Finanziato dall'Unione europea NextGenerationEU | Ministero dell'Università e della Ricerca | Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA

multi-Risk sciEnce for resilienT commUnities undeR a changiNgclimate

Codice progetto MUR: **PE00000005** – B73C22001220006

Deliverable title: Report on Machine learning codes (proof-of-concept level) and pattern recognition - final report

Deliverable ID: DV 3.2.3

Due date: 30/3/2026

Submission date: 09/03/2026

AUTHORS

UNIPA: Alessandro Aiuppa, Marcello Bitetto, Luciano Curcio, Laura Insinga, Joao Pedro Lages Noguera, Giovanni Lo Bue Trisciuzzi; Angelo Vitale

OGS: Alessandro Vuan, Waed Abed, Luca Moratto, Andrea Carducci, Angela Saraò, Monica Sugan,

UNINA: Rosa Di Maio, Paola Petrosino

UNIBA: A. Siniscalchi, A. Tallarico, G. Romano, S. Tripaldi

UNIROMA SAPIENZA: Eugenio Carminati; Mattia Crespi; Chris Marone, Valeria Belloni, Cristiano Collettini, Laura Giaccio, Augusto Mazzoni, Michela Ravanelli, Roberta Ravanelli, Elisa Tinti

Technical references

Project Acronym	RETURN
Project Title	multi-Risk sciEnce for resilienT commUnities undeR a changiNg climate
Project Coordinator	Domenico Calcaterra UNIVERSITA DEGLI STUDI DI NAPOLI FEDERICO II domcalca@unina.it
Project Duration	December 2022 – March 2026 (40 months)

Deliverable No.	DV3.2.3
Dissemination level*	PU
Work Package	WP3.2 - Innovation in multiparametric monitoring on volcanoes
Task	T3.2.1 - Artificial intelligence time-series analysis
Lead beneficiary	UNIPA
Contributing beneficiary/ies	UNINA, UNIBA, UNIROMA1, OGS

* PU = Public
PP = Restricted to other programme participants (including the Commission Services)
RE = Restricted to a group specified by the consortium (including the Commission Services)
CO = Confidential, only for members of the consortium (including the Commission Services)

Document history

Version	Date	Lead contributor	Description
0.1	16.11.2023	Alessandro Aiuppa (UniPa)	First draft - first year intermediate report – M12
0.2	2.02.2026	All contributors	Critical review and proofreading
0.3	28.02.2026	All contributors	Edits for approval
1.0	9.03.2026	Alessandro Aiuppa (UniPa)	Final version

ABSTRACT

This report describes the results achieved in task 3.2.3: - Artificial intelligence time-series analysis. This task was devoted to setting up novel numerical methodologies for multivariate-multiparametric analysis of time series streamed by instrumental volcano networks, including forecasting approaches based on spatiotemporal pattern/precursor recognition through semiautomated computer analysis (machine learning). Within this task, major efforts have been spent to develop new tools and methods, eventually using machine learning and artificial intelligence, for near-real time processing and analysis of monitored signals at volcanoes. Here, we report on the obtained results in the following topics:

i. New statistical tools for processing volcanic gas datasets, and hence identify new composite indexes for the probabilistic quantification of the likelihood of a volcanic eruption. These techniques have been applied to volcanic gas records from Stromboli volcano. New interpretative tools of volcanic unrest, based on the use of volcanic gases, have also been identified.
ii. New methods for advanced real-time analysis and modelling of GNSS data;
iii. An integrated geophysical–numerical approach to identify permeable fault and fracture systems controlling natural CO_2 degassing in tectonically active environments
iv. A Random Forest Machine Learning algorithm to predict electrical resistivity in natural degassing systems;
v. Modelling of non-double-couple source mechanisms in volcanic seismicity
vi. On the use of Electrical Resistivity Tomography (ERT) for the understanding of mud volcanoes and their associated risks;
vii. on novel methods for the near-real-time identification of significant velocity variations in Global Navigation Satellite Systems (GNSS) records;
viii. on AI methods for rapid and precise earthquake location.

3. Table of contents

List of Figures

4.Results obtained

4.1 Novel methodologies for the analysis of volcanic gas time series

Volcanic gas geochemistry has long been recognised as a potentially very useful volcano monitoring tool. Magmatic volatiles have finite solubilities in magmas, causing them to partition into a separate volatile phase (gas bubbles) at upper mantle to crustal conditions. The intrinsically mobile nature of such gas bubbles implies that they can separate and escape from deeply stored magma to be eventually emitted at the surface well before magma. However, the use of volcanic gas observations as truly exploitable eruption forecasting tools has traditionally been challenged by a lack of quantitative data analysis testing the statistical robustness of any identified gas "precursor". This partly reflects the fact that, given to limitations and immaturity of gas sensing tools and networks, complete (temporally continuous) long-term volcanic gas datasets are limited to only a few, well-instrumented volcanoes. This has led to monitoring gas data being insufficient (in number, continuity, and quality) to systematically test patterns indicative of a certain volcanic behaviour.

In order to fill this gap of knowledge, and foster a major advance in the use of volcanic gases as volcano monitoring tools, we have develop novel methods for volcanic gas time-series analysis and pattern/precursor recognition (also assisted by machine learning), and have applied these methods to volcanic gas datasets obtained at Stromboli (within the context of volcano monitoring projects funded by DPC). New interpretative tools of volcanic unrest, based on the use of sulphur isotopes in volcanic gases and SO_2 fluxes, have also been identified.

4.1.1 Pattern recognition of volcanic gas time-series at Stromboli.

Stromboli volcano in Sicily (Southern Italy) hosts a well- instrumented volcanic gas monitoring network and, owing to its relatively steady and well-understood eruptive behaviour, is an ideal test site to quantitatively explore the link between volcanic gas variations and volcano behaviour. The volcano's generally mild (Strombolian) activity is punctuated by the periodic occurrence of sudden, larger-than-normal explosive eruptions referred as "major explosions". These explosions occur (2 to 4 times a year on average) without any obvious precursory change in surface activity, and therefore pose a threat to volcanologists/monitoring staff, and to the population in the most severe cases. Past work has found a link between periods of high CO_2/SO_2 ratios in the volcanic gas plume and the occurrence of such explosions, but this association has never been statistically verified.

In Return, we analysed a nearly continuous record of volcanic gas plume CO_2 and SO_2 concentrations and ratios in the plume, as well as SO_2 and CO_2 flux observations, recorded in a four-year (2020 to 2023) activity period of Stromboli volcano, during which 22 such major explosions took place (Fig. 4.1.1).

This dataset was then processed with novel tools never applied to volcanic gas time-series before (at least to our knowledge), which included the Conditional Logit model and the Receiver Operating Characteristic (ROC) analysis. From this analysis, we identified a characteristic pattern in volcanic gas flux/composition time-series prior to Stromboli's major explosions, in which most major explosions in 2020 to 2023 were preceded by periods (of days to weeks) of reduced plume SO_2 concentrations (highly significant), low SO_2 fluxes and increased CO_2/SO_2 ratios (significant). Analysis shows these variations are statistically significant (p values range from 0.002 to 0.02) over temporal timescales of days to weeks (Fig. 4.1.2). The statistical association

between the occurrence of major explosions and periods of reduced SO_2 concentrations (very significant) and fluxes, and high CO_2/SO_2 ratios (significant) is interpreted using a conceptual model in which major explosions are explained as caused by gas bubble accumulation at a rheological discontinuity, resulting from deceleration of the shallow convecting magma that supplies the "regular" Strombolian activity.

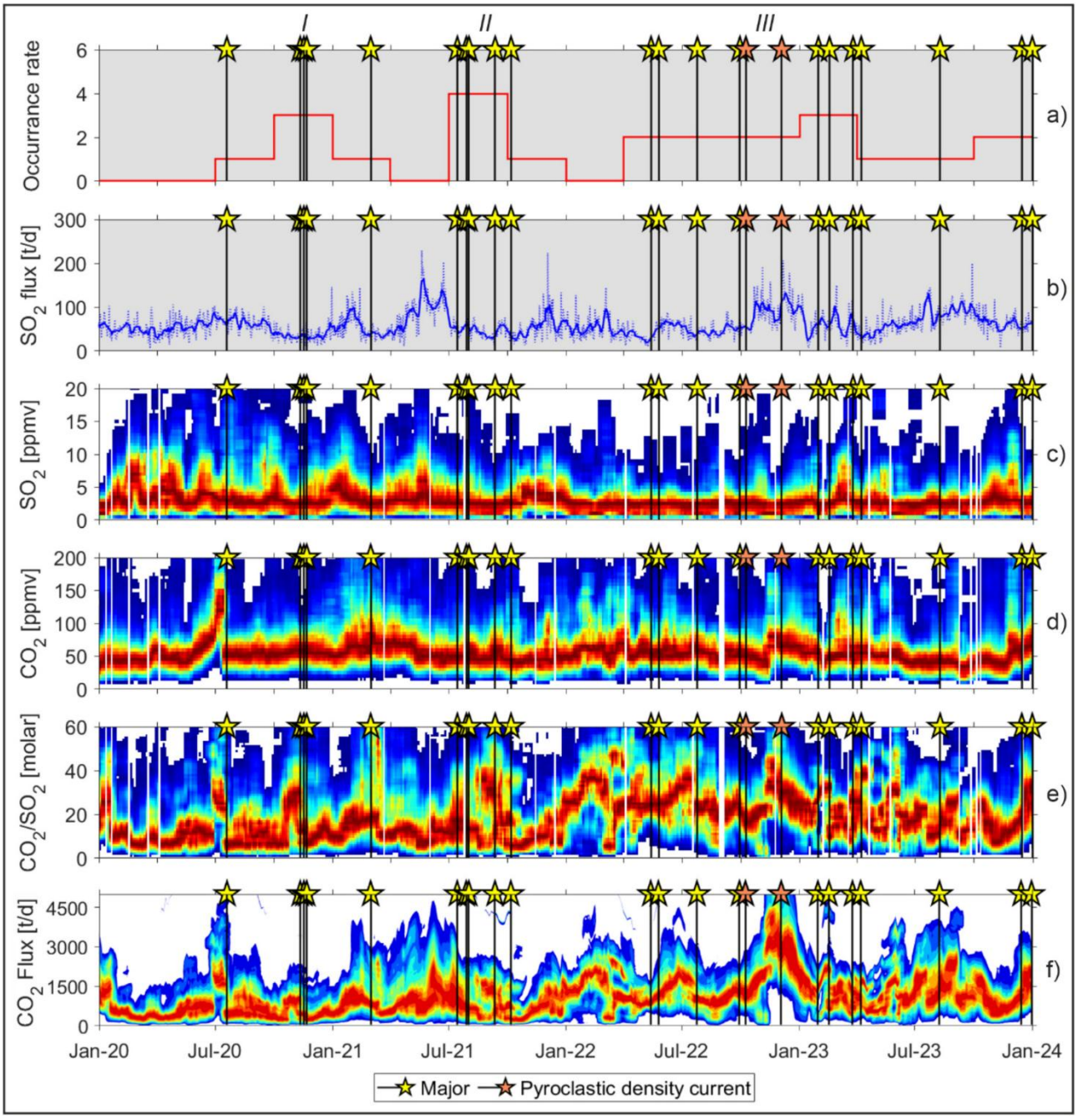


Figure 4.1.1: Four years (2020-2023) of gas observations at Stromboli. (a) Rate of occurrence (number of events/3 months) of major explosions. Yellow and orange stars identify major explosions and collapse/pyroclastic density currents, respectively; (b) Daily averaged SO_2 fluxes (t/d); (c) to (f), temporal plots of (normalized) frequency distribution diagrams of (c) SO_2 and (d) CO_2 concentrations in the plume, CO_2/SO_2 (molar) ratios (e) and CO_2 fluxes (f). Taken from Aiuppa et al (2025).

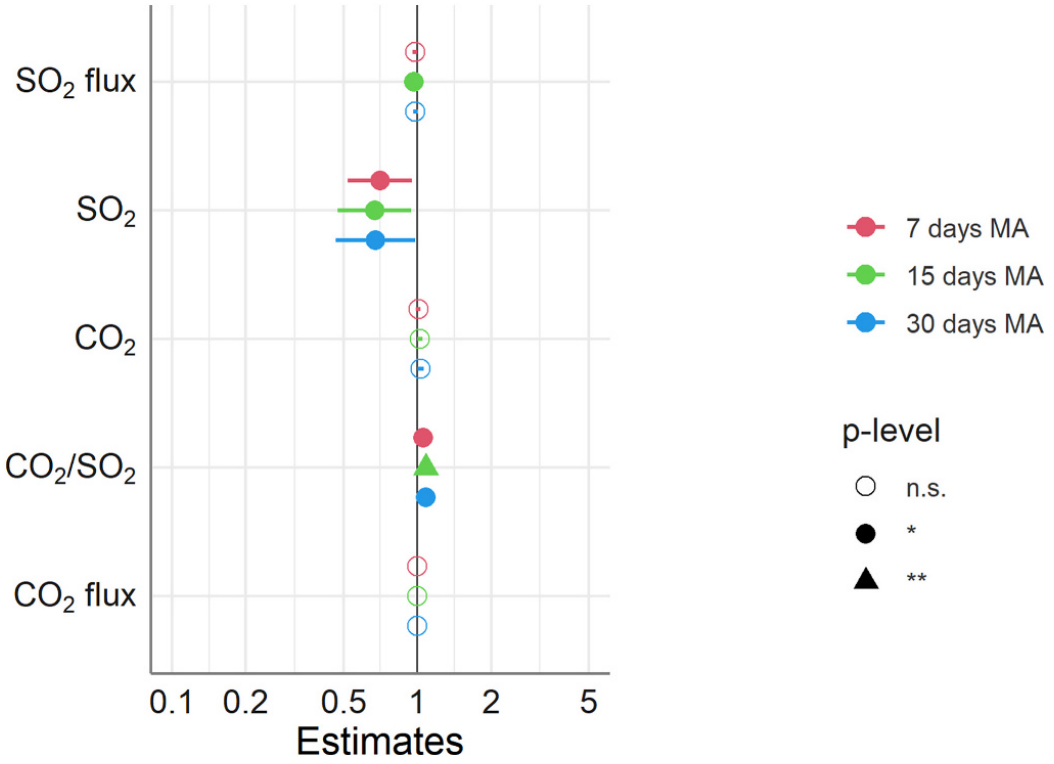


Figure 4.1.2: Odd ratio and 95 % confidence intervals for 7, 15 and 30 days moving averages of the four parameters analysed. p-levels: Open symbols, n.s., not significant; filled circles, statistically significant. Taken from Aiuppa et al (2025).

Using results of statistical analysis, we developed a volcanic gas-based Composite indicator (Fig. 4.1.3) that expresses a gas-based prediction - on a 0 to 1 scale - of having a major explosion, based on the linear combination of relative risks obtained for the statistically significant parameters (SO_2 concentrations and CO_2/SO_2 ratios). Results (Fig. 4.1.3) show that the Composite indicator varied widely during 2020–2023 (range, 0.08–0.90), and that major explosions were recurrently anticipated by increases to >0.48 (the optimal threshold) days to several weeks beforehand. The confusion matrix (Fig. 4.1.3, right) shows that the algorithm successfully forecasts (by a-posteriori analysis) 71 % of the events on timescales of week(s), yet being associated with a large (32 %) False Positive rate.

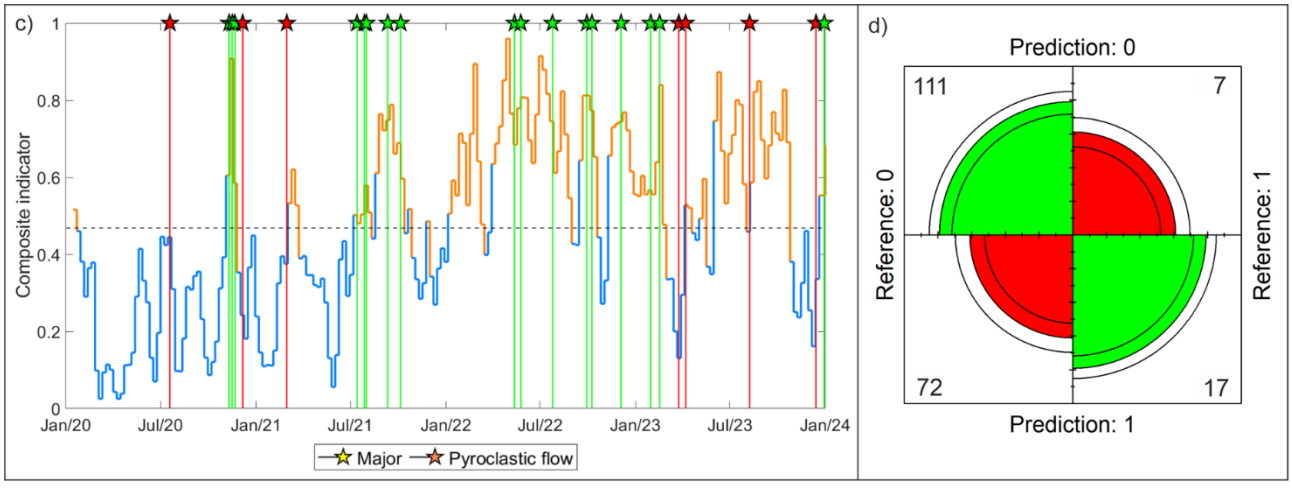


Figure 4.1.3: A composite indicator for estimating the likelihood of occurrence of "major explosions". At Stromboli. The confusion matrix is shown to the right. Taken from Aiuppa et al (2025).

