UNIVERSITÀ DEGLI STUDI DI PALERMO

Analysis of the spatial distribution and temporal variation of the Completeness Magnitude and b-value in the Gutenberg-Richter law

PhD Thesis

XXXVII Cycle
Department of Earth and Marine Sciences

Author:
Anna Figlioli

Supervisor:
Raffaele Martorana

Co-supervisors:
Antonino D'Alessandro
Attilio Sulli

a.c 2021-2024

Abstract

This Ph.D. thesis presents a detailed study the spatial distribution and temporal variation of the Magnitude of Completeness (Mc) and *b*-value parameters, fundamental to the Gutenberg-Richter Distribution (GR) within a seismically active region. The main objective is to provide a nuanced understanding of local seismicity through statistical analysis of recorded earthquake data. The research begins with a rigorous construction and validation of seismic catalogues, ensuring a robust foundation for further analysis. Mc, which defines the minimum magnitude at which events are reliably recorded, is a critical parameter for catalogue accuracy and is studied here with advanced estimation techniques, such as Maximum Curvature and other tests. These methods allow for a precise delineation of Mc across temporal and spatial domains, identifying completeness thresholds critical for seismic modelling. The second focus, the *b*-value, reflects the relative frequency of smaller to larger earthquakes and is essential for assessing regional tectonic stress and seismic hazard potential. Through advanced statistical analysis techniques and computational methods, including Maximum Likelihood Estimation (MLE) and *b*-positive, this study evaluates *b*-value variations across different geodynamic settings. A comprehensive comparative analysis examines the Mc and *b*-value across Italy, Taiwan, and Iceland, each characterized by unique tectonic activities. Italy's complex plate interactions, Taiwan's convergent boundary dynamics, and Iceland's mid-ocean ridge environment highlight how local geodynamic conditions shape seismic patterns. The study further applies the Epidemic Type Aftershock Sequence (ETAS) model to distinguish mainshocks from aftershocks, enhancing the understanding of clustering patterns within the seismic data of Sicily. This application supports the exploration of spatiotemporal seismic behaviour and correlates with regional geodynamic factors. Furthermore, this thesis introduces the *Tremors* software tool developed to streamline the calculation and visualization of the Mc and *b*-value, particularly addressing the challenges posed by short-term aftershock incompleteness. This innovative tool not only aids in effective data processing, but also adapts to varied seismic conditions, contributing a valuable resource for both academic and applied seismology. This research provides a comprehensive framework for applied seismology, advancing the understanding of seismicity in complex tectonic regions. Its findings have implications for seismic risk assessment and offer foundational insights for future seismological research.

Dedication

To myself,
for having taken this long course with determination, even when the challenges seemed impossible. I chose not to give up, and step by step I found the strength to continue believing in myself and in my dream.

Every experience, from the moments of solitary study to the satisfactions gained at conferences, has contributed to my personal and professional growth. Here I am today, in front of a goal that I have always wanted: to obtain a Ph.D. And this is only the beginning, with the hope of a future in research.

This thesis represents not only my achievements, but also the path that has made me the person I am today. To me, for my passion, perseverance, and courage to always go forward.

"There is no lift to success, you have to take the stairs." cit. *Zig Ziglar*

To my four-legged love,
I have dedicated nothing to him until now. But I do feel like dedicating this achievement to her, together with her I have supported everything from the first exam to the first submission of a manuscript. She has been with me every step of the way.

I thank you for everything, my beloved Peggy.

Acknowledgements

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The study of seismicity has long been a fundamental aspect for understanding Earth's dynamic processes. One of the core principles guiding this field is the Gutenberg-Richter law, which describes the statistical distribution of earthquake magnitudes. This thesis, titled "Study of Spatial Distribution and Temporal Variation of the Completeness Magnitude and the b-value of the Gutenberg-Richter Law" aims to provide a comprehensive analysis of these critical parameters, utilizing an extensive seismic catalogue as a foundational resource.

The completeness magnitude and the b-value serve as essential indicators for assessing seismic hazard and understanding the underlying mechanics of earthquake occurrence. This research begins with a detailed examination of a curated seismic catalogue, allowing for the identification of spatial and temporal patterns in seismic activity. By integrating various databases and employing advanced statistical techniques, this study seeks to clarify the factors influencing both the b-value and completeness magnitude across different geological settings.

In addiction, this work introduces innovative software tools and methodologies specifically designed for seismic analysis, which will enable more accurate calculations and interpretations of the Gutenberg-Richter Law parameters, yielding deeper insights into the data. Through rigorous analysis and the application of these new tools, this thesis aims to broaden our understanding of seismic behaviour and contribute to the ongoing discourse on earthquake prediction and risk assessment.

Ultimately, this study not only refines current methodologies, but also lays the groundwork for future research in seismology, providing a robust framework for analysing seismic data and addressing the complexities of seismic phenomena.

1.1 Research background

Earthquakes, a powerful and often devastating geological phenomenon, have left an indelible mark on the surface of our planet and throughout human history. They serve as a stark reminder of the dynamic nature of our Earth, constantly reshaping landscapes, influencing geological processes, and having a profound impact on human civilizations. In the incessant movement of tectonic plates which is the basis of the geological evolution of our planet, earthquakes are evidence of the dynamism of our planet.

An earthquake is a rapid movement of the earth's crust that is induced by the release of energy accumulated over time. The consequences of such seismic events can range from minor vibrations that are difficult to perceive to catastrophic impacts on human life. The defining characteristic of an earthquake lies in its unpredictable nature, an event that can strike with little warning and with sometimes devastating consequences.

Understanding these fault types is essential in comprehending the mechanics of

earthquakes. Throughout history, the world has witnessed some of the most catastrophic earthquakes that have left an indelible mark on human society. From the iconic 1906 San Francisco earthquake to the 1964 Alaska earthquake to the 2004 Indian Ocean earthquake and tsunami, the world has seen the raw power of seismic events.

In Italy, a country geologically located on the edge of the Eurasian and African tectonic plates, earthquakes have had a profound impact on its history, with events such as the 1908 Messina earthquake, or 2006 in Aquila and the 2016 central Italy earthquakes remaining etched in the collective memory.

To mitigate the effects of an earthquake, seismic risk is taken into account. Seismic risk is defined as the product of vulnerability, exposure, and seismic hazard.

$$Seismic\,Risk = Seismic\,Hazard \times Seismic\,vulnerability \times Exhibition$$

Seismic hazard (P): coincides with the seismicity of the place. It represents the probability that earthquakes of a certain size will occur in a specific area and in a specific period.
Seismic vulnerability (V): measures the disposition of a building to suffer damage due to an earthquake and its capacity or resistance to seismic actions.
Exhibition (E): focused on the protection of human life, it hypothesizes the set of human lives and material resources that can be lost as a result of the earthquake.

Seismic risk assessment is a critical aspect of earthquake studies, involving the evaluation of the probability of ground shaking in specific locations over a defined period. The seismic risk formula combines the likelihood of ground shaking with the intensity of ground motion, allowing a quantification of the risk faced by a particular region. This approach is especially relevant in a country such as Italy, where the seismic hazard is considerable due to its complex tectonic setting.

Italy, situated at the convergence of the African and Eurasia tectonic plates, is one of the most seismically active regions in Europe. Over the centuries, it has experienced numerous powerful earthquakes that have significantly shaped the country's history, infrastructure, and response strategies. The regions most affected by seismic activity include the Apennine Mountains, Sicily, and areas of northern Italy, such as Friuli and Veneto. These zones have been the epicenter of some of the deadliest earthquakes in Italian history, many of which have had long-lasting social and economic impacts.

One of the most catastrophic events was the 1908 Messina and Reggio Calabria earthquake, which had an estimated magnitude of 7.1. It caused widespread destruction, killing between 80,000 and 100,000 people and triggering a deadly tsunami. This disaster highlighted the need for improved building practices and disaster preparedness, leading to the first significant seismic regulations in Italy. Similarly, the 1980 Irpinia earthquake (magnitude 6.9) caused over 2,700 deaths and extensive damage in southern Italy. The slow response to this disaster spurred reforms in Italy's civil protection system, leading to the creation of modern emergency management protocols.

Earlier, the 1915 Avezzano earthquake (magnitude 7.0) devastated the central region of Marsica, leaving around 30,000 dead. This event remains one of the deadliest in Italy's history and emphasized the necessity for better urban planning and construction standards in seismic zones. The 1976 Friuli earthquake (magnitude 6.5), which struck northern Italy, caused approximately 1,000 deaths, but its aftermath is often regarded as a success story in terms of recovery and rebuilding, with rapid and efficient reconstruction efforts guided by the local communities.

In more recent times, the 2009 L'Aquila earthquake (magnitude 6.3) and the 2016-2017 Central Italy earthquake sequence (with magnitudes between 5.5 and 6.5) demonstrated that despite advances in seismic technology, the country remains vulnerable. These events caused significant damage to cultural heritage sites and highlighted the ongoing challenge

of preparing for future seismic events.

Italy's seismic history is well-documented in catalogs like the CPTI15 (Parametric Catalogue of Italian Earthquakes) and the DBMI15 (Italian Macroseismic Database). These resources compile historical records dating back over a millennium, offering detailed information on past earthquakes, their magnitudes, and their impacts.

Furthermore, these historical earthquakes have driven advancements in Italy's seismic building codes, which have evolved over time to integrate lessons learned from past disasters. Today, seismic risk models not only incorporate geological and seismological data but also benefit from the rich historical catalogue of Italian earthquakes. This has led to improved predictions of ground shaking, guiding better construction practices and urban planning in high-risk areas.

Italy's seismic risk assessment efforts are deeply rooted in its rich seismic history. The frequent and often devastating earthquakes the country has experienced over the centuries have shaped both its physical landscape and its approach to disaster mitigation. The integration of historical knowledge with modern science has proven vital in reducing future risks, safeguarding both lives and the nation's cultural heritage.

1.2   Scope of the research

In the field of seismology, the application of the Gutenberg-Richter law is a fundamental pillar for understanding the seismicity of a specific region. This law provides a statistical description of the distribution of magnitudes of seismic events, which is essential for assessing the seismicity of a specific area over time.The use of a seismic catalogue as a basic data set plays a central role in this context. These catalogues, compiled through years of seismic monitoring, collect detailed information on each recorded event, including parameters such as magnitude, depth and geographical location.

The accuracy and comprehensiveness of such catalogues are crucial for a correct application of the Gutenberg-Richter law, allowing the magnitude of completeness (Mc) and b-value to be reliably identified.In this context, this study aims to explore the importance of using reliable seismic catalogues as a database, underlining their direct influence on the quality and robustness of analyses of the spatial distribution and temporal variations of key seismic parameters.

The seismic catalogue will be the main topic of this thesis. The potential of a seismic catalogue, which can be real or synthetic, will be analysed. A seismic catalogue to be a good catalogue should be complete. In real catalogues, this completeness is missing. The effects of this incompleteness will be studied. So one will study the performance of a seismic network, what its areal coverage is, why it represents spatial gaps, whether it collects good quality data based on seismic noise. It will study seismic sequences and how they affect the seismic catalogue and the detection of seismic events of magnitude smaller than 3. Whether seismic sensors become saturated following a seismic event of magnitude greater than 4. This saturation phenomenon depends on several factors, such as: the proximity to the epicentre of the earthquake, i.e. how far away a seismic station may be from the seismic event. Another factor could be related to the type of sensor being considered, generally velocimeters rather than accelerometers have this problem.

In the literature to carry out analyses for magnitude of completeness and b-value we find techniques, now known worldwide, such as MAXC,MBASS, GFT [42]. But it is possible to find, which it will apply to this study, statistical analyses based on probabilistic calculations to understand how the magnitude of completeness varies over time as the performance of seismic networks increases; or stochastic analyses such as ETAS [12] or point process method [63].

Another useful analysis to understand the b-value parameter, in particular, is to study

time series. Those can provide valuable feedback on the variations of the parameter. The techniques adopted for this calculation are: maximum likelihood (MLE) [2], the b positive [18]. This parameter will also be studied spatially. A b-value map can provide information on the variation of the geodynamics of a study area.

For the analysis of the magnitude of completeness and b-value, representative geographical areas were selected, such as Italy, Taiwan, Iceland and, in particular, Sicily. These regions present a wide range of seismic activity, from frequent medium to low-magnitude earthquakes in Italy and Sicily, to larger and more frequent earthquakes in Taiwan and Iceland. This allows the magnitude of completeness and b-value to be analysed in different geodynamic scenarios. Namely, Italy and Sicily are located along the boundary between the African and Eurasian plates, while Taiwan is located along the boundary between the Eurasian and Philippine plates, and Iceland is situated over the mid-Atlantic ridge. These areas have a large quantity of historical and recent seismic data, thanks to well-developed seismic networks. The quantity and quality of available data are crucial for accurate analysis.

The figure n. 1.1 illustrates the workflow of a PhD thesis titled "Study of spatial distribution and temporal variation of the Completeness Magnitude and the b-value of the Gutenberg-Richter law."

The process begins with the collection of real data, specifically earthquake catalogues, followed by a preliminary analysis where the data is filtered and organized. Next, the Gutenberg-Richter law is applied to understand the frequency distribution of earthquakes, leading to the study of the spatial and temporal distribution of the completeness magnitude (Mc) and the b-value.

The final stage focuses on the results obtained from both real and synthetic data. These results are applied to forecasting potential seismic activity, creating risk maps, and evaluating the performance of the seismic network in monitoring and predicting earthquakes. Ultimately, the workflow outlines a comprehensive approach to understanding seismic behaviour through both empirical and simulated data, with the goal of improving predictive capabilities and risk assessment.
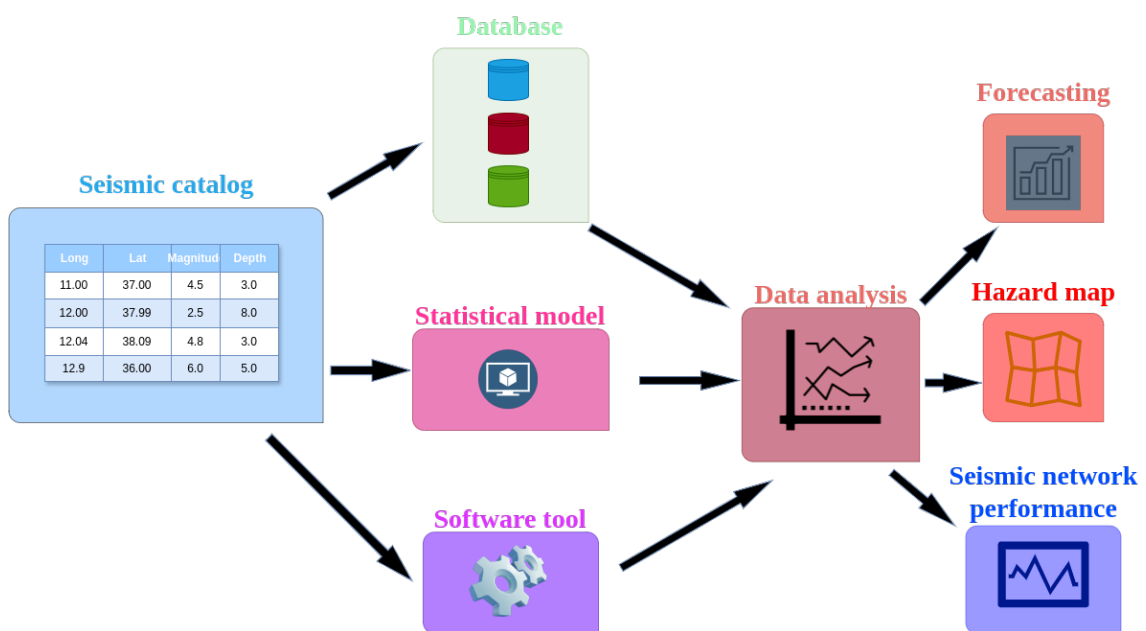
Figure 1.1: Doctoral thesis summary chart

Chapter 2

Statistical techniques and models

This chapter constitutes the core of the thesis, focusing on the main themes of the research. In particular, on the Gutenberg-Richter law, a fundamental principle of seismology, developed by the famous scientists Gutenberg and Richter [25] [29], this law provides a crucial framework for understanding the frequency and distribution of earthquakes of different magnitudes within a specific geographical region. Two key parameters, that are essential to characterize the seismicity of any area, will be examined in detail: the magnitude of completeness and the b-value.

In seismology, the use of statistical models is fundamental for interpreting seismic phenomena, improving earthquake forecasting, and assessing seismic risk. These models offer an essential analytical framework to handle the complexity of seismological data and reveal relationships between key seismic variables. As such, the application of statistical methodologies represents a cornerstone of modern seismic research.

This chapter explores the application of statistical models to seismology, emphasizing the methodologies and techniques developed to tackle the unique challenges presented by seismic data. Starting from the basics of applied statistics, we will explore how statistical models are implemented to analyse earthquake distributions, predict magnitudes, and model seismic risks. Furthermore, we will discuss critical steps in developing these models, such as selecting relevant seismological variables, estimating model parameters, and validating results.

For this purpose, two statistical models are investigated in depth: the Epidemic Type Aftershock Sequence (ETAS) model [50] [51] and the Probability Magnitude Completeness (PMC) model [61]. The ETAS model helps to understand the temporal clustering of earthquakes, specifically aftershock sequences, while the PMC model addresses the estimation of the minimum magnitude at which seismic activity is fully detectable in a region. Together, these models provide powerful tools for analysing seismicity and improving earthquake forecasting in the study region.

2.1 Gutenberg-Richter law

The Gutenberg-Richter relationship ,[25][29], is an empirical law of seismology that expresses the relationship between the number of earthquakes per unit times greater than a given magnitude for a defined volume. The first definition was proposed by Gutenberg and Richter in 1944 [25] in which this law is verified on a global scale for a complete catalogue. The application of this law is fundamental for processing and interpreting a seismic database.

The GR law models the relationship between earthquakes magnitudes and their frequencies in a given area. It states that the logarithm of the number of earthquakes greater than or equal to a given magnitude is inversely proportional to that magnitude. This relationship is quantified by a characteristic parameter known as the b-value, which

represents the relative frequency of small to large earthquakes [29]. Mathematically. it can be expressed as,[48]:

$$\log(N) = a - b(M - M_{\min}) \tag{2.1}$$

where:

- N represents earthquakes with a magnitude greater than or equal to a specific magnitude (M).

- a is the intercept parameter of the theoretical line in the semilogarithmic graph

- b is the slope of the theoretical line in the semilogarithmic graph, representing the relationship between large earthquakes and small earthquakes.

- *Mmin* is the magnitude of completeness of the catalogue [53].

Assuming that establishes a linear relationship between the earthquake magnitude M and the logarithm of the cumulative number of earthquakes with magnitudes greater than M in a given region, it is feasible to correctly estimate key parameters such as the b-value, a-value, and completeness magnitude. The completeness magnitude *Mc* is defined as the minimum magnitude above which all earthquakes within a specified area are consistently recorded. This parameter, also referred to as the limiting magnitude threshold, is theoretically characterized as the lowest magnitude at which 100% of the earthquakes within a defined space-time volume are detected [43].

Graphically, the b-value represents the slope of the curve derived from the frequency-magnitude distribution (FMD), as show figure n. 2.1. A higher b-value indicates a relatively larger proportion of smaller earthquakes, while a lower b-value suggests a greater occurrence of larger earthquakes. The a-value, on the other hand, governs the overall number of events in a given dataset. It represents the seismic activity rate in a region, with higher a-values corresponding a higher productivity.
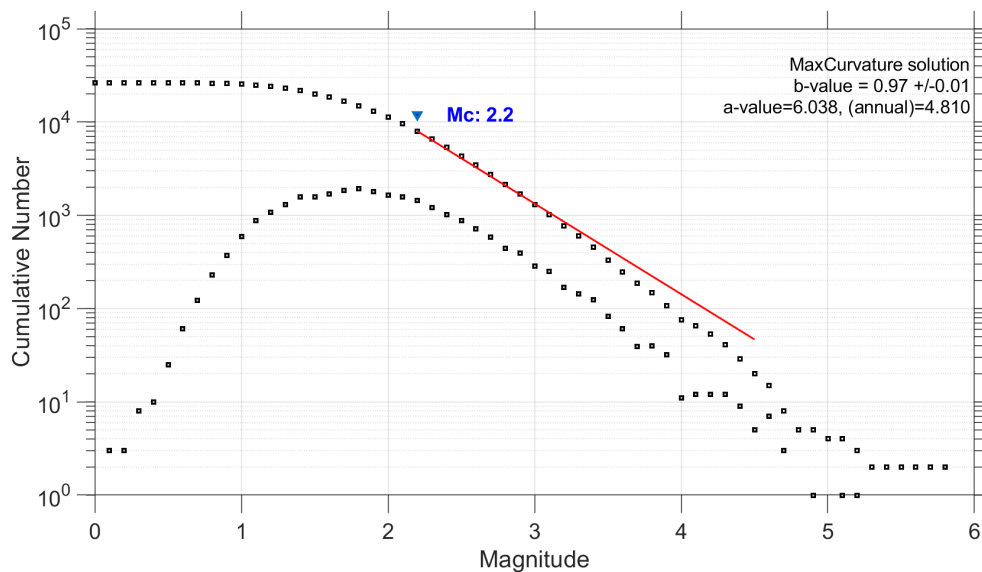
Figure 2.1: Frequency magnitude distribution of case study, Sicily. Range from 2005 to 2022

A series of new sophisticated calculations [64–66] [21] have been developed over the past two decades, starting with GR's law. The focus is mainly falling on the magnitude of completeness and the b-value. There are many techniques to calculate for the completeness

magnitude, the most used are the following: MAXimum Curvature(MAXC) [74], Mc by B-value Stability (MBS) [11], Goodness of Fit Test (GFT) [74], Entire Magnitude Range (EMR) [75], Median-Based Analysis of the Segment Slope (MBASS) [5]. These techniques are based on different algorithms. In recent years, further techniques have been developed that could make computing evolve in a much more sophisticated way [64] [21]. The estimating of *b*-value is important for the scientific community [23] [22] [67]. It is thought that the *b*-value has many significant interpretative ways out, among these there is the opportunity to use it as a precursor of possible seismic sequences [36]. Other application of *b*-value is to research the place of magma storage [47]; it has been correlated to focal mechanism [36].

2.1.1 Evaluation of the Computational Methods for Declustering in Seismic Catalogues

In seismic research, catalogues often contain both independent seismic events and related sequences, such as aftershocks or earthquake swarms. These related events, if left unfiltered, can mask the true seismic patterns and lead to skewed statistical interpretations. To address this, declustering techniques are employed to remove these correlated events and isolate the independent earthquakes. This process is crucial for improving the accuracy of seismic analyses, especially when applying models like the Gutenberg-Richter law, which relies on the assumption of independent event distributions.

Declustering enables seismologists to focus on primary seismic events, providing a clearer view of the earthquake-generating process and improving the predictive models used in hazard assessment.

The Spatially Inhomogeneous Temporally Homogeneous Poisson (SITHP) model is a statistical framework frequently used in seismology to describe earthquake distributions after declustering [37]. The SITHP model assumes that earthquakes occur at a constant rate over time, following a Poisson process, while acknowledging that their spatial distribution is non-uniform. This spatial clustering reflects geological features, such as fault lines, where seismic activity is concentrated. Declustering allows the SITHP model to more accurately represent the independent seismic events by eliminating secondary events that could otherwise lead to an overestimation of earthquake frequency in specific regions. By separating the main events from aftershocks, researchers can better analyse seismic risks and model future earthquake behaviour.

To confirm that a declustered seismic catalogue conforms to the assumptions of the SITHP model, various statistical tests are applied. These tests primarily focus on the temporal independence of seismic events, ensuring that the declustered catalogue follows a Poisson process over time.

Statistical testing techniques include the following:

The Kolmogorov-Smirnov (KS) test is a non-parametric method used to compare the observed inter-event times in the declustered catalogue with a theoretical exponential distribution, which characterises a Poisson process. By assessing the maximum difference between the observed and expected distributions, the KS test helps determine whether the declustered catalogue exhibits the temporal randomness expected under the Poisson assumption. If the test suggests a good fit, it indicates that the declustering process has successfully removed temporally correlated events, leaving behind independent seismic occurrences.

The chi-squared test offers an additional method for assessing temporal independence. It divides the time intervals into discrete bins and compares the observed frequency of events in each bin with the expected frequency under a Poisson distribution. This test is particularly useful for identifying deviations from the expected inter-event times, which could indicate residual clustering in the catalogue. A strong agreement between

observed and expected frequencies supports the conclusion that the declustered events are independent and follow a Poisson process. While temporal independence is crucial, spatial clustering must also be evaluated to confirm the assumptions of the SITHP model. In seismology, events are often spatially clustered due to geological structures like fault zones.

The Ripley's K function is a statistical tool used to assess the degree of spatial clustering within a seismic catalogue. It measures the number of events occurring within a specific radius around a given event and compares this with the expected number under a random spatial distribution. By applying Ripley's K function to a declustered catalogue, researchers can evaluate whether the spatial clustering observed in the data is consistent with the inhomogeneous distribution assumed by the SITHP model. If the function shows a plateau at larger distances, it suggests that events become independent beyond a certain spatial scale, indicating that the declustering process has effectively isolated the independent seismic events while retaining the natural spatial clustering at smaller scales.

The Gardner-Knopoff (GK) algorithm is one of the most established methods for declustering seismic catalogues. It operates by defining spatial and temporal windows around larger seismic events, within which smaller, related events are classified as aftershocks or secondary events and are removed from the catalogue. This algorithm relies on empirically derived parameters to set the size of these windows, which are typically functions of the magnitude of the main event. The GK algorithm works by iterating through the catalogue, starting with the largest events and progressively filtering out smaller, correlated events that fall within the predefined windows. The result is a catalogue that contains only the independent seismic events, which can then be analysed to assess seismic risk or model earthquake recurrence rates. The effectiveness of a declustering technique, such as the GK algorithm, is assessed by the extent to which it produces a catalogue that conforms to both the temporal and spatial assumptions of the SITHP model. A properly declustered catalogue should exhibit temporal homogeneity, with inter-event times following a Poisson process, and spatial heterogeneity, reflecting the clustered distribution of seismic events around fault zones and other geological features. Various statistical tests, including the KS and chi-squared tests, are used to measure temporal homogeneity, while Ripley's K function evaluates spatial clustering. If a declustering technique produces a catalogue that passes these tests, it can be considered effective. However, the performance of the declustering method depends heavily on the selection of parameters, and different choices can significantly impact the results. Therefore, careful optimization of these parameters is essential to ensure the best possible separation of independent seismic events from correlated sequences.

Declustering is a critical process in seismic data analysis, as it enables the removal of correlated events that can distort statistical models. By applying techniques like the Gardner-Knopoff algorithm and validating the results with statistical tests, researchers can generate cleaner, more accurate seismic catalogues. These declustered catalogues allow for better application of models like the SITHP, improving the accuracy of seismic hazard assessments and enhancing our understanding of earthquake dynamics.

In sum, declustering techniques are not only essential for refining seismic data, but also for ensuring that models like the SITHP accurately reflect the underlying seismic activity. Through careful application and testing, these methods contribute to more reliable and meaningful insights into earthquake processes.

2.1.1.1   Results of Declustering Analysis

The catalogue taken into consideration for declustering is the Sicilian catalogue, with the time range from 2005 to 2022. In the process of declustering seismic catalogues, the study employed various parameter combinations for the temporal and spatial windows, aiming to optimize the removal of secondary events. The primary objective was to evaluate the

conformity of declustered catalogues to a Poissonian process by applying statistical tests to the inter-event times.

The testing involved several key steps:

1. Testing a range of parameter values for time (t1, t2) and distance (d1, d2)

2. Applying a magnitude completeness filter to ensure that only events above magnitude 2.2 were included

3. Running two critical tests: one to verify temporal independence (Kolmogorov-Smirnov test) and another to examine spatial clustering (Ripley's K function)

The declustered catalogues were assessed based on their ability to reproduce realistic *b*-values from the Gutenberg-Richter law, which characterizes the relationship between earthquake magnitude and frequency.
The following findings were observed across different declustering methods:

For the ***Gutenberg-Richter law*** parameters, it is possible to observe, see figure n. 2.2, 2.3:

- Uhrhammer Method: This method consistently produced higher *b*-values (above 1) with minimal variability. While this suggests a certain rigidity in the model, it indicates that the window configuration used by Uhrhammer may not be suitable for accurately detecting realistic variations in *b*-values. This method also did not differentiate based on magnitude when defining spatio-temporal windows, resulting in less flexible clustering.

- Gardner-Knopoff (GK) and Gruenthal Methods: Both methods showed more distributed *b*-values, ranging from approximately 0.7 to 1.03 for Gardner-Knopoff and 0.6 to 0.95 for Gruenthal. These methods provided more realistic and variable *b*-values, aligning better with expected seismic behaviour. Compared to Uhrhammer, these methods demonstrated more flexibility and adaptability in capturing the diversity of seismic events.
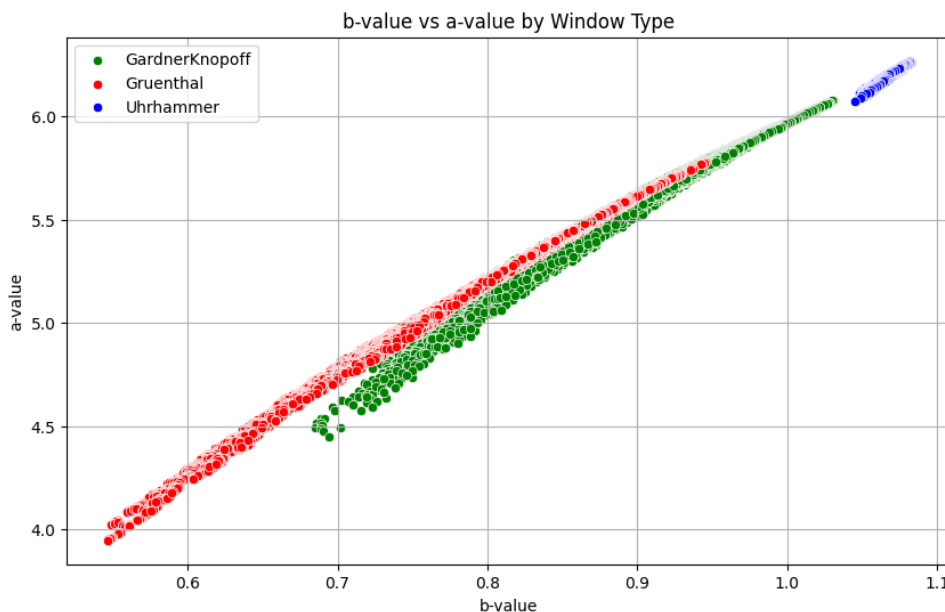
Figure 2.2: Relationship between a-value and *b*-value by different window type
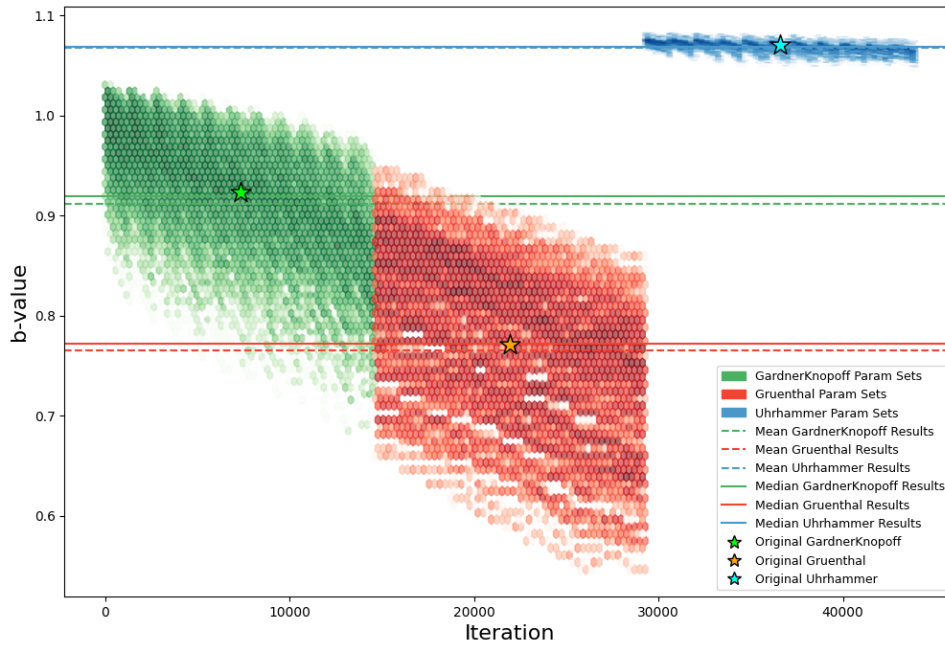
Figure 2.3: Variations of *b*-value with iterations. Each individual point represents a catalogue examined according to the number of iterations. The stars represent the best iteration obtained for the three windows

The 3D plot, figure n. 2.4 representing ***the Kolmogorov-Smirnov (KS) test*** results provided a clear visualization of the distribution of statistical values across the different declustering methods. The key observations include:

- Gardner-Knopoff and Gruenthal Methods: Both methods displayed results with KS statistics close to the expected distributions under a Poisson process. The majority of *p*-values from these methods were above the standard rejection threshold ($p = 0.05$), suggesting that for most parameter configurations, the null hypothesis of temporal independence could not be rejected. These methods were successful in producing declustered catalogues where inter-event times largely followed an exponential distribution.

- Uhrhammer Method: This method showed consistently higher KS values, with many results falling below the $p = 0.05$ threshold. This indicates that the Uhrhammer method frequently led to rejection of the null hypothesis, meaning that the declustered catalogues produced by this method did not conform well to a Poisson process, likely due to insufficient removal of correlated events.
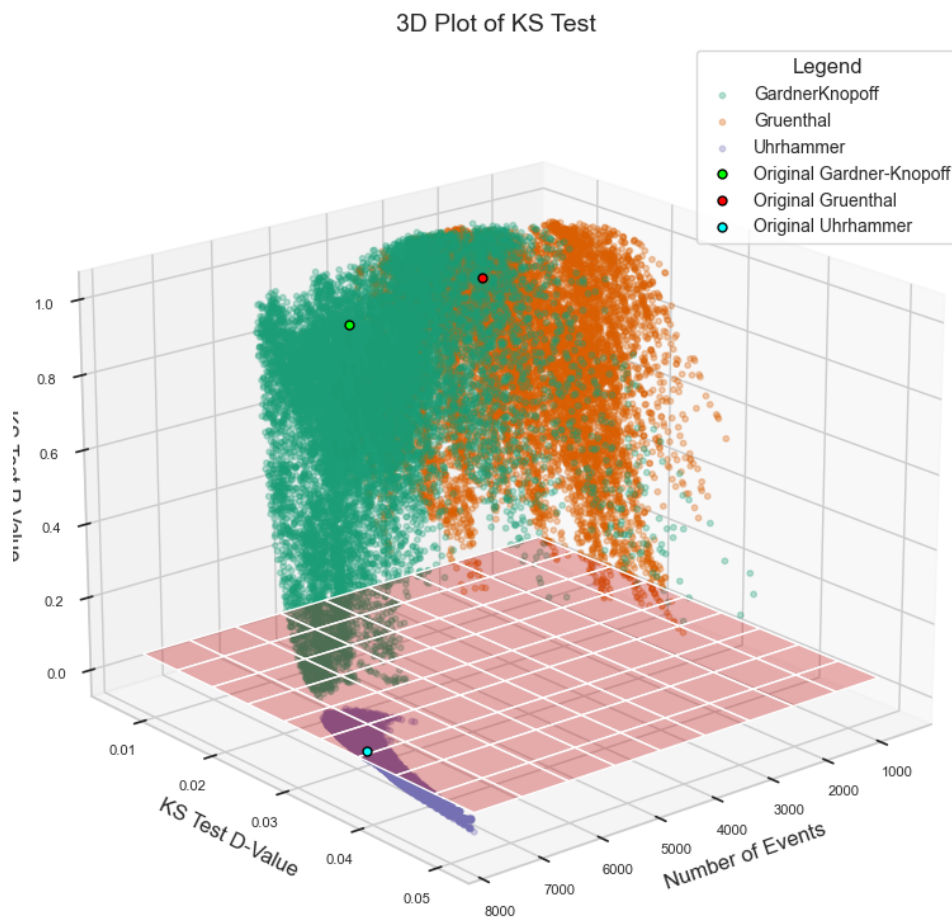
Figure 2.4: 3D graph representing the Kolmogorov-Smirnov (KS) test results. The red grid represents the rejection plane, where the p-value is set to 0.05.

Additionally, the number of events retained in the declustered catalogues did not show a direct correlation with the KS test results, implying that increasing the number of events does not necessarily improve conformity to a Poissonian distribution. Instead, the choice of spatio-temporal windows plays a more significant role in influencing the results.

The Chi-squared test results further corroborated the findings of the KS test, see figure n.2.5 :

- Gardner-Knopoff and Gruenthal Methods: These methods showed a concentration of p-values above the rejection threshold, indicating that the observed frequency of inter-event times was largely consistent with the expected frequencies from a Poisson process. This reinforces the conclusion that these methods perform well in removing correlated seismic events.

- Uhrhammer Method: Similar to the KS test, the Uhrhammer method consistently produced chi-squared statistics that led to rejection of the null hypothesis. This indicates a persistent deviation from the expected Poisson process, suggesting that this method did not effectively decluster the catalogue for the data set under study.
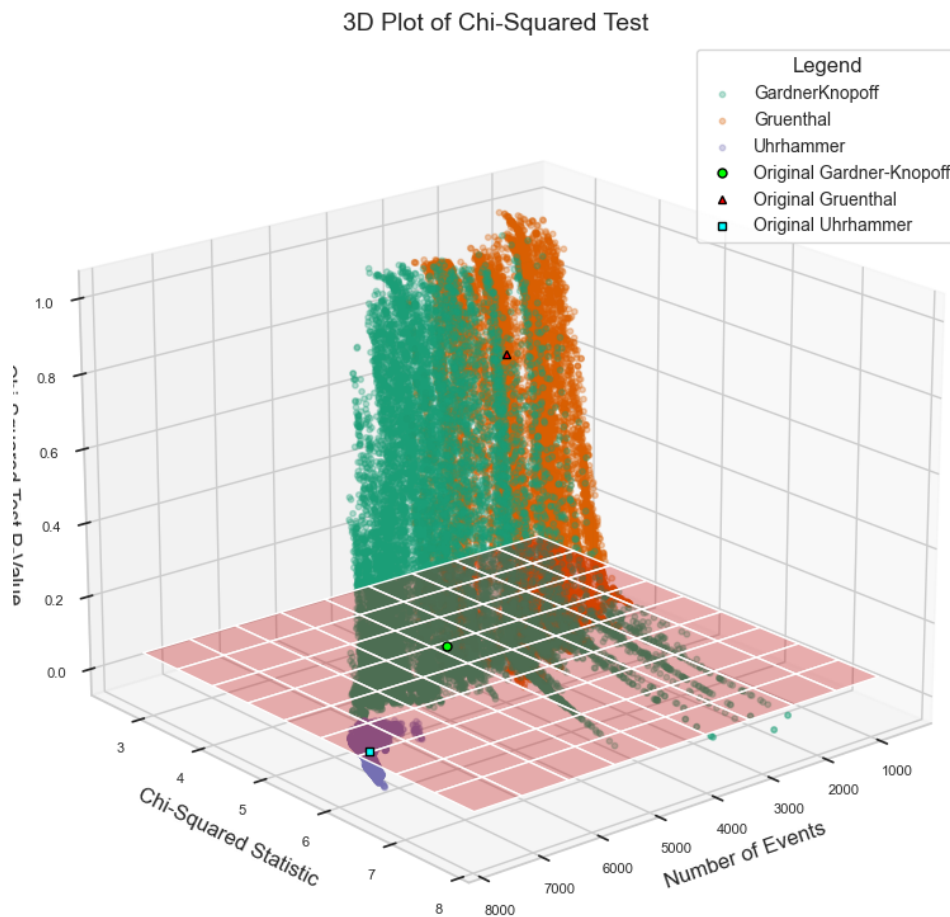
Figure 2.5: 3D graph representing The Chi-squared test results. The red grid represents the rejection plane, where the p-value is set to 0.05.

The analysis of **Ripley's K function** provided insights into the spatial distribution of seismic events after declustering, see figure n.2.6:

- Small-scale clustering (0-5 km): All methods revealed strong clustering at smaller spatial scales, as expected, reflecting the presence of natural earthquake clustering

in areas of active seismicity. This result is consistent with the known geological structures in the region.

- Larger distances: Beyond approximately 5 km, the K function plateaued for all methods, indicating that events became spatially independent at larger scales. However, the Uhrhammer method produced slightly higher K function values at these larger distances compared to the other methods, suggesting that it was less effective at removing spatially correlated events, which aligns with the test results for temporal independence.
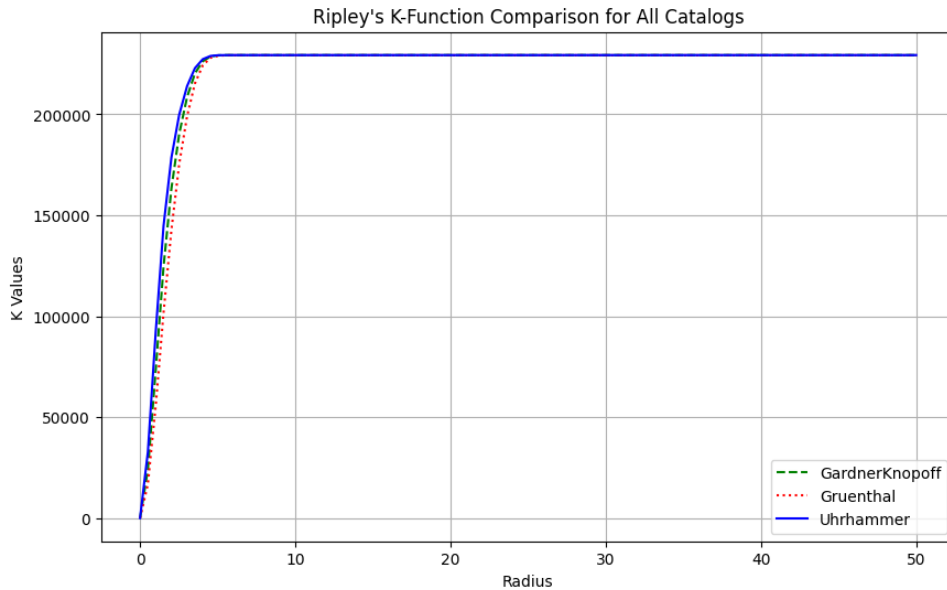
Figure 2.6: Graphical representation of Ripley's K function comparing the three windows.

2.2   Completeness Magnitude

The accurate determination of the completeness magnitude is critical, as it directly influences the reliability and robustness of subsequent seismic analyses. This parameter delineates the threshold above which all seismic events within a specified area and period are consistently and comprehensively recorded. Consequently, it serves as a foundational element for ensuring the quality of seismic catalogues. Furthermore, the completeness magnitude plays a crucial role in defining the homogeneity of the seismic catalogue: a homogeneous seismic catalogue is characterized by consistent recording of seismic events above the completeness magnitude, devoid of temporal or spatial biases. This homogeneity is essential for meaningful statistical analyses and for deriving robust seismological models. Inconsistent recording, arising from variations in seismic network sensitivity or changes in detection capability over time, can lead to misleading conclusions about seismic activity and earthquake risk.

The analysis of the magnitude of completeness is strongly influenced by various factors, such as:

- the areal coverage of the seismic network[41, 60].

- events of magnitude greater than 4, since such events tend to saturate the stations near the hypocentre, and this causes gaps that are extended over time [31].

Short-term Aftershock Incompleteness (STAI) [31] refers to the temporary under-recording of aftershocks following a large earthquake. After a significant seismic event, the

detection of smaller aftershocks may be incomplete due to the overwhelmed capacity of seismic networks, increased background noise, or delayed calibration of instruments.

During this period, smaller aftershocks may go undetected, leading to gaps in the seismic catalogue, thus correcting for STAI is important in seismic analysis. Methods such as adjusting detection thresholds or using statistical models help mitigate this incompleteness in aftershock data.

As previously mentioned, this parameter has been calculated and studied with various techniques: I here illustrate the main ones. The MAXC [74] is a technique based on computing the maximum value of the first derivate of the frequency-magnitude curve. This method is fast and simple, and is able to provide a stable estimate, even when applied to small datasets. Moreover, this method provides satisfactory estimates of Mc when we have a graduate curve of FMD, but it is referring to a local region. The GFT [74] is a method based on the difference between the observed and synthetic FMD. The calculation is based on the parameter R which represents the absolute difference in the number of events in each magnitude bin between the observed and synthetic GR distributions. Below is the formula for the GFT method:

$$R(a, b, M_\infty) = 100 - \left(\frac{\sum_{i=M_\infty}^{M_{\max}} \left| \frac{B_i - S_i}{\sum_i B_i} \right| 100}{\sum_i B_i} \right) \qquad (2.2)$$

where:

- Bi and Si are the observed and predicted cumulative number of events in each magnitude bin.

- Mmax is the maximum magnitude represented in the catalogue

Mc is found at the first magnitude cutoff at which the observed data for $M \geq Mco$ is modelled by a straight line (in log-lin plot) for a fixed confidence level, e.g. R = 90% or 95%.
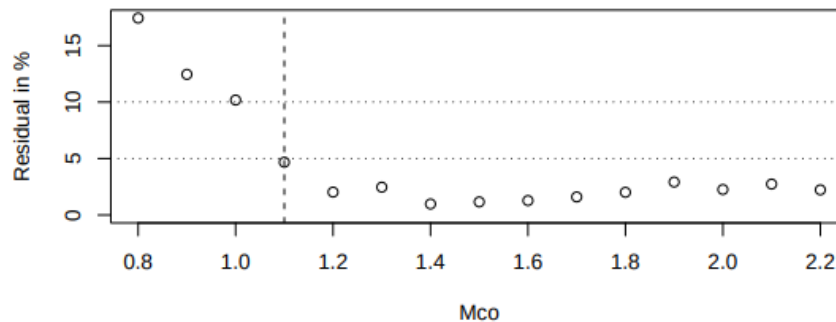
Figure 2.7: GFT tecniques [42]

Mc by MBS is a technique created by Cao and Gao [11]. This technique uses the stability of the b-value in relation to the Mco. It represents the optimal magnitude of completeness obtained by analysing the stability of the b-value. It corresponds to the magnitude value above which the b-value remains constant and accurately represents the frequency-magnitude distribution.If Mco is less than M, the b-value estimates go up and for Mco it is greater than or equal to M then b remains constant. Only if Mco greater than M is not valid. The closer Mco approaches M, the b-value tends to be real and a plateau forms in the distribution. Graphically, a plateau should be obtained that represents the stability of the b-value in correspondence with the value of Mco.

Mc from EMR, proposed by [75], it is a technique that makes use of the entire catalogue even considering the events below the magnitude of completeness. The proposed technique
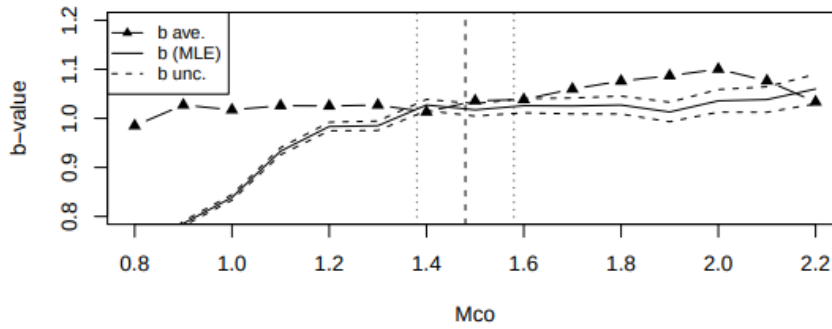
Figure 2.8: MBS technique, [42]

is developed by dividing it into two parts the GR distribution: the G-R law for the complete part, and the cumulative normal distribution for the incomplete part of the non-cumulative FMD. The model attempts to reproduce the entire frequency-magnitude distribution, thus fitting the incompletely observed part.

The MBASS, proposed by [5], is a non-parametric technique to estimate Mc. It is based on an iterative method designed to search for multiple change points in the non-cumulative FMD.

Kolmogorov-Smirnov Test is a non-parametric test and compares a sample distribution with a reference distribution (either theoretical or another empirical dataset). In seismic studies, the Kolmogorov-Smirnov test is employed to evaluate if the distribution of observed magnitudes matches a theoretical model, such as a Gutenberg-Richter law for magnitude-frequency distribution [66].

Lilliefors Test, a modification of the Kolmogorov-Smirnov test, is applied to assess the goodness-of-fit for a normal distribution when the population parameters (mean and variance) are unknown [33]. In seismology, this test can be useful to check if earthquake magnitude data follows a normal distribution after certain transformations [28, 66].
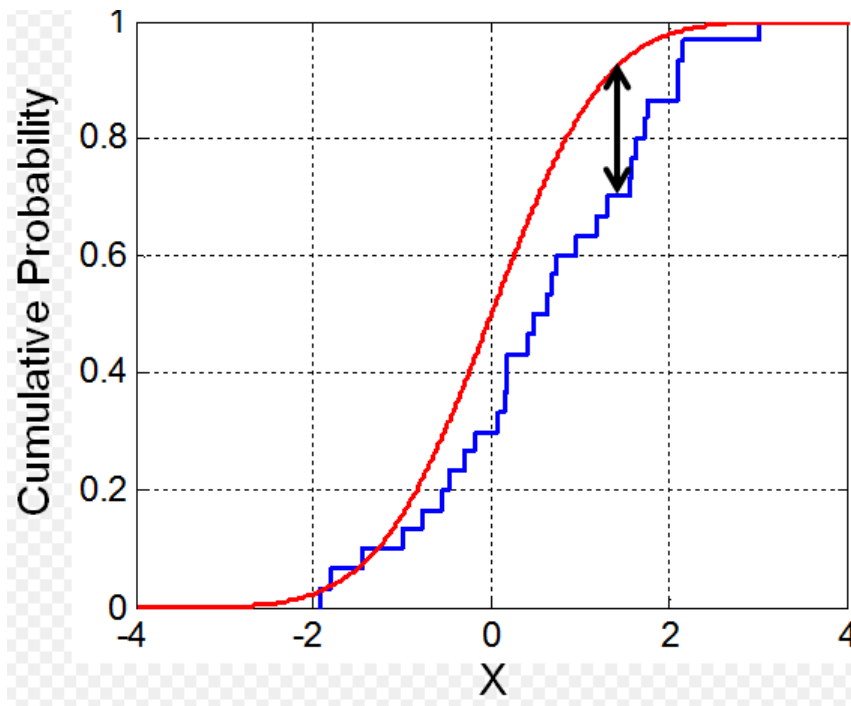
Figure 2.9: Lilliefors Test. The X-axis represents the observed values of the random variable you are testing for normality in the Lilliefors test. These X-values correspond to the observed data, sorted in ascending order and normalized to a continuous interval.

The choice of method for estimating the completeness magnitude and its implementation requires meticulous consideration of the specific characteristics of the seismic network and the study region. This consideration is vital to ensure that the resulting estimates are both accurate and meaningful to the particular seismotectonic environment.

2.2.1   Analysis algorithms

A preliminary phase of the doctoral research involved comparing the main analysis techniques for calculating Mc, using synthetic tests and different seismic catalogues, to identify how the use of each technique influences the estimation of Mc and how variations in geostructural conditions may affect the applicability of specific methods. In particular, the study highlights that different geodynamic contexts may require adjustments in spatial and temporal windows, as standard optimization techniques for declustering and statistical analysis may not perform uniformly across varying tectonic environments.

The results of this comparative study have been described and discussed in the manuscript by Guastella et al., which was submitted in August 2024 to 'Journal of Seismology', and the draft is presented in Appendix A. My contribution primarily consisted of supervising data analyses, processing seismic catalogues, and writing parts of the article. This manuscript aims to be a kind of guideline on which technique to apply and in which context to apply it. Every analysis technique has its peculiarities, its limitations and its advantages.

The manuscript uses a detailed seismic catalogue of southern Italy, covering a significant period of 18 years.

Various declustering techniques are applied, including:

- Gardner-Knopoff

- Gruenthal

- Uhrhammer

For the analysis of Mc, several catalogue-based methodologies are employed, such as:

- Maximum Curvature (MAXC)

- Goodness-of-Fit Test (GFT)

- Median-Based Analysis of Segment Slope (MBASS)

- Entire Magnitude Range (EMR)

- b-value stability (MBS)

Figure n. 2.10, summarizes the work carried out for the manuscript Guastella et al., *submitted*. First, a reference catalogue was taken as the area of Sicily from 2005 to 2022.

