Felice seismic sequence, marked in Figure 3.14 with yellow dots (IDX 1) and featuring the highest number of both detections and relocated events (~ 800 and 250 events respectively), we estimated location uncertainties using a bootstrap strategy, realizing 200 independent double difference location runs on subsets of events within the sequence (see Section 3.2.2.1). Each subset was obtained by randomly extracting 150 detected events, 60 % of which belong to the machine learning catalog, and we evaluated the location uncertainties by considering the distance of each event from the respective cluster's centroid in all the runs where it was located. This strategy led to median values of location uncertainties of 91 m, 31 m and 105 m in the East, North and vertical directions respectively, comparable to the robust median values extracted from the SVD inversion. The cross-section (right panel of Figure 3.14) shows that the seismicity patterns feature clear alignments and a high degree of clustering, highlighting km-sized structures that share similar dips. For all sequences, the spatial extent of the sequences depicted by relocations is much greater than what is expected from the total released seismic moment, extending for 1–3 km along the depth, longer than what is suggested by scaling laws for $M < 4$ earthquakes, and features a NW-SE-dipping plane. Integrating the visualization of the absolute (Figure 3.11) and relative (Figure 3.14) hypocenter location, when mapping the seismic sequences at depth, their location is generally not compatible with the faults that hosted the 1980 event - based on either the fault trace at the surface (Westaway & Jackson, 1987) or the fault dip and geometry estimated from seismic and levelling data (Bernard & Zollo, 1989; Amoroso et al., 2005). This indicates that seismic sequences ruptured small patches of secondary, sub-parallel segments, as compared to the main structure of the M 6.9 earthquake.

Although the relocated catalog is only twice as large as the manual catalog, the improvement in cluster definition and spatial resolution is much more significant, allowing the identification of alignments and structures at a kilometric scale, that were not clearly illuminated from the manual catalog (Palo et al., 2023a), owing to the wide increase in the number of differential travel-times. As an example, for the Rocca San Felice sequence, we used 97.5 k catalog differential times (47k for the P phase, 50.5k for the S phase) and 85k CC delay times (31k for the P phase, 54k for the S phase), which was more than one order of magnitude as compared to the differential travel-times extracted from the manual catalog (Festa et al., 2021).

The high-resolution earthquake location allows us to track the spatiotemporal evolution of the seismic sequence. An interesting case is represented by the Rocca San Felice seismic sequence, which was activated through the occurrence of a M_l 3.0 earthquake and featured a similar magnitude event (M_l 2.8) almost two days later. When zooming on the sequence location (Figure 3.16 - left panel), the epicenters clearly suggest the presence of two clusters, at a distance of 5 km from each other. The projection of the seismicity along the vertical plane oriented N40°E (Figure 3.16 – right panel), indicates that the two clusters feature similar orientations but occurred at different depths: the shallower one is mostly confined between 6 km and 9 km, and the deeper one between 9.5 km and 11 km. The two clusters were activated at different times during the sequence, as shown in Figure 3.16 – right panel, where colors denote the occurrence time relative to the mainshock. The foreshocks (pink circles) and the events that occurred within the first two days of the mainshock (indicated with a black star) illuminate a first 4-km-long segment with a dip of 55°, coherently with the focal mechanism estimated by Festa et al. (2021). Two days after the mainshock, the occurrence of a M_l 2.8 event activated the deeper secondary patch, featuring a slightly shorter extent.

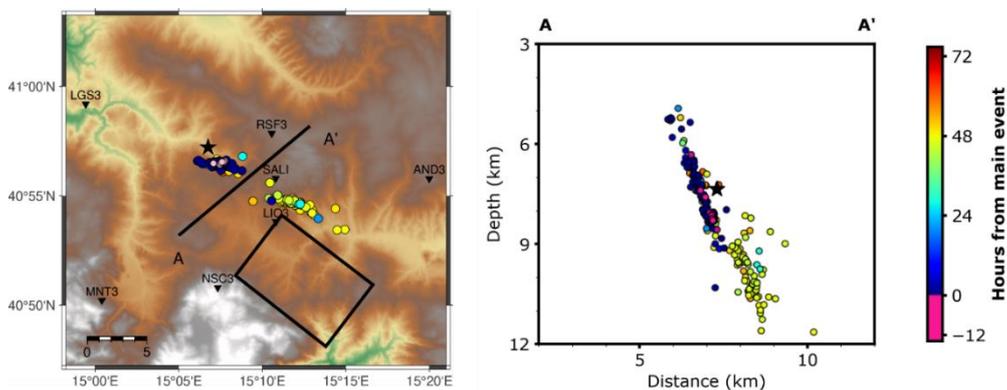


Figure 3.16 Left panel: Spatio-temporal evolution of the epicenters for the Rocca San Felice seismic sequence (IDX 1), colored according to the occurrence time from the main event. The main event is represented by a black star. Right panel: Cross-section colored according to the occurrence time from the main event.

The presence of two separated clusters was not recognized in the previous work of Festa et al. (2021) and is supported by the change in the first station recording the P wave arrival, that occurred at the station RSF3, for events in the first cluster, and at LIO3, for events in the second one. The improvement in the double difference relocation, as compared to the results of Festa et al. (2021), who analysed the same seismic sequence but using just the 48 highest magnitude events, comes from the combination of a deeper event catalog, but also from a larger number of picks per event. Indeed, for most of the events we retrieved picks and waveforms also for the accelerometric station SALI, located close to the centroid's epicenter, indicating that strong motion sensors can provide useful information even for microseismic events if their sensitivity is high enough (for SALI it is 4.0 V/m/s^2). In Chapter 4, we will characterize the source properties for these sequences, investigating the source parameters and the stress release model, aiming to understand the evolution of the seismic sequences along the fault plane.

3.3.2 Using ultra-microseismic events for illuminating fault segments

In Section 2.3.2 we generated enhanced seismic catalogs from seismic records collected within a temporary experiment in Southern Apennines consisting in the deployment of 200 seismic stations organized in 20 dense arrays. We used the approach of Scotto di Uccio et al. (2023), integrating machine learning and template matching techniques on overlapping subnetworks of 6 arrays, aiming to identify very low-magnitude events. The resulting catalog from the one-year monitoring was composed of $\sim 3.6\text{k}$ earthquakes, increasing the existing manual catalogs of the same period by a factor higher than 8 and including as many earthquakes as there are in the manual catalog from 15 years of continuous monitoring in the area using the ordinary seismic network.

In this section, we perform earthquake relocation for the enhanced catalog using the same strategy proposed in Section 3.3.1. We initially obtain absolute locations of the events using NLLoc and a 1D layer velocity model using the automatic phase arrival times. A pick refinement strategy was applied, based on cross-correlation and hierarchical clustering (Muzellec et al., 2024), for improving the quality of the phase arrival times. We thus relocate the events using the refined phase arrival times, finalizing the approach with the double difference relocation of the events using HYPODD.

For earthquake location, we used $\sim 47\text{k}$ P and $\sim 72\text{k}$ S automatic phase arrival times. However, the spatial distribution of the stations providing the arrival times of seismic waves is not homogeneous, but some arrays tend to contribute more to the pick dataset. In Figure 3.17, we report the distribution of the P (left panel) and S (right panel) picks for the declared events among the arrays. A red pixel indicates the presence of the phase arrival time for the considered station, otherwise a blue pixel indicates the absence of a pick. From Figure 3.17, we note that arrays 06, 09 and 11 typically populate the set of

phase arrival times for most of the declared events. If we recall the spatial position of the arrays (Figure 2.23), we observe that these arrays were deployed in the central region of the Irpinia region. Therefore, this can be an indication of a higher seismicity rate in the region and/or a lower noise level around the installation sites, which led to a higher signal-to-noise ratio for low-magnitude earthquakes. Moreover, we note that the density of the phase arrival times in the northern sector of the Irpinia (arrays 01 to 05) is much scarcer than the corresponding values in the central and southern regions. This might indicate a lower number of earthquakes that occurred in the northern region. This observation is supported by the earthquake location in the existing manual catalog as reported in Figure 3.18.

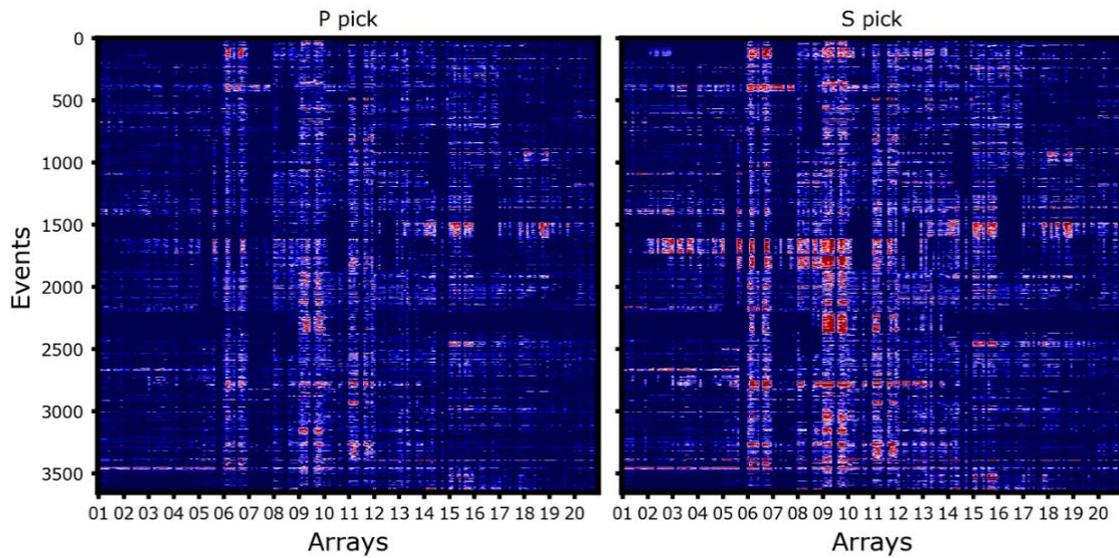


Figure 3.17 Distribution of the P (left panel) and S (right panel) phase arrival times for the 20 deployed arrays. Most of the phase arrival times are observed for the arrays 06, 09 and 11, which are deployed in the central region of the Southern Apennines. A red pixel indicates the presence of the phase arrival time for the considered station, otherwise a blue pixel indicates the absence of a pick.

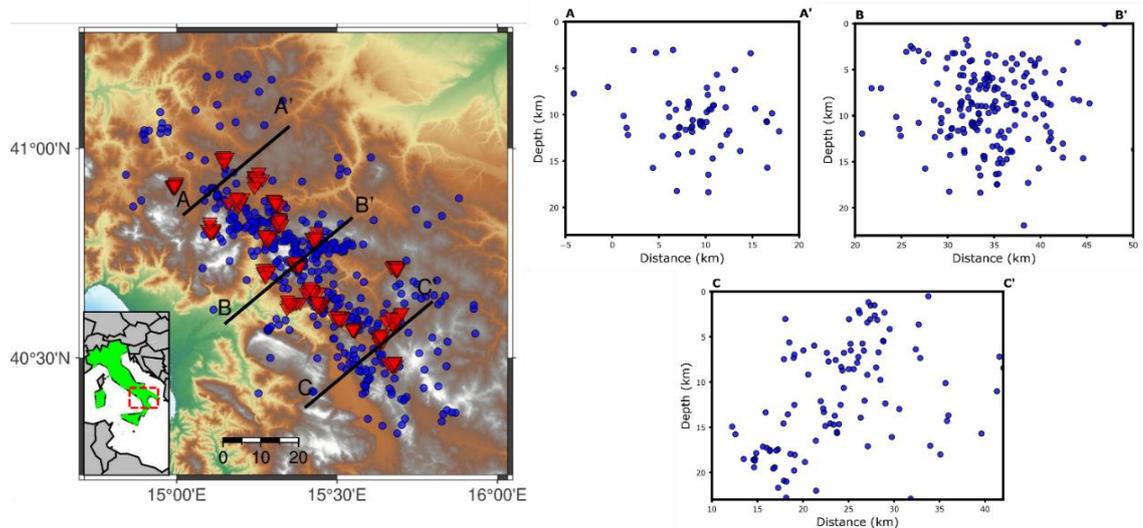


Figure 3.18) Absolute location of the earthquakes in the manual catalog from September 2021 to August 2022, obtained by selecting the stations of the ordinary seismic network. Most of the seismicity occurred in the central and southern section of the Irpinia region. The cross-sections on the right panel represent the projection of the hypocenters along vertical planes oriented N40E. For each sector, only hypocenters within 15 km from the vertical planes are considered.

Moreover, all the cross-sections for the three considered sectors do not feature any significant seismicity alignment. The seismicity is more confined at shallow depths in the northern and central sectors of the area, whilst deepening toward the southern sector of the Irpinia region. Analysing the manual catalog, a lack of seismicity emerges in the southeastern edge of the region (within planes B-B' and C-C'), which was also identified by De Landro et al. (2015) and interpreted to be due to the contact between units with different rheological behaviour in response to the extensional NE–SW stress regime acting in the chain.

Using the automatic phase arrival times and the parameterization of NLLoc as from Section 3.3.1, we obtained absolute hypocentral locations for the enhanced catalog, which are reported in Figure 3.19. The representation of the epicenters from the enhanced catalogs confirms the different seismicity density in the different sectors of the Irpinia region, being considerably higher in the central and southern sectors. The cross-section representations are much more populated than the analogous projections from the manual catalog, marking the presence of shallow (depth < 5 km) and deeper (depth > 5 km) seismicity. Some clustered and grouped seismicity is visible in the cross-section representations along the B-B' and C-C' vertical planes, but the kilometeric spatial location uncertainty (mean values for the horizontal and vertical hypocentral position of 1.7 km and 2.5 km, respectively) and the high azimuthal gap for the lowest magnitude events, due to the low number of arrays at which they are recorded, prevent us from

clearly identifying seismicity patterns. Some artificially shallow seismicity (depth shallower than 3km) and the concentration of hypocentres on specific depth discontinuities of the velocity model might be due to the lack of sufficient data to constrain the focal depths. However, from the representation of the epicenter in the left panel of Figure 3.19, we can still appreciate the lack of seismicity in the southeastern edge, as from the manual catalog (Figure 3.18) and De Landro et al. (2015, Figure 2).

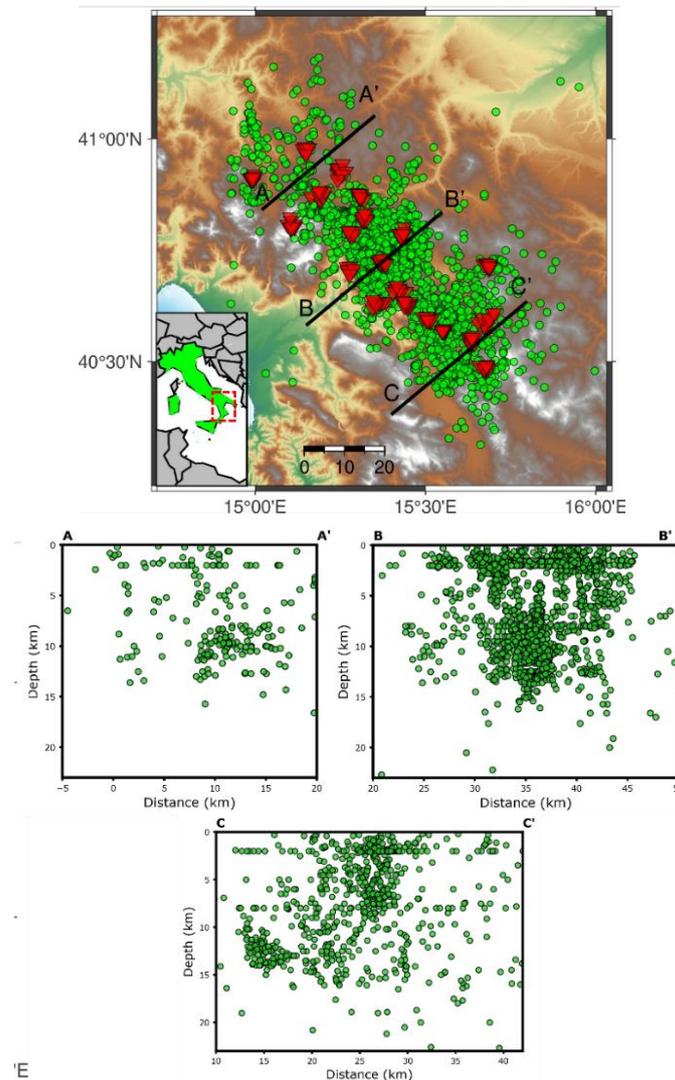


Figure 3.19) Upper panel: Representation of the epicenters from the absolute location of the earthquakes in the enhanced catalogs. The seismicity appears to occur mainly in the central and Southern sector of the area. Lower panels: characteristic cross-section for the area. Absolute locations do not highlight seismicity alignments with activated structures.

To improve the accuracy of earthquake location, we performed double difference relocation of the enhanced catalogs. More than 80% of the automatic phase arrival times were refined (Muzellec et al., 2024), with an average corrective term within ± 0.1 s for both P and S phases, implying the possibility to finely correct the phase arrival times for achieving decametric scale accuracy. The corresponding distribution of the phase residuals is reported in Figure 3.20.

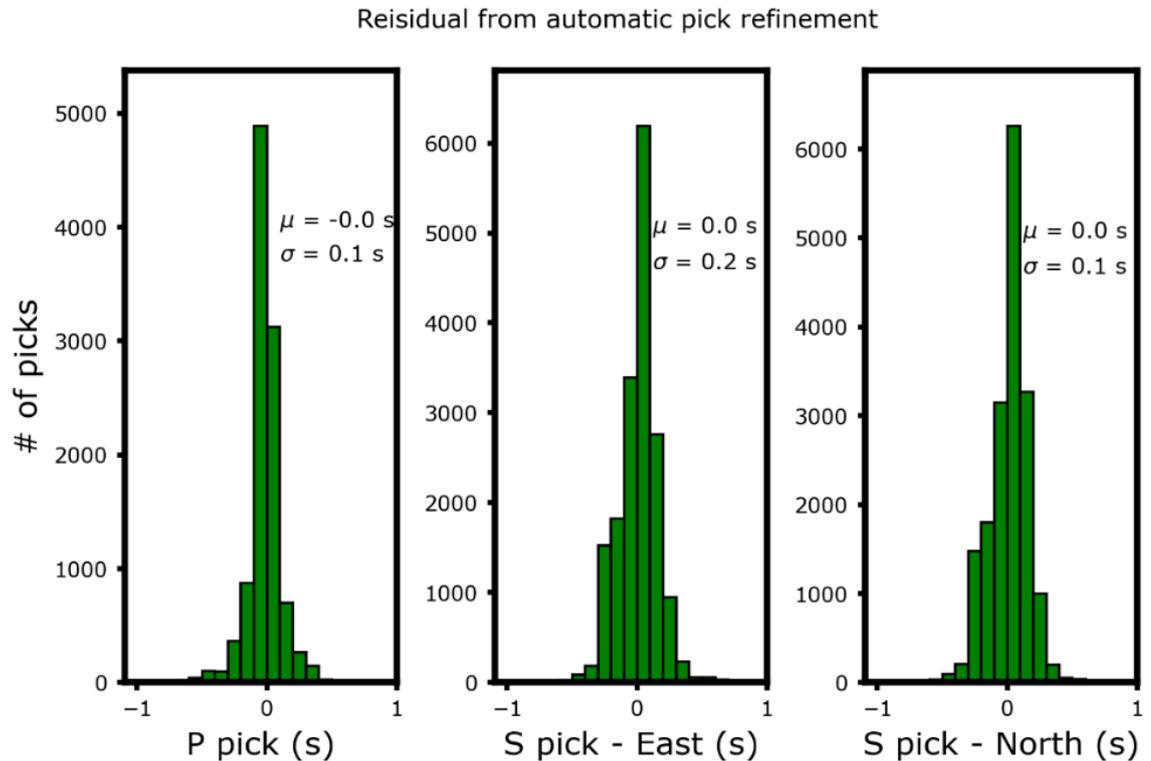


Figure 3.20) Time residuals among the automatic and phase arrival times refined through hierarchical clustering and cross-correlation alignment. The three distributions suggest that the automatic identifications are coherent but slight corrections are necessary to appreciate the small-scale spatial distance among the events.

Using the refined phase arrival times, the absolute hypocentral locations were computed in a 3D model for the area (De Landro et al., 2015) to account for lateral velocity variations, and we selected these locations as starting positions for the double difference relocations. We obtain double-difference relocation for 2248 earthquakes, which represents $\sim 65\%$ of the detected earthquakes in the enhanced catalog. This percentage is almost twice as large as the fraction of the relocated earthquakes in the enhanced catalogs for seismic sequences generated by adopting the same detection strategy and using the ordinary seismic network (Scotto di Uccio et al., 2024a). We address this result with the

use of dense arrays for seismic monitoring, providing a larger number of absolute phase arrival times even for low-magnitude events, translating into a huge number of catalog and cross-correlation differential travel times to be used during the event relocation. In Figure 3.21, we report the epicentral distribution for the relocated seismicity, coloring the cross-sections in the right panel according to the occurrence time of the earthquakes and splitting the dataset into two time periods to avoid overcrowding in the projections of the hypocenters.

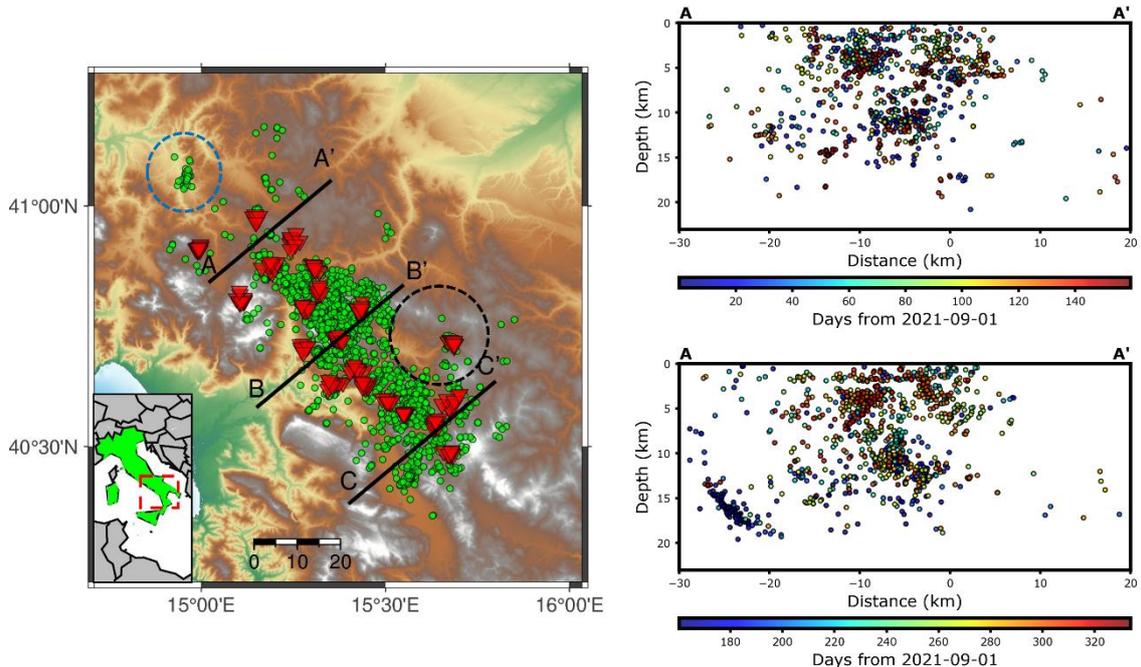


Figure 3.21) Double difference relocation of the enhanced catalog for the DETECT experiment, projected along the A-A' vertical plane. In the left panel, the epicentral representation confirms higher seismicity in the central and Southern sector of the area, while strongly marking the lack of seismicity in the southeastern edge, within the black dashed circle. The cross-sections in the right panel reports the presence of sparse shallow seismicity within 5 km of depth, while the deeper seismicity tends to appear more clustered in space and time. The cluster within the cyan circle represents a swarm-like sequence of 30 earthquakes, associated with the occurrence of a M_l 2.1 event in Dentecane (AV), located outside the array deployment

The epicentral representation of the relocated earthquakes in the enhanced catalog (left panel of Figure 3.21) further marks the observations that came out from the corresponding representation of the absolute locations. The seismicity in the central and southern sectors of the area is much more active than the corresponding seismicity in the Northern part. The seismicity in this latter section is mainly dominated by isolated events and a single swarm-like sequence of 30 earthquakes, associated with the occurrence of a M_l 2.1 event in Dentecane (AV), located outside the array deployment and marked by a cyan circle in

Figure 3.21. We performed a clustering analysis using DBSCAN to identify seismicity clustered in time and space, requiring a minimum number of 10 members for declaring a family. We declared 22 seismicity clusters, with the three most populated families composed of 110, 92 and 85 earthquakes, which occurred in the central and southern sectors within three swarm-like sequences culminated that with M_l 1.7, M_l 1.8 and M_l 1.9 events, respectively. While the shallow seismicity appears to be sparse and diffuse, without being systematically clustered in space and time, the deeper seismicity mainly occurs close in space, illuminating seismicity patterns, such as the brown cluster in Figure 3.22. We found that almost 50 % of the earthquakes that occurred deeper than 5 km belong to a seismic cluster, while this percentage is reduced to 20 % when considering the shallow seismicity.

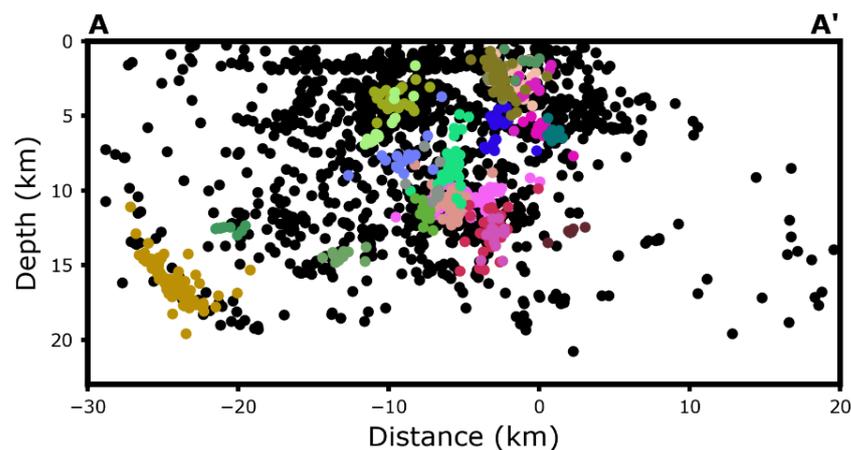


Figure 3.22) Clustering analysis using DBSCAN for identifying seismicity clusters. We declared 22 seismic clusters populated by at least 10 earthquakes.

To better understand the seismicity in the different sections, we provide similar cross-sections for each region of the Irpinia in Figure 3.23, reporting only the earthquakes occurring within 15 km from each vertical plane, splitting the dataset in two time periods. As depicted by the double difference relocation of earthquakes in the DETECT catalog, the seismicity varies in depth and location in a SE-NW direction. To the SE, seismicity extends well within the Apulian carbonate platform along and around a previously identified, SE-dipping, long-lived and multiply reactivated major fault (Amoruso et al., 2014; Amoruso et al., 2005, 2011; Figure 3.24). However, seismicity becomes shallower northwestwards, and at the same time it steps to the right, before deepening again further north. This is interpreted as a result of fault segmentation, which involves the occurrence of a major dextral fault step to the NW (Figure 3.24), as also discussed in Camanni et al., (2024). The deep DETECT seismicity to the north illuminates the multiply reactivated, right-stepping, northern segment of the fault, which represents the third rupture episode

of the 1980 Irpinia earthquake. However, shallow seismicity in the fault step area is likely associated with extension-related detachment of rock volumes along the melange sitting on top of the Apulian carbonate platform (Mazzoli et al., 2001). This fault step is also clearly illuminated by seismic tomography (Amoroso et al., 2014). On the other hand, the fault associated with the first segment of the 1980 Irpinia earthquake (Ascione et al., 2013), being located in its hanging wall, is very poorly illuminated during the DETECT (Figure 3.24).

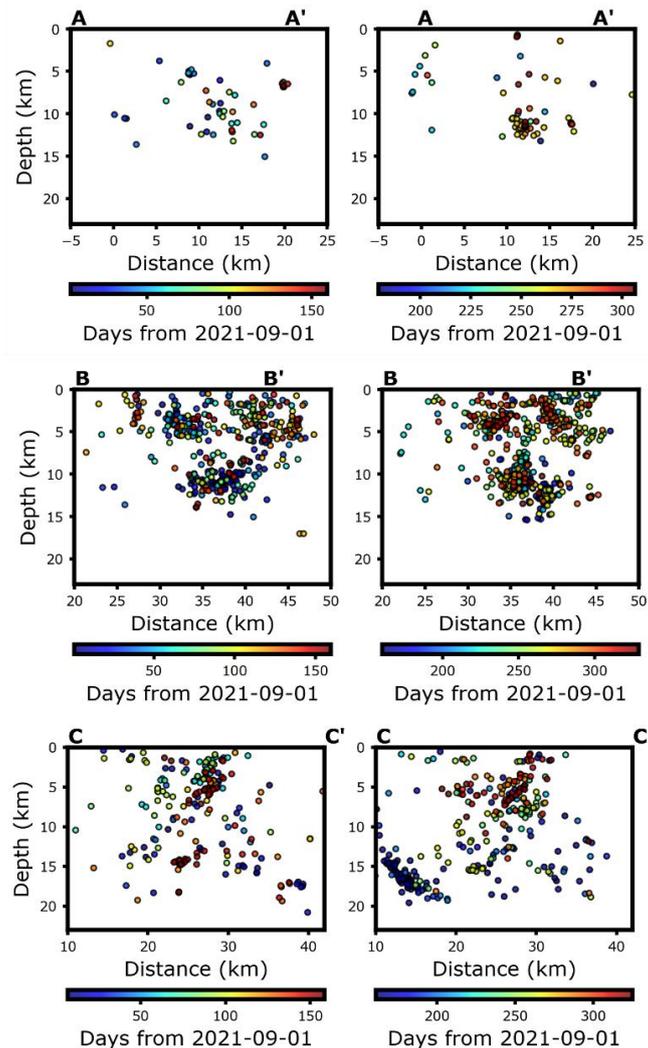


Figure 3.23: Cross-sections along vertical planes oriented N40E for the Northern, Central and Southern region of the Irpinia, as in Figure 3.18 and Figure 3.19. The left panels refer to the period from September 1st 2021 to February 8th 2022, while the right panels contain the event from this latter date up to the end of the experiment.

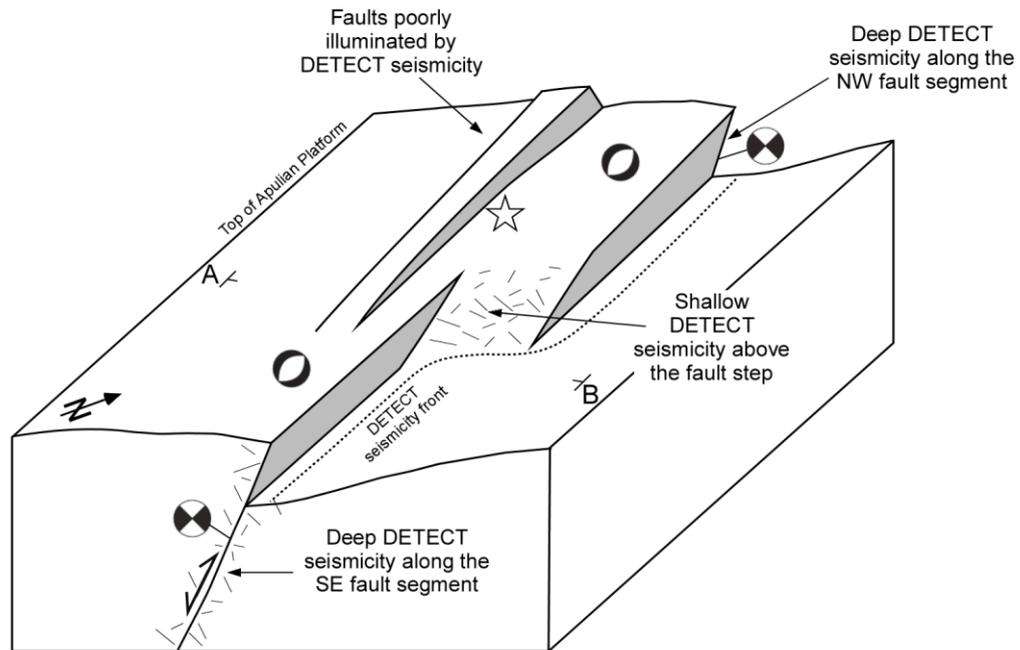


Figure 3.24) Schematic representation of the step fault identified from the analysis of the DETECT seismicity

3.3.3 Delineation of fault zones activated during the 2014–2024 unrest at the Campi Flegrei Caldera (Southern Italy)

The Campi Flegrei caldera is an Italian high-risk volcano experiencing a progressively more intense long-term uplift, accompanied by increasing seismicity and geochemical emissions over the last two decades (Giudicepietro et al., 2024). The Campi Flegrei volcano is characterized by a nested caldera structure (Orsi, 2022; Vitale & Isaia, 2014), produced by two large explosive eruptions, referred to as the Campanian ignimbrite (CI) and the Neapolitan yellow tuff (NYT), at 39 ka and 14.5 ka, respectively (Orsi et al., 1992; Rosi et al., 1996). A series of ground uplift-subsidence episodes (bradyseism), associated with seismic activity, affected the central area of Pozzuoli since the early 1950s (Del Gaudio et al., 2010), with the two most rapid uplift phases occurring in 1970–72 and 1982–84, producing over 20,000 shallow earthquakes during the 1982–84 crisis (D'Auria et al., 2011). A long subsidence phase occurred between 1985 and 2005, with a relatively rare seismicity (Gaeta et al., 2003). Since 2005 a new monotonic uplift phenomenon started with unsteadily accelerating seismicity (Bevilacqua et al., 2022), especially from

2014 onwards, featuring even rates of 15 mm/month, which produced a clear increase in the number of seismic events and maximum magnitude, as reported in Figure 3.25.

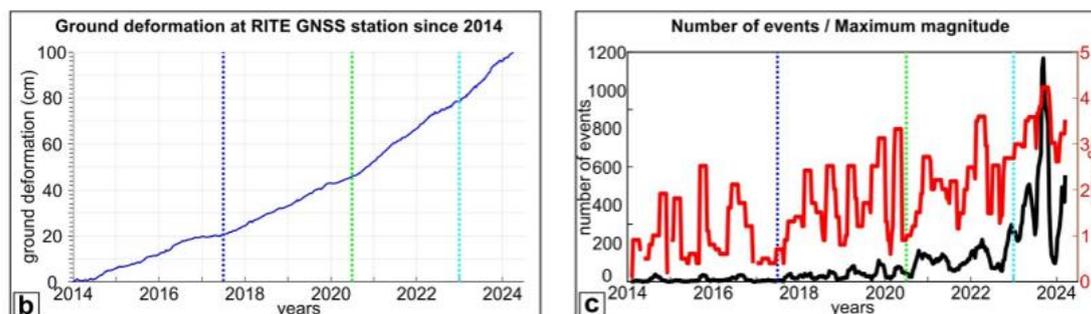


Figure 3.25) Left panel: Ground deformation recorded at the RITE GNSS station from 2014 up to March 2024. The ground uplift reached the value of 100 cm. Vertical dashed lines mark the increase in the uplift rate, experienced in mid-2017, mid-2020 and end of 2022. Right panel: Temporal distribution of the number of earthquakes (black curve) and their maximum magnitude (red curve) in the same time interval. Both curves feature a marked increase, leading to seismic rates of several hundreds of events per month. Figure from Scotto di Uccio et al. (2024b).

The cause of the bradyseism is strongly debated (e.g., Troise et al., 2019). The main hypotheses are that the deformation is either directly caused by pressure and/or volume changes induced by magma emplacement and intrusion at shallow depths beneath the caldera (Macedonio et al., 2014; Woo & Kilburn, 2010) or that it is due to the poroelastic response of the shallow hydrothermal system to changes in pore pressure and fluid content (Bonafede, 1991; Nespoli et al., 2023; Todesco, 2021).

The Campi Flegrei volcano hosts a highly advanced, permanent multiparametric monitoring system (Bianco et al., 2022), including a dense seismic network consisting of 21 inland stations that span, with variable density, through the whole caldera area. The highest density of stations occurs around Solfatara and Pisciarelli, sources of most of the historically recorded seismicity. The southern portion of the caldera is submerged and has represented a break in the network topology until the deployment of 4 marine underwater multi-parametric stations, which are part of the MEDUSA marine monitoring system (Iannaccone et al., 2018; see <http://portale.ov.ingv.it/medusa>). In the central part of the caldera, where most of the current seismicity occurs, the network can also locate earthquakes with a duration magnitude close to -1 (Bianco et al., 2022).

Most of the earthquakes in the caldera occur at depths shallower than 3 km, showing a near-elliptical epicentral distribution at the broad caldera scale, as depicted from the reference catalog of the Istituto Nazionale di Geofisica e Vulcanologia - Osservatorio Vesuviano (INGV-OV, <https://terremoti.ov.ingv.it/gossip/flegrei/2024/index.html>). Most of the seismicity occurs inland, beneath the Solfatara-Pisciarelli area and north of

Inspecting the content of the manual catalog, we note that most of the events feature shallow depth and low magnitude values. In Figure 3.27 we report the depth and magnitude distributions for the considered earthquakes.

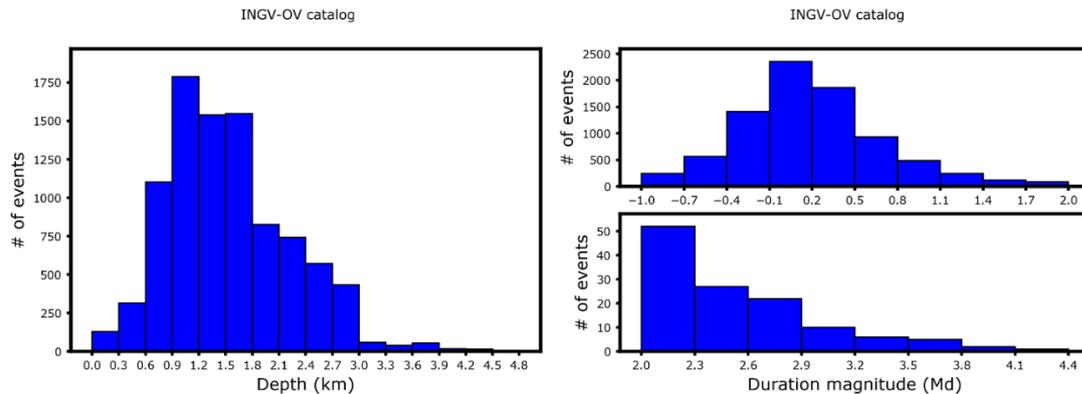


Figure 3.27 Left panel: Depth distribution for the located earthquakes in the INGV-OV catalog from 01/01/2014 to 14/03/2024. Right panel: Duration magnitude (Md) distribution for the earthquakes that occurred within the same time interval.

We observe that more than 75 % of the earthquakes feature depths shallower than 2 km, with a clear different distribution for the onshore and offshore seismicity. While the onshore seismicity mainly occurs within the Solfatara-Pisciarelli area and features a median depth of 1.4 km, the offshore seismicity is typically twice deeper, with a median depth of 2.5 km. The median horizontal and vertical uncertainties in the INGV-OV hypocentral coordinates are 200 m, although they are crudely characterized with steps of 100 m. Moreover, grid effects emerge when just zooming on the epicentral representation, suggesting the necessity of improving the location quality using advanced location techniques for highlighting seismogenic areas. On the other hand, the magnitude distribution suggests that most of the seismicity is represented by microseismic events, featuring duration magnitudes $M_d < 1$ (92% of the located earthquakes in the INGV-OV catalog). 114 events featured $M_d > 2$, while only 16 earthquakes were characterized by $M_d > 3$. However, the generally shallower hypocentral depths, as compared to the tectonic seismicity, make these earthquakes widely felt, raised concerns among the population (850k inhabitants within the area affected by the ground deformation) and public authorities about the impact of seismic activity on buildings and infrastructure in the area.

We used P and S arrival-times from the earthquake catalog provided by the INGV—Osservatorio Vesuviano from 01/01/2014 to 14/03/2024 (<https://terremoti.ov.ingv.it/gossip/flegrei>), accessing the information for individual events. For these events, M_d ranges between -1.1 and 4.2 , with the $M_d 4.2$ event (27-09-2023 01:35:34) having the largest number of picks (18 P-, 6 S-picks). Lower

magnitude events ($M_d < 2$) typically show 6–10 P, 2–4 S arrival times. We extracted arrival times from 18 stations of the INGV network (triangles in Figure 3.26), located within 15 km from the epicenters. Unfortunately, since not all the station recordings are publicly available, we recovered vertical component waveforms from nine velocimetric stations available on EIDA portal (<https://eida.ingv.it>; yellow triangles in 3.26) for evaluating the waveform coherence (Lomax & Savvaidis, 2022).

We obtained multi-scale high-precision earthquake relocations with NLL-SSST-coherence, which combines source-specific, station traveltime corrections (SSST) and stacking of probabilistic locations for nearby events based on inter-event waveform coherence (Lomax and Savvaidis, 2022). We generated SSST corrections which vary smoothly throughout a 3D volume to specify a source-position-dependent correction for each station and phase type. These corrections account for 3D variations in the velocity structure and corresponding distortion in source-receiver ray paths.

In a second relocation stage, NLL-SSST-coherence reduces aleatoric location error by consolidating information across event locations based on waveform coherency between the events, expressed in terms of cross-correlation coefficient among couples of events located within 2 km after NLL-SSST relocation. The stacking over probabilistic locations for nearby events can be used to reduce the noise in this information and improve the location precision for individual, target events. In Figure 3.28 we show the vertical component of velocity records for a cluster of events, having cross-correlation above 0.5 in the frequency band [2 – 10] Hz, recorded at the station COLB. We characterized the events in terms of the cross-correlation coefficients with the master event, displayed in the upper panel. Adopting NLLoc-SC, we relocated ~9000 earthquakes in the INGV-OV catalog from 01/01/2014 to 14/03/2024, using a smooth version of the velocity model extracted from Calò & Tramelli (2018). In Figure 3.29, we report the epicentral distribution of the relocated earthquakes, color-coded and scaled according to the duration magnitude.

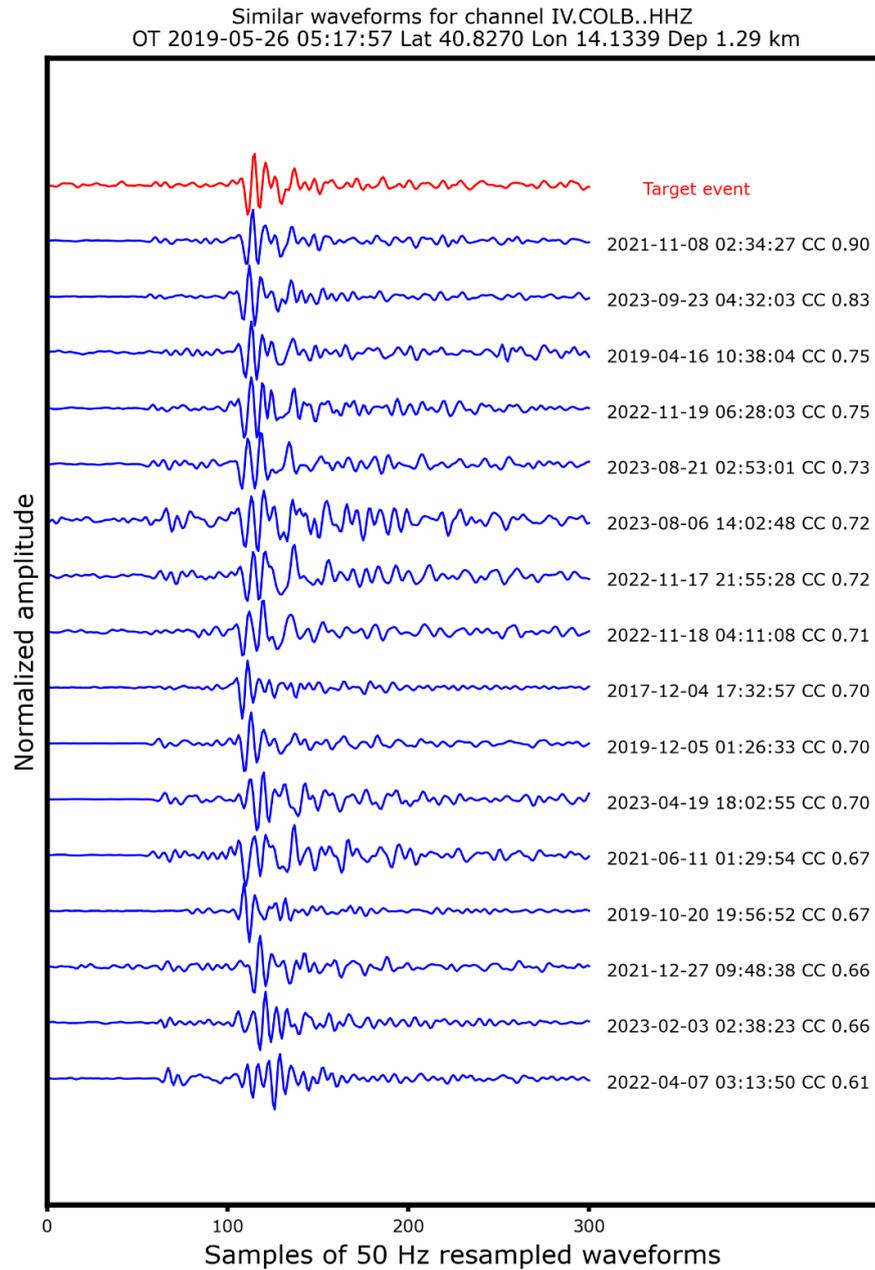


Figure 3.28) Cluster of similar earthquakes recorded at station COLB. The coherency is evaluated in terms of the cross-correlation coefficient with the master event, reported in the upper panel of the figure.