Notably, and even better success rate 77 %) was obtained by processing the same time-series using a Machine Learning algorithm (SVC, Support Vector Classifier) (Fig. 4.1.4).

Ultimately, our results prove the existence of a recurrent volcanic gas temporal pattern prior to Stromboli's major explosions, and provide quantitative support to the use of volcanic gas observations for quantifying the likelihood of larger-than-normal basaltic explosive eruptions (Aiuppa et al., 2025b). This latter aspect has been further explored within the project by in-depth analysis of volcanic gas records obtained at Stromboli (Lo Bue Trisciuzzi et al., 2024; Delle Donne et al., 2026) and Etna (Aiuppa et al., 2023; Lo Bue Trisciuzzi et al., 2026), and by the characterization of the driving processes and mechanisms that control the generation of these events (Stix et al., 2025).

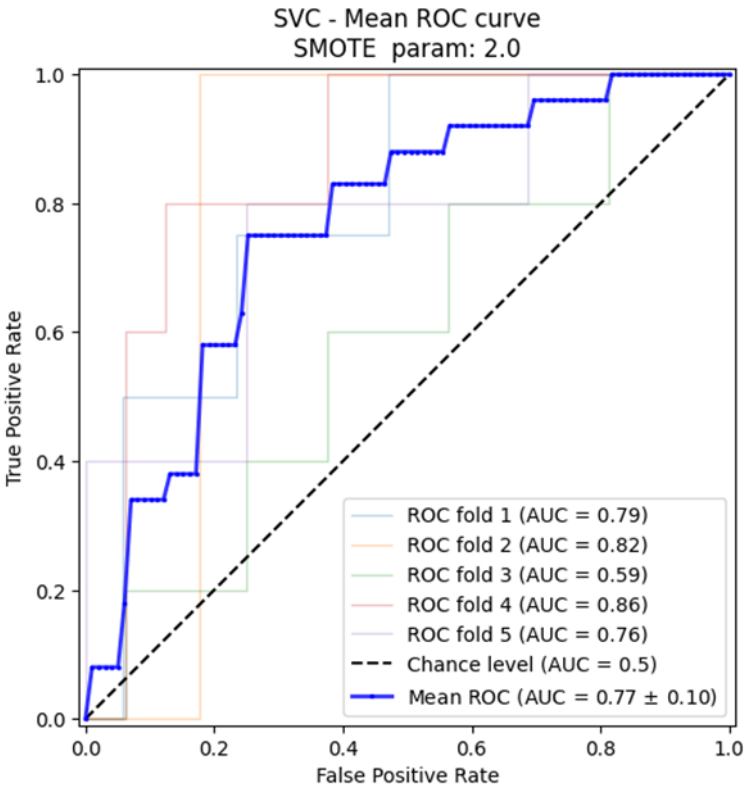


Figure 4.1.4: AOC curve showing the success rate of a Support Vector Classifier applied to the Stromboli volcanic gas dataset.

4.1.2 New interpretative tools of volcanic unrest

The long-lived dormancy periods of hydrothermal volcanoes and calderas are occasionally interrupted by episodes of escalating seismicity, ground deformation and changes in the chemistry/flux of hydrothermal fluid emissions. These unrest episodes are of special concern because they can represent signs of potential volcano reawakening. A remarkable example is offered by the resurgent Campi Flegrei (CF) caldera, in the suburban metropolitan area of Napoli (Italy) that, after four centuries of dormancy and subsidence following the 1538 eruption, has entered a phase of unrest since the 1950s, totally a net total uplift of ~4 m during 4 main ground uplift episodes (1950–1952, 1969–1972, 1982–1984 and 2000–present), accompanied by intense seismicity and escalation in degassing activity. Deformation/seismicity are thought to be caused by elastic/brittle response of the shallow crust to pressure build-up at the base of the hydrothermal system (at ~3-4 km depth), caused by escalating magmatic fluid transport. However, debate remains on whether magma is intruding at shallow depth or remains stationarily at depth, hence challenging interpretations of the future evolution of the unrest.

Emerging evidence indicates that, during unrest, hydrothermal fumaroles exhibit by an increasing transport of sulphur (Aiuppa et al., 2025b). However, the causes driving this escalating sulphur release have remained so far poorly understood, because this element can either be sourced by degassing magmas in the crust or by remobilization of hydrothermal minerals and brines.

Within the activities in RETURN, we contributed to developing novel interpretation models of unrest by modelling the recent escalation in hydrothermal sulphur (in the form of hydrogen sulphide, H_2S) observed at the Solfatara crater of CF since 2018 (Fig. 4.1.5) (Caliro et al., 2025).

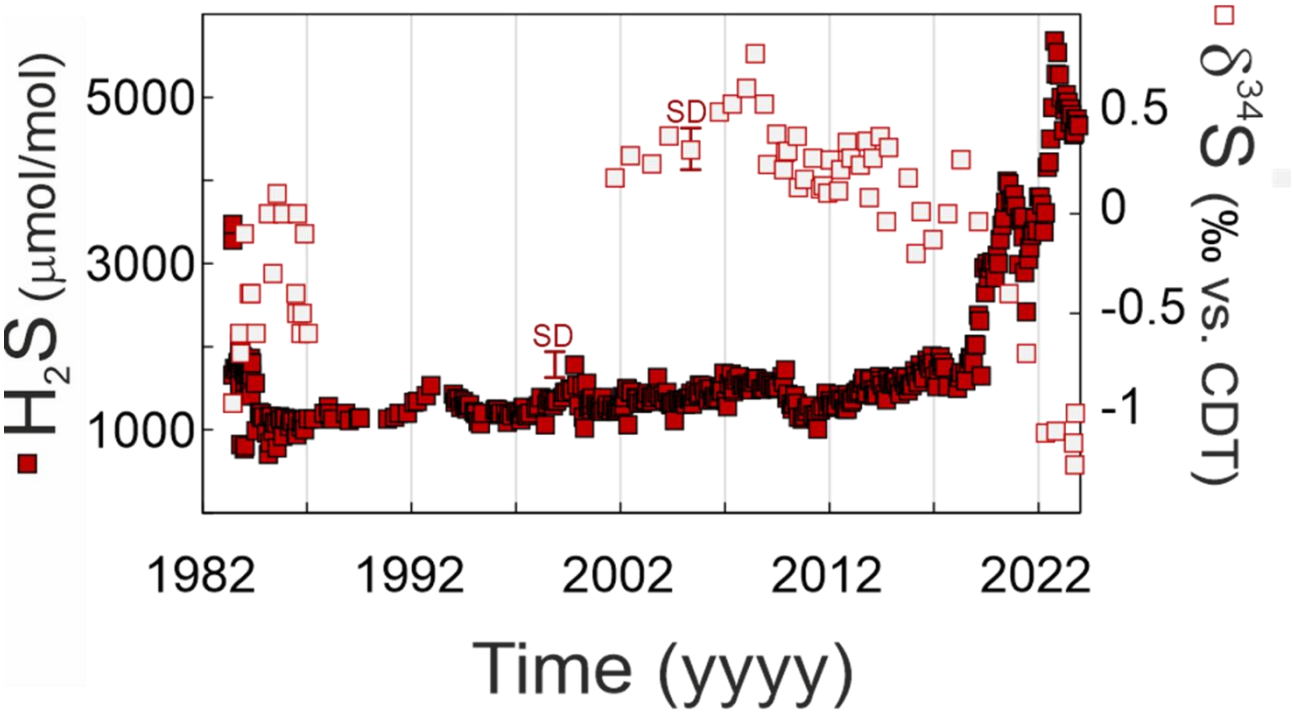


Figure 4.1.5: Temporal evolution of H_2S concentrations and sulphur isotope composition (expressed as $\delta^{34}S$) in the Bocca Grande fumarole inside Solfatara crater, CF. Data from regular monitoring of the Osservatorio Vesuviano, INGV. Modified from Caliro et al (2025).

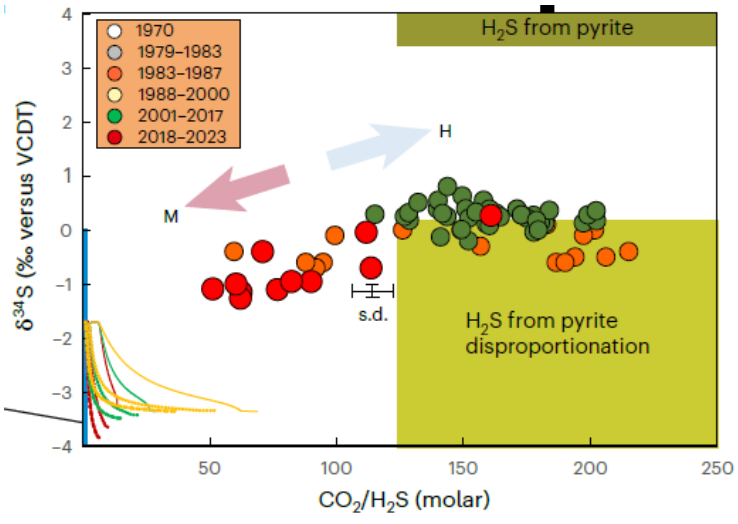


Figure 4.1.6: The compositional vs. sulphur isotope composition (expressed as $\delta^{34}S$) of Bocca Grande fumarole point to a shift form hydrothermal (H) to magmatic (M) signature. Modified from Caliro et al (2025).

An important contribution arising from the project is the demonstration that isotopic composition of sulphur is an especially useful parameters for interpreting unrest. In the specific CF case, results (Fig. 4.1.5) show that the increase in fumarolic H_2S abundance was paralleled by a systematic isotopic shift towards negative sulphur isotope compositions ($\delta^{34}S < 0$), that correspond to the isotopic signature of magmatic sulphur at CF (as inferred from the sulphur isotopic composition in olivine-hosted trachybasaltic melt inclusions from two of the most mafic eruptions of the recent CF activity; these constrain the CF magmatic source to have $\delta^{34}S$ of ~–1.7 ± 1.1‰, identical to those seen in the

fumaroles. It was hence concluded from modelling the escalation in sulphur transport at CF is consistent with an increasing transport of magmatic sulphur from a magmatic source located at depths of ≥6 km (Fig. 4.1.6).

4.2 Methodological Advances in Real-Time GNSS, Numerical Modelling, and Machine Learning for Volcano Monitoring

4.2.1 Real-Time GNSS Monitoring of Geohazard-Prone Areas Using Trimble CenterPoint RTX™

Cecere G., De Martino P., Riccardi U., Di Maio R., 2025. Evaluation of Trimble CenterPoint RTX correction service for Real-Time GNSS Monitoring: A Field-Based Comparison with RTK positioning. Discover Applied Sciences, 7, 1331, https://doi.org/10.1007/s42452-025-07889-z

An evaluation of the Trimble CenterPoint Real Time eXtended (RTX™) GNSS correction service was conducted to assess its suitability for high-precision, real-time monitoring of geohazard-prone areas and critical infrastructures, including volcanoes, landslides, seismic zones, and engineering structures. Real-time geodetic measurements are essential for detecting ground deformation with centimetre-level accuracy, but conventional methods such as single-baseline RTK require nearby reference stations and are limited by distance-dependent biases. RTX, based on Precise Point Positioning (PPP), delivers global, real-time corrections without reliance on local infrastructure, potentially overcoming these limitations (e.g., Kouba et al., 2017; Krasuski et al., 2018).

A series of controlled field tests were designed to simulate realistic ground displacement scenarios. A geodetic GNSS antenna was connected to two receivers simultaneously: one operating in RTK mode, referenced to a base station 3 km away, and one using the RTX service. Horizontal and vertical displacements were manually applied via a micrometric sled and telescopic rod, covering amplitudes of 1–50 cm and rates from 5×10^{-6} m/s to 5×10^{-4} m/s. Additional "blackout" tests simulated temporary signal loss. Data acquisition produced 108 sessions totaling 286,116 one-second observations, allowing robust comparison of positioning precision (STD, 2DRMS, CEP50) and accuracy (RMSE, MAE).

Results indicate that RTX consistently achieves precision and accuracy comparable to or exceeding RTK, with particular improvement for vertical displacements (<1 cm) and rapid deformation scenarios. As summarized in Figures 4.2.1, RTX performance remains robust across both moderate and rapid displacement rates, and under signal loss conditions, demonstrating its reliability for both slow-moving deformations (e.g., bradyseism or soil subsidence) and rapid events such as landslides. RTX performance also shows reduced error accumulation over longer sessions compared to RTK.

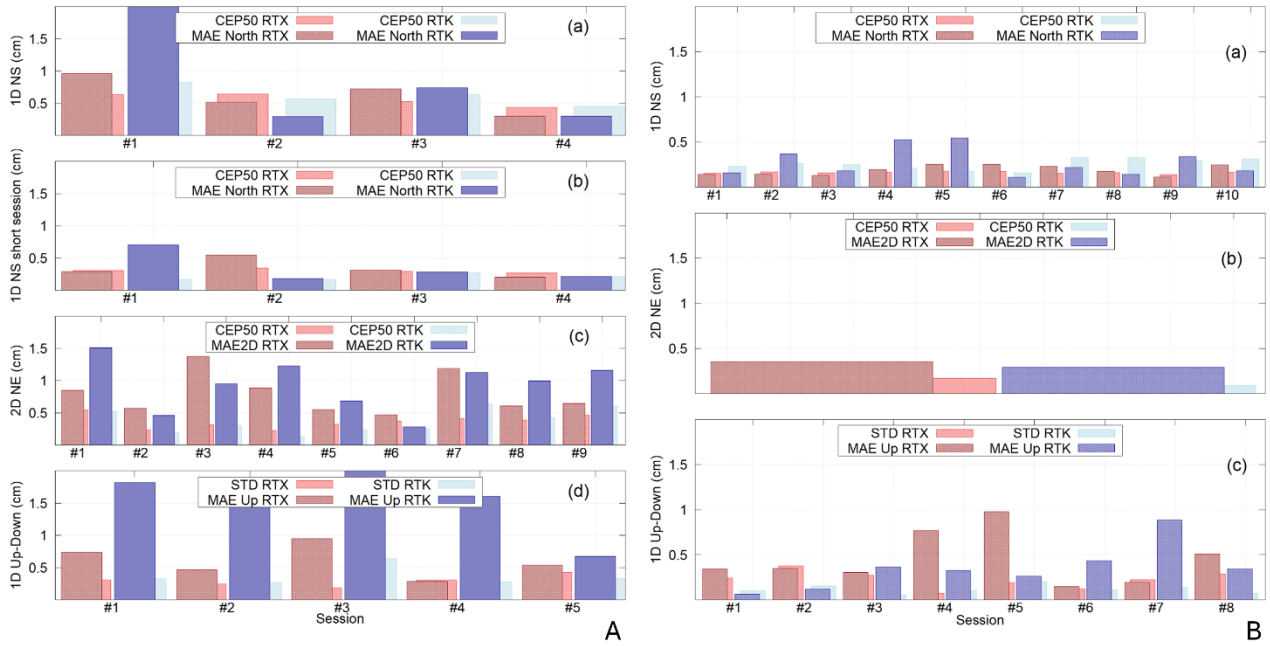


Figure 4.2.1: Accuracy and precision metrics for GNSS displacement tests. (A) Moderate-rate deformation: 1D North-South (long and short sessions, a–b), 2D North–East (c), and 1D Up–Down (d). (B) Rapid-rate deformation: 1D North–South (a), 2D North–East (b), and 1D Up–Down (c); metrics include MAE, CEP50, and STD.

The study highlights RTX as a reliable, scalable solution for real-time geodetic monitoring, with potential integration into AI-based time-series analysis frameworks for multiparametric geohazard surveillance. Future work will explore expanded experimental setups, including full 3D motion control, network RTK comparisons, and long-duration stability assessments, to further enhance real-time monitoring capabilities in dynamic geological environments.