All the previously described analysis techniques were applied to it. Subsequently, three types of declustering analysis were applied and the magnitude of completeness analysis techniques were applied to the three new catalogues.
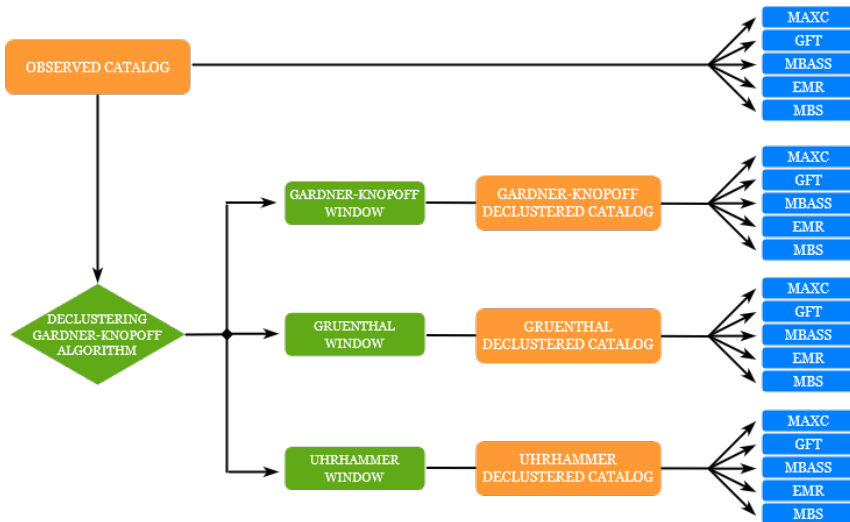
Figure 2.10: Graphic summary of techniques used for Guastella at..,

We also compared the various FMDs obtained from the four new catalogues, as shown in figure n. 2.11 We then created the panels shown in figures n. 2.12,2.13,2.13,2.14,2.15, where all the FMDs for each technique for each catalogue are shown. The distinction between the various panels is according to the type of correction, with the bootstrap effect obtained.
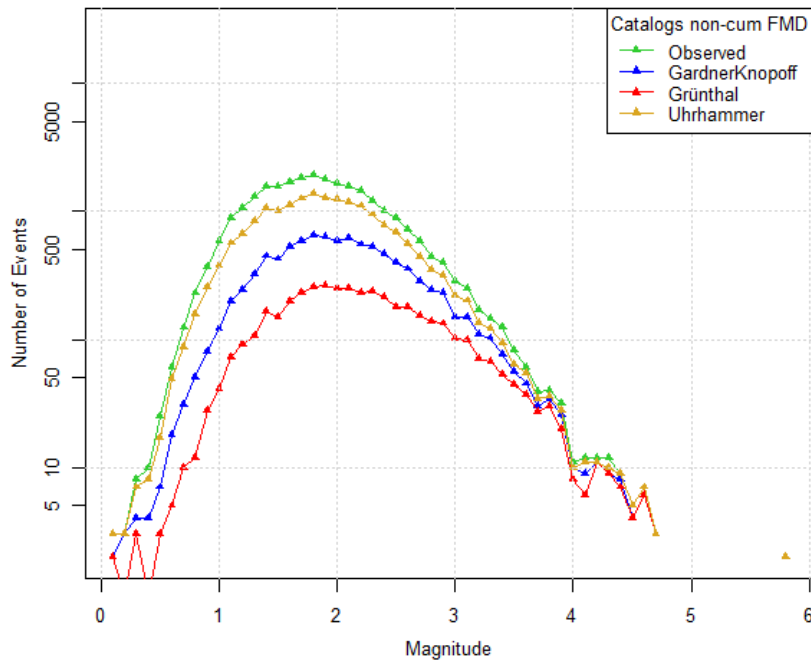
Figure 2.11: FMD calculated for Base Catalogue and three applied technique
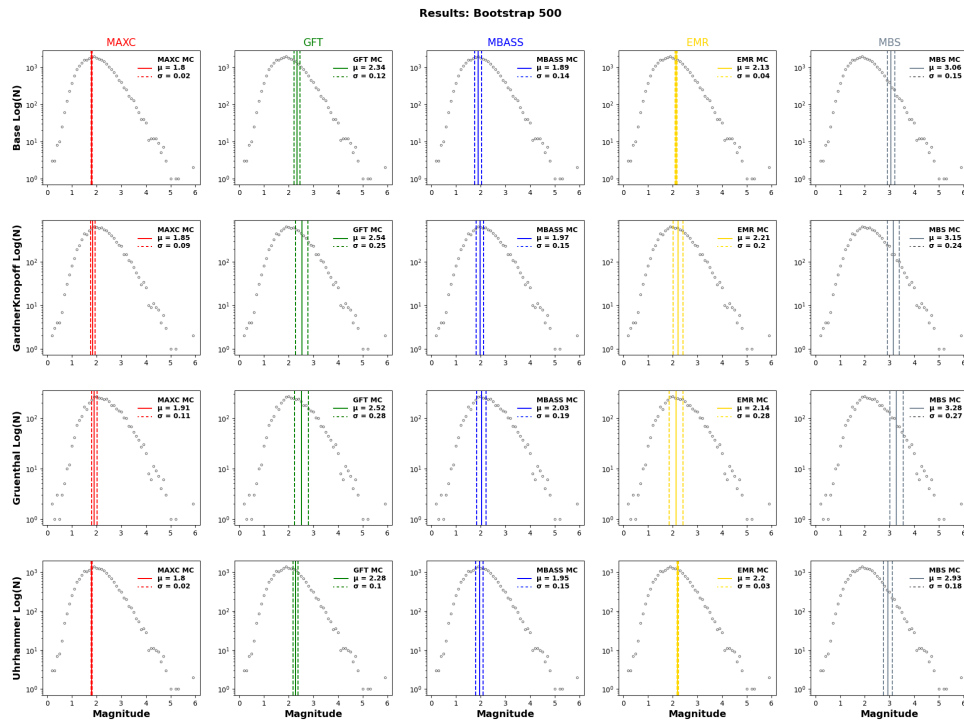
Figure 2.12: Incremental FMDs across each catalogue (rows) and the methods applied per catalogue (columns). The figure presents completeness results (vertical lines) and the corresponding uncertainties (dashed lines), derived from 500 bootstrap samples
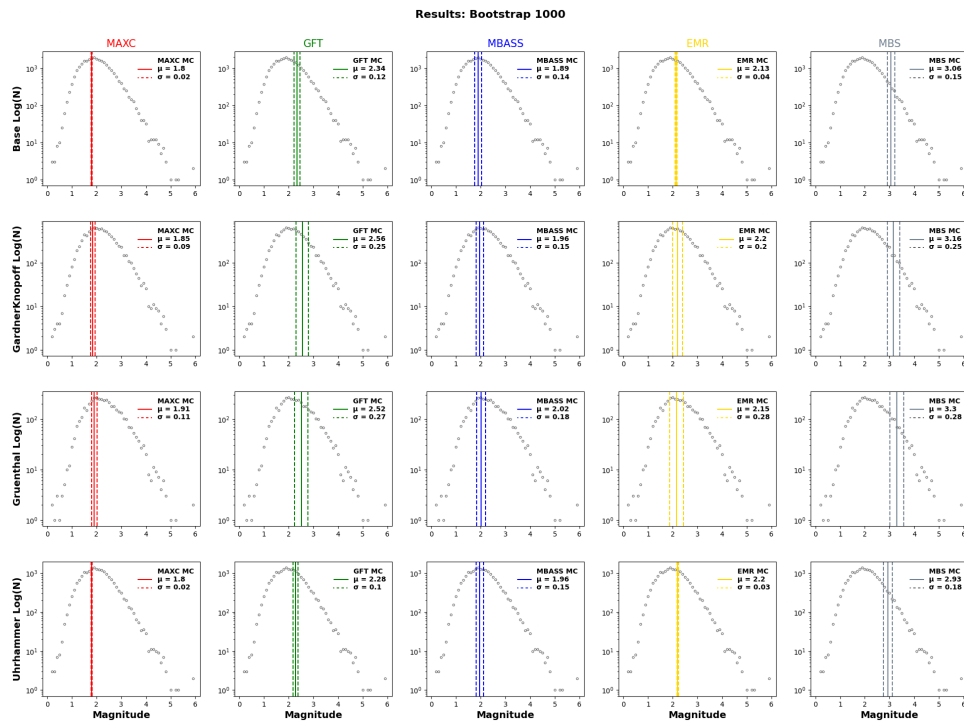
Figure 2.13: Incremental FMDs across each catalogue (rows) and the methods applied per catalogue (columns). The figure presents completeness results (vertical lines) and the corresponding uncertainties (dashed lines), derived from 1000 bootstrap samples
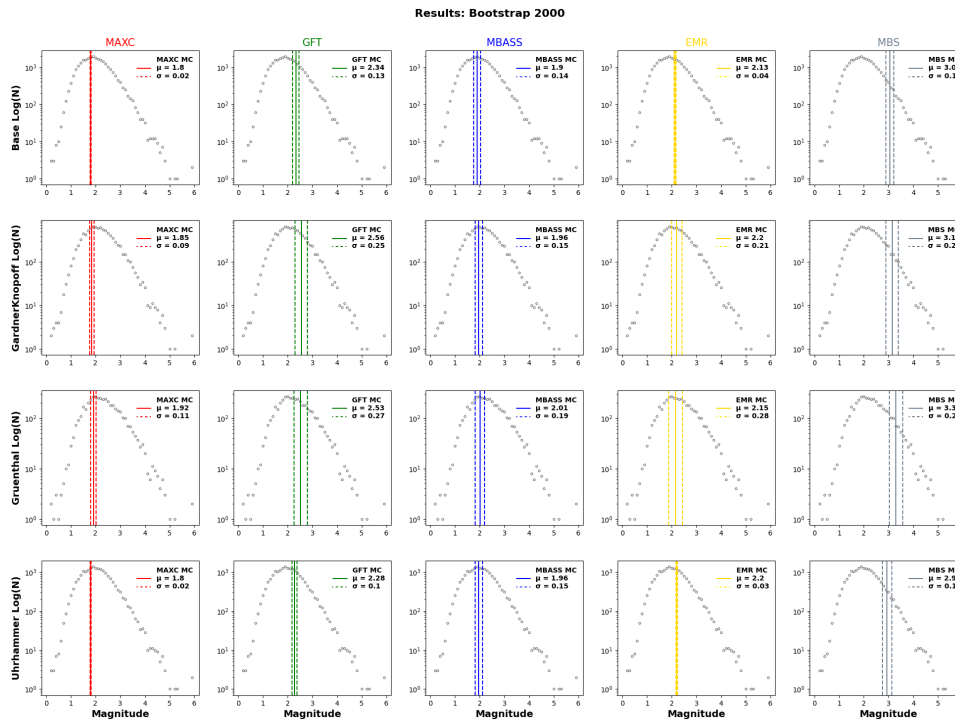
Figure 2.14: Incremental FMDs across each catalogue (rows) and the methods applied per catalogue (columns). The figure presents completeness results (vertical lines) and the corresponding uncertainties (dashed lines), derived from 2000 bootstrap samples
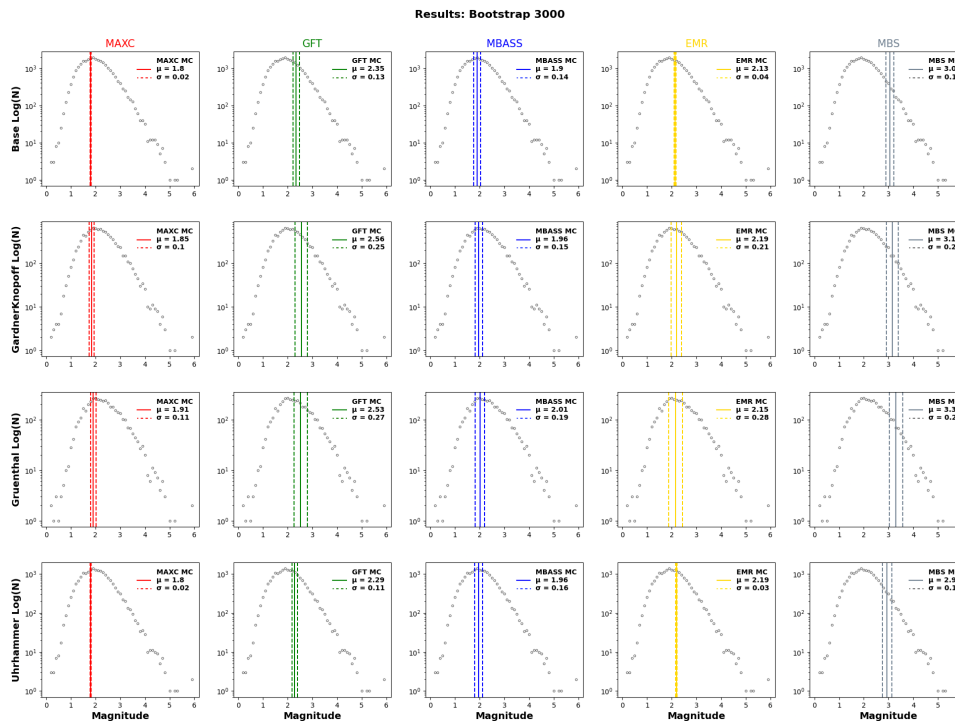
Figure 2.15: Incremental FMDs across each catalogue (rows) and the methods applied per catalogue (columns). The figure presents completeness results (vertical lines) and the corresponding uncertainties (dashed lines), derived from 3000 bootstrap samples

Further results of this study, not included in the paper, are reported in the chapter 5.

2.3   *b*-value

The *b*-value parameter, which represents the slope of the frequency-magnitude distribution, is computed after the determination of the completeness magnitude (Mc). The accuracy and reliability of the *b*-value are intrinsically linked to the precise estimation of Mc. Thus, the estimation of Mc is a decisive factor in the computation of the *b*-value, ensuring the robustness of seismic analyses and interpretations.

Typically, a *b* of about 1.0 is observed globally, indicating a balanced distribution of small and large earthquakes. However, deviations from this mean can provide significant insights into the stress state and tectonic characteristics of a region [3, 26, 44, 59, 70].

A higher *b* value (> 1.0) suggests a relative abundance of smaller earthquakes, which may indicate a lower-stress environment or a more fragmented fault system. In contrast, a lower *b* value (<1.0) implies a higher proportion of larger earthquakes, potentially indicative of a higher stress regime or a more coherent fault structure. Understanding these variations is critical for assessing seismic hazard and anticipating the potential occurrence of large, destructive earthquakes.

This parameter can be applied to various domains to enhance our understanding of seismic processes and to improve earthquake hazard assessments. The fields of application are mainly the following::

- seismic hazard assessment: By analysing the *b*-value across different regions, seismologists can identify areas with higher probabilities of large seismic events. A lower *b*-value indicates a higher likelihood of significant earthquakes, which is critical for developing accurate seismic hazard maps [14, 46].

- earthquake forecasting: Changes in the *b*-value over time can signal variations in the stress state of a region, potentially indicating an impending seismic event. By monitoring *b*-value trends, seismologists can develop probabilistic models to forecast the likelihood of future earthquakes [9, 27, 38, 55].

- tectonics and geological studies: In tectonic and geological studies, the *b*-value provides insights into the stress regime and fault characteristics of a region. Different tectonic environments exhibit distinct *b*-value patterns [13, 39, 45].

- volcanic seismology: In volcanic regions, the *b*-value is used to monitor and forecast volcanic activity. Variations in the *b*-value can reflect changes in volcanic stress regimes, magma movement, and the potential for volcanic eruptions. A significant decrease in the *b*-value may indicate increasing pressure within the magma chamber, suggesting an elevated risk of eruption [21, 47].

- aftershock analysis: The *b*-value is instrumental in aftershock analysis, helping to understand the aftershock sequence following a major earthquake. Typically, aftershocks exhibit a higher *b*-value compared to the mainshock sequence, reflecting the stress relaxation and fault healing processes[16, 32].

2.3.1   Computational techniques

The accurate calculation of the *b*-value is crucial for its various applications in seismology, as previously discussed. Several methodologies are employed to compute the *b*-value from seismic data, each with its advantages and specific use cases. This subsection outlines the most common techniques for calculating the *b*-value.

The most implemented is the one proposed by Aki [2], called maximum likelihood technique (MLE).

$$b = \frac{\log_{10}(e)}{\overline{M} - M_c} \tag{2.3}$$

where:

- log 10 (e) is the decimal logarithm of Euler's constant

- M is the mean magnitude of earthquakes above the magnitude of completeness (Mc).

This method requires accurate estimation of Mc and is particularly effective for large datasets, providing a reliable b-value that accounts for the variability in earthquake magnitudes. This equation cannot be used for seismic catalogues with spatio-temporal variations in tMc. This represents a strong limitation of this approach, since almost all catalogues exhibit a variation of Mc. In fact, usually in seismic catalogues, the magnitude of completeness decreases over time.

The formula proposed by Aki [2], was rewritten by [75], who modified it in this form:

$$b = \frac{\log_{10}(e)}{\left[\overline{M} - \left(M_c - \frac{\Delta M_{\text{bin}}}{2} \right) \right]} \tag{2.4}$$

where $\overline{M}$ is the average magnitude e ΔM_{bin} is the binning width of the catalogue.

This method like others, despite being widely used, is difficult to apply to carry out the classic tests of comparison of the values obtained. Another method useful to estimate the b-value is the Weichert (1980) approach [71]. This is based on the Poisson distribution, it allows calculating the annual trend of the completeness magnitude and the b-value through the MLE approach. Although this method is also used to calculate seismic risk maps, it is not a good method for making classical b-value comparisons.

An innovative b-value estimatior, called b-positive, has been recently introduced [18]. The author shows that the distribution of magnitude differences is identical to the distribution of magnitudes, but with no reference to the minimum magnitude, and that the positive subset of the differences between successive earthquakes are minimally biased by changing catalogue completeness. This new approach greatly improves the robustness of continuous b-value measurements during active earthquake sequences, as well as in historical catalogues with unknown or variable completeness. In this method b is estimated by the following formula:

$$b = \frac{1}{\delta} \cot^{-1} \left(\frac{1}{\delta} \left(\overline{M} - M_c + \delta \right) \right) \tag{2.5}$$

where δ is a constant, $\overline{M}$ is average magnitude, e M_c is completeness magnitude.

This estimator can thus allow the calculation of b-value on datasets with a certain amount of incompleteness as well as on datasets with variable levels of completeness in time, although some pre-cut of the dataset can be necessary in case of high incompleteness [24, 68].

The b-positive approach to spatial and temporal analysis focuses on improving the estimation of the b-value in earthquake sequences, which indicates the frequency distribution of earthquake magnitudes. However, conventional methods of estimating the b-value are often biased due to the transient incompleteness of earthquake catalogues, especially in the early stages of an aftershock sequence when smaller earthquakes go undetected.

Furthermore, the b-positive is robust in both historical sequences and real-time applications, such as during an aftershock sequence. By focusing only on positive magnitude differences, it eliminates the need to filter out early periods of incomplete data, which is common in traditional approaches. This method has proven to be effective in several

earthquake sequences, confirming the reduction of *b*-values after a few foreshocks, but at more subtle levels than previously estimated.

In synthesis, the *b*-positive increases the robustness of the *b*-value analysis by focusing on magnitude differences that are less sensitive to catalogue incompleteness, improving the accuracy of earthquake forecasting in both spatial and temporal dimensions.

Obtained results are reported in the chapter 5.

2.4   Epidemic Type Aftershock Sequence model

The Epidemic Type Aftershock Sequence (ETAS) model is one of the most widely used models in statistical seismology to describe the spatio-temporal distribution of earthquakes. This model was introduced by Ogata in 1988 [49] and is based on the theory of branching point processes. In the ETAS model, each seismic event can be viewed either as a main event or as an event triggered by a previous earthquake, thus allowing seismic activity to be modelled as a series of aftershocks generated by main events. The conditional intensity function of the ETAS model, $\lambda_\theta(s, t | \mathcal{H}_t)$, represents the expected frequency of events in a unit of time and space, conditional on the history of the process up to time . The conditional intensity function can be expressed as:

$$\lambda_\theta(s, t | \mathcal{H}_t) = \mu f(s) + \sum_{t_j < t} \nu_\phi(t - t_j, s - s_j | m_j) \tag{2.6}$$

The estimation of ETAS model parameters is usually carried out using the maximum likelihood (ML) method. However, the simultaneous estimation of parametric and non-parametric components represents a significant challenge, both statistically and computationally.

In the R etasFLP package, described by Chiodi and Adelfio (2017) [12], a mixed estimation approach is implemented that combines non-parametric estimation techniques for background seismicity and ML techniques for induced seismicity. This approach uses an iterative algorithm that alternates between estimating the background component and the induced component until convergence. The results obtained are presented in chapter n 5.

2.5   Probabilistic detection of completeness magnitude

The probabilistic detection of completeness magnitude was first introduced by Schorlemmer (2008) [61]. This method estimates the probability of detecting an earthquake using empirical data, including phase data, station information, and network-specific attenuation relationships.   Unlike traditional approaches, it models detection probability and completeness as functions of the seismic network and its stations, rather than the earthquake events themselves. Detection probability distributions are derived for each station, describing the likelihood of recording an earthquake of a given magnitude at varying distances from the station. Here, detection refers specifically to the ability to record a seismic phase.

This approach avoids assumptions regarding earthquake occurrence, such as event-size distribution, making it particularly suitable for regions with sparse data where conventional methods may be unreliable.

From individual station probabilities, we determine the likelihood of detecting an event at four or more stations—representing the threshold for seismic network activation, as at least four stations are required to initiate the location process. Network configurations change frequently, with stations being added, removed, or replaced, resulting in shifts in detection probabilities. Thus, each probability map is valid only for a specific network configuration.

This method relies entirely on empirical data, which includes: (1) station data (location, activation, and deactivation times), (2) phase picks indicating where events were recorded, and (3) the attenuation relationship used for magnitude estimation.

The procedure consists of two main phases. First, detection probabilities for each station are calculated as functions of space and magnitude. Second, probability maps are generated to display the likelihood of detecting an event of a given magnitude.

Calculating detection probabilities involves six key steps: (1) identifying a consistent data recording period for completeness analysis, (2) loading station data, (3) importing event data with phase picks, (4) associating recorded and unrecorded events with stations, (5) addressing station aliases and doublets, and (6) calculating detection probabilities as a function of magnitude and distance.

Data Requirements

- Station Data: Includes location and operational times (on/off) for each station in the network.

- Phase Picks: Recorded phases of each earthquake at various stations.

- Attenuation Relation: Used for determining the magnitude based on seismic data.

Data Recording:

- Ensure the period selected has consistent triggering conditions and magnitude definitions.

- Avoid periods with significant network changes or large clusters of aftershocks.

For import station data, it is need to select all operational stations within the defined period.Knowledge of on and off times is essential to ensure accurate detection probability calculations.

For import event data, Include all earthquakes that occurred within the homogeneous recording period. Use phase picks (typically P-wave arrivals) that are relevant for automatic triggering.

Assign Recorded and Not Recorded Events to Stations: For each event, generate data triplets indicating if the event was picked at a station, the hypocentral distance, and the event magnitude. Plus triplets represent picked events, and minus triplets represent non-picked events.

Resolve Station Aliases and Doublets: Merge data from stations with aliases or minor relocations to preserve data integrity. If the site conditions are significantly different due to relocation, treat them as separate stations.

Compute Detection Probabilities: Use the station data triplets to derive detection probabilities as a function of magnitude and distance. Apply a metric to measure the distance in the magnitude-distance space, typically defined by the attenuation relation:

$$\Delta M = |M - M_0| \tag{2.7}$$

Here, M translates distance into magnitude units using the attenuation relation:

$$\Delta L = |L - L_0| \tag{2.8}$$

$$L_M = \sqrt{\Delta M^2 + \Delta M'^2} \tag{2.9}$$

$$M = c_1 \log A + c_2 \log L + c_3 \tag{2.10}$$

Calculate detection probabilities PD(M,L) as the ratio of plus triplets +N+to the total triplets N:

$$P_D(M, L) = \frac{N}{N+} \tag{2.11}$$

Synthesis of Detection and Completeness Maps Detection Probability Maps: Combine individual station detection probabilities to determine the probability of detecting an earthquake at four or more stations.

Completeness Magnitude Maps: Derive maps indicating the completeness magnitude MP(x,t), which represents the magnitude above which the network can reliably detect all events.

This method offers several significant advantages. By focusing on network properties rather than relying on earthquake samples, it provides a more relevant analysis of completeness. This approach is particularly beneficial in regions with sparse earthquake data, where traditional methods may prove ineffective. Additionally, the method is highly adaptable, capable of accommodating changes in the seismic network such as the addition or removal of stations, and can even evaluate hypothetical scenarios. Overall, it offers a comprehensive framework for estimating earthquake detection probabilities and completeness magnitudes, enhancing our understanding of seismic network capabilities and facilitating more effective assessment and planning for seismic monitoring efforts.

The results obtained are presented in chapter n 5.

Chapter 3

Software tools

This chapter is dedicated to the description and use of software tools that enable the effective analysis of seismic data, with a focus on magnitude of completeness and b-value. Firstly, a software tool is presented which is already well established in the scientific literature and which allows the accurate calculation of these two fundamental parameters for the characterization of seismicity. Thanks to this tool, it is possible to perform detailed statistical analyses on various seismic catalogues, facilitating the interpretation of data and improving the understanding of the distribution of seismic events as a function of their magnitude. Next to this, Tremors, a software designed by the candidate to address a specific challenge related to the analysis of seismic catalogues, is introduced. Tremors offers an innovative approach to filtering data and handling the STAI (Short-Term Aftershock Incompleteness) problem, providing practical support for researchers to obtain more accurate and reliable results. With this tool, the quality of seismic data can be improved, optimizing the analysis process and reducing the impact of incomplete aftershock sequences.

3.1 ZMAP: A Software Package to Analyze Seismicity

ZMAP is a set of tools, designed to analyse the quality of seismic catalogues and support the scientific community in managing certain routines on seismic networks. The first version was developed in 1994. ZMAP version 6.0 in 2001. It was designed with a graphical user interface (GUI), as shows figure n. 3.1 in the Matlab environment.
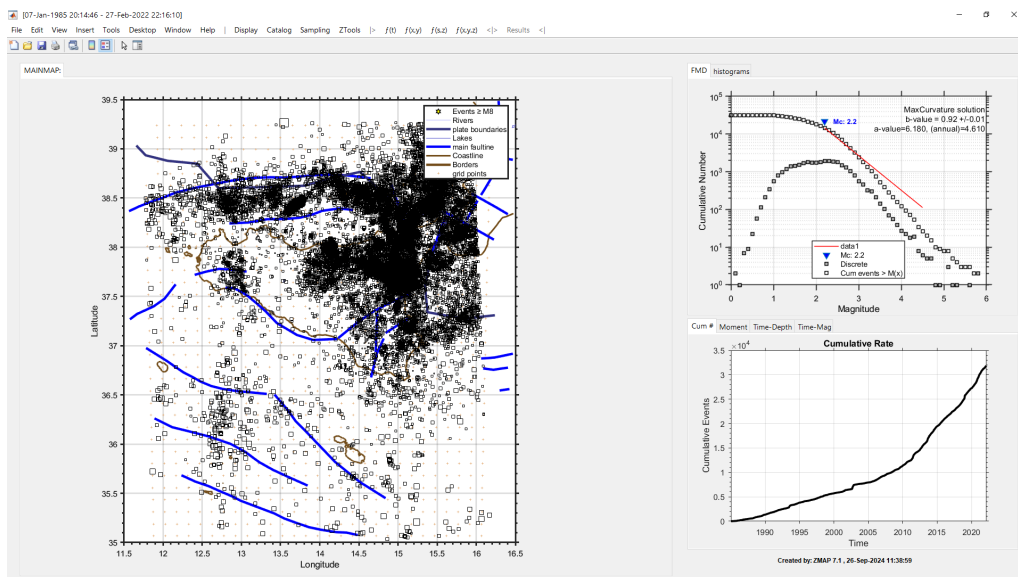
Figure 3.1: ZMAP graphic screen.

It was developed to facilitate the visualization and interpretation of seismic data, helping seismologists to detect changes or trends in seismicity that could be indicative of earthquake or volcano hazards.

ZMAP uses a series of algorithms to analyse seismic catalogues and produce maps and graphs representing the distribution and characteristics of seismic events in time and space.

These algorithms include:

- Frequency-magnitude distribution analysis: ZMAP uses algorithms to calculate the b-value locally, i.e. in specific regions or at specific depths, to identify spatial variations that may indicate anomalies in seismicity. It also makes it possible to calculate the magnitude of completeness.

- Declustering: To separate independent of dependent seismic events (such as aftershocks), ZMAP employs declustering algorithms based on the Reasenberg algorithm [54]. This is essential for identifying seismic clusters and analysing only the background seismicity.

- Inversion of the stress tensor: ZMAP includes tools for inverting the stress tensor to determine the orientation of tectonic forces in a region. These algorithms help map stress variability in temporality and spatially, providing useful information on the dynamics of the Earth's crust.

The physical basis of ZMAP is based on seismology and rock fracture mechanics. Some key concepts include:

- Gutenberg-Richeter Law, that is described in chapter n. 2.1

- tensor of tensions

- fracture theory

Thanks to the use of the ZMAP software, the first maps for the magnitude of completeness, b-value and a-value were produced.

3.2   Tremors: A Software App for the Analysis of the Completeness Magnitude

During the development of the doctoral thesis, a new tool for calculating the magnitude of completeness was designed, using new methodologies [2, 28, 66].This software was published in a special issue of Geosciences, entitled Advances in Statistical Seismology. The paper can be found in the appendix A. In particular, this tool tries to solve the STAI problem, explained in the previous subchapter n. 2.2.

The Tremors software represents an innovative application developed to facilitate the analysis of seismic catalogues, with a particular focus on the estimation of magnitude of completeness (Mc) and the problem of short-term aftershock incompleteness (STAI).
This tool, realized in the MATLAB environment but also available in a stand-alone version, offers the possibility of analysing seismic data with modern methodologies and managing the time incompleteness problems associated with major seismic events.

Tremors is designed to be used even by less experienced users thanks to a simple graphical interface that makes it easy to navigate through the software's various functions, figure n. 3.2. One of its most relevant features is that it does not require a MATLAB licence, being available as a stand-alone application. This makes it accessible to a larger scientific community, especially for those with limited resources.
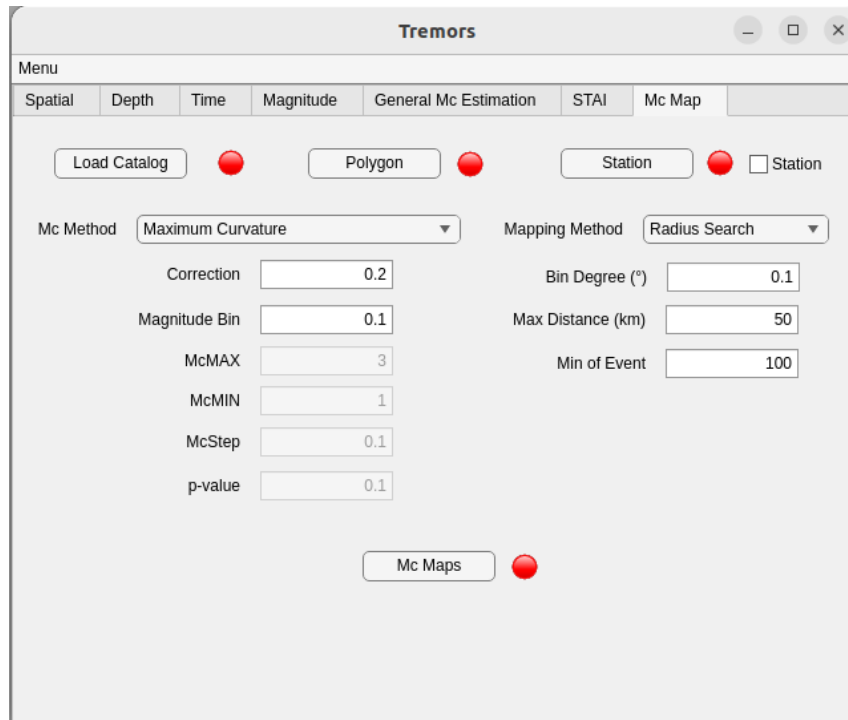
Figure 3.2: Tremors app graphic screen.

One of the main advantages of Tremors compared to existing software, such as ZMAP, is the implementation of advanced methods for estimating the magnitude of completeness. In addition to the classical methods, Tremors uses the Lilliefors method, which is known for its statistical robustness. This approach allows Mc to be determined by testing the exponentially of magnitudes, providing greater accuracy and reliability than traditional methods. In addition, the software includes Taroni's (2023) method [66], which is particularly effective for catalogues with variations in *b*-value, adapting to situations where seismic parameters change depending on the geodynamic zone.

One of the most significant challenges in analysing seismic catalogues is the phenomenon of short-term aftershock incompleteness (STAI), which occurs after large magnitude seismic events. Tremors introduces an innovative approach to handle this problem by temporarily increasing the magnitude of completeness after large shocks, rather than removing STAI periods altogether. This method preserved valuable information that would otherwise be removed, improving data quality and analysis reliability.

The software allows users to filter the seismic catalogue according to spatial, temporal and depth criteria. This flexibility is crucial for the analysis of extensive catalogues, allowing the precise identification of the study area, the selection of specific time periods (e.g. before and after seismic network updates) and the definition of depth intervals, which are essential for accurate Mc estimation.

Another important aspect is that it generates high quality figures and georeferenced maps in *.tif* format, which can be exported and used with other software for further analysis. This makes the application a versatile tool that can be adapted to different data visualization and presentation requirements.

Tremors also allows spatial analysis of magnitude completeness, generating maps that show how Mc varies within a region. This functionality is crucial for understanding gaps in seismic network coverage.

The innovation introduced by Tremors consists in its ability to combine modern methods of estimating completeness magnitude with tools for handling aftershock incompleteness, making it a versatile tool for seismologists and researchers. The software is designed to

optimize workflow, reducing the time required for pre-processing seismic catalogues and improving the reliability of results.

Chapter 4

Comparative analysis of seismic databases across three different global regions

In this chapter, the diversity of geodynamic and seismic scenarios is examined using seismic databases from three different regions: Italy, Taiwan and Iceland. The aim is to explore how the magnitude of completeness and the b-value, two fundamental parameters for the study of earthquakes, vary across different geodynamic environments.

Each region has unique geophysical characteristics, which influence both the distribution of seismic events and how energy is released. By comparing these three geographical contexts, this analysis reveals how variability in local tectonic dynamics impacts seismological parameters, offering insights into how statistical models can be tailored to distinct geodynamic conditions. In Italy, marked by complex plate interactions Taiwan, characterized by intense tectonic activity linked to a convergent plate margin, and Iceland, located on a mid-ocean ridge, different configurations of magnitude of completeness and b-value are observed. This analysis highlight environmental influences on these parameters and elucidates how they vary in response to the specific characteristics of each geodynamic systems.

4.1 Italy

Italy is characterized by significant seismic activity due to its location along the convergent boundary between the Eurasian and African tectonic plates. This tectonic setting results in frequent earthquakes, particularly in regions such as the Apennines Mountains, the eastern Alps, and Sicily, where the risk of seismic events is notably high. Historical records and geological studies indicate that Italy has experienced numerous destructive earthquakes. Notable events include the 1908 Messina earthquake, which resulted in approximately 100,000 fatalities, and the 1980 Irpinia earthquake, which caused significant destruction and loss of life in southern Italy. More recently, the 2009 L'Aquila earthquake and the 2016 Central Italy earthquake sequence highlighted the ongoing seismic hazard in the region.

In this context, the importance of maintaining complete and accurate seismic catalogues becomes evident. Seismic catalogues are essential for understanding seismicity patterns, assessing seismic risks and developing effective mitigation strategies. They provide a systematic record of seismic events, including information on the location, magnitude, depth, and time of each event.

4.1.1 Italian catalogues

In Italy, thanks to the Istituto Nazionale di Geofisica e Vulcanologia (INGV) it is possible to manage data from different seismic catalogues and databases.

For seismic catalogue products, these includes:

- CPTI (Parametric Catalogue of Italian Earthquakes): This is a comprehensive parametric catalogue that provides detailed information on significant historical and recent earthquakes in Italy. The CPTI catalogue is crucial for long-term seismic hazard assessments and for understanding the historical seismicity of the regions [57]. The figure n. 4.1shows the catalogue plotted in the Italian territory.
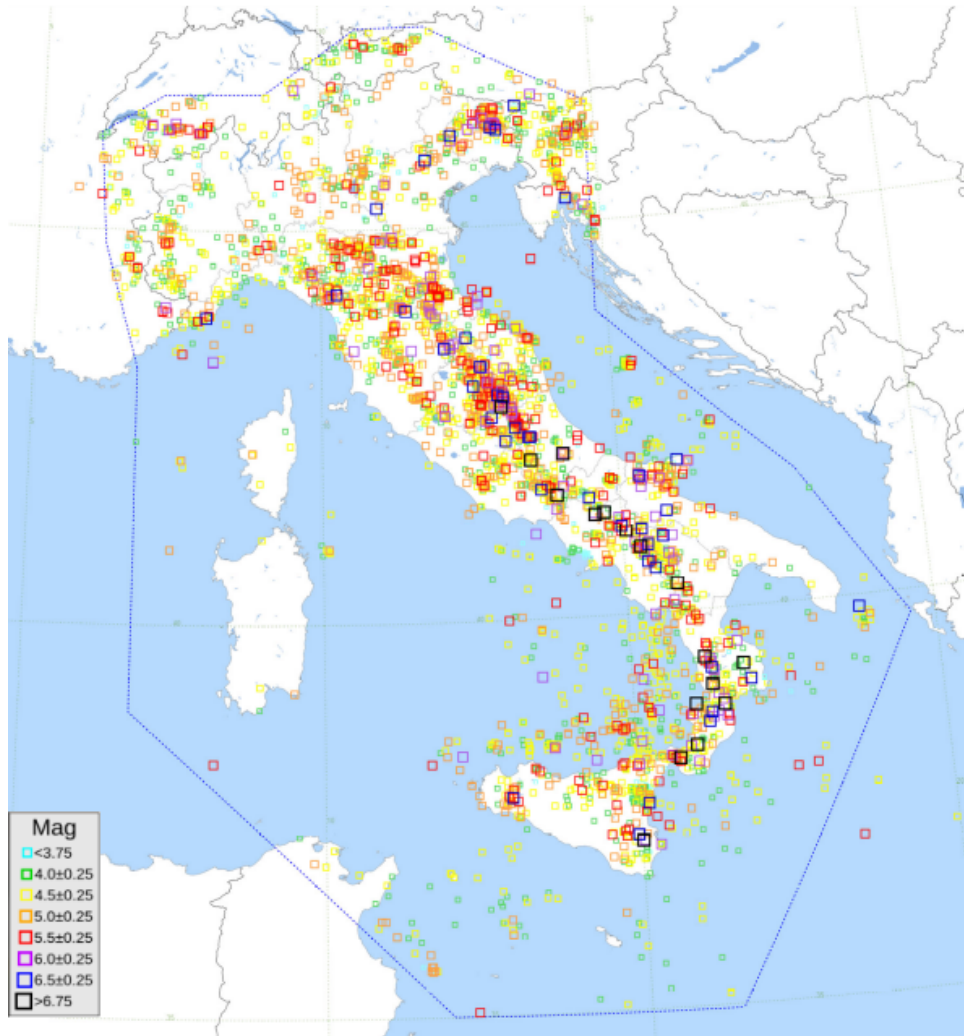
Figure 4.1: Parametric Catalogue of Italian Earthquakes

- Bollettino Sismico Italiano (BSI): It is a periodic bulletin that reports recent seismic activity in Italy, including detailed descriptions and analyses of significant events [34].

As far as database products are concerned, it is included:

- ISIDe (Italian Seismological Instrumental and Parametric Data-Base): It is an instrumental catalogue that contains real-time data from the national seismic network operated by INGV. It provides high-resolution information on seismic events detected by the network's seismometers. This catalogue is essential for monitoring ongoing seismic activity and for conducting detailed seismotectonic studies [30].

- The Historical Italian Earthquake Database (DBMI): It compiles historical earthquake data derived from historical documents and previous studies. It spans several centuries, providing a long-term perspective on seismic activity in Italy [35]. The figure n. 4.2 shows the catalogue plotted in the Italian territory. This is the database from which the previously mentioned CPTI catalogue is generated.
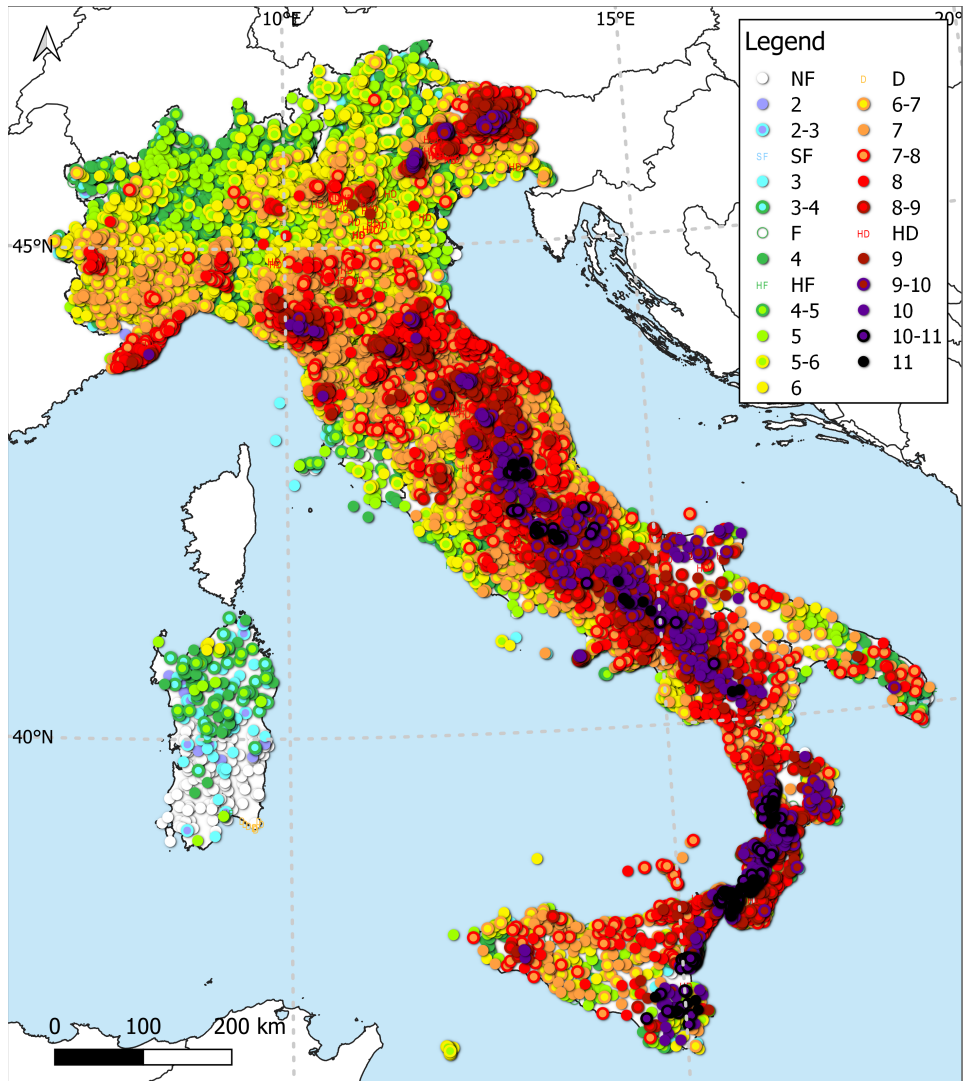
Figure 4.2: Historical Italian Earthquake Database

INGV's most important infrastructure is the National Seismic Network. Thanks to it, we manage to detect and trace all earthquakes and create the databases and catalogues described previously.

The National Seismic Network (Rete Sismica Nazionale, RSN) was established in the 1980s following the devastating earthquakes in Irpinia and Friuli Venezia Giulia. At that time, the network was known as the National Centralized Seismic Network (RSNC) and was managed by the National Institute of Geophysics (ING). Initially, the network comprised about 30 stations distributed across the country, operated by the seismic observatories of the time. The technology used in the seismic stations was rather basic: the seismometers were short-period types capable of recording only the vertical ground movement (Geotech S-13). The signal obtained was amplified and filtered to minimize interference, and then transmitted via telemetry using dedicated telephone lines or military radio links.

Over the years, thanks to economically significant projects and technological

advancements (such as MedNet and Cesis), and through synergy with Civil Protection, the seismic network has been expanded and improved using state-of-the-art technology. A significant expansion of the network occurred in 2005.

Today, the RSN consists of more than 600 stations scattered across the country. The histogram shown in figure n. 4.3 illustrates the evolution of the number of installed stations. As early as 2008, Amato [4] explained the evolution of the seismic network and how it has changed since the early 1980s. Today's technology is far more sophisticated than it was in the 1980s.

The latest generation of sensors is used, including various types of velocimeters, such as Nanometrics Trillium Compact 120s PH, Trillium-40s, Lennartz LE3D-20s, Trillium 360s GNS, Lennartz LE3D-5s, Trillium-120s, Lennartz LE3D-1s, Guralp CMG-3EX-120s, and Tellus 5s.

The accelerometers include Nanometrics Titan, LK-307 full scale 2G, and Episensor 2g, while acquisition systems such as Nanometric Centaur, Guralp, GAIA II, or AtlasF-6ch are also used.

Data transmission and power supply have been significantly improved: efforts are made to supply all stations with 220V power, supported by solar panels. Transmission is preferably done via LTE, taking full advantage of 5G technology when possible, or alternatively via satellite antennas or radio transmission. Data is transmitted in real-time, using secure protocols (TPC-IP), to the seismic operations room in Rome. This marks a significant advancement for the RSN infrastructure. Figure n. 4.4, shows an example of a multi-parameter station, containing seismic sensors and a geodetic sensor with a parabolic transmission. The reference station is ROVR, located in Rovere Veronese (VR). While in the other figure n. 4.5, we show an example of focus relative to the seismic sensors and the acquisition equipment, in this case we have a GAIA2-6 channel digitaliser and a velocimeter and an accelerometer. The reference station is SALO, located in Salò (BS).
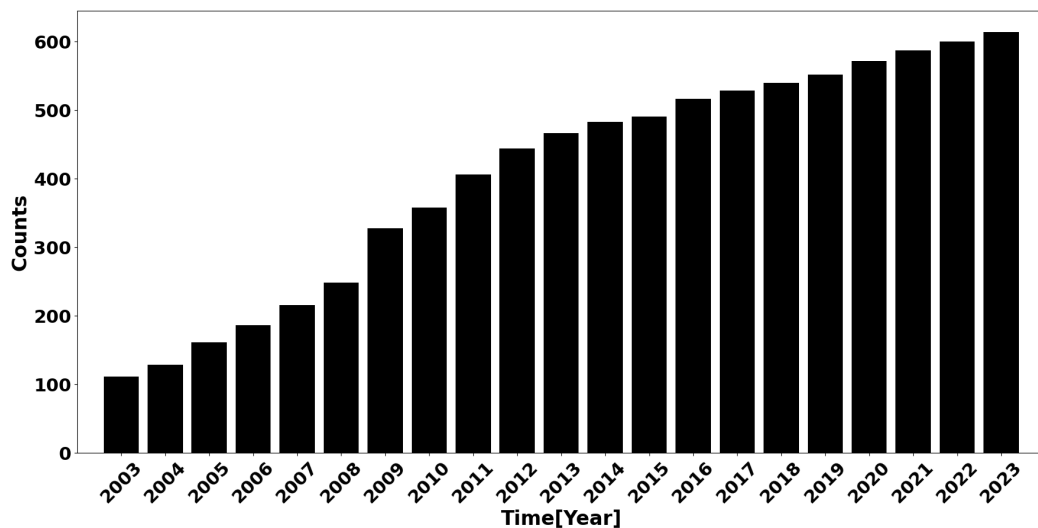
Figure 4.3: Evolution of Rete Sismica Nazionale from 2003 to 2023
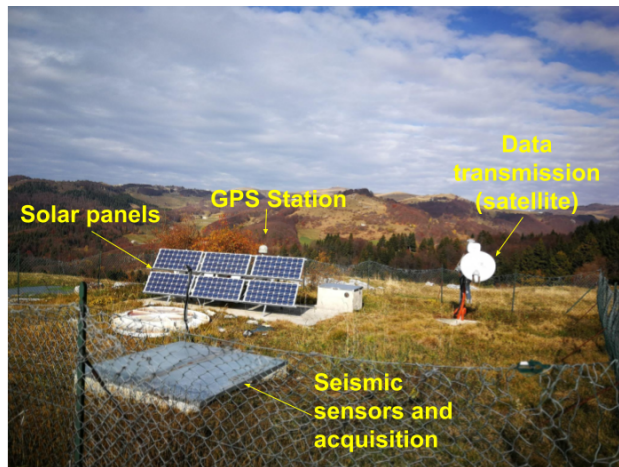
Figure 4.4: Example of a seismic station complete with seismic sensors and geodetic station. In this example, data transmission is by satellite and the power supply is supported by solar panels. This station is part of the National Seismic Network, with code ROVR.
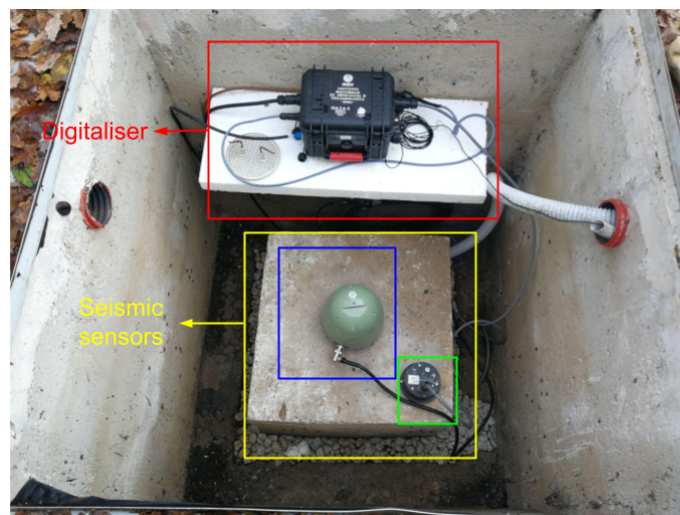
Figure 4.5: Example of seismic sensors and acquisition of a seismic station. The reference station is IV.SALO. The blue square represents a Trillium-120s velocimeter, while the green square represents an Episensor accelerometer. The red square shows a GAIA2-6 channel.

To achieve optimal seismic coverage, the decision to place new sensors must consider the need to primarily record teleseismic events rather than local or regional earthquakes. Teleseisms are earthquakes that occur at great distances from the detection point, usually over 1,000 km, and are useful for studying the Earth's internal structure. Local and regional earthquakes, on the other hand, occur at shorter distances (less than 100 km) and intermediate distances (up to 1,000 km) from the sensor, respectively, providing crucial data for assessing seismic risk in specific areas.

However, it is essential to make this choice carefully to create a sensor array with a sensible and logical geometry. The spatial distribution of the sensors must be carefully planned to ensure that the array can detect a wide range of seismic events without excessive overlap or uncovered areas.