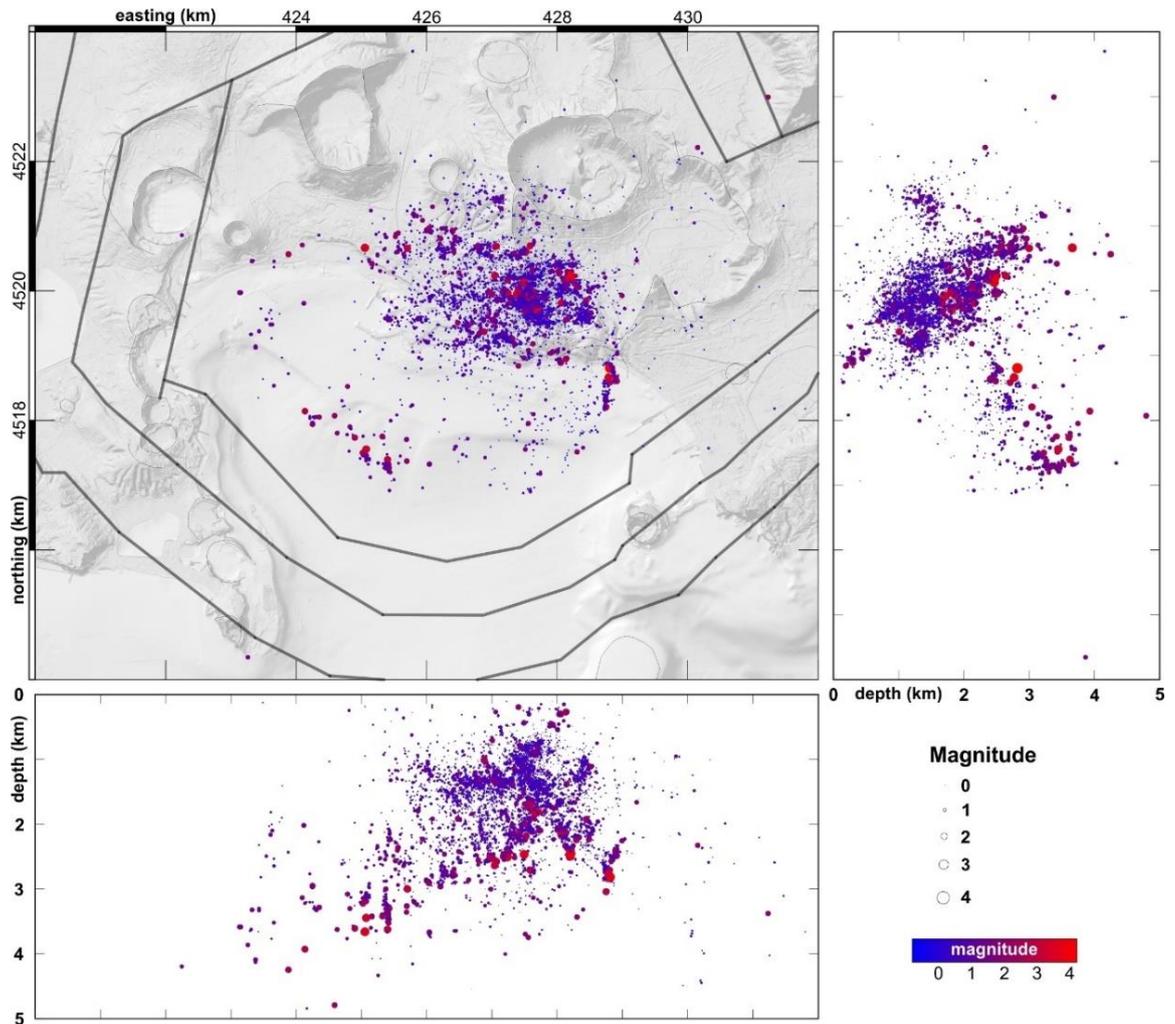


Figure 3.29) Relocated NLL-SC seismicity 2014–2024 ($M_d \geq -0.8$, 8133 events) represented as circles color-coded and scaled according to M_d .

Before proceeding with the interpretation of the seismicity clusters, we inspected the location uncertainties within each relocation stage, which are reported in the following Figure 3.30. We observe a significant decrease in the RMS of travel time residuals when comparing NNLoc to NNLoc-SSST results, indicating that station corrections tailored on source positions (SSST) significantly reduce the misfit of travel times, measured by the RMS. However, the spatial location uncertainties do not diminish significantly and show similar median values. When integrating waveform coherence in the procedure, we observe a dramatic reduction of location uncertainties by a factor of 4 (median values of 0.13 km and 0.15 km on horizontal and vertical errors, respectively), while the RMS of

travel time residuals is similar to the value from NLoc locations. We ascribe this latter behavior to the fact that NLoc-SSST-coherence does not minimize the fit of the arrival times for single events, as it happens in the SSST stage. Instead, it combines multiple location results stacking the PDFs. This yields a minimization of the average residuals for events having similar waveforms. Indeed, the single event RMS of travel time residuals in this case is even less informative of location quality, because it does not concern the cost function to be minimized (Husen & Hardebeck, 2010; Michele et al., 2019).

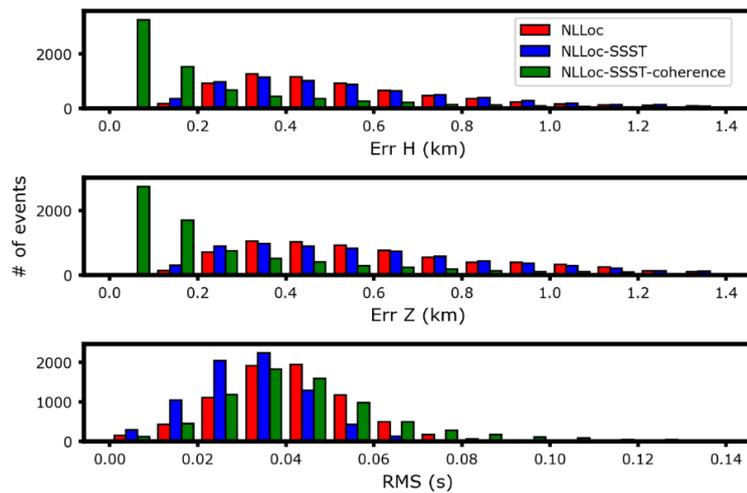


Figure 3.30) Histograms of location uncertainties measured by vertical (top panel) and horizontal (middle panel) errors in km and RMS of travel time residuals in s (bottom panel). We found a dramatic reduction in the location errors due to the waveform coherence step of the location procedure.

Nevertheless, waveform similarity allows to cluster events, significantly reducing spatial relocation uncertainties. We finally compared the location uncertainties obtained within this study with those extracted from the catalog of De Siena et al. (2024) (accessible at <https://osf.io/9bdh2>), which contains located events from 2005 to 2019, through the NLoc software, using the 3D velocity model of Battaglia et al. (2008). We extracted ~1100 common earthquakes in the 2014 – 2019 time window, whose distributions are reported in Figure 3.31. We report location errors compatible values with those resulting from the use of the 3D velocity model. Furthermore, the histograms of the spatial location uncertainties obtained using NLoc-SSST-coherence are mostly populated within the first bin ($[0 - 100]$ m), while the uncertainties estimated in De Siena et al. (2024) mainly populate the $[100 - 200]$ m bin. This result supports the role and the effect of the SSST correction in the relocation stage as an alternative to the 3D velocity models.

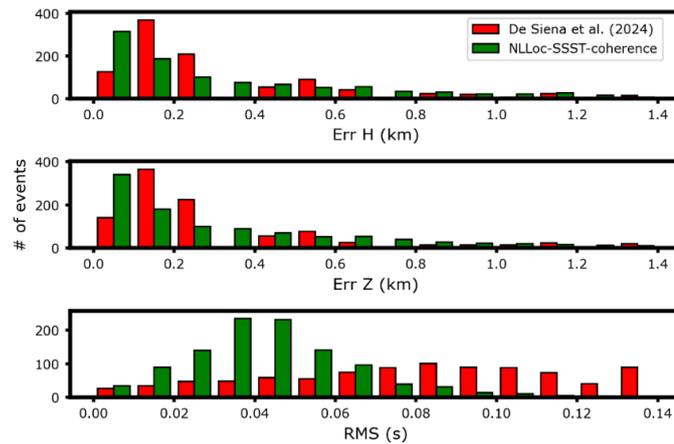


Figure 3.31) Histograms of location uncertainties from this study and from De Siena et al. (2024), for the earthquakes that occurred within 2014 and 2019, located using NLLoc and the 3D velocity model of Battaglia et al. (2008).

The high-precision NLL-SC locations delineate several clusters and alignments of seismicity produced during the ongoing unrest at Campi Flegrei. Most of the seismicity concentrates in the shallow region around the Solfatara-Pisciarelli area. Here, epicenters define a $\sim 1 \text{ km} \times 1 \text{ km}$, horseshoe-shaped structure, opening and deepening toward the northeast beneath the Agnano Plain, and slightly larger than the $\sim 0.5 \text{ km}$ diameter of the Solfatara crater. Smaller-scale shallow seismicity clusters, with a typical size of 100–300 m, occur south and southwest of the Solfatara, along the coast of Pozzuoli. The most recent magnitude Md 3.6+ events, except for the largest magnitude Md 4.2, also occurred in the Solfatara-Pisciarelli area, beneath the horseshoe-shaped seismicity, at depths between 2 and 3 km. Northwest of the Solfatara crater, the seismicity depicts an E–W trending, 1.5–2.0-km-long structure composed of clusters at depths 2–3 km comparable to those of the major events in the Solfatara. Southeastward, off the coast of Bagnoli, a $\sim 1 \text{ km}$ -long, sub-vertical alignment, trending \sim N–S is well defined by the relocated seismicity. This alignment hosts an Md 4.2 event which ruptured an area with characteristic size of 400–700 m, according to the estimated source radius (Scotto di Uccio et al., 2024b). Offshore, to the southwest, the seismicity deepens, down to $\sim 5 \text{ km}$, forming a WNW-oriented alignment offshore of Bacoli, and a N–S alignment off the coast of Monte Nuovo/Baia. When observed on a large spatial scale, this seismicity forms a near-elliptical shape, punctuated by the alignments and clusters containing the larger magnitude (Md > 3) events, with a lack or decreasing rate of seismicity offshore in the southeastern and northwestern sectors of the caldera. To understand the spatio-temporal evolution of the seismicity, we report the epicentral map for different time intervals in Figure 3.32, guided by the seismic rates observed in Figure 3.25.

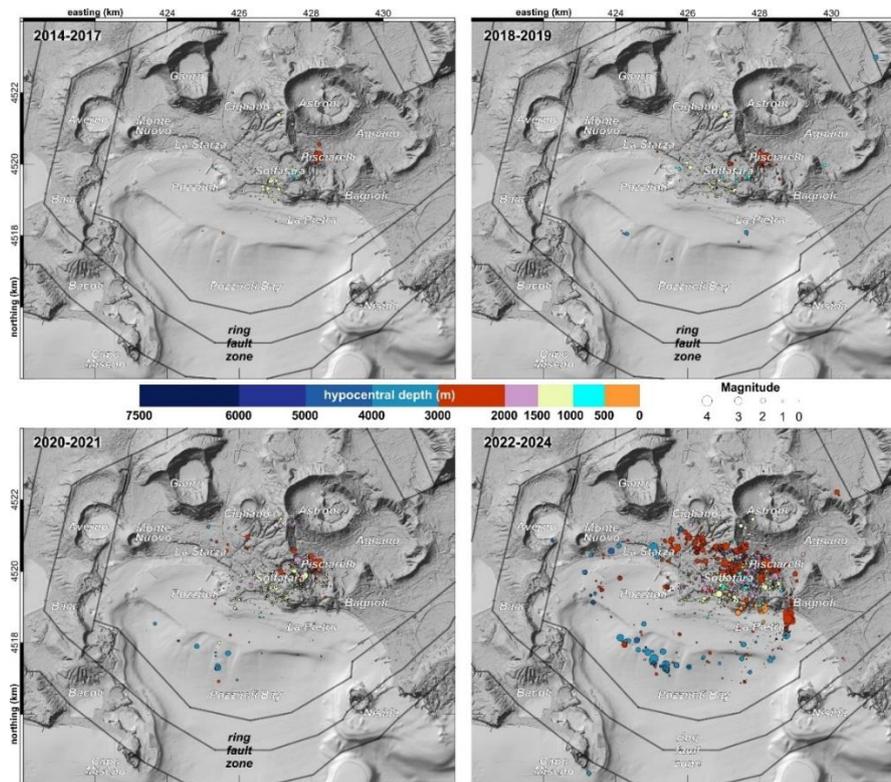


Figure 3.32) Spatiotemporal evolution of the seismicity in periods 2014–2017, 2018–2019, 2020–2021 and 2022–2024. The epicenters, scaled according to M_d , are reported for earthquakes featuring $M_d \geq -0.8$ and ellipsoid major axis ≤ 2.0 km.

The time evolution of the seismicity shows an increase in both the number of events and the maximum magnitude. In the period 2014–2019 a low seismicity rate is observed, mostly characterized by small magnitude ($M_d < 2$) events occurring at depths shallower than 3 km. These events are located within a 1–2 km radius from the Solfatara crater which hosts, together with the adjacent Pisciarelli fumarolic field, the most vigorous hydrothermal activity in the caldera (Chiodini et al., 2017; Tamburello et al., 2019). The inspection of the 2014–2017 seismicity highlights a pattern coherent with the transfer structure discussed by Petrosino and De Siena (2021), which connects the uplifting central sector to the western Agnano plain bounding faults. During 2019–2024 the seismicity deepened, extended offshore and increased in maximum magnitude, while in the last 2 years (2022–2024), the seismicity spreads to a larger area, forming the elliptical, ring-like structure, extending from inland north of Solfatara southwards through Bagnoli, westward toward Bacoli and Monte Nuovo. Uplift velocity rather than cumulative uplift seems to control localized seismicity production with the progressive activation of relatively long fracture zones at the margin of the uplifting resurgent dome (Bevilacqua

et al., 2022; Tramelli et al., 2022). In Figure 3.33, we report the spatial distribution of relocated seismicity for an integrated geo-structural interpretation based on up-to-date reconstructions, integrating the known geological features and the focal mechanisms estimated for the $M_d > 3$ earthquakes.

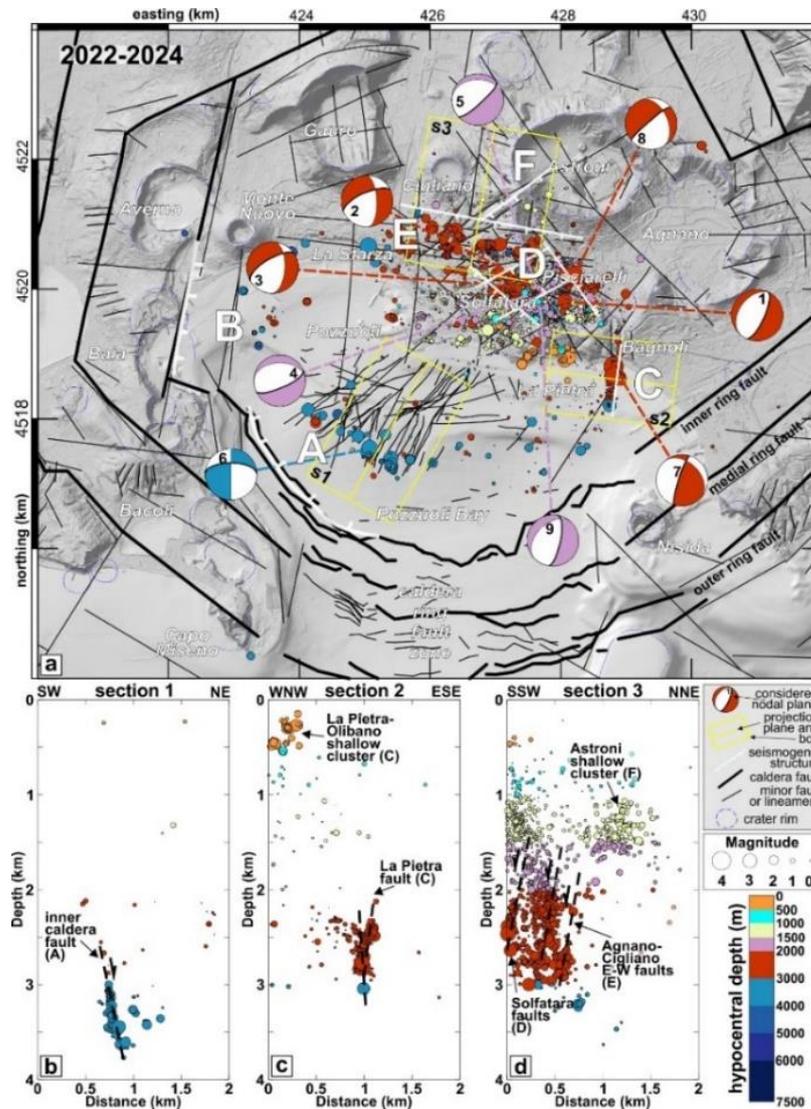


Figure 3.33) Simplified structural map showing the relationship between the epicentral distribution of relocated seismicity in the 2022–2024 period with the elliptical pattern and the main volcano-tectonic structures. Black lines map known geological features. Focal mechanisms solutions for selected 2023 $M_d > 3$ events (ID according to Scotto di Uccio et al., 2024b) with their color coded by depth. The cross-sections for each vertical planes (panels b, c and d) report earthquakes occurred within the corresponding yellow box (s1, s2 and s3).

The near-elliptical shape observed at the global scale of the caldera qualitatively resembles that of the 1982–84 crisis (e.g., Scarpa et al., 2022), whose seismicity distribution was related to a central collapsed portion of the caldera (De Natale et al., 2006). Only a part of the relocated seismicity, occurring in the offshore sector (Features A and B in Figure 3.33), is compatible with the caldera ring fault zone (e.g., Natale et al., 2022). This offshore part appears related to the innermost portion of the composite caldera ring fault zone, recently mapped with high-resolution seismic reflection profiles. Overall, only the deepest offshore seismicity, between 3 and 4 km depth, appears to fit and approximate the downward propagation of the south–western inner ring fault (Feature A in Figure 3.33), where the most frequent dip angles are between 60° and 80° (Natale et al., 2022). This is consistent with a steep ($\sim 70^\circ$) inward-dipping fault structure that justifies the 1.2 km horizontal gap between the surface projection of the mapped inner-ring fault and the ~ 4 km deep epicenter locations.

A novelty of the unrest is the activation of a \sim N–S trending sub-vertical fault structure just offshore La Pietra (Feature C in Figure 3.33), which generated the largest magnitude (Md 4.2) event and is overall producing earthquakes between 2 and 3 km depth. This structure was not identified to date, since it lies in a region where no deep-penetrating reflection profiles are available, and there was no particular seismicity clustering in the 1982–84 crisis (e.g., Scarpa et al., 2022). From spectral modeling of seismic displacement, the average seismic moment and corner frequency of the event indicate a southward rupture extending over ~ 750 m (Scotto di Uccio et al., 2024b) which is consistent with the rupture size delineated by nearby seismicity, and the calculated focal mechanism (event 7 in Figure 3.33, Scotto di Uccio et al., 2024b). The stress drop estimated for the Md 4.2 event (2–3 MPa) is quite large for its hypocentral depth, suggesting a high strength of rocks in the shallow caprock or underlying volcano-sedimentary basement (Vanorio & Kanitpanyacharoen, 2015).

In the Solfatara area (Figure 3.33, Feature D) the relocated seismicity matches well several fault arrays mapped on the surface and subsurface geology. These fault arrays are related to the maar-diatreme structure of Solfatara, whose polygonal shape is controlled by the main NW–SE and NE–SW faults, locally cross-cut by smaller E–W faults exposed at Pisciarelli fumarole field across the western rim of Agnano caldera (Isaia et al., 2021). Hence, the horseshoe distribution of seismicity, deepening northeastward, fits with the presence of such an array of faults at depth (Isaia et al., 2021), which strongly controls the hydrothermal circulation in the area (Troiano et al., 2019).

An approximately E–W-trending fault bounds the distribution of the relocated seismicity NE of the Solfatara crater (Figure 3.33, Feature E), on which a series of spatially and temporally correlated seismicity bursts occurred between 2 and 3 km depth. This structure corresponds to a south-dipping normal fault with a left-lateral component, with noticeable

surface expression at Agnano and Cigliano and corroborated by structural field data (Vitale et al., 2022). This extensional structure roughly matches a similar feature stably detected by ambient noise polarization (Petrosino & De Siena, 2021). The bursts of seismicity occur along a 5-km-long structure, which reaches La Starza marine terrace to the west (Vitale et al., 2019), and is buried beneath the Agnano plain (Troiano et al., 2022). Given its extension and assuming a stress drop as for the Md 4.2 earthquake, this structure might accommodate up to $M_w \sim 5.1$ earthquakes.

The NE–SW seismicity alignment (Feature F) on the western end of the Astroni crater is associated with the occurrence of seismicity bursts ($\sim 50\%$ of events occurred within 12 hr from the previous ones) eventually caused by pressurized fluids moving along NE–SW faults within the shallow (1.0–1.5 km) portion of the hydrothermal system.

We can conclude that several structures delineated by the ongoing seismicity have correspondence in the shallow geological fault record, whose formation was not necessarily related to the same volcanic-tectonic process (i.e., dome resurgence), but rather generated by other, more energetic processes, such as caldera collapse, minor volcano-tectonic collapses, magma intrusions/migration. In general, the stress changes caused by the ongoing uplift of the central caldera appear to concentrate on weaker, pre-existing structures that are reactivated by small-to-moderate, sub-kilometric fractures. All the Md 3.6+ earthquake ruptures, apart from the largest Md 4.2 event, have nucleated along segments of the complex SW–NE and SE–NW fault system array at the margins of the Solfatara crater. As for the Md 4.2 event, the evidence for relatively high-stress drops and average slip (2–3 MPa, 3–5 cm; Scotto di Uccio et al., 2024b) suggests a possible effect of fluid-driven, pore-pressure increase at these faults that could favor the development of larger size fractures.

Chapter 4

Microseismicity characterization

4.1 Introduction

In the previous chapters, we focused on the identification of earthquakes within continuous ground motion records using advanced detection techniques (Chapter 1) and we determined their locations within the crust using absolute and relative location methods (Chapter 2). Despite these analyses allowing us to infer crucial properties of the earthquake distribution, also in a statistical framework, and to identify seismogenic structures within complex fault systems, a clearer view of the mechanical properties of these structures can only be achieved through an accurate source characterization. In the earthquake location of enhanced catalogs for seismic sequences in the Southern Apennines (Section 3.3.1) we identified kilometeric-size seismicity patterns from hypocenter determination. However, to infer the smaller-scale connection between the occurrence of nearby earthquakes and the evolution of the seismic sequences, we should compare the spatial hypocentral distances with the characteristic extensions of the involved ruptures. In this framework, the analysis of the source parameters can reveal the magnitude of the earthquakes, in terms of moment magnitude M_w , along with its characteristic source dimension r . Source characterization also allows for the interpretation of the geometrical properties of the seismic source. In Chapter 2, we discussed the orientation of the seismogenic structures that were depicted by earthquake relocation, in terms of epicentral alignment and dipping of the identified pattern. These deductions can be confirmed by the investigation of focal mechanisms associated with seismic sources, which can reveal the orientation of the fault and direction of the slip for individual earthquakes within each cluster.

In this chapter, we characterize seismic sources in different frameworks:

- We investigate the source parameters for the relocated earthquakes in the enhanced catalogs of seismic sequences in Southern Apennines, determining the moment

magnitude, the rupture length and the stress conditions characterizing individual earthquakes (Section 4.3.1).

- We build stress release models for the seismic sequences in Southern Apennines, based on the estimated source parameters, for inferring the spatiotemporal evolution and mechanical connection of earthquakes in the seismic sequences (Section 4.3.2).
- We infer the focal mechanisms of low-magnitude events recorded by the dense array deployment, generating an enhanced catalog of geometrical features for the background seismicity in Southern Apennines using a machine learning technique (Section 4.3.3). We use the catalog of focal mechanisms to retrieve the stress field acting in the area (Section 4.3.4)

4.2 Source characterization

4.2.1 Source parameters

The characterization of the source parameters of small to moderate earthquakes is an important step in understanding the general mechanisms of earthquake nucleation and propagation, since it shines a light on the physical processes involving faults over different space and time scales. The analyses of the source parameters have depicted crucial features for understanding the occurrence of earthquakes in different contexts, revealing the magnitude and the area involved in the seismic ruptures (Abercrombie & Leary, 1993; Kwiatek et al., 2011; Zollo et al., 2014; Supino et al., 2019; Picozzi et al., 2022a). Different techniques have been proposed for determining the source parameters, living either in the time domain (Urbancic et al., 1996; Colombelli & Zollo, 2015; Nazeri et al 2019; Al-Ismaïl et al., 2023; Longobardi et al., 2023) or in the frequency domain (Nakano et al., 2008; Picozzi et al., 2017; Supino et al., 2019; Shible et al., 2022). Regardless of the domain of development, these techniques aim to determine the moment magnitude and the rupture length of the considered earthquakes.

In a seismological framework, the fault is a planar surface, which ideally separates two blocks approaching failure due to remote tectonic load. Along the fault surface, the fracture produces a discontinuity in the relative displacement with respect to the two lips of the surface itself (Figure 4.1). It is possible to describe the dislocation process related to the fracture through a function, known as the source function, which represents the relative displacement of two points belonging to the different blocks. Mathematically, the source function can be expressed as the relative displacement of the points along the points on the two sides of the fault as:

$$\Delta u(\xi, t) = u(\xi, t)|^{\Sigma^+} - u(\xi, t)|^{\Sigma^-} \quad (4.1)$$

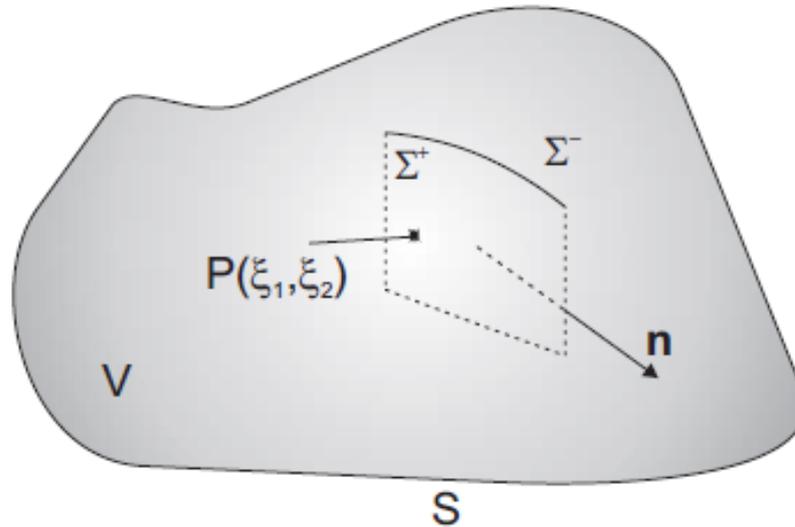


Figure 4.1) Schematic representation of a fault, with area Σ and normal vector \mathbf{n} . The point P is indicated, in a reference frame along the fault, by the coordinates (ξ_1, ξ_2) and the dislocation occurs when the displacement on the Σ^+ side differs from the displacement on the Σ^- side. Figure extracted from Zollo & Emolo (2011)

Quantities such as the fractured area, average dislocation of the fault $\langle \Delta u \rangle$ and the average rupture velocity characterized the fracture process. For seismic waves whose characteristic period is greater than, or comparable to, the duration of the rupture process and for wavelengths higher than the source dimension, it is possible to replace the real, complex rupture physics with its simpler average dislocation model. Thus, in this framework in which the source is assumed to be a geometrical point, the entire dislocation process can be described by a single average dislocation function associated with the barycenter of the fault surface. This latter model is simple enough to be described by an equivalent system of an orthogonal double-couple of forces, oriented in the direction of the slip and normal to the fault surface (Figure 4.2). The moment associated with one of the two couples is referred to as the seismic moment, which the moment magnitude scale is based on (Hanks & Kanamori, 1979).

To describe the seismic source, we can express the observed seismogram at the recording station in \mathbf{x} , $U^*(\mathbf{x}, t)$ as the convolution of the effective source function $S(\xi, t)$, a term accounting for the propagation of the seismic waves within the medium $G(\mathbf{x}, \xi, t)$ and the instrument response $I(\mathbf{x}, t)$:

$$U^*(\mathbf{x}, t) = S(\xi, t) * G(\mathbf{x}, \xi, t) * I(\mathbf{x}, t) \quad (4.2)$$

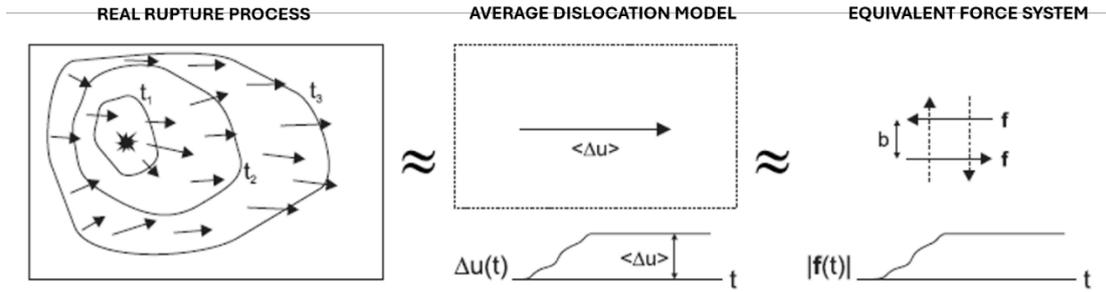


Figure 4.2) The real and complex rupture process can be simplified, under specific conditions, with an average dislocation model, which is analogous to the effect of a system of a double-couple of forces acting along and perpendicular to the fault plane.

Through the convolution theorem, in frequency domain the former relation can be simply expressed as the product of the Fourier transform of the individual operators:

$$U^*(\mathbf{x}, \omega) = S(\xi, \omega) G(\mathbf{x}, \xi, \omega) I(\mathbf{x}, \omega) \quad (4.3)$$

Since the instrumental response is often known, we can remove the effect of the instrument from the recorded ground motion (with U being the corrected displacement) and obtain a formulation for the Fourier transform of the source function as:

$$S(\xi, \omega) = \frac{U(\mathbf{x}, \omega)}{G(\mathbf{x}, \xi, \omega)} \quad (4.4)$$

In this analysis, we adopted the probabilistic inversion approach proposed by Supino et al. (2019) for retrieving the seismic moment and the rupture length from the S-wave displacement amplitude spectra of relocated events. This technique is based on a Bayesian inversion of the spectra and allows for an exploration of the correlations among parameters with a robust uncertainty estimation. The far-field displacement spectra of seismic signals feature characteristic properties for most of the real earthquakes: at low frequencies, the amplitude spectrum presents a plateau level, while, at higher frequencies, the amplitude spectra start decaying following almost a power law. A characteristic displacement spectrum is reported in logarithmic scale in Figure 4.3. The plateau is observed in the frequency domain for which the source is seen as a point. In that domain, interaction among different points of the source cannot be seen due to wavelength resolution, and the source acts as a point with a unique behavior; this is called the flat level of the spectrum Ω_0 , and mathematically, it can be written as the limit:

$$\lim_{\omega \rightarrow 0} U(\omega) = \int_{-\infty}^{\infty} U(t) dt = \Omega_0 \quad (4.5)$$

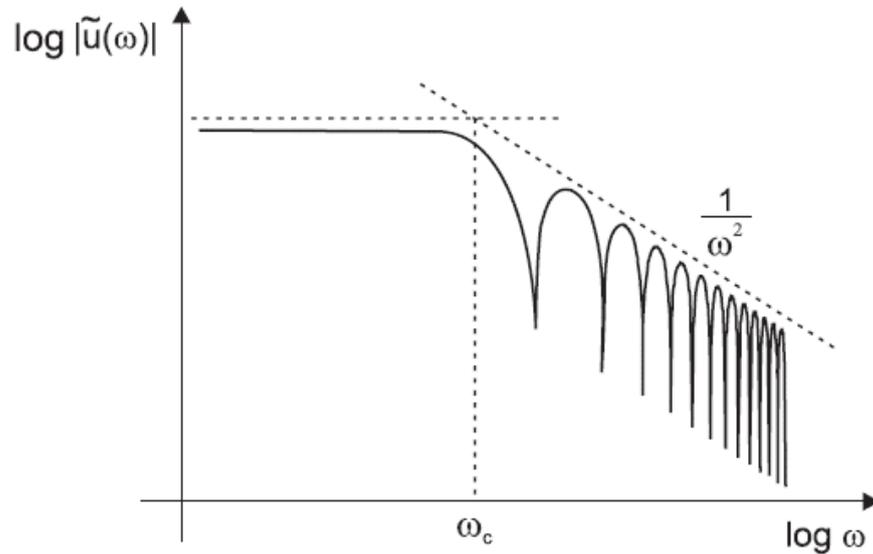


Figure 4.3) Theoretical displacement spectrum for an earthquake. At low frequency, the spectrum presents a plateau level, while at higher frequencies it starts decaying.

In particular, in the point source approximation, Ω_0 can be shown to be related to the total moment released by an earthquake, M_0 (Aki & Richards, 2002). Beyond a certain frequency, namely the corner frequency f_c , the observer has the resolution to distinguish the interactions from different points of the source; this interference effect can be modeled by a power-law. The observed corner frequency is therefore related to the effective size of the source. The simplest expression for the depicted spectral function can be expressed as (Brune 1970):

$$U(f, f_c, \Omega_0) = \frac{\Omega_0}{1 + \left(\frac{f}{f_c}\right)^\gamma} \quad (4.6)$$

If we consider a source function having the shape of a triangle, the representation of the far-field displacement spectrum confirms this relation, with the corner frequency

inversely proportional to the width of the triangle and the seismic moment proportional to the area of the triangle.

In the case of the application of a double-couple point force, under the condition of isotropic and homogenous medium and far-field ($r \gg \lambda$), the radiated wavefield can be expressed as the sum of the contribution of the primary and secondary waves:

$$U_{FF}(r, t) = U_{FF}^P(r, t) + U_{FF}^S(r, t) \quad (4.7)$$

where

$$U_{FF}^P(r, t) = \frac{1}{4\pi\rho\alpha^2} \frac{1}{r} R_P S\left(t - \frac{r}{\alpha}\right) \quad (4.8)$$

$$U_{FF}^S(r, t) = \frac{1}{4\pi\rho\beta^2} \frac{1}{r} R_S S\left(t - \frac{r}{\beta}\right) \quad (4.9)$$

In these relationships, α and β are the velocity of the P and S waves, respectively, S represents the source function and R_P and R_S refer to the radiation patterns of the two waves, respectively. Moreover, these relations suggest that the amplitude of the seismic waves attenuates while propagating from the source to the receiver by a factor of $1/R$, referred to as the geometrical spreading. However, when a wave propagates through a real medium, wave amplitudes further attenuate as a result of the internal heat dissipation of the material; real propagation media are partially anelastic, and anelastic attenuation must be taken into account to model the propagation from the source of an earthquake to the receiver. The effect of the internal friction can be described by an anelastic attenuation factor (quality factor), Q_α , defined as follows:

$$\frac{1}{Q_\alpha(f)} = -\frac{\Delta E}{2\pi E} \quad (4.10)$$

where ΔE represents the energy lost due to non-elasticity, and E is the peak deformation energy of the medium at a given frequency. The anelastic quality factor can be evaluated in the frequency domain through the analysis of the spectral decay. This approach exploits the a priori known spectral shape of a seismic event, reported in Equation 4.6. In this formulation, the low frequency behavior of the displacement spectra typically features a plateau level (Figure 4.3). If we consider the seismic signal recorded at a distance r_1 , which has undergone anelastic attenuation during its propagation from the source, and we

investigate the displacement spectra in the low frequency range, it should report a functional form as:

$$|u(r_1, f)| \cong k e^{-\frac{\pi r_1 f}{\beta Q_a}} \quad (4.11)$$

with k containing information on the geometrical spreading. Extracting the logarithm of the former relation, we can obtain a linear relation expressing the amplitude of the displacement spectra at low frequencies as a function of the frequency itself:

$$\ln(|u(r_1, f)|) = \ln k - \frac{\pi r_1}{\beta Q_a} f \quad (4.12)$$

From the slope of the linear relation, known the propagation time from the source to the receiver and in the low frequency domain, it is possible to estimate the quality factor Q_a . The ambiguities in modelling and isolating the source component from seismograms led to the development of the empirical approach in which a small, co-located smaller earthquake can be used as an empirical Green's function for resolving the propagation effects of higher magnitude events (e.g. Mueller, 1985; Mori & Frankel, 1990). Performing the deconvolution requires no assumptions about source shape, path or site effects, and it can be applied at individual stations to investigate azimuthal variation in the source radiation. It enables the calculation of spectral ratios and deconvolved source time functions, making both spectral- and time-domain source modelling possible. Unfortunately, the requirement for a sufficiently closely located EGF earthquake, that is large enough to have adequate signal-to-noise ratio, but small enough compared to the target event to be an effective Green's function, significantly limits the number of events that can be studied using this method. Furthermore, the need for both earthquake spectra to have signal above the noise level at all analysed frequencies, limits the available frequency range. Also, corner frequencies near the edge of the frequency range will have the largest uncertainties and their estimation will most likely be biased.

Moreover, it is possible to include site effects (Anderson & Hough, 1984), by employing an exponential decay term, $e^{-\pi k f}$, whose characteristic coefficient could be extracted for individual stations and phases.

We can, thus, merge the individual propagation terms (geometrical spreading, anelastic attenuation and site effects) to build the propagation operator G for the S waves as:

$$G(Q_a, k, f) = \frac{R_s F}{4\pi\rho\beta^3 r} e^{-\frac{\pi f T}{Q_a}} e^{-\pi k_s f} \quad (4.13)$$

Where T is the source-receiver travel time, Q_a is the quality factor related to anelastic attenuation and k_s accounts for the site effect. R_s is the average radiation pattern for the S waves (e.g. Boore & Boatwright, 1984) and F is the free surface reflection coefficient.

From the source parameters, we can infer further source properties as the source radius r and the static stress drop $\Delta\sigma$. Madariaga (1976) defined relationships between the observed corner frequency f_c from the displacement spectrum for the P and S waves and the source radius. By analyzing the spectra of the seismic radiation emitted at different angles with respect to the normal vector to the fault surface, Madariaga (1976) obtained average coefficients connecting f_c to the source radius. In particular, for the S waves, he observed that

$$r = K \frac{\beta}{f_c} \quad (4.14)$$

with $K = 0.21$ and β the velocity of the S waves. This relation indicates that the larger the source radius, the lower the resulting corner frequency. Several authors have proposed alternative rupture models, resulting in different values of the K coefficient ($K = 0.37$, Brune, 1970; $K = 0.26$, Kaneko & Shearer, 2014), therefore providing different values for the source radius. Also, different kinematic and dynamic source models have been proposed to infer the stress drop from observations, such as a circular rupture (Brune, 1970; Sato and Hirasawa, 1973; Madariaga, 1976) or a one-dimensional rupture (Haskell, 1964). The static stress drop coincides with the total stress variation during the fracture process, and it is provided by the difference of stress between the initial stage and the stress when the rupture is at the end. Analyzing the properties of the dynamic propagation of the rupture under the condition of uniform static stress release, Madariaga (1977) extracted a relation connecting $\Delta\sigma$ to the average value of the final dislocation along the fault $\langle \Delta u \rangle$:

$$\Delta\sigma = \frac{\mu \langle \Delta u \rangle}{C_\sigma W} \quad (4.15)$$

In the former relation, μ refers to the shear modulus, W is the minor dimension of the fractured surface and C_σ is shape factor related to the geometry of the ruptured area. For circular fractures, an assumption holding for describing the source geometry of microseismic earthquakes, $C_\sigma = 16/7\pi$. The simplest model of a static circular crack is obtained under the assumption of a constant stress drop $\Delta\sigma$ (e.g. Keilis-Borok, 1959), can be expressed, accounting for the definition of the seismic moment M_0 , as:

$$\Delta\sigma = \frac{7}{16} \frac{M_0}{r^3} \quad (4.16)$$

The interpretation and implications of the stress drop is still challenging and controversial within the community. Aki (1967) found the earthquake source to be relatively scale invariant, reported in Figure 4.4, with constant stress drop (albeit scattered across three orders of magnitude), although potentially with some dependence on faulting type or location (Kanamori & Anderson, 1975). Since then, numerous studies of earthquakes over a wide range of magnitudes, in various tectonic settings, involving increasing quantity and quality of seismic data to try and understand better the controlling factors, have led to considerable controversy and uncertainty. For example, some studies find earthquakes to be scale invariant over a wide magnitude range (e.g. Abercrombie, 1995; Kwiatek et al., 2011), whose conclusions agree with the representation of Figure 4.4. On the other hand, other studies concluded that the stress drop is higher for larger magnitude events (Malagnini et al., 2014; Bindi et al., 2020)

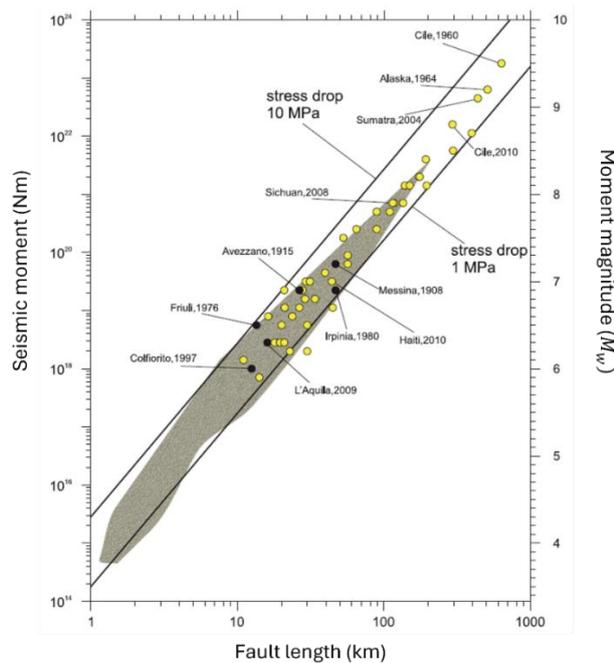


Figure 4.4) Seismic moment as a function of the fault length for different ranges of magnitudes. Continuous black lines mark the theoretical constant stress drop scaling of $\Delta\sigma = 1 \text{ MPa}$, 10 MPa . Yellow and black dots refer to the $M > 6$ earthquakes, while the grey cloud refers to the moderate magnitude seismicity analysed in Madariaga et al. (1991). Source parameters for most of the earthquakes fall within the continuous black lines, supporting the hypothesis of constant stress drop scaling (figure extracted from Zollo & Emolo, 2011)

In this analysis, we adopted the approach proposed by Supino et al. (2019) for estimating the source parameters in the frequency domain, by analyzing the S wave displacement spectra. The authors developed a probabilistic framework based on the conjunction of states of information between data and model, to jointly retrieve earthquake source

parameters and anelastic attenuation factor from inversion the of displacement amplitude spectra. The evaluation of the joint probability density function (PDF) enables us to consider the correlations between the parameters in their final estimations and their related uncertainties. Following this approach, the algorithm searches for the maximum of the a posteriori PDF through the basin hopping technique that couples a global exploration built on a Markov chain with a local deterministic maximization. Then statistical indicators (mean, variance and correlation coefficients) are computed on source parameters and anelastic attenuation through the integration of the PDF in the vicinity of the maximum likelihood solution. The definition of quality criteria based on the signal-to-noise ratio (SNR) and similarity of the marginal PDFs with a Gaussian function, enable us to define the frequency domain for the inversion and to get rid of unconstrained solutions.