4.2.2 Numerical Modelling of Permeable Fracture Networks

Salone R., Carbonari R., De Paola C., Iezzi F., Di Maio R., 2025. Numerical modelling to identify permeable fractures from geophysical imaging of natural degassing areas. Example from the Matese Fault system (Italy). Tectonophysics, 917, 230950, https://doi.org/10.1016/j.tecto.2025.230950

An integrated geophysical–numerical approach was developed to identify permeable fault and fracture systems controlling natural CO_2 degassing in tectonically active environments. In natural degassing areas, accurately constraining the structural pathways of deep fluids is essential for understanding gas migration processes, assessing environmental impacts, and improving hazard evaluation. Although 3D geophysical imaging can delineate subsurface discontinuities, it does not directly discriminate between structurally mapped faults and hydraulically active conduits.

To address this limitation, the study implemented 3D thermo-fluid dynamic simulations constrained by high-resolution Electrical Resistivity Tomography (ERT), Self-Potential (SP) measurements, and geological–geochemical data. The objective was to establish a physically consistent relationship between geophysically imaged structures and the dynamic processes governing H_2O–CO_2 multiphase flow.

The methodology was tested in the Solopaca degassing area within the Matese Fault System (Southern Apennines, Italy). A 3D petrophysical model derived from the integration of the geophysical modelling with the results of geological-structural and geochemical studies (Vitale et al., 2023) was implemented in the

multiphase flow simulator TOUGH2 (EOS2 module), using PetraSim as graphical interface. Simulations reproduced the ascent of a deep H_2O–CO_2 mixture along hypothesized fault zones and were iteratively validated against measured surface CO_2 flux data.

The initial conceptual model (Figure 4.2.2a), which assumed multiple permeable discontinuities, significantly overestimated surface fluxes and showed weak spatial agreement with observations (Figure 4.2.2b). Following iterative refinement—restricting permeability to two NW–SE normal faults supported by resistivity contrasts and SP anomalies—the model (Figure 4.2.2c) successfully reproduced both the order of magnitude and the spatial distribution of measured CO_2 emissions (Figure 4.2.2d). Statistical indicators (normalized Root Mean Square Error and low Spearman correlation) showed substantial improvement, confirming the robustness of the revised structural configuration.

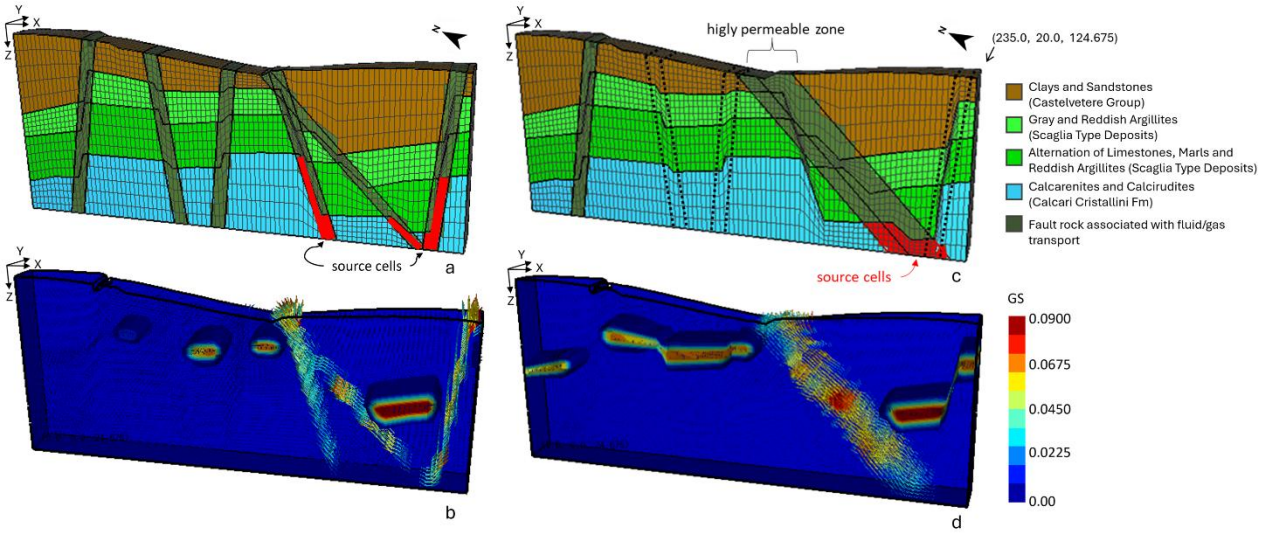


Figure 4.2.2: (a) Initial 3D petrophysical model of the Solopaca area used to simulate ascending CO_2 flux. (b) Corresponding distribution of CO_2 flow vectors and gas saturation (GS) isosurfaces after 100 years (quasi steady-state). (c) Refined 3D model including two hydraulically active fault zones (dark green cells). (d) Simulated CO_2 flow vectors and GS isosurfaces after 100 years for the refined configuration. In both models, red cells indicate the deep H_2O–CO_2 source region.

Furthermore, simulated gas saturation patterns within low-permeability stratigraphic units were consistent with high-resistivity volumes observed in the 3D ERT model, demonstrating internal coherence between numerical predictions and geophysical imaging. The results indicate that only a subset of mapped discontinuities actively controls geogenic CO_2 migration, while other structures are currently hydraulically inactive.

The proposed workflow establishes an iterative validation framework in which geophysical observations constrain the structural model, numerical simulations test hydraulic assumptions, and statistical comparison with surface flux measurements guides model refinement. This approach enhances the reliability of conceptual models in complex fractured systems and provides a quantitative basis for discriminating active from inactive fault zones.

Despite limitations—such as the assumption of homogeneous fault properties and the exclusion of biogenic CO_2 contributions—the methodology is transferable to other geological settings, including volcanic systems, geothermal reservoirs, and CO_2 storage sites.

Overall, the study demonstrates that integrating 3D geophysical imaging with thermo-fluid dynamic modelling significantly improves the structural and dynamic interpretation of natural degassing systems, transforming resistivity-derived geometries into physically validated models of subsurface fluid circulation and supporting predictive hazard assessment frameworks.

4.2.3 Machine Learning–Based Prediction of Electrical Resistivity in Natural Degassing Systems

Carbonari R., Salone R., De Paola C., Di Maio R., 2025. Predicting electrical resistivity in natural degassing geological systems through petrophysical and thermodynamic data: a machine learning approach. Pure and Applied Geophysics, https://doi.org/10.1007/s00024-025-03882-0

An innovative machine learning–based approach is developed to predict electrical resistivity in natural degassing geological systems, which are critical for human safety, groundwater quality, and volcanic hazard assessment. Although resistivity is highly sensitive to water saturation, gas content, and fluid temperature, its quantitative interpretation is complicated by the nonlinear relationships linking thermo-petrophysical and hydraulic parameters to geophysical measurements (e.g., Glover, 2015).

To overcome the limitations of traditional empirical models (e.g., modified Archie's law), the study implements a Random Forest (RF) algorithm trained on thermo-petrophysical numerical simulations. The objective is to establish a predictive relationship between simulated physical variables and 3D resistivity models derived from Electrical Resistivity Tomography (ERT) surveys.

The methodology was tested at two CO_2-dominated degassing sites in southern Italy (Solopaca and Ciorlano). For each site, a 3D ERT-derived resistivity model and a thermo-petrophysical numerical model were developed and spatially harmonized for training (Di Maio et al., 2021; Salone et al., 2025). The RF model achieved high predictive performance (R^2 = 0.80 at Solopaca and 0.94 at Ciorlano, Fig. 4.2.3), significantly outperforming conventional Archie-based approaches (Figure 4.2.4), particularly in complex geological settings with variable porosity and saturation.

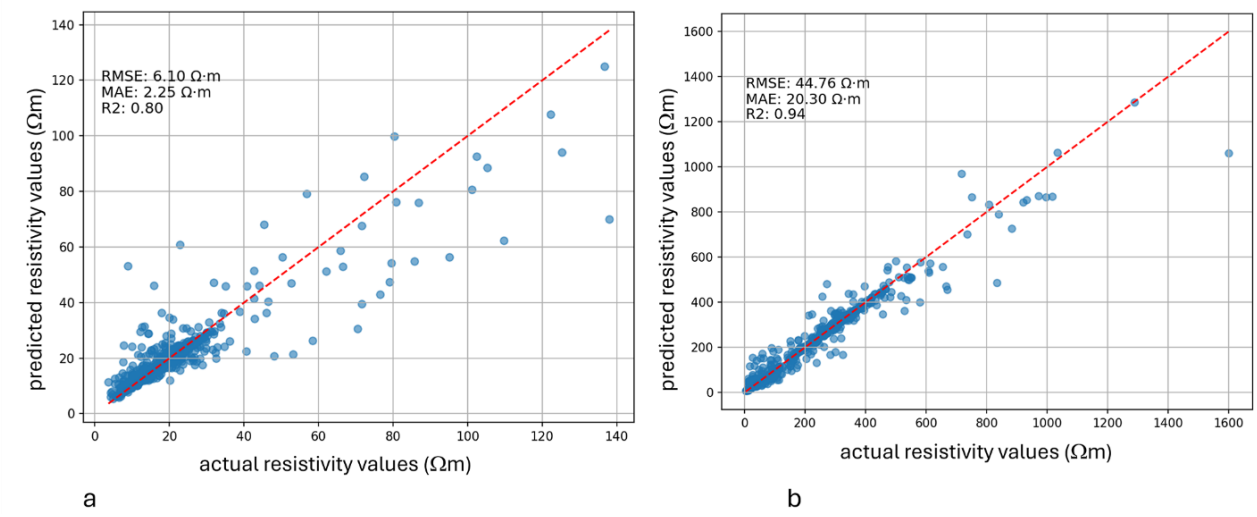


Figure 4.2.3: Resistivity prediction using the Random Forest algorithm for the (a) Ciorlano and (b) Solopaca model.

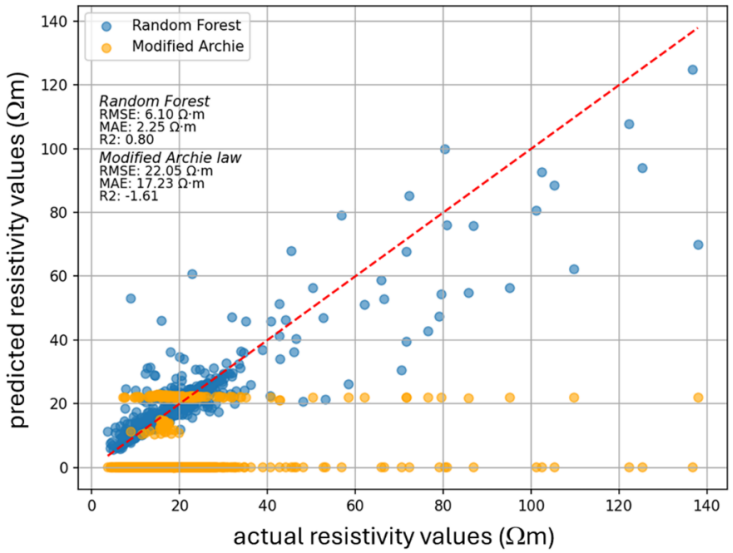


Figure 4.2.4: Comparison of resistivity predictions for the Solopaca model obtained using the Random Forest algorithm (blue dots) and the modified Archie's law (orange dots).

Moreover, feature-importance analysis showed that the most influential predictors are directly linked to system dynamics (heat flow, gas migration, thermo-hydraulic conditions), demonstrating the model's ability to learn physically meaningful and generalizable relationships. This enables iterative integration of geophysical monitoring data with numerical models, refining thermo-petrophysical parameters until simulated and observed resistivity values converge within a multi-method monitoring framework.

Despite limitations—such as spatial resolution mismatches between models and the need for detailed site-specific input data—the framework is transferable to other geological contexts, including volcanic hydrothermal systems, geothermal reservoirs, and CO_2 storage sites. The approach also opens perspectives for near–real-time applications through continuous integration of ERT and magnetotelluric data.

Overall, the study demonstrates that integrating thermo-petrophysical simulations with machine learning significantly enhances the quantitative and predictive monitoring of natural degassing systems, transforming resistivity measurements into a dynamic tool for constraining subsurface processes and supporting risk-informed management strategies.

4.3 Non double couple source earthquakes in the Etna seismicity

4.3.1 Abstract

The study of non-double-couple source mechanisms in volcanic seismicity is essential to unravel the complexity of volcanic processes, gain insights into magma dynamics, understand fluid-induced seismicity, characterize fracture and fault systems, explore volcano-tectonic interactions, and improve hazard assessment in volcanic regions.

The determination of the non-double-couple components requires advanced seismological techniques, as seismic signals from volcanic events can be very complex. Moment tensor inversion is generally used to analyze seismic waveforms and determine the components of the moment tensor that can be decomposed into a double couple and a non-double component associated with the seismic event.

In this study, we report a preliminary analysis of the non-double-couple components of a selected dataset of the highest energy seismic events (ML ≥ 3.4) that occurred in the Mt. Etna area in the period 2005-2022.

The solutions obtained by moment tensor analysis show that the double-couple components dominated the source mechanisms during the study period. In contrast, in the Pernicana fault system, mechanisms not related to double-coupling were predominantly observed, probably due to the presence of fluid circulation. These initial results, obtained in collaboration with researchers from the Osservatorio Etneo -INGV, form the basis for further studies on the interpretation of non-double mechanisms in the context of volcanic dynamics, which will be the subject of next year's work.

The moment magnitude (MW) determined for the 2005-2020 data set contributed to the definition of a data-driven ML - MW relationship for the Etnean region. Accurate quantification of the energy released by the seismic activity in volcanic regions is an important factor for improving seismic and volcanic hazard and risk assessment. This is of crucial importance for densely populated areas, as in the case of Etna.

The results presented in this report are part of a research study described in the paper:

Saraò, A., Moratto, L., Giampiccolo, E., & O Cocina, (2023). Moment magnitude for earthquakes in the Etna volcano area, Geophysical Journal International, 234, 2519- 2533, https://doi.org/10.1093/gji/ggad257

4.3.2 Introduction

The study of earthquake sources in volcanic areas provides important insights into the mechanisms driving seismicity associated with magma movement, fluid migration, and tectonic interactions. The seismic moment tensor describes fault geometry, seismic moment, and force system, allowing estimation of moment magnitude and, in volcanic environments, potentially constraining magma intrusion volumes, eruption probability, and the energetic state of an eruption (e.g., Meyer et al., 2021).

The full moment tensor can be decomposed into a double-couple (DC) component, associated with shear faulting, and non-double-couple (non-DC) components, namely the compensated linear vector dipole (CLVD) and the isotropic (ISO) term, which may reflect crack activation, fluid motion, or volume changes (e.g., Julian et al., 1998). However, non-DC components may also result from inversion artefacts, particularly affecting the CLVD term (e.g., Panza and Saraò, 2000), and may be influenced by structural complexity, noise, or limited station coverage; therefore, careful analyses are required before attributing them to physical source processes.

In Italy, moment tensor studies have investigated low-magnitude seismicity in the Campanian volcanic areas (Guidarelli et al., 2002; 2006) and at Etna. At Vesuvius, non-DC components decrease with depth, whereas at Campi Flegrei they appear correlated with seismicity rate. For Etna, analyses of the 1991–1993 and 2001 unrest episodes showed an increase in non-DC components prior to eruptions, possibly reflecting magma ascent, degassing, and the interaction between tectonic stress and volcanic activity (Saraò et al., 2001; 2010).

Following the development of the broadband seismic network managed by INGV-Osservatorio Etneo, full moment tensors were computed in near real time for earthquakes with 3.4 ≤ ML ≤ 4.8 occurring between 2005 and 2022 (Saraò et al., 2016; 2023; Cocina et al., 2022). In this study, we present this full moment tensor dataset, discuss the non-DC solutions, and refine the MW–ML relationship for Etna by integrating independent moment magnitude estimates derived from source spectral analysis.

4.3.3 The seismic moment tensor

The non-double couple investigated in this study was calculated for a selection of the highest energy earthquakes (3.4≤ ML ≤ 4.8) using the full-time domain moment tensor technique described in Minson and Dreger (2008). The minimum magnitude of earthquakes was defined by the limits imposed by background noise in the frequency passband. The Green functions used in the inversion were computed for a homogeneous stratified velocity structure retrieved for the Etna region (Alparone et al., 2012; Martinez-Arevalo et al., 2005). The fit between observed and synthetic seismograms was quantified by the variance reduction (VR). The source depth was determined iteratively by finding the solution that maximizes VR. The higher the VR the better the solution. The full solution of the moment tensor provides the orientation parameters (strike, dip, and rake) of the two nodal planes, the scalar seismic moment, and the percentage of the double couple (DC) and non-double couple, i.e., the compensated linear vector dipole (CLVD) and volumetric (ISO) part. In our analysis, we selected the solutions that minimize the non-DC components. A more detailed description of the method and its application can be found in Saraò et al., (2023).