Figure 4.6 shows the distribution and geometry of the RSN to date. This network is managed solely by the INGV, each section has its own number of stations to maintain. This network together with the other networks, as seen in the figure 4.7, contribute to the localization of earthquakes which is carried out in the seismic room in Rome.
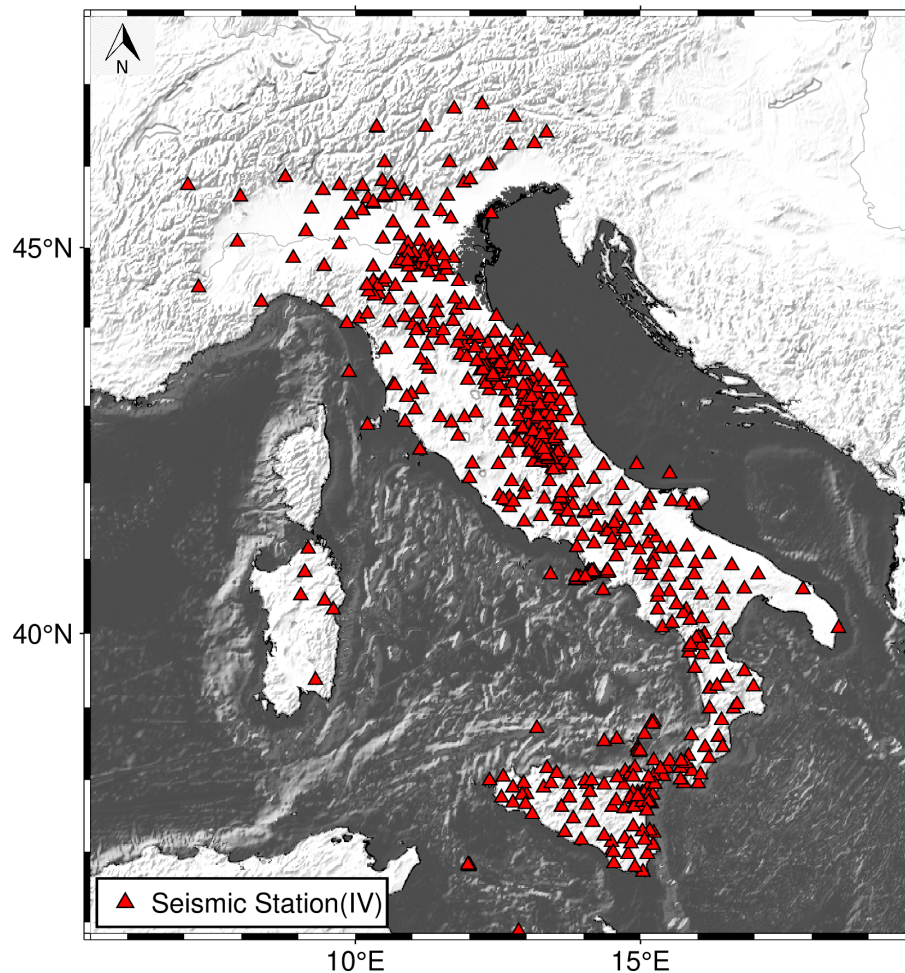
Figure 4.6: Italian Seismic network distribution. The red triangles represented the seismic stations. This map was created using Generic Mapping Tool6 (GMT)[72].
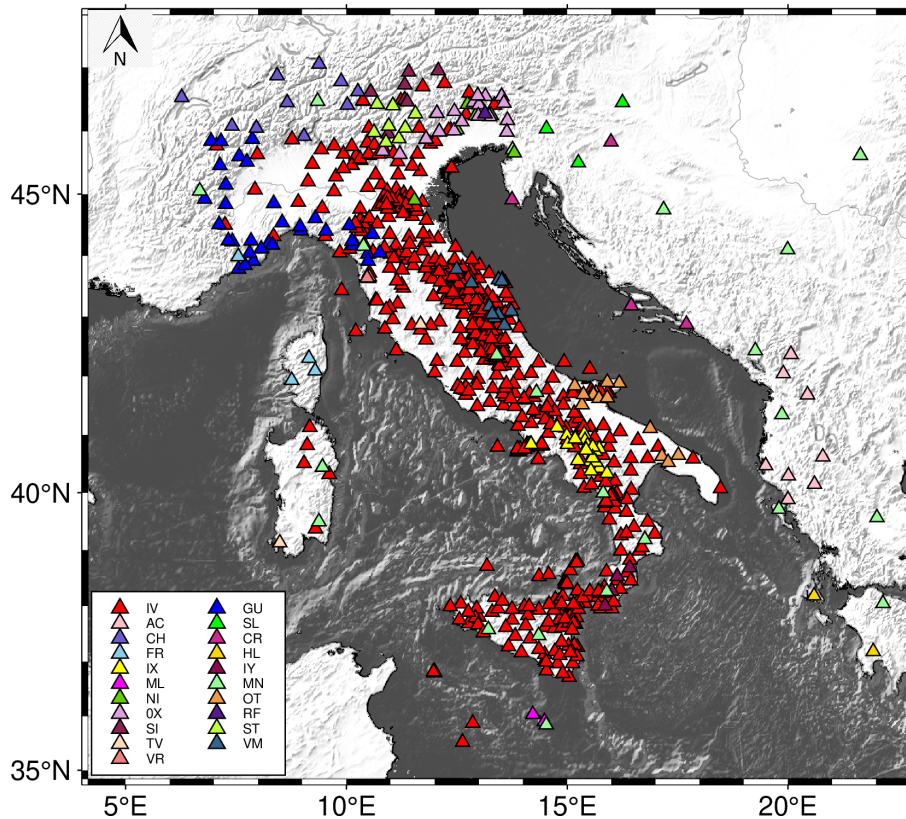
Figure 4.7: Distribution of seismic networks, including national (IV) but also other Italian and foreign networks.This map was created using Generic Mapping Tool6 (GMT)[72].

4.1.2   Italian database analysis for the Gutenberg-Richter law

In the course of the thesis, the Italian database was used to perform the completeness magnitude and b-value analyses.

Initially it was thought to consider and analyse the catalogue created by the INGV called INSTANCE, acronym of The Italian seismic dataset for machine learning [40]. This dataset had:

- 54,008 earthquakes for a total of 1,159,249 3-channel waveforms;

- 132,330 3-channel noise waveforms;

- 115 metadata for each waveform providing information on station, trace, source, path, and quality;

- 19 networks;

- 620 seismic stations

By calculating the frequency distribution magnitude, an anomalous trend relative to the completeness magnitude was noted, 4.8.
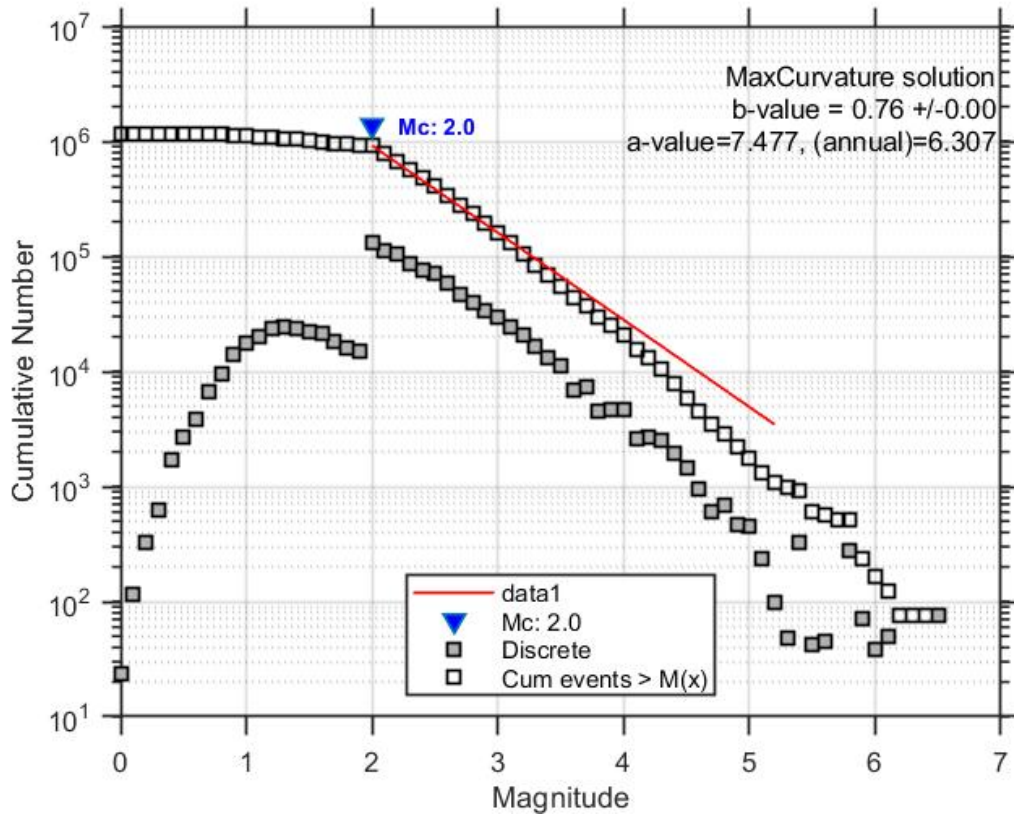
Figure 4.8: Frequency magnitude distribution of Instance Catalogue calculated of Italy seismicity

Going to read the manuscript relative to the dataset, it was noticed that the authors took into consideration different events than those of interest relating to the subject of the thesis. The authors took into consideration events with high magnitude, greater than 4 and / or greater than 5, to balance the number of minor events with the major ones. Furthermore, the localized events are not all falling within the Italian area, but we have events from other geographical parts.

For the purpose of this thesis, this catalogue homogenization technique is not good. So it was decided to fetch the dataset generated by the National Seismic Network. So the events were selected in the time range from 1985 to the present day. The catalogue was downloaded from the official earthquake website of INGV (`http://terremoti.ingv.it/`). Before carrying out the real analysis of the catalogue, the completeness magnitude of the whole range was plotted in time. As can be seen from figure n. 4.9, the value of Mc drops drastically from 2005 onwards.
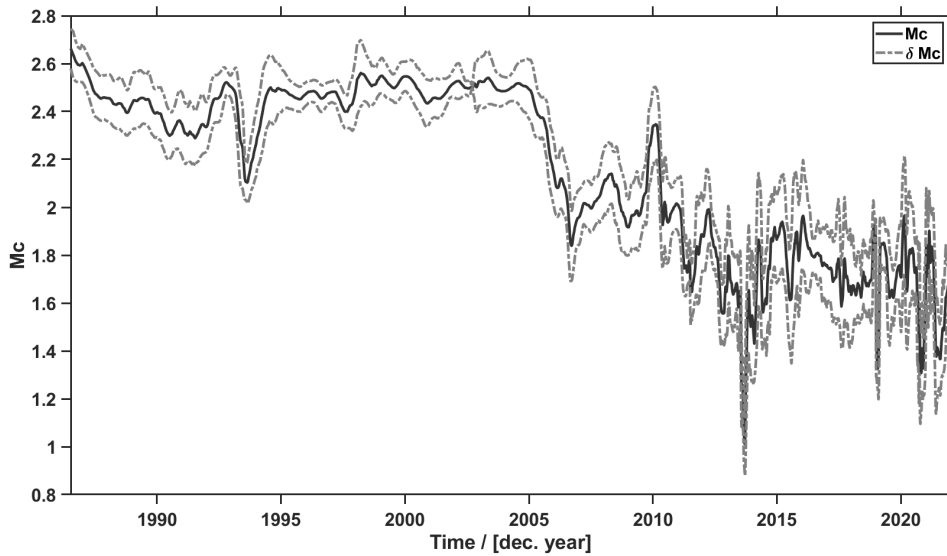
Figure 4.9: Times series of completeness magnitude from 1985 to 2022 for Italy

After some preliminary analysis, it was taken into consideration, a new catalogue, namely the Homogenized Instrumental Seismic Catalogue (HORUS). This catalogue takes into consideration all earthquakes, recorded and published, in ISIDE and the magnitude is recalculated. In this way, you will have a homogeneous catalogue with a moment magnitude. In the ISIDE catalogue, in most cases, the magnitude is local. But it can also happen to have the duration magnitude.

In this way, it was possible to update the seismic catalogue until 16 January 2024.

Given the range of analysis from April 2005 to 16 January 2024, the mainshock event with the highest magnitude was searched for, in order to eliminate it from the analysis. Once identified, the table within the work was consulted [20], where it allows us to keep track of how long a mainshock affected the entire seismic catalogue and at what distance. The new catalogue will therefore run from 28 December 2007 until 2024, as show figure n. 4.10.
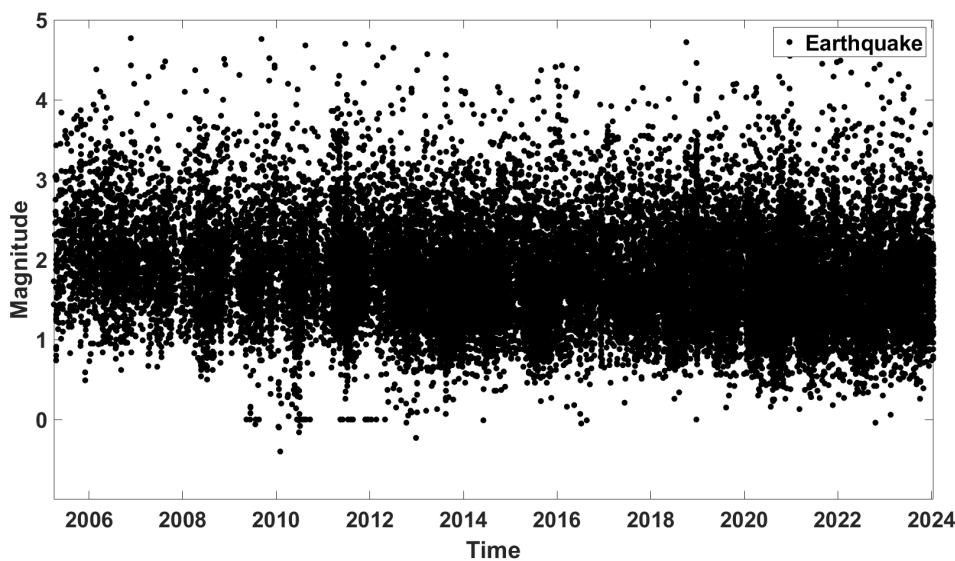
Figure 4.10: Time series of Sicilian seismicity considering the HORUS catalogue, from 2007 to 2024. The various black dots represent the various magnitudes in the catalogue.

In order to understand whether the study area had undergone geodynamic changes, the catalogue was divided into two parts.

48

The first catalogue starts from 2007 until 26 December 2015,as shows figure n. 4.11, the second from 27 December 2015 until 2024, as shows figure n. 4.12.
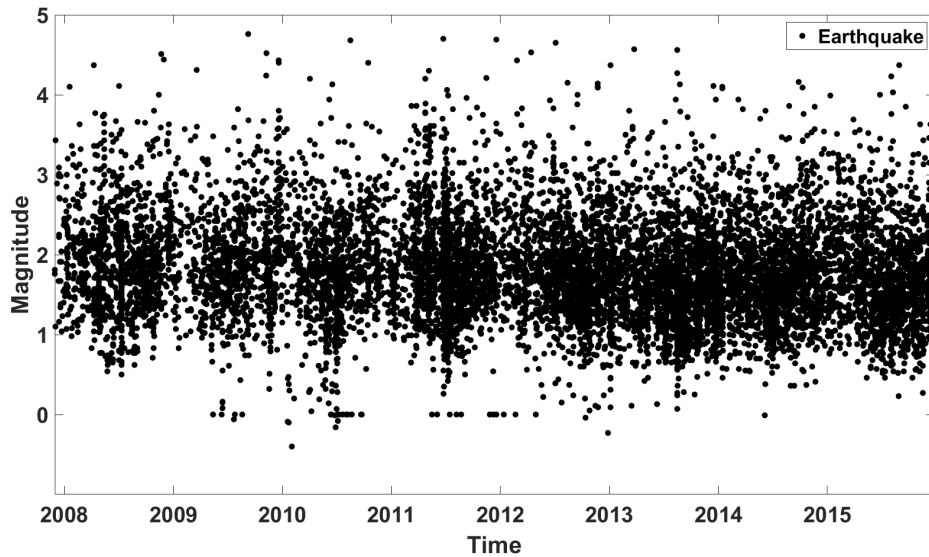
Figure 4.11: Time series of Sicilian seismicity considering the HORUS catalogue, from 2007 to 2015. The various black dots represent the various magnitudes in the catalogue.
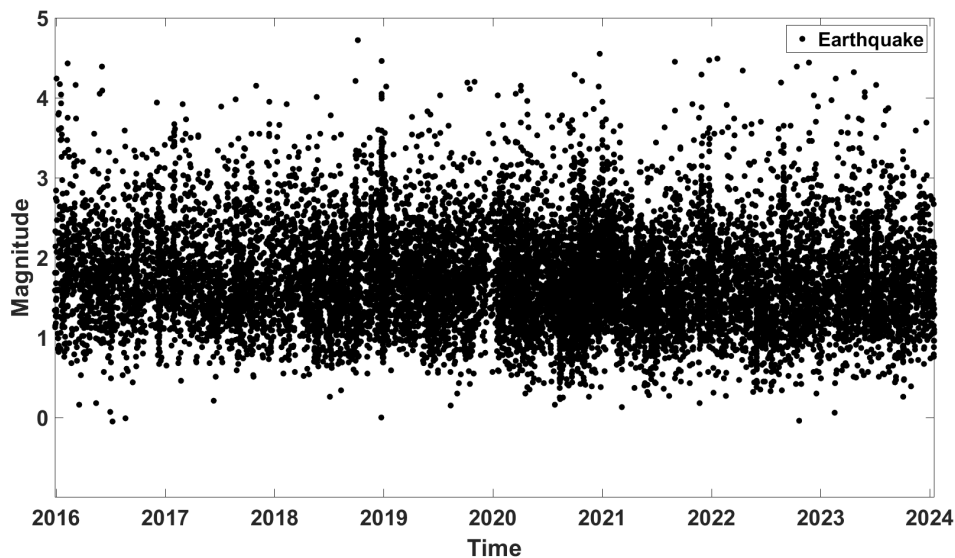
Figure 4.12: Time series of Sicilian seismicity considering the HORUS catalogue, from 2015 to 2024. The various black dots represent the various magnitudes in the catalogue.
All analyses performed are reported in chapter n. 5.

4.2   Taiwan

In this work, we conducted a comprehensive evaluation of the earthquake monitoring network coverage in Taiwan, a region that is historically and seismologically active. Taiwan's complex tectonic setting results in frequent seismic events, making an effective and optimally distributed seismic network essential for accurate monitoring and hazard mitigation. The latest significant earthquake, with a magnitude above 5, occurred on May 6, 2024. The seismic data analysed in this study were obtained from the Central Weather Administration Seismological Center `https://scweb.cwa.gov.tw/en-US`, which maintains a robust database of seismic events for the region.

The network's performance was assessed through a combination of descriptive spatial statistics and point pattern techniques, allowing for a detailed examination of the spatial distribution of the monitoring nodes in relation to seismicity, completeness magnitude (Mc), active seismogenic sources, seismic hazard, and population distribution. Specifically, we focused on two key components of the Taiwanese seismic network: the "Real-time Seismic Monitoring Network" and the "Strong-Motion Earthquake Observation Network," each designed with distinct objectives. These networks were compared with external information relevant to their purposes, such as seismic activity, identified seismogenic zones, and areas of high seismic risk.

The completeness magnitude was calculated using the MAXimum Curvature (MAXC) method in accordance with the Gutenberg-Richter law, as shows figure n.4.13. This analysis was based on 30 years of seismic data, encompassing approximately 790,000 events with magnitudes ranging from 0 to 7.3, as shows figure n. 4.14. The resulting Mc map, see figure n. 4.15, which indicates the detection capability of the network across different regions, was subsequently compared to the Kernel intensity estimation of the network's spatial distribution, providing a direct correlation between network density and seismic detection performance.
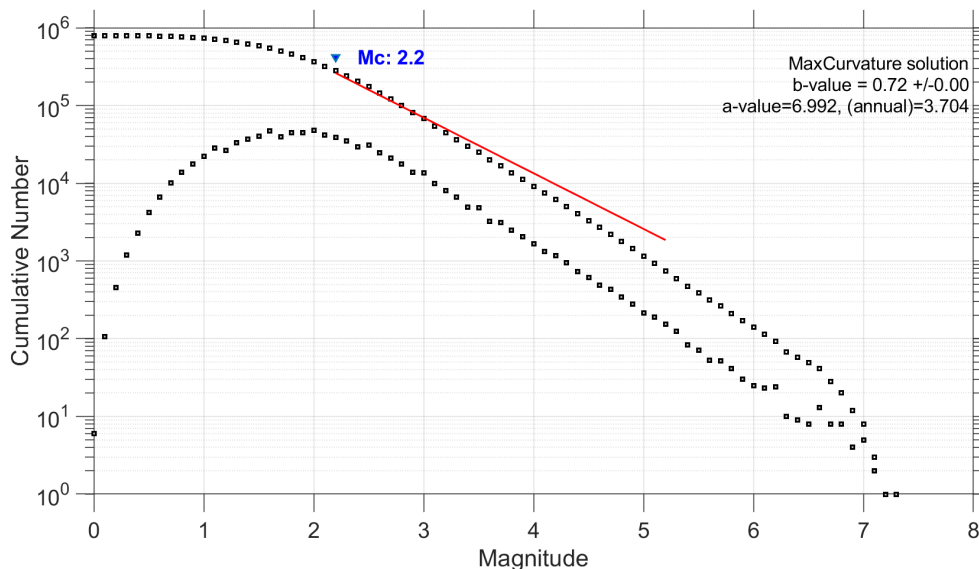
Figure 4.13: Frequency distribution magnitude about Taiwan catalogue.

Our approach, combining statistical analyses with real-world data, provides quantitative insights into the efficiency of Taiwan's seismic network. The results reveal a generally high level of coverage, though certain critical areas were identified for potential improvement. The correlation between the Mc map and the network's spatial distribution, visualized in blue for positive correlations and red for negative, highlights regions where the network performs optimally and areas where it may require future development. This method offers a reliable framework not only for assessing the current state of Taiwan's seismic network, but also for guiding its future expansion and optimization to ensure comprehensive earthquake monitoring.

This work, which has been published in the Journal of Seismology in 2023 [62], demonstrates the effectiveness of Taiwan's monitoring networks and their critical role in earthquake hazard assessment and mitigation. The draft is in the Appendix A.
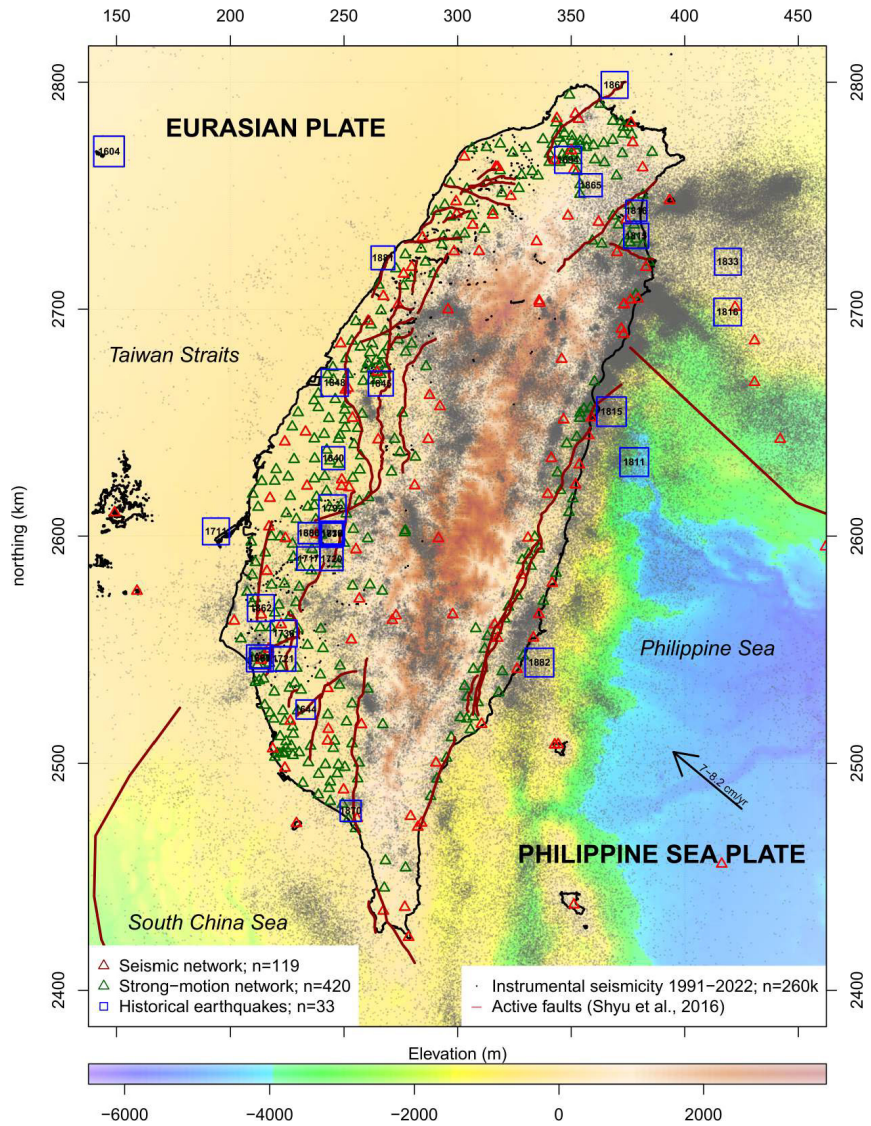
Figure 4.14: Seismicity of the island of Taiwan from the early 1990s until 2022
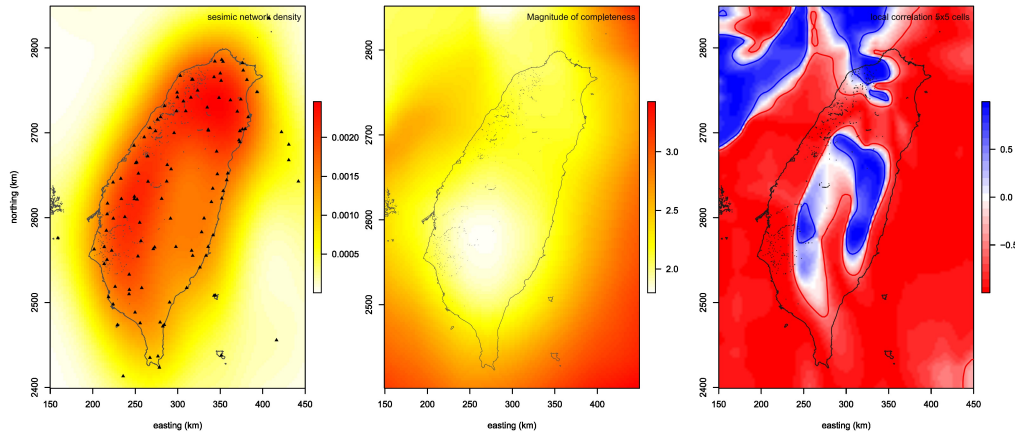
Figure 4.15: Panel showing on the left the density of the seismic network, in the centre the calculation of the magnitude of completeness and on the right the correlation between the two maps.

4.3 Iceland

Iceland is situated on the border between the North American and Eurasian plates. This singular location makes the island an ideal natural laboratory for the study of volcanic and seismic phenomena. In recent years, particularly since 2020, Iceland has experienced a series of volcanic and seismic events of considerable intensity. In 2020, a series of seismic swarms occurred on the Reykjanes Peninsula, some of them consisting of thousands of smaller seismic events concentrated in short periods of time. This sequence marked the beginning of a phase of unique seismic and volcanic activity in the region, previously unseen for about 800 years. In March 2021, after months of intense seismic activity, the Fagradalsfjall volcano on the Reykjanes peninsula itself erupted. This eruption, which lasted several months, was preceded by intense seismic activity with thousands of small earthquakes, some of which reached a magnitude of over 5.0. At the end of 2023/beginning of 2024, there were further seismic swarms of similar magnitude to the previous ones.

In order to study this seismicity, the database was downloaded from the Icelandic Meteorological Office (IMO) website (`https://en.vedur.is/`), which manages the seismic network distributed across Iceland. The database covers a time range from 2020 to 2024.
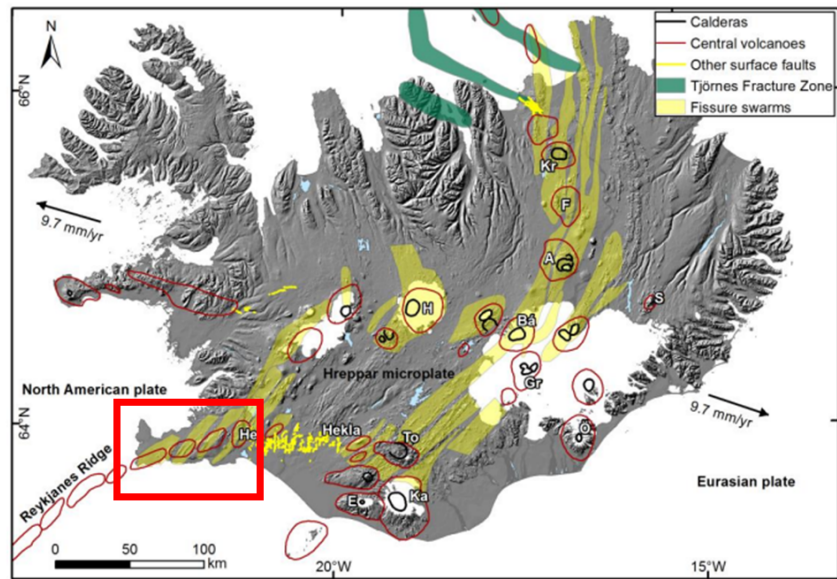
Figure 4.16: The island of Iceland, with its main tectonic and volcanic structures. The rectangle in red represents the study area i.e. Reykjavík
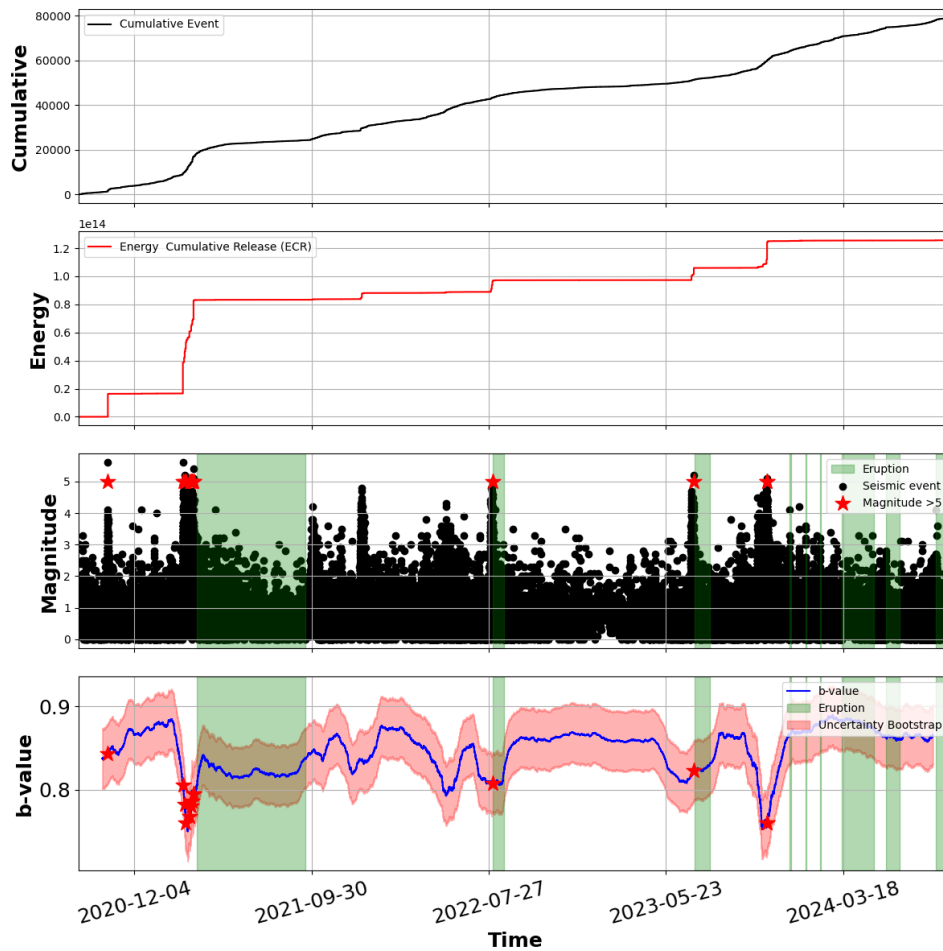
Figure 4.17: The panel shows four figures respectively: the top one shows the cumulative curve, the second the released energy, the third the seismicity and the fourth the b-value time series. The third and fourth show the red stars representing earthquakes of magnitude greater than 5 and the green rectangles the beginning and end of the eruptions.

The graph appears to show a correlation between high-magnitude seismic events and eruption periods. The cumulative energy released (ECR) increases with these higher magnitude events. The value of b appears to decrease just before large earthquakes and increase soon after, suggesting that it may be a predictive indicator for significant seismic or eruptive activity. In summary, the graph compares different aspects of volcanic and seismic activity in a specific area, indicating an interaction between seismic energy and eruptive events.

Focusing on specific time ranges, it can be observed:

December 2020 to May 2021:In this interval, the b-value starts above 0.9, but quickly drops to around 0.7 just before a strong earthquake (indicated by the red star at the beginning of 2021). After the event, the value of b starts to rise slightly, but still remains around 0.8. Several seismic events of varying magnitude are observed, with one significant event of magnitude greater than 5 at the beginning of the period. The energy released increases dramatically with the magnitude greater than 5 events at the beginning of 2021, leading to a sharp increase in the ECR. An eruptive phase (in green) is observed during this period, coinciding with the lowering of the b-value and the increase in earthquake magnitude.

May 2021 to February 2022: After the first eruptive phase, the b-value stabilizes slightly around 0.8. Between September and October 2021, there is a second decrease in the b-value, which falls again towards 0.7, coinciding with another significant seismic event. In this interval, a new event of magnitude greater than 5 occurs around July 2021. Seismicity increases around September 2021, with numerous small and medium magnitude events. An increase in cumulative energy occurs in July 2021, coinciding with the seismic event of magnitude greater than 5. After this increase, the energy stabilizes again. A second eruptive phase occurs between June and October 2021, coinciding with a new drop in b-value and increased seismic activity.

February 2022 to December 2022: During this period, the b-value rises and remains above 0.8 until mid-year, but a significant drop is observed from June-July 2022, falling below 0.7 again. This drop precedes another eruptive and seismic event. Between May and July 2022, another phase of high seismicity occurs, with numerous seismic events of varying magnitude. However, no events with a magnitude greater than 5 are recorded in this specific interval. The cumulative energy released remains relatively stable throughout most of 2022, with no significant increase until the end of the year. A brief eruptive phase is observed around July 2022, coinciding with a decrease in b-value and an increase in seismicity.

December 2022 to October 2023: From the beginning of 2023, the b-value remains relatively stable at around 0.8, with only small fluctuations. However, between May and June 2023, there is another drop in the b-value, falling to around 0.7, again coinciding with an eruptive phase. Another significant seismic event with a magnitude above 5 occurs in mid-2023. Overall seismicity remains high throughout the year, but with few significant fluctuations in magnitude. There is a slight increase in cumulative energy during this phase, but not as marked as in previous years. Another eruptive phase occurs between May and June 2023, accompanied by a decrease in b-value and an increase in higher magnitude seismic events.

October 2023 to March 2024: In this period, the b-value drops sharply below 0.7 towards the end of 2023, probably signalling the approach of a new significant eruptive or seismic event. Numerous seismic events are observed, with another earthquake of magnitude greater than 5 expected in early 2024. Seismic activity appears to be intensifying in the latter phase. Cumulative energy does not have a large increase at this time, but it is possible that it will increase following the next large magnitude event. Although no eruptive phases are observed at present, the b-value trend and the increase in seismicity suggest a

possible imminent eruptive phase.

The b-value regularly drops below 0.7 before significant seismic events and eruptive phases, while it stabilizes around 0.8 or more after the release of energy. Seismic events of magnitude greater than 5 are rare, but tend to occur in correspondence with drops in b-value and increases in cumulative energy. Volcanic eruptions are closely associated with decreases in b-value and increases in seismic activity.

4.4   Features comparison of seismic catalogues and regions studied

The three regions analysed, Sicily, Taiwan, Iceland, have profoundly different seismotectonic and geophysical characteristics, which necessitated the use of different methodological approaches for each seismic catalogue.

In Sicily, the methodology was more diversified. The analysis required the integration of different Italian seismic databases, with the aim of providing as complete a framework as possible of the seismicity of the area. Traditional seismicity analysis techniques were applied, such as the application of the Gutenberg-Richer law, but also more advanced methodologies to identify seismic clusters and event sequences. This allowed us to understand how the various techniques applied can influence the analysis of the law parameters, and to understand the geodynamics of the area and the quality of the national network.

In Taiwan, the focus was on the magnitude of completeness (Mc), which was compared with the density of Taiwan's seismic network. The comparison of network density and magnitude of completeness provided valuable information on where the seismic coverage could be improved to obtain a more accurate estimate of local seismicity.

Lastly, in Iceland, the analysis focused on the evolution of the b-value, an essential parameter for understanding changes in seismicity related to volcanic phenomena. The b-value reflects the distribution of magnitudes within a seismic catalogue, and significant changes in this parameter may indicate changes in the stress regime or ongoing geodynamic processes. In a context such as Iceland, where seismicity can be strongly influenced by volcanic eruptions, this analysis was crucial to correlate seismicity with volcanic phenomena and mainshock tectonic events. In particular, it has been observed how, in correspondence with eruptions or magmatic intrusions, the b-value can undergo significant variations, indicating a change in the size and frequency of earthquakes.

Chapter 5

Statistical modelling of seismological parameters: Completeness Magnitude, *b*-value, and aftershocks

This chapter explores various data analysis techniques through the application of statistical models on key seismological parameters, with a focus on magnitude of completeness and *b*-value. The main objective is to analyse how these models can contribute to a better understanding of the distribution and frequency of seismic events. In particular, the first part of the chapter focuses on using statistical models to estimate the magnitude of completeness, an essential parameter for determining the threshold above which all seismic events are reliably detected. Next, we move on to the analysis of the *b*-value, which represents the rate of decay of the frequency of seismic events as a function of their magnitude, using statistical models to explore its variations. Another central theme is the application of the Epidemic Type Aftershock Sequence (ETAS) model, which allows the description and prediction of aftershock time sequences after a main earthquake, improving the understanding of the dynamics of secondary seismic events. Finally, an accurate probabilistic analysis for detecting the completeness magnitude is presented, an innovative approach compared to traditional methods, which offers more flexibility in identifying the completeness threshold.

5.1 Application of statistical models on magnitude of completeness

5.1.1 Structural geology of Sicily

The structural geology of Italy is fundamental in understanding the tectonic processes that have shaped the Italian territory over the past millennia. Italy is characterized by a complex network of geological structures that reflect the interaction of tectonic plates, the convergence of the Eurasian and African plates, and associated seismic activity. Sicily, the largest island in the Mediterranean Sea,is a micro laboratory of this structural complexity, with unique geological features reflecting both regional and local tectonics, as shows figure n. 5.1.

Northern Italy is characterized by the Alps, a major mountain range resulting from the collision of the Eurasian and African plates. Tectonic deformation resulted in folds and reverse faults on a regional scale. The Apennines, on the other hand, cross the Italian

peninsula in a north-south direction and reflect a complex tectonic history characterized by extension, compression, and transcurrent movements, with the presence of normal, reverse, and strike-slip faults.
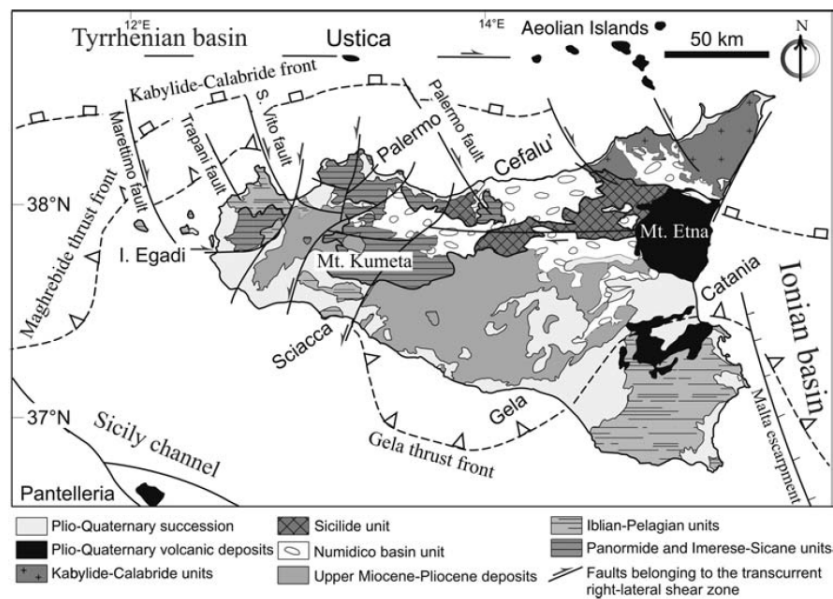
Figure 5.1: Map of the main tectonic structures in Sicily [69]

The structural setting of northern Sicily results from the Cenozoic collision between the North African continental margin and the Sardinia–Corsica block. The main tectonic units derived from the deformation of the northern margin of the African plate show southward-verging folds and thrusts extending for tens of kilometers. The region is characterized by several phases of tectonic activity:

Thrust tectonics from the Early Miocene. Extensional tectonics and crustal thinning from the Late Miocene. Strike-slip tectonic activity during the Plio-Pleistocene, reactivating inherited structures in the region separating the southern Tyrrhenian Sea from northern Sicily. The Sicilian-Maghrebian Chain, located along the northern coast of Sicily, is affected by these processes and has developed a right-lateral shear zone since the Pliocene, known as the Ustica–Eolie and Kumeta–Alcantara shear zones. These strike-slip fault systems, including the Marettimo, Trapani, San Vito, and Palermo faults, form a structural complex that controls the deformation in the area.

Mount Etna, one of the most active volcanoes in the world, plays a significant role in the geological landscape of Sicily. It is a stratovolcano formed by a complex interplay of tectonic and volcanic processes, including the subduction of the African plate beneath the Eurasian plate. The presence of the Ionian slab, which is the remnant of the subducting African plate, significantly influences the volcanic activity and tectonic stability of the region. The Iblean-Maltese escarpment, characterized by steep cliffs and tectonic uplift, provides insights into the geological evolution of the area and the effects of regional tectonics on volcanic activity.

5.1.2 Seismicity of Sicily

The seismicity in Sicily is characterized by significant variability, with several notable historical events underscoring the region's earthquake activity. The 1968 Belice earthquake (Mw 5.4) remains the strongest historical event in the region. More recently, the 2002 Palermo seismic sequence (Mw 5.9) was recorded along the northern coast, marking the largest seismic event in the area. Other significant historical earthquakes include the

1908 Messina earthquake (Mw 7.1), which devastated eastern Sicily, and the 1693 Sicily earthquake, which had a magnitude of approximately 7.4 and caused extensive damage in Catania and surrounding areas.

Several more recent seismic events are as follows:

- 2002 - Southern Tyrrhenian Sea Earthquake: On 27 September 2002, a magnitude 4.3 earthquake occurred in the southern tyrrhenian sea.

- 2007 - Earthquake of 5.2: On 5 July 2007, an earthquake with a magnitude of 5.2 occured near the Eolian Islands.

- 2009 - Palermo Earthquake: On 8 November 2009, an earthquake occurred in the Cerami area (EN) with a magnitude of 4.4.

- 2012 - An earthquake occurred on 13 April 2012 off the coast of Palermo, with a magnitude of 4.3.

- 2016 - An earthquake occurred on 8 February 2016 in Palazzo Acreide(SR) with a magnitude of 4.2.

- 2018 - Earthquake of 4.6: On 6 October 2018, a magnitude 4.6 earthquake struck Ragalna area (CT). This event resulted in minor injuries to buildings but no significant damage.

- 2018 - Viagrande Earthquake: On 26 December 2018, a magnitude 4.9 earthquake occurred near Viagrande, close to Mount Etna. The tremor was felt throughout the Catania province, causing concern among the population but resulting in no significant damage to structures.

- 2019 - Milo Earthquake: An earthquake occurred on 9 January 2019 in Milo(CT) with a magnitude of 4.1.

The figure below 5.2 shows the seismicity of the instrumental catalogue taken into account for the completeness magnitude and b-value analyses. As can be seen, the seismicity is distributed over almost the entire Sicilian territory. We have a prevalence of surface seismicity, about 50 km deep, and another type of deeper seismicity, caused by the deep subduction of the Ionic slab [52, 58]. in addition, the main geological structures, taken from the database of individual seismogenic sources (DISS) [17] and also the national seismic network distributed in Sicily and part of Calabria, are shown.
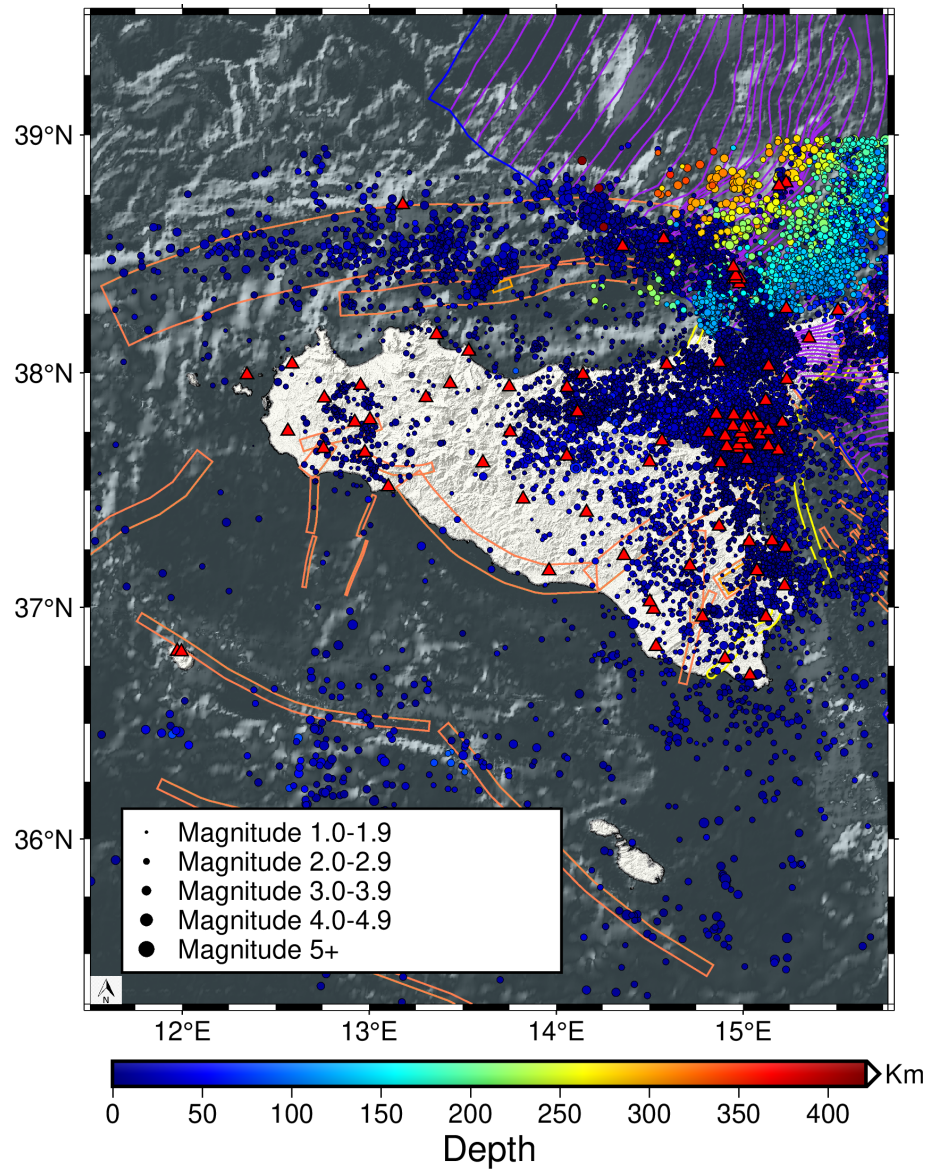
Figure 5.2: Distribution of seismicity in Sicily from 2005 to 2022.The red triangles represent the seismic stations. The change in colour is a function of earthquake depth, the size a function of magnitude.This map was created using Generic Mapping Tool6 (GMT)[72].

5.1.3 Results

The first calculation performed was the analysis of the frequency magnitude distribution using the MAXC method, as shows in figure n.5.3. This calculation was corrected with a corrective value of +0.4.
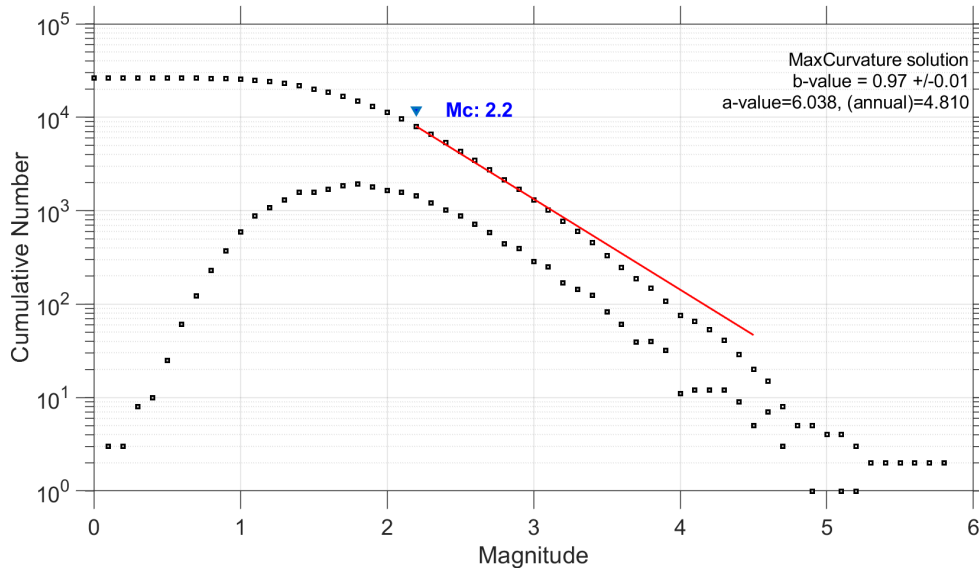
Figure 5.3: Frequency magnitude distribution of the ISIDE catalogue calculated in Sicily from 2005 to 2022.

Preliminary maps, created during the first year of the doctoral research, are presented in figure n.5.4 and figure n.5.5. For the analysis conducted in the case study, various maps were generated, focusing on the spatial distribution of the b-value and the magnitude of completeness (Mc). The initial maps were produced using a constant radius of 100 events and a node spacing of 10 km, with a grid resolution of 5 km.

After performing several experiments and adjusting input parameters, a final map was produced. The optimal configuration uses a grid resolution of 4x4 km, 100 events per node, and a constant radius of 80 events, with a correction of Mc by 0.2.
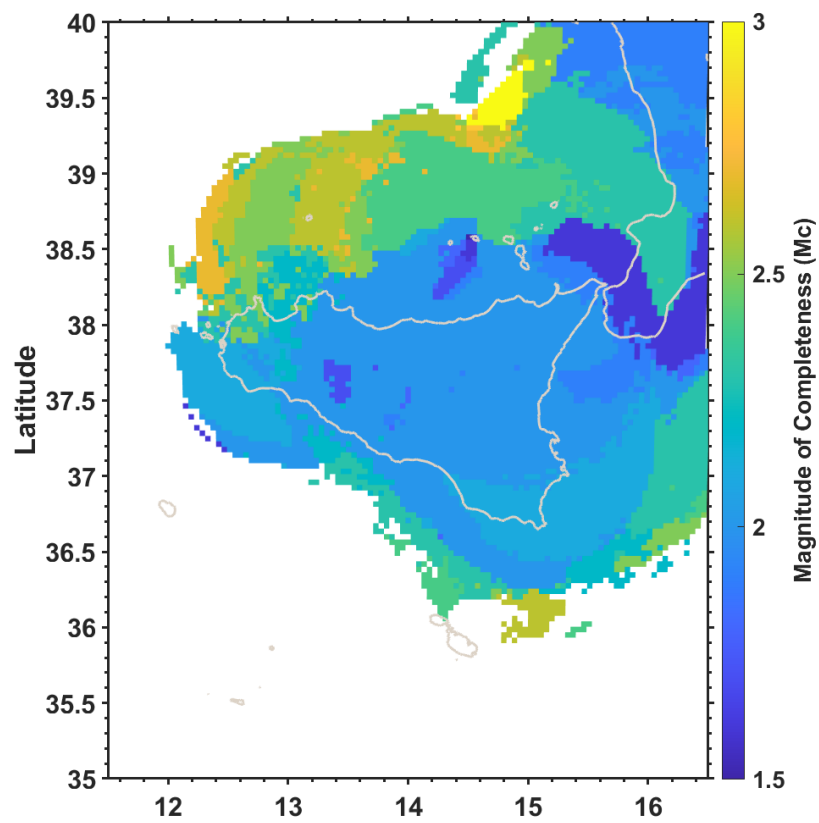
Figure 5.4: Map of magnitude of completeness having a resolution of 4X4 km and a constant radius of and a minimum number of 100 by node.The calculation time is given by the interval from 2005 to 2024, considering the ISIDE catalogue.
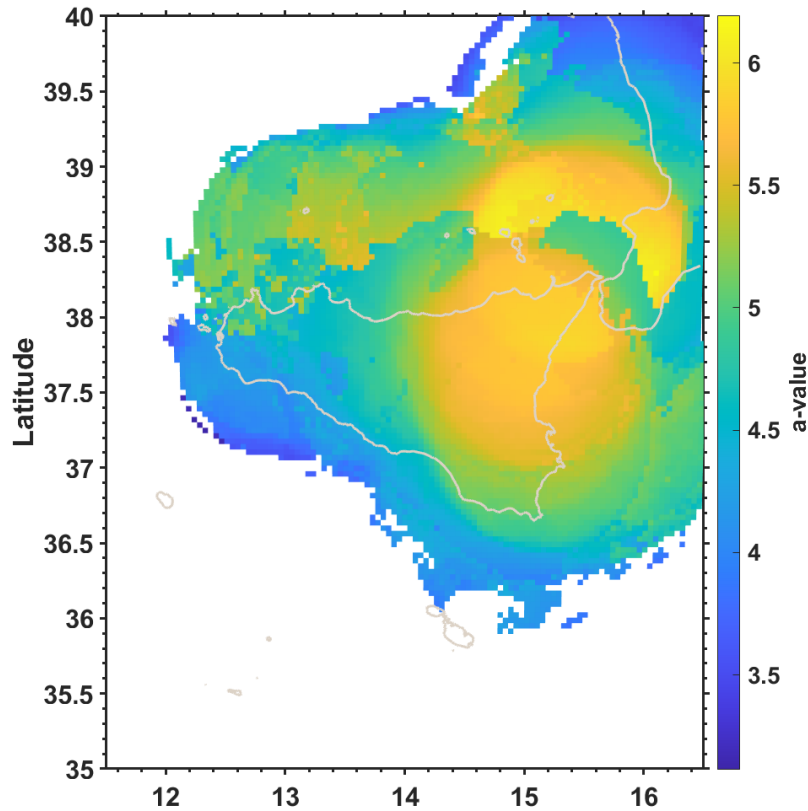
Figure 5.5: Map of a-value having a resolution of 4X4 km and a constant radius of and a minimum number of 100 by node.The calculation time is given by the interval from 2005 to 2024, considering the ISIDE catalogue.