4.2.2 Stress release modelling

The evaluation of the stress perturbation on the fault plane due to the occurrence of earthquakes is crucial for assessing the evolution and the mechanical interactions between seismic events. Smaller earthquakes that sometimes occur in the vicinity of the main event's hypocentre might provide insights into its nucleation process. One view is that they are triggered by aseismic slip over an extended area surrounding the eventual mainshock hypocentre (pre-slip model; Ellsworth & Beroza, 1995). If correct, the underlying aseismic slip may be a precursor to the earthquake (Bouchon et al., 2013). Alternatively, they might occur by neighbour-to-neighbour stress transfer between one foreshock and another one without an aseismic slip component (cascade model; Wyss & Brune, 1971; Ellsworth & Beroza, 1995; Ellsworth & Bulut, 2018). Under this interpretation, the foreshocks are no different than any other set of clustered earthquakes, and the mainshock is just a random outcome of triggering (Helmstetter & Sornette, 2003). One approach to discriminating between these end-member alternatives focuses on the spatial and temporal evolution of the foreshocks as revealed by their seismograms (Ellsworth & Bulut, 2018). If aseismic slip drives the foreshock-aftershock process, their hypocentres would be expected, on average, to populate the pre-slip zone. Under the cascade hypothesis, on the other hand, events would cluster in close enough proximity to one another to permit event-to-event triggering from either dynamic and/or static stress changes.

To introduce the stress function, we consider a circular crack with the final slip being maximum at the center of the rupture and smoothly decreasing to zero, with the relationship generalized from the dynamic frictional model of Burridge and Halliday (1971), as proposed by Andrews (1980):

$$D(r) = \begin{cases} D_0 \left[1 - \left(\frac{r}{a} \right)^2 \right]^{3/2} & r < a \\ 0 & r \geq a \end{cases} \quad (4.17)$$

where D_0 is the maximum slip, a is the source radius and r is the distance from the center of the fault. A representation of the final slip is in Figure 4.5.

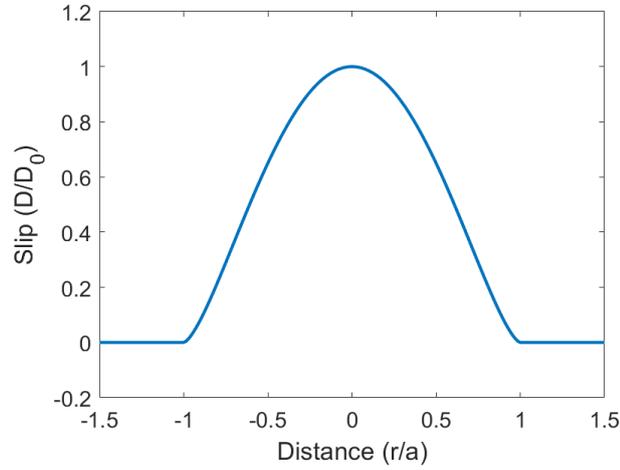


Figure 4.5) Slip as a function of the distance, in the model of Burridge & Halliday (1971).

We compute the 2D Fourier transform of the slip function. Considering the wavenumber $\mathbf{k} = (k_x, k_y) = k(\cos \varphi, \sin \varphi)$, the 2D Fourier transform is

$$D(\mathbf{k}) = \frac{1}{2\pi} \iint_{\mathbb{R}^2} D(x, y) e^{-i\mathbf{k} \cdot \mathbf{r}} dx dy \quad (4.18)$$

Using the azimuthal symmetry of the slip function, we can use polar coordinates to solve the integral and get

$$D(\mathbf{k}) = \frac{1}{2\pi} \int_0^a r D(r) dr \int_0^{2\pi} e^{-ikr(\cos \theta \cos \varphi + \sin \theta \sin \varphi)} d\theta \quad (4.19)$$

since $\mathbf{r} = (x, y) = r(\cos \theta, \sin \theta)$. We note that

$$\int_0^{2\pi} e^{-ikr(\cos \theta \cos \varphi + \sin \theta \sin \varphi)} d\theta = \int_0^{2\pi} e^{-ikr \cos \theta} d\theta = 2\pi J_0(kr) \quad (4.20)$$

where J_0 is the zero-order Bessel function of first kind. Thus, we get that the Fourier transform of the slip only depends on the modulus of the wavenumber and it can be written as

$$D(k) = \int_0^a r J_0(kr) D(r) dr \quad (4.21)$$

which results into the Hankel transform of the function $D(r)$. The integral can be solved analytically, for the function 3.17, yielding

$$D(k) = \frac{D_0}{a^3 k^5} [3(3 - k^2 a^2) \sin(ka) - 9ka \cos(ka)] \quad (4.22)$$

The formula 4.22 can be also written as (Andrews, 1980):

$$D(k) = 2\sqrt{2} \frac{D_0 a^2}{(ka)^{5/2}} \Gamma\left(\frac{5}{2}\right) J_{5/2}(ka) \quad (4.23)$$

Using the isotropic representation, as an approximation of the stiffness function (Andrews, 1980):

$$S(\mathbf{k}) = -\frac{1}{2} \mu k \quad (4.24)$$

This leads to the following representation of the shear stress:

$$S\sigma(\mathbf{k}) = S(\mathbf{k})D(\mathbf{k}) = -\sqrt{2} \frac{D_0 \mu a}{(ka)^{3/2}} \Gamma\left(\frac{5}{2}\right) J_{5/2}(ka) \quad (4.25)$$

where μ is the shear modulus. Since the stress only depends on k , we can use the Hankel anti-transform to get the shear stress in the space domain:

$$\sigma(r) = \int_0^{+\infty} \sigma(k) J_0(kr) k dk = -\sqrt{2} D_0 \mu \Gamma\left(\frac{5}{2}\right) \int_0^{+\infty} \frac{J_{5/2}(ka) J_0(kr)}{(ka)^{1/2}} dk \quad (4.26)$$

This integral can be solved analytically as:

$$\sigma(r) = \frac{D_0 \mu}{2a} \Gamma\left(\frac{5}{2}\right) \begin{cases} \frac{\sqrt{\pi}}{2} \left(3 \frac{r^2}{a^2} - 2\right) & r < a \\ \frac{1}{\sqrt{\pi}} \left[\arcsin \frac{a}{r} \left(\frac{3r^2}{a^2} - 2\right) - 3 \sqrt{\frac{r^2}{a^2} - 1} \right] & r \geq a \end{cases} \quad (4.27)$$

A representation of the function 4.27 is shown Figure 4.6.

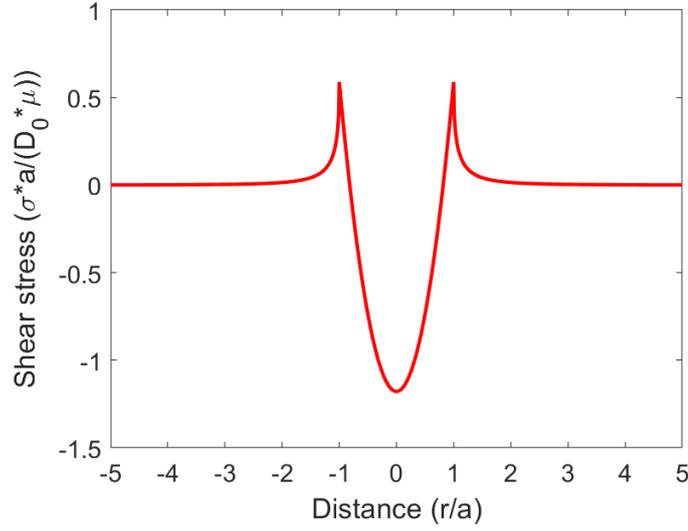


Figure 4.6) Static shear stress from the slip model of Figure 4.5.

We can add the correction to Equation 4.24, to account for the non-isotropic component of the stiffness vector. According to Andrews (1980) the complete stiffness vector is:

$$\begin{aligned} S_c(\mathbf{k}) &= -\frac{1}{2} \frac{\mu}{k} \left(k^2 + \frac{1}{3} k_x^2 \right) = S(\mathbf{k}) - \frac{1}{6} \mu k \frac{k_x^2}{k^2} = S(\mathbf{k}) - \frac{1}{6} \mu k \cos^2 \varphi \\ &= S(\mathbf{k}) + S_a(\mathbf{k}) \end{aligned} \quad (4.28)$$

where x is the direction of the slip and we have assumed a Poissonian medium ($\lambda = \mu$). The stress in the wavenumber domain writes:

$$\sigma_c(\mathbf{k}) = S_c(\mathbf{k})D(\mathbf{k}) = \sigma(\mathbf{k}) + S_a(\mathbf{k})D(\mathbf{k}) \quad (4.29)$$

Because of the linearity of the Fourier transform, we have in the space domain that the stress is obtained as the summation of the contribution from 4.27, plus a correction which comes from the anti-transform of the function

$$\sigma_p(\mathbf{k}) = S_a(\mathbf{k})D(\mathbf{k}) = -\frac{\sqrt{2}}{3} \frac{D_0 \mu a}{(ka)^{3/2}} \Gamma\left(\frac{5}{2}\right) J_{5/2}(ka) \cos^2 \varphi \quad (4.30)$$

The Fourier anti-transform is of this function is

$$\sigma_p(r, \theta) = -\frac{\sqrt{2}}{6\pi} D_0 \mu \Gamma\left(\frac{5}{2}\right) \int_0^{+\infty} \frac{J_{5/2}(ka)}{\sqrt{ka}} dk \int_0^{2\pi} e^{ikr \cos(\varphi-\theta)} \cos^2 \varphi d\varphi \quad (4.31)$$

The inner integral is

$$\int_0^{2\pi} e^{ikr \cos(\varphi-\theta)} \cos^2 \varphi d\varphi = \frac{1}{2} \int_0^{2\pi} e^{ikr \cos(\varphi-\theta)} [1 + \cos(2\varphi)] d\varphi \quad (4.32)$$

$$= \pi J_0(kr) - \pi J_2(kr) \cos(2\theta)$$

The outer integral is composed of two parts, leading to

$$\sigma_p(r, \theta) = \frac{1}{6} \sigma(r) + \frac{\sqrt{2}}{6} \cos(2\theta) D_0 \mu \Gamma \left(\frac{5}{2} \right) \int_0^{+\infty} \frac{J_5(ka) J_2(kr)}{\sqrt{ka}} dk \quad (4.33)$$

$$= \frac{1}{6} \sigma(r) + \sigma_2(r, \theta)$$

Therefore, the final expression can be reported as

$$\sigma_c(r, \theta) = \frac{7}{6} \sigma(r) + \sigma_2(r, \theta) \quad (4.34)$$

with $\sigma(r)$ represented in 4.27 and

$$\sigma_2(r, \theta) = -\frac{1}{24a\sqrt{\pi}} \cos(2\theta) D_0 \mu \Gamma \left(\frac{5}{2} \right) \begin{cases} \frac{3\pi r^2}{2a^2} & 0 < r < a \\ -\frac{(3r^2 + 2a^2)\sqrt{r^2 - a^2}}{ar^2} + 3\frac{r^2}{a^2} \arcsin \frac{a}{r} & r > a \end{cases} \quad (4.35)$$

The solution for the stress is shown in Figure 4.7.

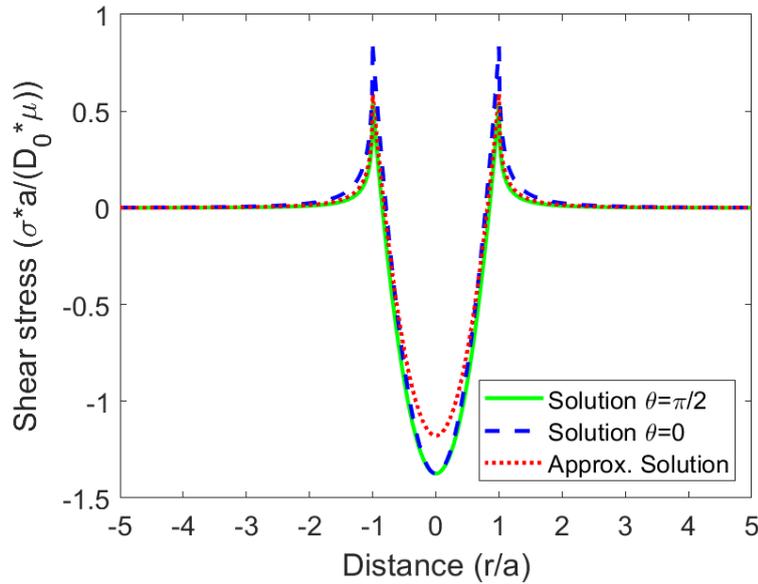


Figure 4.7) Complete stress solution, computed for $\theta = 0$ and $\theta = \pi/2$, compared to the approximate solution of Figure 4.6.

Within the analysis of the stress release model for seismic sequences in Southern Apennines (Section 4.3.1), we adopted the stress formulation accounting for the non-isotropic component of the stiffness.

4.2.3 Focal mechanisms estimation

Focal mechanisms are geometrical or mathematical representations of faulting during an earthquake. Fault planes can be characterized in terms of their geometrical properties, according to their orientation and direction of the slip vector on them. For determining the orientation of the fault plane in a reference frame, two angles are required: the strike angle (Φ_s) and the dip angle (δ). The slip direction is specified through another angle, referred to as the rake angle, (λ). The following Figure 4.8 schematically represents a fault plane with the relative reference angles.

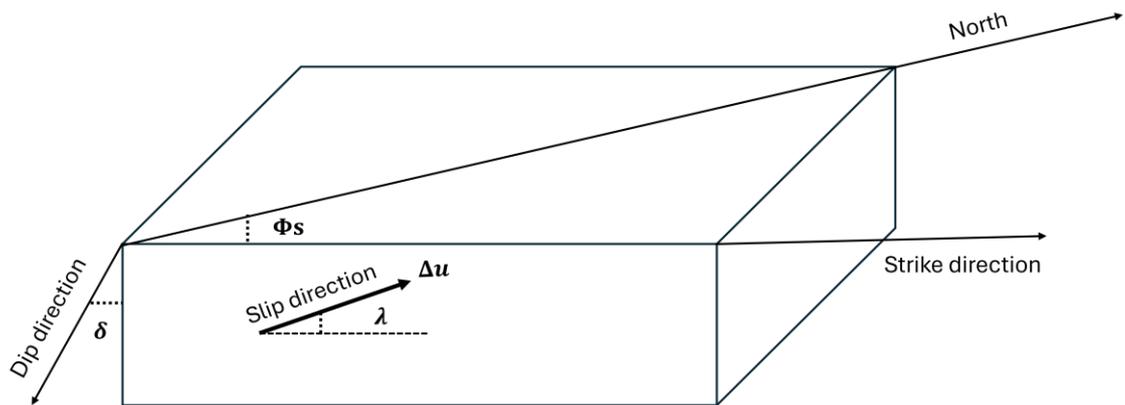


Figure 4.8) Definition of the fault orientation parameters (strike Φ_s and dip δ angles) and the rake angle λ indicating the orientation of the slip vector along the fault.

The strike angle Φ_s is the angle, measured clockwise, between the fault trace and the North. Its value ranges in $[0, 2\pi]$. The strike angle marks the strike direction and looking in that direction, the hanging wall of the fault appears on the right side. On the other hand, the fault dip δ is the angle formed between the fault plane and the Earth's surface, in the vertical plane orthogonal to the strike. The domain of the dip angle is bounded by 0 and $\pi/2$. The orientation of the slip vector $\Delta\mathbf{u}$, which indicates the displacement direction of the hanging wall relatively to the footwall, is described through the rake angle λ , formed between the strike direction and the slip vector. The rake angle λ lives in the domain $[-\pi, \pi]$ and its value characterizes the faulting type. If δ is different from 0 and $\pi/2$ and $\lambda = \pi/2$, the fault is known as a pure inverse fault (or thrust), while if $\lambda = -\pi/2$ the fault is referred to be as pure normal (or direct) fault. For intermediate values of the rake angle, oblique-normal and oblique-reverse faulting can be defined. A strike-slip fault is

characterized by a horizontal slip vector ($\lambda = 0$ or $\lambda = \pi$) and, particularly, a vertical strike slip fault is further characterized by $\delta = \pi/2$. For this fault type, two distinctions can be performed according to the value of the rake angle: a right-lateral strike-slip fault features $\lambda = \pi$ while a left-lateral strike-fault is characterized by $\lambda = 0$. The following Figure 4.9 summarizes the fault types that were previously described.

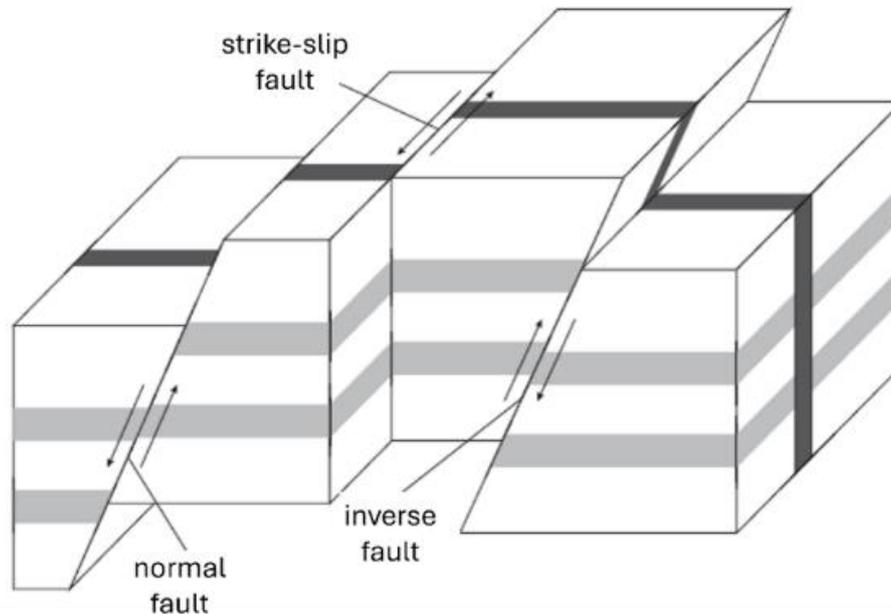


Figure 4.9) Schematic representations of different fault types: normal, strike-slip and inverse fault.

For determining the focal mechanism, the introduction of the focal sphere is a crucial element. The focal sphere is defined as a sphere centred in the earthquake hypocenter, characterized by a unitary radius. The radiation pattern, which is the azimuthal variation of the amplitude of the motion produced by a seismic wave at a fixed distance, is represented along the surface of the focal sphere. In the focal mechanism estimation, the wave amplitude is backprojected along the ray path from the receiver to the source, to identify the point where the seismic ray intersects the focal sphere. It thus appears that the knowledge of the location of the hypocenter, along with the velocity distribution along the ray path, plays a critical role in the accuracy of the focal mechanism estimation. A point on the focal sphere can be identified by a set of polar coordinates (i_ξ, Φ) in a reference system centred in the earthquake hypocenter, in which $i_\xi = 0$ corresponds to the vertical direction oriented downward and Φ represents the azimuth respect to the North direction (Figure 4.10). Since the focal sphere lies within the source near field, it

is not trivial to understand how the radiation pattern of the far field might correctly represent the displacement occurring near the source.

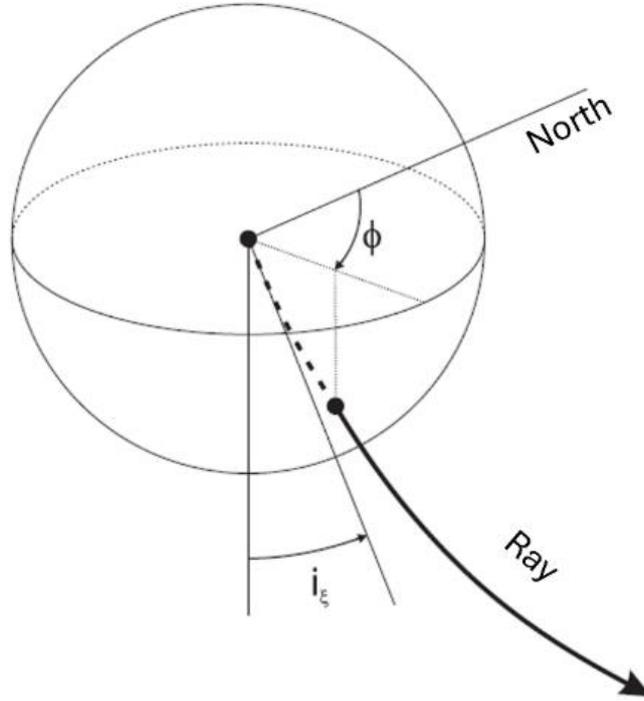


Figure 4.10) Definition of the focal sphere and local polar coordinates in a reference frame centred in the earthquake hypocenter.

For a point source, it is possible to show that near (NF), intermediate (IF) and far (FF) fields of the displacement in the position \mathbf{r} , solution of the elastodynamic equation, might be expressed as a function of the time-dependent seismic moment $M_0(t)$ as (Aki & Richards, 2002):

$$\begin{aligned}
 u(r, t) = & \frac{1}{4\pi\rho} R^{NF} \frac{1}{r^4} \int_{\frac{r}{v_p}}^{\frac{r}{v_s}} \tau M_0(t - \tau) d\tau & (4.36) \\
 & + \frac{1}{4\pi\rho v_p^2} R^{IF_P} \frac{1}{r^2} M_0\left(t - \frac{r}{v_p}\right) + \frac{1}{4\pi\rho v_s^2} R^{IF_s} \frac{1}{r^2} M_0\left(t - \frac{r}{v_s}\right) \\
 & + \frac{1}{4\pi\rho v_p^3} R^{FF_P} \frac{1}{r} M_0\left(t - \frac{r}{v_p}\right) + \frac{1}{4\pi\rho v_s^3} R^{FF_s} \frac{1}{r} M_0\left(t - \frac{r}{v_s}\right)
 \end{aligned}$$

where ρ , v_p and v_s represent the density, P and S wave velocity, while r refers to the source-receiver distance. We can consider cartesian and spherical reference frames for the analysis of the radiation pattern associated with a shear dislocation occurring along the fault surface, as in Figure 4.11.

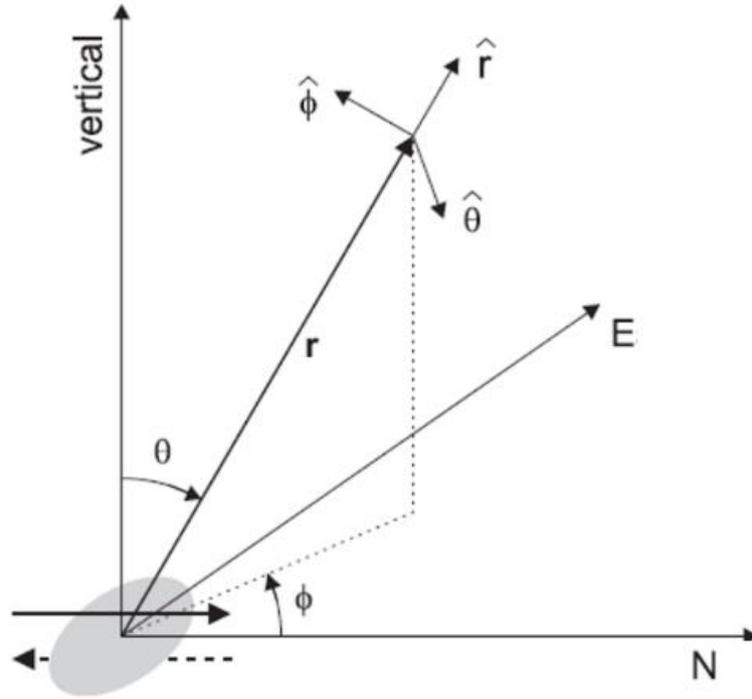


Figure 4.11) Cartesian and spherical reference frames for the analysis of the radiation pattern associated with a shear dislocation occurring along the fault surface.

In this reference frame, the radiation pattern coefficients R are given by:

$$\begin{aligned}
 R^{NF} &= 9 \sin 2\theta \cos \Phi \hat{r} - 6(\cos 2\theta \cos \Phi \hat{\theta} - \cos \theta \sin \Phi \hat{\Phi}) \\
 R^{IFP} &= 4 \sin 2\theta \cos \Phi \hat{r} - 2(\cos 2\theta \cos \Phi \hat{\theta} - \cos \theta \sin \Phi \hat{\Phi}) \\
 R^{IFs} &= -3 \sin 2\theta \cos \Phi \hat{r} + 3(\cos 2\theta \cos \Phi \hat{\theta} - \cos \theta \sin \Phi \hat{\Phi}) \quad (4.37) \\
 R^{FFP} &= \sin 2\theta \cos \Phi \hat{r} \\
 R^{FFs} &= \cos 2\theta \cos \Phi \hat{\theta} - \cos \theta \sin \Phi \hat{\Phi}
 \end{aligned}$$

From Equation 4.36, we can obtain the final value of displacement associated with a shear dislocation of seismic moment $M_0(t)$, evaluating the limit $t \rightarrow \infty$ and assuming that the seismic moment features a constant value M_0 for $t \rightarrow \infty$. We therefore obtain:

$$u(r, t) = \frac{M_0}{4\pi\rho r^2} \left[\frac{1}{2} \left(\frac{3}{v_s^2} - \frac{1}{v_p^2} \right) \sin 2\theta \cos \Phi \hat{r} + \frac{1}{v_p^2} (\cos 2\theta \cos \Phi \hat{\theta} - \cos \theta \sin \Phi \hat{\Phi}) \right] \quad (4.38)$$

If we consider the displacement radiation pattern for a P wave associated with a shear dislocation in the reference frame of Figure 4.11, we expect that a particle of the medium that belongs to one of the quadrants deployed around the fault undergoes a compressive (directed toward the receiver) or extensional (directed towards the source) first motion. From Equation 4.37, the displacement for the P wave is proportional to $\sin 2\theta \cos \Phi$. When $\Phi = 0$, u_r is proportional to $\sin 2\theta$, which corresponds to a diagram with four lobes (left panel of Figure 4.12), reflecting the alternation of signs as in the right panel of Figure 4.12.

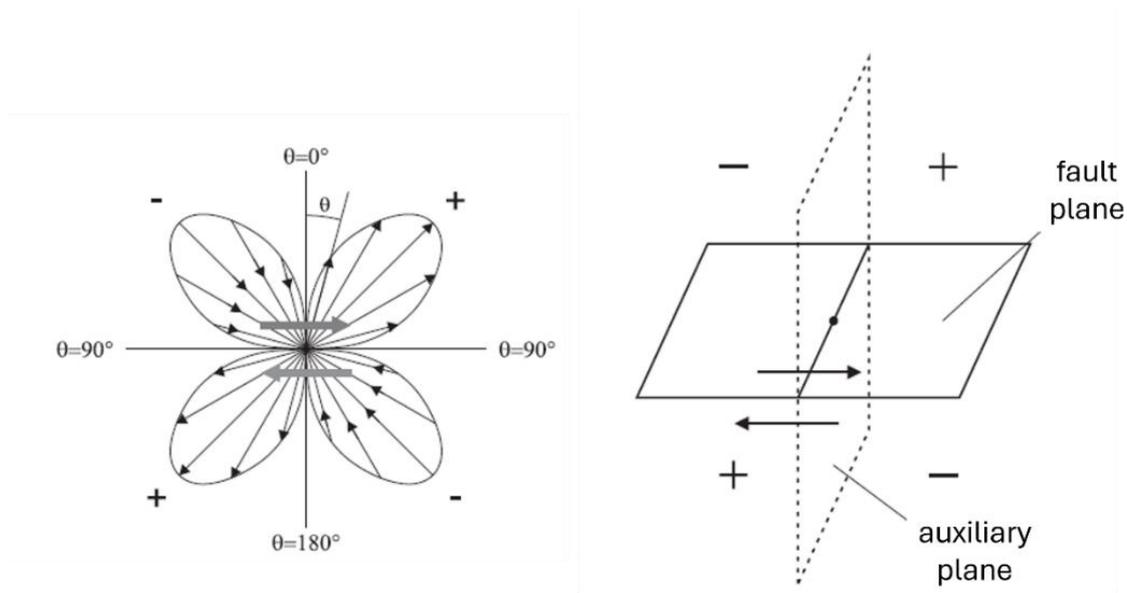


Figure 4.12) Left panel: P wave radiation pattern for a vertical strike-slip fault. Grey arrows indicate the slip direction with respect to the fault plane. Red arrows indicate the P wave amplitude for different θ angles. Right panel: First motion amplitude for the P wave respect to the fault and auxiliary planes.

Since the function expressed in Equation 4.38 varies smoothly, it is straightforward to imagine that the polarity flip occurs when the amplitude of the first motion becomes zero.

Therefore, outside the area affected by the dislocation, there is a continuous transition from motion directed towards the source to motion directed in the opposite direction. The maximum amplitude for the P wave displacement is expected halfway of each of the four quadrants, therefore at 45° from the fault plane.

The sign of the first motion for the P wave is preserved along the ray path towards every receiver. Thus, if enough readings of the P wave first motion signs are available, backprojecting the wave amplitudes from the receivers to the source it is possible to determine the orientation of the fault plane. However, the intrinsic symmetry in the four-lobes radiation patterns makes it impossible to distinguish univocally the fault from the auxiliary plane, on which the slip features the opposite direction. The analysis of the hypocentral locations of the aftershocks of a high-magnitude event can help in tackling this issue and in determining the orientation of the fault plane. Indeed, aftershocks typically occur on the fault plane that has generated the higher magnitude event, disentangling the ambiguity among the two nodal planes. Moreover, distinguishing the fault from the auxiliary plane is intrinsically easier for strike-slip faults, since the planes feature an orthogonal orientation often directly visible on the Earth's surface (Jones, 1998; Prejean et al., 2002; Karasözen et al., 2014). Identifying the fault plane on normal or inverse faults is a more complex task that typically requires accurate hypocentral locations of the aftershocks (Chiaraluce, 2012; De Matteis et al., 2012; Tarantino et al., 2024).

Information on the focal mechanisms is typically represented through the beachball diagram, which is built according to a lower hemisphere stereographic projection of the first motion P polarities observed at multiple seismic stations deployed around the epicenter. In the beachball representation, the colored quadrants indicate the volume of the medium around the source which undergoes extension, whereas the white quadrants refer to the segments which undergo compression. Figure 4.13 reports the beachball for the main fault types which were previously discussed. In these representations, the P and T axes, the principal axes of compression and extension respectively, are reported as points in the white and colored segments, respectively. Associating the orientation of the tectonic compression and extension axes to the P and T axes determined from the focal mechanisms of multiple earthquakes makes it possible to obtain the actual orientation of the regional stress field (Pasquale et al., 2009; De Matteis et al., 2012; Martinez-Garzon et al., 2014b).

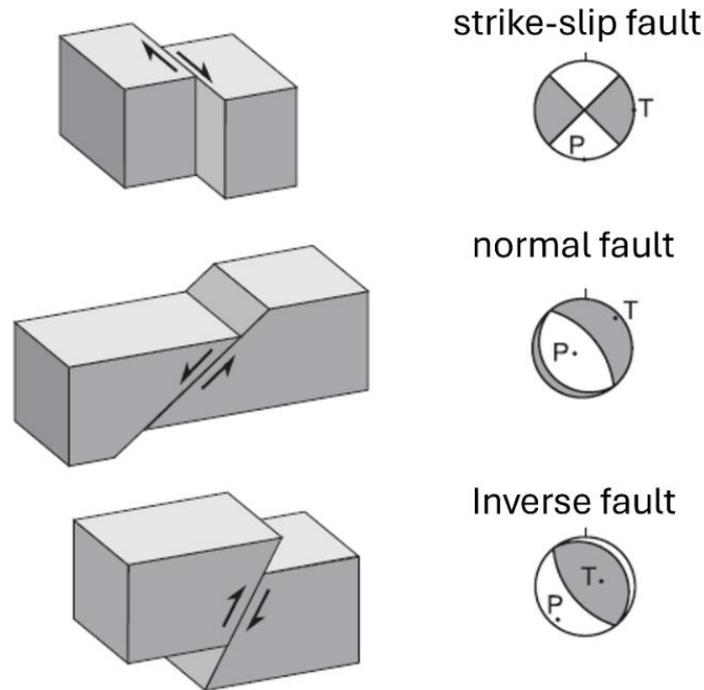


Figure 4.13) Beachball representation for the focal mechanisms associated with the main fault types. The coloured quadrants indicate the volume of the medium around the source which undergoes extension, whereas the white quadrants refer to the segments which undergo compression.

4.2.4 Stress field inversion

Earthquakes are processes associated with the sudden rupture of rocks along cracks, fractures or faults exposed to stress field in the Earth's crust. If the stress reaches a critical value exceeding the strength of faults or fractures in rocks, the accumulated energy of elastic deformation is partially spent for anelastic deformations in the focal zone and partially released and radiated in the form of seismic waves. Stress in the Earth's crust causing earthquakes can be of tectonic or non-tectonic origin (Ruff, 2002). The main source of non-tectonic stress within the Earth is gravitational loading. This stress is vertical with the largest lateral variations near the Earth's surface and being more homogeneous at depth. On the other hand, tectonic stress is mostly horizontal and originates in forces driving the plate motions (Heidbach et al., 2008). Principal stress directions in the Earth's crust are frequently close to vertical and horizontal directions. This led Anderson (1951) to develop a simple scheme connecting the basic stress regimes in the Earth's crust with the type of faulting on a pre-existing fault in the crust. Anderson (1951) distinguishes three possible combinations of magnitudes of principal stresses: the vertical stress is maximum, intermediate or minimum with respect to the horizontal

stresses. If the vertical stress is maximum, the hanging wall is moving downwards with respect to the footwall and the normal faulting is observed along a deeply steeping fault. If the vertical stress is minimum, the crust is in horizontal compression and the hanging wall is moving upwards with respect to the foot wall and reverse faulting is observed along a shallow dipping fault. Finally, if the vertical stress is intermediate, the foot and hanging walls are moving horizontally and strike-slip faulting is observed along a nearly vertical fault. Obviously, the Anderson's classification is simple and does not cover all observations but still it proved to be valid for many seismically active regions and helpful for a rough assessment of the stress regime (Simpson, 1997; Hardebeck & Michael, 2006). In Figure 4.14, we visually summarize the faulting regimes according to the module of the acting stress.

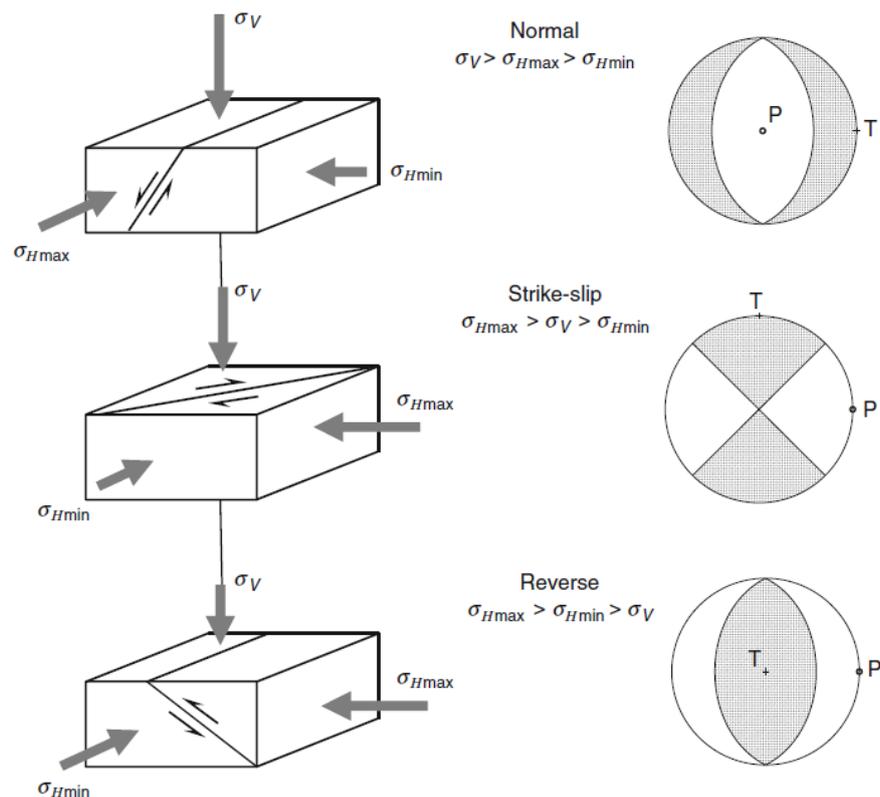


Figure 4.14 Anderson's classification scheme of stress in the Earth's crust (left) and corresponding faulting regimes (right). The focal mechanisms with the P and T axes are shown in the lower-hemisphere equal-area projection.

Stress describes forces acting on a unit surface in a body. Since the acting force and the normal of the unit surface are vectors, the stress is a tensor described by nine components

$$\boldsymbol{\tau} = \begin{bmatrix} \tau_{11} & \tau_{12} & \tau_{13} \\ \tau_{21} & \tau_{22} & \tau_{23} \\ \tau_{31} & \tau_{32} & \tau_{33} \end{bmatrix} \quad (4.39)$$

The force acting on surface S with normal \mathbf{n} is called the traction \mathbf{T} , and is expressed as:

$$T_i = \tau_{ij}n_j \quad (4.40)$$

with its normal and shear components σ_n and τ

$$\sigma_n = T_i n_i = \tau_{ij}n_i n_j \quad (4.41)$$

$$\tau N_i = T_i - \sigma_n n_i = \tau_{kj}n_j(\delta_{ik} - n_i n_k) \quad (4.42)$$

where \mathbf{N} is the direction of the shear component τ and lies on the surface S . Since stress is defined as that part of forces in a body which causes its deformation but not rotation, the stress tensor must be symmetric ($\tau_{ij} = \tau_{ji}$), being described by six independent components only.

The values of the stress tensor components depend on the system of coordinates, in which the components are measured. The coordinate system can always be rotated in the way that the stress tensor diagonalizes

$$\boldsymbol{\tau} = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{bmatrix} \quad (4.43)$$

where σ_1, σ_2 and σ_3 are called the maximum, intermediate and minimum principal stresses (compression is positive):

$$\sigma_1 \geq \sigma_2 \geq \sigma_3 \quad (4.44)$$

and the vectors defining this special coordinate system are called the principal stress directions or principal stress axes. Mathematically, the principal stresses and their directions are found by calculating the eigenvalues and eigenvectors of the stress tensor. The normal and shear components σ_n and τ of traction T (also called the normal and shear stresses) read in the system of principal stress directions

$$\sigma_n = \sigma_1 n_1^2 + \sigma_2 n_2^2 + \sigma_3 n_3^2 \quad (4.45)$$

$$\tau^2 = \sigma_1^2 n_1^2 + \sigma_2^2 n_2^2 + \sigma_3^2 n_3^2 - \sigma_n^2 \quad (4.46)$$

If principal stresses σ_1, σ_2 and σ_3 are fixed, then normal and shear stresses σ_n and τ are just functions of the normal \mathbf{n} to a fault and can be plotted in the Mohr's circle diagram (Figure 4.15). All permissible values of σ_n and τ must lie in the shaded area of the diagram (Mavko et al., 2020).

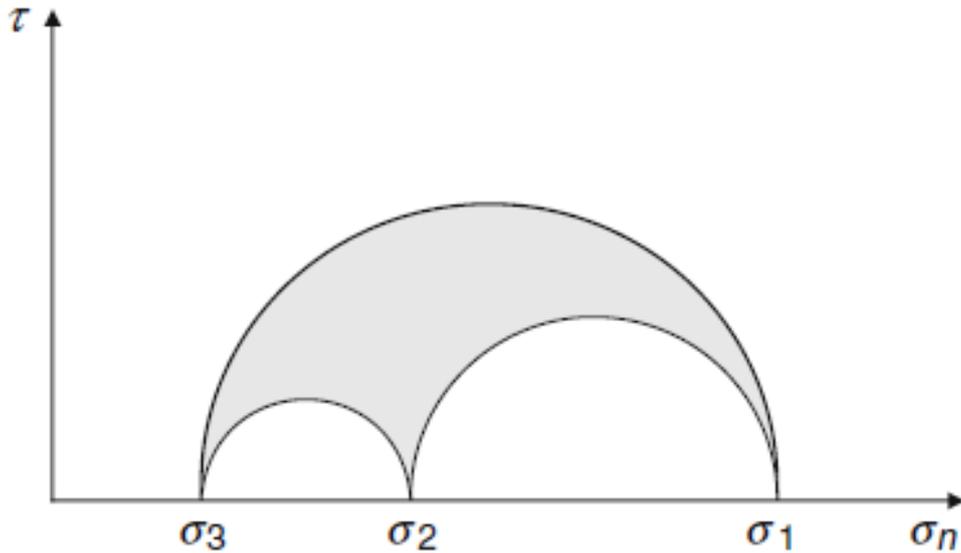


Figure 4.15) Mohr's circle diagram. Quantities σ_n and τ are the normal and shear stresses along a fault, σ_1, σ_2 and σ_3 are the principal stresses. All permissible values of σ_n and τ acting on a fault must lie in the shaded area of the diagram

If a rock is critically stressed in the Earth's crust, the rock is fractured, with the occurrence of the earthquake. In principle, an earthquake can occur on a newly developed fracture or on a pre-existing fault in the Earth's crust which is re-activated. The condition under which fracturing or faulting occurs is described by the so-called failure criteria. According to the Mohr-Coulomb failure criterion (Zoback, 2010; Scholz, 2019), shear stress on an activated fault must exceed the critical value τ_c , which is calculated from cohesion C , fault friction μ , the normal stress σ_n and the pore pressure p :

$$\tau_c = C + \mu(\sigma_n - p) \quad (4.47)$$

If the Mohr-Coulomb failure criterion is satisfied (red area in Figure 4.16), the fault becomes unstable and an earthquake occurs along this fault. The higher the shear stress difference, $\Delta\tau = \tau - \tau_c$, the higher the instability of the fault and the higher the susceptibility of the fault to be activated. A fault most susceptible to failure is called "principal" fault (Vavryčuk, 2011) being defined by the point in which the Mohr-Coulomb failure criterion touches the Mohr's circle diagram (blue point in Figure 4.16).

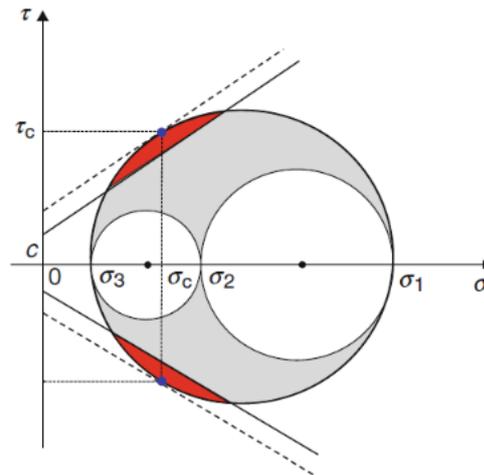


Figure 4.16) Mohr-Coulomb failure criterion. The red area shows all possible orientations of fault planes which satisfy the Mohr-Coulomb failure criterion. The blue dot with shear and normal stresses τ_c and σ_c denotes the principal fault plane which is optimally oriented with respect to stress, and C denotes the cohesion.

A variety in possible orientations of unstable fault planes is demonstrated in Figure 4.17. The left-hand plot of Figure 4.17 shows the Mohr's diagram, the failure criterion and the positions of randomly distributed unstable fault planes satisfying the failure criterion. The middle and right-hand plots of Figure 4.17 show the nodal lines and the P (pressure) and T (tension) axes for the corresponding focal mechanisms, respectively. The nodal lines and P/T axes inform us about the predominant type of faulting and about the scatter in the orientations of the unstable fault planes. Predominant faulting and its scatter are also projected into the scattering of the P/T axes, which form clusters of a specific shape and size (Figure 4.17, right-hand plot).

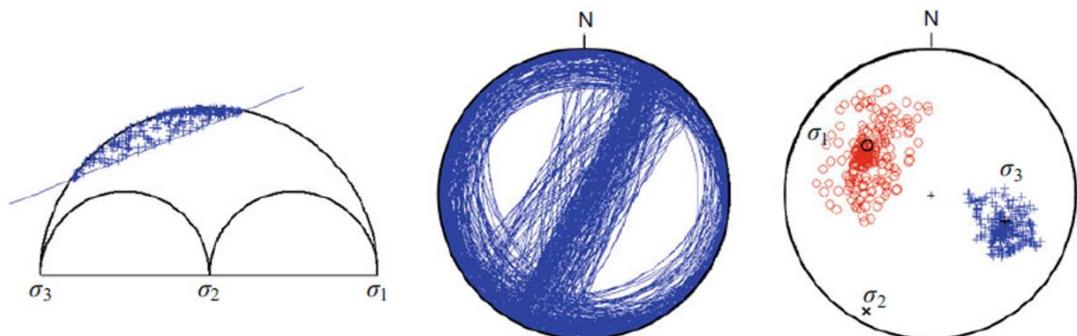


Figure 4.17) Focal mechanisms associated with unstable fault planes. Randomly distributed fault planes inside the unstable area of the Mohr's diagram (left), corresponding nodal lines (in the middle) and the P/T axes (right). The P (pressure) axes are marked by the red circles, the T (tension) axes by the blue crosses.