4.3.4 Data

All data used for this study were provided by the Osservatorio Etneo of the INGV. For the analysis of the moment tensor, 75 earthquakes ML ≥3.4 recorded by the INGV-OE seismic broadband (0.01-40s) network in the period 2005-2020 were selected (see Saraò et al., 2023 for all details) among those with a signal-to-noise ratio greater than 5 (Saraò et al., 2023). The seismic waveforms of the three components were inverted after removing the instrumental response in the frequency range of 0.02-0.10 Hz or 0.02-0.05 Hz.

The dataset of MW for events with ML < 3.4 events was derived from response spectra and source spectral analyses of P waves from a subset of about 400 earthquakes with ML ≥1.0 , corresponding to about 6400 waveforms recorded by the INGV-OE seismic network since 2005. Selected data from the INGV-OE instrumental seismic catalogue (see Saraò et al., 2023 for more details) were examined to eliminate spurious transients and double events. Only data with a signal-to-noise ratio greater than 1.3 in the 1-25 Hz frequency band were included in the analysis of the source spectra.

4.3.5 Results and discussions

The focal mechanisms derived from full moment tensor inversion (Fig. 4.3.1a) show predominant NNW–SSE and NE–SW orientations, consistent with the structural framework of the Etna area, where different source mechanisms coexist, with strike-slip motions prevailing over subordinate normal and reverse components.

The overall distribution of moment tensor solutions (Fig. 4.3.1b) indicates that the DC component dominates across the study area. DC mechanisms are generally associated with the activation of regional tectonic structures compatible with the current stress field, particularly for the deepest events, supporting a tectonic stress-release interpretation. In some cases, DC solutions may also be linked to the formation of eruptive fissures, such as during the 2008 Etna eruption.

Non-DC components, instead, may reflect pore-pressure variations and fluid-related processes. Earthquakes located along the Pernicana Fault are predominantly non-DC (Fig. 4.3.1c), consistent with evidence of fluid circulation from geophysical and geochemical observations (Siniscalchi et al., 2010). However, the link between non-DC components and fluids is complex, as fluid flow may induce pressure imbalances, promote crack opening or closure, increase pore pressure, and reduce frictional resistance. Therefore, a more detailed analysis is required to fully interpret these results within the complex volcano-tectonic dynamics of the area.

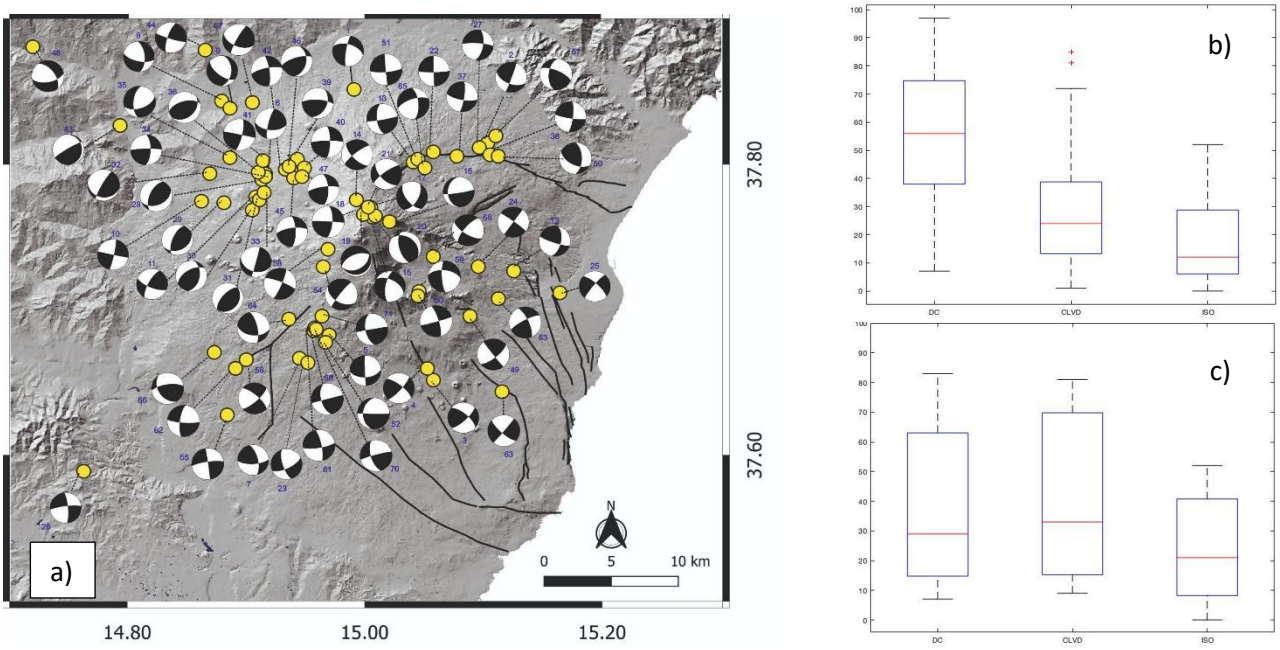


Figure 4.3.1: Map showing the epicenters (yellow circles) and focal mechanisms of the earthquakes analyzed in this study through moment tensor inversion. (b) Boxplots of the relative percentages of the moment tensor components (DC, CLVD, and ISO) for the entire study area; (c) same as (b), but limited to the Pernicana Fault area. In each boxplot, the red line represents the median, the box edges correspond to the 25th and 75th percentiles, and the whiskers extend to the most extreme values excluding outliers.

Finally, the MW values derived from full moment tensor solutions for earthquakes with $3.4 \leq ML \leq 4.8$ occurring between 2005 and 2020 were combined with independent seismic moment and MW estimates for smaller events ($1.0 \leq ML < 3.4$) obtained from source spectral analysis (see Saraò et al., 2023 for details) to calibrate a specific MW–ML relationship for the Etna area (Fig. 4.3.2). The resulting relation is as follows:

$$M_W=0.97(\pm0.01)\ M_L + 0.15(\pm0.01) \qquad \text{for } 1.0 \leq M_L \leq 4.8 \qquad (1)$$

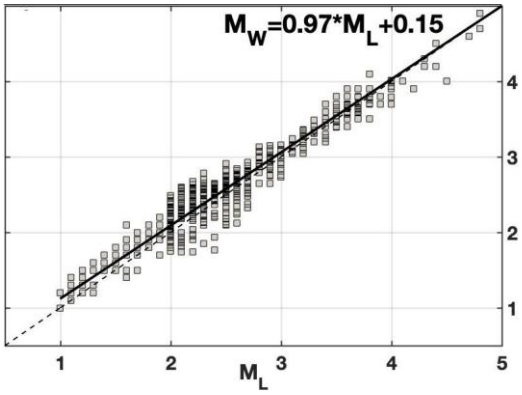


Figure 4.3.2: Map showing the epicenters (yellow circles) and focal mechanisms of the earthquakes analyzed in this study through moment tensor inversion. (b) Boxplots of the relative percentages of the moment tensor components (DC, CLVD, and ISO) for the entire study area; (c) same as (b), but limited to the Pernicana Fault area. In each boxplot, the red line represents the median, the box edges correspond to the 25th and 75th percentiles, and the whiskers extend to the most extreme values excluding outliers.

4.4 Template matching and deep learning applied to time series

4.4.1 Abstract

Template matching and deep learning applied to time series are increasingly contributing to earthquake and volcano monitoring, enabling the detection and classification of weak or complex events in large, heterogeneous datasets. In Task 3.2.3, we first focused on the February 2018 seismic swarm in the Azores, where template matching successfully expanded the event catalog, identified three main event families, and improved the interpretation of fluid injection episodes even under noisy or low-magnitude conditions (Soares et al., 2023). Building on this experience, the approach was applied to Mount Etna, utilizing an original Python-based HPC code developed at OGS. In Task 3.2.3, we also focused on Mount Etna, employing unsupervised deep learning pipelines for clustering and phase discrimination using daily seismic spectrograms, as well as a deep learning-based workflow for automatic seismic phase picking and association, tailored to generate volcano seismicity catalogs.

The template-matching pipeline increases the long-period event catalog of Etna events for the analyzed time window, resolving the persistence and recurrence of LP sources. The Unsupervised Deep Clustering pipeline reliably identifies spectral transitions and the main volcanic regimes (quiescence, LP-dominated, preparatory or recharge, paroxysmal or eruptive activity), achieving a prediction accuracy of over 90% for paroxysmal events during the investigated period (Abed et al., 2026). The workflow-driven deep learning approach achieves recognition rates of about 84% compared to reviewed VT catalogs for the selected time frame, providing further support for volcano monitoring (Carducci et al., 2026).

We collaborated with the Istituto Nazionale di Geofisica e Vulcanologia – Osservatorio Etneo (INGV-OE) and the University of Catania during the project. INGV-OE supported the project by providing essential data and expertise, which were crucial for addressing the scientific challenges at Mount Etna. For testing the Unsupervised Deep Clustering pipeline, the Helmholtz Centre Potsdam, German Research Centre for Geosciences (GFZ) shared code and valuable guidance.

The results presented in this report are part of a public research study described in the following papers:

- Abed, W., Zali, Z., Sciotto, M., Cocina, O., Cannata, A., Picozzi, M., Martínez-Garzón, P., Vuan, A., Saraò, A., Sugan, M. (2026) Hidden Patterns in Volcanic Seismicity: Deep Learning Insights from Mt. Etna's 2020–2021 Activity. Scientific Reports 16, 6155 (2026). https://doi.org/10.1038/s41598-026-36677-x.

- Carducci, A., Cocina, O., Sciotto, M., Cannata, A., Di Gioia, S., Vuan, A., Saraò, A., Tanaka Hernández K., Sugan, M. (2026). Insights into a Deep Learning Workflow for Automatic Detection of Seismic Signals at Mount Etna. Submitted to Seismological Research Letters. Under review.

- Soares, A., Custódio, S., Cesca, S., Silva, R., Vuan, A., Mendes, V.B. (2023). The February 2018 seismic swarm in São Miguel, Azores. Front. Earth Sci. 11:1144151. https://doi.org/10.3389/feart.2023.1144151.

4.4.2 Template Matching: Theory and Implementation

Template matching is a widely used technique in seismology that employs well-located seismic events as templates to identify small-magnitude earthquakes or tremor that may be hidden in noisy continuous data. The PyMPA code (Vuan et al., 2018) used in this project operates similarly to a matched-filter technique (MFT), identifying earthquakes that resemble these template events (e.g., Shelly et al., 2007; Peng and Zhao, 2009; Yang et al., 2009; Kato et al., 2012; Zhang and Wen, 2015).

Built on ObsPy routines (Krischer et al., 2015), the algorithm is highly versatile and supports most commonly used seismic data and earthquake catalog formats. The required inputs are daily continuous seismic waveforms, a parametric catalog of templates, and a 1D velocity model.

While template matching typically requires at least three-component seismic stations, the number of stations can vary. Reducing the number of available stations requires raising the detection threshold and cross-correlation criteria to maintain reliable detections.

The PyMPA code was rewritten to improve performance and scalability (Campanella, 2022), enabling execution during increased seismic activity in a volcanic unrest phase.

Key technical updates for managing computational workloads include:

1. Performance improvements: approximately 200% acceleration in single-threaded mode, near-linear scaling with multiple threads. GPU support, especially with NVIDIA V100, resulted in significant performance gains and enabled processing of tens to hundreds of templates per second per node.

2. Improved usability: enhanced command line interface (CLI), logging, input/output handling, and file system scanning for valid input. Output is provided in a structured binary format with optional compression. Flexibility is increased by supporting arbitrary signal lengths, template durations (per channel), and sampling rates.

3. Robustness and correctness: improved error handling addresses issues such as negative normalization in cross-correlation routines, dropping multiple detections within the template length, and ensuring stable size estimation. Comprehensive error corrections include handling missing data, template data, and improving detection at the beginning or end of the signal.

4. Maintainability: significant reduction of code (-80%) and dependencies, improved readability through refactoring into functions and meaningful variable names. The design emphasizes modularity and facilitates deployment through a registered PyPI package.

The code was applied to the S. Miguel sequence in the Azores archipelago (Soares et al., 2023) and Mount Etna, where a Long Period (LP) event catalog was compiled covering the period from 2019 to mid-2020.

4.4.3 Template matching on S. Miguel sequence in the Azores archipelago

Using the Template Matching automated workflow based on the analysis of continuous waveform data, we built an earthquake catalog for the February 2018 S. Miguel sequence in the Azores archipelago. The catalog includes sub-clusters identified by waveform similarity (Figures 4.4.1 and 4.4.2). The evolution of the cumulative events in our catalog at each stage of the automatic procedure is shown in Figure 3. The temporal evolution of the seismicity in our catalog shows that, although the earthquake sequence may initially appear to be a mainshock-aftershock sequence, a further detailed examination reveals that the highest magnitude earthquake occurred after the onset of the period of most intense activity (Figure 4.4.3).

We assess seismicity clustering and swarm-like behavior by computing different indicators typically used to distinguish between seismic swarms and sequences. The variations of these indicators among different sub-clusters indicate complex behavior, with clear swarm activity in the initial phase and a combination of swarm and mainshock-aftershock sequences during the latter part of the sequence, which includes the most significant events.

The swarm starts with a precursory phase of low-magnitude events, marking the onset of sub-cluster C1 activity. C1 is the deepest sub-cluster, suggesting that the swarm initiated with fracturing at depth, likely caused by the intrusion of volcanic fluids. The high-frequency nature of C1 waveforms supports the interpretation of brittle failure. As C1 activity diminishes, sub-cluster C2 emerges, exhibiting a high seismic rate and containing the largest magnitude earthquakes of the sequence. C2 is located slightly shallower and SE of C1, indicating fracture migration toward the surface and the activation of more superficial faults. Sub-cluster C3 emerged closely after C2 and soon after the earthquake with the highest magnitude. Family C3 is slightly shallower to the SE than C2, suggesting continued migration of earthquakes and the activation of a third, shallower fault system. C3 earthquakes have a very harmonic character, which we interpret as evidence of faulting in a fluid-enriched system.

We hypothesize that the swarm started with the activation of C1, whose earthquakes opened a pathway for fluids to ascend, leading to the activation of C2 and finally C3, consistent with a fluid-enriched environment. Previous studies show that seismicity in volcanic environments is often used to track the fluid propagation within the plumbing system, with recent notable cases involving migration over tens of kilometers and even across the entire crust. The spatial distribution of volcano seismicity with different characteristics, such as varying frequency content of their signals, is often interpreted to understand the plumbing system and other processes along complex fluid migration paths. A key difference between the earthquake families resides in the frequency content of their waveforms. We speculate that the higher-frequency waveform family, C1, indicates brittle failures, while the more harmonic and lower frequency waveforms of C2 and C3 may correspond to earthquakes occurring in the presence of fluids.

The results presented are part of the following paper:

Soares, A., Custódio, S., Cesca, S., Silva, R., Vuan, A., and Mendes, V.B. (2023). The February 2018 seismic swarm in São Miguel, Azores. Front. Earth Sci. 11:1144151. https://doi.org/10.3389/feart.2023.1144151

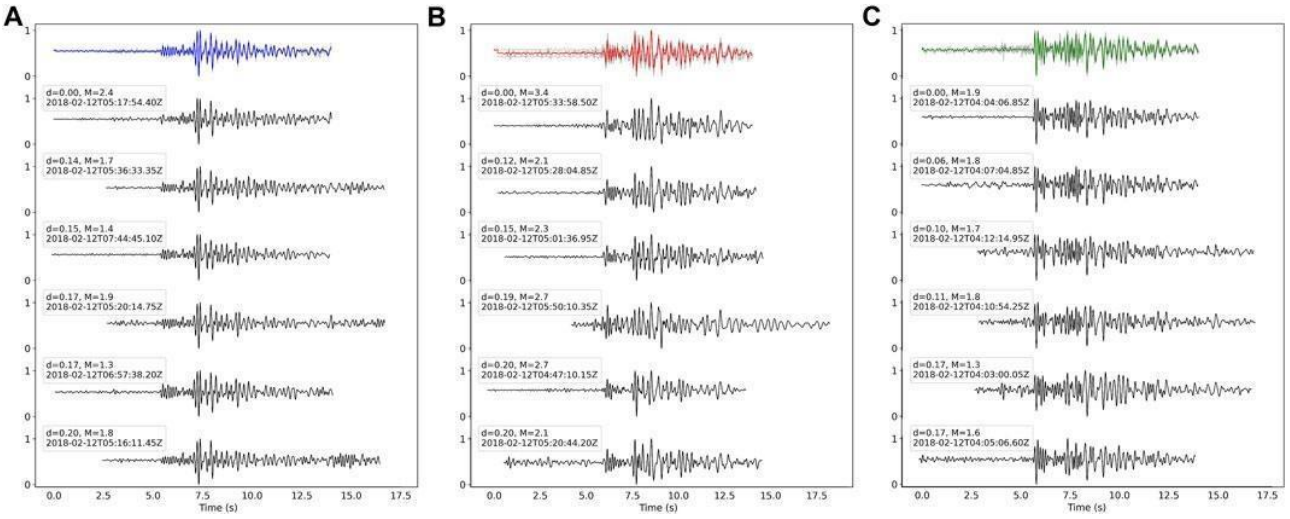


Figure 4.4.1 - Identification of different earthquake families based on the similarity of waveforms recorded at station MESC, vertical component (EHZ). Families C1 (blue), C2 (red) and C3 (green) are shown in panels from (A–C). Colored traces correspond to the stacked waveforms of each family. Selected traces are plotted downward in order of increasing dissimilarity to the top example waveform, where dissimilarity d = 1- CC.