After obtaining preliminary maps using the ZMAP software [73], these maps were revised with the use of the new tool [19].

In this new analysis, a few aspects have been taken into account. In order to better select the best resolution for an accurate analysis of the magnitude of completeness maps, two types of analysis were developed.

The first involves selecting the minimum events of 50 within a cell with a radius (R) that progressively increases by 5 units at a time, as shown in figure n. 5.6. The second involves selecting the minimum events of 100 within a cell with a radius (R) that progressively increases by 5 units at a time, as shown in figure n. 5.7.

Increment of R to 5
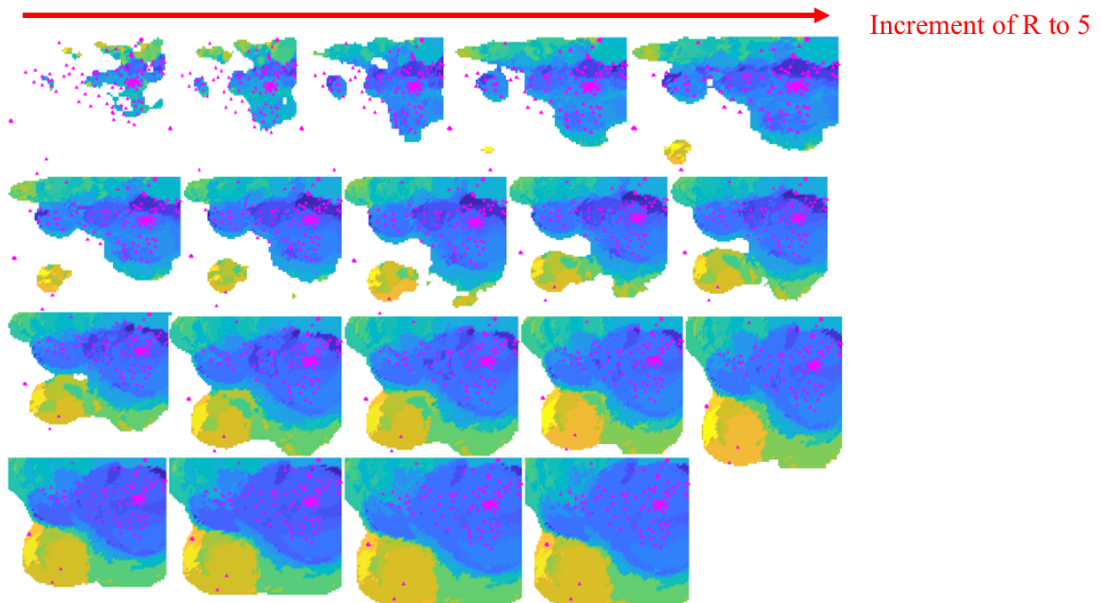
Figure 5.6: Completeness magnitude maps having a minimum number of events of 50, with a progressive radius of 5 units starting from 20.

Increment of R to 5
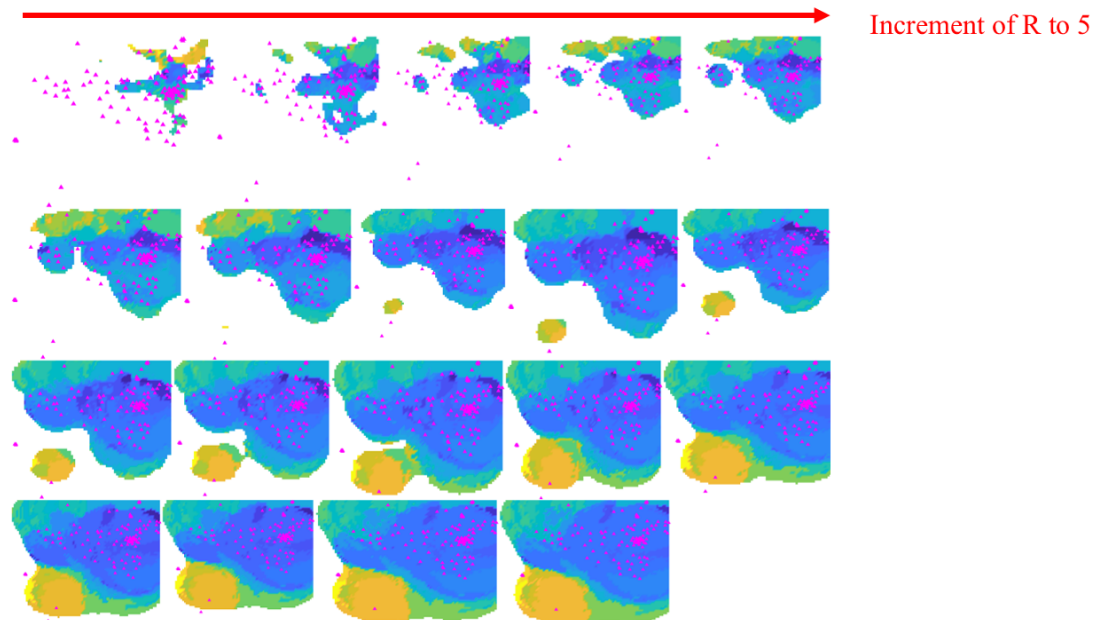
Figure 5.7: Completeness magnitude maps having a minimum number of events of 100, with a progressive radius of 5 units starting from 20.

At the end of the preliminary analysis, it was decided to consider maps with a minimum number of events of 100 and a radius of 20.

The figure n. 5.8 represents the completeness magnitude (Mc) calculated for the entire seismic catalogue over the region of Sicily, based on the data from the seismic network. The red triangles represent the seismic stations. The colour bar on the right indicates the range of Mc values, ranging from approximately 1.4 to 3. High Mc values (in yellow/orange, 2.8-3) are observed mainly in the northeastern part of Sicily (around 38.5°N and 16°E), indicating that the detection threshold for smaller earthquakes is higher in this region. Moderate Mc values (green/light blue, 2.2-2.6) cover most of the central and eastern parts of Sicily, especially between 38°N and 39°N, and extending between 14°E and 16°E. Low Mc values (purple/blue, 1.4-2.0) are concentrated in the western part of the island , like Belice area, and in the eastern part like Messina and Calabria, suggesting that the seismic network in these areas is more sensitive to detecting smaller earthquakes.
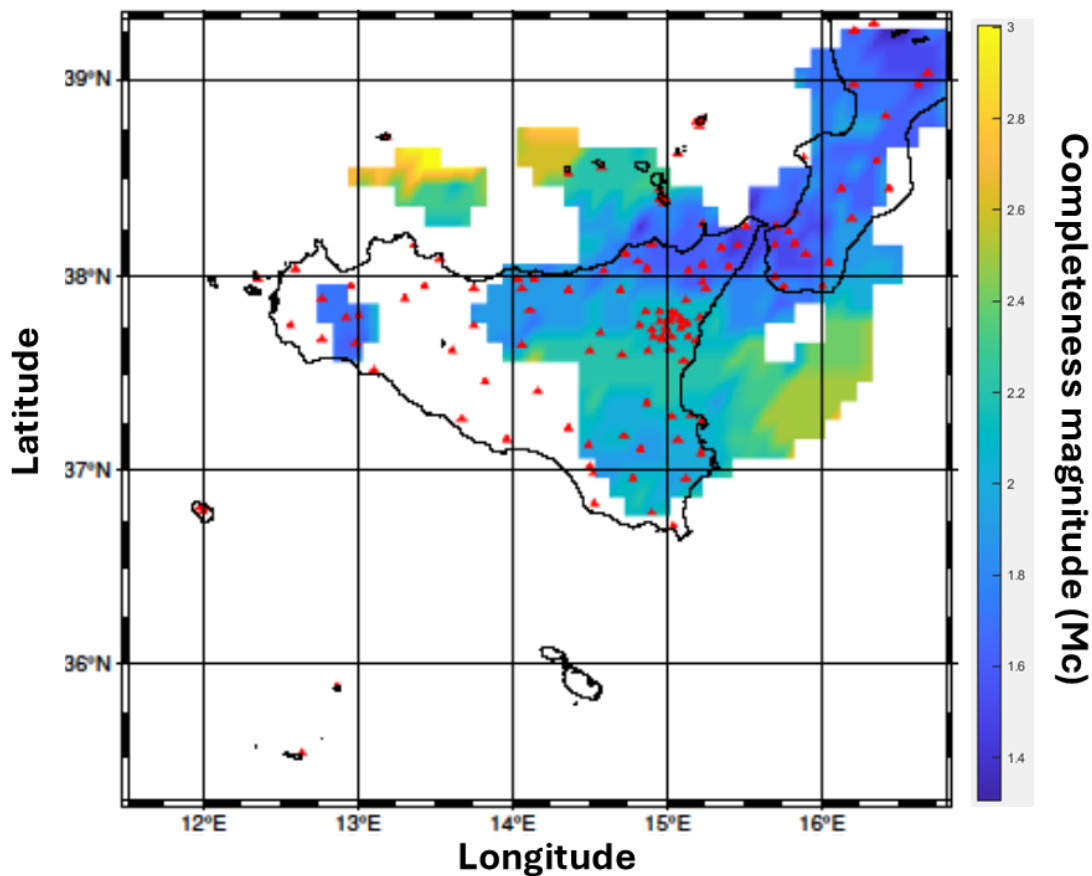
Figure 5.8: Map of completeness magnitude by selecting a minimum number of events of 100 and a maximum radius of 20 km.The analysis was conducted for the entire seismic catalogue (Horus).

For figure n. 5.9, the catalogue from 2007 to 2015 was taken into account. Compared to figure n.5.8, it can be seen that the Mc values tend to be lower, with the largest survey area remaining in eastern Sicily (38°-39°N, 15°-16°E) peaking at around 1.4. For the South-Eastern area(38°-37°N, 15°-16°E), values around 2.4-2.6 are observed. For the Aeolian islands (38°-39°N, 14°-15°E), on the other hand, a Mc of about 2.8 is recorded.
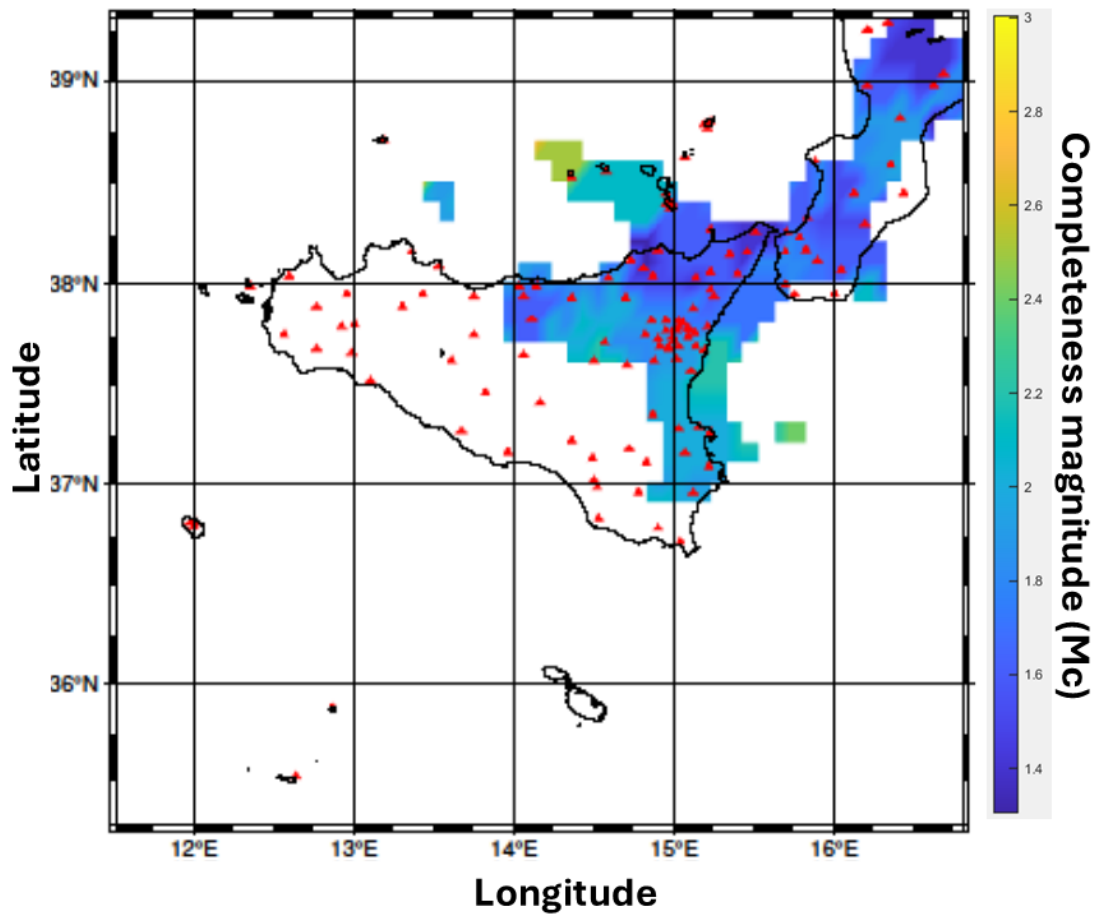
Figure 5.9: Map of completeness magnitude by selecting a minimum number of events of 100 and a maximum radius of 20 km.The analysis was conducted for the entire seismic catalogue (Horus). The analysis was conducted for the time interval from 2007 to 2015.

For figure n. 5.10, the catalogue from 2015 to 2024 was taken into account. Here it is possible to see that the Mc values tend to be lower, with the largest survey area remaining in eastern Sicily (38°-39°N, 15°-16°E) peaking at around 1.4. For the South-Eastern area(38°-37°N, 15°-16°E), values around 2.4-2.6 are observed. For the Aeolian islands (38°-39°N, 14°-15°E), on the other hand, a Mc of about 2.4-2.8 is recorded. In the western part of Sicily(38°-37°N,12°-13°E) it is possible to see a small calculation of Mc, with values ranging from 2.0 to 2.4
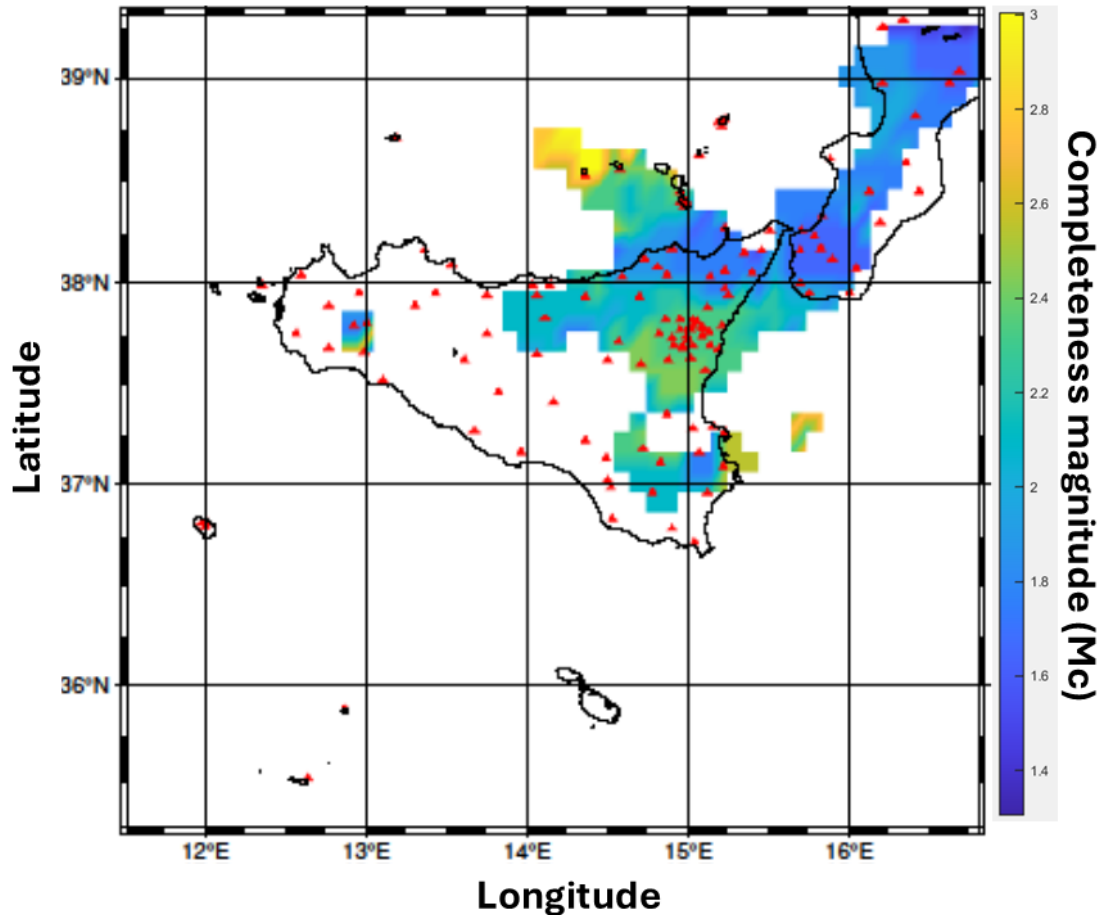
Figure 5.10: Map of completeness magnitude by selecting a minimum number of events of 100 and a maximum radius of 20 km.The analysis was conducted for the entire seismic catalogue (Horus). The analysis was conducted for the time interval from 2015 to 2024.

5.2 Application of statistical models on *b*-value

In this subsection, the results of the spatial and temporal analysis concerning the *b*-value are presented.

First to show how *b* varies spatially in the area under study, Sicily.

The figure n.5.11, is the result of the preliminary analysis carried out at the beginning of the first year of the doctorate. The input parameters are the same as those used and described in subsection n.5.1.3.

Figure n.5.11 illustrates the variation in *b*-value across the entire catalogue, spanning from 2005 to 2022. Notably, higher *b*-values, ranging between 1.1 and 1.3, are observed in regions such as Calabria (38°-39°N,16°-14.5°E), Ragusa(37°-36.5°N,16°-14.5°E), and the southern Tyrrhenian Sea(38°-39°N,16°-12°E). Intermediate *b*-values, ranging from 0.9 to 1.1, are observed in central Sicily (38°-37°N,15,5°-14,5°E), while the lowest values, between

0.6 and 0.7, are found in parts of Calabria(38,5°-38°N,16°-15,5°E).

The following maps were made with the TREMORS tool [19], and are made by first considering the entire current catalogue, i.e. from 2005 to 2024. The second is made considering a time range from 2007 to 2015 and the third from 2015 to 2024.
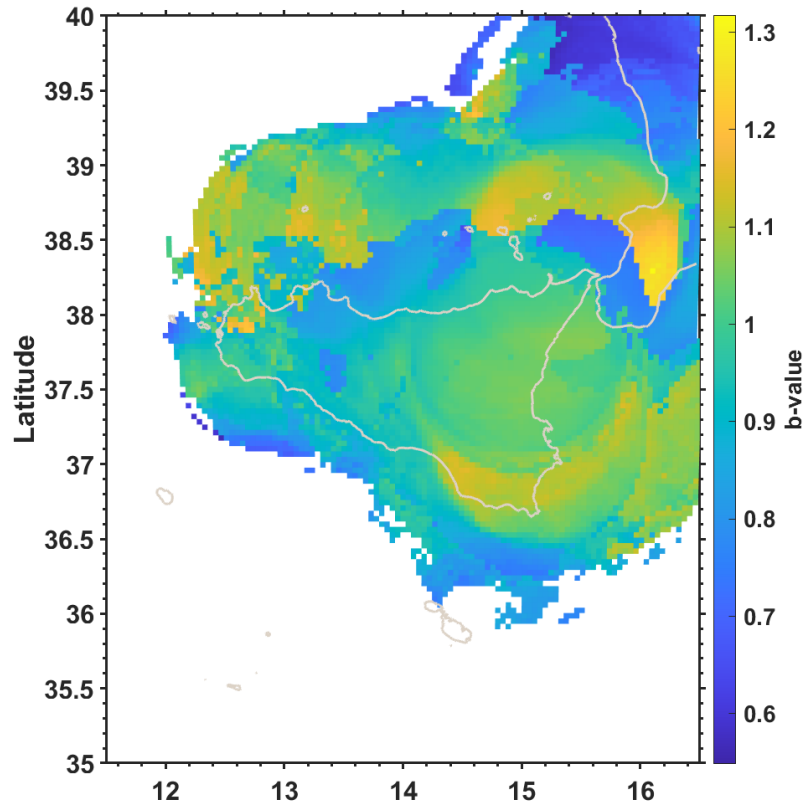
Figure 5.11: Map of *b*-value having a resolution of 4X4 km and a constant radius of and a minimum number of 100 by node.The calculation time is given by the interval from 2005 to 2024, considering the ISIDE catalogue.

In the figure n.5.12 higher *b*-values (0.85-0.95) are concentrated in eastern Sicily(38°-37°N,15,5°-14,5°E), particularly around Mount Etna and north-eastern regions. The central regions of Sicily(38°-37°N,15°-14,9°E) exhibit intermediate *b*-values (0.7-0.8), indicating a more balanced distribution between smaller and larger earthquakes. Lower *b*-values, ranging from 0.6 to 0.7, are observed in southern Sicily(38°-37°N,13°-14°E).
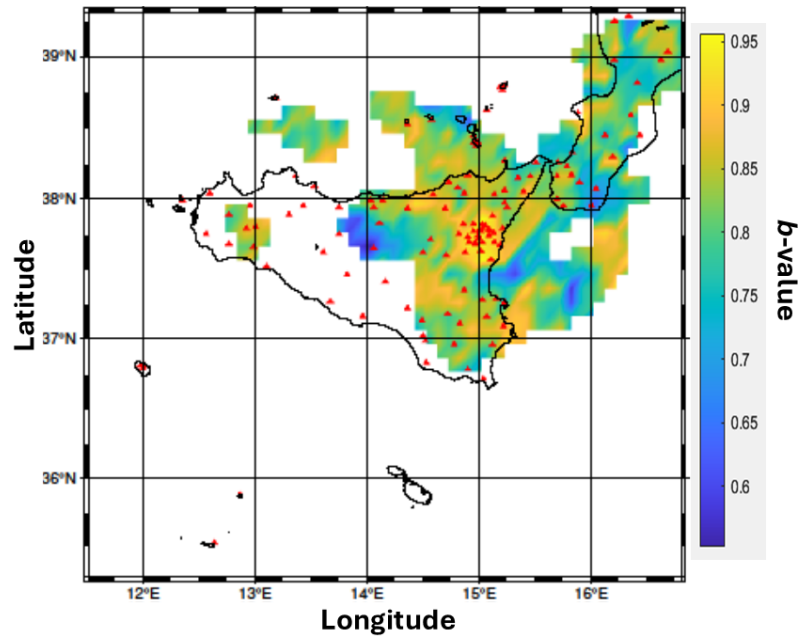
Figure 5.12: Map of *b*-value by selecting a minimum number of events of 100 and a maximum radius of 20 km.The analysis was conducted for the entire seismic catalogue (Horus).

In the figure n.5.13 higher *b*-values (0.85-0.95) are concentrated in eastern Sicily, particularly around Calabria and north-eastern regions(39°-38°N,15°-16°E), suggesting a predominance of smaller seismic events in these areas. Aeolian Islands (39°-38°N,14°-15°E) exhibit intermediate *b*-values (0.7-0.8), indicating a more balanced distribution between smaller and larger earthquakes. Lower *b*-values, ranging from 0.6 to 0.7, are observed in the Ionian Sea (38°-37°N,15°-16°E)and in western sicily (38°-37°N,13,9°-14,2°E).
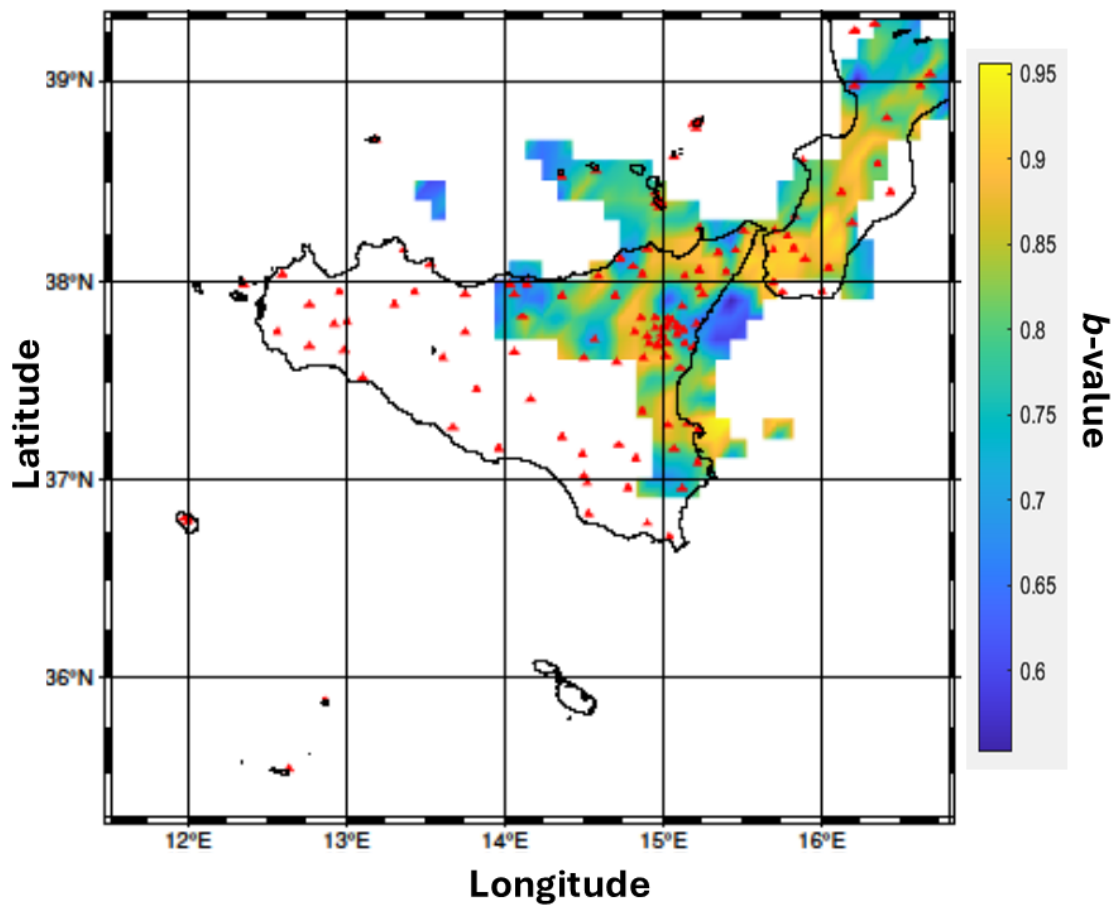
Figure 5.13: Map of *b*-value by selecting a minimum number of events of 100 and a maximum radius of 20 km. The analysis was conducted for the time interval from 2007 to 2015 (Horus).

In the figure n.5.14 higher *b*-values (0.85-0.95) are seen mainly in eastern Sicily, particularly around Mount Etna, the Aeolian Islands, and north-eastern Sicily. These areas are likely experiencing frequent smaller earthquakes, consistent with volcanic activity. Central Sicily and portions of the southern Tyrrhenian Sea show intermediate *b*-values (0.7-0.8). Lower *b*-values (0.6 to 0.7) are concentrated in southern Sicily, including Ragusa and western parts of the island, suggesting a higher probability of larger earthquakes in these areas.
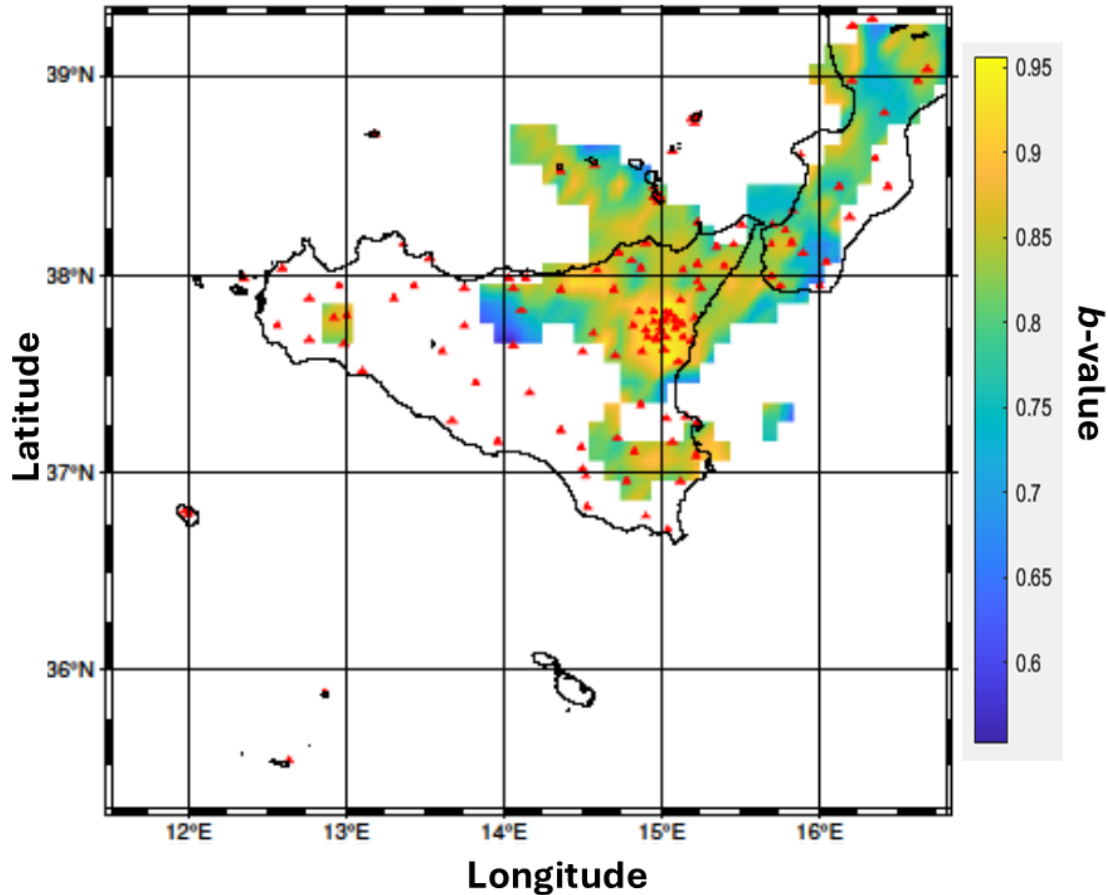
Figure 5.14: Map of *b*-value by selecting a minimum number of events of 100 and a maximum radius of 20 km. The analysis was conducted for the time interval from 2015 to 2024 (Horus).

After the spatial analysis of the *b*-value, the temporal analysis was carried out.

Two geographical areas where there is a greater presence of seismic activity were considered, Etna and Calabria. MLE [29]and *b*-positive [18] techniques were used in this analysis. Entire catalogues were considered by cutting them at a depth of 40 km. Figures n. 5.15 and 5.16, show the time series calculated by the MLE method using a time window of 365 days, a window step of 30 days, a bootstrap of 1000 iterations and a Mc correction factor of 0.4.
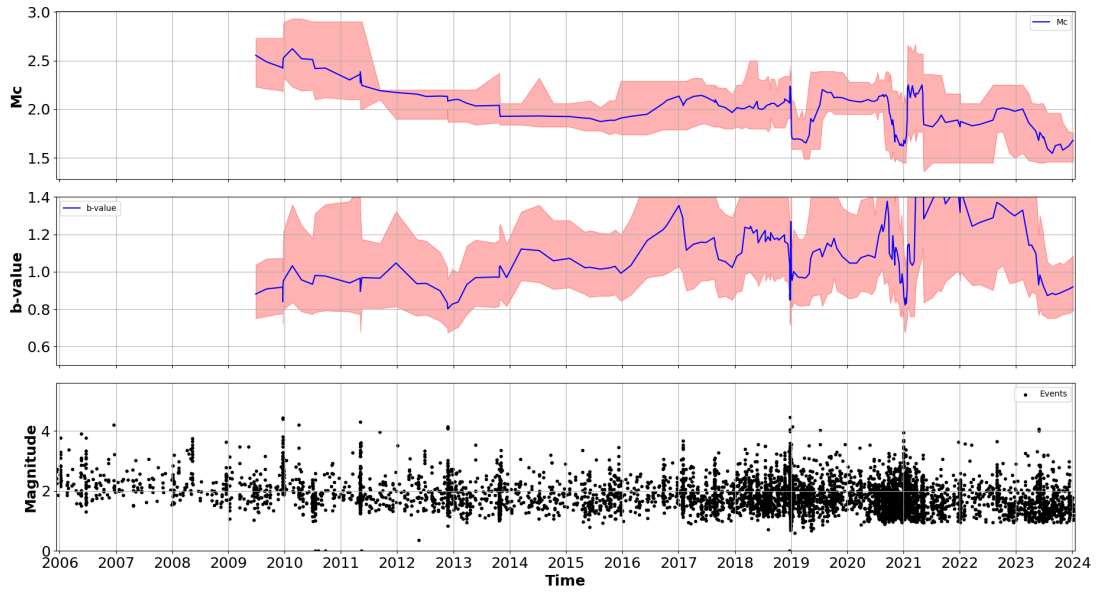
Figure 5.15: Time series calculated using the MLE method on the Etna area. The catalogue is considered for the entire time range (Horus). The first panel represents the magnitude of completeness, the second the *b*-value and the third the seismicity
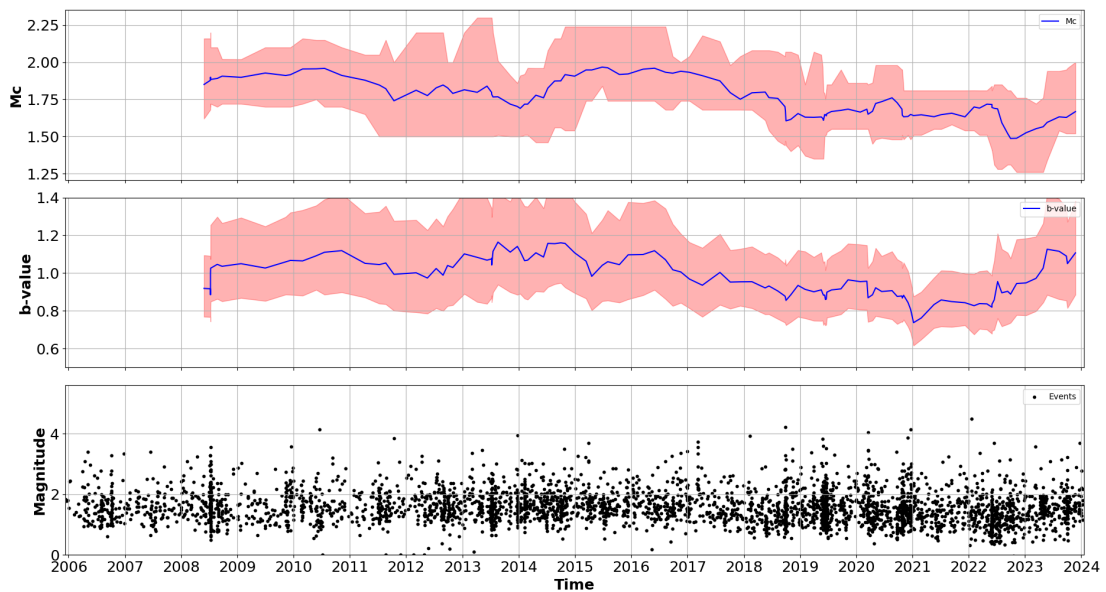
Figure 5.16: Time series calculated using the MLE method on the Calabria area. The catalogue is considered for the entire time range (Horus).

Figures n. 5.17 and 5.18, show the time series calculated by the *b*-positive method using a time window of 365 days, a window step of 30 days, a bootstrap of 1000 iterations and a Mc correction factor of 0.4.
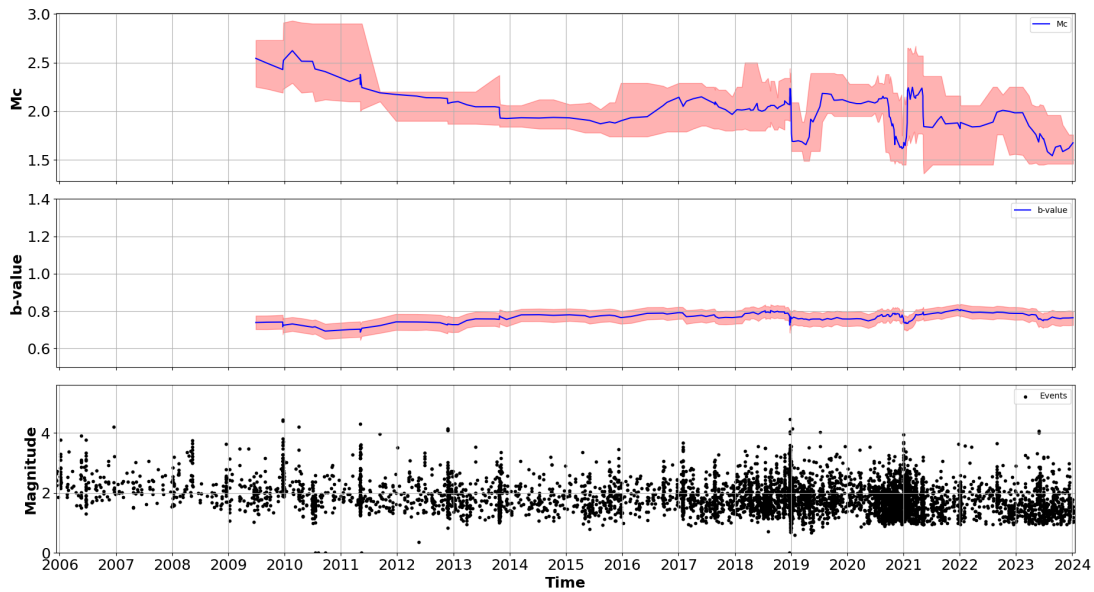
Figure 5.17: Time series calculated using the *b*-positive method on the Etna area. The catalogue is considered for the entire time range (Horus).
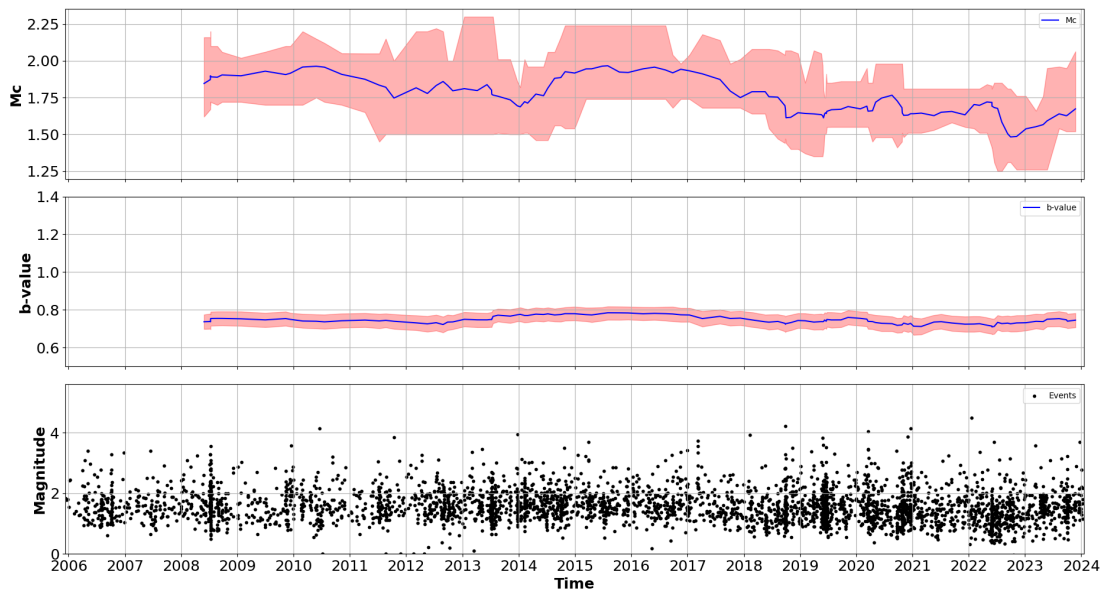
Figure 5.18: Time series calculated using the *b*-positive method on the Calabria area.The catalogue is considered for the entire time range (Horus).

The figures n. 5.19 and 5.20,show the same calculation as the previous ones using the *b*-positive but this time the catalogue has been cut to a magnitude of 1.5 for Etna and 1.0 for Calabria, respectively.
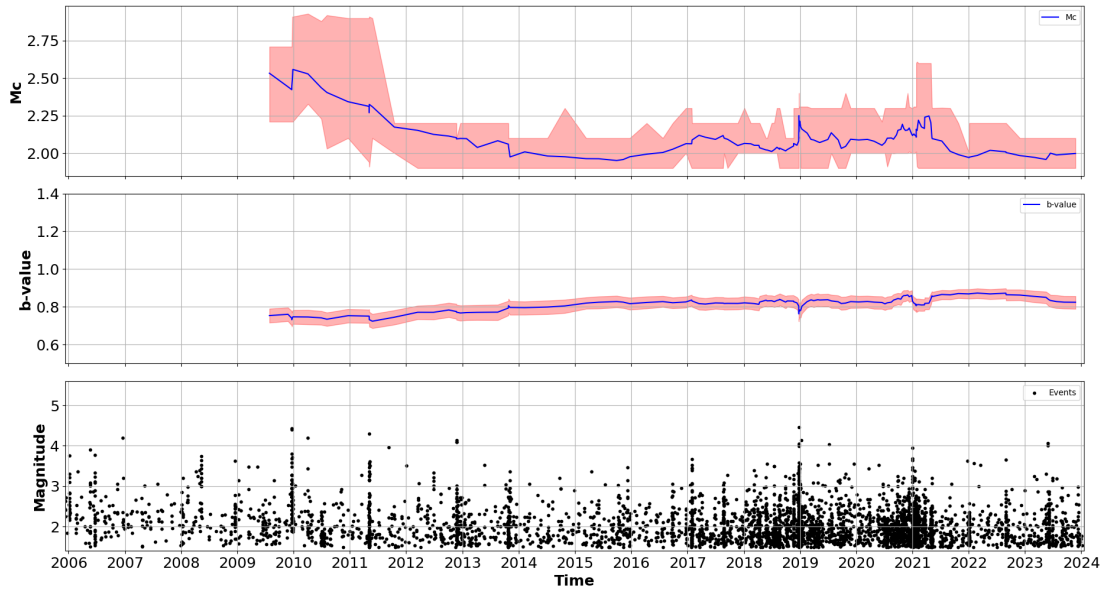
Figure 5.19: Time series calculated using the *b*-positive method on the Etna area, cutting events below 1.5.
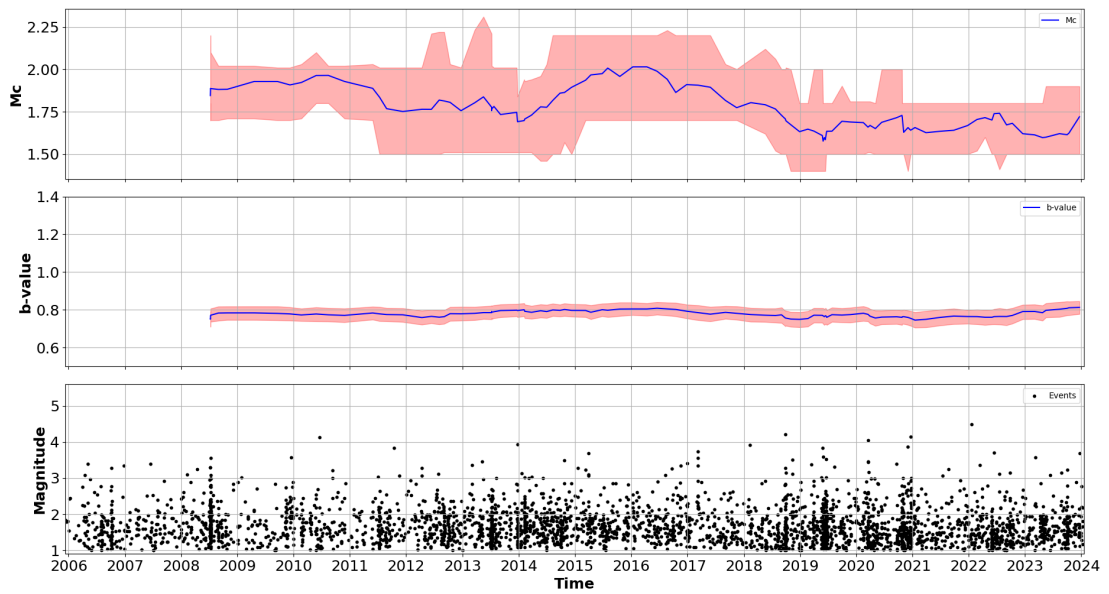
Figure 5.20: Time series calculated using the *b*-positive method on the Calabria area, cutting events below 1.

5.3 Application of the Epidemic Type Aftershock Sequence model

For the application of the Epidemic Type Aftershock Sequence (ETAS) model, the INGV earthquake catalogue was used, specifically focusing on the Sicilian area with a time range from 2005 to 2022. The analyses were conducted using the *etasFLP* R package [12]. This catalogue includes essential information on the event location (latitude, longitude, and depth), magnitude and timing. To ensure optimal application of the ETAS model, additional processing of the INGV catalogue was necessary, not only to obtain estimated values of the main parameters of seismic events, but also to capture the uncertainties associated with these estimates.

These uncertainties are essential for enhancing the robustness of ETAS model predictions, as the accuracy of epicentral location, hypocentral depth and magnitude significantly affects the results.

The integrated uncertainty data included errors in geographical coordinates (latitude and longitude) and uncertainties in focal depth and magnitude.

Other parameters are critical for reliable statistical analysis. Table n. 5.1 below presents all input parameters from the seismic catalogue alongside detailed explanations.

Header	Descrizione
event_id	Event identification number, only serves to uniquely identify an event, as there may be several seismic events with the same hypocentral coordinates.
event_type	Event type.
origin_id	Number identifying the source time, i.e. the instant at which the event started at the source; an ID is used as the same event may be relocated several times or there may be two separate events occurring at the same instant.
version	Code representing the localization version. It can take on various values depending on whether the localization is preliminary or final. A value of 1000 indicates a definitive localization.
ot	Time source of the event, i.e. the instant when the event started at the source.
lon	Longitude of hypocentre in decimal degrees.
lat	Latitude of hypocentre in decimal degrees.
depth	Hypocentral depth expressed in km.
fixed_depth	Indicates whether the depth was estimated from the data or fixed due to insufficient data. 0 = estimated depth, 1 = depth fixed a priori.
err_ot	Uncertainty on the estimate of time of origin expressed in terms of the confidence interval at 69% (1 standard deviation), in seconds.
err_lon	Uncertainty on estimation of longitude expressed in terms of confidence interval to 69% (1 standard deviation), in km.
err_lat	Uncertainty on the latitude estimate expressed in terms of the confidence interval to 69% (1 standard deviation), in km.
err_depth	Uncertainty of the hypocentral depth estimate expressed in terms of the 69% confidence interval (1 standard deviation), in km. The value 'null' is reported if the depth has been fixed.
err_h	Uncertainty of the epicentre estimate expressed in terms of the confidence interval at 69% (1 standard deviation), in km.
err_z	Uncertainty of the hypocentral depth estimate expressed in terms of the 69% confidence interval (1 standard deviation), in km. The value 'null' is reported if the depth has been fixed.
nph_tot	Total number of seismic phases (arrival times of P and S waves) determined.
nph_tot_used	Total number of seismic phases (P and S wave arrival times) actually used in the hypocentral localization process (value less than or equal to nph_tot).
nph_p_used	Total number of P-steps used in the hypocentral localisation process.
nph_s_used	Total number of S-steps used in the hypocentral localisation process.
magnitud_id	Identification of the event by magnitude.
magnitude_type	Type of magnitude, i.e. by which technique it was calculated.
magnitude_value	Magnitude value
magnitude_err	Uncertainty about the magnitude estimate.
magnitude_nsta_used	Number of stations used in magnitude estimation.
pref_magnitud_id	Event identifier for the preferred magnitude.
pref_magnitude_type	Preferred type of magnitude, i.e. by which technique it was calculated.
pref_magnitude_value	Preferred magnitude value.
pref_magnitude_err	Uncertainty about the preferred magnitude estimate.
pref_magnitude_nsta_used	Number of stations used in the estimation of the preferred magnitude.
rms	Mean square deviation, i.e. mean square difference between observed and theoretical arrival times; a measure of the robustness of the hypocentral location.
gap	Expressed in degrees, it indicates the largest angle formed by two stations placing the epicentre at the vertex; it is an inverse measure of the quality of station geometry for good hypocentral location.
source	Web database reference.

Table 5.1: Headers taken into account in the preparation of the seismic catalogue.

This phase of uncertainty collection and verification enhanced the quality of the seismic dataset used, ensuring that all events included in the ETAS modelling contained complete and accurate information.

Following the collection and validation of the seismic catalogue, an analysis of seismicity was conducted to identify regions of particular seismotectonic interest. This categorization aimed to examine how different regions responded to ETAS model application, taking into account their unique geological and seismic characteristic. The identified regions are as follows:

1. **South Tyrrhenian zone:** characterized by significant seismic activity, this region reflects the complex tectonic dynamics of the South Tyrrhenian region.

2. **Belice:** historically known for the destructive 1968 earthquake, the Belice area represents key seismotectonic zone in western Sicily.

3. **Sicily Channel:** located between Sicily and Tunisia, this region is affected by seismic phenomena linked to riftogenesis and volcanic activity.

4. **Iblean Plateaux:** situated in the south-east of Sicily, the Iblei area is a tectonic structure of importance to the island's seismicity, due to the ongoing collision between the African and Eurasian plates.

5. **Madonie Mountains:** located in the northern Sicily, the Madonie Mountains are driven by the compression between the African and European plates, making it as a prominent tectonic uplift zone.

6. **Etna Volcano:** one of the most active volcanic areas, Mount Etna's seismicity is strongly influenced by volcanic processes.

Table n.5.2 shows the results obtained by applying the ETAS model. Some cells are empty. This occurs when the variables taken into account do not come into play in the model. The most influential variables for the model were selected after various tests.

Table 5.2: Estimates for the ETAS model for each cluster considered with the most influential parameters for the model.

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
Magnitude	0.477	0.824	1.01	-0.586	0.616	0.667
Depth						0.015
Longitude error	-0.594	-0.621	-0.038	-1.372		-0.013
Latitude error	-0.557	-0.498	0.152	-1.606		-0.179
H error	0.488	0.653	-0.178	1.807		-0.232
RMS	0.826					
GAP		-0.006				

Figures, from 5.21 to 5.26 show the results of the ETAS model for the six above-mentioned regions. These figures were generated using Surfer software (`https://www.goldensoftware.com/products/surfer/`).
In the top left panels, the regional seismicity can be observed. The top right panels present the same area with the total event density depicted. At the bottom-left panels, the intensity of the background seismicity (aftershock-independent events), is shown, while the bottom-right panels illustrate the triggering as identified by the model.