Several methods have been proposed for the determination of stress from a set of focal mechanisms of earthquakes (Maury et al., 2013). These methods usually assume that (1) the tectonic stress is uniform (homogeneous) in the region, (2) earthquakes occur on pre-existing faults with varying orientations, (3) the slip vector points in the direction of shear stress on the fault (Wallace, 1951; Bott, 1959). Obviously, these conditions might not be always satisfied. In case that stress is not uniform, the area should be subdivided into smaller regions in which the assumption of uniform tectonic stress is reasonable. In this case, the stress inversion methods are capable of determining four parameters of the stress tensor: three angles defining the directions of the principal stress directions, σ_1, σ_2 and σ_3 , and shape ratio R . The stress tensor is usually searched with the normalized maximum compressive stress ($\sigma_1 = 1$) and with zero trace ($Tr(\boldsymbol{\tau}) = \sigma_1 + \sigma_2 + \sigma_3 = 0$). The simplest approach to stress inversion is the method of Michael (1984). This method employs (4.40) expressed in the following form:

$$\tau_{kj} n_j (\delta_{ik} - n_i n_k) = \tau N_i \quad (4.48)$$

For evaluating the right-hand side of the equation, Michael (1984) applies the Wallace-Bott assumption that direction \mathbf{N} of the shear stress component of the traction \mathbf{T} on a fault is identical to the slip direction \mathbf{s} , and he further assumes that shear stress τ on activated faults has the same value for all studied earthquakes. Since the method cannot determine absolute stress values, τ is assumed to be 1. Therefore, 4.40 is expressed in the matrix form:

$$\mathbf{A} \mathbf{t} = \mathbf{s} \quad (4.49)$$

Where \mathbf{t} is the vector of stress components $\mathbf{t} = [\tau_{11}, \tau_{12}, \tau_{13}, \tau_{22}, \tau_{23}]$ and \mathbf{A} is the 3x5 matrix calculated from the fault normal \mathbf{n}

$$\mathbf{A} = \begin{bmatrix} n_1(n_2^2 + 2n_3^2) & n_2(1 - 2n_1^2) & n_3(1 - 2n_1^2) & n_1(-n_2^2 + n_3^2) & -2n_1n_2n_3 \\ n_2(-n_1^2 + n_3^2) & n_1(1 - 2n_2^2) & -2n_1n_2n_3 & n_2(n_1^2 + 2n_3^2) & n_3(1 - 2n_2^2) \\ n_3(-2n_1^2 - n_2^2) & -2n_1n_2n_3 & n_1(1 - 2n_3^2) & n_3(-n_1^2 - 2n_2^2) & n_2(1 - 2n_3^2) \end{bmatrix} \quad (4.50)$$

and \mathbf{s} is the unit direction of the slip vector. Extending the former equation for focal mechanisms of K earthquakes with known fault normal \mathbf{n} and slip directions \mathbf{s} , a system of $3 \times K$ linear equations for five unknown components of the stress tensor is obtained. The system is solved using the generalized linear inversion in the L2-norm (Lay & Wallace, 1995)

$$\mathbf{t} = \mathbf{A}^{-g} \mathbf{s} \quad (4.51)$$

The basic drawback of this method is the necessity to know the orientations of the faults. Usually, when determining the focal mechanisms, the orientations of the two nodal planes are calculated: one nodal plane corresponding to the fault and the other nodal plane (called

the auxiliary plane) defining the slip direction. The inherent ambiguity of the focal mechanisms does not allow distinguishing easily which of the nodal planes is the fault. If the Michael's method is used with incorrect orientations of the fault planes, the accuracy of the retrieved stress tensor is decreased. On the other hand, the method is quite fast and it can be run repeatedly. Therefore, the confidence regions of the solution are determined using the standard bootstrap method (Michael, 1987). If the orientation of fault planes in the focal mechanisms is unknown, each nodal plane has a 50 % probability of being chosen during the bootstrap resampling.

For implementing the stress field inversion, we followed the approach proposed by Martinez-Garzon et al. (2014a), known as MSATSI. Here, the input focal mechanisms can be grouped prior to the inversion into a number of subareas ("grid points") distributed over a number of dimensions, from 0D and 1D (e.g., temporal changes of the stress field), up to 4D (e.g., spatiotemporal distribution). Then, a stress tensor for each grid point is inverted simultaneously using a damped least-squares inversion scheme in order to obtain a smoothed solution. The inversion provides the orientations of the three principal stress axes and a quantity (Φ) that reflects the relative stress magnitude R:

$$\Phi = 1 - R = \frac{\sigma_2 - \sigma_3}{\sigma_1 - \sigma_3} \quad (4.52)$$

in which σ_1 , σ_2 and σ_3 are the magnitudes of the three principal stress axes obtained from the deviatoric stress tensor. The relative stress magnitude quantifies whether the magnitude of the intermediate principal stress σ_2 is closer to the magnitude of the most compressive (σ_1) or the least compressive principal stress (σ_3). The bootstrap resampling method applied to the input focal mechanisms results in the uncertainties of the stress-axes orientations and the relative stress magnitude. In the inversion of the stress field from the focal mechanisms evaluated within the framework of the DETECT experiment, we selected 2000 bootstrap resamplings of the dataset.

4.3 Source characterization for microseismic events in the Southern Apennines

4.3.1 Source parameters estimation of events in enhanced catalogs for seismic sequences in Southern Apennines.

In Section 2.3.1, we produced enhanced catalogs for seismic sequences in the Southern Apennines, integrating advanced detection techniques, which led to an increase in the number of identified earthquakes by a factor of ~ 7 . In Section 3.3.1, we relocated $\sim 30\%$

of the events using absolute and relative location techniques, an increase by a factor of ~ 2 compared to the number of absolute locations in the manual catalogs, which allowed us to highlight the seismicity patterns along small-scale structures where the sequences occurred. In this section, we attempted to estimate the source parameters in the frequency domain for the relocated events in the enhanced catalogs using S waves.

After removing the instrumental response from the seismic records, we can write the amplitude displacement spectra as:

$$U(M_0, f_c, \gamma, Q_a, k, f) = S(M_0, f_c, \gamma, f) G(Q_a, k, f) \quad (4.53)$$

in which the source spectrum is expressed in Equation 4.6 and the complete propagation operator is described by Equation 4.13. The modelling of the spectra requires a joint inversion for source parameters and propagation (here described by the quality factors due to the anelasticity, Q_a , and site k) which are strongly correlated. To reduce this correlation, we tried to evaluate the exponential decay in the Green's function separately from the inversion of source parameters. For each sequence, the events are nearly collocated, preventing the possibility of separating the contribution of the regional anelastic attenuation from the site effects (Strumia et al., 2024). For this reason, we define a comprehensive quality factor Q , such that:

$$\frac{1}{Q} = \frac{1}{Q_a} + \frac{k}{T} \quad (4.53)$$

and:

$$G(Q, f) = K \frac{1}{r} e^{-\frac{\pi f T}{Q}} \quad (4.54)$$

Thus, the Green's function is characterized by the single parameter Q , to be determined, accounting for both anelastic attenuation and site effect.

We considered small events in each sequence as empirical Green's functions (EGF). For those events, the effective (source) corner frequency is much larger than the apparent corner frequency of the Green's function low-pass filter, and sometimes even larger than the Nyquist frequency of the records (in this case $f_{Nyq} = 62.5 \text{ Hz}$). Considering the typical hypocentral distance for the events occurring within the seismic sequences and a regional quality factor for the S waves in the region $Q = 230$ (Zollo et al., 2014), the low-pass filter of the Green's function has characteristic frequencies of [15-20] Hz. Considering the EGF spectra in the domain where $f \ll f_c$, the displacement spectrum can be approximated as:

$$U_{EGF}(Q_{EGF}, f) = K \frac{1}{r} M_0 e^{-\pi f T / Q_{EGF}} \quad (4.55)$$

We selected events featuring local magnitude $M_l < 1$ as EGFs, fitting with a linear model $\log U_{EGF}$ as a function of the frequency to retrieve M_0 and Q_{EGF} . We pre-processed the raw traces by removing the instrumental response, including a 5% Hann taper and a water level regularization during the deconvolution stage. We bracketed the S wave window from 1 s before to 3 s after the phase arrival time. We considered the frequency band that satisfied the condition of SNR larger than 3.5, and evaluated frequency by frequency, between the event and the noise spectra. This latter was computed from a 4s time window extracted before the event origin time. The value of Q_{EGF} is station dependent. Since events in the same sequence share almost the same source-receiver path, we expect a consistency in the Q_{EGF} estimates across the EGFs for the same station. For stations presenting at least 5 estimates of Q_{EGF} , we evaluated the compatibility of the inferred values and used the mean value to correct for the Green's function exponential attenuation. We typically evaluated Q_{EGF} at 4 ± 2 stations for each sequence, deployed within a hypocentral distance of 20 km from the sequence's centroid. This distance represents the threshold below which the events used as EGFs feature a sufficient signal-to-noise ratio.

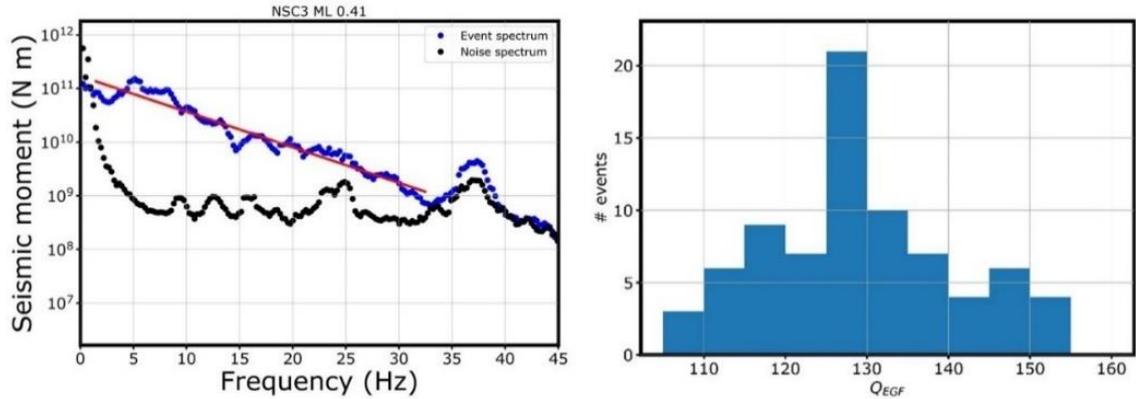


Figure 4.18) Left panel: Q_{EGF} estimation from linear fit (red line) of the logarithm of displacement event spectrum (blue dots) as a function of the linear frequency, for a $M_l = 0.41$ earthquake in the Rocca San Felice seismic sequence. Noise spectrum is reported as black dots. Right panel: histogram of the Q_{EGF} for the events $M_l < 1$ in the Rocca San Felice sequence (IDX 1) at NSC3 station.

We hereby report in Figure 4.18 the results of the evaluation of Q_{EGF} for the station NSC3 for the Rocca San Felice seismic sequence, located at a hypocentral distance of ~ 11 km from the mainshock. We fit the logarithm of the amplitude displacement spectra in the frequency band where $f \ll f_c$ and the signal-to-noise ratio of displacement spectra overcomes the threshold of 3.5. An example of Q_{EGF} distribution (for the station NSC3)

is reported in the right panel of Figure 4.18. We observed a peaked Gaussian-like distribution, typical of stations providing a large number of estimates for the quality factor. We extracted the weighted mean of individual Q_{EGF} values using the inverse of the fit residuals as weighting factors, to describe the quality factor $\langle Q_{EGF} \rangle$ for that station-sequence couple. For the example in Figure 4.4, we estimated $\langle Q_{EGF} \rangle = 130 \pm 12$ for NSC3, which is smaller than, but still within the variability range of the regional estimation provided by Zollo et al. (2014). We typically observe lower values of Q_{EGF} as compared to the regional value suggested by Zollo et al. (2014) suggesting a higher impact of the site effects with respect to the anelastic attenuation for short source-station propagation distances.

For other stations, for which we have an insufficient number of high-quality EGFs, we attempted to estimate a sequence-dependent quality factor Q_{LOC} by exploring different values of the comprehensive quality factor around the average regional estimate $Q_{REG} = 230$ (Zollo et al., 2014). Considering events with $M_l > 1$, we inverted the displacement amplitude spectra, fixing the attenuation to one of the following values $Q = 100, 170, 230, 300, 400$ in different inversion runs. We compared the average residuals resulting from the best solution for the source parameters in each run. We selected as Q_{LOC} the Q value producing the lowest misfit, imposing a minimum number of 5 solutions per station. For each sequence, we typically estimated Q_{LOC} at 4 ± 2 stations, featuring hypocentral distance between 20 and 42 km from the main event in the sequence. We here reported the parametric test performed for the station SFL3 within the Rocca San Felice sequence (IDX 1), at a hypocentral distance of ~ 40 km from the main event. We did not retrieve any solution exceeding the quality threshold for $Q = 100$ and $Q = 170$ (Supino et al., 2019), and we observed the minimum RMSE for $Q = 300$. The choice of the latter quality factor produces, on average, RMSE values from the spectral fitting that are 20 % smaller than those obtained using the regional value $Q = 230$ (Zollo et al., 2014). We report in Figure 4.19 the spectral inversion using $Q = 230$ (left panel) and $Q = 300$ (right panel) for a $M_l 2.2$ event of the Rocca San Felice seismic sequence at the SLF3 station, with the average RMSE for the other Q values for the inversion of all the available events in the bar plot of the inset in the right panel.

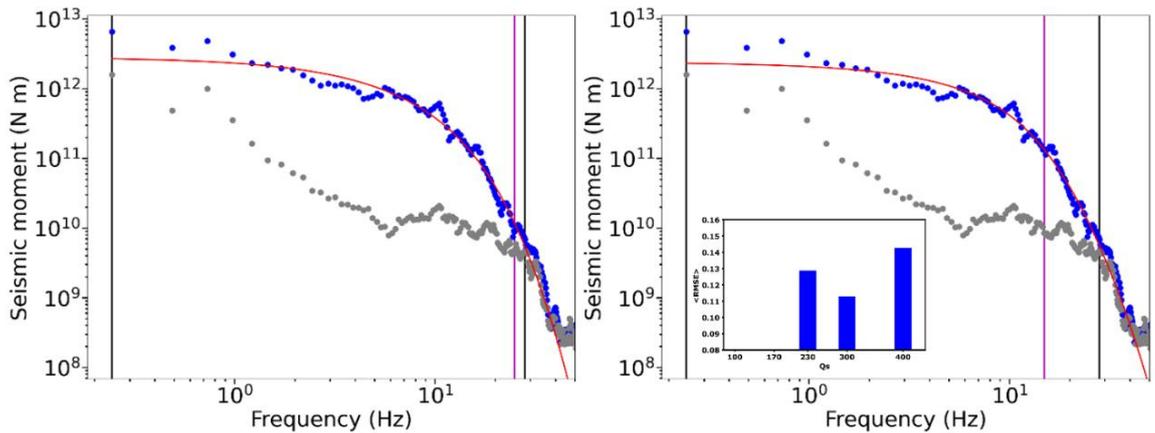


Figure 4.19 Spectral inversion of a $M_l = 2.2$ earthquake at SFL3 station ($r \sim 40$ km from the main event, sequence IDX 1) using different attenuation factors for the station. Left panel: inversion using $Q = 230$. Right panel: inversion using $Q = 300$. In both panels, black vertical lines mark the frequency band in which the spectral fit is performed. The pink vertical lines indicate the estimated corner frequency. Right panel inset: average RMSE for the events at SFL3 for the sequence IDX 1, using different Q_s values in the inversion

We note that in this latter case, the average value ($Q = 230$) provides unreliably large corner frequencies (as compared to the values obtained at other stations), close to the upper limit of the frequency band used for the inversion. We thus select $Q_{LOC} = 300$ for this station, this value proving the least average RMSE in the inversion. However, the possibility of extracting Q_{LOC} depends on the number of events of $M_l > 1$ that occurred within the considered sequence. As an example, for the most populated sequence, we extracted Q_{LOC} for 5 stations with $25 \text{ km} \leq r \leq 45 \text{ km}$

We finally kept $Q = Q_{REG}$ for stations where neither Q_{EGF} nor Q_{LOC} could be evaluated. This is typically the case of the furthest stations, where the signal-to-noise ratio of the seismic records led to an insufficient number of events to be used for estimating the quality factor. We used Q_{REG} on average for 2 ± 1 stations, located at a hypocentral distance of $35 \text{ km} \leq r \leq 58 \text{ km}$.

For a more complete view, we hereby report in Figure 4.20 the results of three spectral inversions for a $M_l 2.8$ earthquake of the Rocca San Felice (IDX 1) at stations characterized by Q_{EGF} ($r \sim 11 \text{ km}$), Q_{LOC} ($r \sim 43 \text{ km}$), and Q_{REG} ($r \sim 48 \text{ km}$), in the left, central and right panel, respectively.

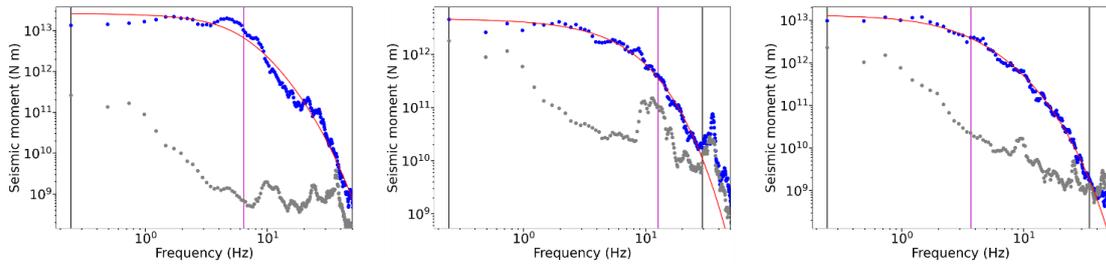


Figure 4.20) Spectral inversion of a $M_l = 2.8$ earthquake using different attenuation factors for the stations. Left panel: inversion using Q_{EGF} at NSC3 station. Central panel: inversion using Q_{LOC} at SCL3 station. Right panel: inversion using Q_{REG} at VDS3

We now discuss the resolution of the source parameters in the inversion of the displacement spectra. In the previous section, we introduced the lowpass filtering effect due to the propagation of the seismic waves from the source to the receiver. We showed that at the characteristic hypocentral distances for the seismic sequences in the Southern Apennines, and given the regional quality factor for the S waves suggested by Zollo et al. (2014), characteristic cutoff frequencies fall within 10 and 20 Hz. Thus, the spectral content at higher frequencies than the low-pass cutoff frequencies of the Earth's Green's function can be severely affected and biased using an incorrect value of quality factor. It is worth noting that f_c for microseismic events (e.g. $M < 2$) typically feature values higher than the [10 – 20] Hz range, thus the estimation of the corner frequency might be biased for events in this magnitude range. We performed synthetic tests, to demonstrate that the plateau level related to the seismic moment is well resolved for these events despite the combined effect of the decay beyond the corner frequency due to the source and the decay due to the anelastic attenuation. We computed synthetic spectra, considering an event with moment magnitude M_w 1.5 and corner frequencies of 15 Hz and 25 Hz. For generating synthetic noise and earthquake spectra, we used the average value for the anelastic attenuation ($Q=230$) at a hypocentral distance of 12 km (an average typical value for earthquakes in this magnitude range for the analyzed sequences), adding 10% of noise to the spectra, following Supino et al. (2019). We then inverted the spectra considering Q values either larger or smaller than the one used in the forward modelling ($Q_{small} = 150$, $Q_{large} = 300$), to account for uncertainty and variability in the anelastic operator, and a frequency band for the inversion 0.5-30 Hz, which represents the largest frequency band available for the inversion of real data in this analysis. We report in Figure 4.21 the results of spectral inversion for the synthetic tests.

Panel	True f_c (Hz)	Q value for the inversion	Retrieved Mw	Retrieved f_c (Hz)
a	15	150	1.52 ± 0.02	26 ± 2
b	15	300	1.49 ± 0.02	13.2 ± 1.0
c	25	150	1.56 ± 0.01	38.6 ± 1.2
d	25	300	1.50 ± 0.02	19.5 ± 1.9

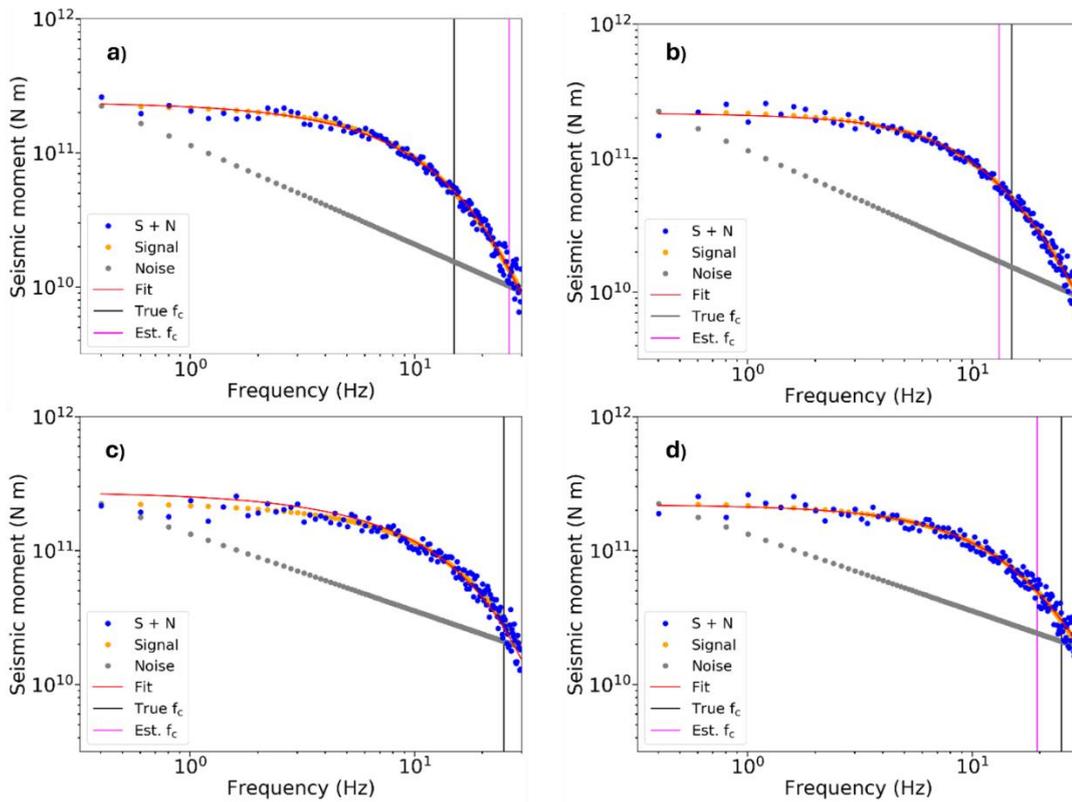


Figure 4.21) Synthetic tests for an event of magnitude Mw 1.5, $Q=230$ and $f_c =15$ Hz - Panels a) and b) - and $f_c = 25$ Hz - Panels c and d. We inverted the spectra assuming $Q=150$ - Panels a) and c) - and $Q=300$ - Panels b) and d). We found that the moment magnitude is well resolved, while the corner frequency estimate is strongly affected by the attenuation and limited bandwidth. In the table we report the retrieved values.

We found that the corner frequency is not well retrieved from the inversion, while the estimation of seismic moment is minimally affected by the change in the quality factor. The final moment magnitude reproduces the original one, with differences smaller than 0.1. These possible biases introduced by the attenuation cannot be reduced in the average estimates, since these events clearly emerge from the noise only at a limited number of stations (<4). Moreover, with decreasing magnitude of the events, the available bandwidth for spectral inversion is also reduced. On the other hand, we demonstrated with synthetics that we can resolve both the moment magnitude and the corner frequency for events with corner frequencies around 10 Hz or smaller, even if we introduce biases in the anelastic attenuation. In Figure 4.22, we report the results obtained for an event of M_w 2.0 and $f_c = 8$ Hz, adopting the same parametrization and inversion strategy of the former synthetic test.

Panel	True f_c (Hz)	Q value for the inversion	Retrieved Mw	Retrieved f_c (Hz)
a	8	150	2.05 ± 0.03	8.7 ± 1.1
b	8	300	1.99 ± 0.02	7.8 ± 0.6

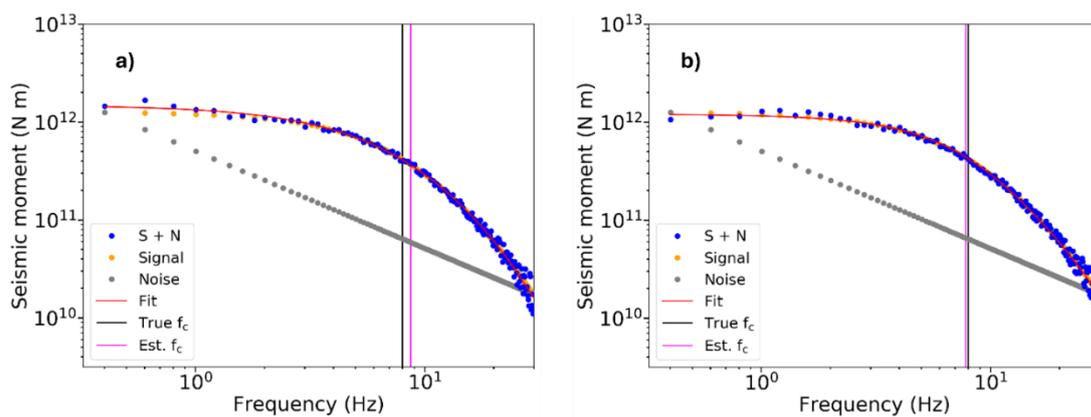


Figure 4.22) Synthetic tests for an event of magnitude M_w 2.0, $Q=230$ and $f_c=8$ Hz. We inverted the spectra assuming $Q=150$ – Panel a) $Q=300$ – Panel b). We found that both moment magnitude and corner frequency are well resolved. In the table we report the retrieved values.

Supported by these results, we estimated the seismic moment M_0 (and the moment magnitude M_w) for all the events, while corner frequencies were only determined for events with $M_l > 2$. The quality of the solutions was checked by analyzing the shape of the a-posteriori probability density function related to the estimated parameters. Solutions not showing peaked probability functions were discarded, following the strategy defined by Supino et al. (2019).

We globally estimated the seismic moment for 236 out of the 550 relocated events ($\sim 60\%$), which correspond to 15% of the starting enhanced catalog of earthquake detections. For the Rocca San Felice seismic sequence, we retrieved the seismic moment M_0 for 45 % of the relocated events. In Figure 4.23 we report the distribution of the moment magnitude M_w (Hanks & Kanamori, 1979), against the local magnitude M_l , as evaluated in Scotto di Uccio et al. (2023), marking with a red line the 1:1 trend between M_l and M_w .

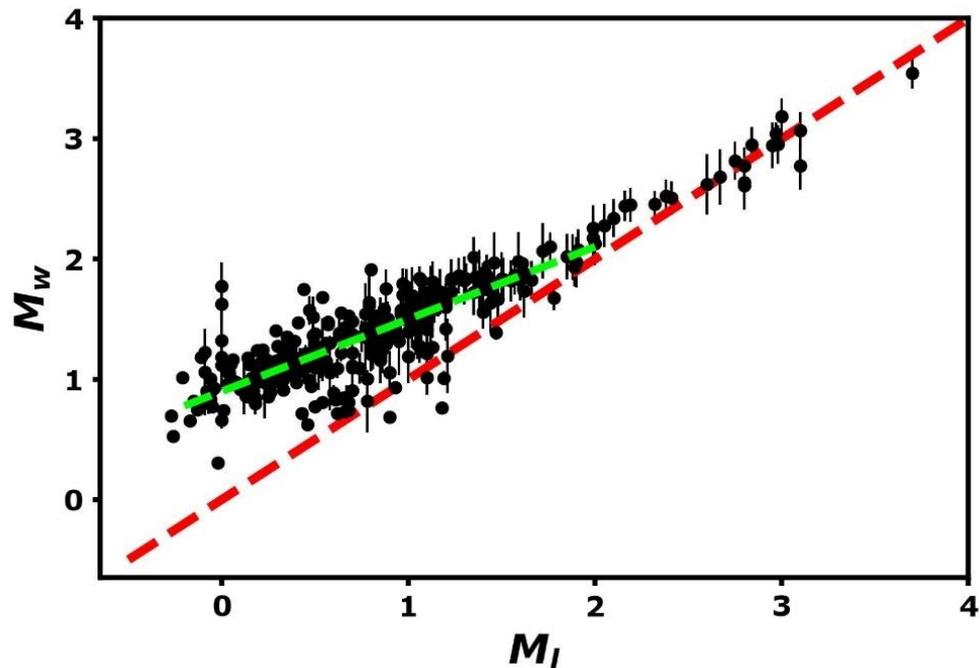


Figure 4.23) $M_w - M_l$ distribution, with the 1:1 scaling (dashed red line). For $M_l < 2$ earthquakes, we observed $M_w = 0.89 (\pm 0.03) + 0.62 (\pm 0.02)M_l$ (green dashed line).

We recognize two trends between the magnitude scales: for $M_l < 2$ the distribution strongly deviates from the 1:1 scaling relation. Evaluating the average value of M_w in different M_l bins of width 0.2 and performing a linear fit between the two quantities, we retrieved $M_w = a + b * M_l = 0.89 (\pm 0.03) + 0.62 (\pm 0.02)M_l$ (fit reported in Figure

4.23 as a green dashed line). The estimated slope agrees with the predictions of Deichmann (2017), which indicated a saturation of the event duration in the local magnitude computation due to the anelastic attenuation, resulting in the scaling $M_w = C + \frac{2}{3}M_l$. For $M_l > 2$, the distribution follows the 1:1 scaling trend between M_l and M_w , as expected from the definition of the moment magnitude (Hanks & Kanamori, 1979), also found by Zollo et al. (2014).

In Figure 4.24, we reported the $\log M_0 - \log f_c$ distribution for the resolved events of individual sequences, with red straight lines marking the theoretical trends obtained assuming characteristic stress drop values of $\Delta\sigma = 100kPa, 1MPa$ and $10 MPa$. The single station corner frequencies have been averaged considering the relative uncertainty of the estimates.

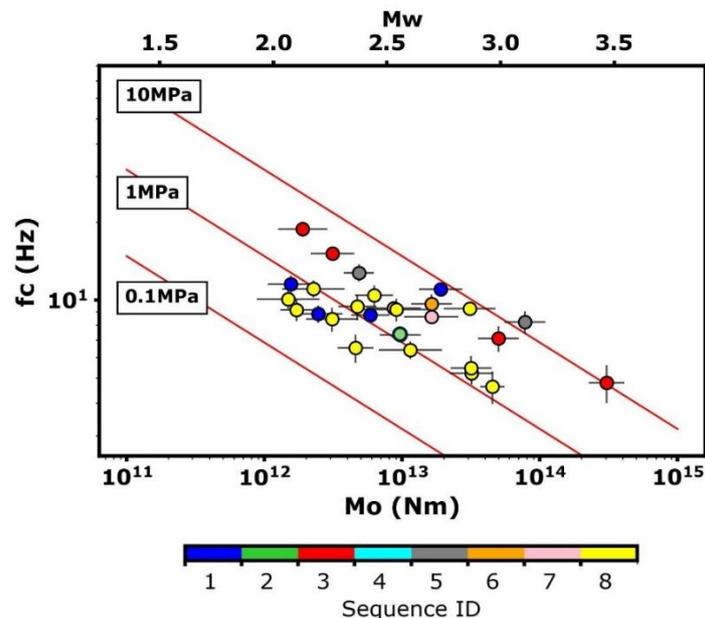


Figure 4.24) Distribution of source parameters for $M_w > 2$, colored according to sequence ID, with theoretical constant stress drops scaling of 0.1 MPa, 1 MPa and 10 MPa (red solid lines)

For the resolved events, the distribution of the corner frequencies with moment appears to follow a nearly linear trend, with stress drop ranging between 1-3 MPa. For the Rocca San Felice sequence, yellow marks in Figure 4.24, the average $\Delta\sigma$ was $\sim 1.0 MPa$. Averaging over all sequences, we found a median stress drop of $\Delta\sigma = 2.2 MPa$; its variability across sequences, estimated by the logarithm of standard deviation is $s_{\Delta\sigma} = 0.3 MPa$. The stress drop found here is one order of magnitude larger than the stress drop retrieved for the background seismicity in the area by Zollo et al. (2014), who used a

similar inversion strategy and an independent catalog not influenced by the sequences studied here. Since they derived the average stress drop using the Madariaga model (Madariaga, 1976), when converting that value into an equivalent Brune's stress drop, they obtained a median value of $\Delta\sigma=0.26$ MPa. This difference indicates that the release of stress during sequences likely occurs in more compact asperities that can be associated with higher coupling than for background seismicity (Chen et al., 2012). The stress drops we retrieved for these seismic sequences are comparable to the estimate of 3.5 MPa of the 1980, M 6.9 Irpinia earthquake (Deschamps & King, 1983, Bernard & Zollo, 1989). Moreover, we found differences in the stress drops associated with sequences, with increasing values moving from North to South in the Irpinia region, as also shown by Picozzi et al. (2022a), whose catalog contains all the events with magnitude larger than 1.5 for the area. We found median stress drops of $\Delta\sigma_N=1.8$ MPa in the Northern sector (Cervialto Fault area, the initial rupturing segment of the 1980 earthquake) and $\Delta\sigma_S=3.4$ MPa in the Southern Sector (San Gregorio Magno Fault area, i.e., on the second rupturing segment of the 1980 earthquake). Tomographic images in velocity (Amoroso et al., 2014; Improta et al., 2014; Vassallo et al., 2016) and anelastic attenuation (Q_p , Q_s ; Amoroso et al., 2017) coupled with rock physics modelling indicate the presence of pressurized fluids in the area of microseismicity. Differences in the stress drops between the two areas could be associated with the different fluid content and fraction. In the Southern sector rock physics modelling indicates the presence of a mixture brine- CO_2 (Amoroso et al., 2017). The large, extended low V_p/V_s anomaly in tomographic images in the northern sector indicates a pressurized reservoir of fluids, associated with the large natural emission of low-temperature CO_2 at the Mefite d'Ansanto, Rocca San Felice site (Chiodini et al., 2010).

4.3.2 Stress release modelling for seismic sequences in Southern Apennines

We evaluated the rupture plane associated with the seismic sequence as the best-fit plane across the hypocenters of the events in the sequence. As an example, for the Rocca San Felice sequence, we obtained the strike and the dip of the plane coherent with the angles estimated from the inversion of first motion polarities (Festa et al., 2021; Palo et al., 2023a). If the locations did not constrain a plane, we used the focal mechanism solutions from Palo et al., (2023a) and selected the plane that was more consistent with the expected orientation of faults in the area. We finally mapped the stress change on the fault plane associated with the sequence, using the rupture model proposed by Andrews (1980), considering a non-isotropic, complete representation of the stiffness. Since in the rupture model neither the slip nor the stress drop is considered constant, we imposed the average stress drop within the crack from the Andrews model coinciding with the event stress drop computed from the source parameters. We evaluated the source size for all the events for which we estimated the seismic moment using Equation 4.16, by considering either

the retrieved corner frequency or assuming self-similarity. The former condition applies to the largest magnitude events in the sequences, the latter for lower magnitude events.

For almost all the sequences, the stress model suggests static stress release as a trigger mechanism, with small events mainly concentrated in or around the area affected by stress changes due to the main events in the sequence. As an example, we report in the left panel of Figure 4.25 the stress release model for a seismic sequence featuring a M_l 2.9 main event (IDX), in which we observe a single km-sized patch mainly oriented along the dip direction, with earthquakes occurring within the volume affected by the main event. For assessing the connection between the earthquakes, we evaluated the nearest neighbor distance for each couple of events, after projecting the earthquake relocations along the estimated fault plane. The distribution of the nearest neighbor distance is reported in the right panel of Figure 4.25. The median nearest neighbor distance was ~ 30 m, with this value being associated with a source radius of a $M_{w_{eq}} = 0.6 / M_{l_{eq}} = -0.3$ earthquake, assuming the median stress drop extracted from the resolved events of the considered sequence and the estimated $M_l - M_w$ scaling law from Section 4.2.1.1. We observe that almost 70 % of the event-couples fell within the sum of the estimated source radius along the fault plane, suggesting a static stress release mechanism.

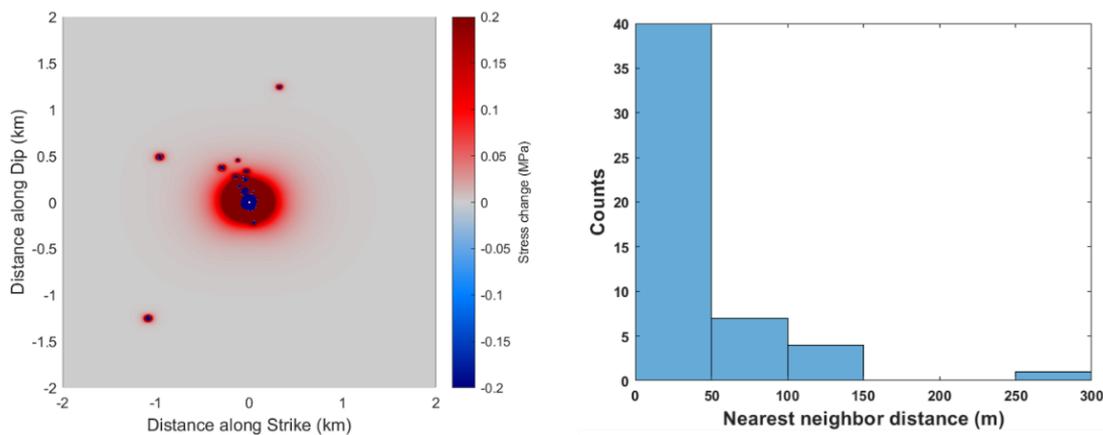
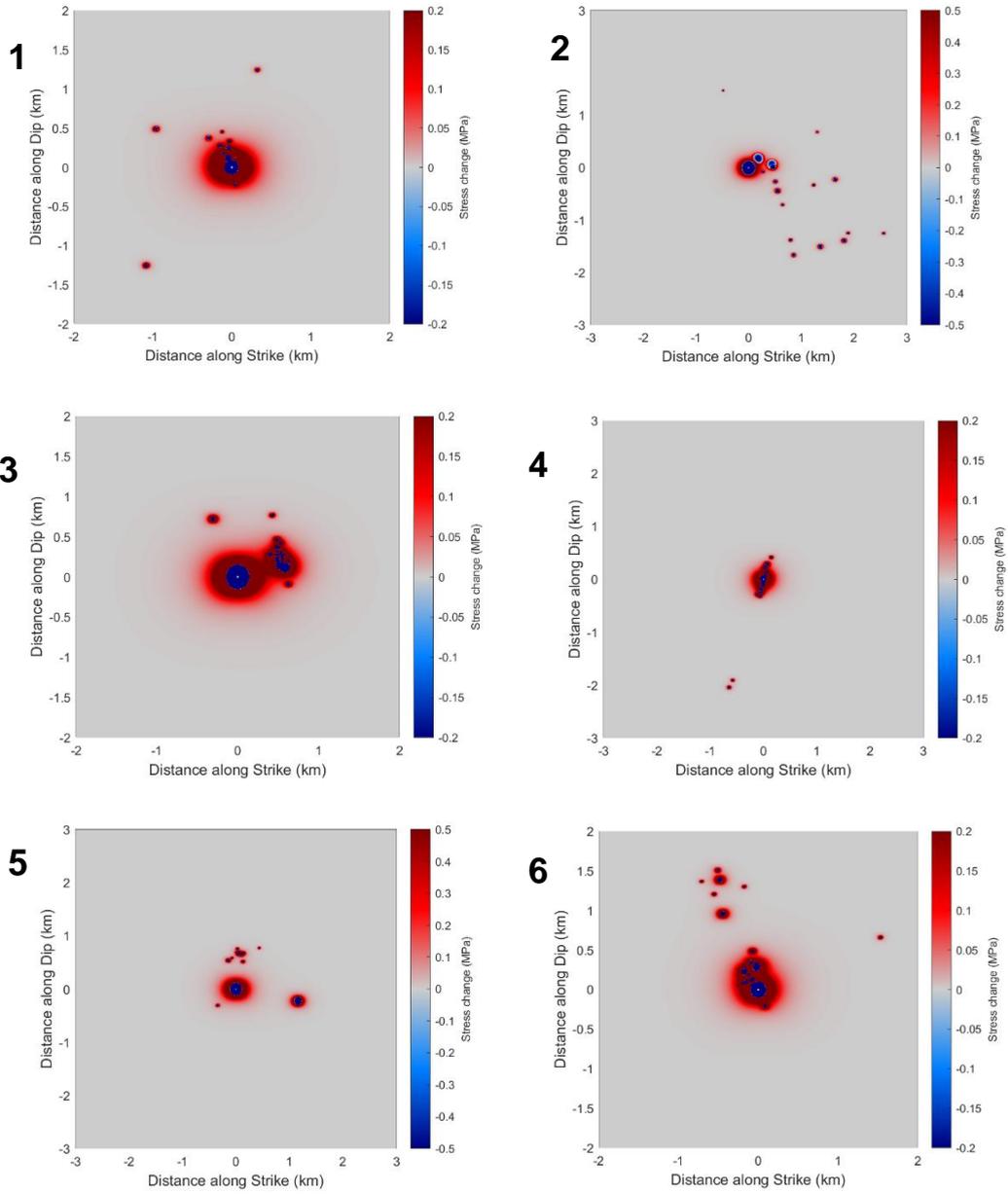


Figure 4.25) Left panel: Stress release model for the sequence IDX 1. Right panel: Histogram of the nearest neighbor distance for the relocated events of the sequence IDX 1.

When considering the other sequences, we observed similar scores, with a mean percentage of 60 ± 12 % of connected events (evaluated for the sequences featuring more than 40 events in the relocated catalog), with stress release model still suggesting static stress release as triggering mechanisms. Moreover, we retrieved similar dip – oriented trends, as for the example of Figure 4.25. In Figure 4.26, we report the stress release for the 8 sequences for which we estimated the stress release model.



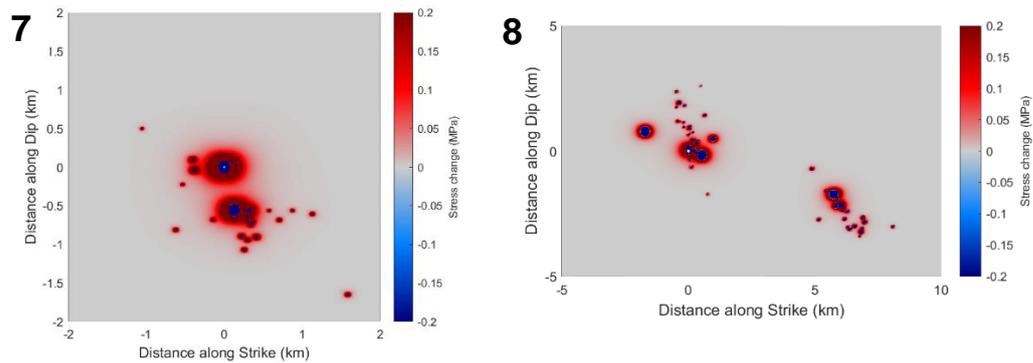


Figure 4.26) Stress release model for the considered sequences. Small events are mainly concentrated in or around the area affected by stress changes due to the main events in the sequences, suggesting stress release as trigger mechanism.

An interesting case is represented by the Rocca San Felice seismic sequence (bottom right panel in Figure 4.26). For this sequence, we observe two seismicity patterns activated at different times (the main event M_l 3.0 involved the leftmost patch, and the seismicity migrated along the rightmost segment almost two days after the mainshock with the occurrence of a M_l 2.8 earthquake). In both clusters we still observe a predominant orientation along the dip direction. Since this sequence features the highest number of events, we used it for assessing the robustness of the obtained stress release models.

As discussed in the Section 4.2.1.1, the estimation of the source radius scales with the selected geometrical shape factor (K term in Equation 4.14). In this analysis, we adopted the shape factor proposed by Brune (1970), which assumed $K = 0.37$. On the other hand, several authors have proposed alternative rupture models, resulting in different values of the K coefficient ($K = 0.21$, Madariaga 1976; $K = 0.26$, Kaneko & Shearer, 2014), therefore providing different values for the source radius and, consequently, stress drops. In particular, when substituting the K factor from the Brune model with the values proposed by Kaneko & Shearer (2014) and Madariaga (1976), the source radius is lowered by factors of ~ 1.4 and ~ 1.8 , respectively. Because of adopting these different rupture models, the stress drop increases by factors of ~ 2.9 and ~ 5.5 , respectively, as compared to the value obtained through the Brune model. For assessing the robustness of the former results, we compared the stress release model obtained through the Brune model with the corresponding one provided by the selection of the end-member Madariaga model. When comparing the stress change induced by a circular rupture modelled with the two end-member cases, the Brune and Madariaga models, in the former we have a larger radius and a smaller stress drop, while in the latter the radius is reduced, but the stress drop is increased significantly. In Figure 4.27 we represent the stress change during an earthquake having a Brune source radius of 100 m and a Brune stress drop of 1

MPa. In the plot, the stress released assuming a Brune rupture was amplified by a factor of 3 for the sake of clarity. Despite the expected differences in the released stress and rupture size, we found that the region experiencing a stress increase after the earthquake has similar size for both models.