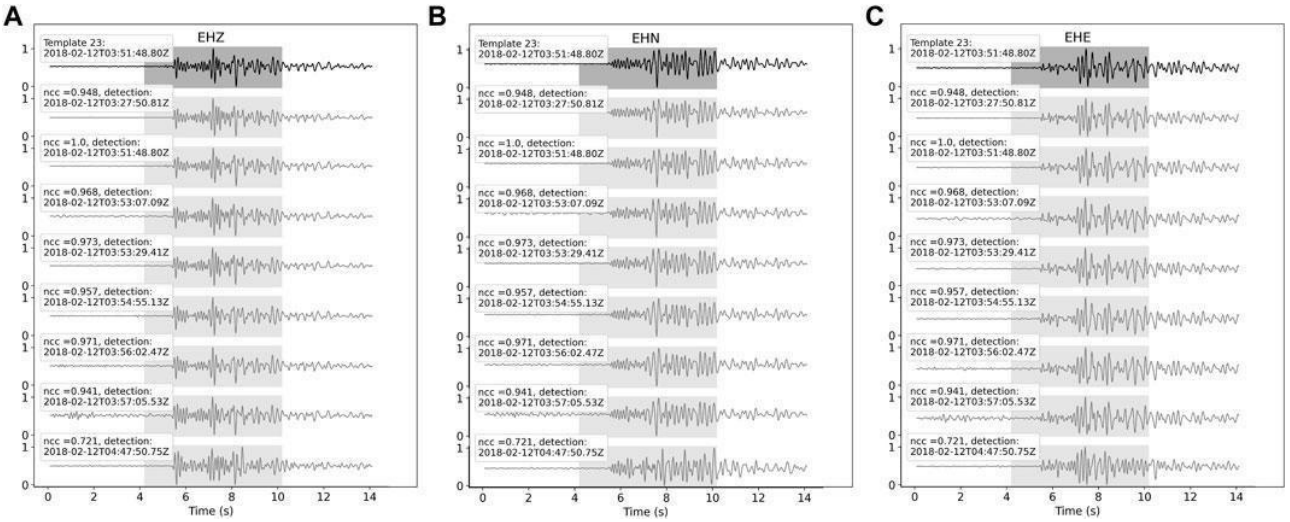


Figure 4.4.2 - Example of the template matching results, showing one template event (black line) and associated detections (gray). As an example, we show the template with origin time 2018-02-12T03:51:48.80. Each rectangle displays different components: (A) vertical (EHZ), (B) north (EHN), and (C) east (EHE) at station MESC. Each detection is assigned a ncc value, where an event with ncc = 1 corresponds to the detection of the template itself.

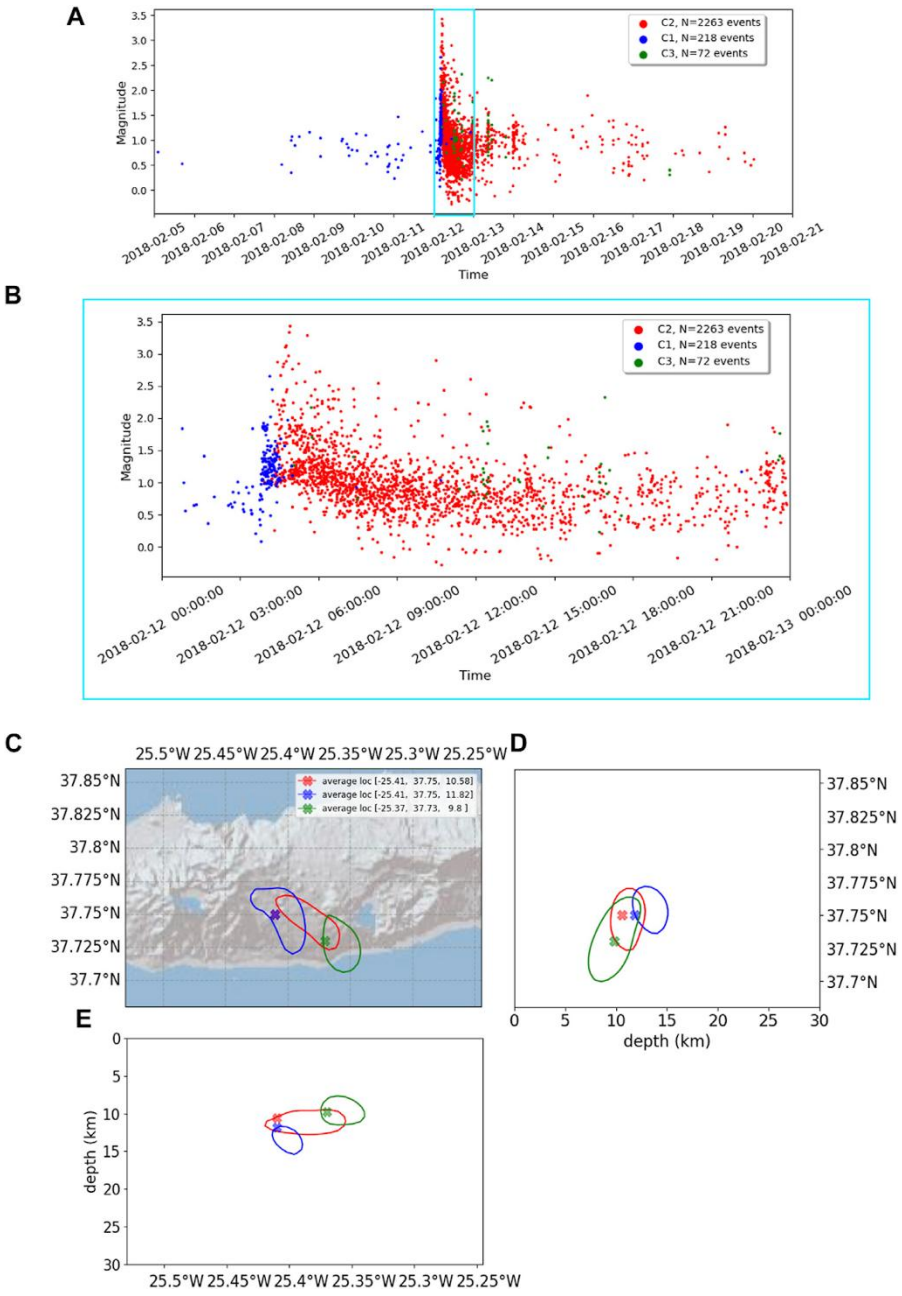


Figure 4.4.3 - Temporal evolution of earthquake's magnitude: (A) two-week period corresponding to the February 2018 swarm. (B) Zoom for February 12th, day with the highest seismic rate. Hypocentral location of the earthquake families: (C) map view and (D, E) cross-sections with contour line surrounding 70% of the events from the final catalog. Events are colored according to families identified by waveform similarity: C1 (blue), C2 (red) and C3 (green).

4.4.4 Template matching on Etna Long-Period Events

Using the Template Matching automated workflow based on the analysis of continuous waveform data, we built a Long Period (LP) event catalog for the period from 2019 to mid-2020 for Mount Etna.

We collected continuous seismic data from the monitoring network installed on the summit of Etna volcano, focusing on both single-station and multi-station approaches. For the single-station approach, we used three-component seismograms from the ECPN seismic station, which is considered a reference station for LPs

related to volcanic activity due to its location near the LP sources (e.g., Cannata et al., 2013; Cannata et al., 2015). For the multi-station approach, we also included other summit stations generally used by INGV-OE for the localization of LP events, such as ECNE, EPDN, and EPLC.

The dataset consisted of 1.5 years of continuous seismic recordings (January 2019 to June 2020), covering different phases of Etna's activity (Sciotto et al., 2022). For the LP catalog, we selected representative templates based on known characteristics and waveforms from previous studies. We used a subset of 30,741 LP events from the dataset described by Sciotto et al. (2022) and Cannata et al. (2025).

We processed the seismic data using standard preprocessing techniques, including downsampling (from 100 Hz to 20 Hz), filtering (0.5–8 Hz), and kurtosis-based analysis to improve signal quality and minimize background noise. We extended the frequency band to 8 Hz to enhance the distinction between LP events and VT events; however, in volcanic areas, this differentiation can be challenging, particularly for extremely shallow earthquakes with dominant low-frequency content (often referred to as VT-B) (Wassermann, 2012). We used 9 seconds of signals, trimmed relative to the theoretical S-wave arrival time (−4 sec, +5 sec).

All candidate templates underwent a kurtosis-based test to exclude noisy signals and those with maximum amplitude at the beginning or end of the trace for different channels. We selected only templates that met the evaluation criteria, using at least two channels for single-station configurations and at least five channels for multi-station configurations. We then applied the template matching algorithm to the preprocessed seismic data for both configurations.

The algorithm systematically compared each segment of the seismic recordings with predefined templates and computed cross-correlation coefficients (CC) to measure similarity. Candidate new LPs were identified using threshold criteria applied to the cross-correlation and daily median absolute deviation (MAD) values. We implemented the original procedure (Vuan et al., 2018; Campanella, 2022), in which the fully normalized correlation coefficient (CC) between the template and incoming data was used to determine a detection, along with a new procedure in which CC is replaced by C|C| (hereafter CC*) in the operational correlation detector, improving detection sensitivity without increasing the probability of false triggering (Gibbons, 2022).

Figure 4.4.4 presents the detections obtained using the multi-station approach. The template-matching method typically identifies more LP events than the classical STA/LTA triggers, and the number of detections generally decreases as the CC* threshold increases.

The ability to track event families using template matching could benefit hazard assessment as well as basic research on the structure of volcano plumbing and the recurrence of specific source mechanisms. The approach has been optimized for computational efficiency in an HPC/parallel environment, possibly enabling future integration into next-generation observatory pipelines.

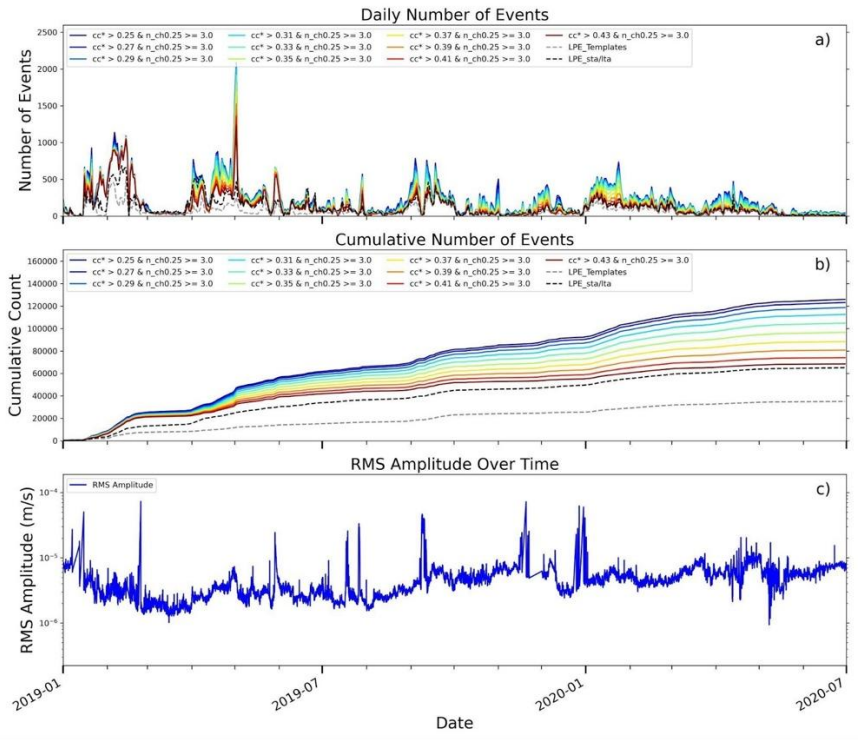


Figure 4.4.4 – Multi-station approach: a) Daily and b) cumulative number of LP events detected between January 2019 and June 2020, along with c) RMS amplitude measurements using the ECPN seismic station. The top panel (a) shows daily event counts for multiple detection thresholds (CC*) with at least three channels having CC* ≥ 0.25 and MAD ≥ 50. Template events (gray dashed line) and STA/LTA-triggered events (black dashed line) provided by INGV-OE are plotted for comparison. The middle panel (b) shows cumulative event counts over the same period, highlighting variations in detection rates. The bottom panel (c) displays the RMS amplitude at ECPN, showing signal fluctuations over time on a logarithmic scale.

4.4.5 The Unsupervised Deep Clustering pipeline at Etna

Deep learning (DL) techniques have recently transformed the analysis of continuous time series in seismic monitoring, enabling pattern recognition that surpasses traditional manual or algorithmic labeling. In this study, we applied autoencoder-based architectures, Clustremor (Zali et al., 2024) and AutoencoderZ (Zali et al., 2025), to identify clusters representing distinct seismic features corresponding to different volcanic phases at Mount Etna.

We used Clustremor and AutoencoderZ to extract compressed feature representations from daily seismic spectrograms at Mount Etna. Clustremor demonstrated its effectiveness by detecting volcanic tremors prior to the 2021 Geldingadalir eruption (Zali et al., 2024), while AutoencoderZ was applied to low-frequency signals preceding the 2023 MW 7.8 Kahramanmaraş earthquake (Zali et al., 2025).

The autoencoder architecture consists of:

1) Encoder – compresses the spectrogram input into a set of learned features using a series of convolutional and fully connected layers;

2) Bottleneck layer – a narrow, central layer that forces the network to learn a low-dimensional representation of the data (Hinton and Salakhutdinov, 2006);

3) Decoder – reconstructs the original input from the bottleneck features to verify the effectiveness of the learned representation.

In AutoencoderZ, the convolutional layers in the encoder are directly linked to their corresponding transposed convolutional layers in the decoder through skip connections. Unlike Clustremor, which does not include this feature, these connections enable AutoencoderZ to retain fine-scale spectral details that might otherwise be lost during compression. This design improves the robustness of feature extraction and allows the model to generalize more effectively across different data types and input sizes. For both the convolutional and fully connected layers, we employ the Exponential Linear Unit (ELU) activation function. In the decoder's final layer, we use a linear activation function. The loss function used for the autoencoder (LR) is the mean squared error (MSE), calculated as the difference between the input and the reconstructed output. To evaluate the efficiency of AutoencoderZ in reconstructing the input spectrograms, we introduced the Relative Bias (RB) metric (Cawley and Talbot, 2010).

We employed a Deep Embedded Clustering (DEC; Mousavi et al., 2019) technique that uses the spectrograms' latent representation for clustering. Clustering groups similar data points, and in this context, the features extracted from the autoencoder represent the compressed version of the spectrograms. This methodology ensures the extraction of the most pertinent data features for the clustering task through the simultaneous optimization of feature extraction and clustering. During the pre-training phase, the autoencoder is trained to reconstruct the output from the latent space close to the input. Subsequently, the features extracted from the bottleneck of the autoencoder are used for clustering with the k-means algorithm. K-means is recognized as one of the fastest and most widely used clustering methods. The algorithm divides the latent space into k clusters, each characterized by a cluster centroid. A cluster centroid, representing the center of a cluster, corresponds to the mean of all the data points within that cluster. To determine the optimal number of clusters, we varied the number of clusters from 2 to 8. We calculated the Davies-Bouldin Index (DBI; Davies and Bouldin, 1979) and the Calinski-Harabasz Index (CHI; Calinski and Harabasz, 1974). DBI measures the average similarity between each cluster and its most similar cluster, considering cluster compactness and separation. Lower DBI values indicate better clustering quality. CHI calculates the ratio of between-cluster dispersion to within-cluster dispersion. Higher CHI values indicate more distinct and well-separated clusters.

Using AutoencoderZ combined with the Deep Embedded Clustering (DEC) method on Etna's summit stations, we identified four clusters corresponding to distinct seismic regimes associated with different volcanic phases from November 2020 to November 2021 (Figure 4.4.5). We compared the resulting clusters with independent volcanic state indicators, including RMS amplitude, LP/VT event catalogs, and lava fountain timing, using the available INGV-OE dataset. Clustering results revealed four regimes:

- Regime 1: Quiescence or background seismicity, characterized by low amplitude and indistinguishable spectral content, attributed to pre- or inter-eruptive quiescence.

- Regime 2: LP-dominated phases, with transient, low-frequency vertical bands in the spectrograms reliably attributed to ongoing magma or gas migration through the shallow conduit system. This regime was most prevalent during periods of apparent recharge and internal pressurization.