Figure n.5.21 shows the result of the ETAS model for the South Tyrrhenian region. Seismic intensity appears concentrated in the southeastern area, where two main clusters and at least two smaller clusters can be observed in the triggering activity. This distribution

aligns well with the geodynamics of the region, as these two main clusters correspond with small tectonic structures in the Tyrrhenian Sea. The region encompasses several tectonic features with independent dynamics, governed by a larger E-W-oriented fault structure extending from the Aeolian Islands to a point near the Aegadian Islands [10]. This structure notably generated the 2002 offshore earthquake near Palermo.
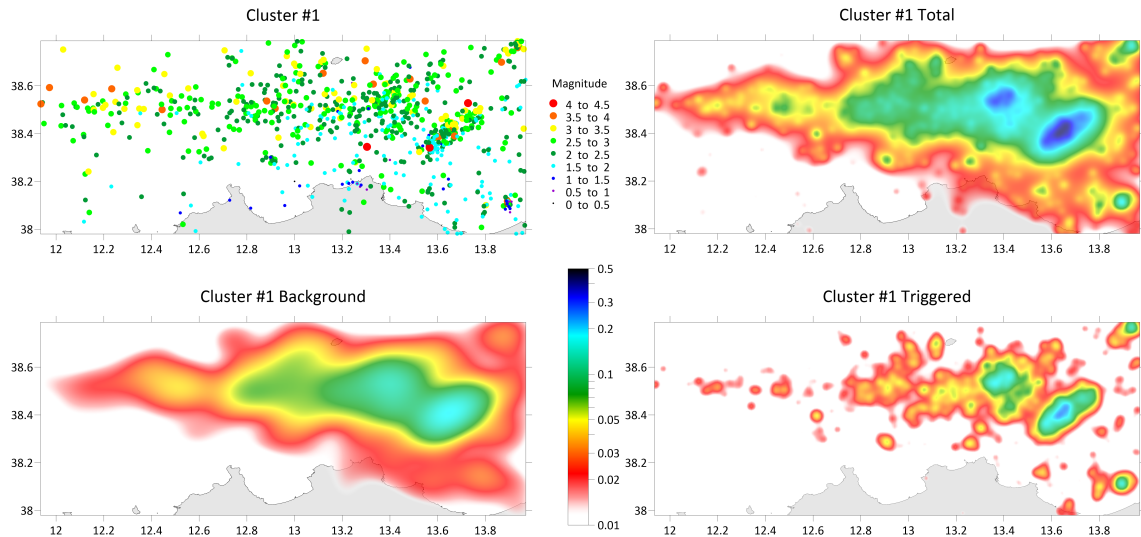
Figure 5.21: Spatial distribution of aftershocks and intensity as modelled by ETAS for the South Tyrrhenian Sea. *Top left:* Observed seismic events, with colours indicating different magnitude classes. *Top right:* Total aftershock density as predicted by the ETAS model, including both background seismic activity and triggered events. *Bottom left:* Background density of seismic events, representing independent seismic activity. *Bottom right:* Density of 'triggered' events (i.e. aftershocks triggered by other events), as estimated by the ETAS model.

Figure n.5.22 shows the result of the Belice ETAS model. The background intensity seems to be higher in the northern area. The triggered intensity reveals three main clusters with indications of a fourth emerging cluster. This region is significant, having experienced a magnitude 6.1 earthquake in 1968. The Belice area lies within the convergence zone of the African and Eurasian plates, where crustal compression has led to the formation of geological structures, including folds and reverse faults. The primary seismogenic structures in the Belice area comprise reverse faults and thrusts. Among these, active faults are often oriented NW-SE or NE-SW, reflecting the direction of maximum horizontal compression [7, 15].
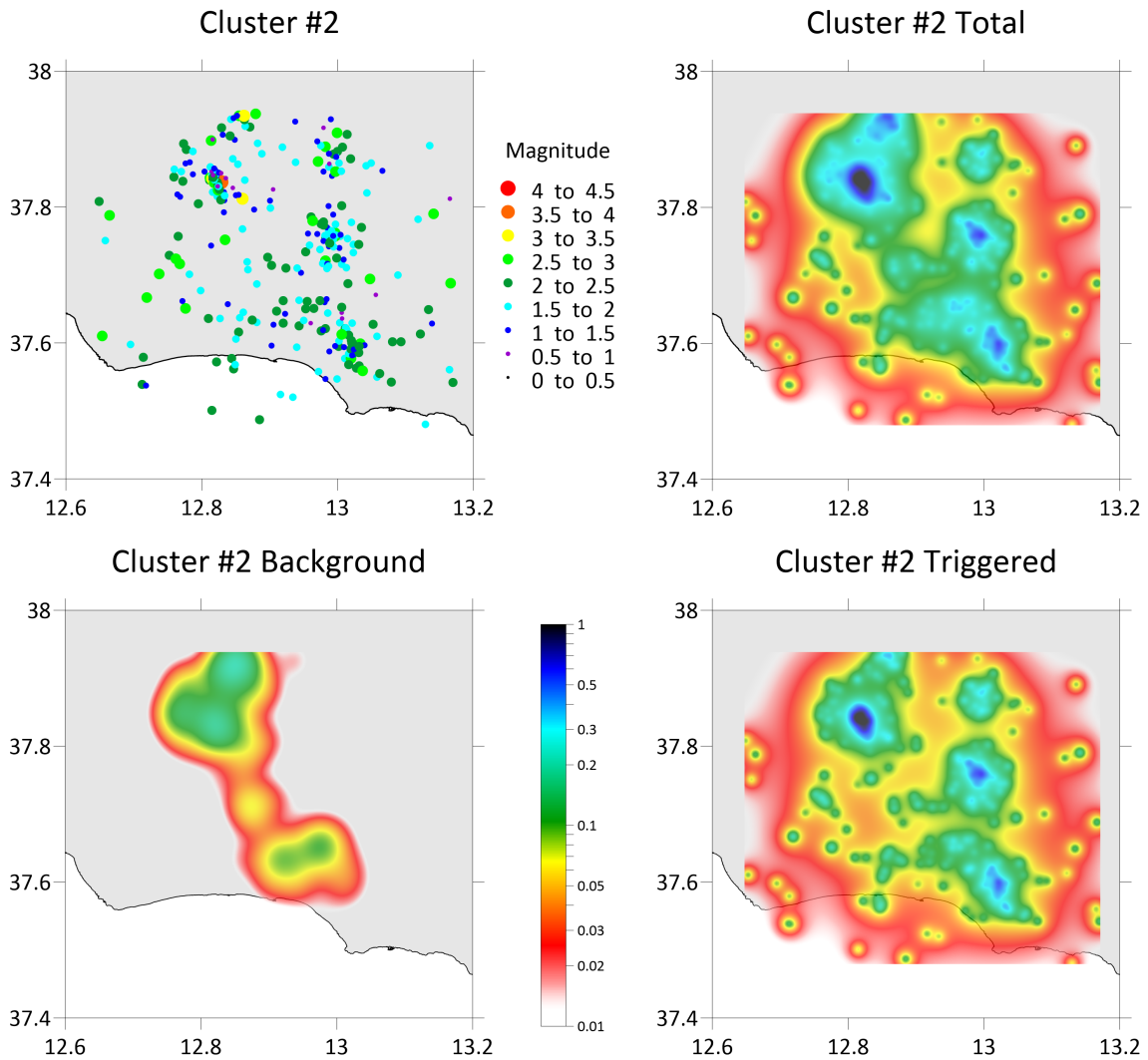
Figure 5.22: Spatial distribution of aftershocks and intensity as modelled by ETAS for Belice. *Top left:* Observed seismic events, with colours indicating different magnitude classes. *Top right:* Total aftershock density as predicted by the ETAS model, including both background seismic activity and triggered events. *Bottom left:* Background density of seismic events, representing independent seismic activity. *Bottom right:* Density of 'triggered' events (i.e. aftershocks triggered by other events), as estimated by the ETAS model.

Figure n.5.23 shows the result of ETAS model for the Sicily Channel. Background seismic intensity appears higher in the south-western portion of the region. Triggered intensity highlights a predominant cluster in the center of the channel along with smaller clusters, some locate near the cost and other offshore. The Strait of Sicily is notable for his extensional tectonic regime, which hosts multiple rifting basins, resulting from back-arc spreading related to the subduction of the African plate beneath the Eurasian plate. Key extensional basins in the Sicilian Channel include the Pantelleria, Malta, and Linosa Basin, where numerous normal faults and graben systems are mainly oriented NW-SE and NE-SW, driving crustal deformation across the region [1].
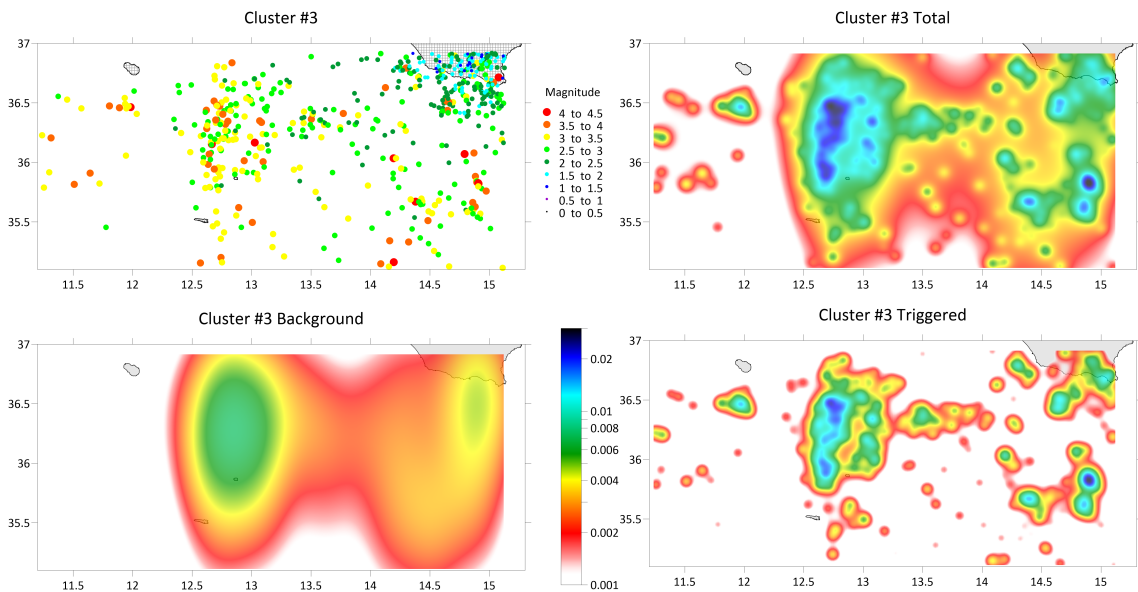
Figure 5.23: Spatial distribution of aftershocks and intensity as modelled by ETAS for Sicily Channel. *Top left:* Observed seismic events, with colours indicating different magnitude classes. *Top right:* Total aftershock density as predicted by the ETAS model, including both background seismic activity and triggered events. *Bottom left:* Background density of seismic events, representing independent seismic activity. *Bottom right:* Density of 'triggered' events (i.e. aftershocks triggered by other events), as estimated by the ETAS model.

Figure n.5.24 shows the result of ETAS model for the Iblean Plateaux. Here, background intensity is higher in the eastern part of the area, with triggered clusters primarily concentrated in the east, through some scattered clusters are observed elsewhere. This region contains significant tectonic structures, notably the Iblean-Maltese escarpment [56]. One of the main tectonic structures is the Scicli Fault, an active extensional fault running parallel to the southeastern Sicily's coast. This fault is associated with extensional tectonic movements and has a significant role in local seismic activity. Additional minor faults and graben systems further influence tectonic deformation and seismicity in this region.
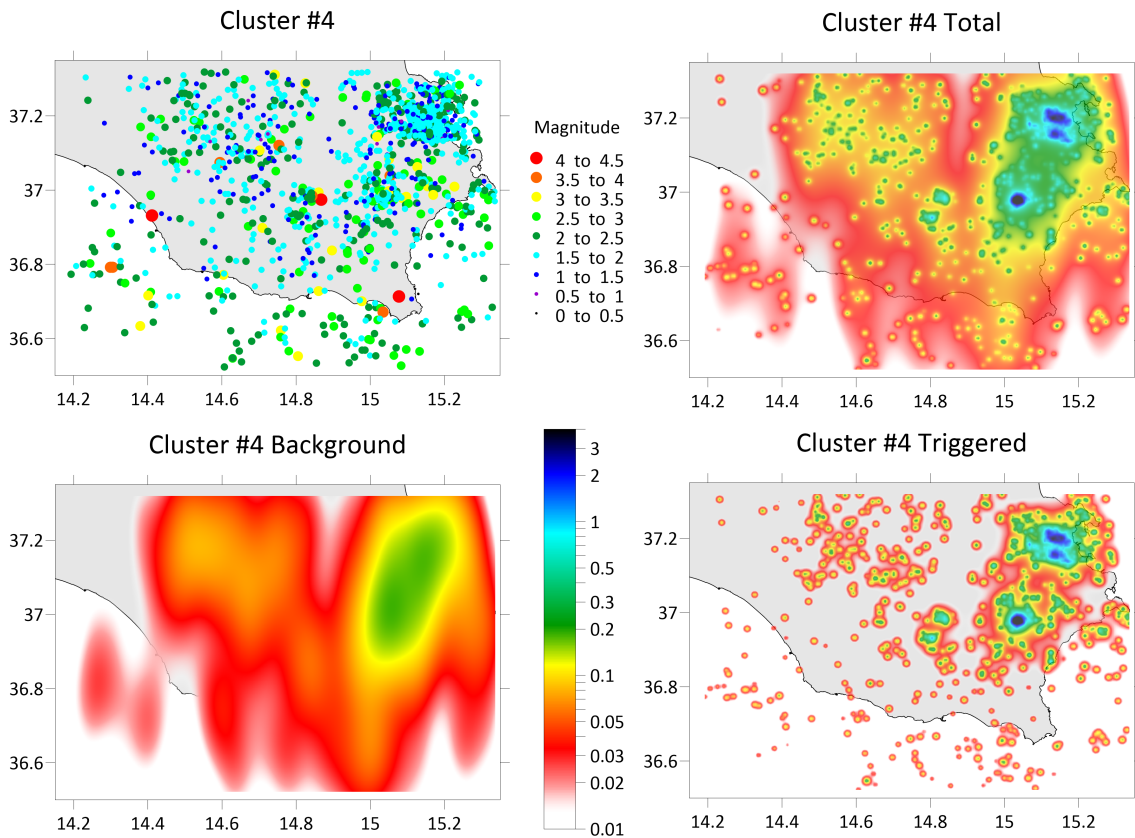
Figure 5.24: Spatial distribution of aftershocks and intensity as modelled by ETAS model for Iblean Plateaux. *Top left:* Observed seismic events, with colours indicating different magnitude classes. *Top right:* Total aftershock density as predicted by the ETAS model, including both background seismic activity and triggered events. *Bottom left:* Background density of seismic events, representing independent seismic activity. *Bottom right:* Density of 'triggered' events (i.e. aftershocks triggered by other events), as estimated by the ETAS model.

Figure n.5.25 shows the results of ETAS model for the Madonie Mountains. Background seismic intensity appears to be extensive throughout the central part of the region. Ttwo main triggered clusters appear in the south-eastern part of the region, with scattered clusters across other areas. The Madonie Mountains. are located in the context of the southern Apennine arc, a mountain range formed by the convergence of the African and Eurasian plates. Primary seismogenic structures include NW-SE oriented reverse faults and thrusts, formed by regional compressive forces. Normal faults are also present. Tectonic dynamics of the Madonie is also influenced by the subduction plate under the Calabrian Arc, contributing to compressional activity in the area [8].
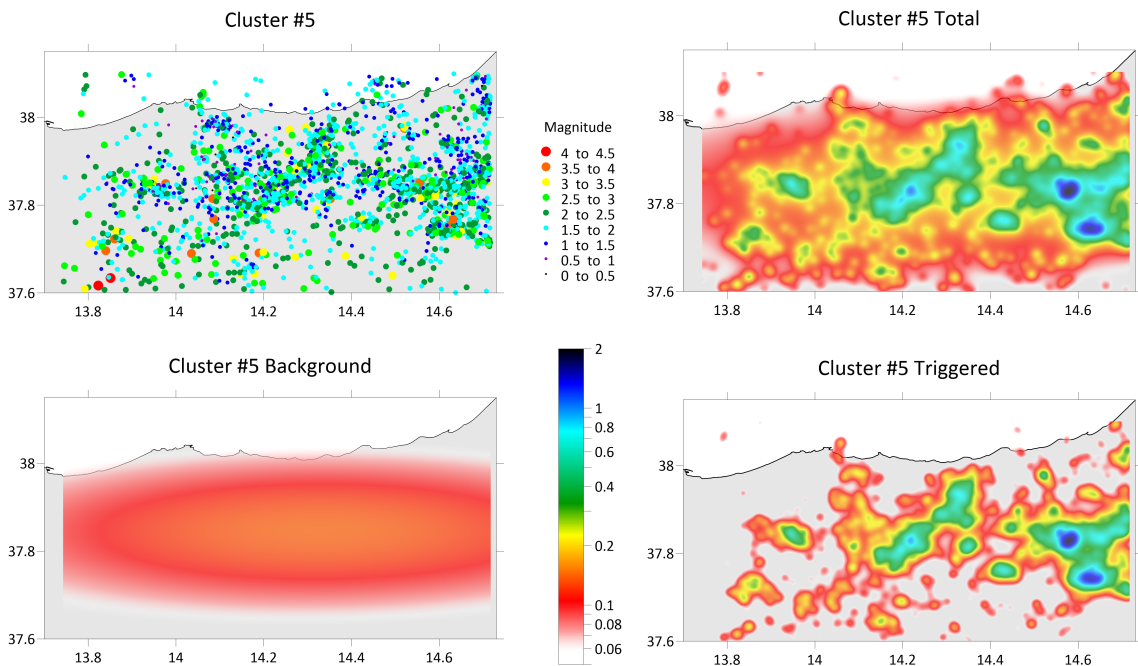
Figure 5.25: Spatial distribution of aftershocks and intensity as modelled by ETAS model for Madonie Mts. *Top left:* Observed seismic events, with colours indicating different magnitude classes.*Top right:* Total aftershock density as predicted by the ETAS model, including both background seismic activity and triggered events. *Bottom left:* Background density of seismic events, representing independent seismic activity. *Bottom right:* Density of 'triggered' events (i.e. aftershocks triggered by other events), as estimated by the ETAS model.

Figure n.5.26 shows the ETAS model results for the Etna volcanic zone. Background intensity is widespread across the central part of the region. Two prominent triggered clusters are in the southeast, and additional scattered clusters are throughout the area. Etna volcano is a volcanically and tectonically. complex zone, with a complex fault system that play a crucial role in the deformation of the volcano and facilitate the magma ascent [6].

Main faults include:

- Pernicana fault: This is one of Etna's most important and active faults, extending approximately 20 km along the northern slope, with a primary dextral strike-slip movement and extensional components. It is associated with notable seismic episodes and surface deformations, oriented E-W.

- Ragalda fault: Situated on Etna's southern slope, this normal fault contributes to the extensional deformation within the volcanic edifice.

Etna also features two primary rifts: the north-east rift, extending from the summit crater towards the north-east and serving as main magma ascent pathway, and the south-west rift,

located on the southwestern slope, representing a zone of significant magmatic extension and ascent.
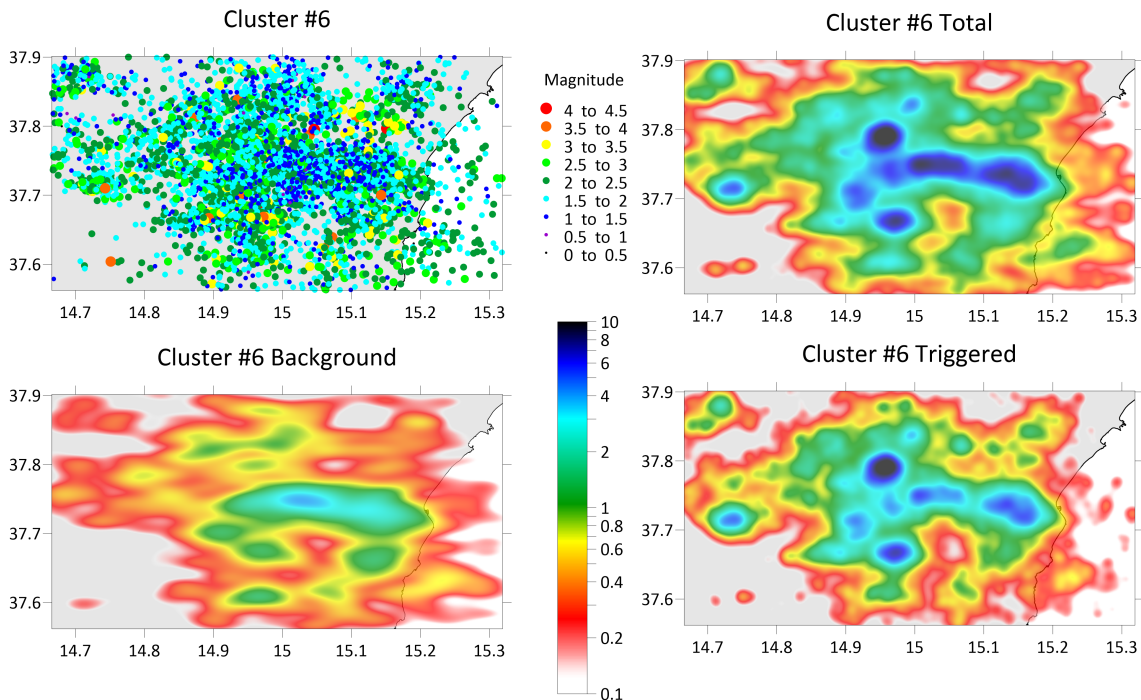
Figure 5.26: Spatial distribution of aftershocks and intensity as modelled by ETAS model for Etna. *Top left:* Map of the distribution of observed seismic events, with colours representing different magnitude classes. *Top right:* Map of total aftershock density predicted by the ETAS model, including both background seismic activity and triggered events. *Bottom left:* Background density of seismic events, representing independent seismic activity. *Bottom right:* Density of 'triggered' events, i.e. aftershocks triggered by other events, as estimated by the ETAS model.

5.4 Application of the Probabilistic detection of completeness magnitude

During my doctoral thesis, I spent a period abroad, including three months of work, in Potsdam at the Helmholtz-Zentrum Potsdam - Deutsches GeoForschungsZentrum (GFZ). I met and collaborated with the author of the previously explained method.

It was chosen to extend the work carried out by Schorlemmer in 2010 [60], applying the method to the case study of the Italian seismic network.

The code was translated entirely, from Python 2 to Python 3. The data preparation and processing phases were then started.

The data utilized in this study was acquired by downloading it from the INGV (Istituto Nazionale di Geofisica e Vulcanologia) website (`https://istituto.ingv.it/it`). The dataset extend from the year 2005 through 2024. Specifically, the data retrieval process involved querying the INGV server to gather comprehensive information, including phase pickings, uncertainties, and arrivals of both P and S waves. The data obtained is in .XML(eXtensible Markup Language) format.This format allows us to read all the information we need and to build the database well.

To build the databases we will use various information deriving from the .xml files, then we build the database in table form. Inside this table, the useful information are:

- the information of stations including the location code

- the on-times of the stations

- event parameters

- the arrivals of P and S waves

- completeness evaluation parameters

- the detection probability distribution per station

We created a file containing all the stations of the Italian seismic network and foreign stations that are also operated by INGV, which are used for earthquake location, figure n. 5.27.
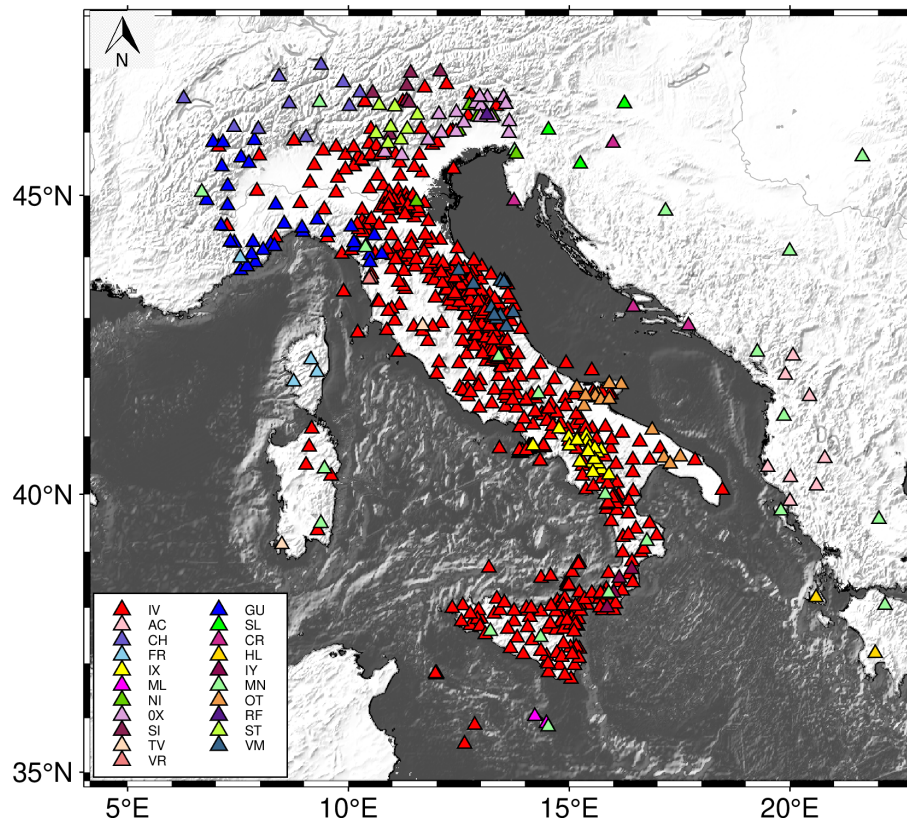
Figure 5.27: Distribution of seismic networks, including national (IV) but also other Italian and foreign networks.This map was created using Generic Mapping Tool6 (GMT)[72].

To have good coverage, additional seismic stations managed in part by INGV for localization, but which belong to foreign seismic networks, were taken into consideration. We have selected 654 stations, of which: 8 belongs to AC network, 10 to CH network, 5 to CR network, 4 to FR network, 2 to HL network, 3 to SL network, 2 to RD network, 2 to OE network, 2 to ML network and 614 to IV network. To better understand the distribution, see figure n. 5.27. For each station, we calculated the operating time for the entire investigated period, as shows figure n. 5.28.
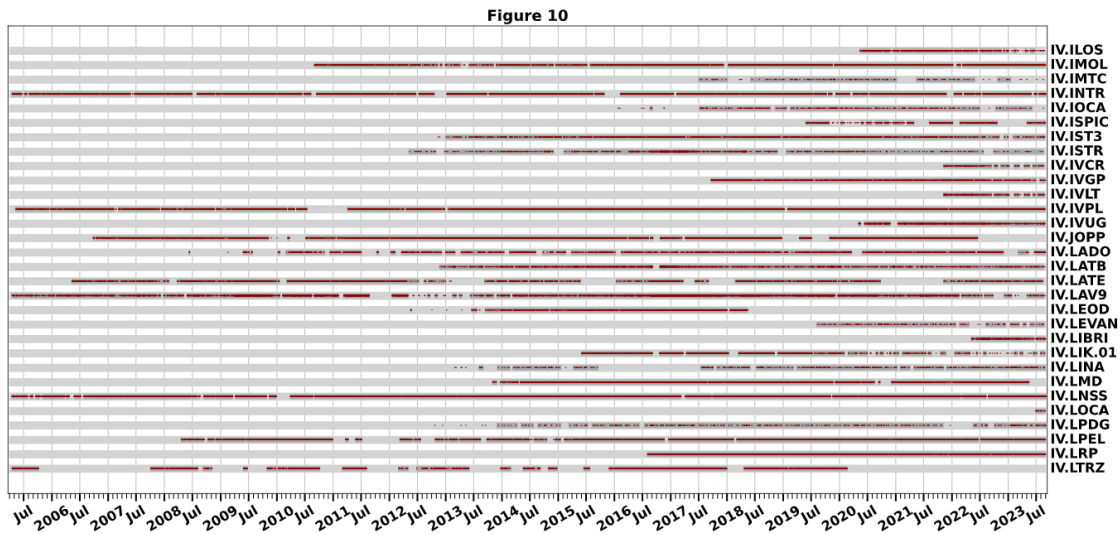
Figure 5.28: On-times calculated for 30 seismic stations showing the station code, station name, station operation. The grey rectangles represent station on and off times, the red dots represent station picks.

Once the entire analysis for all stations has been observed, we go on to develop the probability maps of the magnitude of completeness, then extrapolate the parameter Mp.

What is expected are probability maps of the magnitude of completeness, which vary over time ranges. Considering that, we have more than 20 years of data.

The analyses are still in progress, under review, the candidate hopes to update this part of the thesis and include the maps.

Chapter 6

Discussion and Conclusion

This chapter presents a thorough discussion of the main findings of the research, which explored various aspects of seismicity with a focus on the Gutenberg-Richter law. At the core of this study are the concepts of magnitude of completeness and b-value, both applied to seismic data from Sicily and, more broadly, Italy, through a range of analysis techniques, including both traditional and more recent methodologies. In this chapter, we expand on the data interpretation, particolarly concerning the results for Italy and Sicily.

6.1 Interpretation of the Results

The comprehensive analysis was structured around three main components: constructing and managing the seismological database, applying various data analysis techniques, and utilizing statistical models. Each of these elements contributed to an integrated analysis of the seismic behaviour of the studied region, with a specific focus on Sicily over the 2005-2022 period .

Database construction was a critical phase in ensuring the success of the research. Although the catalogue underwent validation to remove erroneous or duplicate events, some limitations persist, such as the potential underestimation of minor events. Small earthquakes may have been underreported due to the density and effectiveness of the monitoring networks, leading to a possible lack of completeness for lower-magnitude magnitude events in regions with sparse coverage.

Two main aspects were considered when constructing the database and subsequently estimating the magnitude of completeness:

- **Data quality**: Data collected from different catalogues were managed separately, allowing for an assessment of the specific characteristics of each source. This approach made it possible to identify which datasets were more complete and of better quality. Although temporal and spatial coverage was generally satisfactory, the heterogeneity of the sources revealed potential uncertainties in location and magnitude data. Differences in collection methodologies may also have introduced slight distortions in data distribution.

- **Influence of seismic monitoring**: Areas with sparser monitoring networks have lower accuracy in recording minor events, impacting completeness magnitude assessments.

Selecting Sicily as the study area provided an ideal setting for testing the effectiveness of the analysis techniques adopted in this research. The high seismic activity in the region served as a significant test case to evaluate the capabilities and limitations of the methodologies employed, facilitating in understanding how they can contribute to detailed

seismological characterization. The findings offer insights not only on the seismic behaviour of Sicily, but also on the ability of the techniques applied to detect and interpret events of different intensity and frequency. This analysis validated the application of these methods to complex seismicity contests, highlighting the strengths and limitations of the approachs adopted.

The spatial analysis of Mc and *b*-value was conducted on the entire Sicilian seismic catalogue and subsequently replicated on two sub-catalogues. This choice was dictated by the fact of removing the main event that could influence the entire analysis. And also because we wanted to understand how the evolution of the two parameters occurred over time.

A detailed temporal analysis was carried out in two seismically active areas, Etna and Calabria, comparing two methodologies: Maximum Likelihood Estimation (MLE) and b-positive. Results indicate the b-positive method demonstrated greater robustness, being less affected by variations than the MLE method. This suggests that b-positive may be a more reliable approach in tectonically complex areas like Etna and Calabria, where Mc values exhibit significant variability, potentially allowing for a more accurate assessment of the *b*-value even in areas of highly complex seismotectonic regions.

The application of the ETAS (Epidemic Type Aftershock Sequence) model to the Sicilian seismic catalogue enabled an in-depth analysis of local seismicity. helping to distinguish mainshocks and aftershocks. The model revealed significant clustering patterns, highlighting how the distribution of aftershocks and mainshocks affects the spatio-temporal distribution of seismicity. Furthermore, this made it possible to correlate the seismological data with the geodynamics of the study area.

To better understand how a magnitude of completeness and *b*-value analysis can be satisfactory, we compared three different geodynamic areas, namely Sicily, Iceland and Taiwan. All three have the peculiarity of being an island, so they are strongly affected by the limitation of the areal coverage of seismic networks. As far as Taiwan is concerned, the magnitude of completeness was calculated, obtaining a value of 2.2. This analysis was compared with the areal coverage of the seismic network, and areas were found where the network tends to perform lower than in other areas. For Iceland, a study was conducted on the calculation of the *b*-value, correlating it with volcanic eruptions. The phenomenon observed is that the *b*-value tends to vary before significant volcanic eruptions. Thus, we have a possible seismicity that precedes volcanic activity. In this context, *b* plays an important role as it follows seismicity.

6.1.1   Critical Issues of the Analysis

One of the main critical issues concerns the quality of the database. Despite efforts to validate and refine the data, it is possible that some minor seismic events were missed or misclassified, particularly in areas with limited monitoring coverage. This may lead to an overestimation of the completeness magnitude in some region. Additionally, statistical models and declustering techniques present certain limitations, as parameter choices introduce a degree of subjectivity, which may influence the final results. Techniques such as ETAS (Epidemic-Type Aftershock Sequence) or PMC (Probability detection Magnitude Completeness) also exhibit sensitivity to parametric choices or assumptions about the distribution for seismic events.

Another limitations arises from the variability in the density of the seismic monitoring network across regions which impacts the quality and quantity of recorded seismic events. Areas with less dense coverage may underestimate lower-magnitude events, challenging the accurate determination of completeness magnitude.

6.2   Final considerations

The approach taken in this thesis underscores the critical role of data quality and completeness in seismic analyses. The construction of a robust, validated seismic catalogue served as the foundation for examining spatial and temporal variations in Mc and b-value, thereby providing insights into the frequency and intensity of earthquakes across various regions. Limitations related to seismic network coverage and data completeness, particularly in areas with sparse monitoring, have been acknowledged as influencing the accuracy of Mc estimations. This underlines the need for continued expansion and refinement of seismic monitoring networks to achieve higher-resolution assessments of seismic hazards.

Looking ahead, there are significant opportunities for advancing this research. Future work could explore integrating machine learning algorithms into seismic data analysis, which could provide a predictive capability beyond current statistical models. Expanding the application of these techniques to other geodynamically diverse regions would enable a more global understanding of seismicity and potentially uncover universal patterns. Additionally, further refinement of Mc estimation methods and their application to high-resolution catalogues could yield more accurate assessments of seismic hazards, particularly in regions currently under-monitored.

Ultimately, this thesis aims to refine existing seismic methodologies and offer a useful foundation for future research in the field. By enhancing both the analytical tools and understanding of seismicity in diverse geodynamic settings, this work aims to contribute meaningful insights to the ongoing discourse on earthquake prediction and hazard assessment.

Bibliography

1. Agius, M., Magrini, F., Diaferia, G., Kästle, E., Cammarano, F., Faccenna, C., Funiciello, F. & Meijde, M. Shear-Velocity Structure and Dynamics Beneath the Sicily Channel and Surrounding Regions of the Central Mediterranean Inferred From Seismic Surface Waves. *Geochemistry, Geophysics, Geosystems* **23** (**october** 2022).

2. Aki, K. Maximum likelihood estimate of b in the formula $\log N = a - bM$ and its confidence limits. *Bulletin of the Earthquake Research Institute* **43,** 237–239 (1965).

3. Allen, C., Amand, P., Richter, C. & Nordquist, J. Relation between seismicity and geological structure in the southern California region. *Bulletin of the Seismological Society of America* **55,** 752–797 (1965).

4. Amato, A. & Mele, F. Performance of the INGV National Seismic Network from 1997 to 2007. *Annals of geophysics = Annali di geofisica* **51,** 417–431 (**april** 2008).

5. Amorese, D. Amorèse D., 2007, Applying a Change-Point Detection Method on Frequency-Magnitude Distributions, Bull. Seism. Soc. Am. 97, no 5, 1742-1749. *Bulletin of the Seismological Society of America* **97,** 1742–1749 (**october** 2007).

6. Barreca, G., Bonforte, A. & Neri, M. A pilot GIS database of active faults of Mt. Etna (Sicily): A tool for integrated hazard evaluation. *Journal of Volcanology and Geothermal Research* **251,** 170–186 (**august** 2012).

7. Barreca, G., Bruno, V., Cocorullo, C., Cultrera, F., Ferranti, L., Guglielmino, F., Guzzetta, L., Mattia, M. & Pepe, F. Geodetic and geological evidence of active tectonics in south-western Sicily (Italy). *Journal of Geodynamics* **82** (**april** 2014).

8. Barreca, G., Scarfi, L., Cannavò, F. & Koulakov, I. New structural and seismological evidence and interpretation of a lithospheric-scale shear zone at the southern edge of the Ionian subduction system (central-eastern Sicily, Italy): Paleo-STEP. *Tectonics* **35** (**june** 2016).

9. Beskardes, G., Wu, Q., Hole, J., Chapman, M., Davenport, K., Brown, L. & Quiros, D. Aftershock Sequence of the 2011 Virginia Earthquake Derived from the AIDA and Backprojection. *Bulletin of the Seismological Society of America* **109** (**december** 2018).

10. Billi, A., Presti, D., Faccenna, C., Neri, G. & Orecchio, B. Seismotectonics of the Nubia plate compressive margin in the south Tyrrhenian region, Italy: Clues for subduction inception. *Journal of Geophysical Research* **112** (**august** 2007).

11. Cao, A. & Gao, S. Temporal variation of seismic b-values beneath northeastern Japan Island Arc. *Geophysical Research Letters* **29** (**may** 2002).

12. Chiodi, M. & Adelfio, G. Mixed Non-Parametric and Parametric Estimation Techniques in R Package etasFLP for Earthquakes' Description. *Journal of Statistical Software* **76** (**january** 2017).

13. Convertito, V. & Zollo, A. Seismic hazard and b-value changes in Southern Italy based on seismogenic sources. *Journal of Seismology* **15,** 261–276 (2011).

14. DeSalvio, N. D. & Rudolph, M. L. A Retrospective Analysis of b-Value Changes Preceding Strong Earthquakes. *Seismological Research Letters* **93,** 364–375. ISSN: 0895-0695 (**october** 2021).

15. Di Stefano, P., Favara, R., Luzio, D., Renda, P., Cacciatore, M., Calo, M., Napoli, G., Parisi, L., Todaro, S. & Zarcone, G. A regional-scale discontinuity in western Sicily revealed by a multidisciplinary approach: A new piece for understanding the geodynamic puzzle of the southern Mediterranean. *Tectonics* **34** (**september** 2015).

16. Di Stefano, R. & Cattaneo, M. Seismicity of the 2012 Emilia Romagna earthquakes: High-resolution analysis of aftershocks. *Geophysical Journal International* **206,** 798–811 (2016).

17. DISS Working Group. *Database of Individual Seismogenic Sources (DISS), version 3.3.0: A compilation of potential sources for earthquakes larger than M 5.5 in Italy and surrounding areas* 2021. `https://doi.org/10.13127/diss3.3.0`.

18. Elst, N. B-Positive : A Robust Estimator of Aftershock Magnitude Distribution in Transiently Incomplete Catalogs. *Journal of Geophysical Research: Solid Earth* **126** (**february** 2021).

19. Figlioli, A., Vitale, G., Taroni, M. & D'Alessandro, A. Tremors—A Software App for the Analysis of the Completeness Magnitude. *Geosciences* **14.** ISSN: 2076-3263. `https://www.mdpi.com/2076-3263/14/6/149` (2024).

20. Gardner, J. K. & Knopoff, L. Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian? *Bulletin of the Seismological Society of America* **64,** 1363–1367. ISSN: 0037-1106. `https://doi.org/10.1785/BSSA0640051363` (**october** 1974).

21. Godano, C., Convertito, V., Pino, N. A. & Tramelli, A. An Automated Method for Mapping Independent Spatial b Values. *Earth and Space Science* **9,** e02205 (**june** 2022).

22. Gulia, L. & Wiemer, S. The influence of tectonic regimes on the earthquake size distribution: A case study for Italy. *Geophysical Research Letters* **37,** 10305– (**may** 2010).

23. Gulia, L. & Wiemer, S. Real-time discrimination of earthquake foreshocks and aftershocks. *Nature* **574,** 193–199 (**october** 2019).

24. Gulia, L., Wiemer, S., Biondini, E., Enescu, B. & Vannucci, G. Improving the Foreshock Traffic Light Systems for Real-Time Discrimination Between Foreshocks and Aftershocks. *Seismological Research Letters.* ISSN: 0895-0695 (**september** 2024).

25. Gutenberg, B. & Richter, C. F. Frequency of earthquakes in California*. *Bulletin of the Seismological Society of America* **34,** 185–188. ISSN: 0037-1106. eprint: `https://pubs.geoscienceworld.org/ssa/bssa/article-pdf/34/4/185/5300856/bssa0340040185.pdf`. `https://doi.org/10.1785/BSSA0340040185` (**october** 1944).

26. Hatzidimitriou, P., Papadimitriou, D., Mountrakis, M. & Papazachos, B. The seismic parameter b of the frequency-magnitude relation and its association with the geological zones in the area of Greece. *Tectonophysics* **120,** 141–151 (1985).

27. Helmstetter, A. & Sornette, D. Subcritical and Supercritical Regimes in Epidemic Models of Earthquake Aftershocks. *Journal of Geophysical Research: Solid Earth* **107,** 2237 (2002).

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

29. Ishimoto, M. & Iida, K. Observations of earthquakes registered with the microseismograph constructed recently. *Bull. Earthq. Res. Inst.* **17,** 443–478 (1939).

30. ISIDe Working Group. *Italian Seismological Instrumental and Parametric Database (ISIDe) (Version 1)* 2007. `https://doi.org/10.13127/ISIDE`.

31. Kagan, Y. Short-Term Properties of Earthquake Catalogs and Models of Earthquake Source. *Bulletin of the Seismological Society of America California Kern County Joshua Tree–Landers–Big Bear sequence* **94,** 1207–1228 (**september** 2004).

32. Lacidogna, G., Borla, O. & De Marchi, V. Statistical Seismic Analysis by b-Value and Occurrence Time of the Latest Earthquakes in Italy. *Remote Sensing* **15.** ISSN: 2072-4292. `https://www.mdpi.com/2072-4292/15/21/5236` (2023).

33. Lilliefors, H. W. On the Kolmogorov-Smirnov Test for the Exponential Distribution with Mean Unknown. *Journal of the American Statistical Association* **64,** 387–389 (1969).

34. Lisi, A., Mariucci, M. T., Mele, G. R., Castello, B., Montuori, C., Scognamiglio, L. & Lauciani, V. *Bollettino Sismico Italiano (BSI), I quadrimestre 2023 (Version 1) [Data set]* 2024. `https://doi.org/10.13127/BSI/202301`.

35. Locati, M. **andothers**. *Database Macrosismico Italiano (DBMI15), versione 4.0 [Data set]* 2022. `https://doi.org/10.13127/dbmi/dbmi15.4`.

36. Lombardi, A. A Normalized Distance Test for Co-Determining the Completeness Magnitude and b -Value of Earthquake Catalogs. *Journal of Geophysical Research: Solid Earth* **126** (**march** 2021).

37. Luen, B. & Stark, P. B. Poisson tests of declustered catalogues. *Geophysical Journal International* **189,** 691–700 (2012).

38. Marzocchi, W. & Sandri, L. A review and new insights on the estimation of the b-value and its uncertainty. *Annals of Geophysics* **46** (2009).

39. Marzocchi, W. & Sandri, L. Forecasting the effects of earthquake clustering on seismic hazard. *Bulletin of the Seismological Society of America* **99,** 2516–2523 (2009).

40. Michelini, A., Cianetti, S., Gaviano, S., Giunchi, C., Jozinović, D. & Lauciani, V. INSTANCE – the Italian seismic dataset for machine learning. *Earth System Science Data* **13,** 5509–5544. `https://essd.copernicus.org/articles/13/5509/2021/` (2021).

41. Mignan, A., Werner, M., Wiemer, S., Chen, C. & Wu, Y. Bayesian Estimation of the Spatially Varying Completeness Magnitude of Earthquake Catalogs. *AGU Fall Meeting Abstracts* **101** (**december** 2010).

42. Mignan, A. & Woessner, J. Estimating the magnitude of completeness for earthquake catalogs. *Community Online Resource for Statistical Seismicity Analysis* (**april** 2012).

43. Mignan, A. & Woessner, J. Estimating the magnitude of completeness for earthquake catalogs. *Community Online Resource for Statistical Seismicity Analysis* (**april** 2012).

44. Mogi, K. Regional variation in magnitude-frequency relation of earthquake. *Bulletin of the Earthquake Research Institute* **45,** 313–325 (1967).

45. Montaldo, V., Convertito, V. & Zollo, A. Probabilistic seismic hazard analysis in Italy: the role of the b-value. *Geophysical Journal International* **163,** 861–875 (2005).

46. Mousavi, S. Mapping seismic moment and b-value within the continental-collision orogenic-belt region of the Iranian Plateau. *Journal of Geodynamics* **103** (**december** 2016).

47. Murru, M., Rodolfo, C., Falcone, G., Montuori, C. & Sgroi, T. Spatial mapping of the b value at Mount Etna, Italy, using earthquake data recorded from 1999 to 2005. *Journal of Geophysical Research* **112,** 15 (**december** 2007).

48. Nava, F., Márquez-Ramírez, V., Zúñiga, F. & Lomnitz, C. Gutenberg–Richter b-value determination and large-magnitudes sampling. *Natural Hazards* **87** (**may** 2017).

49. Ogata, Y. Statistical Models for Earthquake Occurrences and Residual Analysis for Point Processes. *Journal of the American Statistical Association* **83,** 9–27. `http://dx.doi.org/10.1080/01621459.1988.10478560` (1988).

50. Ogata, Y. Statistical Models for Earthquake Occurrences and Residual Analysis for Point Processes. *Journal of the American Statistical Association* **83,** 9–27 (1988).

51. Ogata, Y. Space-time point-process models for earthquake occurrences. *Annals of the Institute of Statistical Mathematics* **50,** 379–402 (1998).

52. Pasquale, V., Verdoya, M. & Chiozzi, P. Thermal Structure of the Ionian Slab. *Pure and Applied Geophysics* **162,** 967–986 (**may** 2005).

53. Radziminovich, N. A., Miroshnichenko, A. I. & Zuev, F. L. Magnitude of completeness, b-value, and spatial correlation dimension of earthquakes in the South Baikal Basin, Baikal Rift System. *Tectonophysics* **759,** 44–57. ISSN: 0040-1951. `https://www.sciencedirect.com/science/article/pii/S004019511930112X` (2019).

54. Reasenberg, P. Second-order moment of central California seismicity. *Journal of Geophysical Research* **90,** 5479–5495 (1985).

55. Rivière, J., Lv, Z., Johnson, P. & Marone, C. Evolution of b-value during the seismic cycle: Insights from laboratory experiments on simulated faults. *Earth and Planetary Science Letters* **482,** 407–413. ISSN: 0012-821X. `https://www.sciencedirect.com/science/article/pii/S0012821X17306726` (2018).

56. Romagnoli, G., Catalano, S., RIGANO, A., TORRISI, S., Tortorici, G. & Tortorici, L. Tettonica estensionale quaternaria del Plateau Ibleo. *Rendiconti Online Societa Geologica Italiana* **1,** 148–152 (**january** 2008).

57. Rovida, A., Locati, M., Camassi, R., Lolli, B., Gasperini, P. & Antonucci, A. *Catalogo Parametrico dei Terremoti Italiani (CPTI15), versione 4.0* 2022. `https://doi.org/10.13127/cpti/cpti15.4`.

58. Scarfi, L., Graziella, B., Barreca, G., Cannavò, F., Koulakov, I. & Patanè, D. Slab narrowing in the Central Mediterranean: The Calabro-Ionian subduction zone as imaged by high resolution seismic tomography. *Scientific Reports* **8** (**march** 2018).

59. Scholz, C. H. The frequency-magnitude relation of micro-fracturing in rock and its relation to earthquakes. *Bulletin of the Seismological Society of America* **58,** 399–415 (1968).

60. Schorlemmer, D., Mele, F. & Marzocchi, W. A Completeness Analysis of the National Seismic Network of Italy. *Journal of Geophysical Research* (2010).

61. Schorlemmer, D. & Woessner, J. Probability of Detecting an Earthquake. *Bulletin of The Seismological Society of America - BULL SEISMOL SOC AMER* **98,** 2103–2117 (**october** 2008).

62. Scudero, S., D'Alessandro, A. & Figlioli, A. Evaluation of the earthquake monitoring network in Taiwan. *Journal of Seismology* **27,** 643–657 (2023).

63. Siino, M., Scudero, S., Greco, L. & D'Alessando, A. Spatial analysis for an evaluation of monitoring networks: examples from the Italian seismic and accelerometric networks. *Journal of Seismology* **24** (**december** 2020).

64. Taroni, M. Back to the future: Old methods for new estimation and test of the Gutenberg-Richter b-value for catalogues with variable completeness. *Geophysical Journal International* **224.** Cited by: 4, 337–339. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85096478066&doi=10.1093%2fgji%2fggaa464&partnerID=40&md5=a20e8e30ec92929a9ef8f0d06b25069e (2021).

65. Taroni, M. The effect of magnitude uncertainty on the Gutenberg–Richter b-value estimation and the magnitude–frequency distribution: 'what hump?' *Geophysical Journal International* **231,** 907–911 (**june** 2022).

66. Taroni, M. Estimating the Magnitude of Completeness of Earthquake Catalogs Using a Simple Random Variable Transformation. *The Seismic Record* **3,** 194–199. ISSN: 2694-4006. https://doi.org/10.1785/0320230017 (**july** 2023).

67. Taroni, M., Zhuang, J. & Marzocchi, W. High-Definition Mapping of the Gutenberg–Richter b-Value and Its Relevance: A Case Study in Italy. *Seismological Research Letters* **92,** 3778–3784 (**may** 2021).

68. Tinti, S. & Gasperini, P. The estimation of b-value of the frequency–magnitude distribution and of its 1 intervals from binned magnitude data. *Geophysical Journal International* **238,** 433–458. ISSN: 1365-246X (**may** 2024).

69. Tondi, E., Zampieri, D., Giunta, G., Renda, P., Alessandroni, M., Unti, M., Giorgianni, A. & Cello, G. Active faults and inferred seismic sources in the San Vito lo Capo peninsula, northwestern Sicily, Italy. *Geological Society, London, Special Publications* **262,** 365–377 (**september** 2006).

70. Tsapanos, T. b-value of two tectonic parts in the circum-Pacific belt. *Pure and Applied Geophysics (PAGEOPH)* **143,** 229–242 (1990).

71. Weichert, D. H. Estimation of the earthquake recurrence parameters for unequal observation periods for different magnitudes. *Bulletin of the Seismological Society of America* **70,** 1337–1346. ISSN: 0037-1106. eprint: https://pubs.geoscienceworld.org/ssa/bssa/article-pdf/70/4/1337/5328751/bssa0700041337.pdf. https://doi.org/10.1785/BSSA0700041337 (**august** 1980).

72. Wessel, P., Luis, J., Uieda, L., Scharroo, R., Wobbe, F., Smith, W. & Tian, D. The Generic Mapping Tools Version 6. *Geochemistry, Geophysics, Geosystems* **20** (**november** 2019).

73. Wiemer, S. SA software package to analyze seismicity: ZMAP. *Seismological Research Letters* **72,** 373–382 (**may** 2001).

74. Wiemer, S. & Wyss, M. Minimum Magnitude of Completeness in Earthquake Catalogs: Examples from Alaska, the Western United States, and Japan. *Bulletin of the Seismological Society of America* **90,** 859–869. ISSN: 0037-1106 (**august** 2000).

75. Woessner, J. & Wiemer, S. Assessing the quality of earthquake catalogues: Estimating the magnitude of completeness and its uncertainty. *Bulletin of the Seismological Society of America* **95,** 684–698 (2005).

Appendix A

Appendix

A.1 Guastella M., Figlioli A., Martorana R., D'Alessandro A.
Quantifying the Impact of Declustering Techniques on Completeness Magnitude Estimation.
Submitted to *Journal of Seismology*

Quantifying the Impact of Declustering Techniques on Completeness Magnitude Estimation

Massimiliano Guastella[1,2*], Anna Figlioli[2,3], Raffaele Martorana[2], Antonino D'Alessandro[4]

[1*]Dipartimento d'Ingegneria civile ed ambientale, Sapienza Università di Roma, Piazzale Aldo Moro, 5, Roma, 00185, Italy, Italy.
[2]Dipartimento di Scienze della Terra e del Mare, Università degli studi di Palermo, Via Archirafi, Palermo, 90123, Italy, Italy.
[3]Sezione di Milano, Istituto Nazionale di Geofisica e Vulcanologia, Via Alfonso Corti, Milano, 20133, Italy, Italy.
[4]Osservatorio Nazionale dei Terremoti, Istituto Nazionale di Geofisica e Vulcanologia, Via di Vigna Murata, Roma, 00143, Italy, Italy.

*Corresponding author(s). E-mail(s): massimiliano.guastella@uniroma1.it;
Contributing authors: anna.figlioli@unipa.it; raffaele.martorana@unipa.it;
antonino.dalessandro@ingv.it;

Abstract

Seismicity analysis and earthquake hazard assessment require accurate estimation of the magnitude of completeness (Mc), which represents the threshold magnitude below which seismic events are likely to be missed in a catalog. The application of declustering techniques, which aim to remove dependent earthquakes and improve catalog quality, can significantly influence the estimation of Mc. This study focuses on quantifying the impact of application of different declustering methods on Mc estimation. A comprehensive earthquake catalog covering a significant temporal and spatial extent in South Italy is utilized for the analysis. Multiple catalog-based methods for Mc estimation are employed, and various declustering algorithms are applied to assess their influence on Mc estimates. The study highlights the importance of considering declustering effects when estimating Mc. Different declustering techniques yield varying Mc estimates, indicating the sensitivity of the results to the applied method. The findings emphasize the need for a comprehensive approach that incorporates multiple catalog-based methods to validate and assess Mc estimation. This approach ensures robustness and minimizes potential biases in the results. The study contributes to the understanding of the impact of declustering techniques on Mc estimation. By quantifying the influence of declustering on Mc estimates, it provides valuable insights for seismic researchers and practitioners involved in seismicity analysis and earthquake hazard assessment.