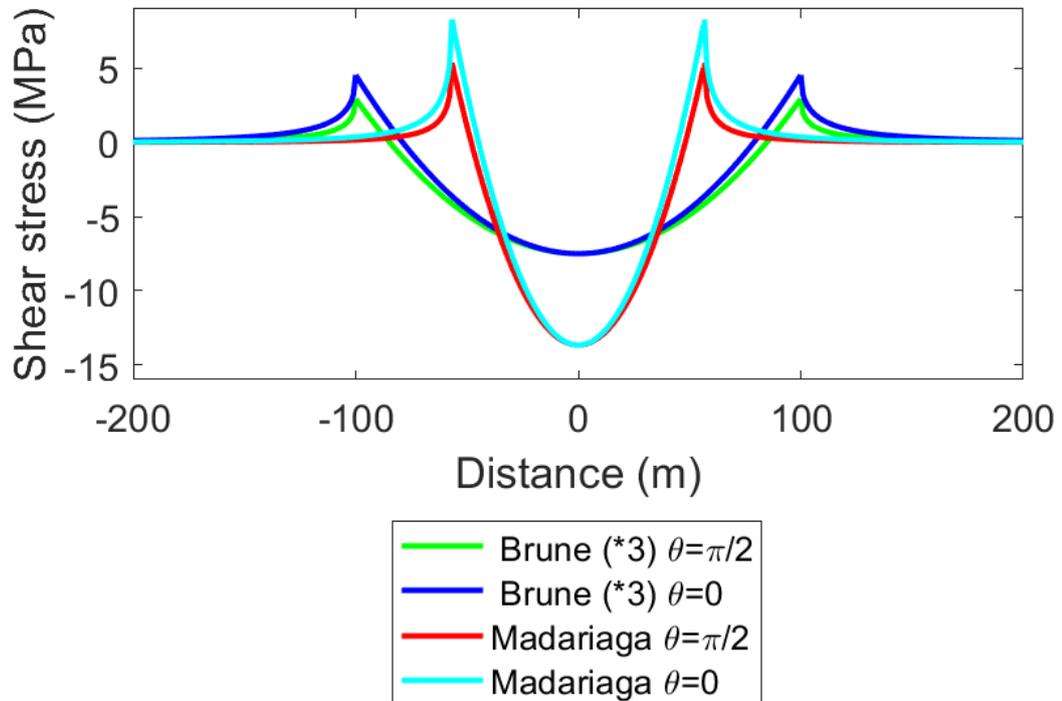


Figure 4.27 Comparison of the stress released by an earthquake with Brune radius and stress drop of 100 m and 1 MPa respectively, using Brune and Madariaga models. The solutions are represented along the directions parallel ($\theta=0$) and orthogonal ($\theta = \pi/2$) to the slip. The solution from the Brune model is amplified by a factor of 3. We found that the size of the region experiencing a stress increase after the earthquake is of the same order of magnitude for both models.

When comparing the stress change on the fault plane, obtained using Brune (Figure 4.28 – left panel) and Madariaga (Figure 4.28, right panel) models for the Rocca San Felice seismic sequence, the two models show very similar patterns in terms of stress release at the fault scale, with the main differences being related to areas where the stress decreases. This indicates the robustness of the results, even when changing the source model. However, when zooming along the fault plane and using appropriate normalized scales, we appreciate the differences between the two models (Figure 4.29).

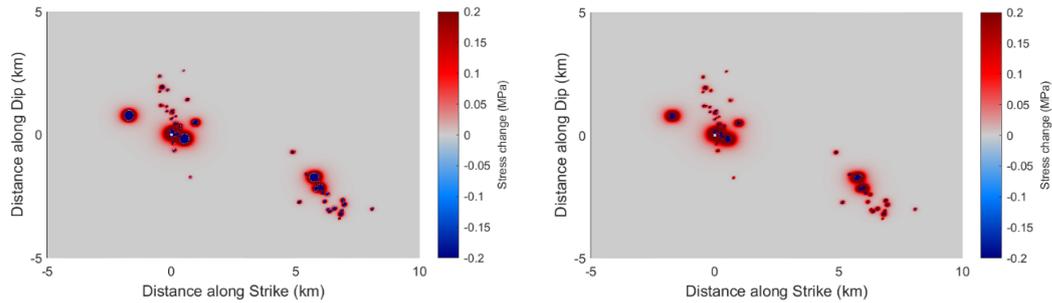


Figure 4.28) Comparison of stress changes using Brune (left panel) and Madariaga models (right panel) with the same scale for the stress for the Rocca San Felice sequence. At the scale of the fault, the two images appear very similar, with the main differences around the hypocenters of the events, within the source radius.

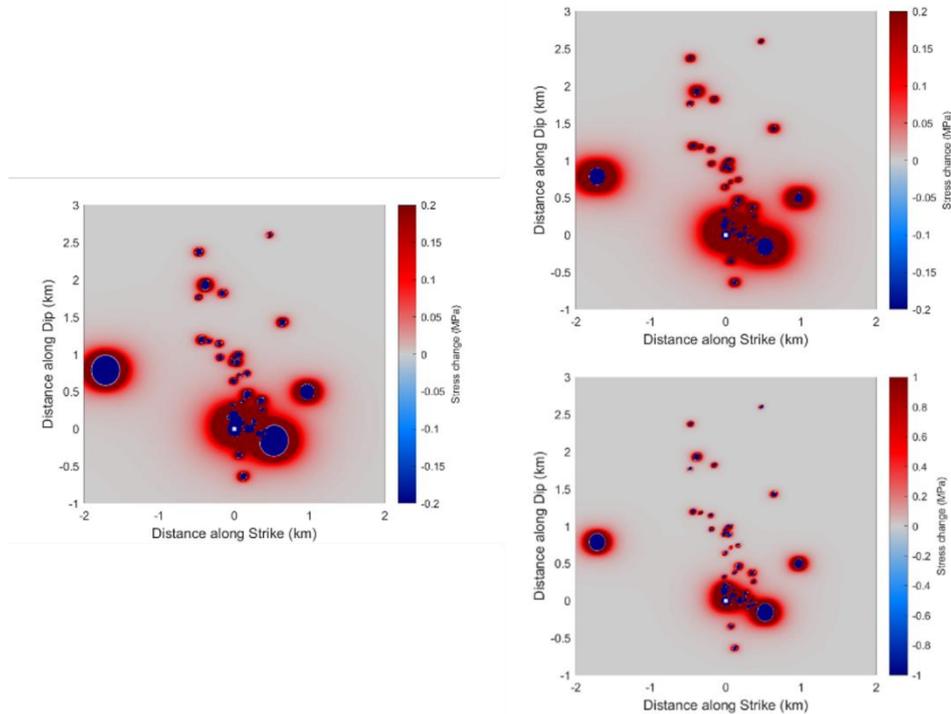


Figure 4.29) Zooming on the Figure 4.28, representing now the left cluster of the Rocca San Felice sequence. Left panel - Stress change with the Brune model; Central Panel - Stress change with the Madariaga model represented with the same scale as the Brune model; Right Panel - Stress change with the Madariaga model, represented with a wider scale that helps to identify the maximum amplitudes.

In the evaluation of the source size for low-magnitude events within each sequence, we used the median stress drop inferred from all the events of the corresponding sequence having $M > 2$, for which we computed source parameters. Although self-similarity was also claimed for the background seismicity in the area by Zollo et al., (2014), we are aware that self-similarity is not guaranteed across the scales explored by events of magnitude M_w 0.5-3.6 investigated in this study. To assess the sensitivity of the results against the self-similarity assumption, we also tested a different scaling model $M_0 \propto f_c^{-(3+\epsilon)}$, where the ϵ parameter was selected according to Picozzi et al. (2022a) for events in the Irpinia area. In Figure 4.30 we report the stress change for the Rocca San Felice sequence assuming self-similarity (left panel) and the scaling of Picozzi et al. (2022a), using $\epsilon = -0.39$, as proposed by the authors for characteristic hypocentral depths of this sequence (right panel). We observe no significant differences between the two models, due to the fact that changes in stress drops appear as changes in the source radius at a power of $-1/3$.

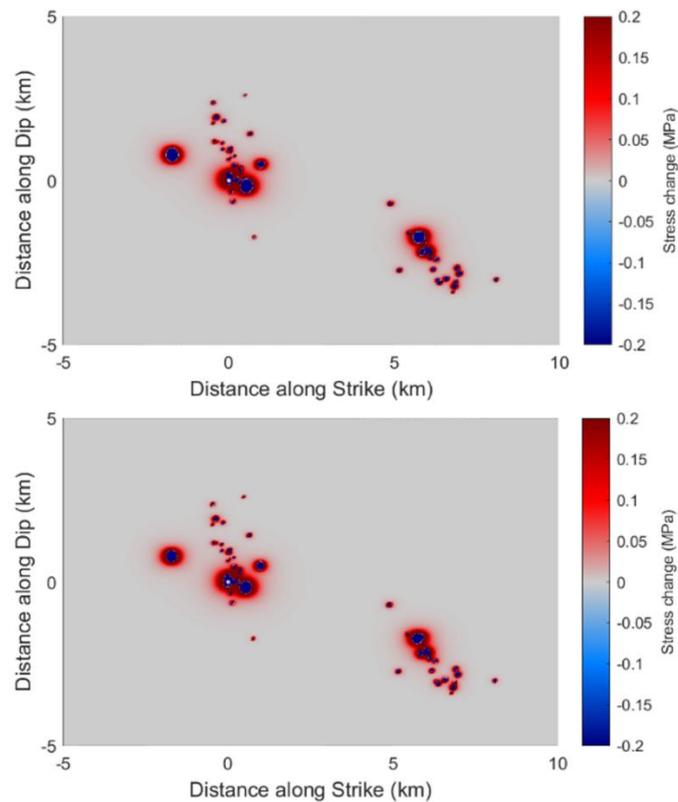


Figure 4.30 Stress changes obtained assuming self-similarity (left panel) or the scaling model $M_0 \propto f_c^{-(3+\epsilon)}$, where the $\epsilon = -0.39$ according to Picozzi et al. (2022a) for the Rocca San Felice sequence. The two maps appear very similar, with differences around the hypocenters of small events.

An important feature retrieved within this analysis is that the distribution of the events is not isotropic around the main events of the sequences, but small events tend to align dominantly along the dip direction, which also corresponds to the slip direction, for normal faults. A schematic representation for the lineations is reported in Figure 4.31.

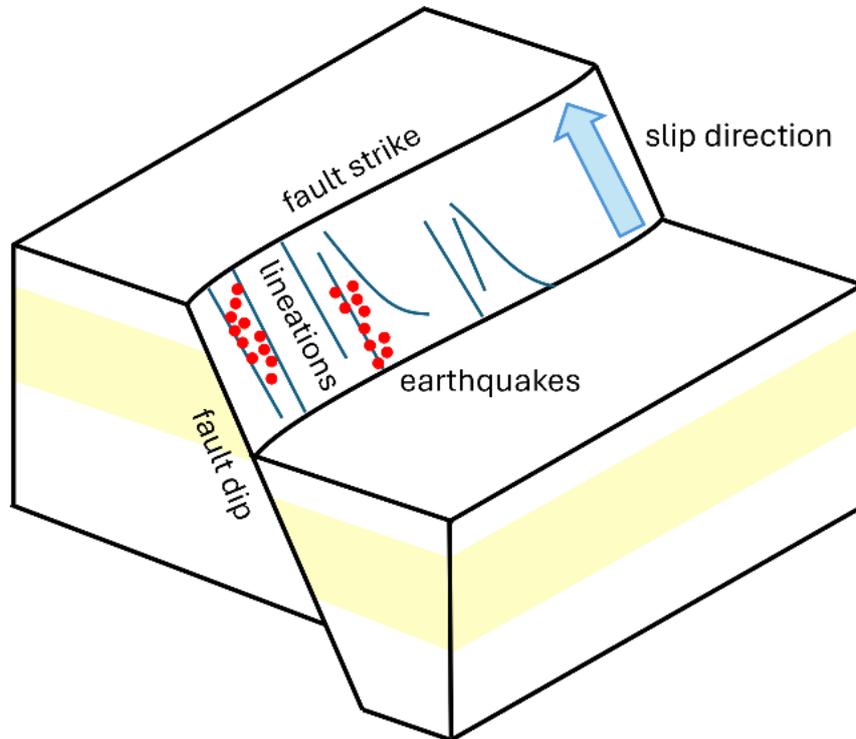


Figure 4.31) Schematic representation of seismicity lineations, oriented along the dip direction, which for a normal fault coincides with the slip direction.

Specific patterns for sequences along the direction of the slip have been observed in strike-slip environments (Rubin et al., 1999; Shearer 2002). Lineations of the seismicity along the major faults in California have been interpreted as the boundary between locked and creeping domains (Rubin et al., 1999; Rubinstein & Beroza 2007). In the normal fault environment of the Southern Apennines, the evolution of the seismicity during the sequences is also controlled by slip and cannot be explained by the anisotropic stress release after the event (Andrews, 1980). Fault roughness, modulated by repeated stick-slip episodes may determine predominant patterns at the scale of the microseismicity observed here (10 – 100m), with striations mainly oriented along the dip direction (Candela et al., 2011). Corrugated faults behave as geometrical asperities and can localize deformation hosting stick-slip episodes at small scales (a few centimeters of slip) (Resor & Meer, 2009). Fault roughness and geometrical barriers at this scale may also impede small events from growing into larger magnitude earthquakes (Sagy et al., 2007; Marshall

& Morris, 2012). These strips can also favor upward migration of fluids, although we cannot discern a signature of diffusion-dominated processes from the space-time evolution of the sequences. At generally larger scales, a M4.0 event was reported to precede and trigger the M 6.1 L'Aquila earthquake in Central Apennines (Cabrera & Poli, 2023).

The occurrence of aseismic slip episodes near the lineations could also be the cause for the along-dip evolution of the seismicity and might explain the longer extent compared to the released seismic moment. Aseismic transients have already been observed in normal fault environments during the occurrence of larger seismic sequences (Gualandi et al., 2017; Kaviris et al., 2021). However, for the sequences analyzed here, geodetic data has not detected aseismic transients at this space-time scale during the sequences analyzed in this study. The Rocca San Felice sequence shows the activation of two parallel clusters, oriented along the dip direction, but about 5 km apart. The two clusters featured mainshocks of similar magnitude (M_l 3.0 and M_l 2.8, respectively), a kilometeric size extension along the dip (4 km and 2 km), with the first evolving preferentially up-dip and the second one, activated about two days late, downdip. As shown in Figure 4.28, the stress perturbation associated with the first sequence cannot be responsible for the activation of the second patch. Also, the lack of seismicity between the two segments does not support the hypothesis of fluid migration as being responsible for triggering the second cluster. According to the rate of occurrence of independent events with $M_l > 2.5$ in the northern part of the region ($\lambda = 2.1 * 10^{-3} ev/day$), we estimated the probability of occurrence of two independent events within 2 days to about 0.4%. We, therefore, tested the hypothesis of aseismic slip between the two seismicity clusters of this sequence. We assess the evolution of the displacement at the three closest GPS stations SNAL, MTMR, ANG1, the first two belonging to the INGV-RING network, the latter to the Regione Campania. The time series of daily coordinates (see D'Agostino et al., 2020 for details of GPS data processing) at the three stations have been checked for possible offsets across the seismic sequence. Evaluating the average positions in North, East and vertical coordinates, before and after the Rocca San Felice sequence, we could not find significant static offsets within the estimated error bounds (Figure 4.32).

We assessed the maximum average slip allowed on a deep dislocation whose displacement on the surface would not emerge from the noise level at the three GPS stations. We centered the potential aseismic dislocation between the two clusters and used the fault geometry and kinematics inferred by the composite focal mechanism solution of Festa et al. (2021), calculating the surface displacements using the Okada techniques (Okada, 1992). We tested a range of uniform slip on the dislocation between 25 and 75 mm, assuming constant strain drop of 10^{-5} . The relatively deep position of the dislocation centroid between the two clusters (~8 km) allows slip on the deep dislocations up to 50 mm without detection at the surface (Figure 4.32). For slip larger than 50 mm, the non-

linear increase of cumulative seismic moment determines surface displacements outside the range of allowed offsets. Thus, an aseismic event of $M_w \sim 5.0$ could have occurred during the sequence, transferring stress across the two asperities without producing a signal that would have been visible at the GPS stations.

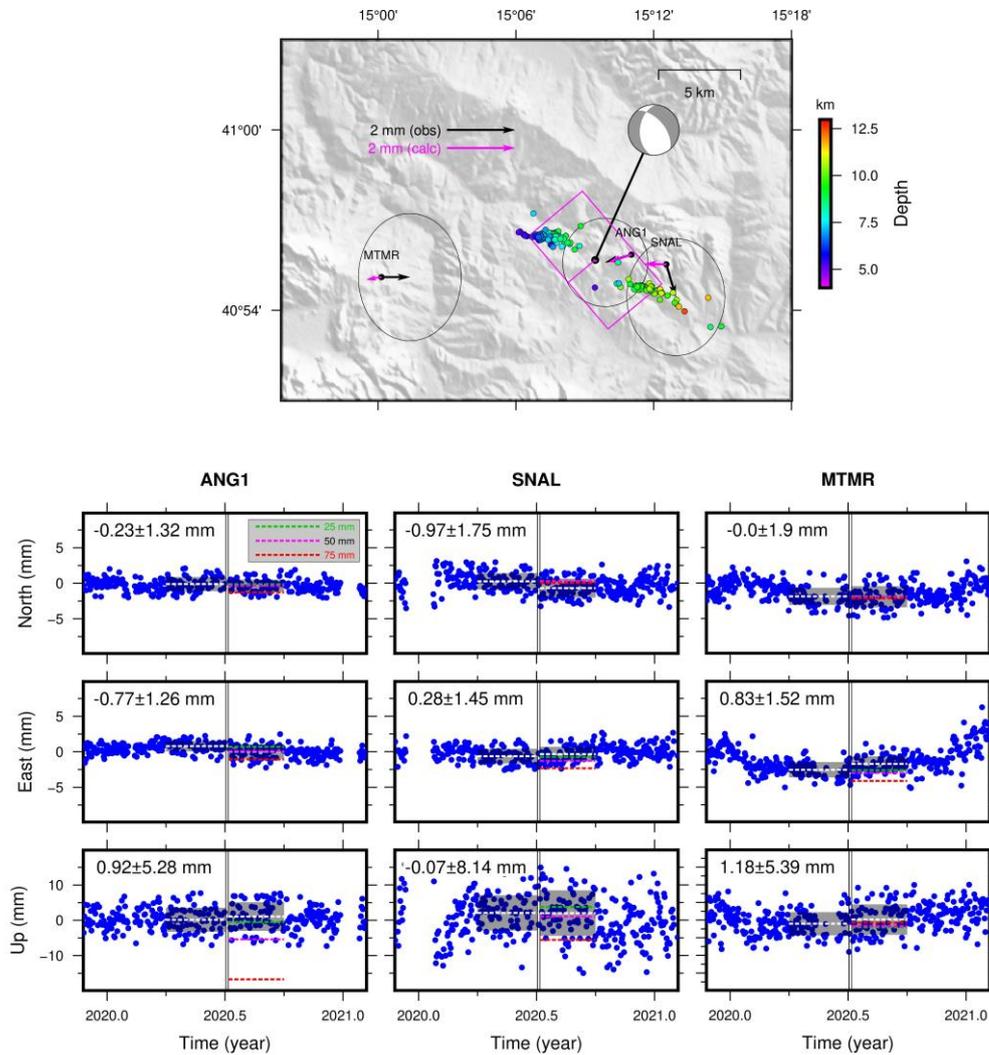


Figure 4.32) Upper panel: map of the GPS stations near the Rocca San Felice seismic sequence and associated displacements. The vectors show the horizontal displacements with 1-sigma error ellipses, from the static offset during the seismic sequence. Three different synthetic scenarios were evaluated assuming increasing slip (25, 50, 75 mm), with geometry and kinematics from Festa et al., 2021. The surface displacements produced by a slip of 50 mm on the fault drawn with purple lines is shown in the map (purple arrows). Lower panel. GPS position time series and average positions (white dashed lines) before and after the seismic sequence, with a 1-sigma error shown with grey shading. Calculated displacements are shown (dashed lines) for the 25mm (green), 50mm (purple) and 75 mm (red) dislocation scenarios.

4.3.3 Focal mechanism estimation for earthquakes in Southern Apennines using dense array deployments

In Section 2.3.2, we described the temporary deployment of 20 dense seismic arrays of 10 stations each in the Southern Apennines, which continuously recorded ground motion velocity from September 2021 to August 2022. In this analysis we aim to produce an enhanced catalog of focal mechanisms for the area, integrating the stations from the ordinary seismic network and the dense survey (231 stations).

Southern Apennines are dominantly characterized by a normal-faulting tectonic regime, subject to an extensional regional stress field (Bernard & Zollo, 1989; Montone et al., 1999; De Matteis et al., 2012; Festa et al., 2021; Tarantino et al., 2024). The area was struck by the destructive 1980 M 6.9 Campania-Lucania earthquake (Westaway & Jackson, 1987; Bernard & Zollo, 1989), a pure normal-faulting event that occurred on an approximately 60-km-long NW-SE-striking fault segment with three main rupture episodes at 0, 18 and 39 seconds after the first nucleation. Since 1980, the largest event that occurred within the epicentral area of the 1980 earthquake was the 3 April 1996 M_l 4.9 earthquake, also characterized by a normal-faulting mechanism (Cocco et al., 1999). Two moderate magnitude seismic sequences occurred between 1990 and 1991 in the Potenza region, located about 40 km southeast of the 1980 Irpinia aftershock area (Ekstrom, 1994). The two mainshocks (M_l 5.2 and M_l 4.7) and the larger events of the sequences were characterized by strike-slip faulting mechanisms, with the preferred fault planes having an east–west orientation (Di Luccio et al., 2005). In Figure 4.33, we present the focal mechanisms for the discussed earthquakes (from De Matteis et al., 2012).

Several authors have inspected microseismicity in the area for inferring information on the focal mechanisms. Pasquale et al. (2009) selected 2352 $M < 4$ aftershocks of the 1980 M 6.9 earthquakes, which resulted in a catalog of 139 fault plane solutions (6 % of the starting earthquake list). Most of the retrieved focal mechanism solutions indicate normal component faulting (pure normal faulting and oblique-normal faulting). Only some solutions show strike-slip or inverse faulting. De Matteis et al. (2012) selected 1312 earthquakes with magnitudes ranging from 0.9 to 3.1, that occurred within August 2005 and April 2011 and were recorded by the INFO and INGV networks, which are much denser than the seismic network that recorded the 1980 M 6.9 aftershocks used in Pasquale et al. (2009). The authors estimated the fault planes for 118 earthquakes (9% of the starting earthquake list). The majority of focal mechanisms still show a dominant normal-faulting mechanism with pure normal faulting and oblique-normal faulting. Only a few solutions show pure strike-slip or reverse faulting. As also observed by Pasquale et al. (2009), in the Potenza area most focal mechanisms show strike-slip kinematics.

Therefore, the authors claimed that the background microearthquake activity in the studied sector of the Apenninic chain is controlled by two major fault systems: a sequence of subparallel northwest–southeast-trending normal faults and an approximately east–west oriented strike-slip fault transversely cutting the chain (De Matteis et al., 2012).

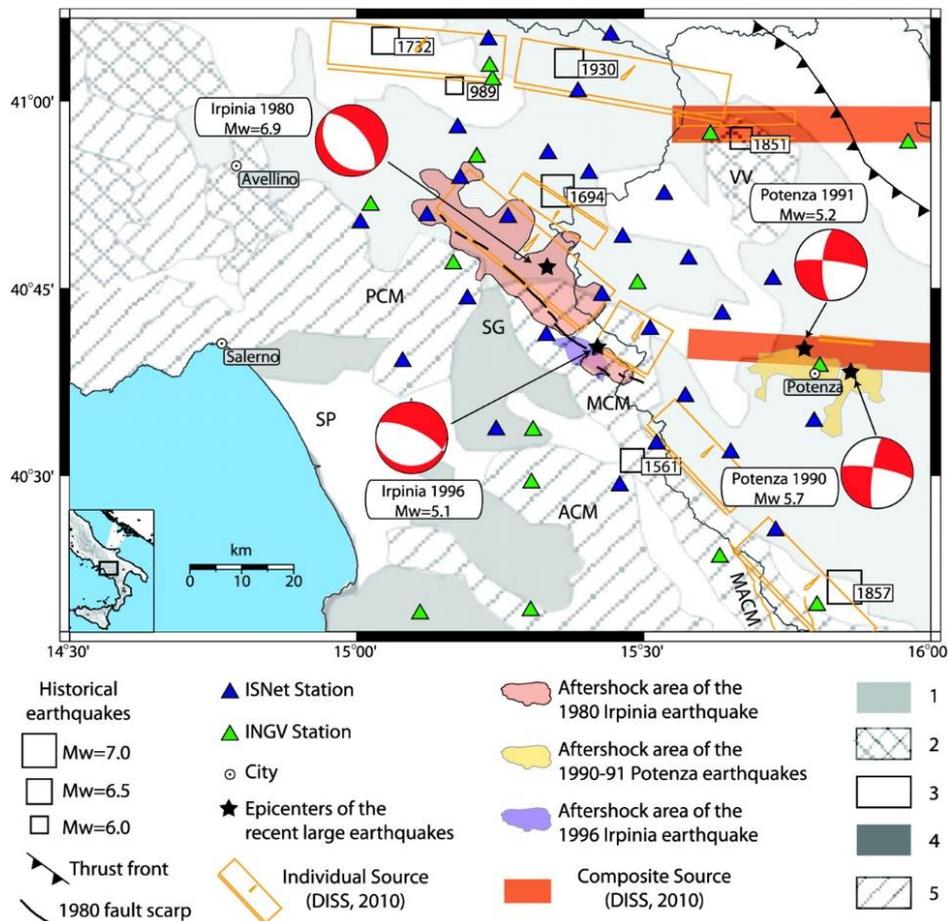


Figure 4.33) Focal mechanisms for the main earthquakes occurred in the Southern Apennines. The destructive 1980 M 6.9 and the 1996 M 5.1 Irpinia earthquakes feature pure normal-faulting mechanism, while the events of the 1990-1991 Potenza seismic sequence mainly report east-west oriented mechanisms. Figure extracted from De Matteis et al. (2012).

Although in Section 2.3.2 and Section 3.3.2 we generated and located enhanced detection catalogs using solely the stations of the dense arrays, producing a catalog of 2.2k located earthquakes with a magnitude of completeness close to $M_l 0$, we do not expect to obtain focal mechanisms for most of these ultra-low magnitude events. Indeed, the signals of

most of these earthquakes typically emerge at only one or a few arrays, having a low signal-to-noise ratio and leading to a complexity in the evaluation of the first motion polarities to be used for obtaining the geometrical properties of the source, as well as an inadequate coverage of the focal sphere. We therefore selected the events which were included in the existing manual INFO catalog (<http://isnet-bulletin.fisica.unina.it/cgi-bin/isnet-events/isnet.cgi>), which contains ~400 earthquakes in the considered time window and features a magnitude of completeness of M_c 1.1. In Figure 4.34, we report the magnitude distribution of the events in the manual INFO catalog.

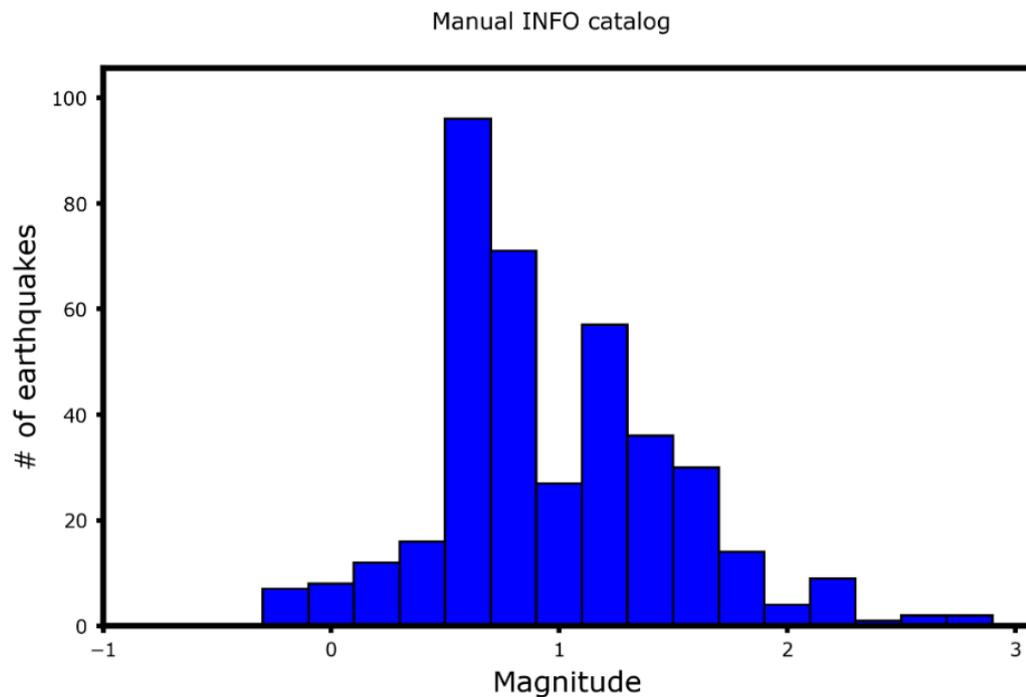


Figure 4.34) Magnitude distribution of the events in the manual INFO catalog from September 2021 to August 2022.

Within the manual INFO catalog, the focal mechanism has been evaluated for 6 events only (~2% of the catalog), whose M_l ranges between 1.8 and 2.4. The number of used first motion polarities for these events ranges from 6, which also represents the minimum number of readings required for the inversion in the standard procedures of the network, and 11, this latter value was observed for the M_l 2.4 event. Among the 6 earthquakes for which the focal mechanism was estimated, 4 events report a normal focal mechanism, coherent with the main fault segments in the area, while the remaining earthquakes reported a less reliable strike-slip or inverse faulting. In Figure 4.35, we present the focal mechanisms for the manual INFO catalog, along with the main fault segments (DISS,

2021) and the stations of the ordinary seismic network with red triangles. The size of the reported beachball is proportional to the magnitude of the earthquakes.

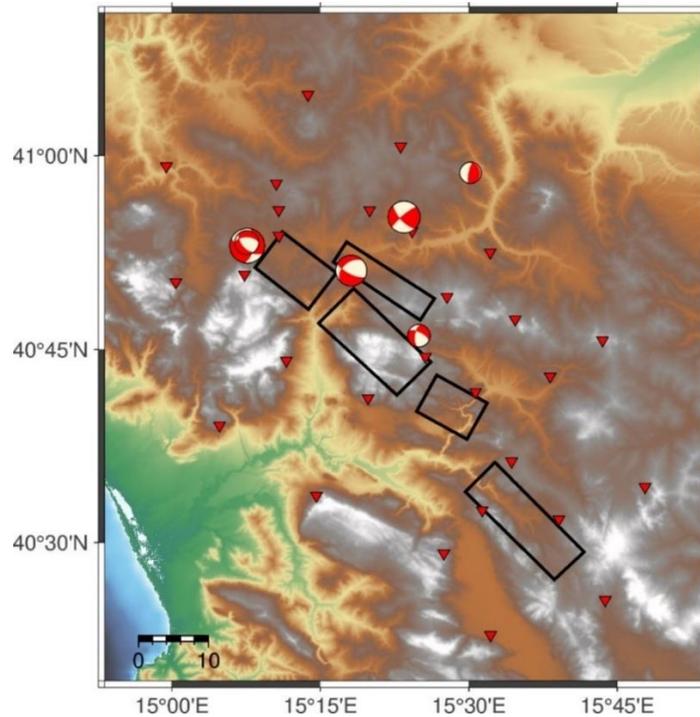


Figure 4.35) Focal mechanisms for the events in the manual INFO catalog (6 events). Among the 6 earthquakes for which the focal mechanism was estimated, 4 events report a normal focal mechanism, coherent with the main fault segments in the area, while the remaining earthquakes reported a less reliable strike-slip and inverse faulting.

To estimate the first motion polarities, we followed the deep-learning approach proposed by Ross et al. (2018), with a convolutional neural network (CNN) to identify the sign of the first motion P wave polarity. Each convolution layer consists of a set of learnable filters that are convolved with the outputs of a previous layer to identify patterns of interest anywhere within that data subset. In this case, learnable means that the coefficients of the filters are optimized along with all the other coefficients of the network during the training process. After convolution, pooling layers are commonly used to decimate the convolution output so that subsequent layers learn attributes of a rescaled representation of the original input data. This helps in recognizing variants of the same objects with different sizes, and it leads to an indirect connection between the only locally connected neurons of any individual pair of layers with the distant neurons of more shallow layers. As a training dataset, a massive dataset of 273,882 earthquakes recorded by the SCSN (Southern California Earthquake Data Center, 2013) from 2000 to 2017 at 692 stations was used. The selected stations are all deployed at an epicentral distance up

to 120 km and are composed of both accelerometer and velocimeter sensors. These seismograms are associated with 2,530,857 first-motion polarities assigned by SCSN analysts. Each record is labelled as up (U), down (D), or unknown (K), based on whether a first-motion polarity was assigned by the analyst for the station of interest. Records for which an analyst has determined a P wave arrival time but not assigned a first-motion polarity are assigned to the label K. In Figure 4.36 we report an example of waveforms manually labelled as U, D and K by SCNS analysts, along with the signal-to-ratio (SNR) distribution of each class in the training dataset. We can observe that the histogram of the SNR for the assigned polarities is typically more populated at higher SNR values with respect to the analogous distribution of the unknown polarities, indicating an easier estimation of the first motion amplitude by SCSN analysts.

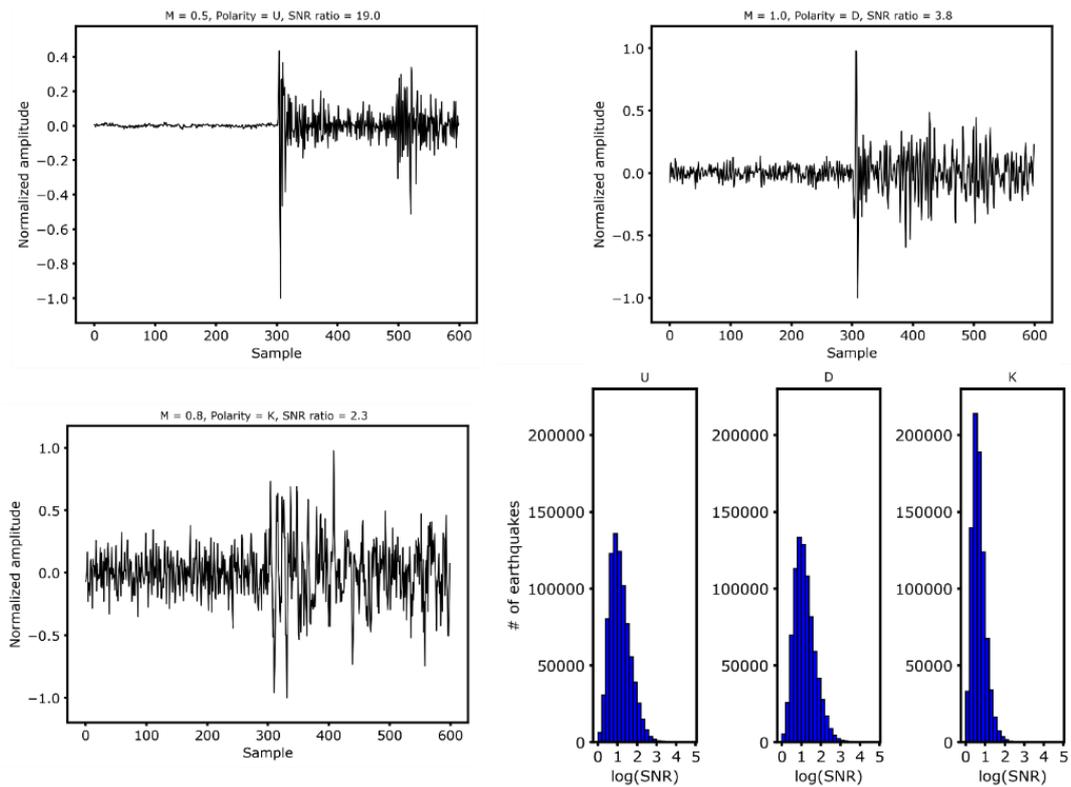


Figure 4.36) Instance of seismograms related to polarity U (upper-left panel), D (upper-right panel) and K (lower-left panel), respectively. Lower-right panel: distribution of the signal-to-noise ratio associated with each of the three polarities classes (Ross et al., 2018).

Before training the CNN, data have been down-sampled to 100 Hz, detrended, filtered with a causal Butterworth filter between 1 and 20 Hz and trimmed for generating a snippet

of a 4-s-long centered on the P wave arrival. Next, the amplitudes in each seismogram window are normalized by the peak absolute amplitude in the window, to suppress the influence of amplitude variations with magnitude or distance. Finally, the training dataset is split into 75% for training and 25 % for validation. We hereby report in Figure 4.37 the precision and recall curves as a function of the SNR, distance and magnitude values for the trained model. On the training dataset, the model typically features a precision above 90 % for each class where the polarities were manually assigned, over all the SNR, distance and magnitude domains. On the other hand, recall results higher than 80% in attributing the correct polarities for all the events featuring $SNR > 4$ while, despite being almost independent of the magnitude ranges, the recall starts dropping for stations with a distance range higher than 80 km, this value being larger than the investigated spatial scale, especially for low magnitude events.

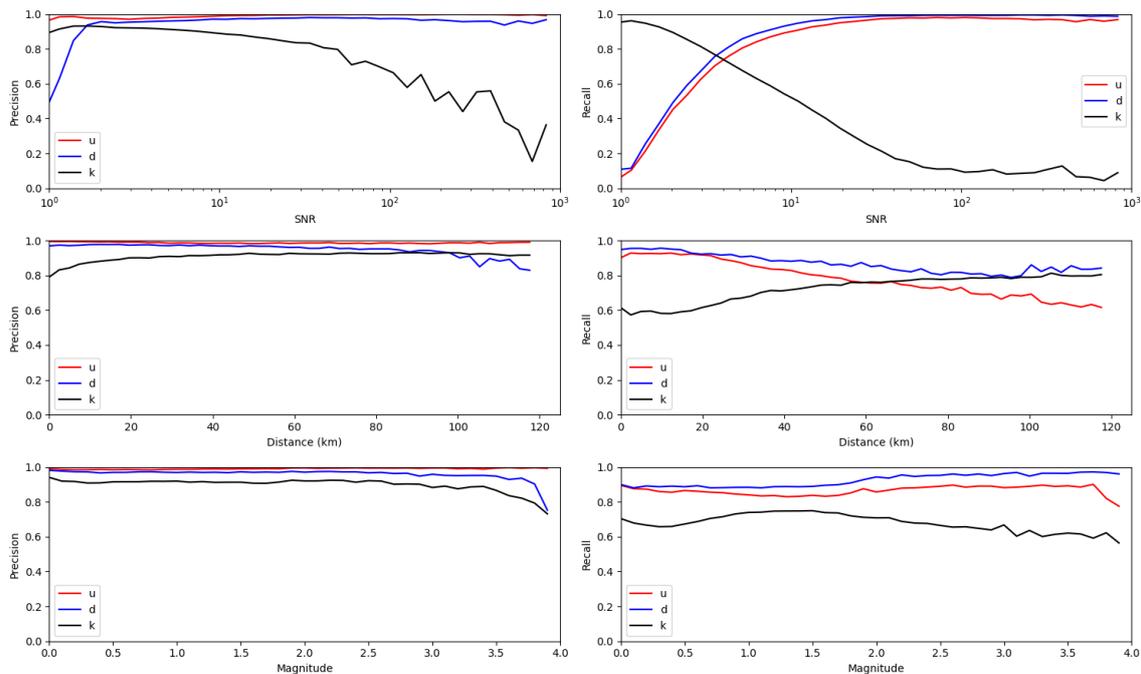


Figure 4.37) Precision (left column) and recall (right column) curves from the model training as a function of the SNR (upper panels), distance (central panels) and magnitude (bottom panels), respectively.

We finally estimated the focal mechanisms using FPFIT (Reasenberg & Oppenheimer, 1986), which inverts a set of observed first motion polarities to obtain the double-couple fault-plane solution (source model), minimizing the weighted sum of first-motion polarity discrepancies. In addition to finding the minimum-misfit solution, FPFIT provides alternative solutions corresponding to significant relative minima in misfit, which typically correspond to different faulting mechanisms from the best solution. These

solutions might be preferred after the consideration of the a-priori knowledge of the tectonic environment. As input, FPFIT requires earthquake location and take-off angles, which we computed using phase arrival times obtained in Section 2.3.2 and through the NLLoc algorithm (Lomax et al., 2000), while we used the polarities determined by the deep learning model for assigning the sign of first motion amplitude to the stations featuring the P phase arrival times.

As a first test case, we applied the deep learning technique by Ross et al. (2018) on the M_l 2.4 earthquake, which represents one of the 6 events for which the focal mechanism was also estimated in the manual catalog, inverting 9 identifications of the first motion amplitudes. Since there are no ground truth polarity estimations for the stations of the DETECT survey, we manually assigned a polarity to the stations featuring a P phase arrival time from Section 2.3.2. For the ordinary seismic stations, we extracted the polarity provided by the manual catalog. When comparing the accuracy of the automatic polarities for this event, we observe that 88 out of the 121 assigned polarities match the manual label (72%), a value significantly lower than the precision claimed by Ross et al. (2018), who obtained precision values higher than 90 % in the magnitude and distance ranges representative of this test case. Moreover, the deep learning technique assigns unknown polarities to 16 instances which were manually labelled (expected 8 U and 8 D), while it assigns a wrong polarity to 22 instances (3 manual D polarities labelled as U, 19 manual U polarities labelled as D). We therefore investigated the families of wrong and missed polarities to assess the cause of the low classification performance. In Figure 4.38, we report two instances for which the deep learning model attributed a wrong polarity estimate. From Figure 4.38 it clearly appears that, despite the high signal-to-noise ratio of the first motion, the deep learning model fails to recognize the correct P wave polarity. However, both seismograms feature a similar characteristic. While in the training dataset the correct onset for the P wave was exactly placed in the center of the analyzed window, in these two instances the real first motion amplitude precedes the center of the seismogram, despite this latter being trimmed around the automatic P arrival time. Therefore, these two instances are characterized by a delay in the phase picking, which may cause a lower performance of the deep learning classifier. Thus, these results suggest that the actual model does not generalize well, and it appears to be strongly sensitive (at the scale of the samples) to the accuracy of the phase arrival time

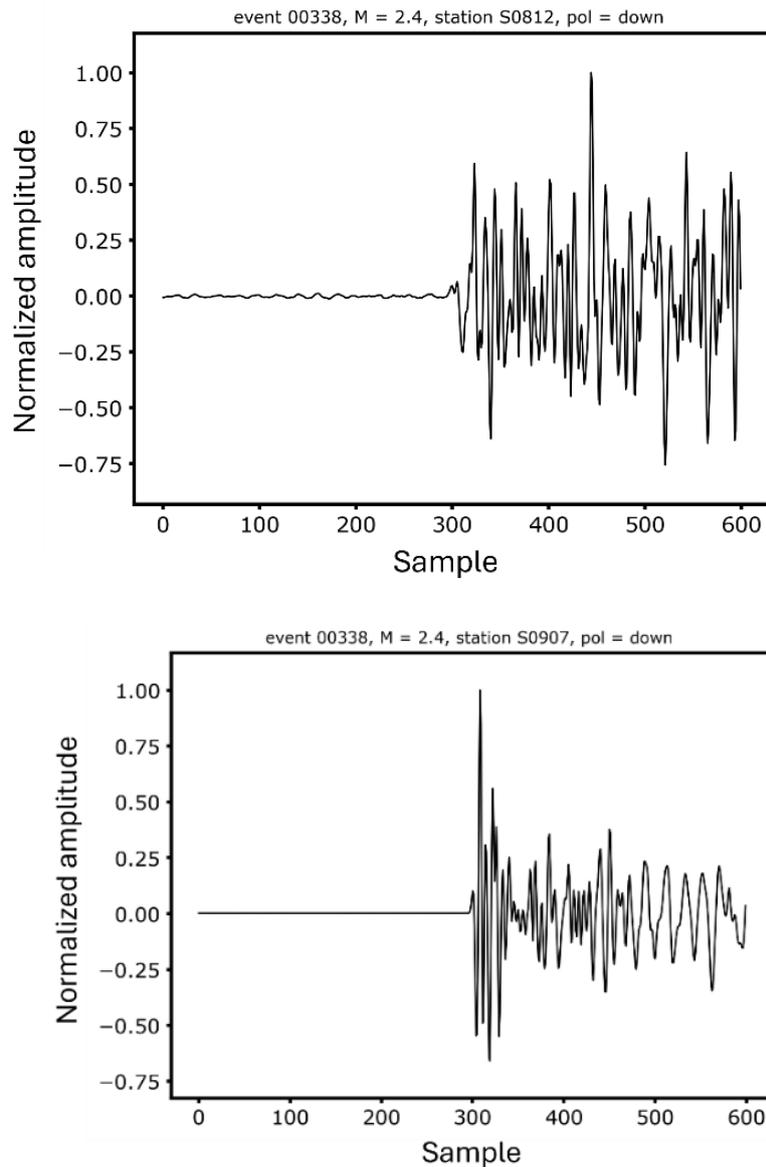


Figure 4.38) Examples of wrong polarity provided by the deep learning model. Despite the seismograms featuring a clear positive first motion amplitude for both stations, the model assigns polarity D for both instances.