- Regime 3: Preparatory phase, characterized by sustained high-amplitude tremor (consistent energy 0.5–5 Hz) and low LP event rates. These periods preceded recognized eruptive sequences, with a marked increase in shallow VT seismicity and RMS amplitude, representing a possible precursor signature.

- Regime 4: Paroxysmal or eruptive activity, closely coinciding with cataloged lava fountains and associated with broadband, high-energy spectrogram features, capturing the most intense eruptive episodes throughout the study period.

The results presented are part of the following paper:

Abed, W., Zali, Z., Sciotto, M., Cocina, O., Cannata, A., Picozzi, M., Martínez-Garzón, P., Vuan, A., Saraò, A., Sugan, M. (2026) Hidden Patterns in Volcanic Seismicity: Deep Learning Insights from Mt. Etna's 2020 - 2021 Activity. Scientific Reports 16, 6155 (2026). https://doi.org/10.1038/s41598-026-36677-x.

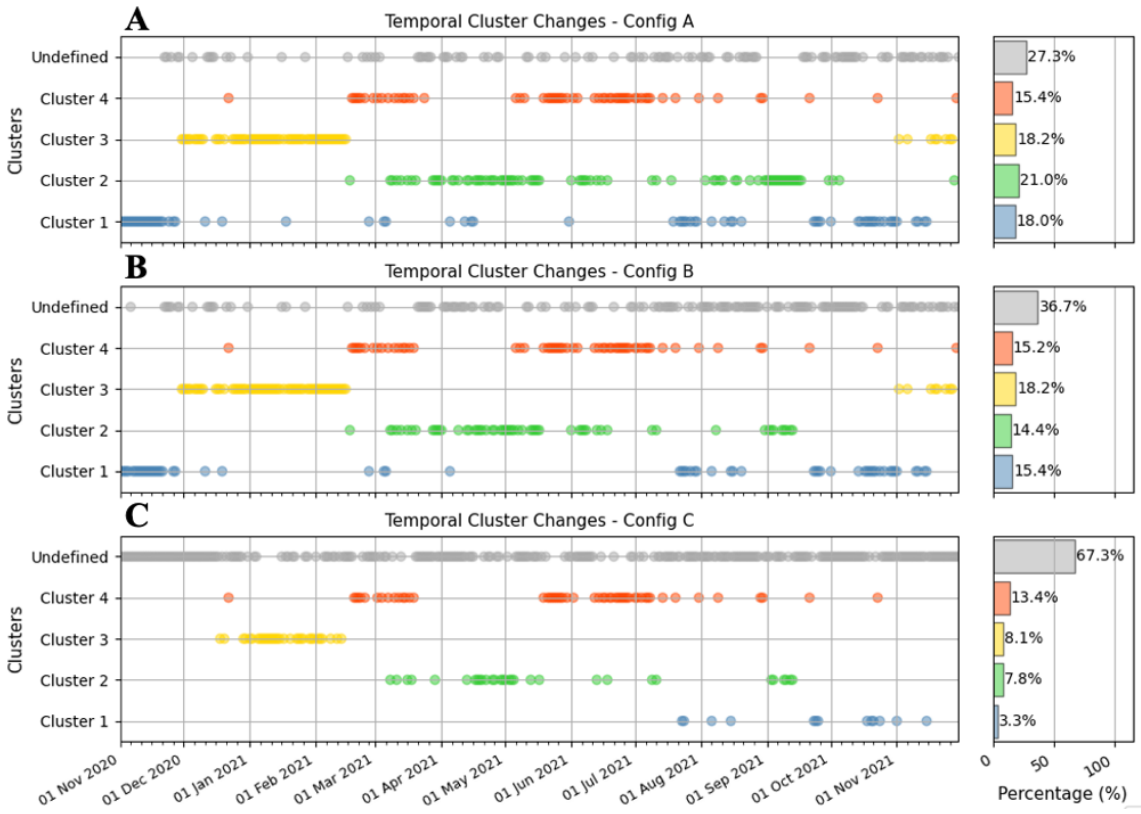


Figure 4.4.5 - Temporal evolution of clusters derived from multi-station analysis in relation to independent volcanic state indicators. Panels (A), (B), and (C) show different configurations of seismic station combinations used for clustering. In each row, the left plot shows the temporal distribution of cluster assignments over the analysis period, with each dot representing a day; the center plot displays the proportion of time each cluster was active during the analyzed period (modified from Abed et al., 2026).

4.4.6 Workflow-driven deep learning catalogue at Mount Etna

Using a deep learning (DL)-based workflow for automatic seismic phase picking and association, we constructed a catalog of Volcano-Tectonic (VT), Long-Period (LP), Low Spectrum (LS), and Mixed events spanning January 1, 2019, to June 30, 2020, for the Etna region. The waveform dataset was recorded by 38 seismic stations operated by INGV-OE. To establish a reliable baseline, we used two independent event catalogs provided by INGV-OE: the VT Earthquake Catalogue (Alparone et al., 2020; Barberi et al., 2020) and the LP Event Catalogue (Sciotto et al., 2022; Cannata et al., 2025), both limited to events within the Mount Etna domain.

We developed and optimized a distributed workflow to automate seismic phase detection, association, and classification. The process consists of three main stages:

1. Phase picking: P- and S-phase identification using the open-source toolkit SeisBench (Woollam et al., 2022) and a suite of pre-trained ML-based pickers (PhaseNet: Zhu and Beroza, 2019; EQTransformer: Mousavi et al., 2020).

2. Phase association: Grouping ML-predicted phases into events using GaMMA (Zhu et al., 2022).

3. Event classification: Assigning events to categories based on a spectral Frequency Index (FI) approach.

To identify the best-performing phase-picking models available in SeisBench, we conducted benchmarking tests using seismic waveforms recorded during a period of elevated activity (April–July 2019). Ten pre-trained configurations, combinations of PhaseNet and EQTransformer trained on various datasets, were evaluated against manually labeled P- and S-picks.

Performance assessment on this four-month dataset identified PhaseNet Volpick, EQTransformer Volpick, and PhaseNet Original as the top-performing models. All three exhibited high recall, exceeding 93% for P-phase picks and 76% for S-phase picks, based on INGV-OE VT events.

Following phase association with GaMMA, recall rates of about 84% were achieved for VT events. Only events detected by at least three independent stations (each providing both P and S picks) were retained, significantly reducing spurious associations.

Spectral classification was then performed using a modified Frequency Index (FI), defined as the logarithmic ratio between maximum spectral amplitudes in predefined frequency bands. This approach effectively differentiates LP, VT, low spectrum (LS), or Mixed types. In this study, the LS class corresponds to events with predominantly low-frequency spectral content and no detectable contribution from the summit seismic stations in the FI calculation. To calibrate the FI frequency bands, we analyzed mean amplitude spectra from 100 manually selected high-SNR VT and LP events. Based on the stacked spectra, we defined the low-frequency band (0.5–3.0 Hz) characteristic of LP events and the high-frequency band (3.0–20 Hz) characteristic of VT events.

Applying the three best-performing models over 1.5 years of data revealed that configurations trained on the Volpick dataset, particularly EQTransformer Volpick, consistently delivered superior results. However, periods of strong volcanic tremor reduced model accuracy, occasionally obscuring VT and LP signals. Across all models, a decline in detection and classification performance was observed beginning in September 2019, coincident with increased volcanic tremor activity at Etna (Sciotto et al., 2022). These tremors (1–5 Hz) overlap the frequency ranges of both VT and LP events, likely degrading SNR and impairing picking and classification accuracy. Notably, EQTransformer demonstrated greater robustness to such noise, maintaining high recall up to an FI size of 3.

Our findings (Carducci et al., 2026) highlight the importance of volcano-specific training datasets, such as Volpick, for improving the performance and generalizability of DL models in volcanic contexts. In particular, EQTransformer Volpick shows strong potential for monitoring, event association, and seismic signal classification at active volcanoes. Future integration of advanced spectral analysis techniques may further refine event-type discrimination and reduce classification ambiguities, contributing to improved understanding of volcanic seismicity (e.g., Zali et al., 2024; Abed et al., 2026).

The results presented are part of the following paper:

Carducci, A., Cocina, O., Sciotto, M., Cannata, A., Di Gioia, S., Vuan, A., Saraò, A., Tanaka Hernández, K., Sugan, M. (2026). Insights into a Deep Learning Workflow for Automatic Detection of Seismic Signals at Mount Etna. Submitted to Seismological Research Letters. Under review.

4.5 The Contribution of Geoelectrical Tomography to Hazard Reduction at Nirano Salse

4.5.1 Introduction

The Nirano Salse represents one of the most significant mud volcano fields in Italy, attracting over 50,000 visitors annually. However, these structures are inherent geohazards, capable of violent mud and hydrocarbon expulsions, soil degradation, and dangerous "quicksand" or liquefaction effects. By Electrical Resistivity Tomography (ERT) we provided a critical scientific foundation for reducing these risks by imaging the subsoil down to 250 meters for the first time.

4.5.2 Model results

ERT models (Fig. 4.5.1) provides several key contributions to local risk mitigation:

1. Identification of Localized Reservoirs:

ERT imaging identified shallow conductive zones (50–100 m deep) that act as fluid reservoirs. Crucially, the data revealed that these reservoirs are small and sparse, proving there is no massive, unstable "mud caldera" beneath the entire area. This finding limits the scope of potential catastrophic collapses.

2. Mapping Migration Conduits:

The study successfully mapped subvertical discontinuities, such as faults and fractures, which serve as the primary routes for rising fluids. By understanding exactly where these conduits are located—particularly in the more heterogeneous southern sector—authorities can predict where new vents or ground failures are most likely to occur.

3. Distinguishing Stability Zones:

The geophysical models highlight a stark contrast between the stable northern sector, characterized by regular stratification and multilayer aquifers, and the unstable southern sector, which is dominated by faults.

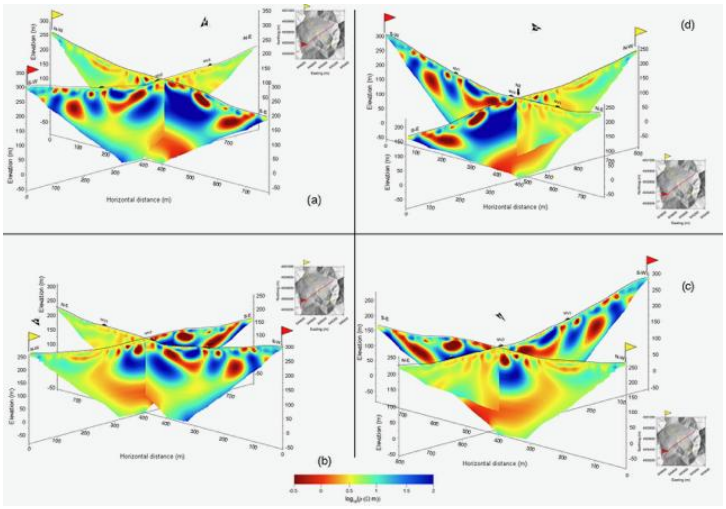


Figure 4.5.1 - The 3D arrangements of the tomographic results. Starting from panel (a), an anticlockwise rotation of the tomographic result is performed. The yellow and red flags on top of each section indicate the starting point of the geoelectrical surveys. The inset in each panel shows the location of the survey.

4.5.3 Impact

ERT model insights directly translate into public safety strategies. The ability to pinpoint the exact location of shallow reservoirs allows for fencing off high-risk areas prone to liquefaction and optimizing tourist routes to ensure they remain on stable ground.

In conclusion, ERT provides a non-invasive but highly effective way to characterize the "plumbing system" of mud volcanoes. By moving beyond surface observations to deep structural mapping, this methodology enables a targeted approach to hazard assessment, ensuring safer public access to this unique geological site.

4.6 Novel methodologies for processing of GNSS datasets

We defined and tested a methodology for the near-real-time identification of significant velocity variations, on large datasets of historical series of coordinates estimated from the processing of Global Navigation Satellite Systems (GNSS) measurements and Synthetic Aperture Radar (SAR) images, which may be representative of the triggering of deformational phenomena potentially associated with volcanic events.

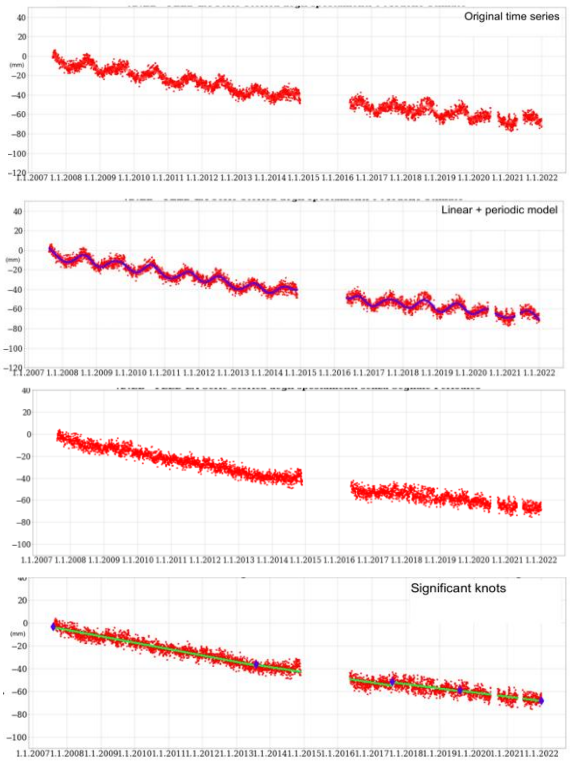
In particular, with reference to even very long time series (over 20 years) of GNSS coordinates, a pre-analysis and analysis strategy has been defined, implemented and tested, which considers temporal autocorrelations and identifies significant velocity variations, estimating a piecewise linear model linear (FEED 2.0 - Fast Elbow Effect Detection 2.0).

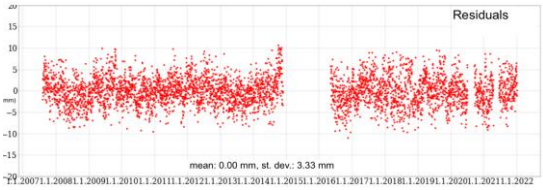
The obtained results have been published in:

Colosimo, G., Ravanelli, R., Mazzoni, A., Crespi, M., FEED - Fast Elbow Effect Detector 2.0, Kuaternion (Internal Report), 2022

4.6.1 Results.

FEED 2.0 is an original methodological approach designed and implemented to analyze complex time series in the context of geophysical monitoring. The algorithm is specifically optimized to process data from GNSS (coordinate sets or displacement data) and SAR (line-of-sight (LOS) or directional displacements, e.g., West-East and Up-Down components). FEED 2.0 is engineered to provide robust analysis under challenging data conditions. Its main capabilities include: (i) data integrity management (it handles time series characterized by outliers and significant interruptions); (ii) non-linear analysis (it is capable of modeling non-linear behaviors in the data); (iii) velocity variation detection (the algorithm identifies significant changes in velocity





FEED PP - 365 giorni			
Periodo		velocità [mm/y]	sigma [mm/y]
20/06/2002	- 20/06/2003	-6.65	0.53
20/06/2003	- 19/06/2004	-1.50	0.38
19/06/2004	- 19/06/2005	0.93	0.37
19/06/2005	- 19/06/2006	-2.92	0.39
19/06/2006	- 19/06/2007	0.67	0.39
19/06/2007	- 18/06/2008	-3.94	0.36
18/06/2008	- 18/06/2009	3.44	0.36
18/06/2009	- 18/06/2010	-6.28	0.36
18/06/2010	- 18/06/2011	0.52	0.34
18/06/2011	- 17/06/2013	-4.47	0.16
17/06/2013	- 17/06/2014	-3.64	0.33
17/06/2014	- 17/06/2015	-0.46	0.36
17/06/2015	- 16/06/2016	-3.08	0.36
16/06/2016	- 16/06/2017	0.29	0.35
16/06/2017	- 16/06/2018	-2.11	0.36
16/06/2018	- 16/06/2019	-4.77	0.31
16/06/2019	- 31/12/2021	-2.88	0.33

Figure 4.6.1: Example of GNSS time series modelled through FEED 2.0.

and estimates a piecewise linear model to represent these shifts); (iv) periodicity estimation (it detects and estimates annual periodic signals, allowing for amplitude changes over time); (v) precision estimation (it accounts for temporal autocorrelations to provide statistically correct velocity precisions).

FEED 2.0 was tested on several GNSS and SAR time series (an example in Figure 4.6.1) and it represents a comprehensive tool for high-precision monitoring of volcanic and earthquake-prone areas. By integrating GNSS and SAR data and employing advanced piecewise modeling with rigorous statistical validation, it provides a reliable means to detect subtle changes in Earth surface dynamics.