Keywords: Completeness Magnitude, Earthquake Catalog, Statistical Analysis, Declustering Algorithm

1

052
053
054
055
056
057
058
059
060
061
062
063
064
065
066
067
068
069
070
071
072
073
074
075
076
077
078
079
080
081
082
083
084
085
086
087
088
089
090
091
092
093
094
095
096
097
098
099
100
101
102

1 Introduction

Earthquake catalogs are important datasets to understand seismicity, earthquake physics and earthquake hazard (Liu et al., 1996; Huang et al., 1994; Xu and Gao, 2014). Catalogs provide valuable information about the occurrence, location, and magnitude of earthquakes over a specific time period. However, it is widely recognized that earthquake catalogs are inherently incomplete, meaning that they do not capture the full extent of seismic activity in a given area. This incompleteness stems from various factors, including instrumental limitations, network coverage, data processing techniques, and the influence of natural phenomena on earthquake detection. Furthermore, the occurrence of large-magnitude earthquakes may lead to a decreased number of detected earthquakes in the immediate aftermath. This can often be attributed to the "blind time" immediately after a large seismic event, during which seismic networks miss smaller successive events due to the saturation of seismic sensors and the general high level of seismic noise (Kagan, 2004; Helmstetter et al., 2006). This blind time produce an increase in the completeness of the catalog due to the loss of data during a seismic sequence. To ensure the accuracy and completeness of earthquake catalogs, it is essential to define the magnitude of completeness (Mc). This threshold magnitude represents the lower limit of earthquake magnitudes that can be reliably detected and recorded by the seismic network. Any earthquake with a magnitude below this threshold may either go unnoticed or be underrepresented in the catalog. Accurate estimation of completeness allows researchers to obtain reliable statistical analysis and interpretation of earthquake patterns and behaviour. Moreover, it provides for evaluating seismic hazard, determining earthquake recurrence rates, and assessing the potential for future seismic events. There are primarily two approaches to estimating the completeness magnitude: the catalog-based methods which rely solely on catalog data, and the waveform-based methods. These latter methods have been exemplified in studies by Sereno and Bratt (1989), and Gomberg (1991). Waveform-based methods that require estimating the signal-to-noise ratio for numerous events at many stations are time-consuming and cannot generally be performed as part of a particular seismicity study Wiemer and Wyss (2000). In the vast majority of instances, methods relying on catalogs ascertain the Mc-value by identifying the threshold beneath which the cumulative Frequency-Magnitude Distribution (FMD) diverges from the Gutenberg-Richter (G-R) law (Gutenberg and Richter, 1944). In this study, we investigate catalog based methods both parametric, which rely on the G-R law validity, and non-parametric, which take no prior

3

assumptions about the distribution. These methods include the the goodness-of-fit test (Wyss et al., 1999; Wiemer and Wyss, 2000), the Mc by b-value stability (Cao and Gao, 2002), the entire magnitude range (Woessner and Wiemer, 2005), and the median-based analysis of the segment slope (Amorese, 2007). Declustering is also performed to investigate the effects of aftershock sequences and analyse the impact on the magnitude of completeness estimation.

Earthquake catalogs are considered incomplete due to various factors and limitations inherent in the data collection process. Understanding the reasons behind catalogue incompleteness is crucial for accurate seismic analysis and estimating the Mc. Instrumental limitations may result in the inability to detect small-magnitude earthquakes, which may be associated with lower energy release and weaker ground shaking. Detection thresholds vary depending on the seismic network and instrumentation used, such as seismometers. These thresholds are determined by factors like noise levels, filtering techniques, and the sensitivity settings of the instruments. Additionally, the spatial distribution of seismic monitoring networks contributes to catalog incompleteness. Uneven network coverage, particularly in remote or inaccessible regions, can lead to gaps in earthquake detection and reporting. Sparse monitoring stations in certain areas result in reduced seismic data availability, leading to incomplete catalogs for those

regions (Schorlemmer et al. (2010), Abd el aal (2012),Mignan and Chouliaras (2014), Gonzalez (2017),D'Alessando et al. (2021)). Data processing techniques also play an important role in catalog completeness. Signal processing algorithms and filtering methods applied to raw seismic data can introduce biases and affect the detection and identification of earthquakes, particularly in case of small-magnitude events. Under these circumstances, the selection of an appropriate processing parameters is crucial to optimize event detection and minimize data loss. Furthermore, natural phenomena can also have an impact on catalogs completeness. High levels of background noise from environmental factors, such as oceanic waves, cultural noise, or anthropogenic activities, can also mask or obscure seismic signals.

In a wide range of magnitudes within seismogenic zones, the distribution of earthquake sizes can typically be effectively represented through a power law relationship (Abercrombie, 1995). The G-R law models the relationship between earthquakes magnitudes and their frequencies in a given area. It states that the logarithm of the number of earthquakes greater than or equal to a given magnitude is inversely proportional to that magnitude. This relationship is quantified by a characteristic parameter known as the b-value, which represents the relative frequency of small to large earthquakes (Ishimoto and Iida, 1939;

Gutenberg and Richter, 1944). Mathematically. it can be expressed as:

$$\log(N) = a - b(M - M_{\min}) \tag{1}$$

where:

- N represents earthquakes with a magnitude greater than or equal to a specific magnitude (M).
- a is the intercept parameter, indicating the level of seismicity at a reference magnitude.
- b is the slope parameter, describing the rate of decrease in earthquake frequency with increasing magnitude.

The a-value, or "b-value intercept" reflects the seismicity rate at a specific magnitude. A higher a-value suggests a higher seismic activity level, while a lower value indicates relatively fewer earthquakes. The b-value, or slope parameter, characterizes the relative distribution of earthquake magnitudes. A high b value indicates a larger proportion of small earthquakes, a lower b-value indicates a larger proportion of large earthquake (Schorlemmer et al. (2005)).

The accuracy and validity of the G-R law depend on correctly determining the magnitude range over which the relationship holds true. This is where the magnitude of completeness becomes crucial. The magnitude of completeness is usually defined as the lowest magnitude at which the earthquakes in a space-time volume are 100% detected (Rydelek and Sacks (1989)). While this definition emphasizes the seismic network's ability to detect earthquakes, establishing a practical threshold founded on the network coverage and sensitivity, in most catalog-method based Mc is inferred from the expected behavior of seismicity by fitting a G-R model to the observed frequency-magnitude distribution (FMD) (Mignan and Woessner, 2012). In the context of G-R law, the magnitude at which the lower end of the observed FMD departs from the G-R model is taken as an estimate of Mc (Zuniga and Wyss, 1995). By determining the Mc, the lower limit of the magnitude range for which the G-R law is applicable is defined. Below this value, events are considered incomplete in the catalog and are not representative of the entire seismicity in the region. Therefore, including these incomplete magnitudes in the analysis can introduce biases and distort in the characteristics of the earthquake FMD. It is also important to take into account the shape of the FMD and its implications in the estimation of Mc.

Understanding the characteristics of the FMD, therefore employing appropriate methods for estimating Mc, are essential in order to obtain reliable and meaningful results. The gradual curvature has been explained by different detection function

205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255

5
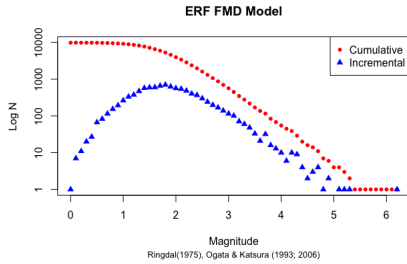
Fig. 1: Synthetic gradually curved FMD shape ($\mu = 2.0$, $\sigma = 0.5$)

formulations, including the cumulative normal distribution, also called or ERror Function (ERF) (Ringdal, 1975; Ogata and Katsura, 1993; 2006). In the context of this model, the incremental FMD is characterized by a gradual curvature below the completeness magnitude. This curvature indicates that smaller earthquakes are less likely to be detected compared to larger ones.

The ERF model uses the parameters μ and σ to describe this gradual curvature. Here, μ represents the mean detection threshold magnitude, indicating the central point of the detection function, while σ represents the standard deviation, indicating the spread or variability of detection probabilities around this mean (figure n. 1).

2 Materials and Methods

The seismic catalog, that was downloaded at the site of Istituto Nazionale di Geofisica e Vulcanologia (https://www.ingv.it/), covers a range of 18 years, from 2005 to 2022. It comprises a comprehensive record of 26.300 earthquakes that occurred within this timeframe. The catalog encompasses a wide range of magnitudes, starting from 0.1 and reaching a maximum magnitude of 5.8. Geographically, it captures seismic activity in the region of Sicily, Italy, with coverage extending from longitude 11.6420 E to 16.4700 E and latitude 35.0020 N to 39.6957 N, as shown in figure n. 2. The study area is located in the mid-west of the Mediterranean Sea, between the African and Euro-Asian plates. Its seismicity is characterized by two types: a superficial one, with depths from 0 to 40 km, and a deeper one, from 40 to 600 km (figure n. 3). A distinction is therefore made between crustal seismicity and seismicity related to the subduction of the Ionian slab. To evaluate the influence of declustering on completeness estimation, the catalog underwent further analysis. Before estimating the completeness magnitude, three additional datasets were derived from the original catalog using the Gardner-Knopoff algorithm (Gardner and Knopoff, 1974). These datasets were generated to explore different windowing approaches: Gardner-Knopoff (Gardner and Knopoff, 1974), Gruenthal (Van Stiphout et al., 2012), and Uhrhammer (Uhrhammer, 1986). Finally, a simulation analysis was conducted utilizing synthetically generated catalogs to delve into the differences between the results obtained from Mc estimates in the actual catalog. Due to the gradually curved shape of the studied FMD, synthetic catalogs were based on the ERF

6

model. Magnitude of completeness computational calculations were performed using the statistical software R based on Mignan and Woessner, 2012 and Amorese, 2007. Additionally, simulation analysis was conducted using, rseismNet R package. The declustering operations, for which the python software were employed, were carried out using OpenQuake Engine package (GEM Foundation). All the work described in the following chapters is summarised graphically in the workflow, represented in fig n. 4.

2.1 Declustering Algorithm: Gardner-Knopoff

In standard Probabilistic Seismic Hazard Analysis, the occurrence of previous earthquakes is modeled using a Poisson process, (Cornell, 1968) where the occurrence of future events is independent from the past ones. Declustering of an earthquake catalog is a process that aims to identify and separate seismic events into their respective mainshocks, foreshocks, and aftershocks. The objective is to remove the influence of aftershocks and other associated seismic activity from the catalog in order to analyze the primary seismicity and accurately estimate various seismic parameters. During a seismic sequence, earthquakes occur in a temporal and spatial pattern. The mainshock represents the largest event in the sequence,

often accompanied by a series of smaller earthquakes that occur before and after the main event. Foreshocks and aftershocks are respectively earthquakes that precede and follow the mainshock. One common approach is based on the time and distance separation between events in the catalog, these methods are known as windowing techniques. Within this category, the most widely used is the Gardner-Knopoff algorithm, (Gardner and Knopoff, 1974) which consists of an iterative process that involves the following steps:

1. The Earthquakes are sorted in descending order according to Magnitude and initially considered as potential mainshocks.

2. The iteration starts and each event is evaluated individually to determine if it meets the criteria to be classified as a mainshock, foreshock, or aftershock.

3. The first event corresponds to the maximum magnitude value in the catalog, meaning for sure that it is a mainshock.

4. When a mainshock is detected, the time-space window is applied both forward and backward. At this point, an assessment is made to determine if other events fall within the spatial temporal distance defined by the window. If an earthquake satisfies the distance criterion and falls within the time window of a larger event, it is labeled as an aftershock or foreshock.

358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
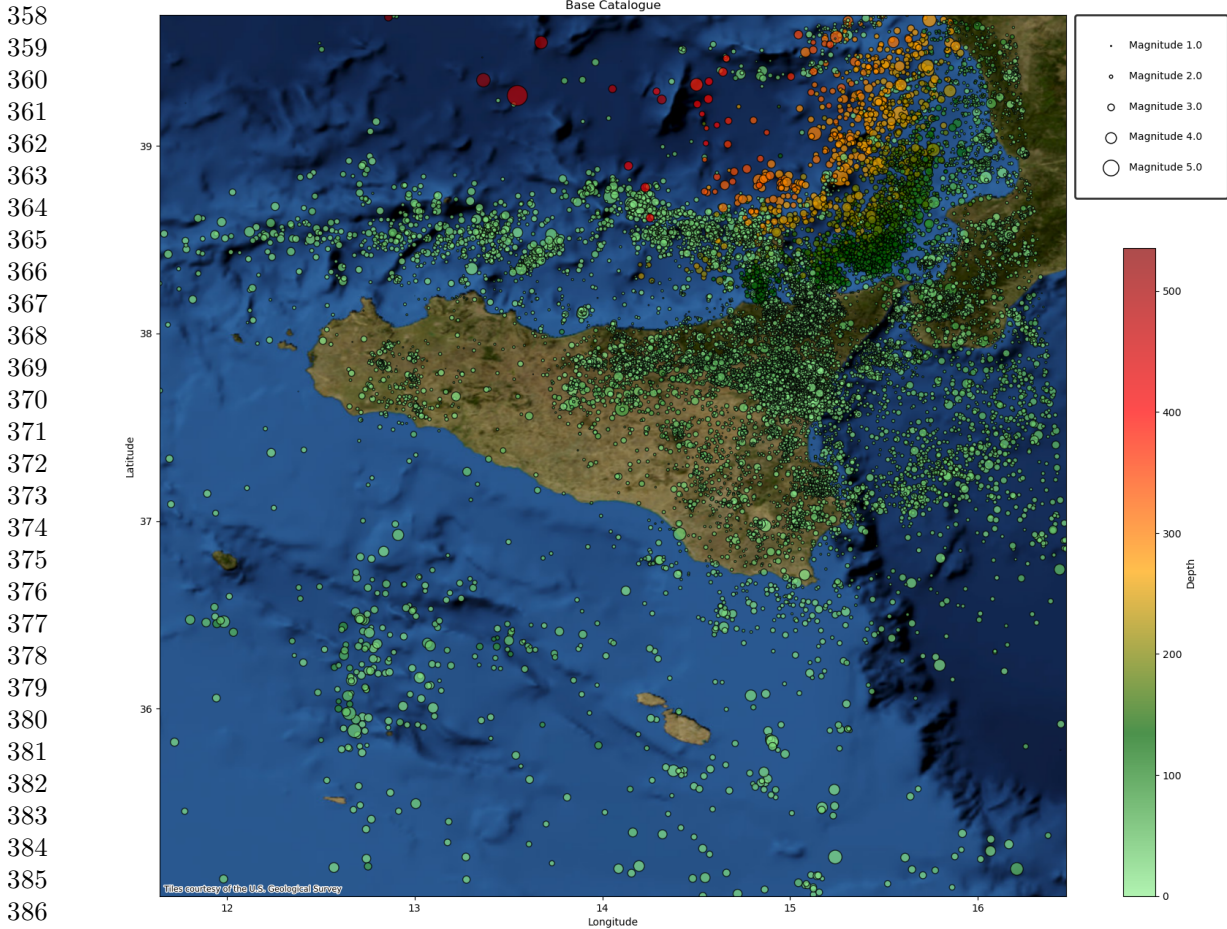
Fig. 2: Spatial distribution of the analysed seismic catalogue. The dots are color-coded by depth, with their size scaled on the magnitude

5. The process resumes from the last labeled mainshock event, repeating the procedure until all events in the series have been classified.

The declustering of a seismic catalog typically produces two datasets: a copy of the original catalog with each event classified based on the employed method, and a sub-catalog containing only the events considered as independent. Declustering algorithms act by filtering events as well as the completeness magnitude, although on a different basis. The balance between necessities of the study and sample size should be always pursued. The parameters of the three different space-time windows applied in this study are illustrated in table n. 1 (Gardner and Knopoff, 1974; Uhrhammer, 1986;Van Stiphout et al., 2012). The choice to employ these three different space-time windows, specifically Gardner-Knopoff, Gruenthal, and Uhrhammer window, was conducted by the aim to explore the variability of declustering

8
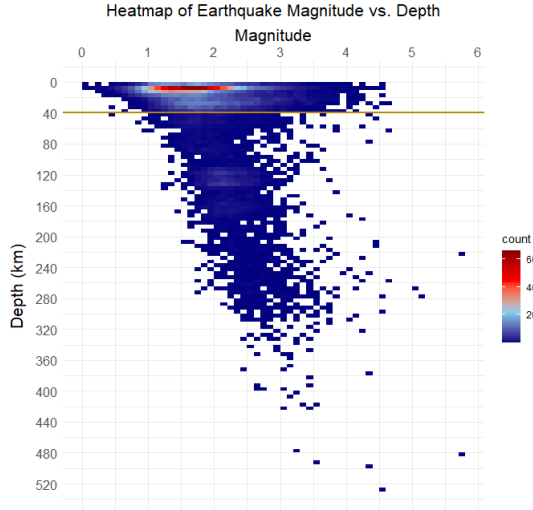
(a)
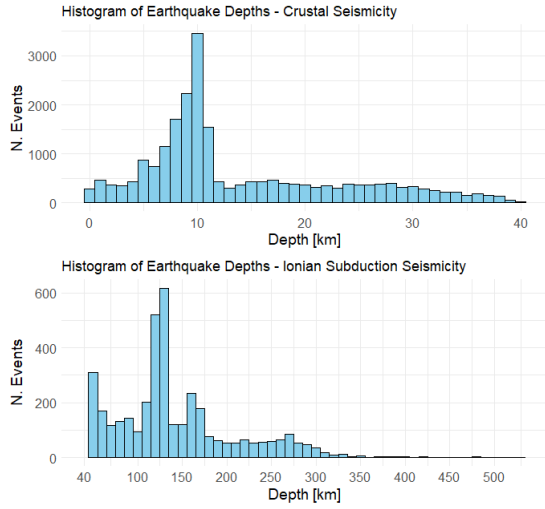
(b)

Fig. 3: (a) Heatmap of Earthquake Magnitude vs. Depth: illustrates the relationship between earthquake magnitude and depth.The orange line represents the delineation between superficial and deep seismicity. (b) Histograms of the Distribution of Depths: displays the frequency distribution of earthquakes depth for both classes.

results to different parameter settings on a Sicilian earthquake catalog. This comparison has been applied in other study areas, such as Juellyan et al. (2023) Kim (2022).
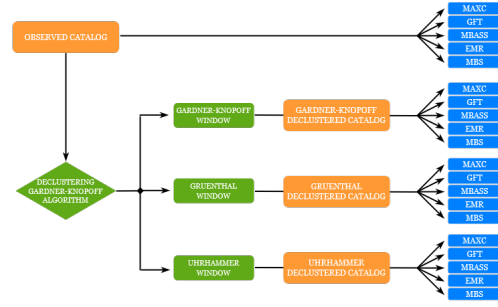
Fig. 4: Workflow overview of applied methodological pipeline.

In figure n.5 is depicted a heatmap showcasing the distribution of earthquakes frequencies, classified in 0.5 year interval and 0.1 magnitude increments before applying any declustering process, whereas the results of pruning the origin catalog are shown in figure n.6 and, more significantly, in figure n.7. The comparison of incremental FMDs visually outlines at which magnitude values the most substantial event reductions occurred. Due to the fact that maximum magnitude observed in the original catalog does not reach the magnitude threshold set by the authors, the Gardner-Knopoff and Gruenthal time window definitions for $M \geq 6.5$ are not reported as not needed.

Metric	G-K	GRU	UHR
Dist.	$10^{0.12M+0.98}$	$e^{(1.8+1.06M)^2}$	$e^{-1.02+0.8M}$
Time	$10^{0.54M-0.55}$	$10^{2.8+0.02M}$	$e^{-2.87+1.24M}$
Mainshock	9498	4226	18847
Foreshocks	5635	9046	875
Aftershocks	11167	13028	6578

Table 1: Space-time parameters definition for each windowing approach employed and the relate number of events obtained after pruning the catalog.

409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459

460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
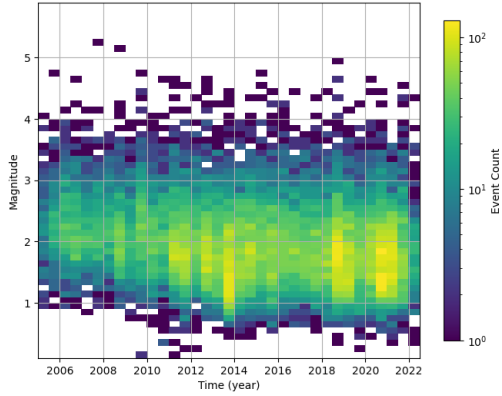
Fig. 5: Seismic event density over time: Earthquake occurrences grouped by 0.5-Year Intervals and 0.1 magnitude increments
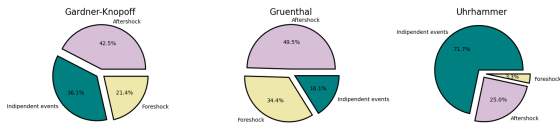
Fig. 6: Number of events obtained after pruning the original catalog for each time window applied

2.2 Catalog-based methods to estimate the completeness magnitude

The MaXimum Curvature (MAXC) technique (Wyss et al., 1999; Wiemer and Wyss, 2000) is a catalog-based method used to estimate the magnitude of completeness in earthquake catalogs. It operates by calculating the maximum value of the first derivative of the frequency-magnitude curve. In practical terms, this method identifies the magnitude that corresponds to the mode of the non-cumulative frequency-magnitude distribution. By identifying the magnitude associated with the highest frequency of events in the FMD, it
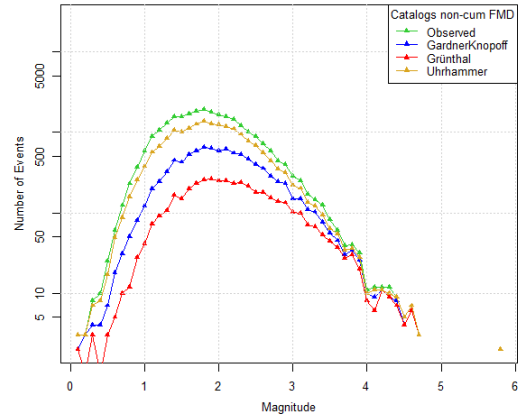
Fig. 7: Incremental FMD comparison between the original and the three declustered catalogs extracted

aims to capture a significant point of deviation from the linear cumulative part of the Gutenberg-Richter law. The Goodness-of-fit (GFT) method, developed by Wiemer and Wyss (2000), is a parametric method used to assess quantitatively the differences between the observed and synthetic distributions, providing a measure of how well the G-R law fits the input data. The procedure begins by selecting a magnitude cut-off range, which represents the potential magnitude of completeness values. The G-R law is then applied to calculate the theoretical distribution of earthquake magnitudes using maximum likelihood estimation of the parameters a and b from the observed data. The next step involves comparing the observed cumulative number of events (Bi) in each Mco with the predicted cumulative number of events (Si) from
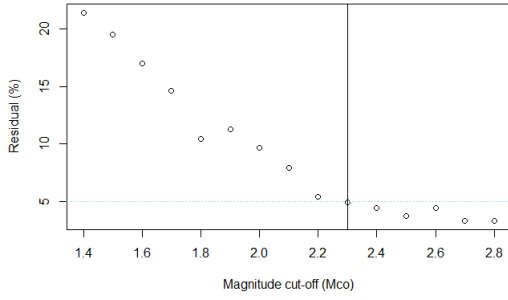
Fig. 8: GFT method: Residual are plot as a function of magnitude cutoff interval. The completeness is identified (vertical black line) as the first value that achieve the confidence level (in this study R = 95 % is achieved).

the theoretical distribution. The absolute difference R between Bi and Si is calculated for each magnitude cut-off using the equation n. 2.

$$R(a, b, M_{\mathrm{co}}) = 100 - \left(\frac{\sum_{M_{\mathrm{co}}}^{M_{\mathrm{max}}} |B_i - S_i|}{\sum_i B_i} 100 \right) \tag{2}$$

By evaluating R at different magnitude thresholds, the test identifies the Mc value that satisfies the desired confidence level. If the observed and theoretical distributions exhibit large differences, indicating a poor fit, a higher Mc value will be selected. Conversely, if the observed and theoretical distributions align well, the Mc value will be lower.

The completeness magnitude is defined as the first magnitude cut-off at which the residual falls below the horizontal line corresponding to the

desired confidence level (e.g.. 95% fit) as presented in figure n. 8 . If the desired confidence level cannot be achieved, 90% may be used as a compromise. If even the 90 confidence level cannot be reached, the MAXC estimate is used instead.

Amorese (2007) introduced the median based analysis of the segment slope (MBASS) method as a non-parametric technique for estimating the magnitude of completeness in earthquake catalogs. It is based on the change-point detection principles to identify multiple discontinuities in the non-cumulative frequency-magnitude distribution, with the primary discontinuity corresponding to completeness. The MBASS method builds upon the multiple change-point procedure developed by Lanzante (1996), which has demonstrated success in climate data analysis by employing resistant, solid, and non-parametric approach. In the context of earthquake catalog, the MBASS algorithm follows an iterative process to estimate Mc. Initially, the frequency magnitude distribution is divided into segments, defined by magnitude cut-offs. For each segment, the slope is calculated and then iteratively tested for change points, which represents potential Mc values. The null hypothesis states that the frequencies are equal. The acceptance or rejection of the null hypothesis of no change at the identified change-point is based on the Wilcoxon-Mann-Whitney test (Wilcoxon, 1945), a widely used non-parametric statistical test. The significance level for evaluating the

511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561

11

statistical significance of each change-point is typically chosen in advance. In this study a significance level of 5 has been chosen. The change-point test is iteratively applied to the magnitude series as long as the statistical significance of each new change-point remains below the specified significance level. Each iteration generates a list of N+1 change-points and prior to applying the change-point test in subsequent iterations, the magnitude series is adjusted by subtracting the median of each segment slope from the corresponding points. The final Mc estimate is determined based on the identified change points, typically using the magnitude associated with the best test statistic as the Mc value which corresponds to the the main discontinuity. In figure n.9 is depicted the identification of the main discontinuity obtained through 3000 bootstrap sample for the origin catalog. Other discontinuities may correspond to upper-magnitude breakpoints (Wesnousky, 1994). The MBASS procedure cannot be applied to cumulative FMD data because, in common with many other statistical procedures, the Wilcoxon-Mann-Whitney test requires independence within groups (Hollander and Wolfe, 1973).

The Entire Magnitude Range (EMR) method (Woessner and Wiemer, 2005), offers a parametric approach to estimating the completeness magnitude by employing a two-part model that combines the Gutenberg-Richter law and a cumulative normal distribution function. This method
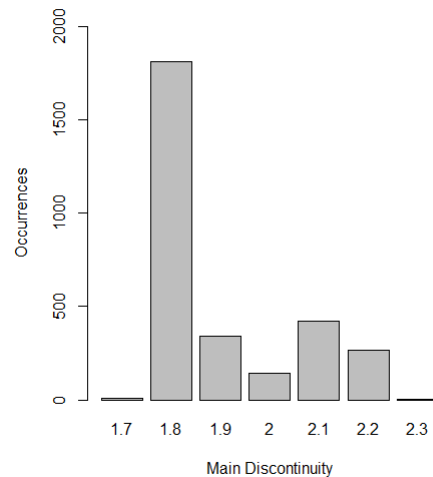
Fig. 9: MBASS method – Occurrences of the Mco identified as the main discontinuity obtained for the origin catalog through 3000 bootstrap samples

accounts for both the complete and incomplete parts of the magnitude frequency distribution. In the EMR model, the first component of the model incorporates the G-R law and represents the part of the earthquake catalog where earthquake detection is assumed to be reliable and complete, capturing the entirety of seismic activity above the completeness. The second component of the model incorporates a cumulative normal distribution function, denoted as q, which characterizes the detection capability of the seismic network. This component addresses the incomplete part in the catalog, modelling the incremental frequency magnitude below Mc, acknowledging that some events might not be detected. Figure n.10 shows the reconstruction of the FMD model performed by the EMR method.

The choice of the normal cumulative distribution function is based on visual inspection and modeling of a variety of catalogs, as well as comparisons to other possible functions, but is not based on physical reasoning. Thus, cases exist for which the choice of another function might be more appropriate. However, synthetic tests endorse that estimates of Mc can be correct even if this assumption is violated (Woessner and Wiemer, 2005).

The probability of detecting an earthquake with magnitude M, given a specific detectability level determined by the parameters μ and σ, can be quantified by the following function:

$$q(M|\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{M_c} \exp\left(-\frac{(M-\mu)^2}{2\sigma^2}\right) dM$$
$$\text{for } M < M_c \tag{3}$$

$$q(M|\mu,\sigma) = 1 \quad for \quad M \geq M_c \tag{4}$$

The parameter μ represents the magnitude at which 50% of the earthquakes are detected. The parameter σ represents the standard deviation, reflecting the width of the magnitude range over which partial detection occurs. A larger σ corresponds to a wider range of magnitudes over which the detection capability gradually changes.
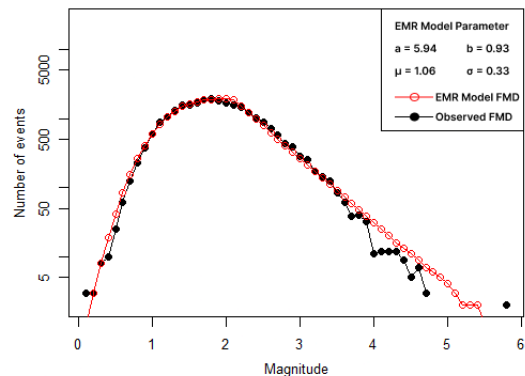
Fig. 10: Observed and predicted incremental FMD. The parameters of the EMR model were obtained through maximum-likelihood ($b = 0.93$, $\mu = 1.06$, $\sigma = 0.33$)

By explicitly incorporating Mc into the equation n.3 and n.4, as proposed by the authors, the EMR method enables completeness estimation through the minimization of maximum likelihood function (figure n.11). This statistical technique aims to identify the values of the parameters (μ and σ) that maximize the likelihood of observing the seismicity data.

The magnitude of completeness by b-value stability (MBS), initially proposed by Cao and Gao (2002) and further developed by Woessner and Wiemer (2005), offers an approach to estimate the completeness magnitude based on the stability of the b-value with respect to the cut-off magnitude. The MBS method operates under the assumption that the b-value estimates exhibit an increasing trend for Mco values lower than Mc and remain

13
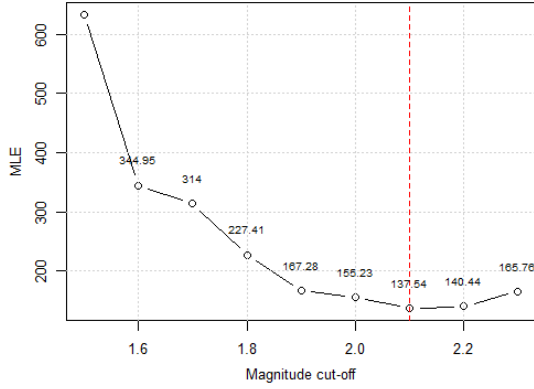
Fig. 11: EMR method – Maximum likelihood estimate of M*c*

constant for M*co* values greater than or equal to M*c*. Deviations from this trend indicate inaccurate *b*-value estimates when M*co* is below M*c*. As M*co* approaches M*c*, the *b*-value converges to its true value and plateaus, reflecting the stable seismicity characteristics. In the original formulation, M*c* was arbitrarily defined as the magnitude at which the change in *b*-value between successive magnitude bin is less than 0.03. However, Woessner and Wiemer identified the instability of this criterion due to variations in the frequency of events within individual magnitude bins. To address this instability and introduce an objective measure, they proposed employing the *b*-value uncertainty (δb) as the criterion. The *b*-value uncertainty is computed as depicted in equation n. 5.

$$\delta b = 2.3b\sqrt{\frac{\sum_{i=1}^{N}(M_i - \overline{M})^2}{N(N-1)}} \qquad (5)$$

As displayed in equation n. 5 *b* represents the estimated *b*-value, $\overline{M}$ is the mean magnitude, N denotes the number of events, and Mi represents individual cut-off magnitudes. This formulation stabilizes the MBS method numerically. According to the MBS method, M*c* is defined as the first magnitude increment at which the absolute difference between the average *b*-value (*bave*) and the individual *b*-value (*b*) is less than or equal to the *b*-value uncertainty (δb). The average *b*-value (*bave*) is calculated by averaging the *b*-values of successive cutoff magnitudes within a half magnitude range (*dM* = 0.5) using a bin size (Δm) of 0.1 (figure n.12). The MBS method aims to identify M*c* by monitoring the changes in the *b*-value as M*co* increases. Once completeness is reached, the *b*-value exhibits a decline for smaller magnitudes, indicating a shift in earthquake occurrence patterns. Compared to other techniques, the MBS method tends to yield higher M*c* values. as evidenced in the findings presented by Woessner and Wiemer (2005) along with the greater uncertainty.

14
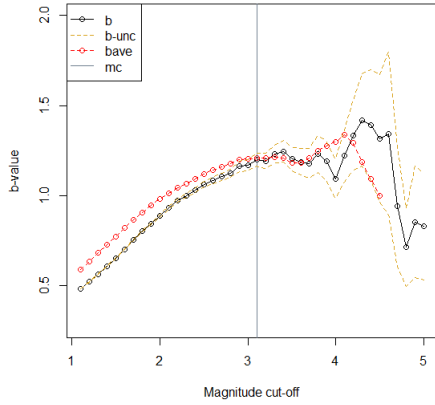
Fig. 12: MBS method – b–value estimate as a function of minimum magnitude cutoff Mco with the relate uncertainty. Mc (vertical line) is the first magnitude increment at which the condition $\Delta b \leq \delta b$ is met.

2.3 Evaluating Bias in Mc Estimation Methods for Curved Datasets

To further investigate the bias of magnitude of completeness estimation methods in case of curved datasets, synthetic catalogs were generated on the ERF model assumptions (Ringdal, 1975; Ogata and Katsura, 1993;2006) for different parameter sets and the Mc estimation was computed. The parameters were varied systematically to span a broad range of potential real-word scenarios, covering integer μ values from 1 to 3, while σ values ranging from 0.1 to 3.0 with the increments being in steps of 0.1. To ensure the simulations accurately reflected the logarithmic relationship between earthquake magnitude and frequency as

mandated by the Gutenberg-Richter law, the b-value was set to 1. The choice of four possibilities for the number of events parameter (n.eq 5000, 10000, 20000, 25000) was guided by the characteristics of both the original and the declustered catalog as it is depicted in figures n.17, 18, 19 and 20. The comparative analysis between synthetic datasets and real earthquake catalogs concerning the estimation of Mc could give assistance in validating outcomes and enabling the assessment of the accuracy of various techniques. This direct comparison is instrumental in helping determining which methods may result more reliable and precise in estimating completeness for a specific case study. Moreover it is crucial in identifying any biases or limitations inherent in the estimation methods. Through such analysis, it becomes apparent whether certain techniques are better suited for particular types of seismic activities or specific regions, and how the complexities of real-world data impact Mc accuracy Mignan, 2012. The simulation test was primarily focused on MAXC, GFT and MBASS. The exclusion of the EMR method from this analysis stems from its fundamental assumption of a normal distribution for the incomplete segment of the dataset, which directly corresponds to the simulated parameters (μ, σ). On the contrary a comparison between EMR model and a synthetic catalog built on similar μ and σ may provide insight on the performance of the method.

15

Additionally, the MBS method was not considered owing to to its notable lack of stability. It has been observed that the MBASS and GFT methods might underestimate or overestimate Mc, particularly on certain magnitude samples. To mitigate this issue, it is advised to employ bootstrapping techniques (Efron, 1979). Bootstrapping uses resampling with replacement (also known as Monte Carlo resampling) to estimate the statistic's sampling distribution (Haukoos and Lewis, 2005). It is used to assess the accuracy of an estimator such as the standard error, a confidence interval, or the bias of an estimator. The technique may be useful for analysing smallish expensive-to-collect data sets where prior information is sparse, distributional assumptions are unclear, and where further data may be difficult to acquire (Henderson, 2005). Adopting a cautious stance involves utilizing the mean plus one to three standard deviations, enhancing the reliability of Mc estimation in the context of curved FMDs Mignan and Woessner, 2012. Given the extensive range of simulations conducted and the multitude of results obtained, we decided to report particularly those that align with the σ and μ obtained with the EMR method for each different number of events tested. Accordingly, synthetic catalogs as illustrated in figures 13 though 16 have been selected.
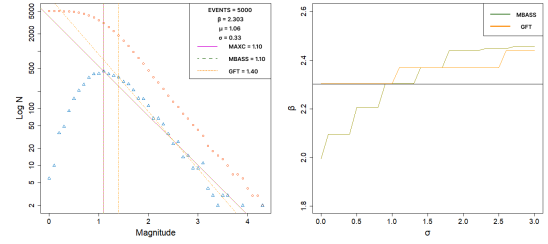
Fig. 13: Left side: Synthetic curved FMD (n.events: 5000, μ: 1.06, σ: 0.33. Vertical lines indicate Mc estimates by different methods: MAXC (purple), MBASS (dashed yellow), and GFT (dotted green), with related bias. Right side: Variations obtained with MBASS and GFT through bootstrap (sample dimension equal to 500).
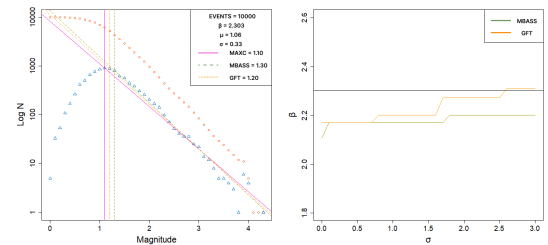
Fig. 14: Left side: Synthetic curved FMD (n.events: 10000, μ: 1.06, σ: 0.33. Vertical lines indicate Mc estimates by different methods: MAXC (purple), MBASS (dashed yellow), and GFT (dotted green), with related bias. Right side: Variations obtained with MBASS and GFT through bootstrap (sample dimension equal to 500).
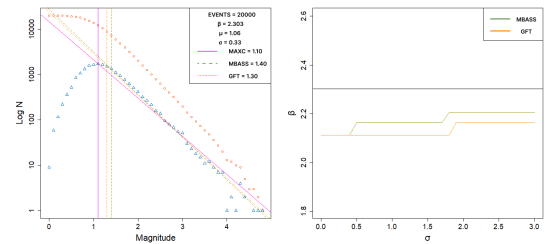
Fig. 15: Left side: Synthetic curved FMD (n.events: 20000, μ: 1.06, σ: 0.33. Vertical lines indicate Mc estimates by different methods: MAXC (purple), MBASS (dashed yellow), and GFT (dotted green), with related bias. Right side: Variations obtained with MBASS and GFT through bootstrap (sample dimension equal to 500).
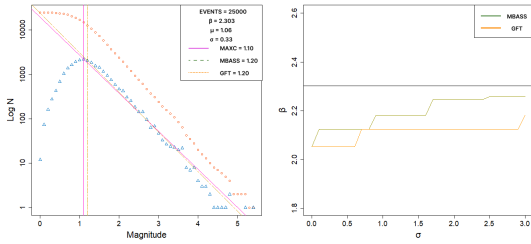
Fig. 16: Left side: Synthetic curved FMD (n.events: 25000, μ: 1.06, σ: 0.33. Vertical lines indicate Mc estimates by different methods: MAXC (purple), MBASS (dashed yellow), and GFT (dotted green), with related bias. Right side: Variations obtained with MBASS and GFT through bootstrap (sample dimension equal to 500).

3 Results and Discussion

The results presented are derived from an exhaustive analysis of methodologies acknowledged by the scientific community for calculating the magnitude of completeness. The catalog under investigation (INGV, https://www.ingv.it/) spans 18 years, from 2005 to 2022. It includes records of 26300 events during this period covering seismic activities in Sicily, Italy (fig n. 2). Employing diverse catalog-based methods for estimation, our results span a range of Mc values from 1.8 to 2.3. In details, the MBS method deviates significantly, yielding a markedly higher value of 3.1. These results are represented in figure n.17.

The entirety of the results derived from our analysis is outlined in tables n.2 through n.6. These tables collectively encapsulate the quantitative findings of our study, systematically presenting the data from single computations to resampling analyses. Starting with table n.2, we

introduce the baseline comparisons across various methodologies for the base and the declustered catalogs.

Concerning about declustering through windowing techniques, four different possibilities can be considered (Luen and Stark, 2012). The Gardner-Knopoff algorithm implement the simplest way to achieve its task, searching events in magnitude-descending order and remove any event within the window of the largest event. No distinction is made between primary aftershocks and aftershocks those originating due to the previous aftershocks, but only relies on the assumption that everything occurs within the window (Van Stiphout et al., 2012). When events available for the analysis are considerably low, the bootstrap resampling may be a helpful tool to achieve stable estimation of completeness. Furthermore, individual computations can introduce an additional amount of bias, leading to overestimation or underestimation of completeness (Mignan, 2012).

Progressing through to Table n.6, that demonstrates the stability, variability, and methodological robustness of our Mc estimations under different catalog conditions.

Given the scope of the regional catalog, starting with a simpler method like MAXC for an initial estimation of Mc could be practical, especially when the FMD shows a gradual curve (Mignan and Woessner, 2012). This method provides a baseline understanding of seismic completeness,

817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867

setting the stage for more detailed analyses. The MAXC technique offers several advantages that make it time-saving and easy to handle. Firstly, it relies on simple mathematical assumptions, specifically the calculation of derivatives. This makes it computationally efficient and relatively straightforward to implement. Secondly, despite the fact that MAXC tends to underestimate sistematically Mc in case of gradual curved shape FMD (Mignan and Woessner, 2012), it is widely employed as starting point to define the range of the potential completeness magnitudes for other methods.

The MAXC method for all the catalogues was around 1.8-1.9, similar to the results obtained with the MBASS method. However, concerning the results obtained from declustered catalogs, MAXC was not quite influenced by declustering and remained stable across all the catalogs, albeit with significant differences in terms of events. Especially in the catalog where the space-temporal window proposed by Gruenthal (figure n.19) was applied, which presented a considerably lower number of events.

Comparably, the MBASS approach demonstrated no notable susceptibility related to the number of events in order to obtain stable estimation of Mc, as a matter of fact the discrepancies relevated are on the order of 0.1. The MBASS shows a not significantly lowering of the Mc as the number of events increased, and case of underestimation and overestimation verify when resampling

is not taken into account. This phenomenon is illustrated in the simulation analysis, where the variations attributable to computational biases are mapped out, showcasing the potential for deviation from true values. Since MBASS is an iterative procedure, it has shown a higher Mc uncertainty compared to MAXC, even in catalogs with the highest number of events the Mc uncertainty remain quite similar.

The similarity in the results between MBASS and MAXC necessitates a cautious interpretation, particularly in light of MAXC's propensity to underestimate Mc in scenarios with gradually curved FMD. In our study MBASS behaves increasingly similar to MAXC as the curvature becomes more gradual. The changes caused by cleaning the origin catalog, which graphically results in a lowering of the curve, do not affect the identification of discontinuities. This may be closely related to the variation point approach, which is non-parametric and independent of any assumptions regarding the distribution and MAXC. This resemblance raises a potential concern about MBASS's effectiveness in accurately determining Mc in certain parts of our catalog, especially where seismic events are less frequent or smaller in magnitude. Despite MBASS's strengths in handling complex seismic data through its non-parametric and iterative approach, the possibility of underestimation must be carefully considered

in light of the similar behaviour of MAXC and MBASS showed in the simulation's analysis.

The EMR method has demonstrated noteworthy sensitivity to the number of recorded events, but only concerning the uncertainty. The most robust results were obtained for catalogs with higher number of events, in these cases the uncertainty recorded was less than 0.05. The estimated completeness obtained through single computation remained nearly unchanged for all different resampling size adopted, as it is clear looking at figures n.22 through n.25. Increasing the level of resampling did not alter the mean and standard deviation values of Mc obtained, however it is preferable to adopt resampling techniques, if doubts arise concerning undersampling, especially considering the difference in the results obtained for the catalog with the lowest number of events (n.19) between single computation and bootstrap. Specifically, in the catalogs with larger number of events, the uncertainty is quite similar to the results obtained with MAXC. On the contrary, in low event scenarios (n.18 and n.19) it recorded the highest uncertainty in the whole study. Despite the expected discrepancies observed for lower magnitude values (approx. M ¡ 1.0), the model generated by EMR method is able to fairly well describe the curvature of the FMD compared to the observed data. As a result, it provides a reasonable estimate of Mc, as well as the best compromise between a b-value tending to 1

and the lowest possible value of Mc among those obtained (10).

Considering the full spectrum of earthquake magnitudes it achieves the success in modelling the incomplete part. However, despite the capacity of the EMR model to fit the breadth of seismic activity within the catalog, the disparities in Mc values between the synthetic and the real catalogs indicate limitation in its effectiveness, particularly when delineating the threshold at which complete detection of seismic events is maintained. The method's apparent ability to trace the gradually curved shape of the FMD does not fully compensate for the discrepancies in Mc values. This discrepancy implies that the EMR model, while statistically robust in fitting the catalog data, may not fully encapsulate the operational aspects of seismic monitoring that influence the detection of smaller magnitude events. The authors themselves recommend stopping using this method cause of ineffectiveness and lack of foundation (Mignan and Woessner, 2012).

GFT and MBS produced the most conservative estimates of Mc obtained in this study, along with high uncertainty values.

The GFT results range between 2.3 and 2.5, the second highest Mc estimations of the entire study. It must be noted that individual computations can overestimate or underestimate the magnitude of completeness. This becomes evident when observing the effects of declustering and

919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969

19

resampling. The same result was obtained through individual computations for the declustered Gruenthal and Uhrhammer catalogs, which differ significantly in the number of events, as depicted in figures n.19 and n.20. However, a higher estimate was obtained for the Gardner-Knopoff catalog (18), showing similarity with the MBS. The use of resampling, even with a minimum bootstrap value, has shown to provide stable results (figure n.22). Specifically, it has been demonstrated that GFT underestimated Mc in the catalog with the lowest number of events (figure n.19), while overestimated in the Gardner-Knopoff catalog (figure n.18), which approximately had twice the number of events. The results obtained for the catalogs with higher number of events remained stable as illustrated from figures n.23 to n. 25.

The application of the Goodness-of-Fit Test provides valuable insights into the alignment of the observed seismic data with the expected statistical distribution according to the Gutenberg-Richter law. This alignment or lack thereof can help understanding the representativeness of our catalog in capturing the range of seismic activities. As a matter of fact, it is also used as indirect assessment of upgrades in the seismic network. Despite the well known limit of selecting the first Mco that achieves the test statistic, the results it provides offer a good compromise between managing undersampling and being conservative. It appears reasonable to consider slightly higher

value of the Mco that demonstrate improved performance in statistical testing as completeness.

The completeness estimated by MBS was the highest overall and exceeded 3.0, markedly diverging from other methods results (17). Initially, the computation failed to identify the completeness. Consequently, we focused on expanding the range of potential completeness values from 20 to 40. The initial range of potential values adopted from Mignan and Woessner (2012) was centred around the value of MAXC (1.8 for the original catalog), resulting in a maximum and minimum value of Mco of 1.1 and 3.0, respectively. At the first attempt after doubling the number of potential values, MBS immediately identified Mc above 3.0, confirming the initial suspicion and demonstrating that caution should be exercised in defining the values of Mco to be tested when applying this methodology. In declustered catalogs, the effect of event removal impacts MBS estimation, the most conservative results were obtained in catalogs with fewer events, as expected. However, there is a notable exception. Despite the original catalog having a greater number of events compared to the declustered Uhrammer catalog (figures n.17 and n.20), MBS produced a more conservative estimate of Mc for the non-declustered catalog.

The MBS method, which is based on the stability of the b-value, has shown limitations in our study area. Certain zones within the catalog coverage, particularly those with sparse seismic events,

may not be well-represented by this method. The absence of events in some areas suggests that the MBS method might not be entirely suitable for our scenario, where the seismic completeness just for the entire region is evaluated.

Mc results recorded by MBS clarify that it is not solely the number of events that influences the estimates but rather the characteristics of the FMD, as expected due to the scope of the catalog. Furthermore, the instability of the b-value for magnitude greater than Mc confirm a lack of larger events at the upper tail of the distribution.

4 Conclusions

Catalog-based methods offer a simple instrumental approach to compute Mc estimations. They are time-saving and easy to implement, as a matter of fact the computation involves only the magnitude vector to infer the completeness from the expected behaviour of seismicity (Mignan and Woessner, 2012).

The range of Mc values obtained from different techniques may not always overlap. Such discrepancies arise because these methods do not necessarily adhere to a uniform definition of Mc (Mignan and Woessner, 2012).

In summing up the findings of the study, it is crucial to acknowledge that even when a particular technique yields a low uncertainty estimate for the magnitude of completeness, it does not inherently affirm the reliability of the estimation itself. Rather, it suggests that the estimated output is robust within the context of the dataset analysed.

Declustering may have a direct impact on Mc estimation. Lowering the number of events did not always produce more conservative estimates, suggesting the influence of the shape of the FMD and the assumptions of the method employed.

The utilization of resampling techniques, has proven to be highly effective. By resampling the catalog data, it was possible to assess the stability and uncertainty of completeness estimates. Even with a minimum bootstrap value, the completeness results remained stable.

Different methods offer varied analytical perspective of Mc, emphasizing that no single approach is universally applicable. Therefore, choosing the appropriate methodology must be contingent upon the intended purpose of the study, ensuring that the assessment of Mc aligns with the context and the unique geological features of the region under investigation.