To test this hypothesis, we retrained the model allowing a random shift of the window for extracting the seismogram in $[-0.5, 0.5]$ seconds ($[-50, 50]$ samples), to account for eventual errors in the automatic estimation of the phase arrival times. This range is even larger than the corresponding one obtained when considering the typical mean and

standard deviation values for the residuals between automatic and manual identification of phase arrival times (Figure 2.34; Figure 3.4; Figure 3.20) and thus represents a more conservative choice. However, on the one hand we might expect to limit the effect of fluctuations of the phase arrival times through a random shift of the position of the P onset within the considered seismogram, gaining generalization of the model but, on the other hand, we might achieve lower precision and recall values for the trained model. We hereby report in Figure 4.39, the precision and recall curves as a function of the SNR, distance and magnitude values for the novel version of the trained model.

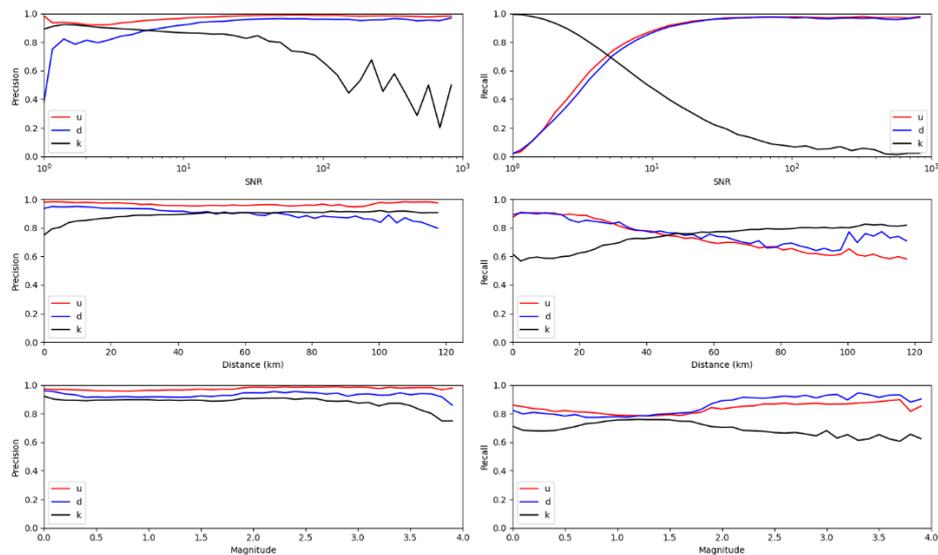


Figure 4.39 Precision (left column) and recall (right column) curves from the model trained by randomly shifting the position of the P wave onset within the seismogram, as a function of the SNR (upper panels), distance (central panels) and magnitude (bottom panels), respectively.

As compared to the respective curves reported in Figure 4.37 describing the model trained by fixing the position of the P wave onset in the centre of the analysed seismograms, the novel version of the model features slightly lower precision values among the SNR, magnitude and distances domains, but still close to 95% and 90 % for the U and D polarities, respectively. Conversely, recall starts degrading above 60 km, featuring values close to 60 %. However, this distance range is still wider than the typical distance affected by low-to-moderate magnitude events. Despite the described slightly lower classification performances, the novel version of the model is able to handle eventual shifts in the position of the P wave onset within the fed seismogram, featuring higher generalization as compared to the former model when applied to data with an automatic identification of the phase arrival times. To assess the performance of the novel version of the model, we applied it on the instances reported in Figure 4.38, characterized by a first motion

amplitude not exactly placed in the center of the seismogram due to a delay in the identification of the phase arrival time and classified with the wrong polarity class by the former version of the model. In Figure 4.40, we report the classification of the two versions of the model, indicating the probability score associated with each of the three classes.

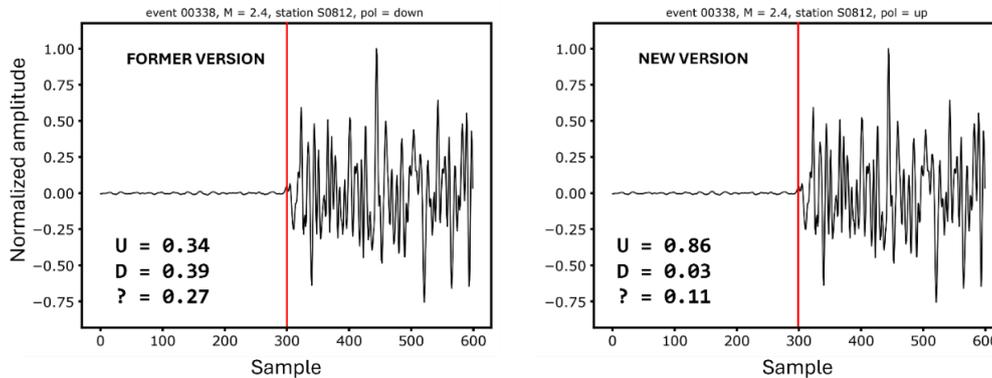


Figure 4.40) Application of the model trained by fixing the position of the P wave onset in the centre of the seismogram (left panel) and the model trained allowing a random shift of the onset in $[-0.5, 0.5]$ s in the seismogram (right panel) to an instance featuring a delay in the identification of the P phase arrival time. While the former version of the model misclassified the polarity, the novel version of the model is able to handle the earlier onset of the P wave, correctly associating the class label with a high probability score.

Despite the seismograms being extracted in the same way for both models, while the former version of the model misclassified the polarity, the novel version of the model is able to handle the earlier onset of the P wave, correctly attributing the class label with a high probability score. This result indicates that the model trained to allow a random shift of the P wave onset in the seismograms is more indicated for attributing the polarity class to data where the phase arrival times are identified through automatic techniques. When summarizing the classification results for the M_l 2.4 earthquake against the manual polarity identification, reported in Table 4.1, we noted strongly improved results compared to the former application.

Table 4.1) Summary of the classification results for the two versions of the trained model. The model trained randomizing the position of the P wave onset in the seismogram provided improved polarity classification performances compared to the former version.

	FORMER VERSION	NOVEL VERSION
SHARED POLARITIES	88/121 (72%)	111/121 (92%)
WRONG POLARITIES	22	5
MISSED POLARITIES	16	5
NEW POLARITIES	11	13

The novel version of the model correctly classified 111 out of the 121 polarities assigned manually, increasing by 20% the percentage of retrieved polarities compared to the former version of the model. Moreover, the number of missed and wrong polarities strongly diminished, from 22 to 5 and from 16 to 5 for the instances incorrectly classified and erroneously classified as unknown polarity, respectively, reducing the erroneous classification by more than 70 %. Following these results, we thus adopted the model trained by randomizing the P wave onset within the seismogram in the analysis of focal mechanisms.

Similarly to location algorithms, FPFIT allows the definition of a weighting scheme for defining the importance of the single polarity estimation in the focal mechanism determination. In particular, the algorithm accepts a discrete weighting encoding from 0 to 3, with decreasing importance in the inversion procedure. To evaluate a possible correlation between the impulsivity of the P wave onset and the estimated probability scores, we inspected the seismograms and the outputs of the classifications. In Figure 4.41, we show three instances of the same earthquake, recorded at different stations and featuring different impulsivity of the first motion amplitude. We also reported the probability score associated with each of the three classes. From Figure 4.41, we can clearly note that the probability score of the correct class associated with seismograms featuring strong impulsivity is much higher than the ones associated with more emergent onsets. Despite this feature appears straightforward, this was not the case with the probability scores associated with the deep learning pickers (Chapter 1), where no direct connection between the accuracy of the phase arrival time and the probability scores is observed (Park et al., 2023; Pita-Slim et al., 2023). To extend this observation, we investigated the possibility to extract in a linear relationship between the probability score and the signal-to-noise ratio of the P wave onset. Since this latter quantity could be directly connected to the reliability of the polarity estimation, verifying the existence of a relation between the two quantities can lead to the definition of a weighing scheme for assigning importance within the inversion procedure. For each earthquake, we performed a linear regression between the logarithm of the signal-to-noise ratio and the probability scores associated with the polarity estimations at the different stations. We then verified the consistency of the estimates in histograms for the slope and the intercept of individual linear regressions, which are reported in Figure 4.42.

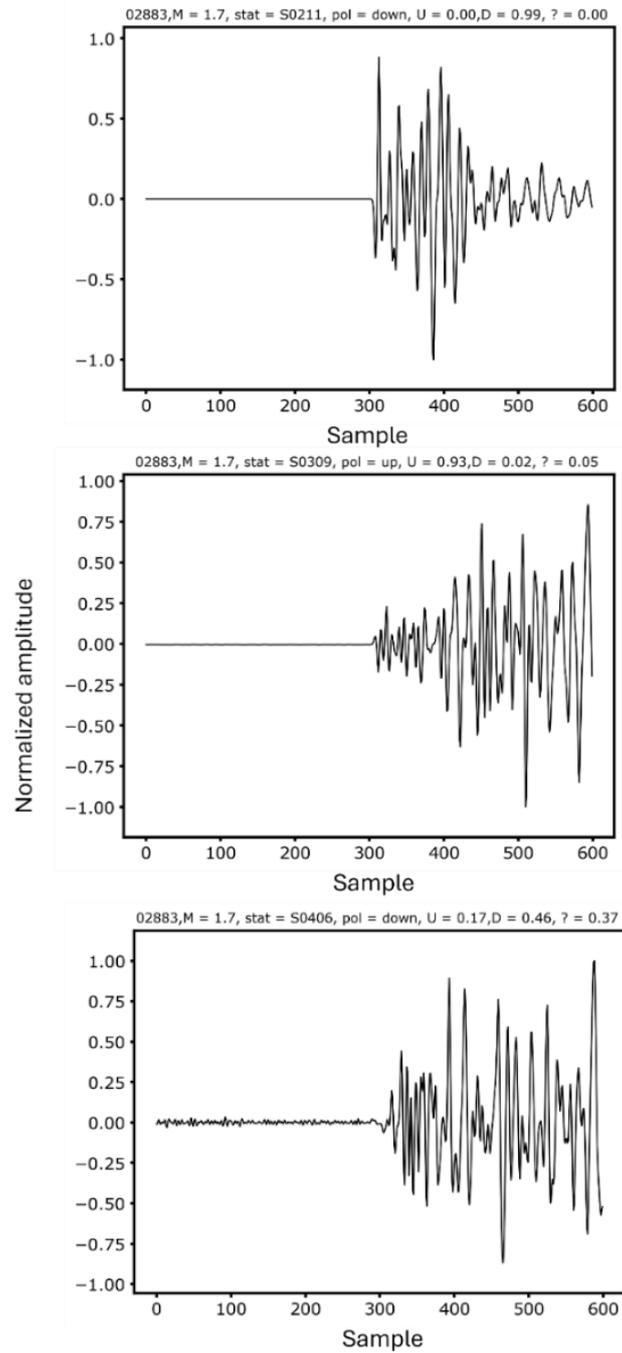


Figure 4.41) Representation of the classification outputs for three stations reporting different impulsivity of the P wave onset (decreasing from the top to the bottom). The higher the impulsivity of the P wave onset, the higher the probability score associated with the class.

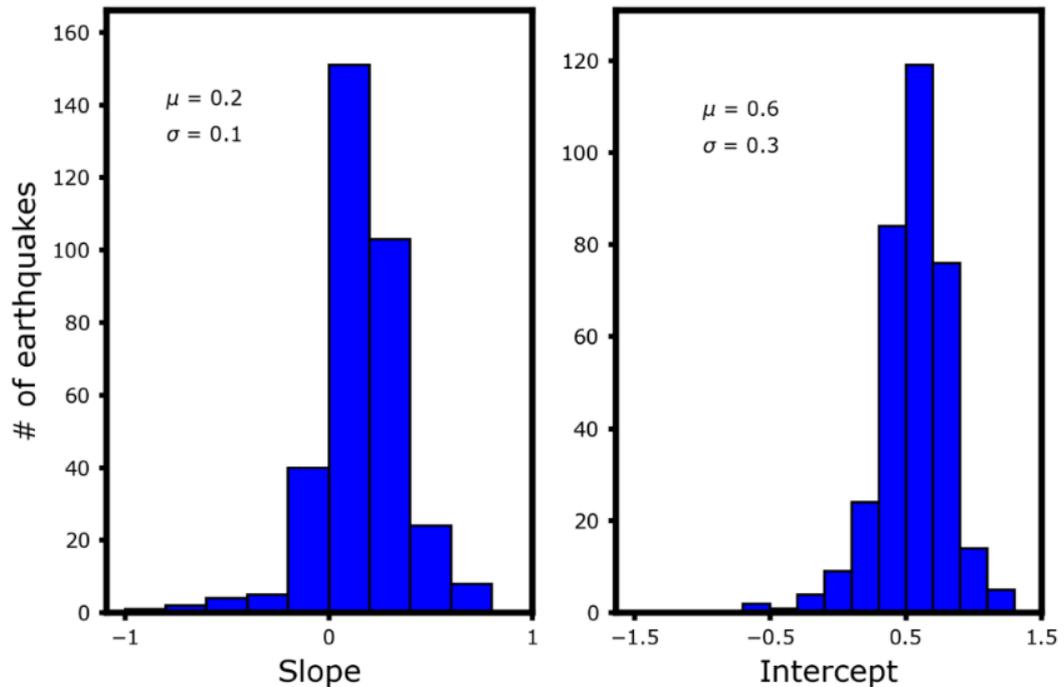


Figure 4.42) Histograms of the slopes (left panel) and intercepts (right panel) of the linear regression between the logarithm of the signal-to-noise ratio and the probability scores predicted by the classifier for the assigned polarities for each event. Coherence in the slope and the intercept suggests a correlation between the two parameters.

The coherence observed in the slope and the intercept for individual earthquakes suggests a correlation between the two parameters, leading to the possibility of the definition of a weighting scheme based on the probability scores yielded by the deep learning classifier. Moreover, the positiveness of the slopes indicates that higher accuracy can be attributed to high probability scores, due to the higher signal-to-noise ratios. We thus define a weighting scheme for the inversion procedure, as in Table 4.2. The higher the associated discrete weight, the lower the importance of the inversion procedure.

Table 4.2) Proposed weighted scheme based on the conversion of class probability scores into discrete weights.

	WEIGHT
$\text{PROB} \geq 0.85$	0
$\text{PROB} \in [0.7 - 0.85[$	1
$\text{PROB} \in [0.55 - 0.7[$	2
$\text{PROB} \leq 0.55$	3

We used this conversion scheme to attribute the importance of the polarities in the inversion scheme and we estimated the focal mechanisms using FPFIT. We hereby report in Figure 4.43 the application for a M_l 1.7 earthquake.

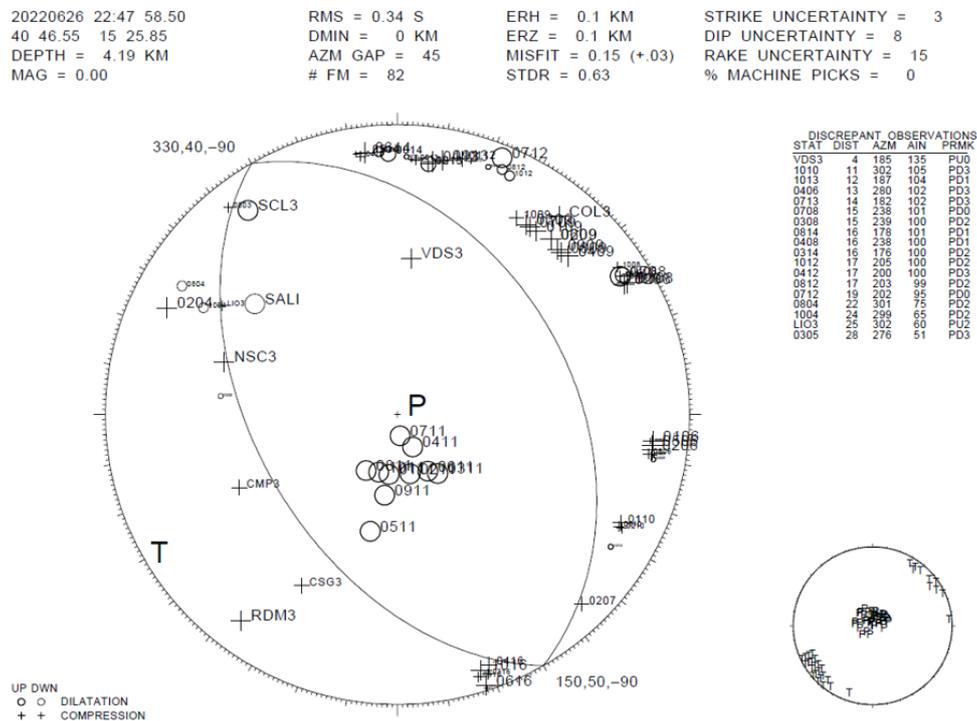


Figure 4.43) Focal mechanism estimation for a M_l 1.7 earthquake, featuring 82 polarity estimates. The focal mechanism resulted in a pure normal faulting.

For this event, 82 polarity estimates were provided by the deep learning classifier among the stations featuring a P wave phase arrival time. The focal mechanism indicates pure normal faulting, coherently with the orientation of the main fault segments in the area. However, we still observe a 15 % of discrepant polarities between the observed and expected first motion pulse predicted by the theoretical radiation pattern of the estimated focal mechanism. Although this percentage is compatible with those observed in the earthquakes of SCSN catalog and in Ross et al. (2018), we attempted to extract a tentative cutoff probability score for discarding polarities, aiming to limit the percentage of discrepant observations. For the same event, we therefore recomputed the focal mechanism selecting only polarities featuring probability scores higher than 0.70, 0.80 and 0.90, respectively. In Figure 4.44, we summarize the results of the analysis.

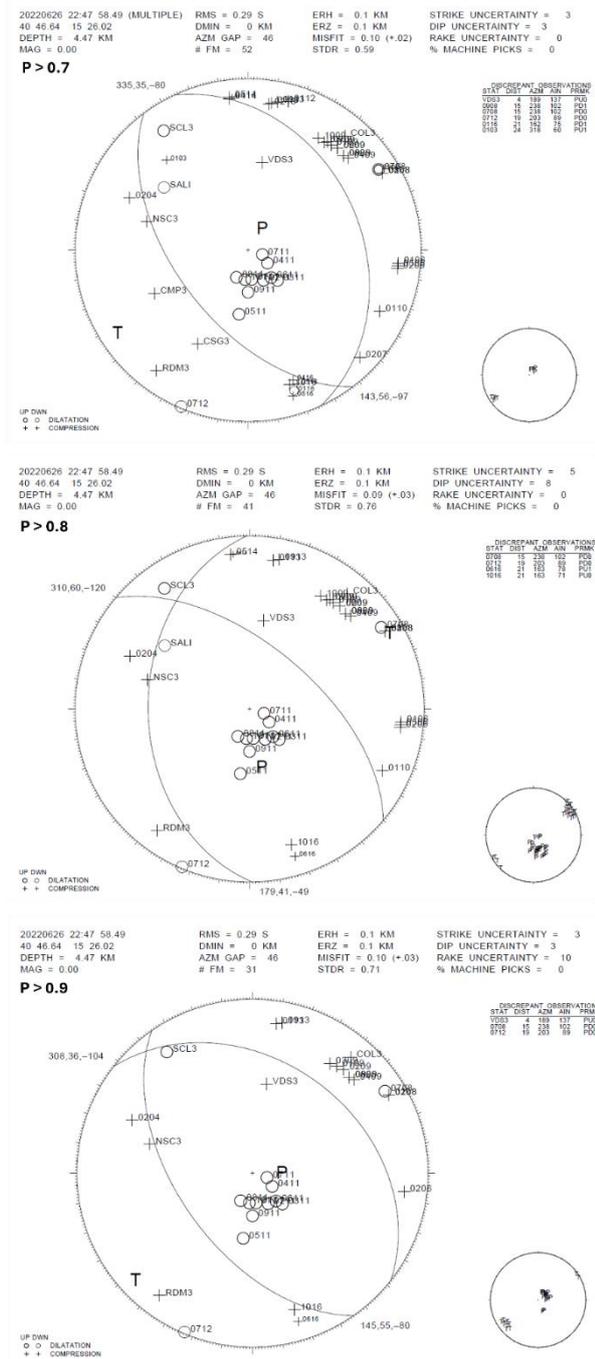


Figure 4.44) Focal mechanism selecting only polarities featuring probability scores higher than 0.70, 0.80 and 0.90, respectively.

When considering the focal mechanisms obtained by selecting only the polarities characterized by a probability score above a certain threshold, we note a progressive lowering of the number of observations from 82 polarities to 52, 41 and 31 when selecting as probability thresholds the value 0.7, 0.8 and 0.9, respectively. However, although at first sight the solutions might appear similar to the starting estimation, the loss of stations might lead to a different position of the nodal planes. Indeed, when inspecting the focal mechanisms obtained by selecting the polarities with a probability score higher than 0.8, we note a higher strike-slip component as compared to the pure normal faulting characterizing the solution obtained by inverting all the available polarities. The presence of a major strike-slip component in the focal mechanism seems to be driven by discarding the station S0207, whose U polarity constrained the position of the nodal plane in the former focal sphere. Indeed, as reported in Figure 4.45, the mentioned station features a faint P wave onset, which suggests the closeness of the receiver to a nodal plane on the focal sphere.

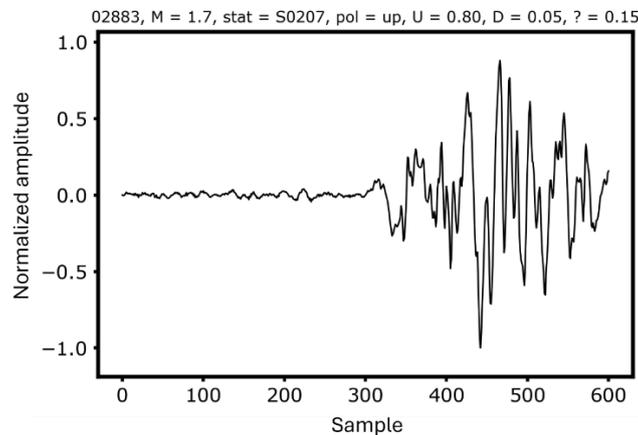


Figure 4.45) Seismogram related to the station S0207, whose polarity is discarded when considering probabilities higher than 0.80. However, this station is expected to appear on the focal sphere nearb a nodal plane, better constraining the mentioned plane.

Therefore, discarding stations due to lower probability scores might lead to a loss of crucial information for constraining the geometrical properties of the fault. Moreover, although the number of discrepant observations compared to the expected first motion amplitude is reduced by imposing a higher probability score, the percentage of discrepant polarities oscillates around similar values. While in the starting solution the percentage of discrepancy was around 15%, the respective ratios become 12% when adopting a probability threshold of 0.7, and 10% when selecting 0.8 and 0.9 as a lower limit. Thus, given the similar percentages of discrepant polarities and no significant improvement in the estimation of the focal mechanisms, we decided not to set a probability threshold for estimating the geometrical properties of the faults.

We extended the analysis of the focal mechanisms for the entire set of available earthquakes. We estimated the focal mechanisms for 223 earthquakes in the starting catalog (around 55%), which is higher than the number of fault planes estimated in the area by De Matteis et al. (2012), covering almost 5 years of microseismic events. Moreover, the percentage of events for which the focal mechanism was estimated is much higher than those observed in previous studies in the area, increasing by a factor of 9 and 5.5 the percentages of focal mechanisms determined in De Matteis et al. (2012) and Pasquale et al. (2009), respectively. Therefore, the use of dense array deployments can provide higher resolution in determining the fault planes compared to the use of ordinary seismic networks, strongly reducing the minimum magnitude for which the geometrical properties of the fault can be evaluated. We observed different percentages of determined focal mechanisms among the magnitude classes: we estimated fault planes for 75% of the events with $M_l > 1$, while the percentage of characterized events approaches 40% for the events featuring $0.5 < M_l \leq 1$. As expected, the percentage of characterized events with $M_l \leq 0.5$ strongly drops, achieving a value slightly larger than 25%. This results from the characteristic number of estimated polarities for earthquakes in each magnitude range. In Figure 4.46, we report the histograms of the number of estimated polarities for events in each of the three magnitude classes.

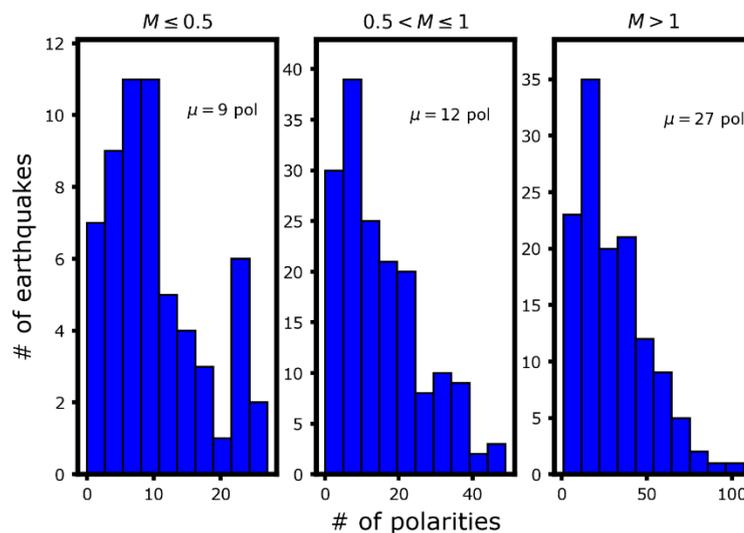


Figure 4.46) Distribution of the number of polarities used within the inversion of focal mechanisms for earthquakes in three magnitude classes: $M_l \leq 0.5$, $0.5 < M_l \leq 1$ and $M_l > 1$. Within each subplot, the median number of polarities is reported.

We estimated a median number of polarities of 9, 12 and 27 for earthquakes featuring $M_l \leq 0.5$, $0.5 < M_l \leq 1$ and $M_l > 1$. For this latter family, the number of available

polarities range from 3 to 5 times the analogous value as for the focal mechanisms determined using the manual polarities extracted from the ordinary INFO network.

In Figure 4.47, we report the catalog of estimated focal mechanisms, along with the known fault traces. As for 3.28, the size of the beachballs is proportional to the magnitude of the events and the red and white sectors refer to the compressive and extensional quadrants of the focal sphere, respectively. When multiple solutions were provided for an earthquake, we chose the most reliable solution according to the a-priori knowledge of the fault system in the area.

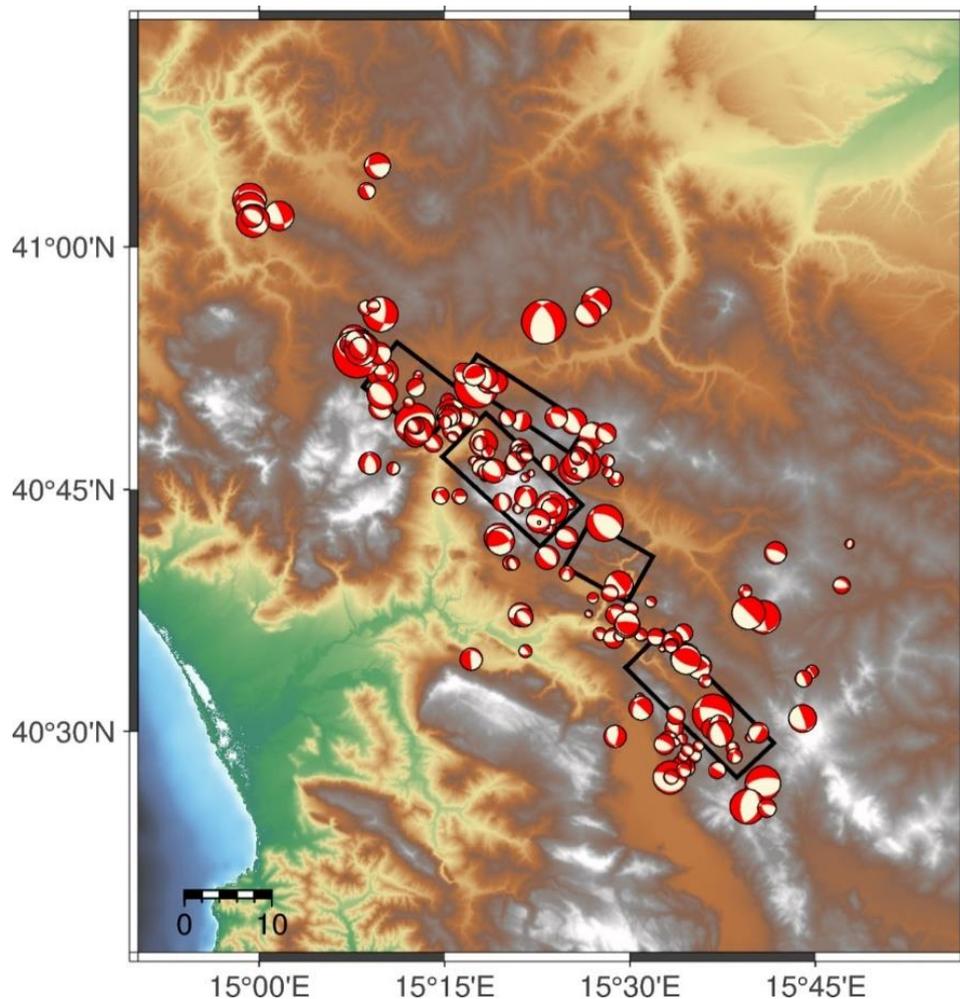


Figure 4.47) Focal mechanisms obtained from the inversion of the polarities from the ordinary seismic network and the dense array deployments. The size of the beach balls is proportional to the magnitude of the earthquakes. The 223 obtained fault planes mainly show pure normal faulting mechanisms, with a slight percentage of strike-slip component and inverse faulting.

The 223 estimated focal mechanisms mainly show a close-to-pure normal faulting mechanism, with a slight percentage of strike-slip component or inverse faulting for a low percentage of the fault planes. Colocated events typically share similar geometric properties of the source. In particular, three main clusters can be identified from this representation: two in the Northern sector of the Irpinia region, while the other occurred in the Southern sector of the investigated area. The northernmost cluster, composed of 4 earthquakes that occurred within the same day slightly outside the dense array deployment and whose magnitude ranged between 1.2 and 1.8, reported a normal faulting mechanism with a minor strike-slip component. The average geometrical angles resulted to be $\Phi = 285 \pm 15^\circ$, $\delta = 43 \pm 6^\circ$, $\lambda = -113 \pm 14^\circ$. Regarding the cluster of the Southern sector, composed of 13 earthquakes, belonging to a seismic swarm that lasted one day and featured a mainshock of M_l 1.8, the angles are similar, with a steeper dip and an even lower strike-slip component. The average angles for the cluster are $\Phi = 300 \pm 20^\circ$, $\delta = 56 \pm 17^\circ$, $\lambda = -105 \pm 17^\circ$. To summarize the results obtained for the entire catalog, we report in Figure 4.48 the rose diagram for strike, dip and rake angles for the preferred solutions.

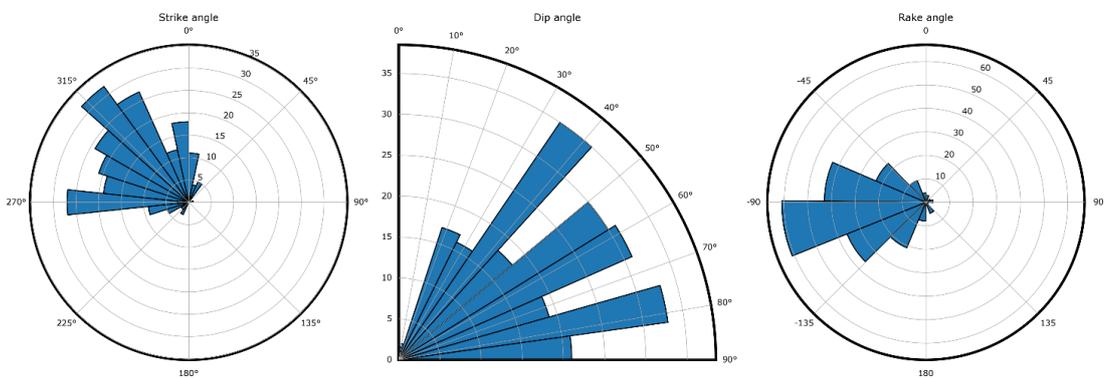


Figure 4.48) Rose diagram reporting the strike (left panel), dip (central panel) and rake (right panel) angles resulting from the 223 extracted focal mechanisms. These histograms support the observations of a main normal faulting environment, with a slight component of strike-slip.

The histograms reported in Figure 4.48 support the observations of a main normal faulting environment, with a slight component of strike-slip. When extracting the average and standard deviation values, we obtain $\Phi = 310 \pm 35^\circ$, $\delta = 55 \pm 20^\circ$, $\lambda = -84 \pm 35^\circ$, which are also coherent with the main angles estimated for the main events of the M 6.9 1980 Irpinia earthquakes (Bernard & Zollo, 1989). In Figure 4.49, we report the distribution of the strike, dip and rake angle uncertainties for the obtained focal mechanisms. The median uncertainties on the strike, dip and rake angles resulted to be 10° , 8° and 15° , respectively.

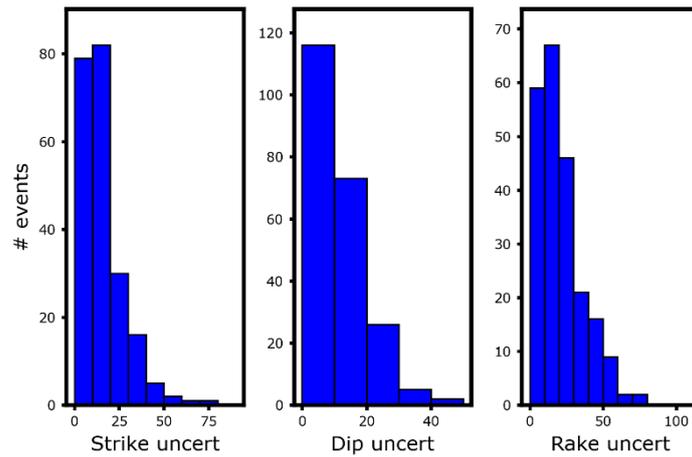


Figure 4.49) Strike (left panel), dip (central panel) and rake (right panel) uncertainties for the 223 estimated focal mechanisms.

Despite scarcely populating the catalog of focal mechanisms, some constrained earthquakes featured inverse faulting. In Figure 4.50, we report the distribution of the focal mechanisms for the events characterized by inverse faulting solutions, marking with a red star the corresponding epicentres.

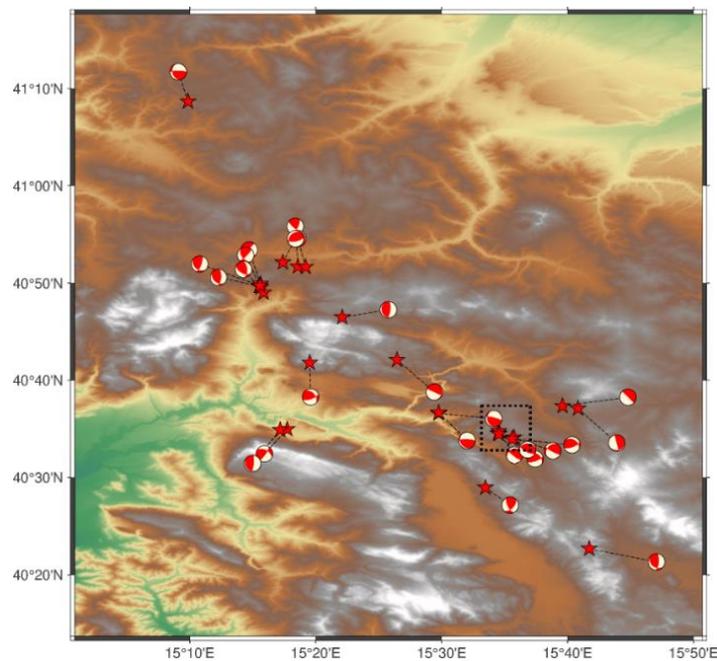


Figure 4.50) Representation of focal mechanisms featuring an inverse faulting. Red stars mark the corresponding epicentres.

Integrating Figure 4.50 with an analysis of the occurrence time of the earthquakes, it appears that most of the events featuring inverse faulting occurred clustered in space and time, with interevent times of minutes (cluster in the northern Sector), hours (couple of events in the central-western sector) or one day (cluster in the southern sector). Moreover, events within the cluster typically share a similar focal mechanism. In Figure 4.51, we show the focal mechanism for three $M_l > 1$ earthquakes belonging to the cluster in the Southern Sector of the area, marked with a black box in Figure 4.50.

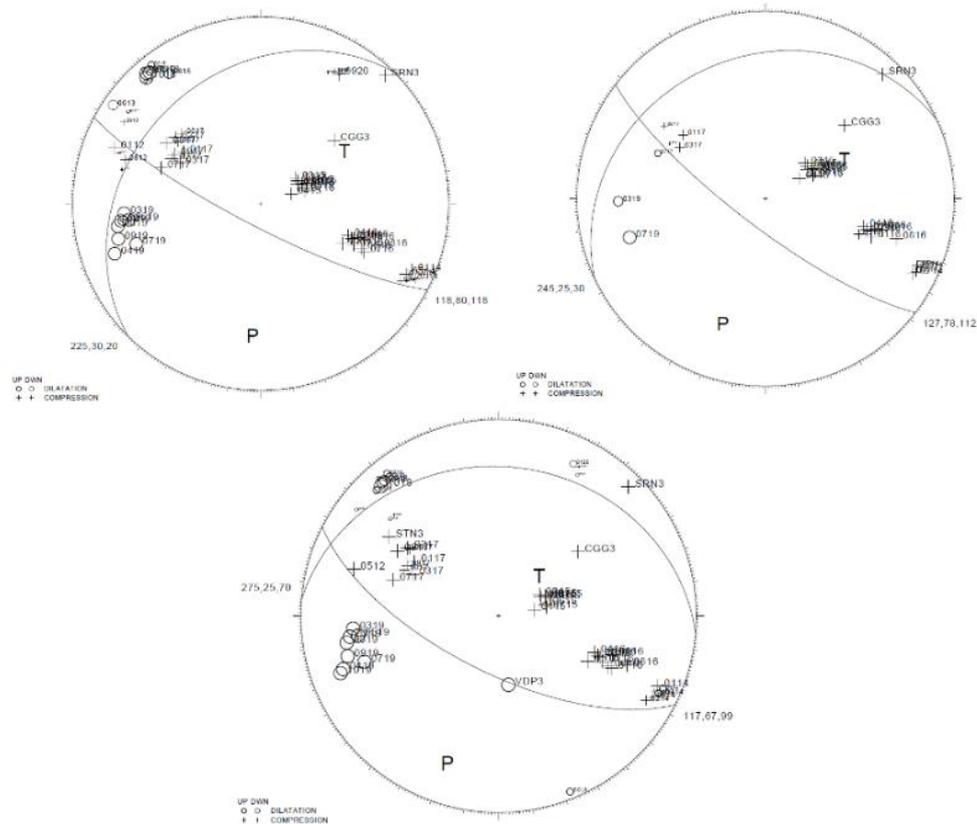


Figure 4.51 Focal mechanism for three clustered earthquakes that occurred within a span of 20 days in the Southern Sector of the region (black box in Figure 4.50). The estimated inverse focal mechanisms appear similar for the considered earthquakes.

The earthquakes in Figure 4.51 occurred between April 30th and May 23rd 2022, featuring close epicentral distances and depths ranging between 6.8 and 7.5 km. All the three focal mechanisms were estimated using a high number of P wave polarities, consisting of 69, 61 and 32 automatic readings of the sign of first motion onset. The closeness and the high number of polarity readings for these events, as for the other obtained inverse focal mechanisms, increase the reliability of the considered focal mechanisms, that were

therefore included in the final catalog. For further verifying the correctness of the estimated inverse fault planes, we compared the results from the fully automatic procedure (deep learning estimation of the phase arrival times, earthquake location and deep learning evaluation of the P wave polarities) with the corresponding manual estimation, by visually identifying the phase arrival times and P wave polarities for some of the aforementioned inverse faulting earthquakes. In Figure 4.52, we compared the focal mechanism obtained for an event reported in Figure 4.51 with the respective one obtained with manual estimation of phase arrival times and P wave polarities. We did not observe any significant differences in the earthquake location, focal depth and estimated focal mechanisms between the fully automatic and fully manual procedure, with the two estimations being compatible within the angle uncertainties. This result further supports the reliability of the obtained fault plane solutions. In the following Section 4.4, we will use the estimated focal mechanisms for evaluating the stress field in the area.

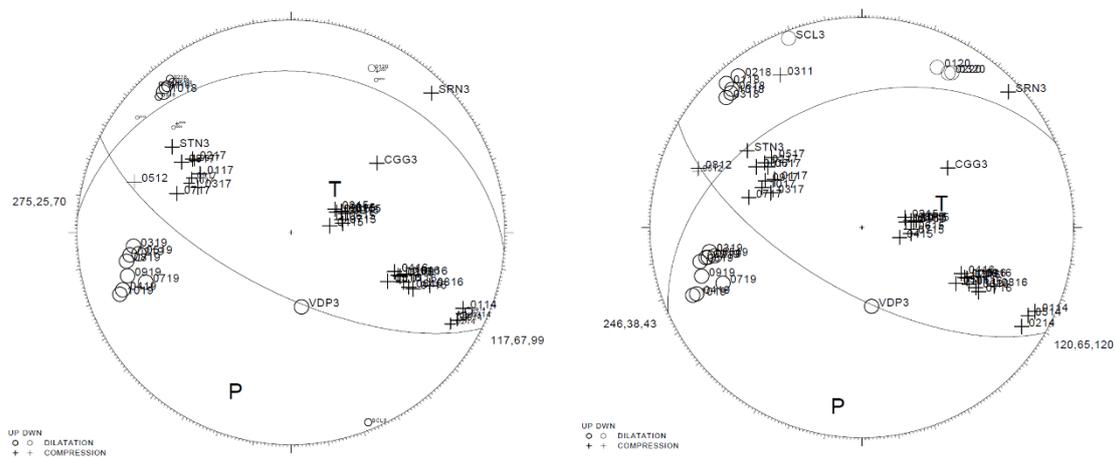


Figure 4.52) Comparison of the estimated focal mechanisms for an earthquake featuring inverse faulting resulting from the fully automatic procedure (left panel) and manual identification of phase arrival times and P wave polarities (right panel).

4.3.4 Stress field inversion from focal mechanisms using dense array deployments

In Section 4.3.3, we produced an enhanced catalog of 223 focal mechanisms from the ISNet manual catalog, exploiting the dense monitoring of the seismic arrays. In the area of the Southern Apennines, few studies have evaluated the stress field from the inversion of focal mechanisms for micro-earthquakes (Pasquale et al., 2009; De Matteis et al., 2012). These mentioned analyses either refer to the aftershocks of the 1980 M 6.9 Irpinia earthquake (Pasquale et al., 2009) or considered earthquakes spanning a time window of several years (De Matteis et al., 2012). However, both studies used a lower number of

focal mechanisms as compared to the results of Section 4.3.3, which provided a twice larger list of fault plane solutions. Using a dataset of 1312 microearthquakes between 2005 and 2011, which resulted into 118 focal mechanisms, De Matteis et al. (2012) obtained a regional stress field characterized by a nearly horizontal northeast–southwest minimum compressive stress axis (σ_3) and a nearly vertical maximum compressive stress axis (σ_1), indicating the presence of a unique normal-faulting regime (left panel of Figure 4.53). These results are compatible with the existence of extension perpendicular to the axis of the belt, which is well known in the study area (Montone et al., 2004). Pasquale et al. (2009), focused on the stress inversion for the analysis of the 2352 aftershocks within 4 months after the M 6.9 Irpinia earthquake, which was performed by inverting 139 focal mechanisms. By inverting all the selected focal mechanisms, the stress field still showed a nearly horizontal NE–SW minimum compressive stress axis (σ_3) and a maximum compressive stress axis (σ_1) that is nearly vertical. Although the axis σ_3 is horizontal, with a small error on its plunge angle, there is a large uncertainty for its trend, while axis σ_1 shows a large error both for trend and plunge.