4.7 Machine learning tool for real-time monitoring solutions for fast and very rapidly evolving volcanic dynamics

Earthquakes represent a great scientific challenge, perhaps one of the most important of all time. They represent a massive natural hazard, with potential for catastrophic loss of life and colossal economic damage. While even a modest scientific advance could have a major impact on public safety, national economic health, and world-wide goals for energy transition, the state-of-the art for seismic hazard assessment and earthquake forecasting is limited. Part of the problem is that we have access to only too short and incomplete historical records rather than detailed models of earthquake physics and local measurements of fault zone stress state. Our work conducted for PE03 Return Spoke 3 VS3-DST - Progetto PE00000005 addresses these limitations.

Existing approaches for seismic hazard assessment are especially lacking in some settings, as we know all too well based on recent devastations in the 2023 Turkey-Syria earthquakes and, closer to home, the 2016 central Italy sequence of Amatrice/Norcia. While we are still far from predicting earthquakes, recent works provide an opportunity for transformational advancement. These studies demonstrate: 1) the ability to reproduce the full range of earthquake fault slip behaviors in controlled lab experiments under geophysical stresses, 2) the possibility to collect observations of foreshocks and precursory changes in fault zone elastic properties prior to laboratory earthquakes, 3) that artificial intelligence methods based on Machine Learning (ML) and Deep Learning (DL) can predict the timing of lab earthquakes, the fault zone stress state and in some cases the magnitude of impending labquakes, and 4) that the lab-based ML/DL techniques can be applied to tectonic faults in nature, at least under some conditions. In the laboratory, we have found that frictional stick-slip failure events – the lab equivalent of earthquakes – are preceded by a cascade of micro-failure events that radiate elastic energy in a manner that foretells impending catastrophic failure. Our recent work suggests that such signatures are also present for earthquakes associated with fluid-injection for energy recovery and storage, where seismic risk is often a limiting factor for progress.

4.7.1 Development of Foundational Models and LLM for Seismic Data.

AI methods are expanding at an incredible rate driven by increases in computational power and their success in diverse problems. As applied to earthquake physics, ML/DL techniques recognize patterns and subtle features that traditional approaches do not. Thus, it is not surprising that applications of AI to seismology have produced an explosion of interest. However, in many cases progress is hindered by the need for labeled data or task-specific datasets. The next wave of progress is likely to be driven by models that use unlabeled waveforms from openly available data. This would follow the ontogeny of LLMs for natural language processing or models for computer vision, which transitioned from task-specific datasets to general-purpose foundation models capable of handling a wide range of tasks with minimal fine tuning. Thus, we have developed such a model for application to seismic data.

4.7.2 Rapid Methods for Estimation of Earthquake Magnitude and Location.

We worked on LLM-based models and a foundation models for seismic data processing. The first step is an LLM model applied to Real Time Estimation of Earthquake (RTEE) source parameters (Bassani et al., 2025, 2026) and the foundation model is applied to seismic waveforms (Laurenti et al., 2026). Both approaches have clear applications for Earthquake Early Warning Systems (EEWS) because they can be fully automated and readily generalizable with minimal fine tuning.

For the RTEE work we adopted a pre-trained LLM (TinyLlama) and are using it to estimate earthquake magnitude and location (Figure 4.7.1). The LLM, TinyLlama is a compact LLM pretrained on > 1 trillion tokens. It is built on the architecture and tokenizer of Llama 2 using open-source, community input methods to improve computational efficiency. Despite its relatively small size, TinyLlama has produced impressive results. We fine-tuned it on a customized version of INSTANCE (The Italian Seismic Dataset for Machine Learning) and applied it to RTEE for earthquake magnitude and location.

Using LLMs and foundation models eliminates the need to develop and train tailor-made architectures. Our model using TinyLlama selects stations based on P-wave arrival times and uses the station location along with a single parameter extracted from the ground motion waveform: the peak ground velocity (PGV). We use PGV in a 0.21 s interval around the known P-wave arrival based on INSTANCE. Remarkably, the model is able to locate the events and estimate their magnitude with great accuracy (Figures 4.7.2 and 4.7.3). We do not use a crustal velocity model nor S-wave arrival time. In fact, our LLM approach uses only the station location, PGV, and P-wave arrival times in the form of the time difference between stations.

Our model compares favorably with automated methods used by the Italian National Institute of Geophysics and Volcanology (INGV), with similar performance in magnitude estimation and superior accuracy in epicenter, hypocenter and origin time prediction. Our LLM-based model achieves average errors of 6.3 km, 11.1 km, and 1.1 s for epicenter, hypocenter, and origin time estimation, respectively, in contrast to 8.6 km, 15.0 km, and 1.8 s for the INGV automatic solution. This represents an improvement of more than 26% and thus provides a strong direction for future work.

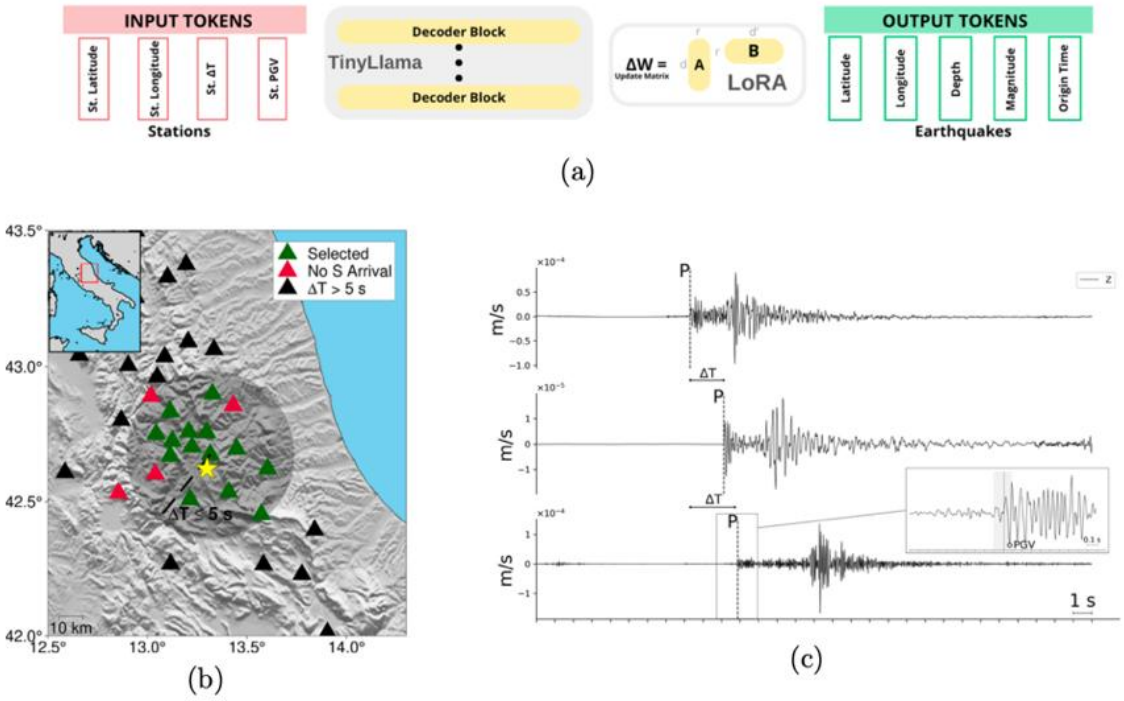


Figure 4.7.1: LLM model structure and workflow. >(a) Model summary; (b) Map of earthquake location (yellow star); c) Seismogram

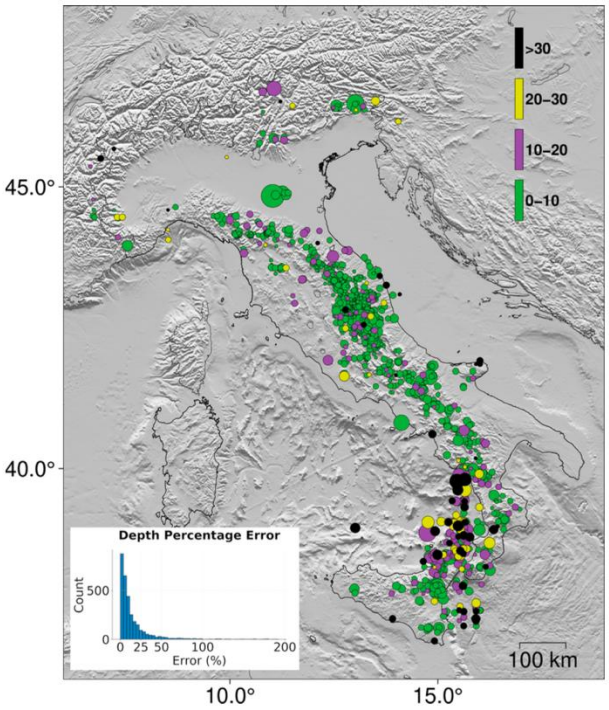


Figure 4.7.2: Map of earthquake epicentres determined by the LLM.

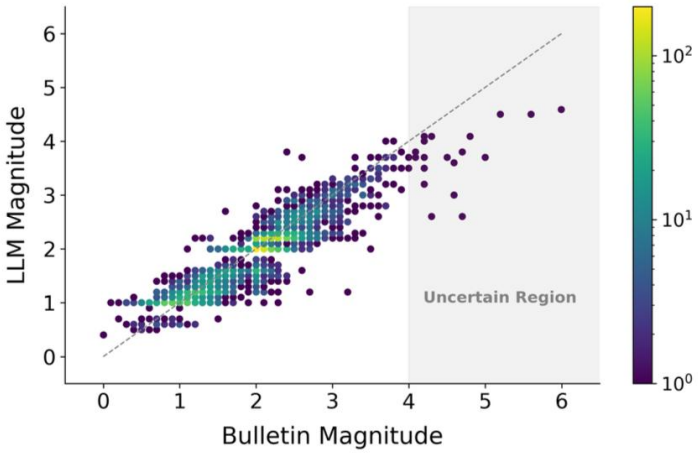


Figure 4.7.3: Comparison of INGV bulletin magnitudes with those estimated by the LLM

5. Conclusions

Research Activities in TK3.2.3 have identified a variety of new tools and methods, eventually using machine learning and artificial intelligence, for near-real time processing and analysis of monitored signals at volcanoes.

New statistical tools have been developed for processing volcanic gas datasets, and hence identify new composite indexes for the probabilistic quantification of the likelihood of a volcanic eruption. These techniques, when applied to volcanic gas records from Stromboli volcano, are now opening new avenues for the potential probabilistic prediction of otherwise unpredictable “major explosions”. New interpretative tools of volcanic unrest, based on the use of sulphur isotopes in volcanic gases and SO_2 flux records, have also been identified. New data processing tools have been developed to analyse and real-time interpret GNNS and electric resistivity datasets.

Monitoring seismicity is a key component of volcano surveillance systems. The identification of source patterns associated with specific volcanic processes contributes to a more reliable interpretation of unrest phases and improves the understanding of volcano–tectonic interactions. In this framework, full moment tensor analysis represents an advanced tool for characterizing fracture systems, fault activation, and possible fluid-related processes.

Within the project activities, we computed full moment tensors for the most energetic earthquakes ($3.4 \leq ML \leq 4.8$) that occurred in the Etna region between 2005 and 2022. The results indicate a clear dominance of the double-couple (DC) component during the investigated period, consistent with the activation of regional tectonic structures. In some cases, DC mechanisms may also be associated with eruptive fissure formation and rock fracturing processes. Conversely, events located along the Pernicana Fault system show a predominance of non-double-couple (non-DC) components, likely linked to fluid circulation.

The relationship between non-DC components and fluid processes is complex and requires further investigation. The results presented constitute the basis for the next phase of the project. Future activities will include a more detailed analysis of non-DC events and the integration of additional earthquakes into the dataset, with the objective of better constraining their role within the volcanic dynamics of the Etna area.

The study of advanced seismic event detection and classification – including template matching, unsupervised deep learning clustering, and deep learning cataloging – shows that integrating signal-processing and AI-based strategies marks a significant advancement in both scientific understanding and operational monitoring of active volcanoes.

Our analysis also highlights both the strengths and limitations of the different automated techniques applied to volcanic seismicity. Template matching proved highly effective in detecting repeating events, but its performance strongly depends on the careful selection of representative templates. In volcanic environments, building such a template dataset requires significant effort, largely because earthquake locations are often uncertain and signal variability is high. Spectrogram based clustering allowed us to identify groups of similar events, yet the method is inherently limited by the difficulty of linking these clusters to specific volcanic processes. Without robust qualitative and quantitative reference observations, assigning physical meaning to the clusters remains challenging. Finally, deep learning detection methods, particularly for distinguishing LP and VT events, require well designed, balanced training datasets that capture the full

range of possible seismic signals. In volcanic settings, where signal types overlap and noise conditions vary, this remains a substantial limitation.

Overall, while these automated approaches offer powerful tools for improving volcanic seismic monitoring, their effectiveness depends critically on high quality datasets, careful model design, and the integration of complementary geophysical observations. A multi method strategy, combining traditional time series analysis, unsupervised feature discovery, and deep learning catalogs, provides a robust path forward for operational volcano monitoring. The experience gained here demonstrates that such an integrated framework is both adaptable and transferable to any region where high resolution seismic and multi parametric data are available, offering significant potential for strengthening volcanic seismic monitoring and hazard assessment in complex volcanic environments.

A novel methodology was defined and tested for the near-real-time identification of significant velocity variations, on large datasets of historical series of coordinates estimated from the processing of Global Navigation Satellite Systems (GNSS) measurements and Synthetic Aperture Radar (SAR) images.

A LLM model has been applied to Real Time Estimation of Earthquake source parameters and a foundation model has been applied to seismic waveforms. Both approaches have clear applications for Earthquake Early Warning Systems (EEWS) because they can be fully automated and readily generalizable with minimal fine tuning.

A possible further improvement could be based on the integration of real-time GNSS-derived PGV through VADASE approach (Colosimo et al., 2011; Branzanti et al., 2013; Fratarcangeli et al., 2018; Ravanelli et al., 2021), both for epicenter and magnitude estimation (Crowell et al., 2013), benefiting from the resilience of GNSS permanent stations with respect to clipping, what allows PGV determination even at or very close to the epicenter.

6. References

Publications arisen from this task

Abed, W., Zali, Z., Sciotto, M., Cocina, O., Cannata, A., Picozzi, M., Martínez-Garzón, P., Vuan, A., Saraò, A., Sugan, M. (2026) Hidden Patterns in Volcanic Seismicity: Deep Learning Insights from Mt. Etna's 2020 - 2021 Activity. Scientific Reports 16, 6155 (2026). https://doi.org/10.1038/s41598-026-36677-x.

Aiuppa, A., Lo Bue Trisciuzzi, G., Alparone, S., Bitetto, M., Coltelli, M., Delle Donne, D., Ganci, G., Pecora, E. A SO2 flux study of the Etna volcano 2020–2021 paroxysmal sequences (2023) Frontiers in Earth Science, 11, art. no. 1115111, DOI: 10.3389/feart.2023.1115111

Aiuppa, A., Bitetto, M., Curcio, L., Delle Donne, D., Lages, J., Lo Bue Trisciuzzi, G., Tamburello, G., Vitale, A., Cannavo', F., Coltelli, M., Coppola, D., Innocenti, L., Insinga, L., Lacanna, G., Laiolo, M., Massimetti, F., Pistolesi, M., Privitera, E., Ripepe, M., Voloschina, M., Cilluffo, G. Volcanic gas changes prior to Stromboli's major explosions are statistically significant (2025) Journal of Volcanology and Geothermal Research, 462, art. no. 108325 DOI: 0.1016/j.jvolgeores.2025.108325

Aiuppa, A., Caudron, C., Chiodini, G., Ingebritsen, S., Viveiros, F. Geochemical Monitoring of Volcanic Fluids in the Twenty-First Century (2025b) Advances in Volcanology, Part F790, pp. 209 – 260 DOI: 10.1007/978-3-031-86841-2_8

Bassani, A., Trappolini, D., Poggiali, G., Tinti, E., Galasso, F., Marone, C., and A. Michelini, Real time estimation of earthquake location and magnitude using large language models, EGU, https://meetingorganizer.copernicus.org/EGU25/EGU25-16157.html, 2025.

Bassani, A., Trappolini, D., Poggiali, G., Tinti, E., Galasso, F., Michelini, A. and C. Marone, Adapting a pretrained llm to rapidly recover earthquake magnitude and location, J. Geophys. Res, 2026.