1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071

1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
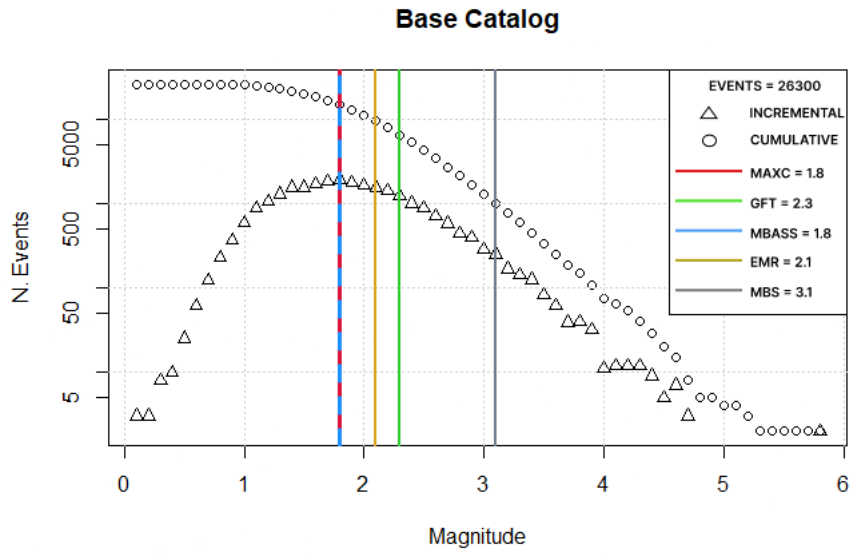
Fig. 17: Cumulative and incremental FMD of the observed catalog and corresponding Mc estimates (vertical lines)
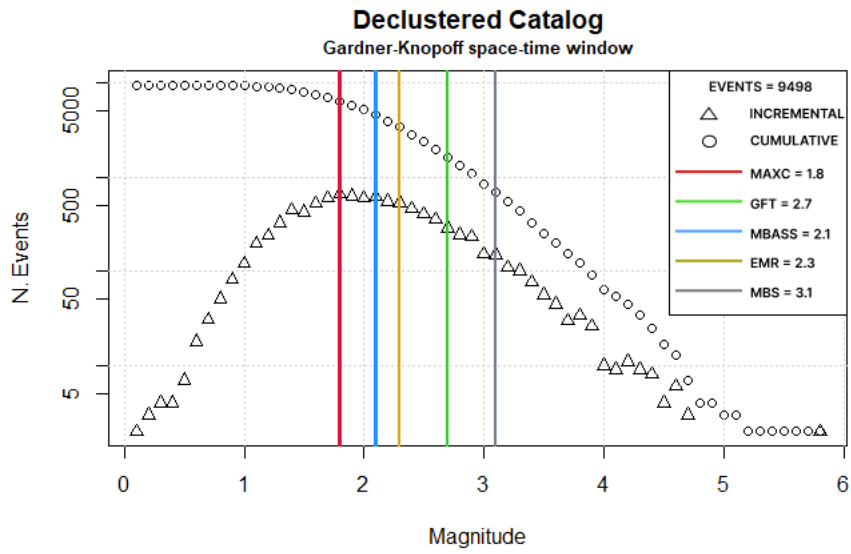
Fig. 18: Cumulative and incremental FMD of the declustered catalog with Gardner-Knopoff time-space window and corresponding Mc estimates (vertical lines)

22
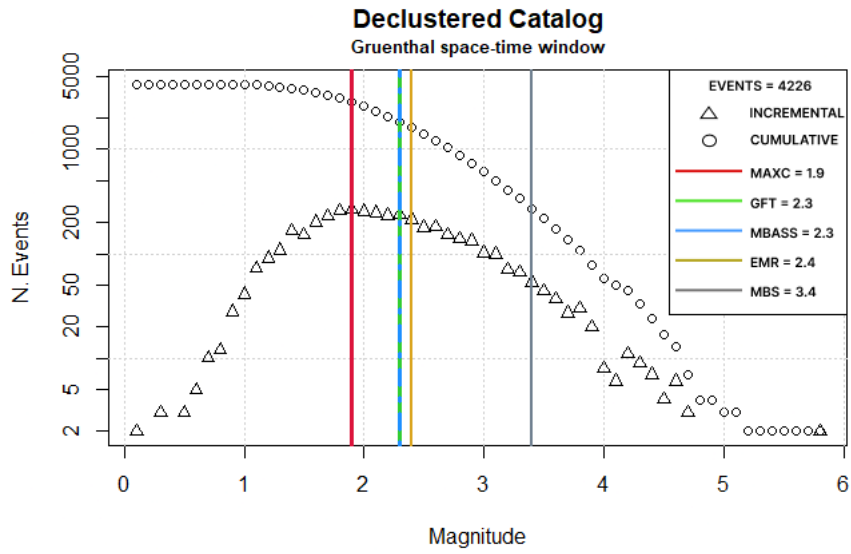
Fig. 19: Cumulative and incremental FMD of the declustered catalog with Gruenthal time-space window and corresponding Mc estimates (vertical lines)
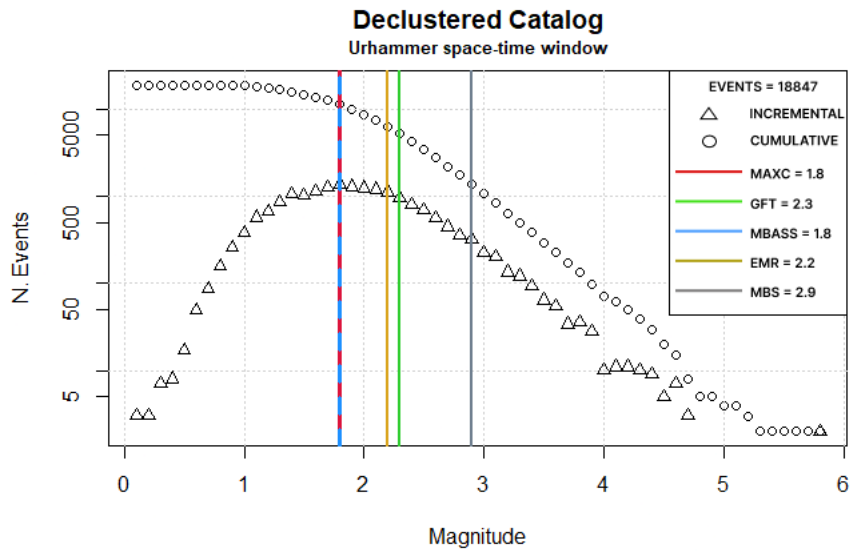
Fig. 20: Cumulative and incremental FMD of the declustered catalog with Uhrhammer time-space window and corresponding Mc estimates (vertical lines)
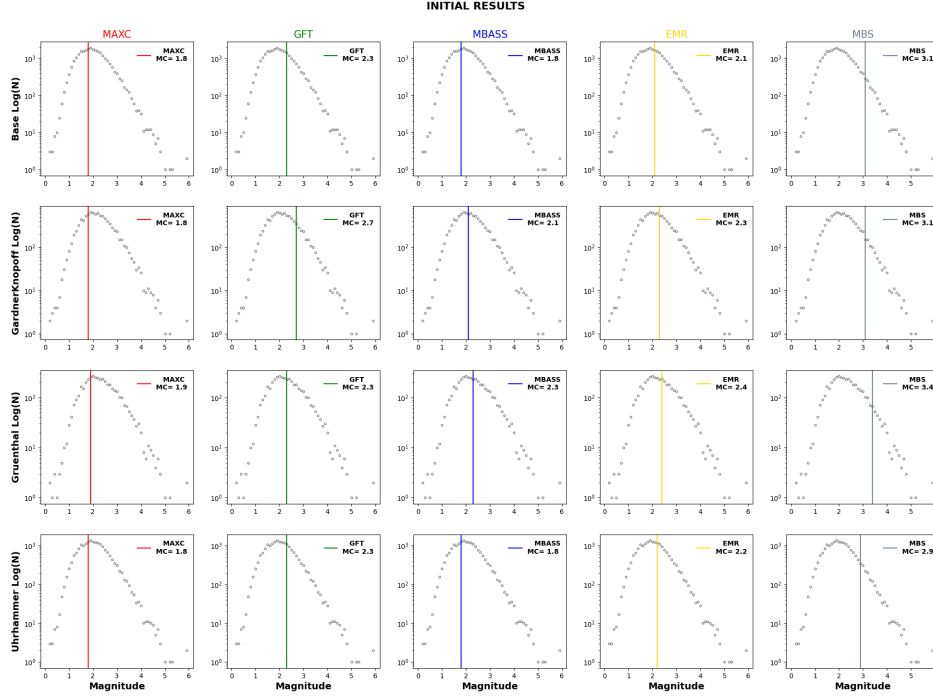
Fig. 21: Incremental FMDs of each catalog (rows) and each catalog-based method employed (columns). The figure shows the completeness results (vertical lines) obtained through single computation.

SINGLE COMPUTATION					
	MAXC	GFT	MBASS	EMR	MBS
Mc	1.80	2.30	1.80	2.10	3.10
b	0.71	0.83	0.80	0.93	0.94
a	4.11	3.73	4.11	5.94	2.89
	MAXC GK	GFT GK	MBASS GK	EMR GK	MBS GK
Mc	1.80	2.70	2.10	2.30	3.10
b	0.62	0.98	0.80	0.87	0.88
a	3.76	3.12	3.59	5.53	2.73
	MAXC GRU	GFT GRU	MBASS GRU	EMR GRU	MBS GRU
Mc	1.90	2.30	2.30	2.40	3.40
b	0.56	0.74	0.74	0.76	0.86
a	3.41	3.21	3.21	5.03	2.33
	MAXC UHR	GFT UHR	MBASS UHR	EMR UHR	MBS UHR
Mc	1.80	2.30	1.80	2.20	2.90
b	0.69	0.82	0.78	0.95	0.89
a	4.00	3.62	4.00	5.89	3.02

Table 2: Results for the origin catalog and the declustered catalogs through single computation for each Mc estimation method employed

24

Fig. 22: Incremental FMDs of each catalog (rows) and each catalog-based method employed (columns). The figure shows the completeness results (vertical lines), along with the relative uncertainty (dashed lines), obtained through bootstrap equal to 500.

1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
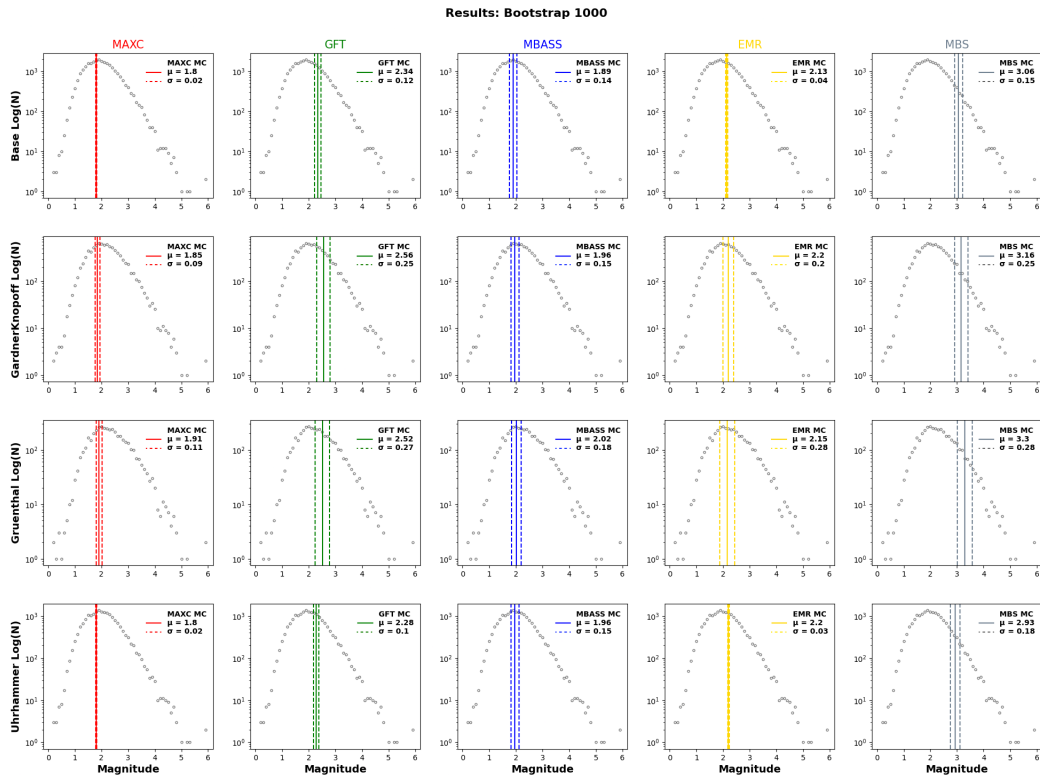
Fig. 23: Incremental FMDs of each catalog (rows) and each catalog-based method employed (columns). The figure shows the completeness results (vertical lines), along with the relative uncertainty (dashed lines), obtained through bootstrap equal to 1000.

Fig. 24: Incremental FMDs of each catalog (rows) and each catalog-based method employed (columns). The figure shows the completeness results (vertical lines), along with the relative uncertainty (dashed lines) obtained through bootstrap equal to 2000.

1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
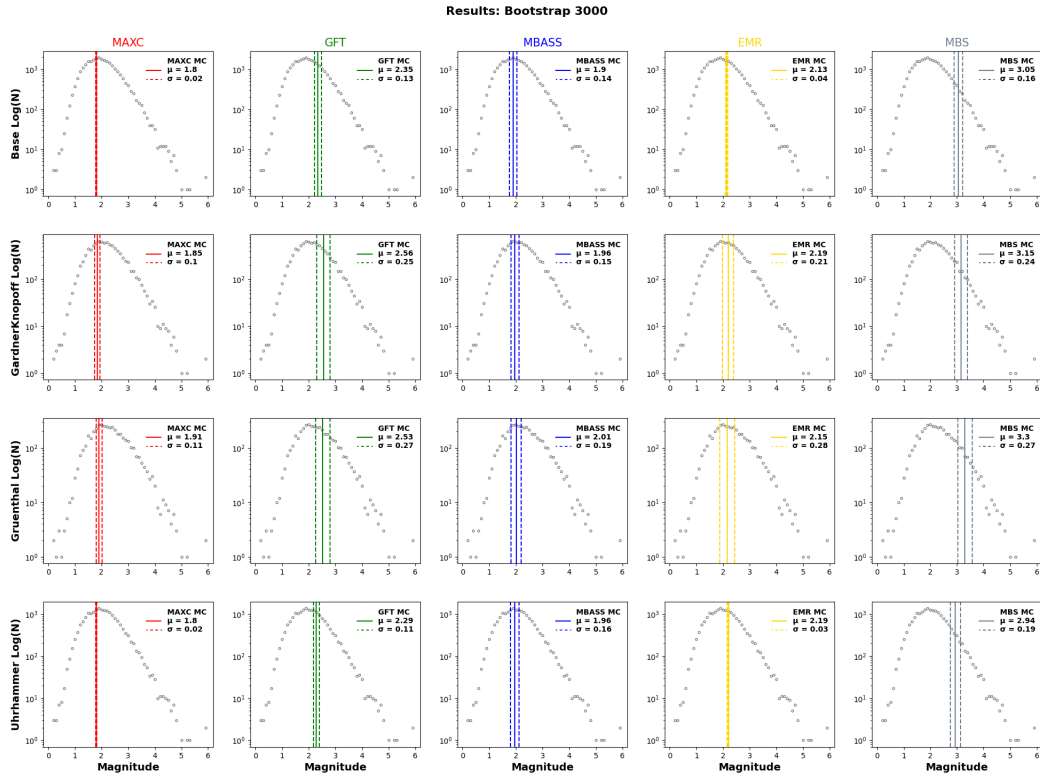
Fig. 25: Incremental FMDs of each catalog (rows) and each catalog-based method employed (columns). The figure shows the completeness results (vertical lines), along with the relative uncertainty (dashed lines) obtained through bootstrap equal to 3000.

1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479

| | BOOTSTRAP ORIGIN | | | | | | | | | |

BOOTSTRAP 500

	MAXC		GFT		MBASS		EMR		MBS	
	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std
Mc	1.80	0.02	2.34	0.12	1.89	0.14	2.13	0.04	3.06	0.15
b	0.71	0.02	0.89	0.07	0.84	0.06	0.94	0.02	1.00	0.11
a	4.12	0.01	3.69	0.11	4.05	0.10	5.96	0.05	2.94	0.17

BOOTSTRAP 1000

	MAXC		GFT		MBASS		EMR		MBS	
	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std
Mc	1.80	0.02	2.35	0.13	1.90	0.14	2.13	0.04	3.05	0.15
b	0.71	0.02	0.89	0.07	0.85	0.06	0.94	0.02	1.00	0.11
a	4.12	0.01	3.68	0.12	4.04	0.10	5.96	0.05	2.95	0.18

BOOTSTRAP 2000

	MAXC		GFT		MBASS		EMR		MBS	
	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std
Mc	1.80	0.02	2.34	0.13	1.90	0.14	2.13	0.04	3.05	0.16
b	0.71	0.02	0.89	0.08	0.84	0.06	0.94	0.02	1.00	0.12
a	4.12	0.01	3.69	0.12	4.05	0.10	5.96	0.05	2.95	0.19

BOOTSTRAP 3000

	MAXC		GFT		MBASS		EMR		MBS	
	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std
Mc	1.80	0.02	2.35	0.13	1.90	0.14	2.13	0.04	3.05	0.16
b	0.71	0.02	0.89	0.08	0.84	0.06	0.94	0.02	1.00	0.12
a	4.12	0.01	3.68	0.12	4.04	0.10	5.96	0.05	2.95	0.19

Table 3: Results for the base catalog through bootstrap sample size (500, 1000, 2000, 3000) for each method used (MAXC, GFT, MBASS, EMR, MBS)

29

		BOOTSTRAP GK									

		BOOTSTRAP 500									
		MAXC GK		GFT GK		MBASS GK		EMR GK		MBS GK	
		Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std
	Mc	1.85	0.09	2.54	0.25	1.97	0.15	2.21	0.20	3.15	0.24
	b	0.65	0.06	0.88	0.10	0.75	0.06	0.84	0.08	0.99	0.16
	a	3.73	0.05	3.25	0.21	3.67	0.09	5.44	0.20	2.69	0.27

		BOOTSTRAP 1000									
		MAXC GK		GFT GK		MBASS GK		EMR GK		MBS GK	
		Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std
	Mc	1.85	0.09	2.56	0.25	1.96	0.15	2.20	0.20	3.16	0.25
	b	0.65	0.06	0.89	0.10	0.75	0.06	0.83	0.08	0.99	0.17
	a	3.73	0.05	3.24	0.21	3.67	0.09	5.43	0.20	2.68	0.28

		BOOTSTRAP 2000									
		MAXC GK		GFT GK		MBASS GK		EMR GK		MBS GK	
		Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std
	Mc	1.85	0.09	2.56	0.25	1.96	0.15	2.20	0.21	3.15	0.25
	b	0.65	0.06	0.89	0.10	0.74	0.06	0.83	0.08	0.98	0.16
	a	3.73	0.05	3.24	0.21	3.67	0.09	5.43	0.21	2.68	0.29

		BOOTSTRAP 3000									
		MAXC GK		GFT GK		MBASS GK		EMR GK		MBS GK	
		Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std
	Mc	1.85	0.10	2.56	0.25	1.96	0.15	2.19	0.21	3.15	0.24
	b	0.65	0.06	0.89	0.10	0.75	0.06	0.83	0.08	0.98	0.16
	a	3.73	0.05	3.24	0.21	3.67	0.09	5.42	0.21	2.69	0.29

Table 4: Results for the declustered catalog with the Gardner and Knopoff (1974) time-space window through bootstrap sample size (500, 1000, 2000, 3000) for each method used (MAXC, GFT, MBASS, EMR, MBS)

30

1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581

				BOOTSTRAP GRU						

				BOOTSTRAP 500						
	MAXC GRU		GFT GRU		MBASS GRU		EMR GRU		MBS GRU	
	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std
Mc	1.91	0.11	2.52	0.28	2.03	0.19	2.14	0.28	3.28	0.27
b	0.58	0.05	0.78	0.10	0.65	0.06	0.69	0.09	0.97	0.18
a	3.41	0.05	3.06	0.19	3.35	0.09	4.82	0.24	2.46	0.28

				BOOTSTRAP 1000						
	MAXC GRU		GFT GRU		MBASS GRU		EMR GRU		MBS GRU	
	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std
Mc	1.91	0.11	2.52	0.27	2.02	0.18	2.15	0.28	3.30	0.28
b	0.58	0.05	0.78	0.10	0.65	0.06	0.69	0.09	0.98	0.20
a	3.40	0.05	3.05	0.19	3.35	0.09	4.83	0.24	2.43	0.29

				BOOTSTRAP 2000						
	MAXC GRU		GFT GRU		MBASS GRU		EMR GRU		MBS GRU	
	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std
Mc	1.92	0.11	2.53	0.27	2.01	0.19	2.15	0.28	3.30	0.27
b	0.58	0.05	0.78	0.10	0.65	0.06	0.69	0.09	0.98	0.18
a	3.40	0.05	3.05	0.19	3.35	0.09	4.83	0.24	2.44	0.28

				BOOTSTRAP 3000						
	MAXC GRU		GFT GRU		MBASS GRU		EMR GRU		MBS GRU	
	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std
Mc	1.91	0.11	2.53	0.27	2.01	0.19	2.15	0.28	3.30	0.27
b	0.58	0.05	0.78	0.10	0.65	0.06	0.69	0.09	0.98	0.19
a	3.40	0.05	3.05	0.19	3.35	0.09	4.83	0.23	2.44	0.29

Table 5: Results for the declustered catalog with the Gruenthal (Van Stiphout et al., 2012) time-space window through bootstrap sample size (500, 1000, 2000, 3000) for each method used (MAXC, GFT, MBASS, EMR, MBS)

BOOTSTRAP UHR

BOOTSTRAP 500

	MAXC UHR		GFT UHR		MBASS UHR		EMR UHR		MBS UHR	
	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std
Mc	1.80	0.02	2.28	0.10	1.95	0.15	2.20	0.03	2.93	0.18
b	0.69	0.02	0.90	0.07	0.85	0.07	0.95	0.02	0.94	0.09
a	4.00	0.01	3.64	0.09	3.89	0.11	5.89	0.04	3.00	0.20

BOOTSTRAP 1000

	MAXC UHR		GFT UHR		MBASS UHR		EMR UHR		MBS UHR	
	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std
Mc	1.80	0.02	2.28	0.10	1.96	0.15	2.20	0.03	2.93	0.19
b	0.69	0.02	0.90	0.07	0.85	0.07	0.95	0.02	0.95	0.10
a	4.00	0.01	3.64	0.09	3.89	0.11	5.89	0.04	2.99	0.20

BOOTSTRAP 2000

	MAXC UHR		GFT UHR		MBASS UHR		EMR UHR		MBS UHR	
	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std
Mc	1.80	0.02	2.28	0.10	1.96	0.15	2.20	0.03	2.94	0.19
b	0.70	0.02	0.90	0.07	0.85	0.07	0.95	0.02	0.95	0.10
a	4.00	0.01	3.64	0.09	3.89	0.11	5.89	0.04	2.98	0.21

BOOTSTRAP 3000

	MAXC UHR		GFT UHR		MBASS UHR		EMR UHR		MBS UHR	
	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std	Mean	Dev.Std
Mc	1.80	0.02	2.29	0.11	1.96	0.16	2.19	0.03	2.94	0.19
b	0.70	0.02	0.90	0.07	0.85	0.07	0.95	0.02	0.95	0.10
a	4.00	0.01	3.64	0.10	3.89	0.11	5.89	0.04	2.98	0.21

Table 6: Results for the declustered catalog with the Uhrhammer (1986) time-space window through bootstrap sample size (500, 1000, 2000, 3000) for each method used (MAXC, GFT, MBASS, EMR, MBS)

References

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

Abd el-aal, A.E.-A.: Very broadband seismic background noise analysis of permanent good vaulted seismic stations. Journal of Seismology **17** (2012) https://doi.org/10.1007/s10950-012-9308-5

Amorese, D.: Amorèse d., 2007, applying a change-point detection method on frequency-magnitude distributions, bull. seism. soc. am. 97, no 5, 1742-1749. Bulletin of the Seismological Society of America **97**, 1742–1749 (2007) https://doi.org/10.1785/0120060181

Cao, A.M., Gao, S.S.: Temporal variations of seismic b-values beneath northeastern japan island arc. Geophys. Res. Lett. **29** (2002)

Cornell, C.A.: Engineering seismic risk analysis. Bulletin of the Seismological Society of America **58**(5), 1583–1606 (1968) https://doi.org/10.1785/BSSA0580051583

D'Alessando, A., Greco, L., Scudero, S., Lauciani, V.: Spectral characterization and spatiotemporal variability of the background seismic noise in italy. Earth and Space Science **8** (2021) https://doi.org/10.1029/2020EA001579

Efron, B.: Bootstrap methods: Another look

33

at the jackknife. The Annals of Statistics **7**(1), 1–26 (1979) https://doi.org/10.1214/aos/1176344552

Gardner, J.K., Knopoff, L.: Is the sequence of earthquakes in southern california, with aftershocks removed, poissonian? Bull. Seismol. Soc. Am. **64**, 1363–1367 (1974)

Gomberg, J.: Seismicity and detection/location threshold in the southern great basin seismic network. Journal of Geophysical Research: Solid Earth **96**(B10), 16401–16414 (1991)

Gonzalez, A.: The spanish national earthquake catalogue: Evolution, precision and completeness. Journal of Seismology **21**(3), 435–471 (2017) https://doi.org/10.1007/s10950-016-9610-8

Gutenberg, B., Richter, C.F.: Frequency of earthquakes in California*. Bulletin of the Seismological Society of America **34**(4), 185–188 (1944) https://doi.org/10.1785/BSSA0340040185 https://pubs.geoscienceworld.org/ssa/bssa/article-pdf/34/4/185/5300856/bssa0340040185.pdf

Henderson, A.R.: The bootstrap: a technique for data-driven statistics. using computer-intensive analyses to explore experimental data. Clinica Chimica Acta **359**(1-2), 1–26 (2005) https://doi.org/10.1016/j.cccn.2005.04.002

Helmstetter, A., Kagan, Y.Y., Jackson, D.D.: Comparison of short-term and time-independent earthquake forecast models for southern california. Bulletin of the Seismological Society of America **96**(1), 90–106 (2006)

Haukoos, J.S., Lewis, R.J.: Advanced statistics: Bootstrapping confidence intervals for statistics with "difficult" distributions. Academic Emergency Medicine **12**(4), 360–365 (2005) https://doi.org/10.1197/j.aem.2004.11.018

Huang, W.Q., Li, W.X., Cao, X.F., *et al.*: Completeness analysis of earthquake catalogs in china main land i. Acta Seismologica Sinica **16**(3), 273–280 (1994). In Chinese

Hollander, M., Wolfe, D.A.: Nonparametric Statistical Methods. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, New York (1973)

Ishimoto, M., Iida, K.: Observations of earthquakes registered with the microseismograph constructed recently. Bull. Earthq. Res. Inst. **17**, 443–478 (1939)

Juellyan, J., Setiawan, B., Hasan, M., Yunita, H., Sungkar, M., Saidi, T.: Comparing gardner-knopoff, gruenthal, and uhrhammer earthquake declustering methods in aceh, indonesia. IOP Conference Series: Earth and Environmental

Science **1245**, 012010 (2023) https://doi.org/10.1088/1755-1315/1245/1/012010

Kagan, Y.Y.: Short-term properties of earthquake catalogs and models of earthquake source. Bulletin of the Seismological Society of America **94**(4), 1207–1228 (2004)

Kim, S.K.: Comparison of earthquake catalogs declustered from three different methods in the korean peninsula. Research Square (2022). PREPRINT (Version 1)

Lanzante, J.R.: Resistant, robust and non-parametric techniques for the analysis of climate data: Theory and examples, including applications to historical radiosonde station data. International Journal of Climatology **16**, 1197–1226 (1996)

Liu, J., Chen, Q.F., Chen, Y.: Completeness analysis of the seismic catalog in north china region. Earthquake **16**(1), 59–67 (1996)

Luen, B., Stark, P.B.: Declustering and poisson tests. Geophysical Journal International **189**, 691–700 (2012)

Mignan, A., Chouliaras, G.: Fifty years of seismic network performance in greece (1964–2013): Spatiotemporal evolution of the completeness magnitude. Seismological Research Letters **85**, 657–667 (2014) https://doi.org/10.1785/0220130209

Mignan, A.: Functional shape of the earthquake frequency-magnitude distribution and completeness magnitude. J. Geophys. Res. **117**(B08302) (2012)

Mignan, A., Woessner, J.: Estimating the magnitude of completeness for earthquake catalogs. Community Online Resource for Statistical Seismicity Analysis (2012) https://doi.org/10.5078/corssa-00180805

Ogata, Y., Katsura, K.: Analysis of temporal and spatial heterogeneity of magnitude frequency distribution inferred from earthquake catalogues. Geophysical Journal International **113**, 727–738 (1993)

Ogata, Y., Katsura, K.: Immediate and updated forecasting of aftershock hazard. Geophysical Research Letters **33** (2006) https://doi.org/10.1029/2006GL025888

Ringdal, F.: On the estimation of seismic detection thresholds. Bulletin of the Seismological Society of America **65**(6), 1631–1642 (1975) https://doi.org/10.1785/BSSA0650061631

Rydelek, P.A., Sacks, I.S.: Testing the completeness of earthquake catalogs and the hypothesis of self-similarity. Nature **337**, 251–253 (1989)

Sereno, T.J.J., Bratt, S.R.: Seismic detection capability at noress and implications for the detection threshold of a hypothetical network

in the soviet union. Journal of Geophysical Research: Solid Earth **94**(B8), 10397–10414 (1989)

Schorlemmer, D., Mele, F., Marzocchi, W.: A completeness analysis of the national seismic network of italy. Journal of Geophysical Research **115**, 04308 (2010) https://doi.org/10.1029/2008JB006097

Schorlemmer, D., Wiemer, S., Wyss, M.: Variations in earthquake-size distribution across different stress regimes. Nature **437**, 539–42 (2005) https://doi.org/10.1038/nature04094

Uhrhammer, R.: Characteristics of northern and central california seismicity. Earthquake Notes **57**(1), 21 (1986)

Van Stiphout, T., Zhuang, J., Marsan, D.: Seismicity declustering. Community Online Resource for Statistical Seismicity Analysis (2012)

Wesnousky, S.G.: The gutenberg-richter or characteristic earthquake distribution, which is it? Bulletin of the Seismological Society of America **84**, 1940–1958 (1994)

Wyss, M., Hasegawa, A., Wiemer, S., Umino, N.: Quantitative mapping of precursory seismic quiescence before the 1989, m7.1 off-sanriku earthquake, japan. Annali Di Geofisica **42**, 851–869 (1999)

Wilcoxon, F.: Individual comparisons by ranking methods. Biometrics Bulletin **1**, 80–83 (1945)

Wiemer, S., Wyss, M.: Minimum Magnitude of Completeness in Earthquake Catalogs: Examples from Alaska, the Western United States, and Japan. Bulletin of the Seismological Society of America **90**(4), 859–869 (2000) https://doi.org/10.1785/0119990114

Woessner, J., Wiemer, S.: Assessing the quality of earthquake catalogues: Estimating the magnitude of completeness and its uncertainty. Bulletin of the Seismological Society of America **95**(2), 684–698 (2005)

Xu, W., Gao, M.: Statistical analysis of the completeness of earthquake catalogs in china mainland. Chinese Journal of Geophysics **57**(9), 2802–2812 (2014)

Zuniga, F.R., Wyss, M.: Inadvertent changes in magnitude reported in earthquake catalogs: Their evaluation through b-value estimates. Bulletin of the Seismological Society of America **85**, 1858–1866 (1995)

**A.2   Figlioli A., Vitale G., Taroni M., D'Alessandro A.
Tremors—A Software App for the Analysis of the
Completeness Magnitude.**
Geosciences. **2024; 14(6):149.**
https://doi.org/10.3390/geosciences14060149

Tremors—A Software App for the Analysis of the Completeness Magnitude

Anna Figlioli [1,*], Giovanni Vitale [2], Matteo Taroni [3] and Antonino D'Alessandro [2]

1 Dipartimento di Scienze della Terra e del Mare, Università degli Studi di Palermo, 90123 Palermo, Italy
2 Istituto Nazionale di Geofisica e Vulcanologia, Osservatorio Nazionale Terremoti, 00143 Rome, Italy; giovanni.vitale@ingv.it (G.V.); antonino.dalessandro@ingv.it (A.D.)
3 Istituto Nazionale di Geofisica e Vulcanologia, Roma 1, 00143 Rome, Italy; matteo.taroni@ingv.it
* Correspondence: anna.figlioli@unipa.it

Abstract: This paper introduces a software tool developed within the MATLAB environment, called Tremors, aimed at streamlining the pre-processing and analysis of seismic catalogues, with a particular emphasis on determining the Magnitude of Completeness. It will outline the criteria for event selection, as well as various techniques to derive the Magnitude of Completeness values, including the recent and widely used Lilliefors statistical method. The study also addresses the important issue of short-term aftershock incompleteness and proposes solutions for managing it. Moreover, the software generates high-quality, customizable figures, and georeferenced raster images in .tif format as output. A standalone version of the App is also available (i.e., the users do not need a MATLAB license on their PC/laptop).

Keywords: magnitude of completeness; b-value; Short-Term Aftershocks Incompleteness (STAI); statistical analysis

1. Introduction

The realm of seismology necessitates a deep comprehension of seismic catalogs and the complex analyses they yield. These catalogs are invaluable to scientific research, aiding in the exploration of seismic processes and enhancing our grasp of the dynamics behind earthquake generation and propagation.

The proper usability of seismic catalogs hinges largely on the precise estimation of the Magnitude of Completeness (Mc). The Mc is the threshold magnitude above which seismic events are considered complete (i.e., no missing events above this threshold).

The determination of Mc is critical for computing subsequent parameters, such as the b-value of the Gutenberg–Richter law. Misjudging the Mc could result in inaccurate b-value estimations or a flawed assessment of the national seismic network's efficacy.

The domain of Mc estimation has undergone significant evolution over recent decades, with the advent of various innovative techniques and methodologies that have markedly impacted seismological studies. Among the methods documented in the literature [1–3], we find the MAXimum Curvature (MAXC), the Goodness of FiT (GFT), and the Absolute Magnitude Based Method (Mbass). These conventional methods have long served as benchmarks in seismology, establishing a reliable foundation for Mc estimation. With technological advancements and analytical algorithm evolution, the Lilliefors method has emerged [4]. This technique provides a statistically robust approach to determining Mc by testing the exponentiality of the magnitudes.

To streamline the intricate process of Mc estimation, we introduce an intuitive software application. Our tool is built to facilitate the calculation of the Magnitude of Completeness (Mc) on modern seismic catalogs, also addressing the challenges posed by time-varying incomplete data.

Moreover, our App strives to offer guidance for optimal catalog processing, enabling users to select data based on classical parameters such as time, magnitude, depth, and survey area. The App focused attention particularly on the problem of Short-Term Aftershocks Incompleteness (STAI) [5,6].

The challenge with STAI is linked to significant events that lead to temporary inflation in Mc, marked by a substantial increase in the seismicity rate. Modeling STAI periods automatically has not been successful yet. To catch them, seismologists must look at the incremental number vs. magnitude plot [4,7,8]. Therefore, this procedure necessitates subjective decisions to determine the number and length of STAI periods. For scientific reproducibility, it is essential to document all subjective decisions made by seismologists when they study the Mc of a seismic catalog, including the parameters used to exclude STAIs [9].

With respect to ZMAP [10], one of the most used software for the computation of the Mc of seismic catalog, our App offers two advantages: it contains modern methods to compute the magnitude of completeness, along with a heuristic technique to catch STAI periods; it can be used as a stand-alone code, without a MATLAB license. This second advantage can be important, especially for those researchers who have fewer funds and cannot afford paid software.

This paper will delve into the App's principal features and demonstrate how it simplifies seismic catalog analysis, rendering it accessible even to novices in the field.

2. The Different Features of Tremors

Tremors has a graphical user interface (see Figure 1) with a horizontal navigation bar at the top that can be used to access the various functions described below.
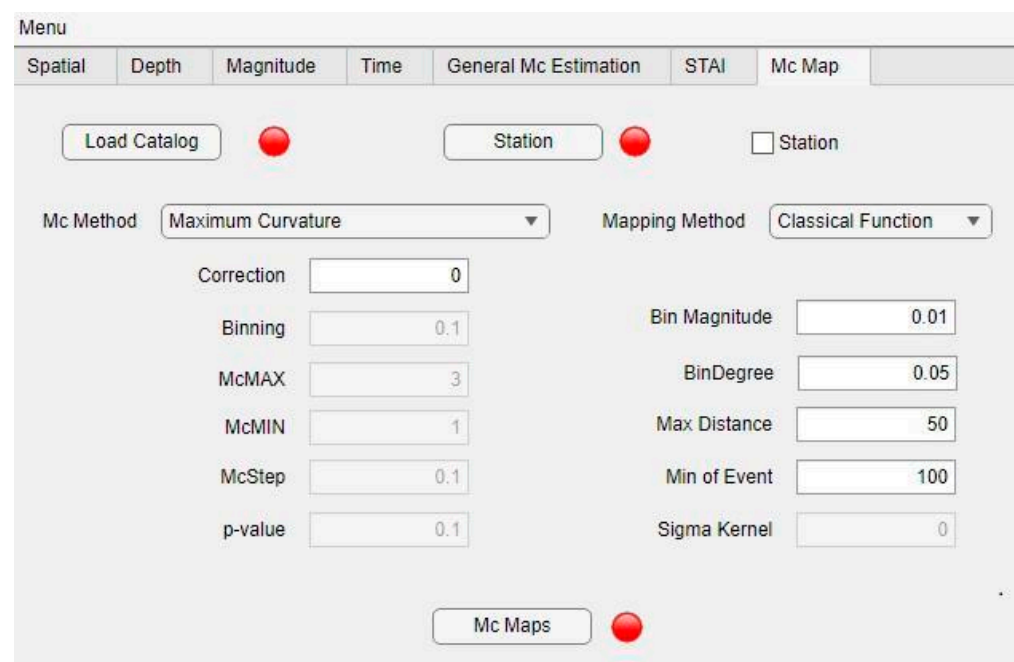
Figure 1. The graphical interface of Tremors.

Spatial Selection: The analysis of seismicity within a specific region requires careful assessment of the spatial distribution of earthquake epicenters. When defining the study area, it is crucial to consider potential disturbances that could lead to an over/under-estimation of the magnitude of completeness. Seismic catalogs generally show increased incompleteness with distance from seismic stations, due in part to the reduced robustness of data for more distant events, which are located using fewer seismic stations. A fundamental principle is to establish a perimeter to work within an area where the data are robust. The

Tremors tool allows users to draw a polygon or load an existing one to select the desired work area. The seismic catalog input format is the 10-column ZMAP format [10], ensuring full compatibility with ZMAP, which has been the predominant seismicity software over the past two decades.

Time selection: Seismic catalogs spanning several decades must be analyzed taking into account possible changes in the seismic network over time. Modern instruments capture a wide frequency range with highly sensitive and nearly real-time capabilities, influenced by the spatial distribution and density of installed stations. If significant dates for network updates are known, the catalog can be temporally segmented to analyze different periods separately, since Mc is strongly dependent on the seismic network.

Depth Selection: Analyzing a seismic catalog also involves considering depth, the third spatial dimension. Therefore, we have to determine a depth range where the location of hypocenters and the detection of seismicity are considered reliable. Human activities or natural events can contaminate the seismic catalog within the first few hundred meters, especially anthropogenic activities like quarry blasts or explosions, which distort seismicity parameters [11,12]. Moreover, as depth increases, so does Mc, due to the growing distance between seismicity and monitoring infrastructure. As a rule of thumb, the end-of-range depth should be determined by looking at the magnitude vs. depth plot.

Magnitude selection: This App also offers a sub-section where it is possible to cut the catalog according to a minimum magnitude threshold. We underline that this minimum magnitude threshold is not the completeness magnitude. This functionality could be useful in the case of catalogs with a very large number of events, to speed up the computations of the other sub-sections.

General Mc Estimation: The core function of this app is to estimate the completeness magnitude of a seismic catalog, taking into account constraints such as area, depth, magnitude, and time. This estimation can be performed using three different methodologies. One of the most recent methods is that of Herrmann and Marzocchi [4], which is based on the Lilliefors test [13], a statistically robust technique. This method checks the compatibility with the Gutenberg–Richter law, i.e., an exponential distribution for the magnitudes [14]. The catalog could be considered complete if the Lilliefors test does not reject the hypothesis of exponentiality for the magnitudes above a certain threshold. This threshold will be the magnitude of completeness of the catalog. This approach does not require any parameter, apart from the *p*-value of the test (we suggest setting the value $p \geq 0.1$, as in Herrmann and Marzocchi [4]). Another approach implemented, similar to the one of Herrmann and Marzocchi [4], is the method of Taroni (2023) [15]. It uses a simple random variable transformation ($\frac{X}{X+Y}$) to transform exponential random variables to Uniform [0, 1] variables, and it is more effective in the case of catalogs containing events with different b-values [15].

Finally, for fast computations, we also included the classical maximum curvature method [16]; this method requires a positive correction of about 0.2 or 0.3 to be reliable. The output of this sub-section is a figure containing the plot of the *p*-value of the test (in the first two approaches), the magnitude frequency distribution along with the estimated magnitude of completeness, and the incremental number vs. magnitude plot [8] (Figure 2). The latter is useful for identifying any time-dependent incompleteness in the catalog. Indeed, the previously described estimation methods are useful to compute the general completeness, i.e., the time-independent completeness threshold of the whole catalog. However, such a Mc value could be ineffective after the strongest events in the catalog. Looking at the lower panel of Figure 2, it is possible to notice areas with a low density of points at the bottom and vertical alignment of points forming small white (empty) triangles: these zones represent the temporal part of the incompleteness. To eliminate this type of incompleteness, we subsequently rely on the STAI-removing technique (next paragraph).
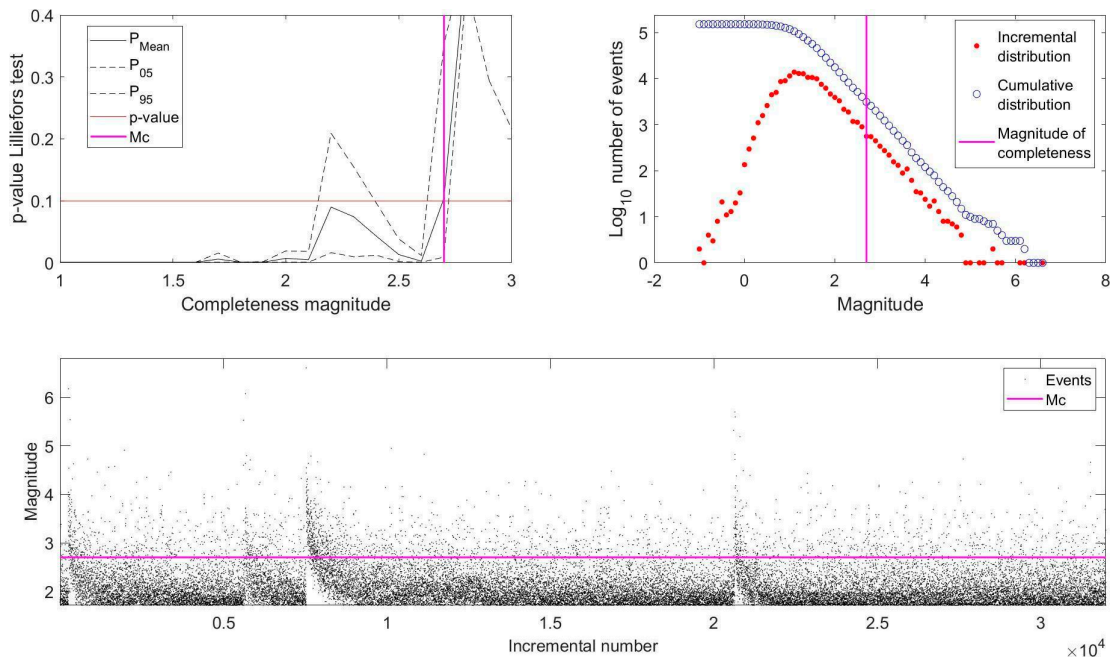
Figure 2. Upper left panel: *p*-value of the test as a function of the magnitude; the vertical magenta line represents the estimated magnitude of completeness; black solid and dotted lines represent the *p*-value of the test with the uncertainties. Upper left panel: incremental (red) and cumulative (blue) distribution of the magnitudes; the vertical magenta line represents the estimated magnitude of completeness. Lower panel: incremental number vs. magnitude plot; the horizontal magenta line represents the estimated magnitude of completeness; empty spaces in the lower bottom part indicate a short-term aftershock incompleteness.

STAI investigation: Handling Short-Term Aftershock Incompleteness (STAI) periods presents challenges due to their complex characteristics. These periods typically follow the strongest seismic events in the catalog but exhibit variability influenced by multiple factors. While one potential solution involves removing STAI periods using methods described in the existing literature [7,17], the Tremors App takes a different approach to prevent data loss. Instead of outright removal, it temporarily increases the magnitude of completeness after significant events in the catalog. This strategy balances the need for robust data while preserving valuable information (Figure 3).
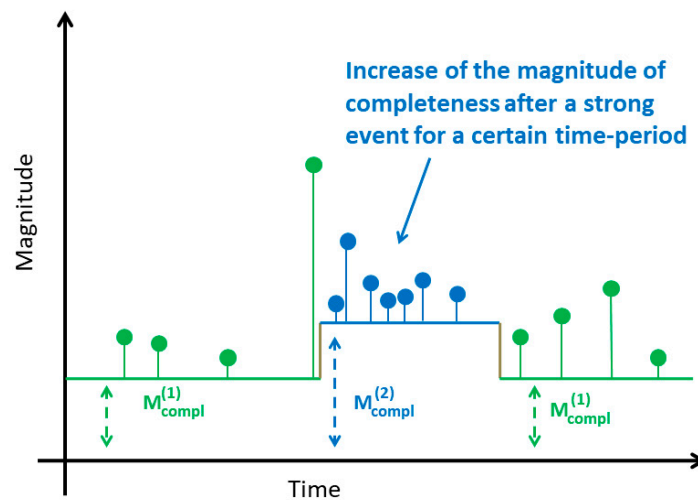
Figure 3. A simple scheme to illustrate the increase in the completeness magnitude (blue color) with respect to the background value (green color) after a strong event.

To properly identify STAI periods we used the incremental number vs. magnitude plot [8]. Our methodology involves the use of three parameters that allow the result obtained to be uniquely reproducible. The first one is M: a magnitude threshold that defines the minimum magnitude of events that generate the STAI. Usually, STAI periods start to become evident after events with a magnitude larger than 5.0, but this threshold also depends on the general magnitude of completeness. The second one is T: it defines the length of the STAI period expressed in days. The last parameter is the Delta Magnitude of Completeness (DMC), i.e., how much increases the completeness level after the strong event M. Once these three parameters are defined, is it possible to remove events in the identified STAI periods. This procedure could be applied repeatedly and for different M magnitude thresholds. Tremors provide three magnitude thresholds, each with its own set of parameters. This is a "try and catch" procedure: the user has to try different parameters to frame the STAI periods. In Figure 4 we show an example of STAI investigation using the catalog of the Amatrice-Norcia 2016-17 sequence [18]. At the end of the processing, a plot of the incremental number versus M-Mc (the magnitude of the events minus the corresponding completeness threshold) will help verify whether all STAI periods are adequately treated [7] (Figure 5). This plot is very similar to the one suggested by [8], but it uses M-Mc instead of magnitudes. When we obtain a plot without "white holes" or vertical alignments of points, we can consider this procedure complete. We underline that a precise description of the STAI periods via the set of parameters used is essential to increase the reproducibility of any seismic analyses.
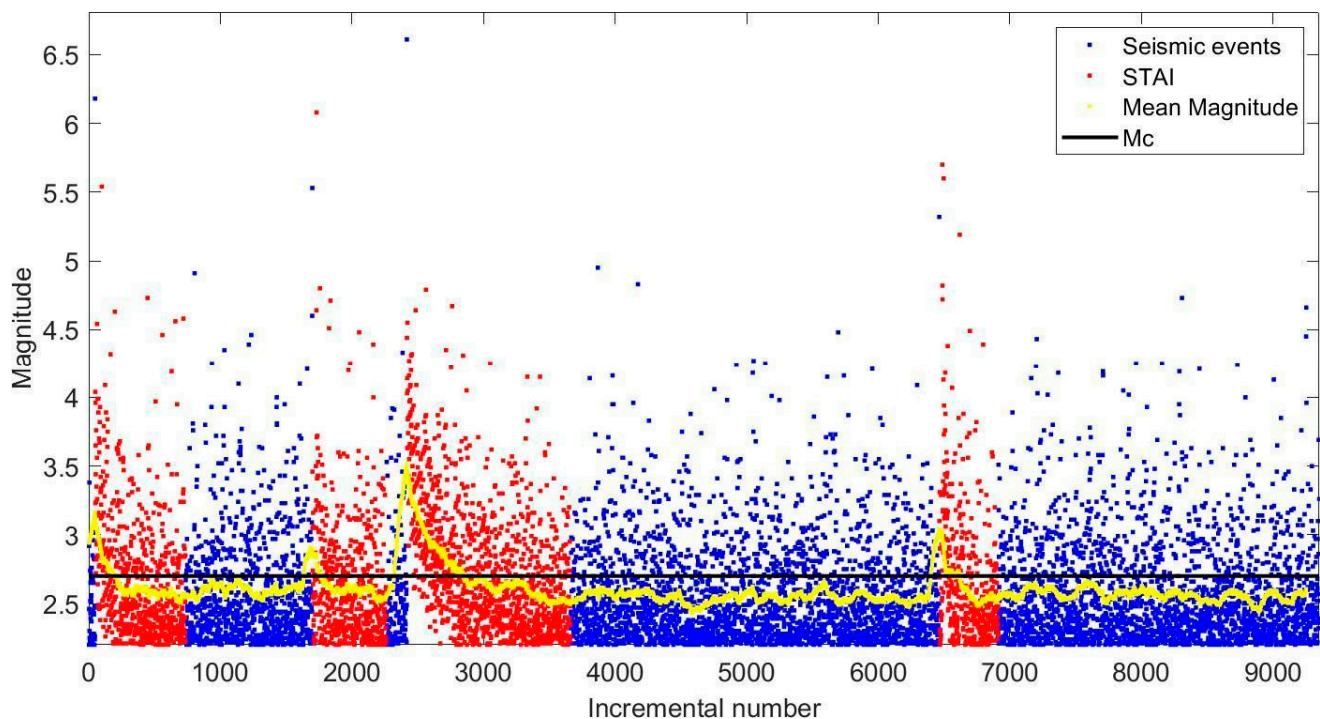
Figure 4. Incremental number vs. magnitude plot. Red dots represent events during STAI periods, while blue dots are relative to normal periods. The black horizontal line is the general Mc computed for the whole catalog. The yellow curve is the rolling-window average magnitude; it increases during STAI periods, and it helps to catch incomplete time periods.

Mc map analysis: In this last section of the App, we show how it is possible to analyze the spatial variation of the magnitude of completeness. Regional catalogs, with a spatial distribution of epicenters spanning hundreds to thousands of km, must be analyzed with techniques that allow for spatially dependent M*c* [16,19]. The spatial analysis of the M*c* allows us to understand whether the national seismic network in that particular

area has significant gaps [20,21]. We propose three types of spatial Mc calculations from the literature [4,16]. Figure 6 shows the result of the analysis performed on the Italian instrumental catalog from 2005 to 2022 [22] with the MAXC+0.2 method. We underline that these results are useful for understanding relative variations of the M*c* but are not reliable for estimating the M*c* in a certain zone (also because this spatial analysis does not consider the STAI problem). In any case, the results obtained from the application of the spatial calculation of the magnitude show that most of the peninsula has a Mc value around 1.4 to 1.8, with most of the lower values in Central Italy, where we have a greater coverage of the national seismic network.
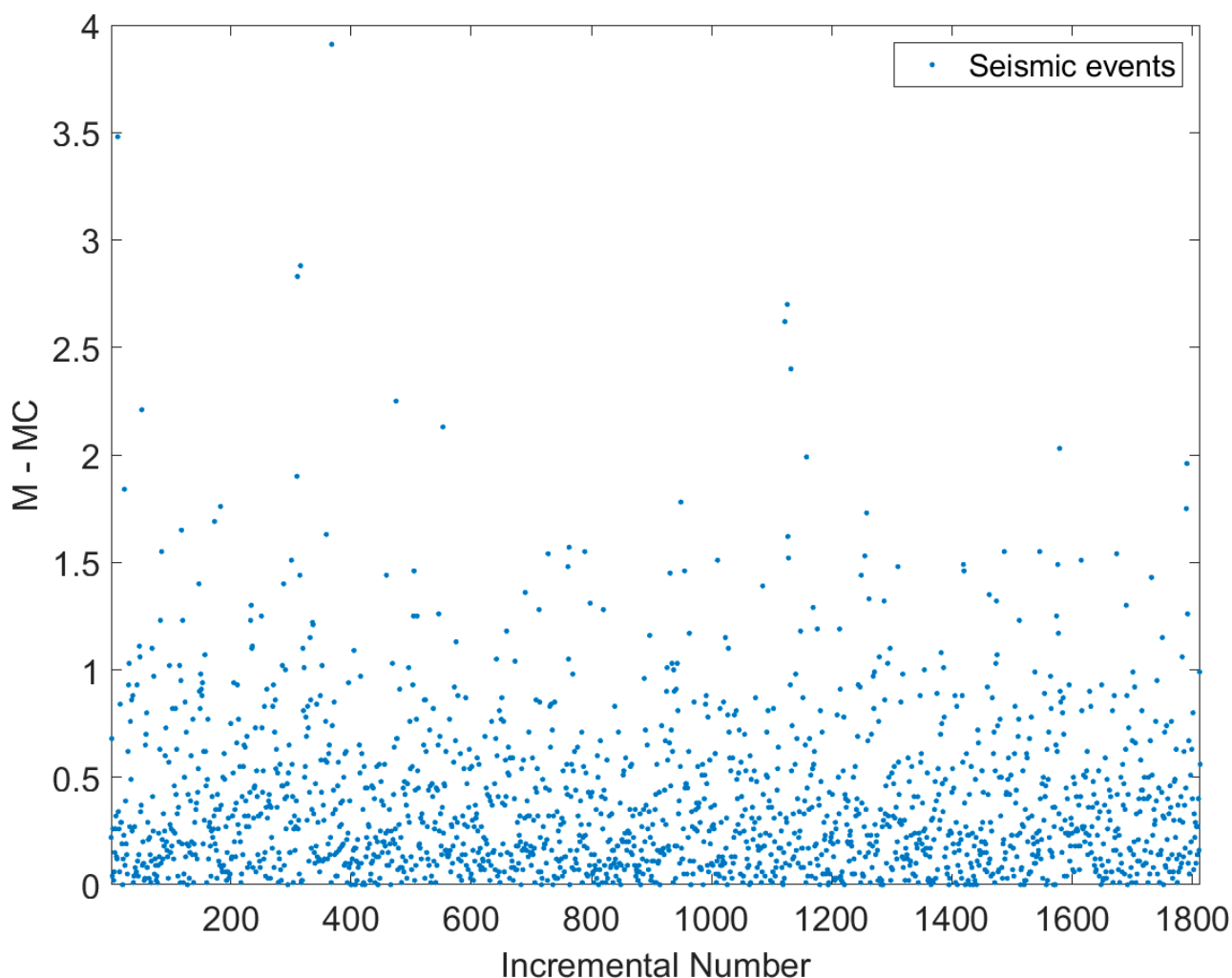
Figure 5. Incremental number vs. M-Mc plot. As the incremental number vs. magnitude plot, it helps to catch possible incompleteness in the catalog. It is useful to understand if the STAI periods were properly treated.

There are some gaps caused by the lack of data and the algorithm cannot estimate the Mc (central part of Piedmont, Sardinia, Southern Apulia). In the Po plain, we notice a very high magnitude of completeness tending to 2.8. In this case, this value is the result of various aspects, including the constant presence of industries that tend to produce considerable seismic noise and the thick sediments that characterize that plain. Other areas where we have a high magnitude of completeness are near the Sicilian offshore, where network coverage is poor and onshore stations cannot detect small events.