We therefore inverted the catalog of 223 focal mechanisms obtained in Section 4.3.3, using the approach proposed by Martinez-Garzon et al. (2014), selecting 2000 bootstrap resamples. The retrieved stress field is reported in the right panel Figure 4.53, where each dot represents the retrieved stress component in individual bootstrap instances.

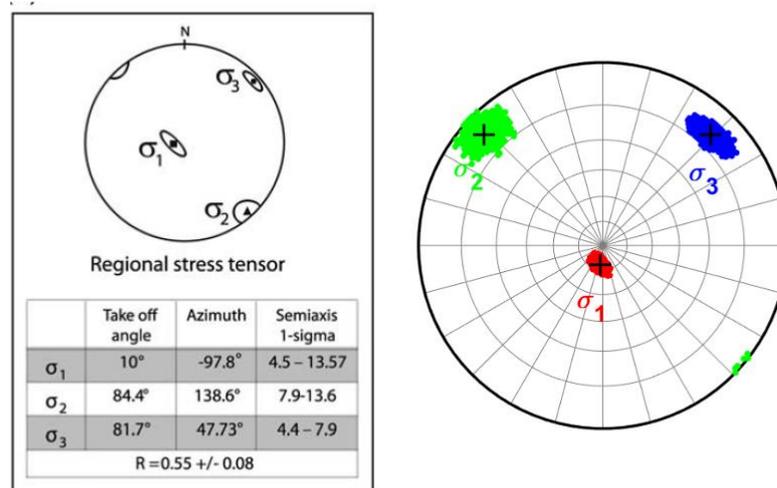


Figure 4.53) Left panel: Stress field inversion from De Matteis et al. (2012). Right panel: Stress field retrieved from the inversion of the 223 focal mechanisms estimated in Section 4.3.3.

We observe coherent results with De Matteis et al. (2012), in terms of the azimuth and plunge angles for the stress components. In Table 4.3, we report the best values, along

with the 95% confidence intervals for the mentioned angles as retrieved by our stress field inversion, compared with the results from De Matteis et al. (2012), reported in red.

Table 4.3) Plunge and azimuth angles for σ_1 , σ_2 , and σ_3 obtained from the inversion of focal mechanisms evaluated in Section 4.3.3. Red values refer to the results of De Matteis et al. (2012). Bracketed values refer to the 95% confidence interval from bootstrap resampling.

	PLUNGE	AZIMUTH
σ_1	78.2°/80° [70.8, 84.2]°	-172.2°/ -97.8° [-198.2, -129.7]°
σ_2	6.8°/5.6° [0.2, 15.9]°	-47.2°/138.6° [-57.5, -36.3]°
σ_3	9.6°/8.3° [4.7, 14.9]°	43.9°/47.7° [34.1, 54.7]°

We note compatibility between the plunge and azimuth angles with De Matteis et al. (2012), with the former values falling within the confidence interval. While we observe coherence among the azimuth of σ_3 , pointing in the direction of the anti-Apenninic extensional field, we note detaching values for the azimuth of σ_1 and σ_2 . Nevertheless, the azimuth of σ_2 appear to be the supplementary as compared to the σ_2 value from De Matteis et al. (2012) and, given the wrapping of the azimuth angles at the edge of the polar representation, these values are actually closer than they appear. On the other hand, we observe a major difference for the azimuth of σ_1 among the two estimates. However, given the high plunge angle of σ_1 , slight variations in the position on the polar plot might result in strong differences in the azimuth angles. Indeed, despite the narrower distribution of the σ_1 azimuth angles (red dots in the left panel of Figure 4.53) as compared to the respective ones of σ_2 and σ_3 , the confidence interval associated to the maximum compressive stress results to be the largest. Moreover, comparing the extension of the confidence interval for σ_1 in Figure 4.53, we can note a wider distribution of the σ_1 azimuth even for the results of De Matteis et al. (2012), which did not explicitly report the confidence interval for the compressive stresses. Therefore, considering the confidence interval also for the results in De Matteis et al. (2012) might lead to a coherency among the estimated stress fields. When focusing on the stress ratio R , defined as $R = (\sigma_1 - \sigma_2)/(\sigma_1 - \sigma_3)$ and representing the relative magnitude of the compressive stress, we observed a coherent value ($R = 0.55$, confidence interval [0.39, 0.71]) as compared to the results of De Matteis et al. (2012), who estimated $R = 0.55 \pm 0.08$. In Figure 4.54, we report the distribution of the R value from each bootstrap resample.

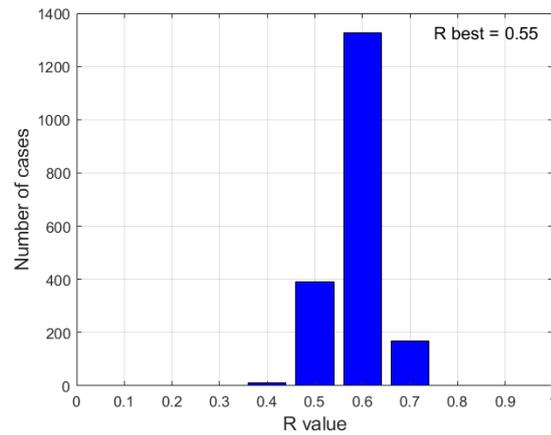


Figure 4.54) Distribution of the R value obtained from each bootstrap instance using the catalog of 223 focal mechanisms.

Given the high number of available focal mechanisms and the spatial distribution of the earthquakes, we performed the stress field inversion for the events occurring in the Northern, Central and Southern sector, respectively, seeking for eventual differences within the stress tensor. When splitting the dataset for the three regions, we got 67, 94 and 62 earthquakes for the stress field inversion in the Northern, Central and Southern sector, respectively. In Figure 4.55, we report the spatial distribution of the considered earthquakes.

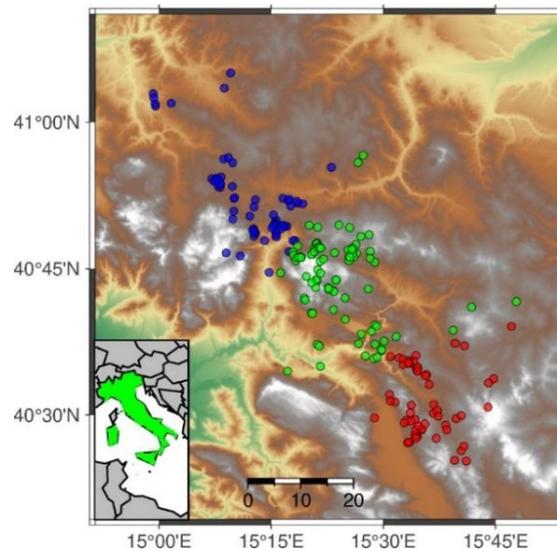


Figure 4.55) Dataset splitting for the inversion of the stress field in the Northern (blue), Central (green) and Southern sector.

In Figure 4.56 we report the stress field retrieved from the inversion of focal mechanisms split as in Figure 4.55, along with the preferred value and 95 % confidence interval of each stress component.

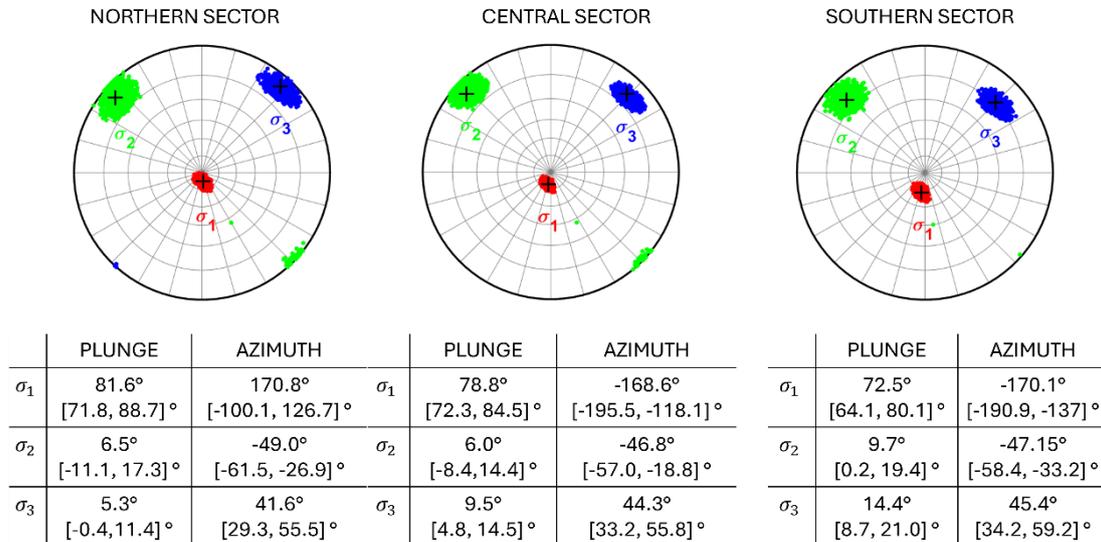


Figure 4.56) Stress field inversion from focal mechanisms belonging to the Northern (left panel), Central (central panel) and Southern (right panel) sector.

When considering the azimuth and plunge angles for the different sectors, we note coherent results among the quantities. Only the plunge angles of σ_3 slightly vary, decreasing moving from the Northern to the Southern sector. However, the obtained values for the considered quantity are still compatible within the 95% confidence interval and the corresponding preferred azimuth remarkably follows the anti-Apeninic direction, as for the inversion of the global dataset. Therefore, this result supports the hypothesis of a quite homogenous stress field in the Irpinia region. Further evidence of a homogenous stress field is provided by the distribution of the R ratio for the different sectors, shown in Figure 4.57. We still observe a homogeneous stress field when considering the earthquakes during the charge/discharge of the karst aquifer, which has been proven to modulate the occurrence of shallow seismicity (D'Agostino et al., 2018; Tarantino et al., 2024). In Figure 4.58, we report the stereonet representation for the events that occurred before (77, left panel) and after (143, right panel) February 1st 2022.

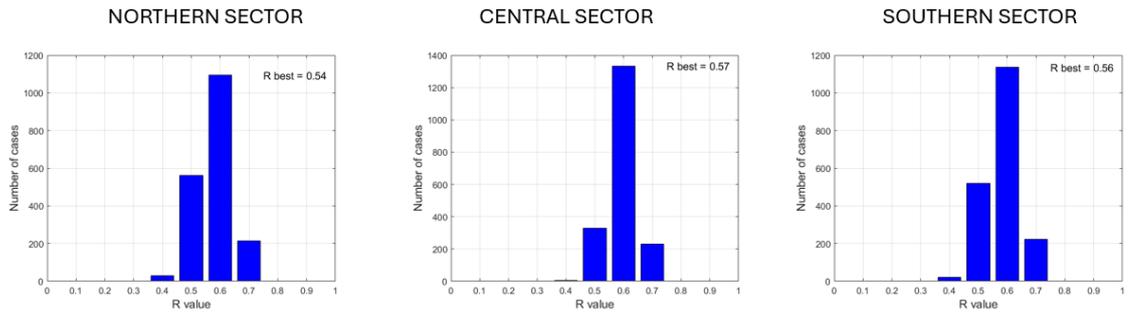


Figure 4.57) Histogram of the R ratio from bootstrapping resample of the focal mechanisms for the Northern (left panel), Central (central panel) and Southern (right panel) sector.

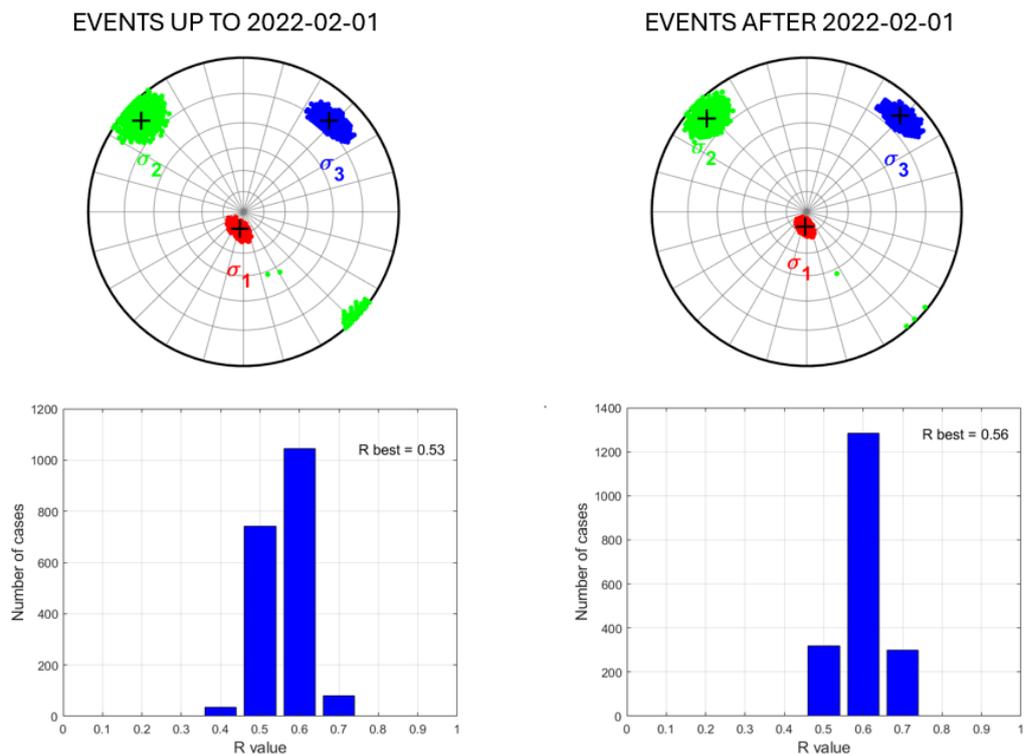


Figure 4.58) Stress field inversion for the events occurred before (left panel) and after (right panel) February 1st 2022.

The use of temporary dense array deployment can provide important insights into the analysis of the stress field from the focal mechanism inversion. Even when focusing on the earthquakes already contained in the manual catalogs, the high number of estimated focal mechanisms allows for the inversion of a constrained set of fault planes, which can

be much larger than the catalogs extracted from multiple years of microseismic monitoring using ordinary seismic networks. Moreover, the spatio-temporal distribution of the earthquakes recorded by the arrays allows the investigation of local features in the stress field, which could have not been analyzed before due to the scarcity of available fault planes. In particular, in this analysis, we showed that the stress field acting in the Irpinia region is, at a global scale, characterized by a near-vertical compression stress σ_1 and horizontal σ_2 and σ_3 , whose latter azimuth agrees with the direction of the extensional regional field pointing towards the anti-Apenninic direction. When looking for spatial variations of the stress field, we note a coherency of the azimuth and plunge angles, along with compatibility of the relative magnitude of the stress components, between the northern, central and southern sectors of the Irpinia region, suggesting the homogeneity of the acting stress field. A similar coherency is also observed when inspecting temporal variations within the characteristic time scale of the karst aquifer cycle (year).

Chapter 5

Conclusions

In this thesis, we characterized microseismic events for achieving insights in the earthquake generation processes. Using advanced detection techniques, such as deep learning models and similarity-based algorithms, we generated enhanced seismic catalogs as compared to the existing manual and standard earthquake lists, proposing the integration of the deep learning detector EQTransformer (Mousavi et al., 2020) and the template matching algorithm EQCorrscan, with the former techniques used to provide a wider set of templates to be used for the similarity-based detection, as best detection strategy. We showed that the number of events occurring within seismic sequences in the Southern Apennines is effectively one order of magnitude higher than the content of the former catalogs (Scotto di Uccio et al., 2023). The improved catalogs can now be investigated through robust statistical analyses, which can reveal the parameters of the Gutenberg-Richter relationship regulating the magnitude-occurrence distribution of earthquakes. In particular, for the considered seismic sequences in the Irpinia region, we observed a lower *b-value* as compared to the analogous parameter extracted for the seismic catalog of the past 15 years in the region, which is mainly populated by background seismicity. Since this parameter 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, likely rupturing compact, sub-kilometric size asperities. With the twofold aim of improving the spatial resolution and lowering the minimum magnitude for detecting earthquakes, 200 seismic stations organized in 20 dense arrays were deployed in the Southern Apennines in the framework of the DETECT experiment, integrating the existing seismic network from September 2021 to August 2022. The proposed detection strategy successfully led to an increase of a factor ~ 8 in the number of detected events as compared to the manual catalog built on the visual inspection of the waveforms recorded by the ordinary seismic network, leading to hundreds of identifications of phase arrival times even for $M_l < 2$ earthquakes. Although no major earthquakes occurred during the deployment of the seismic arrays, with an observed maximum magnitude M_l 2.8, we detected almost 3600 earthquakes, an amount comparable with the number of earthquakes contained in the manual catalog of the ordinary seismic network for 15 years of monitoring. When

analyzing the seismic catalog resulting from DETECT, which mainly accounts for background seismic events, we retrieved a compatible *b-value* with the existing manual catalog, suggesting that the integration of advanced detection strategies and dense short-time monitoring can lead to the same level of statistical information as for multiple years of ordinary seismic monitoring. Within this thesis, automatic detection strategies have been shown to provide improved seismic catalogs on extended time scales even in monitoring the induced seismicity in gas storage facilities (Scala et al., 2022) and searching for potential near-repeating earthquakes (Palo et al., 2023b), opening the way for implementing automatic monitoring systems for earthquake identification. In this direction, we built up a fully automatic detection workflow for supporting the daily (and manual) monitoring activities in Southern Apennines, where visual earthquake detection is performed by network operators. In order to limit the subjective bias and to lower the detection threshold, we integrated the deep learning detector PhaseNet (Zhu & Beroza, 2018) and the phase associator REAL (Zhang et al., 2019), supporting the manual activities providing in advance a detection list along with picked waveforms. We demonstrated the robustness of the implemented detection strategy, which provided an enhanced catalog of a factor ~ 3 , as compared to the manual catalogs of the former years within the same 6 months of operativity. This improvement resulted in a lower magnitude of completeness of the seismic catalog of 0.5 magnitude units. The low number of false detections declared by the workflow, which resulted in less than one declaration per day on average, and the even lower number of missed earthquakes (18), supports the possibility of implementing the proposed detection strategy as a standard monitoring system for seismic networks. Indeed, the comparison between the automatic and manual identification of phase arrival times promisingly reports consistency among the quantities, featuring zero-mean residuals for both P and S picks, being therefore accurate enough for performance automatic earthquake location.

Within this thesis, we further discussed the improvement of the knowledge of seismogenic sources that might be achieved by integrating advanced detection techniques and accurate location algorithms. Exploiting the enhanced catalogs obtained for the seismic sequences in Southern Apennines, we obtained accurate relative hypocenter locations using HYPODD (Waldhauser & Ellsworth, 2000) selecting catalog and cross correlation differential travel times. The analysis revealed small-scale alignment of the seismicity along subparallel segments featuring consistent orientations among the main structures generating the 1980 M 6.9 Irpinia earthquake (Scotto di Uccio et al., 2024a), allowing to finely track the spatio-temporal evolution of the seismic sequences. Although only a limited percentage of the events in the enhanced catalogs can be relocated (around 30 % of the starting detection lists, that is still twice larger than the number of events in the manual catalogs obtained from absolute location techniques), the amount of differential travel times that can be used for relocating the earthquakes might be orders

of magnitude higher compared to the same quantity extracted from standard catalogs, strongly constraining the hypocenter determination. When projecting the relocated hypocenters along the fault planes related to individual seismic sequences, an important revealed feature is that the distribution of the events is not isotropic around the main events of the sequences, but small events tend to align dominantly along the dip direction, which also corresponds to the slip direction, for normal faults and might be interpreted as the boundary between locked and creeping domains (Rubinstein & Beroza, 2007). When considering the seismicity that occurred within the dense array deployment, the hypocenter relocation of the enhanced catalogs reveals similar characteristics. However, the fraction of relocated events is almost twice as large as compared the respective observed for the earthquakes in the seismic sequence and recorded by the ordinary seismic network. This result suggests that a dense monitoring system can provide enough observation for better constraining the hypocenter determination of microearthquakes. Analyzing the spatiotemporal distribution of the relocated hypocenters, we observed differences among the shallow (depth < 6km) and deep seismicity. In particular, shallow earthquakes appear sparse, without clearly illuminating seismicity clusters. This feature has been associated with the effects of karst aquifers, which have been proven to modulate the occurrence of shallow seismicity in the area (D'Agostino et al., 2018, Tarantino et al., 2024). On the other hand, deeper earthquakes are typically clustered in space and time, illuminating seismic alignments along small-scale structures, similarly to the clusters highlighted by the relocation of earthquakes in the enhanced catalogs of seismic sequences. As depicted by the double difference location of earthquakes in the DETECT catalog, seismicity varies in depth and location in a SE-NW direction. To the SE, seismicity extends well within the Apulian carbonate platform along and around a previously identified, SE-dipping, long-lived and multiply reactivated major fault (Amoroso et al., 2014; Amoroso et al., 2005). However, seismicity shallows NW-ward, and at the same time it steps to the right, before deepening again further north. This is interpreted as a result of fault segmentation, which involves the occurrence of a major dextral fault step to the NW. Advanced location techniques can provide significant insight into the seismogenic structures even in volcanic environments, as results from our analysis in the Campi Flegrei caldera. Despite leveraging on the hypocenter characterization for the manual provided by the INGV-OV from 2014 to 2023, the use of an innovative location algorithm, which develops source-specific station travel-time corrections and exploits waveform similarity (NLL-SC, Lomax & Savvaidis, 2022), achieves improved resolution in the seismic source location, with lower uncertainties as compared to standard techniques. The high-precision NLL-SC locations help in unraveling the spatiotemporal evolution of the seismicity, delineating several clusters and alignments of seismicity produced during the ongoing unrest at Campi Flegrei. Most of the seismicity concentrates in the shallow region around the Solfatara- Pisciarelli area. Here, epicenters define a $\sim 1 \text{ km} \times 1 \text{ km}$, horseshoe-shaped structure, opened and

deepening toward the northeast beneath the Agnano Plain, and slightly larger than the ~ 0.5 km diameter of the Solfatara crater. Smaller-scale shallow seismicity clusters, with a typical size of 100–300 m, occur south and southwest of the Solfatara, along the coast of Pozzuoli. The most recent magnitude $M_d 3.6+$ events except for the largest magnitude $M_d 4.2$, also occurred in the Solfatara-Pisciarelli area, beneath the horseshoe-shaped seismicity, at depths between 2 and 3 km, while an approximately E–W trending fault bounds the distribution of the relocated seismicity NE of the Solfatara crater on which a series of spatially and temporally correlated seismicity bursts occurred between 2 and 3 km depths. A portion of the relocated seismicity, occurring in the offshore sector, is compatible with the caldera ring fault zone and appears related to the innermost portion of the composite caldera ring fault zone. Overall, only the deepest offshore seismicity, between 3 and 4 km depth, appears to fit and approximate the downward propagation of the south–western inner ring fault. A novelty revealed by our analysis is the \sim N–S trending sub-vertical fault structure just offshore La Pietra, which generated the largest magnitude ($M_d 4.2$). This structure was not identified to date, since it lies in a region where no deep-penetrating reflection profiles are available and there was no particular seismicity during the 1982–84 crisis.

Characterizing the properties of the seismic source can provide insights into the earthquake size, geometrical characteristics and mechanical interactions of the events. In the analysis of the source size, we focused on the determination of the seismic moment and the corner frequency, which are directly related to the magnitude and the source size, respectively. The proposed approach, leveraging on the use of low-magnitude events for determining the parameters to describe the attenuation processes from the source to the receivers, allowed to retrieve the seismic moment (and thus the moment magnitude) for 60 % of the relocated events. We recognize two trends between the magnitude scales: for $M_l < 2$, the M_l - M_w distribution strongly deviates from the 1:1 scaling relation which is observed for higher magnitude. We attribute this feature to the saturation of the event duration in the local magnitude computation due to the anelastic attenuation. On the other hand, we retrieved the corner frequency (and thus the source radius) for the events with $M > 2$. We demonstrated with synthetic tests that at the characteristic hypocentral distances for the seismic sequences in the Southern Apennines, the spectral content at higher frequencies than the low-pass cutoff frequencies of the Earth Green's function, falling within 10 and 20 Hz, can be severely affected and biased using an incorrect value of the quality factor. Since f_c for microseismic events typically feature values higher than the [10 – 20] Hz range, the estimation of the corner frequency might be biased for events in this magnitude range. For the resolved events, the distribution of the corner frequencies with moment appears to follow a nearly linear trend, with stress drops ranging between 1-3 MPa, while we observed median stress drops of $\Delta\sigma = 1.8$ MPa in the Northern sector and $\Delta\sigma = 3.4$ MPa in the Southern Sector. This difference might be associated with the

different fluid content and fraction. In the Southern sector rock physics modelling indicates the presence of a mixture brine-CO₂ (Amoroso et al. 2017). Assuming self-similarity for low-magnitude events for each sequence, we evaluated the static stress release model within the swarm, attributing a source radius to the events for which we did not estimate the corner frequency using the median stress drop. For almost all the sequences the stress model suggests static stress release as a trigger mechanism, with small events mainly concentrated in or around the area affected by stress changes due to the main events in the sequence, identifying a single km-sized patch mainly oriented along the dip direction. An interesting case is represented by the Rocca San Felice seismic sequence, for which we observe two seismicity patterns oriented along the dip direction activated at different times (the main event M_l 3.0 involved the first patch, and the seismicity migrated along the second segment almost two days after the mainshock with the occurrence of a M_l 2.8 earthquake). To infer the occurrence of an aseismic event we investigated the geodetic data from GNSS stations deployed close to the seismic sequence. Although no clear offsets were observed on the data, a resolution test revealed that an aseismic event with M 5.0, whose size could be comparable to the spatial shift of the seismicity, might have been buried within the noise of the surface GNSS stations.

For assessing the geometrical properties of the seismic source, we estimated the focal mechanism for the events in the manual catalog integrating the ordinary and dense DETECT network, the latter characterizing low-to-moderate magnitude earthquakes with hundreds of P wave arrival times. For evaluating the P wave polarities, whose inversion can reveal the strike, dip and rake angles describing the orientation of the generating fault, we used a convolutional neural network (Ross et al., 2018) trained on millions of manually labeled polarities. The inversion of the retrieved polarities allowed the determination of 223 focal mechanisms, a number twice larger as compared to former studies in the area, which either grounded on manual seismic catalog for the aftershocks of the 1980 M 6.9 earthquake (Pasquale et al., 2009) or on 6 years of microseismic events (Matrullo et al., 2012). Moreover, the percentage of resolved events is more than 5 times higher than the analogous value in the mentioned studies, suggesting that dense short-time array monitoring can provide resolution for constraining the focal mechanisms of low-magnitude events. The analysis of the retrieved fault planes revealed a prevalent normal faulting environment, coherently with the main faults in the area, with a slight component of strike-slip and inverse faulting. Using the obtained focal mechanisms, we inverted the fault planes for retrieving the orientation of the stress field adopting the methodology proposed by Martinez-Garzon et al. (2014), which implemented a bootstrapping approach for estimating the confidence interval for the stress tensor components. When inverting the entire catalog of fault planes, we retrieved near-to-vertical most compressive stress σ_1 and horizontal intermediate and least compressive stress σ_2 and σ_3 , these orientations compatible with the stress associated with normal

faulting. Interestingly, the azimuth of the least compressive stress σ_3 lied in the anti-Apenninic direction, compatibly with the direction of the regional extensional stress field acting in the area. Moreover, we observe agreement with the results obtained by De Matteis et al. (2012), both in terms of azimuth and plunge angles of the deviatoric stress components and also in terms of the relative stress magnitude. Therefore, dense and temporary surveys can provide similar information on the stress field as compared to multiple years of ordinary seismic monitoring. The enhanced catalog of focal mechanisms and their spatiotemporal distribution allowed us to monitor eventual local variations in the stress field. We split the dataset of focal mechanisms selecting events in the northern, central and southern sectors of the Irpinia region, individually inverting for the stress field. We note a strong homogeneity within the retrieved stress field, with slight variations only in the plunge angles of σ_3 , which increase moving from the Northern to the Southern sector. However, the obtained values for the considered quantity are still compatible within the 95% confidence interval and the corresponding preferred azimuth remarkably follows the anti-Apenninic direction. The same level of homogeneity is observed when splitting the dataset according to the charge/discharge trend period of the karst aquifer, which is proven to be responsible for the modulation of the shallow seismicity (D'Agostino et al., 2018), supporting the hypothesis of a quasi-homogenous stress field acting in the Irpinia region.

Within this work, we individually tackled the characterization of the microseismicity in terms of earthquake detection, hypocenter location and seismic source properties. The in-depth analyses we implemented led to the creation of ready-to-use blocks, that at this stage are not interoperable yet. An interesting outlet for this thesis could be the creation of automatized workflows for characterizing the ongoing seismicity, moving towards a near-real-time monitoring. This direction could lead to the definition of standardized approaches to be applied in different seismogenic environments, providing robust and homogenous catalogs. We made a first step in this direction, building an automatic and reliable system for providing detection and phase arrival times in the Southern Apennines grounded on advanced state-of-the-art techniques, whose promising performance might suggest a stand-alone implementation soon. The integration of blocks for source parameters and automatic P wave polarity estimation blocks could generate a first prototypal workflow for a comprehensive analysis of the microseismicity. An opening question in the framework of earthquake detection using deep learning techniques concerns the sensitivity of models trained on worldwide datasets to local seismic features. The existence of manual labels in long-term catalogs might support the training of models for earthquake detection and phase picking entirely leveraging local seismic data, which we believe could be more robust in terms of minimization of false detections. Cross-domain analysis and tests against the performance of consolidated worldwide models can

answer the former question, suggesting the most effective model for the detection analysis in a specific region.

Interestingly, enhanced catalogs might indicate the existence of creeping domains, which might generate slow events over time. Recent observations in normal faulting regimes show that slow events up to magnitude 5.5 can occur also in these tectonic contexts (Northern Apennines, Gualandi et al., 2017; Corinth rift, Kaviris et al., 2021). However, our observational capability seems to be inadequate to detect those events in Southern Apennines, requiring further technological development, with the presence of borehole instrumentation and strainmeters.



Bibliography

Abercrombie, R., & Leary, P. (1993). Source parameters of small earthquakes recorded at 2.5 km depth, Cajon Pass, southern California: implications for earthquake scaling. *Geophysical Research Letters*, 20(14), 1511-1514.

Abercrombie, R. E. (1995). Earthquake source scaling relationships from 1 to 5 ML using seismograms recorded at 2.5-km depth. *Journal of Geophysical Research*, 100(B12), 24015–24036.

Aki, K. (1967). Scaling law of seismic spectrum. *Journal of geophysical research*, 72(4), 1217-1231.

Aki, K., & Richards, P. G. (2002). *Quantitative seismology*.

Al-Ismaïl, F., Ellsworth, W. L., & Beroza, G. C. (2023). A Time-Domain Approach for Accurate Spectral Source Estimation with Application to Ridgecrest, California, Earthquakes. *Bulletin of the Seismological Society of America*, 113(3), 1091-1101.

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

Amoroso, O., Ascione, A., Mazzoli, S., Virieux, J., & Zollo, A. (2014). Seismic imaging of a fluid storage in the actively extending Apennine mountain belt, southern Italy. *Geophysical Research Letters*, 41(11), 3802-3809.

Amoroso, O., Russo, G., De Landro, G., Zollo, A., Garambois, S., Mazzoli, S., ... & Virieux, J. (2017). From velocity and attenuation tomography to rock physical modeling: Inferences on fluid-driven earthquake processes at the Irpinia fault system in southern Italy. *Geophysical Research Letters*, 44(13), 6752-6760.

Amoruso, A., Crescentini, L., D'Anastasio, E., & De Martini, P. M. (2005). Clues of postseismic relaxation for the 1915 Fucino earthquake (central Italy) from modeling of leveling data. *Geophysical Research Letters*, 32(22).

Amoruso, A., Crescentini, L., Di Lieto, B., & Scarpa, R. (2011). Faulting mechanism of the Campania–Lucania 1980 earthquake, Italy, from high-resolution, 3D velocity structure, aftershock relocation, fault-plane solutions, and post-seismic deformation modeling. *Annals of Geophysics*, 54(6).

- Anderson, E. M. (1951). The dynamics of faulting and dyke formation with applications to Britain.
- Anderson, J. G., & Hough, S. E. (1984). A model for the shape of the Fourier amplitude spectrum of acceleration at high frequencies. *Bulletin of the Seismological Society of America*, 74(5), 1969-1993.
- Andrews, D. J. (1980). A stochastic fault model: 1. Static case. *Journal of Geophysical Research: Solid Earth*, 85(B7), 3867-3877.
- Antonioli, A., Piccinini, D., Chiaraluce, L., & Cocco, M. (2005). Fluid flow and seismicity pattern: Evidence from the 1997 Umbria-Marche (central Italy) seismic sequence. *Geophysical Research Letters*, 32(10).
- Ascione, A., Mazzoli, S., Petrosino, P., & Valente, E. (2013). A decoupled kinematic model for active normal faults: Insights from the 1980, MS= 6.9 Irpinia earthquake, southern Italy. *Bulletin*, 125(7-8), 1239-1259.
- 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. *Bulletin of the Seismological Society of America*, 104(1), 394-409.
- Baraschino, R., Cito, Colombelli, S., De Landro, G., Elia, L., Festa, G., Iaccarino, A.I., Iervolino, I. Muzellec, T., Nazeri, S., Russo, G., Scala, A., Scotto di Uccio, F., Strumia, C., & Zollo, A. (2024) The Solfatara Earthquake 20/5/24 Md4.4. https://www.researchgate.net/publication/380824228_The_Solfatara_Earthquake_2052_4_Md44
- Battaglia, J., Zollo, A., Virieux, J., & Iacono, D. D. (2008). Merging active and passive data sets in travelttime tomography: the case study of Campi Flegrei caldera (Southern Italy). *Geophysical Prospecting*, 56(4), 555-573.
- Becker, D., McBrearty, I. W., Beroza, G. C., & Martínez-Garzón, P. (2024). Performance of AI-Based Phase Picking and Event Association Methods after the Large 2023 Mw 7.8 and 7.6 Türkiye Doublet. *Bulletin of the Seismological Society of America* 2024. <https://doi.org/10.1785/0120240017>
- Ben-Zion, Y., Vernon, F. L., Ozakin, Y., Zigone, D., Ross, Z. E., Meng, H., ... & Barklage, M. (2015). Basic data features and results from a spatially dense seismic array on the San Jacinto fault zone. *Geophysical Journal International*, 202(1), 370-380.
- Beroza, G. C., Segou, M., & Mousavi, S.M. (2021). Machine learning and earthquake forecasting—next steps. *Nature communications*, 12(1), 4761.

- Bernard, P., & Zollo, A. (1989). The Irpinia (Italy) 1980 earthquake: detailed analysis of a complex normal faulting. *Journal of Geophysical Research: Solid Earth*, 94(B2), 1631-1647.
- Bevilacqua, A., De Martino, P., Giudicepietro, F., Ricciolino, P., Patra, A., Pitman, E. B., ... & Neri, A. (2022). Data analysis of the unsteadily accelerating GPS and seismic records at Campi Flegrei caldera from 2000 to 2020. *Scientific reports*, 12(1), 19175.
- Bianco, F., Caliro, S., De Martino, P., Orazi, M., Ricco, C., Vilardo, G., ... & Castellano, M. (2022). The permanent monitoring system of the Campi Flegrei caldera, Italy. In *Campi Flegrei: A restless Caldera in a densely populated area* (pp. 219-237). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Bindi, D., Spallarossa, D., Picozzi, M., & Morasca, P. (2020). Reliability of source parameters for small events in central Italy: Insights from spectral decomposition analysis applied to both synthetic and real data. *Bulletin of the Seismological Society of America*, 110(6), 3139-3157.
- Bindi, D., Cotton, F., Picozzi, M., & Zollo, A. (2021). The Irpinia seismic array [Dataset]. GFZ Data Services. <https://doi.org/10.14470/MX7576871994>
- Boore, D. M., & Boatwright, J. (1984). Average body-wave radiation coefficients. *Bulletin of the Seismological Society of America*, 74(5), 1615-1621.
- Bragato, P. L., & Tonto, A. (2005). Local magnitude in northeastern Italy. *Bulletin of the Seismological Society of America*, 95(2), 579-591.
- Broder, A. Z., Charikar, M., Frieze, A. M., & Mitzenmacher, M. (1998). Min-wise independent permutations. In *Proceedings of the thirtieth annual ACM symposium on Theory of computing* (pp. 327-336).
- Brown, J. R., Beroza, G. C., & Shelly, D. R. (2008). An autocorrelation method to detect low frequency earthquakes within tremor. *Geophysical Research Letters*, 35(16).
- Brune, J. N. (1970). Tectonic stress and the spectra of seismic shear waves from earthquakes. *Journal of geophysical research*, 75(26), 4997-5009.
- Bobbio, A., Vassallo, M., & Festa, G. (2009). A local magnitude scale for southern Italy. *Bulletin of the Seismological Society of America*, 99(4), 2461-2470.
- Bonafede, M. (1991). Hot fluid migration: an efficient source of ground deformation: application to the 1982–1985 crisis at Campi Flegrei-Italy. *Journal of Volcanology and Geothermal Research*, 48(1-2), 187-198.

- Bott, M. H. P. (1959). The mechanics of oblique slip faulting. *Geological magazine*, 96(2), 109-117.
- Bouchon, M., Durand, V., Marsan, D., Karabulut, H., & Schmittbuhl, J. (2013). The long precursory phase of most large interplate earthquakes. *Nature geoscience*, 6(4), 299-302.
- Burridge, R., and G. S. Halliday. Dynamic shear cracks with friction as models for shallow focus earthquakes. *Geophysical Journal International* 25.1-3 (1971): 261-283.
- Cabrera, L., Poli, P., & Frank, W. B. (2022). Tracking the spatio-temporal evolution of foreshocks preceding the Mw 6.1 2009 L'Aquila earthquake. *Journal of Geophysical Research: Solid Earth*, 127(3), e2021JB023888.
- Cabrera, L., & Poli, P. (2023). A struggled rupture initiation of the Mw 6.1 2009 L'Aquila earthquake. *Geophysical Research Letters*, 50(6), e2022GL102337.
- Calò, M., & Tramelli, A. (2018). Anatomy of the Campi Flegrei caldera using enhanced seismic tomography models. *Scientific reports*, 8(1), 16254.
- Camanni, G., De Landro, G., Mazzoli, S., Michele, M., Muzellec, T., Schaff, D. P., ... & Zollo, A. (2024). Remobilization of Inverted Normal Faults Drives Active Extension in the Axial Zone of the Southern Apennines Mountain Belt (Italy). *Authorea Preprints*.
- Candela, T., Renard, F., Schmittbuhl, J., Bouchon, M., & Brodsky, E. E. (2011). Fault slip distribution and fault roughness. *Geophysical Journal International*, 187(2), 959-968.
- Chamberlain, C. J., Hopp, C. J., Boese, C. M., Warren-Smith, E., Chambers, D., Chu, S. X., ... & Townend, J. (2018). EQcorrscan: Repeating and near-repeating earthquake detection and analysis in Python. *Seismological Research Letters*, 89(1), 173-181.
- Chen, X., Shearer, P. M., & Abercrombie, R. E. (2012). Spatial migration of earthquakes within seismic clusters in Southern California: Evidence for fluid diffusion. *Journal of Geophysical Research: Solid Earth*, 117(B4).
- Chiarabba, C., De Gori, P., & Boschi, E. (2009). Pore-pressure migration along a normal-fault system resolved by time-repeated seismic tomography. *Geology*, 37(1), 67-70.
- Chiaraluce, L. (2012). Unravelling the complexity of Apenninic extensional fault systems: A review of the 2009 L'Aquila earthquake (Central Apennines, Italy). *Journal of Structural Geology*, 42, 2-18.
- Chiaraluce, L., Di Stefano, R., Tinti, E., Scognamiglio, L., Michele, M., Casarotti, E., ... & Marzorati, S. (2017). The 2016 central Italy seismic sequence: A first look at the mainshocks, aftershocks, and source models. *Seismological Research Letters*, 88(3), 757-771.

- Chiaraluca, L., Festa, G., Bernard, P., Caracausi, A., Carluccio, I., Clinton, J. F., ... & Sokos, E. (2022a). The Near Fault Observatory community in Europe: a new resource for faulting and hazard studies. *Annals of Geophysics*, 65(3), DM316.
- Chiaraluca, L., Michele, M., Waldhauser, F., Tan, Y. J., Herrmann, M., Spallarossa, D., ... & Segou, M. (2022b). A comprehensive suite of earthquake catalogs for the 2016-2017 Central Italy seismic sequence. *Scientific Data*, 9(1), 710.
- 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. *Geophysical Research Letters*, 37(11).
- Chiodini, G., Selva, J., Del Pezzo, E., Marsan, D., De Siena, L., D'auria, L., ... & Petrillo, Z. (2017). Clues on the origin of post-2000 earthquakes at Campi Flegrei caldera (Italy). *Scientific reports*, 7(1), 4472.
- Coccia, S., Lizeur, A., Bigarre, P., Contrucci, I., & Klein, E. (2013, September). Accurate 3D location of mine induced seismicity in complex near-field underground conditions. In 8. International Symposium on Rockbursts and Seismicity in Mines (RaSIM 8) (p. NC). Geophysical Survey of Russian Academy of Sciences. Obninsk.
- Cocco, M., Chiarabba, C., Di Bona, M., Selvaggi, G., Margheriti, L., Frepoli, A., ... & Campillo, M. (1999). The April 1996 Irpinia seismic sequence: evidence for fault interaction. *Journal of seismology*, 3, 105-117.
- Cole, H. M., Yeck, W. L., & Benz, H. M. (2023). MLAAPDE: A machine learning dataset for determining global earthquake source parameters. *Seismological Research Letters*, 94(5), 2489-2499.
- Colombelli, S., & Zollo, A. (2015). Fast determination of earthquake magnitude and fault extent from real-time P-wave recordings. *Geophysical Journal International*, 202(2), 1158-1163.
- 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. *Geophysical Research Letters*, 45(22), 12-253.
- D'Agostino, N., Métois, M., Koci, R., Duni, L., Kuka, N., Ganas, A., Georgiev, A., Jouanne, F., Kaludjerovic, N., & Kandić, R. (2020). Active crustal deformation and rotations in the southwestern Balkans from continuous GPS measurements. *Earth and Planetary Science Letters*, 539, 116246.
- D'Auria, L., Giudicepietro, F., Aquino, I., Borriello, G., Del Gaudio, C., Lo Bascio, D., ... & Ricco, C. (2011). Repeated fluid-transfer episodes as a mechanism for the recent

dynamics of Campi Flegrei caldera (1989–2010). *Journal of Geophysical Research: Solid Earth*, 116(B4).

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 the Monthly Notices of the Royal Astronomical Society*, 203(3), 1821-1831.

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. *Scientific Reports*, 12(1), 1247.

De Matteis, R., Matrullo, E., Rivera, L., Stabile, T. A., Pasquale, G., & Zollo, A. (2012). Fault delineation and regional stress direction from the analysis of background microseismicity in the southern Apennines, Italy. *Bulletin of the Seismological Society of America*, 102(4), 1899-1907.

De Natale, G., Troise, C., Pingue, F., Mastrolorenzo, G., Pappalardo, L., Battaglia, M., & Boschi, E. (2006). The Campi Flegrei caldera: unrest mechanisms and hazards. *Geological Society, London, Special Publications*, 269(1), 25-45.

De Santis, A., Cianchini, G., Favali, P., Beranzoli, L., & Boschi, E. (2011). The Gutenberg–Richter law and entropy of earthquakes: Two case studies in Central Italy. *Bulletin of the Seismological Society of America*, 101(3), 1386-1395.

De Siena, L., Amoroso, A., Petrosino, S., & Crescentini, L. (2024). Geophysical responses to an environmentally-boosted volcanic unrest. *Geophysical Research Letters*, 51(5), e2023GL104895.

Deichmann, N. (2017). Theoretical basis for the observed break in ML/M w scaling between small and large earthquakes. *Bulletin of the Seismological Society of America*, 107(2), 505–520

Del Gaudio, C., Aquino, I., Ricciardi, G. P., Ricco, C., & Scandone, R. (2010). Unrest episodes at Campi Flegrei: A reconstruction of vertical ground movements during 1905–2009. *Journal of Volcanology and Geothermal Research*, 195(1), 48-56.

Deschamps, A., & King, G. C. P. (1983). The Campania-Lucania (southern Italy) earthquake of 23 November 1980. *Earth and Planetary Science Letters*, 62(2), 296-304.

Di Luccio, F., Piscini, A., Pino, N. A., & Ventura, G. (2005). Reactivation of deep faults beneath Southern Apennines: evidence from the 1990–1991 Potenza seismic sequences. *Terra Nova*, 17(6), 586-590.