Branzanti M., Colosimo G., Crespi M., Mazzoni A.,GPS near-real-time coseismic displacements for the great tohoku-oki earthquake. IEEE Geoscience and Remote Sensing Letters, 10 (2), art. no. 6265361, pp. 372 - 376, 2013

Caliro, S., Chiodini, G., Avino, R., Carandente, A., Cuoco, E., Di Vito, M.A., Minopoli, C., Rufino, F., Santi, A., Lages, J., Mangiacapra, A., Monteleone, B., Pappalardo, L., Taracsák, Z., Tramati, C., Vizzini, S., Aiuppa, A. Escalation of caldera unrest indicated by increasing emission of isotopically light sulfur (2025) Nature Geoscience DOI: 10.1038/s41561-024-01632-w

Carbonari R., Salone R., De Paola C., Di Maio R., 2025. Predicting electrical resistivity in natural degassing geological systems through petrophysical and thermodynamic data: a machine learning approach. Pure and Applied Geophysics, https://doi.org/10.1007/s00024-025-03882-0

Carducci, A., Cocina, O., Sciotto, M., Cannata, A., Di Gioia, S., Vuan, A., Saraò, A., Tanaka Hernández, K., Sugan, M. (2026). Insights into a Deep Learning Workflow for Automatic Detection of Seismic Signals at Mount Etna. Submitted to Seismological Research Letters. Under review.

Cecere G., De Martino P., Riccardi U., Di Maio R., 2025. Evaluation of Trimble CenterPoint RTX correction service for Real-Time GNSS Monitoring: A Field-Based Comparison with RTK positioning. Discover Applied Sciences, 7, 1331, https://doi.org/10.1007/s42452-025-07889-z

Colosimo, G., Ravanelli, R., Mazzoni, A., Crespi, M., FEED - Fast Elbow Effect Detector 2.0, Kuaternion (Internal Report), 2022

Colosimo G., Crespi M., Mazzoni A., Real-time GPS seismology with a stand-alone receiver: A preliminary feasibility demonstration, Journal of Geophysical Research: Solid Earth, 116 (11), art. no. B11302, 2011.

Crowell, B. W., Bock, Y., & Melgar, D., Real-time magnitude scaling from GNSS-derived peak ground displacement and velocity. Geophysical Research Letters, 40(24), 6089-6094, 2013.

Delle Donne, D., Lacanna, G., Aiuppa, A., Bitetto, M., Ulivieri, G., Biagioli, F., Lo Bue Trisciuzzi, G., Ripepe, M. The beating sound of passive degassing at an open-vent volcano captured by combined infrasonic and SO2 flux observations (2026) Earth and Planetary Science Letters, 678, art. no. 119846, 10.1016/j.epsl.2026.119846

Fratarcangeli F., Ravanelli M., Mazzoni A., Colosimo G., Benedetti E., Branzanti M., Savastano G., Verkhoglyadova O., Komjathy A., Crespi M., The variometric approach to real-time high-frequency geodesy, Rendiconti Lincei, 29, pp. 95 - 108, 2018.

Laurenti, L., Johnson, C. W., Trappolini, D., Tinti, E., Galasso, F., and C. Marone, Testing audio compression autoencoders for seismology: moving toward foundation models, J. Geophys. Res, doi.org/10.1029/2025JH000787, 2026.

Lo Bue Trisciuzzi, G., Aiuppa, A., Salerno, G., Bitetto, M., Curcio, L., Innocenti, L., Lacanna, G., Lages, J.P., Lo Forte, F.M., Maugeri, S.R., Murè, F., Principato, P., Ripepe, M., Vitale, A., Delle Donne, D. Improved volcanic SO2 flux records from integrated scanning-DOAS and UV Camera observations. (2024) Journal of Volcanology and Geothermal Research, 455, art. no. 108207, DOI: 10.1016/j.jvolgeores.2024.108207

Insinga, L., Voloschina, M., Marianelli, P., Bartolomeo, E., Bertagnini, A., Métrich, N., Rotolo, S.G., Aiuppa, A., Ripepe, M., Pistolesi, M. Magma source, pre-eruptive dynamics and timescales of major explosions at Stromboli volcano (Italy) (2025) Bulletin of Volcanology, 87 (9), art. no. 75 DOI: 10.1007/s00445-025-01862-9

Ravanelli M., Occhipinti G., Savastano G., Komjathy A., Shume E.B., Crespi M., GNSS total variometric approach: first demonstration of a tool for real-time tsunami genesis estimation, Scientific Reports, 11 (1), art. no. 3114, 2021.

Romano, G., Antonellini, M., Patella, D., Siniscalchi, A., Tallarico, A., Tripaldi, S., and A. Piombo (2023). Fluid Conduits and Shallow Reservoir Structure Defined by Geoelectrical Tomography at the Nirano Salse (Italy). Natural Hazards and Earth System Sciences, 23 (8), pp. 2719 - 2735, DOI: 10.5194/nhess-23-2719-2023

Salone R., Carbonari R., De Paola C., Iezzi F., Di Maio R., 2025. Numerical modelling to identify permeable fractures from geophysical imaging of natural degassing areas. Example from the Matese Fault system (Italy). Tectonophysics, 917, 230950, https://doi.org/10.1016/j.tecto.2025.230950

Saraò, A., Moratto, L., Giampiccolo, E., & O Cocina, (2023). Moment magnitude for earthquakes in the Etna volcano area, Geophysical Journal International, 234, 2519- 2533, https://doi.org/10.1093/gji/ggad257

Soares, A., Custódio, S., Cesca, S., Silva, R., Vuan, A., and Mendes, V.B. (2023). The February 2018 seismic swarm in São Miguel, Azores. Front. Earth Sci. 11:1144151. https://doi.org/10.3389/feart.2023.1144151

Stix, J., de Moor, J.M., Aiuppa, A. Understanding and forecasting sudden explosive eruptions (2025) Bulletin of Volcanology, 87 (11), art. no. 99 DOI: 10.1007/s00445-025-01886-1

References for chapter 4.3

Cocina, O., Saraò, A., Cannata, A., Montalto, P.M., (2022). Full Moment Tensor Solutions for ML ≥ 3.5 Etna earthquakes from October 2018 to December 2020 (CMTC_2018-2020) (Version 1) [Data set]. Istituto Nazionale di Geofisica e Vulcanologia (INGV). https://doi.org/10.13127/cmtc/2018_2020

Giampiccolo, E., D'Amico S., Patanè D., Gresta, S., (2007). Attenuation and Source Parameters of Shallow Microearthquakes at Mt. Etna Volcano, Italy. Bulletin of Seismological Society of America, 97, 184-197. https://doi.org/10.1785/0120050252

Guidarelli, M., Saraò, A., Panza, G.F. (2002). Surface wave tomography and seismic source studies at Campi Flegrei (Italy). Physics of the Earth and Planetary Interiors, 134, 157-173. https://doi.org/10.1016/S0031-9201(02)00154-1

Guidarelli, M., Zille, A., Saraò, A., Natale, M., Nunziata, C., Panza, G.F. (2006). Chapter 6 - Shear-wave velocity models and seismic sources in Campanian volcanic areas: Vesuvius and Phlegraean fields. In: Dobran F. (Ed), Developments in Volcanology, Vesuvius Education, Security and Prosperity, Elsevier, Amsterdam, 287-309. https://doi.org/10.1016/S1871-644X(06)80010-0

Martínez-Arevalo, C., Patane, D., Rietbrock, A., Ibáñez, J.M., (2005). The intrusive process leading to the Mt. Etna 2001 flank eruption: constraints from 3D attenuation tomography. Geophysical Research Letters, 32, L21309. https://doi.org/10.1029/2005GL023736

Meyer, K., Biggs, J., Aspinall, W., (2021). A Bayesian reassessment of the relationship between seismic moment and magmatic intrusion volume during volcanic unrest. Journal of Volcanology and Geothermal Research, 419, 107375. https://doi.org/10.1016/j.jvolgeores.2021.107375

Minson, S., Dreger, D., (2008). Stable Inversions for Complete Moment Tensors. Geophysical Journal International, 174, 585–592. https://doi.org/10.1111/j.1365-246X.2008.03797.x

Panza, G.F. and Saraò, A., (2000). Monitoring volcanic and geothermal areas by full seismic moment tensor inversion: are non-double couple components always artefacts of modelling? Geophysical Journal International, 143, 353–364. https://doi.org/10.1046/j.1365-246X.2000.01250.x

Saraò, A., Panza, G.F., Privitera, E., Cocina, O., (2001). Non-double-couple mechanisms in the seismicity preceding the 1991–1993 Etna volcano eruption. Geophysical Journal International, 145, 319–335. https://doi.org/10.1046/j.1365-246x.2001.01375.x

Saraò, A., Cocina, O., Privitera, E., Panza, G.F., (2010). The dynamics of the 2001 Etna eruption as seen by full moment tensor analysis. Geophysical Journal International, 181, 951–965. https://doi.org/10.1111/j.1365-246X.2010.04547.x

Saraò, A., Cocina, O., Moratto L., Scarfì L., (2016). Dynamics and kinematics of the eastern flank from seismological and ground deformation analyses (Etna). Earthquake features through the seismic moment tensor. In Miscellanea INGV (Vol. 29, pp. 1–172). INGV Istituto Nazionale di Geofisica e Vulcanologia. https://doi.org/10.5281/zenodo.5774162

Saraò, A., Moratto, L., Giampiccolo, E., & Cocina, O. (2023). Moment magnitude for earthquakes in the Etna volcano area. Geophysical Journal International, 234, 2519- 2533. https://doi.org/10.1093/gji/ggad257

Siniscalchi, A., Tripaldi, S., Neri, M., Giammanco, S., Piscitelli, S., Balasco, M., Behncke, B., Magrì, C., Naudet, V., Rizzo, E., (2010). Insights into fluid circulation across the Pernicana Fault (Mt. Etna, Italy) and implications for flank instability. Journal of Volcanology and Geothermal Research, 193, 137-142. https://doi.org/10.1016/j.jvolgeores.2010.03.013

References for chapter 4.4

Abed, W., Zali, Z., Sciotto, M., Cocina, O., Cannata, A., Picozzi, M., Martínez-Garzón, P., Vuan, A., Saraò, A., Sugan, M. (2026). Hidden Patterns in Volcanic Seismicity: Deep Learning Insights from Mt. Etna's 2020–2021 Activity. Scientific Reports Scientific Reports 16, 6155 (2026). https://doi.org/10.1038/s41598-026-36677-x

Alparone S. C., Barberi G., D'Amico S., Di Grazia G., Ferrari F., Giampiccolo E., Maiolino V., Mostaccio A., Musumeci C., Scaltrito A., Scarfì L., Sciotto M., Spampinato S., Tusa G., Tuvè T., Ursino A., Zuccarello L. (2020). Mt. Etna Revised and Concise Seismic Catalog from 1999 (EtnaRCSC) [Data set]. Istituto Nazionale di Geofisica e Vulcanologia (INGV). https://doi.org/10.13127/etnasc/etnarcsc

Barberi, G., Di Grazia, G., Ferrari, F., Firetto Carlino, M., Giampiccolo, E., Maiolino, V., Mostaccio, A., Musumeci, C., Scaltrito, A., Sciotto, M., Tusa, G., Tuvè, T., Ursino, A. (2020). Mt. Etna Revised Seismic Catalog from 2020 (EtnaRSC2020) (Version 1) [Data set]. Istituto Nazionale di Geofisica e Vulcanologia (INGV). https://doi.org/10.13127/ETNASC/ETNARSC2020

Calinski, T., Harabasz, J. (1974). A dendrite method for cluster analysis. Communications in Statistics: Theory and Methods, 3, 1–27.

Campanella, S. (2022). EQTM. Zenodo. https://doi.org/10.5281/zenodo.7140418

Cannata, A., Di Grazia, G., Aliotta, M., Cassisi, C., Montalto, P., Patanè, D. (2013). Monitoring seismo-volcanic and infrasonic signals at volcanoes: Mt. Etna case study, Pure Appl. Geophys. 170, 1751–1771. https://doi.org/10.1007/s00024-012-0634-x

Cannata, A., Spedalieri, G., Behncke, B., Cannavò, F., Di Grazia, G., Gambino, S., Gresta, S., Gurrieri, S., Liuzzo, M., Palano, M. (2015). Pressurization and depressurization phases inside the plumbing system of Mount Etna volcano: Evidence from a multiparametric approach. Journal of Geophysical Research: Solid Earth, 120. https://doi.org/10.1002/2015JB012227

Cannata, A., Di Grazia, G., Sciotto, M. (2025). Catalog of Long Period seismic events recorded at Mt. Etna during the time interval 2019-2020 (LPSEC2019_2020), Istituto Nazionale di Geofisica e Vulcanologia (INGV). https://doi.org/10.13127/ETNA/ELPC/2019_2020.

Cawley, G. C., Talbot, N. L. C. (2010). On over-fitting in model selection and subsequent selection bias in performance evaluation. Journal of Machine Learning Research, 11, 2079–2107.

Davies, D. L., Bouldin, D. W. (1979). A cluster separation measure. IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-1, 224–227. https://doi.org/10.1109/TPAMI.1979.4766909

Gibbons, S. J. (2022). The optimal correlation detector? Geophysical Journal International, 228(1), 355–365. https://doi.org/10.1093/gji/ggab344

Hinton, G. E., Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. Science, 313, 504–507. https://doi.org/10.1126/science.1127647

Kato, A., Obara, K., Igarashi, T., Tsuruoka, H., Nakagawa, S., Hirata, N. (2012). Propagation of slow slip leading up to the 2011 Mw 9.0 Tohoku-Oki earthquake. Science, 335, 705–708. https://doi.org/10.1126/science.1215141

Krischer, L., Megies, T., Barsch, R., Beyreuther, M., Lecocq, T., Caudron, C., Wassermann, J. (2015). ObsPy: A bridge for seismology into the scientific Python ecosystem. Computational Science & Discovery, 8, 014003. https://doi.org/10.1088/1749-4699/8/1/014003

Mousavi, S. M., Zhu, W., Ellsworth, W., Beroza, G. (2019). Unsupervised clustering of seismic signals using deep convolutional autoencoders. IEEE Geoscience and Remote Sensing Letters, 16, 1693–1697. https://doi.org/10.1109/LGRS.2019.2909218

Mousavi, S.M., Ellsworth, W.L., Zhu, W. et al.(2020). Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. Nat Commun 11, 3952. https://doi.org/10.1038/s41467-020-17591-w

Peng, Z., Zhao, P. (2009). Migration of early aftershocks following the 2004 Parkfield earthquake. Nature Geoscience, 2, 877–881. https://doi.org/10.1038/ngeo697

Sciotto, M., Cannata, A., Di Grazia, G., Montalto, P. (2022). Volcanic tremor and long period events at Mt. Etna: Same mechanism at different rates or not?, Phys. Earth Planet. Inter., 324, 106850. https://doi.org/10.1016/j.pepi.2022.106850

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

Soares, A., Custódio, S., Cesca, S., Silva, R., Vuan, A., Mendes, V. B. (2023). The February 2018 seismic swarm in São Miguel, Azores. Frontiers in Earth Science, 11, 1144151. https://doi.org/10.3389/feart.2023.1144151

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. https://doi.org/10.1785/0120170106

Wassermann, J. (2012). Volcano seismology. In New Manual of Seismological Observatory Practice 2 (NMSOP-2) (pp. 1–77). GFZ.

Woollam, J., Münchmeyer, J., Tilmann, et al. (2022). SeisBench - A toolbox for machine learning in seismology. Seismological Research Letters, 93(3), 1695-1709. https://doi.org/10.1785/0220210324

Yang, H., Zhu, L., Chu, R. (2009). Fault-plane determination of the 18 April 2008 Mount Carmel, Illinois, earthquake by detecting and relocating aftershocks. Bulletin of the Seismological Society of America, 99, 3413–3420. https://doi.org/10.1785/0120090038

Zali, Z., Mousavi, S. M., Ohrnberger, M., Eibl, E. P. S., Cotton, F. (2024). Tremor clustering reveals pre-eruptive signals and evolution of the 2021 Geldingadalir eruption of the Fagradalsfjall Fires, Iceland. Communications in Earth & Environment, 5, 1. https://doi.org/10.1038/s43247-023-01166-w

Zali, Z., Martínez-Garzón, P., Kwiatek, G. et al. (2025). Low-frequency tremor-like episodes before the 2023 MW 7.8 Türkiye earthquake linked to cement quarrying. Scientific Reports, 15, 6354. https://doi.org/10.1038/s41598-025-88381-x

Zhang, M., Wen, L. (2015). An effective method for small event detection: Match and locate (M&L). Geophysical Journal International, 200, 1523–1537. https://doi.org/10.1093/gji/ggu466

Zhu, W., Beroza, G.C. (2019). PhaseNet: A deep-neural-network-based seismic arrival-time picking method. Geophysical Journal International, 216(1), 261–273. https://doi.org/10.1093/gji/ggy423

Zhu, W., McBrearty, I. W., Mousavi, S. M., Ellsworth, W. L., Beroza G.C. (2022). Earthquake Phase Association Using a Bayesian Gaussian Mixture Model, J. Geophys. Res. Solid Earth, 127, no. 5, e2021JB023249. https://doi.org/10.1029/2021JB023249