Ding, H., Zhou, Y., Ge, Z., Taymaz, T., Ghosh, A., Xu, H., ... & Song, X. (2023). High-resolution seismicity imaging and early aftershock migration of the 2023 Kahramanmaraş (SE Türkiye) MW7. 9 & 7.8 earthquake doublet. *Earthquake Science*, 36(6), 417-432.

DISS Working Group (2021). Database of Individual Seismogenic Sources (DISS), Version 3.3.0: A compilation of potential sources for earthquakes larger than M 5.5 in Italy and surrounding areas. Istituto Nazionale di Geofisica e Vulcanologia (INGV). <https://doi.org/10.13127/diss3.3.0>

Ekstrom, G. (1994). Teleseismic analysis of the 1990 and 1991 earthquakes near Potenza.

Ellsworth, W. L., & Beroza, G. C. (1995). Seismic evidence for an earthquake nucleation phase. *Science*, 268(5212), 851-855.

Ellsworth, W. L., & Beroza, G. C. (1998). Observation of the seismic nucleation phase in the Ridgecrest, California, earthquake sequence. *Geophysical research letters*, 25(3), 401-404.

Ellsworth, W. L., & Bulut, F. (2018). Nucleation of the 1999 Izmit earthquake by a triggered cascade of foreshocks. *Nature Geoscience*, 11(7), 531-535.

Festa, G., Adinolfi, G. M., Caruso, A., Colombelli, S., De Landro, G., Elia, L., ... & Zollo, A. (2021). Insights into mechanical properties of the 1980 Irpinia fault system from the analysis of a seismic sequence. *Geosciences*, 11(1), 28.

Frémont, M. J., & Malone, S. D. (1987). High precision relative locations of earthquakes at Mount St. Helens, Washington. *Journal of Geophysical Research: Solid Earth*, 92(B10), 10223-10236.

Gaeta, F. S., Peluso, F., Arienzo, I., Castagnolo, D., De Natale, G., Milano, G., ... & Mita, D. G. (2003). A physical appraisal of a new aspect of bradyseism: The miniuplifts. *Journal of Geophysical Research: Solid Earth*, 108(B8).

Geiger, L. (1912). Probability method for the determination of earthquake epicentres from the arrival time only. *Bull. St. Louis Univ.*, 8, 60.

Giudicepietro, F., Casu, F., Bonano, M., De Luca, C., De Martino, P., Di Traglia, F., ... & Lanari, R. (2024). First evidence of a geodetic anomaly in the Campi Flegrei caldera (Italy) ground deformation pattern revealed by DInSAR and GNSS measurements during the 2021–2023 escalating unrest phase. *International Journal of Applied Earth Observation and Geoinformation*, 132, 104060.

Goertz-Allmann, B. P., Goertz, A., & Wiemer, S. (2011). Stress drop variations of induced earthquakes at the Basel geothermal site. *Geophysical Research Letters*, 38(9).

- Grigoli, F., Cesca, S., Krieger, L., Kriegerowski, M., Gammaldi, S., Horalek, J., ... & Dahm, T. (2016). Automated microseismic event location using master-event waveform stacking. *Scientific reports*, 6(1), 25744.
- Gomberg, J. S., Shedlock, K. M., & Roecker, S. W. (1990). The effect of S-wave arrival times on the accuracy of hypocenter estimation. *Bulletin of the Seismological Society of America*, 80(6A), 1605-1628.
- Got, J. L., Fréchet, J., & Klein, F. W. (1994). Deep fault plane geometry inferred from multiplet relative relocation beneath the south flank of Kilauea. *Journal of Geophysical Research: Solid Earth*, 99(B8), 15375-15386.
- Gualandi, A., Nichele, C., Serpelloni, E., Chiaraluce, L., Anderlini, L., Latorre, D., ... & Avouac, J. P. (2017). Aseismic deformation associated with an earthquake swarm in the northern Apennines (Italy). *Geophysical Research Letters*, 44(15), 7706-7714.
- Gulia, L., & Wiemer, S. (2019). Real-time discrimination of earthquake foreshocks and aftershocks. *Nature*, 574(7777), 193-199.
- Gutenberg, B., & Richter, C. F. (1944). Frequency of earthquakes in California. *Bulletin of the Seismological society of America*, 34(4), 185-188.
- Hanks, T. C., & Kanamori, H. (1979). A moment magnitude scale. *Journal of Geophysical Research: Solid Earth*, 84(B5), 2348-2350.
- Hardebeck, J. L., & Michael, A. J. (2006). Damped regional-scale stress inversions: Methodology and examples for southern California and the Coalinga aftershock sequence. *Journal of Geophysical Research: Solid Earth*, 111(B11).
- Haskell, N. A. (1964). Total energy and energy spectral density of elastic wave radiation from propagating faults. *Bulletin of the Seismological Society of America*, 54(6A), 1811-1841.
- Hauksson, E., & Shearer, P. (2005). Southern California hypocenter relocation with waveform cross-correlation, part 1: Results using the double-difference method. *Bulletin of the Seismological Society of America*, 95(3), 896-903.
- Heidbach, O., Tingay, M., Barth, A., Reinecker, J., Kurfeß, D., & Müller, B. (2008). The world stress map database release 2008, doi: 10.1594/GFZ. WSM. Rel2008, 9.
- Helmstetter, A., & Sornette, D. (2003). Foreshocks explained by cascades of triggered seismicity. *Journal of Geophysical Research: Solid Earth*, 108(B10).

Herrmann, M., & Marzocchi, W. (2021). Inconsistencies and lurking pitfalls in the magnitude–frequency distribution of high-resolution earthquake catalogs. *Seismological Research Letters*, 92(2A), 909-922.

Husen, S., & Hardebeck, J. (2010). Earthquake location accuracy. Community Online Resource for Statistical Seismicity Analysis.

Iannaccone, G., Guardato, S., Donnarumma, G. P., De Martino, P., Dolce, M., Macedonio, G., ... & Beranzoli, L. (2018). Measurement of seafloor deformation in the marine sector of the Campi Flegrei caldera (Italy). *Journal of Geophysical Research: Solid Earth*, 123(1), 66-83.

Iannaccone, G., Zollo, A., Elia, L., Convertito, V., Satriano, C., Martino, C., ... & Emolo, A. (2010). A prototype system for earthquake early-warning and alert management in southern Italy. *Bulletin of Earthquake Engineering*, 8, 1105-1129.

Improta, L., De Gori, P., & Chiarabba, C. (2014). New insights into crustal structure, Cenozoic magmatism, CO₂ degassing, and seismogenesis in the southern Apennines and Irpinia region from local earthquake tomography. *Journal of Geophysical Research: Solid Earth*, 119(11), 8283-8311.

Isaia, R., Di Giuseppe, M. G., Natale, J., Tramparulo, F. D. A., Troiano, A., & Vitale, S. (2021). Volcano-tectonic setting of the Pisciarelli fumarole field, Campi Flegrei caldera, southern Italy: Insights into fluid circulation patterns and hazard scenarios. *Tectonics*, 40(5), e2020TC006227.

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.

Jones, L. M. (1988). Focal mechanisms and the state of stress on the San Andreas fault in southern California. *Journal of Geophysical Research: Solid Earth*, 93(B8), 8869-8891.

Kanamori, H., & Anderson, D. L. (1975). Theoretical basis of some empirical relations in seismology. *Bulletin of the seismological society of America*, 65(5), 1073-1095.

Kaneko, Y., & Shearer, P. M. (2014). Seismic source spectra and estimated stress drop derived from cohesive-zone models of circular subshear rupture. *Geophysical Journal International*, 197(2), 1002-1015.

Karasözen, E., Özacar, A. A., Biryol, C. B., & Beck, S. L. (2014). Seismicity, focal mechanisms and active stress field around the central segment of the North Anatolian Fault in Turkey. *Geophysical Journal International*, 196(1), 405-421.

- Kato, A., & Nakagawa, S. (2014). Multiple slow-slip events during a foreshock sequence of the 2014 Iquique, Chile Mw 8.1 earthquake. *Geophysical Research Letters*, 41(15), 5420-5427.
- Kaviris, G., Elias, P., Kapetanidis, V., Serpetsidaki, A., Karakonstantis, A., Plicka, V., ... & Bernard, P. (2021). The western Gulf of Corinth (Greece) 2020–2021 seismic crisis and cascading events: First results from the Corinth Rift Laboratory Network. *The Seismic Record*, 1(2), 85-95.
- Keilis-Borok, V. (1959). On estimation of the displacement in an earthquake source and of source dimensions. *Annali di Geofisica*, 12(2), 205–214.
- Kim, W., Hong, T. K., Lee, J., & Taira, T. A. (2016). Seismicity and fault geometry of the San Andreas fault around Parkfield, California and their implications. *Tectonophysics*, 677, 34-44.
- Kwiatek, G., Plenkers, K., Dresen, G., & JAGUARS Research Group. (2011). Source parameters of picoseismicity recorded at Mponeng deep gold mine, South Africa: Implications for scaling relations. *Bulletin of the Seismological Society of America*, 101(6), 2592-2608.
- Kwiatek, G., Martínez-Garzón, P., Becker, D., Dresen, G., Cotton, F., Beroza, G. C., ... & Bohnhoff, M. (2023). Months-long seismicity transients preceding the 2023 MW 7.8 Kahramanmaraş earthquake, Türkiye. *Nature Communications*, 14(1), 7534.
- Lay, T., & Wallace, T. C. (1995). *Modern global seismology*. Elsevier.
- Lee, W. H., & Lahr, J. C. (1975). HYPO71 (revised; a computer program for determining hypocenter, magnitude, and first motion pattern of local earthquakes (No. 75-311). US Dept. of the Interior, Geological Survey, National Center for Earthquake Research.
- Li, Z., & Zhan, Z. (2018). Pushing the limit of earthquake detection with distributed acoustic sensing and template matching: A case study at the Brady geothermal field. *Geophysical Journal International*, 215(3), 1583-1593.
- Lim, C. S., Lapins, S., Segou, M., & Werner, M. J. (2024). Deep Learning Phase Pickers: How Well Can Existing Models Detect Hydraulic-Fracturing Induced Microseismicity from a Borehole Array?. *Authorea Preprints*.
- Lin, G., & Shearer, P. (2005). Tests of relative earthquake location techniques using synthetic data. *Journal of Geophysical Research: Solid Earth*, 110(B4).
- Liu, E., Zhu, L., Raj, A. G., McClellan, J. H., Al-Shuhail, A., Kaka, S. I., & Iqbal, N. (2017). Microseismic events enhancement and detection in sensor arrays using autocorrelation-based filtering. *Geophysical Prospecting*, 65(6), 1496-1509.

- Lomax, A., Virieux, J., Volant, P., & Berge-Thierry, C. (2000). Probabilistic earthquake location in 3D and layered models: Introduction of a Metropolis-Gibbs method and comparison with linear locations. *Advances in seismic event location*, 101-134.
- Lomax, A. (2008). Location of the Focus and Tectonics of the Focal Region of the California Earthquake of 18 April 1906. *Bulletin of the Seismological Society of America*, 98(2), 846-860.
- Lomax, A., Michelini, A., Curtis, A., & Meyers, R. A. (2009). Earthquake location, direct, global-search methods. *Encyclopedia of complexity and systems science*, 5, 2449-2473.
- Lomax, A. (2020). Absolute location of 2019 Ridgecrest seismicity reveals a shallow Mw 7.1 hypocenter, migrating and pulsing Mw 7.1 foreshocks, and duplex Mw 6.4 ruptures. *Bulletin of the Seismological Society of America*, 110(4), 1845-1858.
- Lomax, A., & Savvaidis, A. (2022). High-precision earthquake location using source-specific station terms and inter-event waveform similarity. *Journal of Geophysical Research: Solid Earth*, 127(1), e2021JB023190.
- Longobardi, V., Nazeri, S., Colombelli, S., Rea, R., De Landro, G., & Zollo, A. (2023). Time domain source parameter estimation of natural and man-induced microearthquakes at the geysers geothermal field. *Energies*, 16(3), 1121.
- Macedonio, G., Giudicepietro, F., D'auria, L., & Martini, M. (2014). Sill intrusion as a source mechanism of unrest at volcanic calderas. *Journal of Geophysical Research: Solid Earth*, 119(5), 3986-4000.
- Madariaga, R. (1976). Dynamics of an expanding circular fault. *Bulletin of the Seismological Society of America*, 66(3), 639-666.
- Madariaga, R. (1977). High-frequency radiation from crack (stress drop) models of earthquake faulting. *Geophysical Journal International*, 51(3), 625-651.
- Madariaga, R., Perrier, G., & Allègre, C. (1991). *Les tremblements de terre (Vol. 211)*. Paris: Presses du CNRS.
- Malagnini, L., Munafo', I., Cocco, M., Nielsen, S., Mayeda, K., & Boschi, E. (2014). Gradual fault weakening with seismic slip: Inferences from the seismic sequences of L'Aquila, 2009, and Northridge, 1994. *Pure and Applied Geophysics*, 171, 2709-2730.
- Malagnini, L., & Munafò, I. (2018). On the Relationship between ML and M_w in a Broad Range: An Example from the Apennines, Italy. *Bulletin of the Seismological Society of America*, 108(2), 1018-1024.

- Mancini, S., Segou, M., Werner, M. J., Parsons, T., Beroza, G., & Chiaraluce, L. (2022). On the use of high-resolution and deep-learning seismic catalogs for short-term earthquake forecasts: Potential benefits and current limitations. *Journal of Geophysical Research: Solid Earth*, 127(11), e2022JB025202.
- Marshall, S. T., & Morris, A. C. (2012). Mechanics, slip behavior, and seismic potential of corrugated dip-slip faults. *Journal of Geophysical Research: Solid Earth*, 117(B3).
- Martínez-Garzón, P., Kwiatek, G., Ickrath, M., & Bohnhoff, M. (2014a). MSATSI: A MATLAB package for stress inversion combining solid classic methodology, a new simplified user-handling, and a visualization tool. *Seismological Research Letters*, 85(4), 896-904.
- Martínez-Garzón, P., Kwiatek, G., Sone, H., Bohnhoff, M., Dresen, G., & Hartline, C. (2014b). Spatiotemporal changes, faulting regimes, and source parameters of induced seismicity: A case study from The Geysers geothermal field. *Journal of Geophysical Research: Solid Earth*, 119(11), 8378-8396.
- Martínez-Garzón, P., Beroza, G. C., Bocchini, G. M., & Bohnhoff, M. (2023). Sea level changes affect seismicity rates in a hydrothermal system near Istanbul. *Geophysical Research Letters*, 50(3), e2022GL101258.
- Matrullo, E., De Matteis, R., Satriano, C., Amoroso, O., & Zollo, A. (2013). An improved 1-D seismic velocity model for seismological studies in the Campania–Lucania region (Southern Italy). *Geophysical Journal International*, 195(1), 460-473.
- Maury, J., Cornet, F. H., & Dorbath, L. (2013). A review of methods for determining stress fields from earthquakes focal mechanisms; Application to the Sierentz 1980 seismic crisis (Upper Rhine graben). *Bulletin de la Societe Geologique de France*, 184(4-5), 319-334.
- Mavko, G., Mukerji, T., & Dvorkin, J. (2020). *The rock physics handbook*. Cambridge university press.
- Mazzoli, S., Barkham, S., Cello, G., Gambini, R., Mattioni, L., Shiner, P., & Tondi, E. (2001). Reconstruction of continental margin architecture deformed by the contraction of the Lagonegro Basin, southern Apennines, Italy. *Journal of the Geological Society*, 158(2), 309-319.
- Meng, X., Peng, Z., & Hardebeck, J. L. (2013). Seismicity around Parkfield correlates with static shear stress changes following the 2003 Mw6. 5 San Simeon earthquake. *Journal of Geophysical Research: Solid Earth*, 118(7), 3576-3591.

- Messuti, G., Scarpetta, S., Amoroso, O., Napolitano, F., Falanga, M., & Capuano, P. (2023). CFM: a convolutional neural network for first-motion polarity classification of seismic records in volcanic and tectonic areas. *Frontiers in Earth Science*, 11, 1223686.
- Michael, A. J. (1984). Determination of stress from slip data: faults and folds. *Journal of Geophysical Research: Solid Earth*, 89(B13), 11517-11526.
- Michael, A. J. (1987). Use of focal mechanisms to determine stress: a control study. *Journal of Geophysical Research: Solid Earth*, 92(B1), 357-368.
- Michele, M., Latorre, D., & Emolo, A. (2019). An empirical formula to classify the quality of earthquake locations. *Bulletin of the Seismological Society of America*, 109(6), 2755-2761.
- Michele, M., Chiaraluce, L., Di Stefano, R., & Waldhauser, F. (2020). Fine-scale structure of the 2016–2017 Central Italy seismic sequence from data recorded at the Italian National Network. *Journal of Geophysical Research: Solid Earth*, 125(4), e2019JB018440.
- Michelini, A., Cianetti, S., Gaviano, S., Giunchi, C., Jozinović, D., & Lauciani, V. (2021). INSTANCE—the Italian seismic dataset for machine learning. *Earth System Science Data*, 13(12), 5509-5544.
- Montone, P., Amato, A., & Pondrelli, S. (1999). Active stress map of Italy. *Journal of Geophysical Research: Solid Earth*, 104(B11), 25595-25610.
- Montone, P., Mariucci, M. T., Pondrelli, S., & Amato, A. (2004). An improved stress map for Italy and surrounding regions (central Mediterranean). *Journal of Geophysical Research: Solid Earth*, 109(B10).
- Mori, J., & Frankel, A. (1990). Source parameters for small events associated with the 1986 North Palm Springs, California, earthquake determined using empirical Green functions. *Bulletin of the Seismological Society of America*, 80(2), 278-295.
- 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.
- Mousavi, S. M., Ellsworth, W. L., Zhu, W., Chuang, L. Y., & Beroza, G. C. (2020). Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. *Nature communications*, 11(1), 3952.
- Mousavi, S. M., & Beroza, G. C. (2022). Deep-learning seismology. *Science*, 377(6607), eabm4470.

- Mueller, C. S. (1985). Source pulse enhancement by deconvolution of an empirical Green's function. *Geophysical Research Letters*, 12(1), 33-36.
- Münchmeyer, J., Woollam, J., Rietbrock, A., Tilmann, F., Lange, D., Bornstein, T., ... & Soto, H. (2022). Which picker fits my data? A quantitative evaluation of deep learning based seismic pickers. *Journal of Geophysical Research: Solid Earth*, 127(1), e2021JB023499.
- Muzellec, T., De Landro, G., Camanni, G., Adinolfi, G. M., & Zollo, A. (2024). The complex 4D multi-segmented rupture of the 2014 Mw 6.2 Northern Nagano Earthquake revealed by high-precision aftershock locations. *Authorea Preprints*.
- Nadeau, R. M., Foxall, W., & McEvilly, T. V. (1995). Clustering and periodic recurrence of microearthquakes on the San Andreas fault at Parkfield, California. *Science*, 267(5197), 503-507.
- Nakano, M., Kumagai, H., & Inoue, H. (2008). Waveform inversion in the frequency domain for the simultaneous determination of earthquake source mechanism and moment function. *Geophysical Journal International*, 173(3), 1000-1011.
- Natale, J., Camanni, G., Ferranti, L., Isaia, R., Sacchi, M., Spiess, V., ... & Vitale, S. (2022). Fault systems in the offshore sector of the Campi Flegrei caldera (southern Italy): Implications for nested caldera structure, resurgent dome, and volcano-tectonic evolution. *Journal of Structural Geology*, 163, 104723.
- Nazeri, S., Colombelli, S., & Zollo, A. (2019). Fast and accurate determination of earthquake moment, rupture length and stress release for the 2016–2017 Central Italy seismic sequence. *Geophysical Journal International*, 217(2), 1425-1432.
- NCEDC, 2014. Northern California Earthquake Data Center. UC Berkeley Seismological Laboratory [Dataset].
- Nespoli, M., Tramelli, A., Belardinelli, M. E., & Bonafede, M. (2023). The effects of hot and pressurized fluid flow across a brittle layer on the recent seismicity and deformation in the Campi Flegrei caldera (Italy). *Journal of Volcanology and Geothermal Research*, 443, 107930.
- Okada, Y.(1992) Internal deformation due to shear and tensile faults in a half-space. *Bulletin of the Seismological Society of America*, 82 (2), 1018–1040. <https://doi.org/10.1785/BSSA0820021018>
- Orsi, G., D'Antonio, M., de Vita, S., & Gallo, G. (1992). The Neapolitan Yellow Tuff, a large-magnitude trachytic phreatoplinian eruption: eruptive dynamics, magma

withdrawal and caldera collapse. *Journal of Volcanology and Geothermal Research*, 53(1-4), 275-287.

Orsi, G. (2022). Volcanic and deformation history of the Campi Flegrei volcanic field, Italy. *Campi Flegrei: A Restless Caldera in a Densely Populated Area*, 1-53.

Paige, C. C., & Saunders, M. A. (1982). LSQR: An algorithm for sparse linear equations and sparse least squares. *ACM Transactions on Mathematical Software (TOMS)*, 8(1), 43-71.

Palo, M., Picozzi, M., De Landro, G., & Zollo, A. (2023a). Microseismicity clustering and mechanic properties reveal fault segmentation in southern Italy. *Tectonophysics*, 856, 229849.

Palo, M., Scotto di Uccio, F., Picozzi, M., & Festa, G. (2023b). An Enhanced Catalog of Repeating Earthquakes on the 1980 Irpinia Fault System, Southern Italy. *Geosciences*, 14(1), 8.

Park, Y., Beroza, G. C., & Ellsworth, W. L. (2022). Basement fault activation before larger earthquakes in Oklahoma and Kansas. *The Seismic Record*, 2(3), 197-206.

Park, Y., Beroza, G. C., & Ellsworth, W. L. (2023). A mitigation strategy for the prediction inconsistency of neural phase pickers. *Seismological Society of America*, 94(3), 1603-1612.

Pasquale, G., De Matteis, R., Romeo, A., & Maresca, R. (2009). Earthquake focal mechanisms and stress inversion in the Irpinia Region (southern Italy). *Journal of seismology*, 13, 107-124.

Pavlis, G. L. (1986). Appraising earthquake hypocenter location errors: a complete, practical approach for single-event locations. *Bulletin of the Seismological Society of America*, 76(6), 1699-1717.

Peng, Z., & Zhao, P. (2009). Migration of early aftershocks following the 2004 Parkfield earthquake. *Nature Geoscience*, 2(12), 877-881.

Perol, T., Gharbi, M., & Denolle, M. (2018). Convolutional neural network for earthquake detection and location. *Science Advances*, 4(2), e1700578.

Petrosino, S., & De Siena, L. (2021). Fluid migrations and volcanic earthquakes from depolarized ambient noise. *Nature Communications*, 12(1), 6656.

Picozzi, M., Oth, A., Parolai, S., Bindi, D., De Landro, G., & Amoroso, O. (2017). Accurate estimation of seismic source parameters of induced seismicity by a combined approach of generalized inversion and genetic algorithm: Application to The Geysers

geothermal area, California. *Journal of Geophysical Research: Solid Earth*, 122(5), 3916-3933.

Picozzi, M., Bindi, D., Festa, G., Cotton, F., Scala, A., & D'Agostino, N. (2022a). Spatiotemporal evolution of microseismicity seismic source properties at the Irpinia near-Fault Observatory, southern Italy. *Bulletin of the Seismological Society of America*, 112(1), 226-242.

Picozzi, M., Iaccarino, A. G., Bindi, D., Cotton, F., Festa, G., Strollo, A., ... & Spallarossa, D. (2022b). The DENSE mulTi-paramETriC observations and 4D high resoluTion imaging (DETECT) experiment, a new paradigm for near-fault observations. In *EGU General Assembly Conference Abstracts* (pp. EGU22-8536).

Pita-Sllim, O., Chamberlain, C. J., Townend, J., & Warren-Smith, E. (2023). Parametric testing of EQTransformer's performance against a high-quality, manually picked catalog for reliable and accurate seismic phase picking. *The Seismic Record*, 3(4), 332-341.

Podvin, P., & Lecomte, I. (1991). Finite difference computation of traveltimes in very contrasted velocity models: a massively parallel approach and its associated tools. *Geophysical Journal International*, 105(1), 271-284.

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. *Geophysical Journal International*, 205(3), 1548-1573.

Poupinet, G., Ellsworth, W. L., & Frechet, J. (1984). Monitoring velocity variations in the crust using earthquake doublets: An application to the Calaveras Fault, California. *Journal of Geophysical Research: Solid Earth*, 89(B7), 5719-5731.

Prejean, S., Ellsworth, W., Zoback, M., & Waldhauser, F. (2002). Fault structure and kinematics of the Long Valley Caldera region, California, revealed by high-accuracy earthquake hypocenters and focal mechanism stress inversions. *Journal of Geophysical Research: Solid Earth*, 107(B12), ESE-9.

Priolo, E., Romanelli, M., Plasencia Linares, M. P., Garbin, M., Peruzza, L., Romano, M. A., ... & Fabris, P. (2015). Seismic monitoring of an underground natural gas storage facility: The Collalto Seismic Network. *Seismological Research Letters*, 86(1), 109-123.

Reasenberg, P., & Oppenheimer, D. (1986). FPFIT, FPLOT and FPPAGE: Fortran computer programs for calculating and displaying earthquake fault-plane solutions (Vol. 85, No. 739). US Geological Survey.

Resor, P. G., & Meer, V. E. (2009). Slip heterogeneity on a corrugated fault. *Earth and Planetary Science Letters*, 288(3-4), 483-491.

- Richards-Dinger, K. B., & Shearer, P. M. (2000). Earthquake locations in southern California obtained using source-specific station terms. *Journal of Geophysical Research: Solid Earth*, 105(B5), 10939-10960.
- Romano, M. A., Peruzza, L., Garbin, M., Priolo, E., & Picotti, V. (2019). Microseismic portrait of the Montello thrust (Southeastern Alps, Italy) from a dense high-quality seismic network. *Seismological Research Letters*, 90(4), 1502-1517.
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention—MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18* (pp. 234-241). Springer International Publishing.
- Rosi, M., Vezzoli, L., Aleotti, P., & De Censi, M. (1996). Interaction between caldera collapse and eruptive dynamics during the Campanian Ignimbrite eruption, Phlegraean Fields, Italy. *Bulletin of Volcanology*, 57, 541-554.
- Ross, Z. E., Meier, M. A., Hauksson, E., & Heaton, T. H. (2018). Generalized seismic phase detection with deep learning. *Bulletin of the Seismological Society of America*, 108(5A), 2894-2901.
- Ross, Z. E., Trugman, D. T., Hauksson, E., & Shearer, P. M. (2019). Searching for hidden earthquakes in Southern California. *Science*, 364(6442), 767-771.
- Ross, Z. E., Cochran, E. S., Trugman, D. T., & Smith, J. D. (2020). 3D fault architecture controls the dynamism of earthquake swarms. *Science*, 368(6497), 1357-1361.
- Rubin, A. M., Gillard, D., & Got, J. L. (1999). Streaks of microearthquakes along creeping faults. *Nature*, 400(6745), 635-641.
- Rubinstein, J. L., & Beroza, G. C. (2007). Full waveform earthquake location: Application to seismic streaks on the Calaveras fault, California. *Journal of Geophysical Research: Solid Earth*, 112(B5).
- Ruff, L. J. (2002). State of stress within the Earth. In *International Geophysics* (Vol. 81, pp. 539-557). Academic press.
- Sagy, A., Brodsky, E. E., & Axen, G. J. (2007). Evolution of fault-surface roughness with slip. *Geology*, 35(3), 283-286
- Sato, T., & Hirasawa, T. (1973). Body wave spectra from propagating shear cracks. *Journal of Physics of the Earth*, 21(4), 415-431.

- Satriano, C., Lomax, A., & Zollo, A. (2008). Real-time evolutionary earthquake location for seismic early warning. *Bulletin of the Seismological Society of America*, 98(3), 1482-1494.
- Scala, A., Adinolfi, G. M., Picozzi, M., Scotto di Uccio, F., Festa, G., De Landro, G., ... & 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.
- Scarpa, R., Bianco, F., Capuano, P., Castellano, M., D'Auria, L., Di Lieto, B., & Romano, P. (2022). Historic unrest of the Campi Flegrei caldera, Italy. In *Campi Flegrei: A Restless Caldera in a Densely Populated Area* (pp. 257-282). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Schaff, D. P., & Waldhauser, F. (2005). Waveform cross-correlation-based differential travel-time measurements at the Northern California Seismic Network. *Bulletin of the Seismological Society of America*, 95(6), 2446-2461.
- Schaff, D. P., & Waldhauser, F. (2010). One magnitude unit reduction in detection threshold by cross correlation applied to Parkfield (California) and China seismicity. *Bulletin of the Seismological Society of America*, 100(6), 3224-3238.
- Scholz, C. H. (2015). On the stress dependence of the earthquake b value. *Geophysical Research Letters*, 42(5), 1399-1402.
- Scholz, C. H. (2019). *The mechanics of earthquakes and faulting*. Cambridge university press.
- Schorlemmer, D., Wiemer, S., & Wyss, M. (2005). Variations in earthquake-size distribution across different stress regimes. *Nature*, 437(7058), 539-542.
- Scotto di Uccio, F., Scala, A., Festa, G., Picozzi, M., & Beroza, G. C. (2023). Comparing and integrating artificial intelligence and similarity search detection techniques: application to seismic sequences in Southern Italy. *Geophysical Journal International*, 233(2), 861-874.
- Scotto di Uccio, F., Michele, M., Strumia, C., Supino, M., Beroza, G. C., Chiaraluce, L., D'Agostino, N., & Festa, G. (2024a). Characterization and evolution of seismic sequences in the normal fault environment of the Southern Apennines. *Journal of Geophysical Research: Solid Earth*, 129(8), e2023JB028644.
- Scotto di Uccio, F., Lomax, A., Natale, J., Muzellec, T., Festa, G., Nazeri, S., ... & Zollo, A. (2024b). Delineation and fine-scale structure of fault zones activated during the 2014–

2024 unrest at the Campi Flegrei caldera (Southern Italy) from high-precision earthquake locations. *Geophysical Research Letters*, 51(12), e2023GL107680.

Shearer, P. M. (1997). Improving local earthquake locations using the L1 norm and waveform cross correlation: Application to the Whittier Narrows, California, aftershock sequence. *Journal of Geophysical Research: Solid Earth*, 102(B4), 8269-8283.

Shearer, P. M. (2002). Parallel fault strands at 9-km depth resolved on the Imperial fault, southern California. *Geophysical Research Letters*, 29(14), 19-1.

Shelly, D. R., Hill, D. P., Massin, F., Farrell, J., Smith, R. B., & Taira, T. A. (2013). A fluid-driven earthquake swarm on the margin of the Yellowstone caldera. *Journal of Geophysical Research: Solid Earth*, 118(9), 4872-4886.

Shelly, D. R. (2020). A high-resolution seismic catalog for the initial 2019 Ridgecrest earthquake sequence: Foreshocks, aftershocks, and faulting complexity. *Seismological Research Letters*, 91(4), 1971-1978.

Shible, H., Hollender, F., Bindi, D., Traversa, P., Oth, A., Edwards, B., ... & Gueguen, P. (2022). GITEC: A generalized inversion technique benchmark. *Bulletin of the Seismological Society of America*, 112(2), 850-877.

Simpson, R. W. (1997). Quantifying Anderson's fault types. *Journal of Geophysical Research: Solid Earth*, 102(B8), 17909-17919.

Skoumal, R. J., Brudzinski, M. R., Currie, B. S., & Levy, J. (2014). Optimizing multi-station earthquake template matching through re-examination of the Youngstown, Ohio, sequence. *Earth and Planetary Science Letters*, 405, 274-280.

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

Stabile, T. A., Satriano, C., Orefice, A., Festa, G., & Zollo, A. (2012). Anatomy of a microearthquake sequence on an active normal fault. *Scientific reports*, 2(1), 410.

Strumia, C., Trabattoni, A., Supino, M., Baillet, M., Rivet, D., & Festa, G. (2024). Sensing optical fibers for earthquake source characterization using raw DAS records. *Journal of Geophysical Research: Solid Earth*, 129(1), e2023JB027860.

Stucchi, M., Meletti, C., Montaldo, V., Crowley, H., Calvi, G. M., & Boschi, E. (2011). Seismic hazard assessment (2003–2009) for the Italian building code. *Bulletin of the Seismological Society of America*, 101(4), 1885-1911.

- Sugan, M., Vuan, A., Kato, A., Massa, M., & Amati, G. (2019). Seismic evidence of an early afterslip during the 2012 sequence in Emilia (Italy). *Geophysical Research Letters*, 46(2), 625-635.
- Sugan, M., Campanella, S., Chiaraluca, L., Michele, M., & Vuan, A. (2023). The unlocking process leading to the 2016 Central Italy seismic sequence. *Geophysical Research Letters*, 50(5), e2022GL101838.
- Supino, M., Festa, G., & Zollo, A. (2019). A probabilistic method for the estimation of earthquake source parameters from spectral inversion: application to the 2016–2017 Central Italy seismic sequence. *Geophysical Journal International*, 218(2), 988-1007.
- Supino, M., Scognamiglio, L., Chiaraluca, L., Doglioni, C., & Herrero, A. (2024). Source characterization of the 20th May 2024 MD 4.4 Campi Flegrei caldera earthquake through a joint source-propagation probabilistic inversion. *Seismica*, 3(2).
- Tamburello, G., Caliro, S., Chiodini, G., De Martino, P., Avino, R., Minopoli, C., ... & Capecchiacci, F. (2019). Escalating CO₂ degassing at the Pisciarelli fumarolic system, and implications for the ongoing Campi Flegrei unrest. *Journal of Volcanology and Geothermal Research*, 384, 151-157.
- Tan, Y. J., Waldhauser, F., Ellsworth, W. L., Zhang, M., Zhu, W., Michele, M., Chiaraluca, L., Beroza, G.C & Segou, M. (2021). Machine-learning-based high-resolution earthquake catalog reveals how complex fault structures were activated during the 2016–2017 central Italy sequence. *The Seismic Record*, 1(1), 11-19.
- Tarantino, S., Poli, P., D'Agostino, N., Vassallo, M., Festa, G., Ventafredda, G., & Zollo, A. (2024). Non-linear elasticity, earthquake triggering and seasonal hydrological forcing along the Irpinia fault, Southern Italy. *Nature Communications*, 15(1), 9821.
- Tarantola, A., & Valette, B. (1982). Generalized nonlinear inverse problems solved using the least squares criterion. *Reviews of Geophysics*, 20(2), 219-232.
- Todesco, M. (2021). Caldera's breathing: Poroelastic ground deformation at Campi Flegrei (Italy). *Frontiers in Earth Science*, 9, 702665.
- Tramelli, A., Godano, C., Ricciolino, P., Giudicepietro, F., Caliro, S., Orazi, M., ... & Chiodini, G. (2021). Statistics of seismicity to investigate the Campi Flegrei caldera unrest. *Scientific reports*, 11(1), 7211.
- Tramelli, A., Convertito, V., & Godano, C. (2024). b value enlightens different rheological behaviour in Campi Flegrei caldera. *Communications Earth & Environment*, 5(1), 275.

- Tramelli, A., Giudicepietro, F., Ricciolino, P., & Chiodini, G. (2022). The seismicity of Campi Flegrei in the contest of an evolving long term unrest. *Scientific reports*, 12(1), 2900.
- Troiano, A., Isaia, R., Di Giuseppe, M. G., Tramparulo, F. D. A., & Vitale, S. (2019). Deep Electrical Resistivity Tomography for a 3D picture of the most active sector of Campi Flegrei caldera. *Scientific reports*, 9(1), 15124.
- Troiano, A., Di Giuseppe, M. G., & Isaia, R. (2022). 3D structure of the Campi Flegrei caldera central sector reconstructed through short-period magnetotelluric imaging. *Scientific Reports*, 12(1), 20802.
- Troise, C., De Natale, G., Schiavone, R., Somma, R., & Moretti, R. (2019). The Campi Flegrei caldera unrest: Discriminating magma intrusions from hydrothermal effects and implications for possible evolution. *Earth-science reviews*, 188, 108-122.
- Trugman, D. T., & Shearer, P. M. (2017). GrowClust: A hierarchical clustering algorithm for relative earthquake relocation, with application to the Spanish Springs and Sheldon, Nevada, earthquake sequences. *Seismological Research Letters*, 88(2A), 379-391.
- Uchida, N., Matsuzawa, T., Ellsworth, W. L., Imanishi, K., Shimamura, K., & Hasegawa, A. (2012). Source parameters of microearthquakes on an interplate asperity off Kamaishi, NE Japan over two earthquake cycles. *Geophysical Journal International*, 189(2), 999-1014.
- Uchida, N., & Bürgmann, R. (2019). Repeating earthquakes. *Annual Review of Earth and Planetary Sciences*, 47(1), 305-332.
- Urbancic, T. I., Trifu, C. I., Mercer, R. A., Feustel, A. J., & Alexander, J. A. G. (1996). Automatic time-domain calculation of source parameters for the analysis of induced seismicity. *Bulletin of the Seismological Society of America*, 86(5), 1627-1633.
- Vanorio, T., & Kanitpanyachoen, W. (2015). Rock physics of fibrous rocks akin to Roman concrete explains uplifts at Campi Flegrei Caldera. *Science*, 349(6248), 617-621.
- Vassallo, M., Festa, G., & Bobbio, A. (2012). Seismic ambient noise analysis in southern Italy. *Bulletin of the Seismological Society of America*, 102(2), 574-586.
- 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. *Journal of Geophysical Research: Solid Earth*, 121(6), 4290-4305.
- Vavryčuk, V. (2011). Principal earthquakes: Theory and observations from the 2008 West Bohemia swarm. *Earth and Planetary Science Letters*, 305(3-4), 290-296.

Vidale, J. E., Ellsworth, W. L., Cole, A., & Marone, C. (1994). Variations in rupture process with recurrence interval in a repeated small earthquake. *Nature*, 368(6472), 624-626.

Vitale, S., & Isaia, R. (2014). Fractures and faults in volcanic rocks (Campi Flegrei, southern Italy): insight into volcano-tectonic processes. *International Journal of Earth Sciences*, 103, 801-819.

Vitale, S., Isaia, R., Ciarcia, S., Di Giuseppe, M. G., Iannuzzi, E., Prinzi, E. P., ... & Troiano, A. (2019). Seismically induced soft-sediment deformation phenomena during the volcano-tectonic activity of Campi Flegrei caldera (southern Italy) in the last 15 kyr. *Tectonics*, 38(6), 1999-2018.

Vitale, S., Natale, J., Isaia, R., Tramparulo, F. D. A., & Ciarcia, S. (2022). Evidence of seismic-related liquefaction processes within the volcanic record of the Campi Flegrei caldera (Italy). *Geosciences*, 12(6), 241.

Vuan, A., Sugan, M., Amati, G., & Kato, A. (2018). Improving the detection of low-magnitude seismicity preceding the Mw 6.3 L'Aquila earthquake: Development of a scalable code based on the cross correlation of template earthquakes. *Bulletin of the Seismological Society of America*, 108(1), 471-480.

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. *Geophysical Research Letters*, 47(17), e2020GL089039.

Waldhauser, F., Ellsworth, W. L., & Cole, A. (1999). Slip-parallel seismic lineations on the northern Hayward fault, California. *Geophysical research letters*, 26(23), 3525-3528.

Waldhauser, F., & Ellsworth, W. L. (2000). A double-difference earthquake location algorithm: Method and application to the northern Hayward fault, California. *Bulletin of the seismological society of America*, 90(6), 1353-1368.

Waldhauser, F. (2001). HypoDD-A program to compute double-difference hypocenter locations (No. 2001-113).

Waldhauser, F., & Schaff, D. P. (2008). Large-scale relocation of two decades of Northern California seismicity using cross-correlation and double-difference methods. *Journal of Geophysical Research: Solid Earth*, 113(B8).

Waldhauser, F., Michele, M., Chiaraluce, L., Di Stefano, R., & Schaff, D. P. (2021). Fault planes, fault zone structure and detachment fragmentation resolved with high-precision aftershock locations of the 2016–2017 central Italy sequence. *Geophysical Research Letters*, 48(16), e2021GL092918.

- Wallace, R. E. (1951). Geometry of shearing stress and relation to faulting. *The Journal of geology*, 59(2), 118-130.
- Westaway, R., & Jackson, J. (1987). The earthquake of 1980 November 23 in Campania—Basilicata (southern Italy). *Geophysical Journal International*, 90(2), 375-443.
- Wiemer, S. (2001). A software package to analyze seismicity: ZMAP. *Seismological Research Letters*, 72(3), 373-382.
- Wiens, D. A., McGuire, J. J., Shore, P. J., Bevis, M. G., Draunidalo, K., Prasad, G., & Helu, S. P. (1994). A deep earthquake aftershock sequence and implications for the rupture mechanism of deep earthquakes. *Nature*, 372(6506), 540-543.
- Woo, J. Y., & Kilburn, C. R. (2010). Intrusion and deformation at Campi Flegrei, southern Italy: sills, dikes, and regional extension. *Journal of Geophysical Research: Solid Earth*, 115(B12).
- Woollam, J., Rietbrock, A., Bueno, A., & De Angelis, S. (2019). Convolutional neural network for seismic phase classification, performance demonstration over a local seismic network. *Seismological Research Letters*, 90(2A), 491-502.
- Woollam, J., Münchmeyer, J., Tilmann, F., Rietbrock, A., Lange, D., Bornstein, T., ... & Soto, H. (2022). SeisBench—A toolbox for machine learning in seismology. *Seismological Society of America*, 93(3), 1695-1709.
- Wyss, M., & Brune, J. N. (1971). Regional variations of source properties in southern California estimated from the ratio of short-to long-period amplitudes. *Bulletin of the Seismological Society of America*, 61(5), 1153-1167.
- Yoon, C. E., O'Reilly, O., Bergen, K. J., & Beroza, G. C. (2015). Earthquake detection through computationally efficient similarity search. *Science advances*, 1(11), e1501057.
- Yoon, C. E., Huang, Y., Ellsworth, W. L., & Beroza, G. C. (2017). Seismicity during the initial stages of the Guy-Greenbrier, Arkansas, earthquake sequence. *Journal of Geophysical Research: Solid Earth*, 122(11), 9253-9274.
- Yoon, C. E., Bergen, K. J., Rong, K., Elezabi, H., Ellsworth, W. L., Beroza, G. C., ... & Levis, P. (2019). Unsupervised large-scale search for similar earthquake signals. *Bulletin of the Seismological Society of America*, 109(4), 1451-1468.
- Zhang, M., Tian, D., & Wen, L. (2014). A new method for earthquake depth determination: stacking multiple-station autocorrelograms. *Geophysical Journal International*, 197(2), 1107-1116.

Zhang, M., Ellsworth, W. L., & Beroza, G. C. (2019). Rapid earthquake association and location. *Seismological Research Letters*, 90(6), 2276-2284.

Zhang, M., Liu, M., Feng, T., Wang, R., & Zhu, W. (2022). LOC-FLOW: An end-to-end machine learning-based high-precision earthquake location workflow. *Seismological Society of America*, 93(5), 2426-2438.

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

Zoback, M. D. (2010). *Reservoir geomechanics*. Cambridge university press.

Zollo, A., & Emolo, A. (2011). *Terremoti e onde-metodi e pratica della sismologia moderna*. Liguori Editore.

Zollo, A., Orefice, A., & Convertito, V. (2014). Source parameter scaling and radiation efficiency of microearthquakes along the Irpinia fault zone in southern Apennines, Italy. *Journal of Geophysical Research: Solid Earth*, 119(4), 3256-3275.



Author's publications

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.
2. Palo, M., **Scotto di Uccio, F.**, Picozzi, M., & Festa, G. (2023). An Enhanced Catalog of Repeating Earthquakes on the 1980 Irpinia Fault System, Southern Italy. *Geosciences*, 14(1), 8.
3. **Scotto di Uccio, F.**, Scala, A., Festa, G., Picozzi, M., & Beroza, G. C. (2023). Comparing and integrating artificial intelligence and similarity search detection techniques: application to seismic sequences in Southern Italy. *Geophysical Journal International*, 233(2), 861-874.
4. Trabattoni, A., Biagioli, F., Strumia, C., van den Ende, M., **Scotto di Uccio, F.**, Festa, G., Rivet, D., Sladen, A., Ampuero, J.P., Metaxian, J., & Stutzmann, É. (2023). From strain to displacement: using deformation to enhance distributed acoustic sensing applications. *Geophysical Journal International*, 235(3), 2372-2384.
5. **Scotto di Uccio, F.**, Lomax, A., Natale, J., Muzellec, T., Festa, G., Nazeri, S., Convertito, V., Bobbio, A., Strumia, C., & Zollo, A. (2024). Delineation and fine-scale structure of fault zones activated during the 2014–2024 unrest at the Campi Flegrei caldera (Southern Italy) from high-precision earthquake locations. *Geophysical Research Letters*, 51(12), e2023GL107680.
6. **Scotto di Uccio, F.**, Michele, M., Strumia, C., Supino, M., Beroza, G. C., Chiaraluce, L., D'Agostino, N., & Festa, G. (2024). Characterization and evolution of seismic sequences in the normal fault environment of the Southern Apennines. *Journal of Geophysical Research: Solid Earth*, 129(8), e2023JB028644.