All the figures can be saved in .png or .mat (MATLAB) format, and also in a georeferenced .tiff format; in this way, the user can plot the calculations using other software.
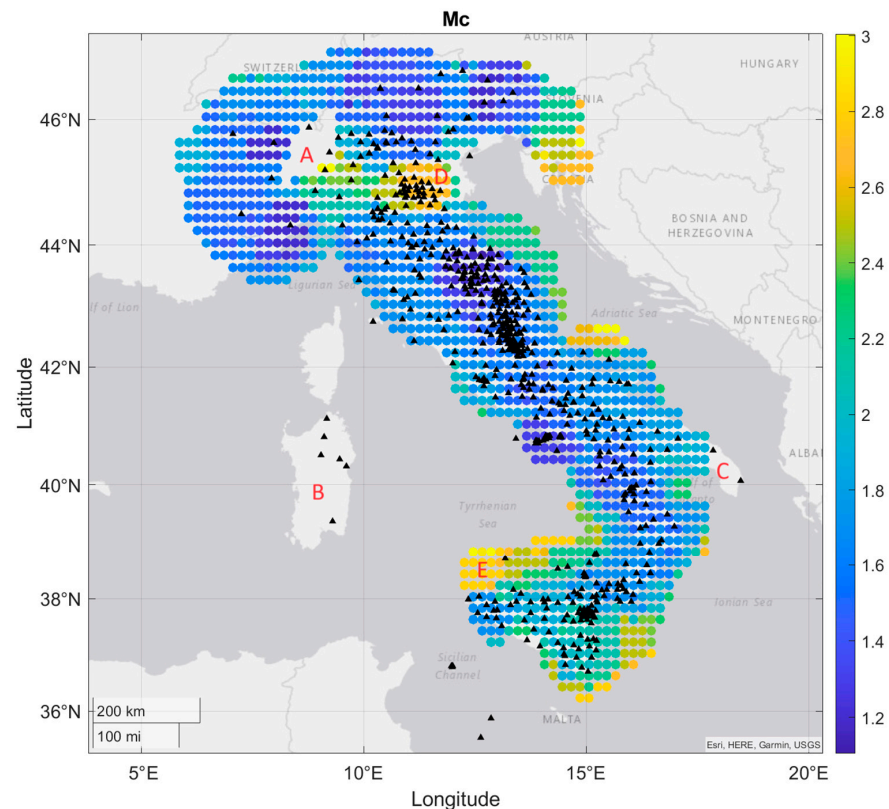
Figure 6. Map of the completeness magnitude analysis, with the stations of the seismic network (black triangles). Red letters indicate the zones cited in the text (A for Piedmont, B for Sardinia, C for Southern Apulia, D for Po plain, E for Sicilian offshore).

3. Conclusions and Future Developments

In conclusion, the Tremors App allows performing all the magnitude of completeness (Mc) estimation steps and facilitates the creation of a catalog for subsequent processing tasks, such as b-value estimation. The App is meticulously designed to augment the user experience by offering functionalities for spatial, depth, magnitude, and temporal selections, complemented by a variety of techniques for calculating Mc. Its user-friendly interface simplifies the workflow for seismologists, enabling them to concentrate on the fundamental aspects of seismic analysis.

Looking to the future, there are numerous exciting opportunities to enhance the app's functionality and utility. A notable forthcoming enhancement is the integration of b-value analysis techniques, both spatially and temporally. This development promises to provide a more comprehensive understanding of seismic activity by examining the variations in b-values across different regions and over various time periods.

Author Contributions: A.F. conceived the app and conducted analysis, G.V. conducted the development of the App, M.T. implemented algorithms and functions for catalog analyses, A.D. analyzed the results. All authors reviewed the manuscript. All authors have read and agreed to the published version of the manuscript.

Additional Information: To download the code, there is a repository named "Tremors" at https://github.com/GammaArietis/Tremors. The app was developed using MATLAB 2022b, with all toolboxes installed, and App designed. A standalone version of "Tremors" is also available in the same repository. For any questions, contact giovanni.vitale@ingv.it, matteo.taroni@ingv.it.

References

1. Cao, A.; Gao, S. Temporal variation of seismic b-values beneath northeastern Japan island arc. *Geophys. Res. Lett.* **2002**, *29*, 48-1–48-3. [CrossRef]
2. Amorese, D. Applying a change-point detection method on frequency-magnitude distributions. *Bull. Seismol. Soc. Am.* **2007**, *97*, 1742–1749. [CrossRef]
3. Wiemer, S.; Wyss, M. Minimum magnitude of completeness in earthquake catalogs: Examples from Alaska, the western United States, and Japan. *Bull. Seismol. Soc. Am.* **2000**, *90*, 859–869. [CrossRef]
4. Herrmann, M.; Marzocchi, W. Inconsistencies and lurking pitfalls in the magnitude–frequency distribution of high-resolution earthquake catalogs. *Seismol. Res. Lett.* **2021**, *92*, 909–922. [CrossRef]
5. Kagan, Y.Y. Short-term properties of earthquake catalogs and models of earthquake source. *Bull. Seismol. Soc. Am.* **2004**, *94*, 1207–1228. [CrossRef]
6. Lolli, B.; Gasperini, P. Comparing different models of aftershock rate decay: The role of catalog incompleteness in the first times after main shock. *Tectonophysics* **2006**, *423*, 43–59. [CrossRef]
7. Taroni, M.; Selva, J.; Zhuang, J. Estimation of the tapered Gutenberg-Richter distribution parameters for catalogs with variable completeness: An application to the Atlantic Ridge seismicity. *Appl. Sci.* **2021**, *11*, 12166. [CrossRef]
8. Zhuang, J.; Ogata, Y.; Wang, T. Data completeness of the Kumamoto earthquake sequence in the JMA catalog and its influence on the estimation of the ETAS parameters. *Earth Planets Space* **2017**, *69*, 36. [CrossRef]
9. Dascher-Cousineau, K.; Lay, T.; Brodsky, E.E. Two foreshock sequences post Giulia and Wiemer. *Seismol. Soc. Am.* **2020**, *91*, 2843–2850.
10. Wiemer, S. A software package to analyze seismicity: Zmap. *Seismol. Res. Lett.* **2001**, *72*, 373–382. [CrossRef]
11. Gulia, L. Detection of quarry and mine blast contamination in European regional catalogues. *Nat. Hazards* **2010**, *53*, 229–249. [CrossRef]
12. Gulia, L.; Gasperini, P. Contamination of frequency–magnitude slope (b-value) by quarry blasts: An example for Italy. *Seismol. Res. Lett.* **2021**, *92*, 3538–3551. [CrossRef]
13. Lilliefors, H.W. On the Kolmogorov-Smirnov test for the exponential distribution with mean unknown. *J. Am. Stat. Assoc.* **1969**, *64*, 387–389. [CrossRef]
14. Aki, K. Maximum likelihood estimate of b in the formula log n= a-bm and its confidence limits. *Bull. Earthq. Res. Inst. Tokyo Univ.* **1965**, *43*, 237–239.
15. Taroni, M. Estimating the magnitude of completeness of earthquake catalogs using a simple random variable transformation. *Seism. Rec.* **2023**, *3*, 194–199. [CrossRef]
16. Woessner, J.; Wiemer, S. Assessing the quality of earthquake catalogues: Estimating the magnitude of completeness and its uncertainty. *Bull. Seismol. Soc. Am.* **2005**, *95*, 684–698. [CrossRef]
17. Herrmann, M.; Piegari, E.; Marzocchi, W. Revealing the spatiotemporal complexity of the magnitude distribution and b-value during an earthquake sequence. *Nat. Commun.* **2022**, *13*, 5087. [CrossRef]
18. Chiaraluce, L.; Michele, M.; Waldhauser, F.; Tan, Y.J.; Herrmann, M.; Spallarossa, D.; Beroza, G.C.; Cattaneo, M.; Chiarabba, C.; De Gori, P.; et al. A comprehensive suite of earthquake catalogues for the 2016-2017 Central Italy seismic sequence. *Sci. Data* **2022**, *9*, 710. [CrossRef]
19. Mignan, A.; Werner, M.J.; Wiemer, S.; Chen, C.-C.; Wu, Y.-M. Bayesian estimation of the spatially varying completeness magnitude of earthquake catalogs. *Bull. Seismol. Soc. Am.* **2011**, *101*, 1371–1385. [CrossRef]
20. Schorlemmer, D.; Woessner, J. Probability of detecting an earthquake. *Bull. Seismol. Soc. Am.* **2008**, *98*, 2103–2117. [CrossRef]
21. Godano, C.; Convertito, V.; Pino, N.A.; Tramelli, A. An automated method for mapping independent spatial b values. *Earth Space Sci.* **2022**, *9*, e2021EA002205. [CrossRef]
22. Lolli, B.; Randazzo, D.; Vannucci, G.; Gasperini, P. The homogenized instrumental seismic catalog (HORUS) of Italy from 1960 to present. *Seismol. Soc. Am.* **2020**, *91*, 3208–3222. [CrossRef]

A.3 **Scudero, S., D'Alessandro A., Figlioli A.**
Evaluation of the earthquake monitoring network in
Taiwan.
Journal of Seismology. **27, 643–657 (2023).**
https://doi.org/10.1007/s10950-023-10162-8

RESEARCH

Evaluation of the earthquake monitoring network in Taiwan

Salvatore Scudero · Antonino D'Alessandro · Anna Figlioli

Abstract In this work, we perform an evaluation of the coverage of the earthquake monitoring network of Taiwan. The capability of a general network is a function of an adequate number of optimally distributed nodes. For this case study, the evaluation is performed with a statistical approach which includes descriptive spatial statistics in combination with point pattern techniques. The spatial distribution of the nodes of the earthquake monitoring network is analyzed in comparison with the distribution of seismicity, completeness magnitude, active seismogenic sources, seismic hazard, and population distribution. All these data can be put in relationship with the objectives of an earthquake monitoring network; therefore, they can be used, in turn, to retrieve information about the consistency of the network itself. In particular, we investigate the "Real-time Seismic Monitoring Network" and the "Strong-Motion Earthquake Observation Network," each one characterized by its own objectives, and therefore respectively compared with external information related to their purposes such as seismicity, seismogenic sources, seismic hazard, and population distribution. This simple and reliable approach reveals the high quality of the networks established in Taiwan. In general, it is able to provide quantitative information on the coverage of any type of network, identifying possible critical areas and addressing their future development.

1 Introduction

A country-scale earthquake monitoring network usually extends over areas in the range of 10^5–10^6 km^2 (D'Alessandro 2019) and, from its initial planning to its final configuration may be necessary for several years to decades. Throughout this process, the final geometrical arrangement of the nodes could turn remarkably different from the initial, planned configuration. For example, the Italian National Seismic Network (D'Alessandro et al. 2011a), the Red Sísmica Nacional in Spain (D'Alessandro et al. 2013), the Romanian network (D'Alessandro et al. 2012), and also the Global Seismographic Network (GSN) took on average 20–30 years (or even more) to reach their current configurations, and they will keep developing further in the future (Mezcua 1995; Gee and

S. Scudero (✉) · A. D'Alessandro
Istituto Nazionale di Geofisica e Vulcanologia, Osservatorio Nazionale, Terremoti, Via di Vigna Murata 605, 00143 Rome, Italy
e-mail: salvatore.scudero@ingv.it

A. Figlioli
Università degli studi di Palermo, Dipartimento di Scienze della Terra e del Mare (DiSTeM), Via Archirafi, 20, 90123 Palermo, Italy

Leith 2011; Popa et al. 2015; Michelini et al. 2016).
In some other cases, a network could be the result
of the merge between several local sub-networks, as
for instance happened in Greece (D'Alessandro et al.
2011b). As a counterpart, in some cases, the long
time span necessary to build up the networks turned
into a benefit because of the concurrent technologi-
cal improvements in sensors, transmission systems,
etc. In all the cases, it should always be opportune to
verify periodically whether the distribution of nodes
within a given network copes with its final objectives,
mainly because the knowledge about the seismicity of
a given area evolves over time.

In this work, we propose the application of the
approach proposed by Siino et al. (2020) for the eval-
uation of the earthquake monitoring network in Tai-
wan. This approach considers the spatial distribution
of the nodes of the network together with the ancillary
information related to the aims of the network itself.
Taiwan island is characterized by a relevant seismic-
ity (Theunissen et al. 2010; Shin and Teng 2001; Wu
et al. 2013), accountable with the active geodynamics
of this region which involves active subduction and
collision (Barrier and Angelier 1986; Sibuet and Hsu
1997; Rau and Wu 1998; Shyu et al. 2005; Chin et al.
2019). The seismic release is very high: more than
one hundred M>6.0 earthquakes are recorded in the
last 25 years; the largest earthquake that occurred in
recent times is the M_w 7.6, 1999 Chi-Chi earthquake
(Shin and Teng 2001; Kim et al. 2010).

The beginning of the instrumental earthquake
monitoring in Taiwan dates back to 1897 when the
very first seismograph was installed (Shin et al.
2013). Since then, a network has been established
and expanded over the whole country, mainly
boosted later by the technological advancements in
the 1970s and 1990s (i.e., digital signals, real-time
data transmission, and telemetry), and it is still
expanding and improving at present (Wang 1989;
Shin et al. 2013). The map in Fig. 1 shows the main
seismotectonic features of the region together with
the locations of the earthquake monitoring network
in Taiwan. For the abovementioned reasons, the
island of Taiwan represents an interesting case study
to evaluate the earthquake monitoring networks.
Moreover, the last paper dealing with this topic in
Taiwan dates back to 1997 (Tsai and Wu 1997). In
detail, we considered separately the part of the net-
work which includes the broad-band velocimeters

(hereafter *Real-time Seismic Monitoring Network*
), from the part of the network which includes the
accelerometers (hereafter *Strong-Motion Earth-
quake Observation Network*). Each of the two spe-
cific networks is then analyzed in comparison with
various data sets (e.g., seismicity, seismogenic
sources, the magnitude of completeness, seismic
hazard, and population distribution) according to
their mutual relationships as suggested by Siino
et al. (2020). Contributing to the assessment of the
degree of coverage of the network, this work can
possibly address also its future optimizations.

2 Data

In this section, we introduce the data sets and their
sources. The main information concerns the spatial
distribution of the earthquake networks which are
processed together with the ancillary information
related to the objectives.

All the information on the earthquake monitor-
ing network in Taiwan (i.e., location, type of sensor,
international code) and about seismicity is available
at the website of the Seismological Centre of the Cen-
tral Weather Bureau (see Data and Resources).

Concerning the observation network, we focused
on two different typologies, namely the 24-bit Real-
time Seismic Monitoring Network (also known as
Central Weather Bureau Seismographic Network,
CWBSN) which counts 119 ground stations and
9 OBS at present, and it is in charge of the seismic
surveillance of the country and the Strong-Motion
Earthquake Observation Network (also known as
Taiwan Strong Motion Instrumentation Program net-
work, TSMIP) which counts 420 monitoring sites
at present. This last network collects full records
of strong earthquakes to understand the seismic
activities under different geological conditions. It
also serves as the basis for the earthquake intensity
assessment and earthquake-resistant design norms
which are tools to effectively reduce and prevent
earthquake disasters. Technical details about these
networks are provided by Tsai and Lee (2005), Hsiao
et al. (2009), Chang et al. (2012), Shin et al. (2013),
Guan et al. (2020), and Central Weather Bureau
(CWB, Taiwan) (2012).

The selected catalogue of instrumental seismicity
covers the period from January 1, 1991, to December
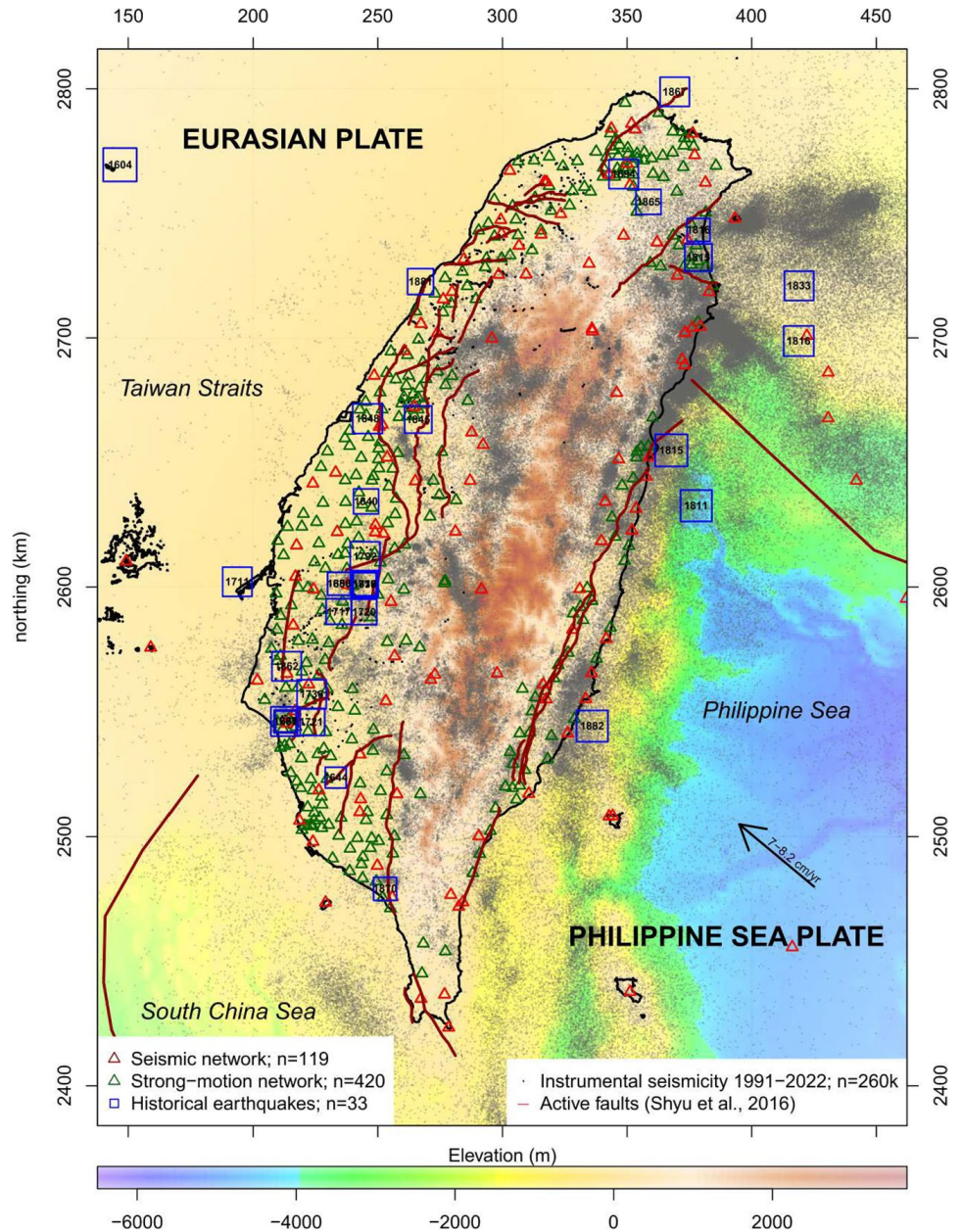
31, 2022, and counts 259,290 events, with earthquake magnitude ranging in the interval 2.2–7.3. The minimum magnitude corresponds to the magnitude of completeness (M_c) calculated for the entire catalogue of 787,674 events with a magnitude within the range of 0–7.3. The values of (M_c) are coherent with the ones for inland Taiwan provided by Mignan et al. (2011).

The catalogue for the historical seismicity results from the combination of two different lists, namely the one from Cheng and Yeh (1989) and the one from Chen and Tsai (2008). In the case of multiple magnitude estimation for a single event,

we took into account the worst scenario considering the highest magnitude for those events with different values. The merged list counts 33 earthquakes with M> 5 that occurred in the period from 1604 to 1988. Besides, the events with $M >= 5.5$ from both instrumental and historical earthquakes have been merged into a unique list of 382 events including all those earthquakes usually considered beyond the damage threshold. For these events, we estimated the seismic moment (M_0) by converting the provided values of local magnitude by means of the relation proposed by Hanks and Kanamori (1979):

$$M_0 = 10^{(M+10.7)} * 1.5$$

The mapped active faults within the study area are retrieved from the global active faults database (Styron and Pagani 2020). They have been collected from several geological and geophysical studies at various scales. Most of the seismic sources are from Shyu et al. (2016) which are also the same that have been used as the base for the evaluation of the seismic hazard model. The map of the active faults contains 44 elements with lengths ranging from ~10 to ~200 km. The main fault systems can be directly framed within the broad-scale geodynamic setting of region 1.

This study also considers the earthquake hazard of Taiwan. Seismic hazard maps provide information about the likelihood of ground shaking in a given area and represent fundamental tools for estimating the likelihood of damage and losses. The more recent Taiwan probabilistic seismic hazard model is the one from Chan et al. (2020) named "TEM PSHA2020" and hereinafter briefly "TEM." The TEM includes the updated earthquake database and seismogenic structure database together with a new set of ground motion prediction equations and site amplification factors. The probabilistic model assesses the likelihood that a given ground acceleration threshold is overcome in a given time span. The TEM model is included in the data analysis considering the probability of exceedance equal to 10% in 50 years, which is the most common reference to display a seismic hazard model.

Finally, we considered the distribution of the population over the study area. Data, in the form of population density (i.e., number of people per map unit), come from official United Nations population estimates and are updated to 2015 (see Data and Resources).

3 Method

In this paper, data are analyzed by means of descriptive spatial statistics and point process methods following the approach proposed by Siino et al. (2020). The methodologies are briefly recalled in the following paragraphs. The analysis is performed with

R statistical software (R Development Core Team 2005), and the functions employed are in the packages *raster* (Hijmans 2018), *spatstat* (Baddeley and Turner 2005), and *spatstat.local* (Baddeley 2018).

3.1 Point process methods

A point process model X is a random finite subset of $W \subset R^d$, where $|W| < \infty$. A spatial point pattern $v = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$, which is a realization of X, is an unordered set of points in the region W, where $N(v) = n$ is the number of points and $d = 2$.

For a spatial point process, the first-order intensity ($\lambda(\mathbf{u})$) is the expected number of points in a small region around a location ($\mathbf{u}$) divided by its area ($d\mathbf{u}$) in the limit as $1d\mathbf{u}$ goes to 0:

$$\lambda(\boldsymbol{u}) = \lim_{|d\boldsymbol{u}| \to 0} \frac{E(N(d\boldsymbol{u}))}{|d\boldsymbol{u}|} \tag{1}$$

where E is the expected number of points. The intensity ($\lambda(\mathbf{u})$) is assumed to be inhomogeneous, so it is not constant over the considered region; for spatial point patterns, the intensity is estimated non-parametrically to understand the spatial trend.

The usual kernel estimator of the intensity function is computed as proposed by Baddeley et al. (2015):

$$\hat{\lambda}(\mathbf{u}) = \frac{1}{e(\mathbf{u})} {}_{i=1}^{n} k(\mathbf{u} - \mathbf{u}_i, \boldsymbol{h}) \tag{2}$$

where $e(\mathbf{u})$ is the edge correction and $k(\cdot)$ is a bivariate Gaussian density distribution function with a standard deviation (smoothing bandwidth) equal to $\boldsymbol{h} = (h_x, h_y)$. The bandwidth controls the degree of smoothing, and a larger set of values gives more smoothing. The estimation of the bandwidth vector $\boldsymbol{h} = (h_x, h_y)$ is computed with Scott's rule (Scott 1992):

$$h_x = n^{-1/6} \sqrt{var(X)} \quad h_y = n^{-1/6} \sqrt{var(Y)} \tag{3}$$

In the proposed approach, the datasets concerning the nodes of the seismic and accelerometric networks and the instrumental and historical seismicity are treated as point patterns, and their spatial intensities are properly computed using the nonparametric estimator in Eq. (2).

Generally, it can be valuable to describe the first-order characteristics of a point pattern by studying its relationship with external variables (also called

covariates). In particular, for a seismic network, the intensity is computed depending on an external covariate, namely the distance to the nearest geological fault, $D(\mathbf{u})$, to understand the spatial displacement of the nodes relative to a seismic source. An inhomogeneous Poisson model with a parametric linear form is assumed following Baddeley et al. (2015):

$$\lambda(\mathbf{u}) = \beta_0 + \beta_1 D(\mathbf{u}) \tag{4}$$

where $D(\mathbf{u})$ is a spatial covariate and β_0 and β_1 are the parameters to be estimated. The parameters in Eq. (4) can be interpreted as follows: β_0 (intercept) gives the intensity at locations close to a geological fault (i.e., $D(\mathbf{u}) \approx 0$), while the intensity changes by a factor of β_1 (slope) for every 1 unit of distance away from the seismic source.

Moreover, we locally infer the spatially varying estimates of these two parameters of the inhomogeneous Poisson process model in Eq. (4). This technique is called "Geographically Weighted Regression" (GWR). GWR is a data exploration technique that allows understanding changes in the importance of different variables over space which may indicate that the model used is mis-specified and can be improved (Fotheringham et al. 1998). Regression models are typically "global": that is, all data are used simultaneously to fit a single model. In some cases, it can make sense to fit more flexible "local" models. Such models exist in a general regression framework (e.g., generalized additive models), where "local" refers to the values of the predictor values. In a spatial context, local refers to a location. Rather than fitting a single regression model, it is possible to fit several models, one for each (out of possibly very many) location.

3.2 Descriptive spatial statistics

In the second step of the analysis, the intensities of the network are compared with the pertinent information to assess their coherence. The estimated intensities of the point patterns, the hazard maps, and the population distribution are considered raster data sets. Pairs of raster maps are then compared to determine if the spatial information within the raster data is locally correlated. The local correlation coefficient for a pair of raster maps is computed considering a grid resolution of 3 km and a neighborhood of 5×5 cells (15×15 km^2) around each cell. This resolution has been

calibrated according to the case study. In addition, as proposed by Siino et al. (2020), we compute some descriptive statistics to infer further insights about the distribution of the earthquake networks with respect to the related data sets.

4 Results

The results are presented considering separately the *Real-time Seismic Monitoring Network* and the *Strong-Motion Earthquake Observation Network*. Each of the two networks is then analyzed in comparison with the various data set described in the previous paragraph. All the maps resulting from the analysis are projected according to the Mercator projection UTM51 in km, and all the gridded maps have pixel size resolution equal to 3×3 km. Results are displayed over an area of 300 km by 450 km. Of course, being the stations mainly located in the mainland area, here results are more consistent. In the eastern off-shore where seismicity is consistent and where OBSs are placed, results are robust in any case. In the NW corner of the study area, results should be considered not reliable. First, we calculated the inhomogeneous intensities with a kernel estimator as shown in Eq. (2), given the spatial point patterns of the two networks. Having fixed the grid resolution to 3 km, the values reported in the intensity plots indicate the number of stations over a 9 km^2 area.

4.1 Real-time seismic monitoring network

The intensity maps of the real-time seismic monitoring network, briefly the "seismic" network is shown in Fig. 2. The intensity appears to be almost constant over the mainland area: relatively higher intensity values mark the northern portion of the country and part of the western. This seismicity intensity map shows very high values of intensity along a portion of the eastern coast (Fig. 2). The relationship (i.e., local correlation) between intensity maps of seismic stations and instrumental seismicity is evaluated considering the pairs of values around each cell for both maps within a square area of 5×5 cells; the correlation coefficient is then recorded at the central cell of the focal area. Both zones with positive and negative correlation emerge (Fig. 2). The positive correlation indicates coherence between the two considered
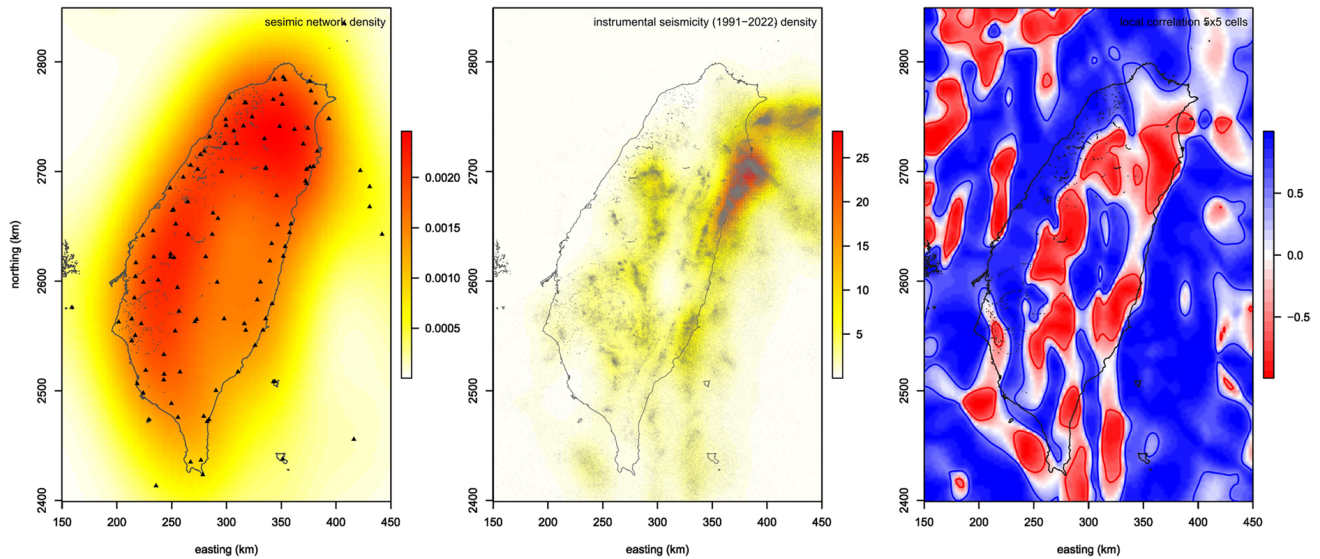
Fig. 2 Left: Kernel intensity estimation for the seismic monitoring network. Center: Kernel intensity estimation for the instrumental seismicity. Right: Local correlation coefficients of the two previous intensity maps. Bandwidth selected with Scott's rule; pixel size resolution is 3×3 km

variables; conversely, the negative correlation indicates discordance between the distributions of the stations and the instrumental seismicity. At this stage, it is not possible to discriminate whether the negative correlations areas are due to the lack of stations in a seismic area or poor seismicity in a well-covered area; however, some considerations will be discussed later.

Furthermore, the cumulative number of events with increasing distance is calculated around each station (Fig. 3). The euclidean distance between each pair of station-epicenter has been considered. We selected the 10% of stations with the highest number of events at a distance of 5 km (green triangles) and the 10% stations with the lowest number of events at a distance of 100 km (red triangles). These threshold values of distance and percentage are taken from Siino et al. (2020) and have the purpose to distinguish which stations are best and worst located with respect to the instrumental seismicity. A map of all the stations ranked according to the cumulative number of events at a distance of 50 km is also shown in Fig. 3.

To further investigate the relationship between the seismic network and the seismicity, we also mapped the magnitude of completeness (M_c) over the study area. M_c is defined as the minimum earthquake magnitude above which the earthquakes are reliably detected. The completeness magnitude has been mapped for the time interval from 1991/01/01 to 2022/12/31 which includes approximately 790,000 events with magnitude ranging from 0 to 7.3. The estimation of M_c usually resorts to the assessment of the for Guternber-Richter (GR) law (Gutenberg and Richter 1944) for the earthquake catalogue. In particular, assuming that the GR law establishes a linear relationship between the magnitude M and the logarithm of the number of earthquakes that have a magnitude greater than M in a given area, it is possible to estimate M_c as the lower threshold of the linear behavior. Based on this assumption, there are many techniques for the calculation of the completeness magnitude (Wiemer and Wyss 2000; Cao and Gao 2002; Wiemer and Wyss 2000; Woessner and Wiemer 2005; Amorese 2007; Taroni 2021; Godano et al. 2022). In this study, M_c has been assessed by computing the maximum curvature of frequency-magnitude distribution (Wiemer and Wyss 2000) using the ZMAP software (Wiemer 2001). To map the distribution of M_c over the area, we fixed a minimum number of 100 events within a constant radius equal to 50 km around each map unit. The M_c map has been finally smoothed with a moving average window (Fig. 4, center). M_c ranges from ~1.8–2.0 in the mainland, decreasing outwards, reaching values greater than ~3.5 in the peripheral areas. This pattern is consistent with

Fig. 3 Top: cumulative
number of instrumental
earthquakes located at
increasing distance from
each seismic station and
corresponding intensity.
Bottom: Top 10% of all
stations with the high-
est cumulative number of
earthquakes at a distance of
5 km (green triangles) and
the top 10% of the stations
with the lowest cumulative
number of earthquakes at
a distance of 100 km (red
triangles); seismic stations
classified according to
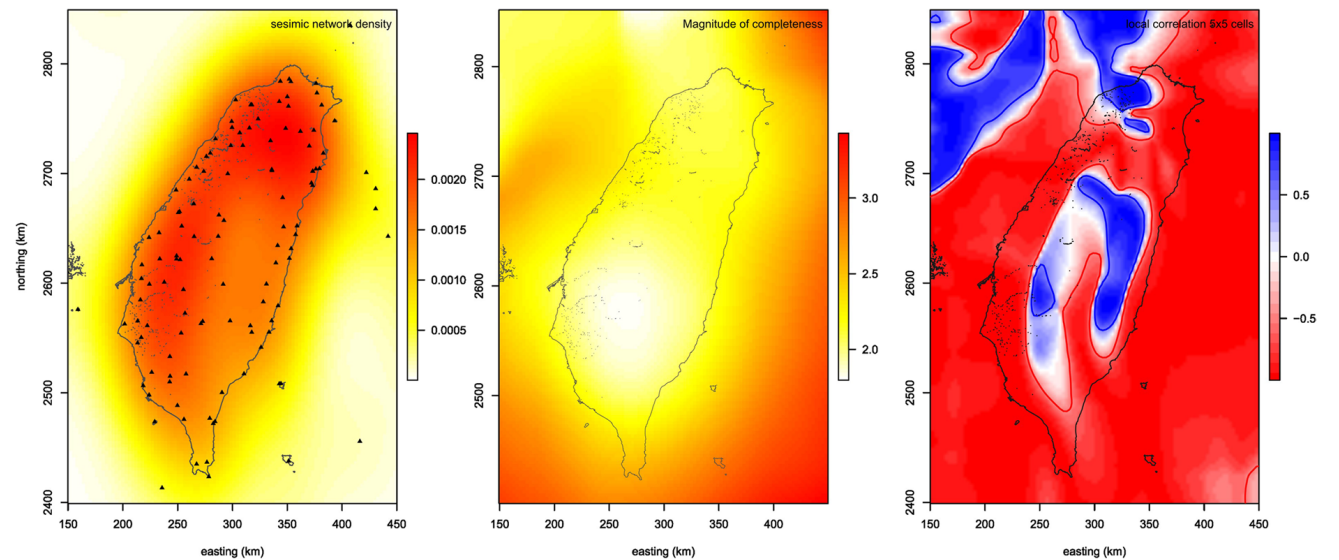the cumulative number of
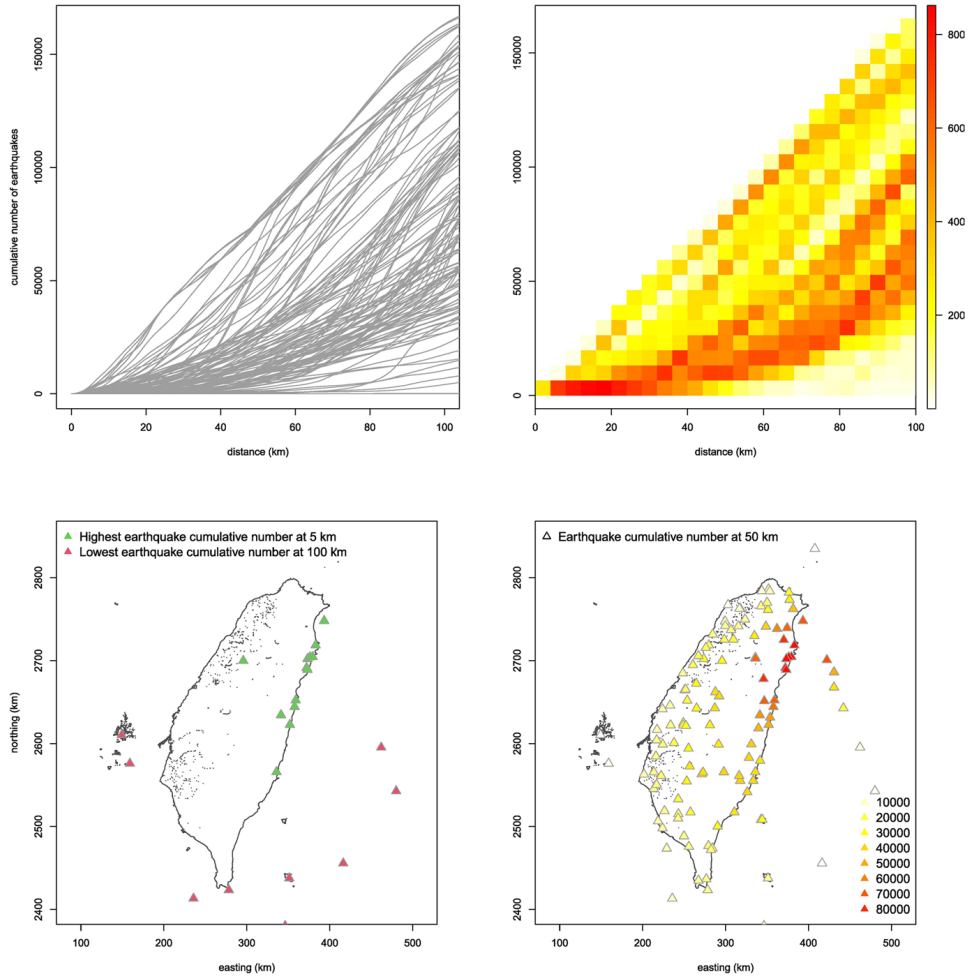earthquakes at a distance of
50 km

Fig. 4 Left: Kernel intensity estimation for the seismic moni-
toring network. Center: Smoothed map of the completeness
magnitude (M_c). Right: Local correlation coefficients of the
two previous intensity maps. Bandwidth selected with Scott's
rule; pixel size resolution is 3×3 km

the M_c map provided by Mignan et al. (2011). The spatial relationship between the nodes of the seismic network and the active faults has been also explored. The distance from each node (i.e., seismic station) to the nearest fault is computed, and vice versa, and the corresponding cumulative density functions are shown in Fig. 5.

Among all the seismic stations, approximately 55% are located within 10 km from a fault, and approximately 90% are less than 25 km away from the nearest fault (Fig. 5). In total, about 70% of the faults are closely monitored by at least one seismic station, and 90% of the faults are monitored from a distance shorter than 10 km. (Fig. 5).

As explained in the previous section, the seismic station intensity is also computed as a function of an external covariate, namely the distance to the nearest fault. The main goal is to assess whether the stations are spatially arranged to occur more frequently near mapped faults. The geological information (i.e., red lines in Fig. 1) is transformed into a spatial variable defined at all locations $\mathbf{u} \in W$, namely, $D(\mathbf{u})$, which is the distance to the nearest active fault (Fig. 6). We expected that, increasing the distance to the nearest fault, the station intensity decreases (i.e., negative coefficient of the geographically weighted regression). However, positive values of the slope coefficient mark large portions of the study area, especially

Fig. 5 Left: cumulative density function of the distance to the nearest fault of each seismic station. Right: cumulative density function of the distance to the nearest seismic station of each fault
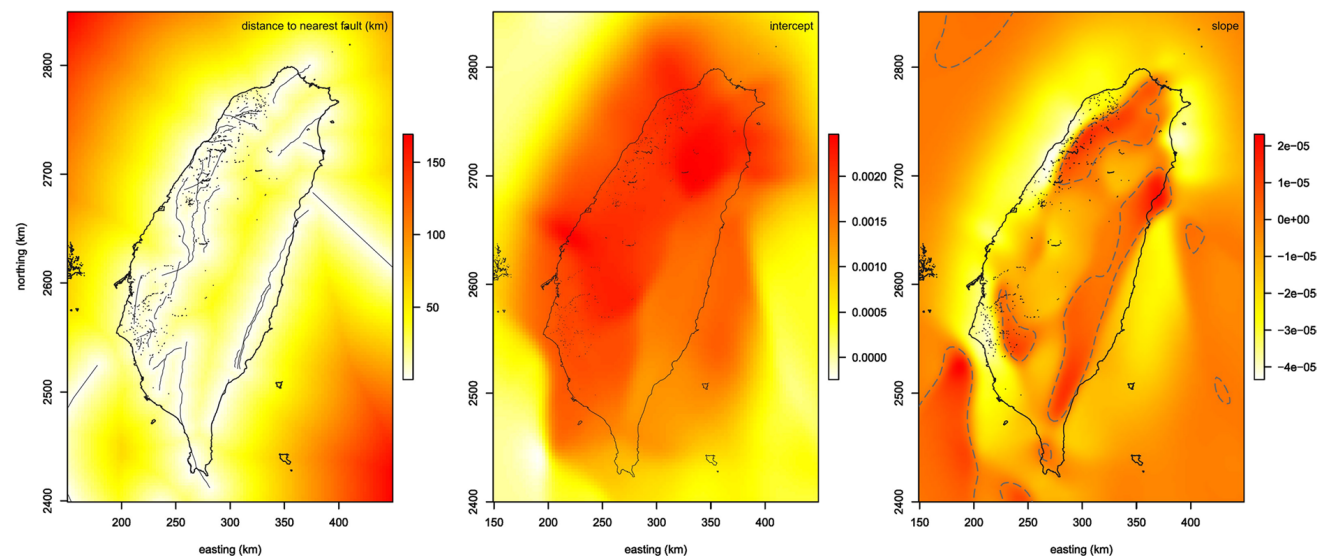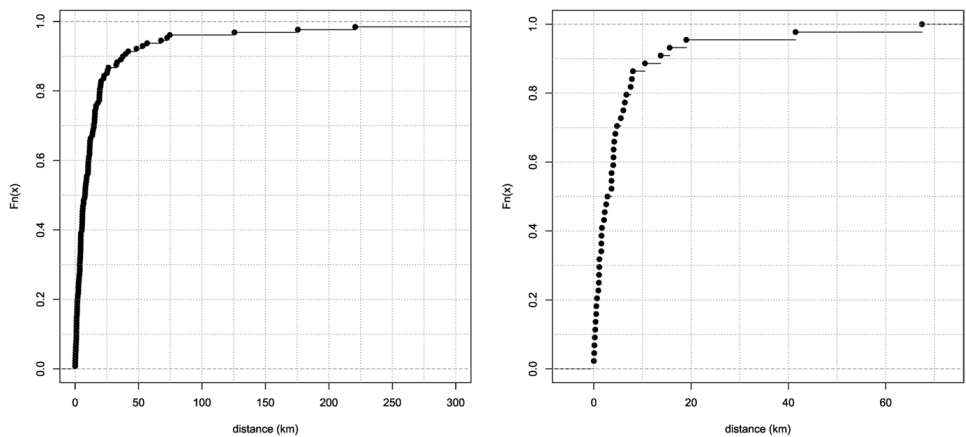

Fig. 6 Left: distances (in km) to the nearest fault for each grid point. Center: spatial distribution of β_0 (i.e., intercept) of the geographically weighed regression model. Right: spatial distribution of β_1 (i.e., slope) of the geographically weighed regression model; contour line (dashed gray line) corresponds to $\beta_1 = 0$

in the mainland, indicating, at first glance, an unexpected increase in the number of stations at increasing distance from the faults' positions.

4.2 Strong-Motion Earthquake Observation Network

The intensity maps of the Strong-Motion Earthquake Observation Network, briefly "strong-motion" network is shown in Fig. 7. The intensity is clearly higher in the western and northern parts of the country, while it is sensibly lower in the inner and southern parts. The intensity map of the $M > 5.5$ earthquakes shows a gradient moving from west to east (Fig. 7). The relationship (i.e., local correlation) between intensity maps of strong-motion stations and potentially damaging seismicity has been calculated considering the pairs of values around each cell in both maps. This focal square area is equal to 5×5 cells, and we recorded the correlation value at the central cell of the focal area (Fig. 2, right). The positive correlation indicates coherence between the two considered variables; conversely, the negative correlation indicates discordance between the distributions of the stations and the historical seismicity. At this stage, it is not possible to discriminate whether the negative correlations areas are due to the lack of stations in a historically seismic area or, on the contrary, to the poor

seismicity in a well-covered area; however, some considerations will be discussed later.

Similarly to the instrumental seismicity, the cumulative seismic moment M_0 with increasing distance is computed around each node of the strong-motion network. The obtained curves provide information about the coherence between the location of each station and the historical seismic release (Fig. 8). Among all the stations, we selected 5% having the highest values of M_0 at 5 km (green triangles) and those having the lowest values at 100 km (red triangles). Again, such thresholds are taken from Siino et al. (2020) and have the purpose to recognize the best and worst–placed strong-motion stations with respect to the historical seismic release. A map of all the stations ranked according to the cumulative M_0 at 50 km is also presented (Fig. 8), providing a comprehensive explanatory framework for the effectiveness of the strong-motion network. The relationships between the spatial distribution of strong motion stations and the hazard maps and population have been also evaluated. In fact, because one of the goals of a strong motion network is to assess the earthquake intensity (i.e., the effects at the surface), this network should be also planned according to the distribution of people to the expected ground motion. The local correlation between the strong-motion station intensity and the
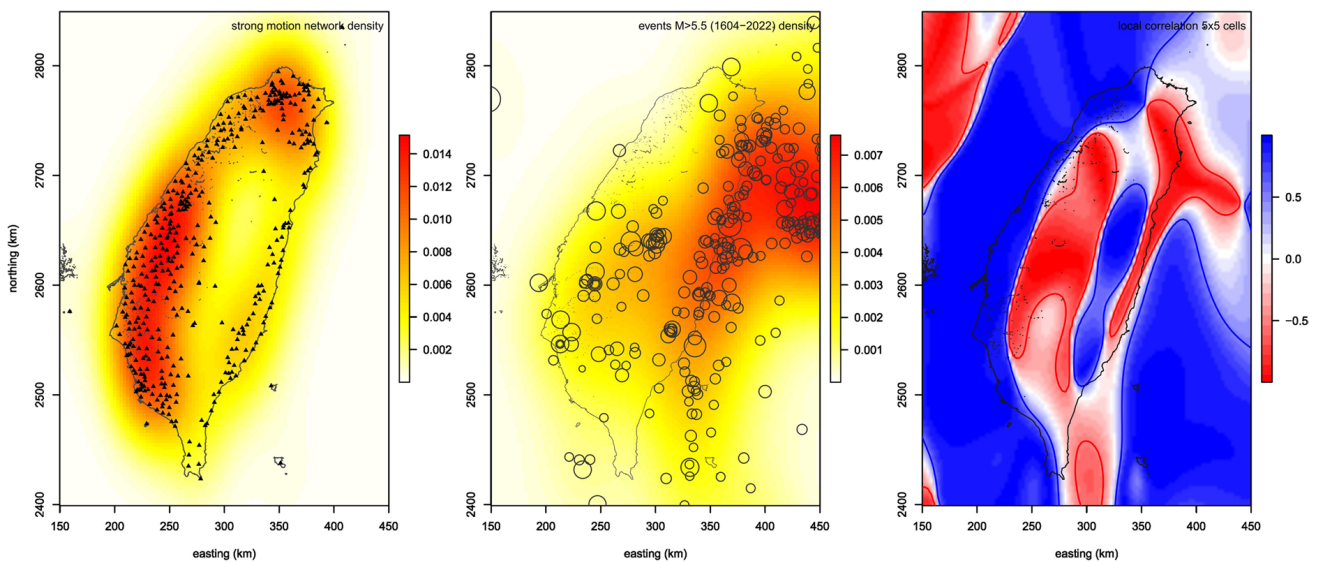

Fig. 7 Left: Kernel intensity estimation for the strong motion network. Center: Kernel intensity estimation for the events with $M >= 5.5$. Right: Local correlation coefficients of the two previous intensity maps. Bandwidth selected with Scott's rule; pixel size resolution is 3×3 km

Fig. 8 Top: cumulative M_0 of events $M >= 5.5$ located at increasing distance from each strong motion station and corresponding intensity. Bottom: Top 5% of all the strong-motion stations with the highest cumulative M_0 at a distance of 5 km (green triangles) and top 5% of the stations with the lowest cumulative M_0 at a distance of 100 km (red triangles); strong-motion stations classified according to the cumulative M_0 at a distance of 50 km
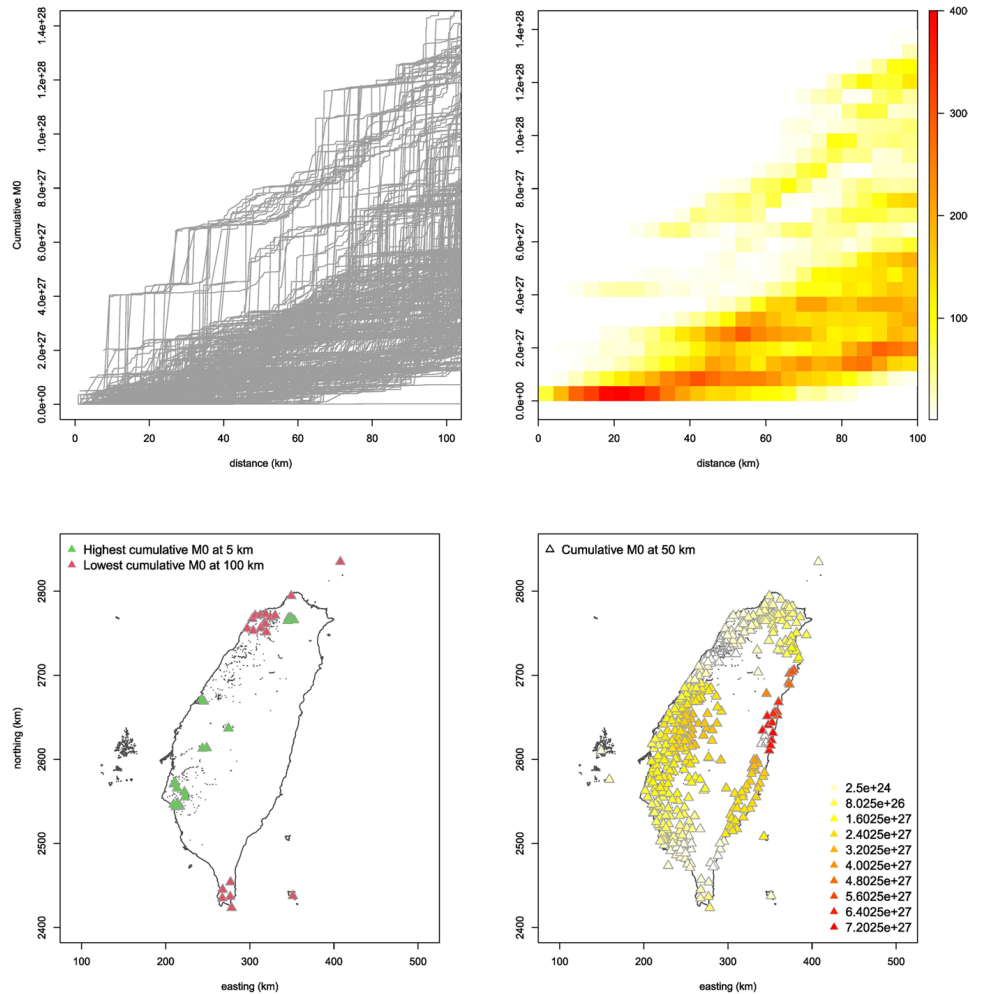
population density (Fig. 9a, b) shows a high spatial variability of the area with positive and negative correlation with short wavelength. Moreover, we have also calculated the cumulative number of people living around each station with increasing distance (Fig. 10). By selecting the 5% of stations with the highest number of people at a distance of 5 km (green triangles) and the 5% stations with the lowest number of people at a distance of 100 km (red triangles), it is possible to distinguish which stations are placed best and worst with respect to the population distribution. A map of all the stations ranked according to the cumulative number of people at 50 km is presented (Fig. 10).

Finally, the local correlation between the strong-motion station intensity and the expected peak ground acceleration in 50 years has been computed (Fig. 9). The results indicate a general positive correlation, whereas small-scale regions of negative correlation occur especially in the northern half of the island.

5 Discussion

The assessment of the quality of earthquake monitoring networks is necessary to recognize both weak and strong points and to address the continuous development. We performed this task for the seismic and strong motion networks in Taiwan following the approach proposed by Siino et al. (2020) which apply point-process analysis together with descriptive spatial statistics.

The characteristics of the seismic network have been assessed taking into account the instrumental seismicity, the M_c, and the active faults. Results of the local correlation between the station intensity and the instrumental seismicity highlight zones with negative correlation (red areas in Fig. 2). Considering the almost uniform station intensity (Fig. 2), most of these negative areas are likely the result of a significant network coverage in zones with limited instrumental seismicity. The positions of the stations

Fig. 9 Top row: population density (displayed as log_{10} of the number of people per map unit) and local correlation with the strong motion network. Bottom row: expected PGA from the TEM model considering 10% exceedance probability in 50 years (Chan et al. 2020) and local correlation with the strong motion network. The pixel size resolution is 3×3 km
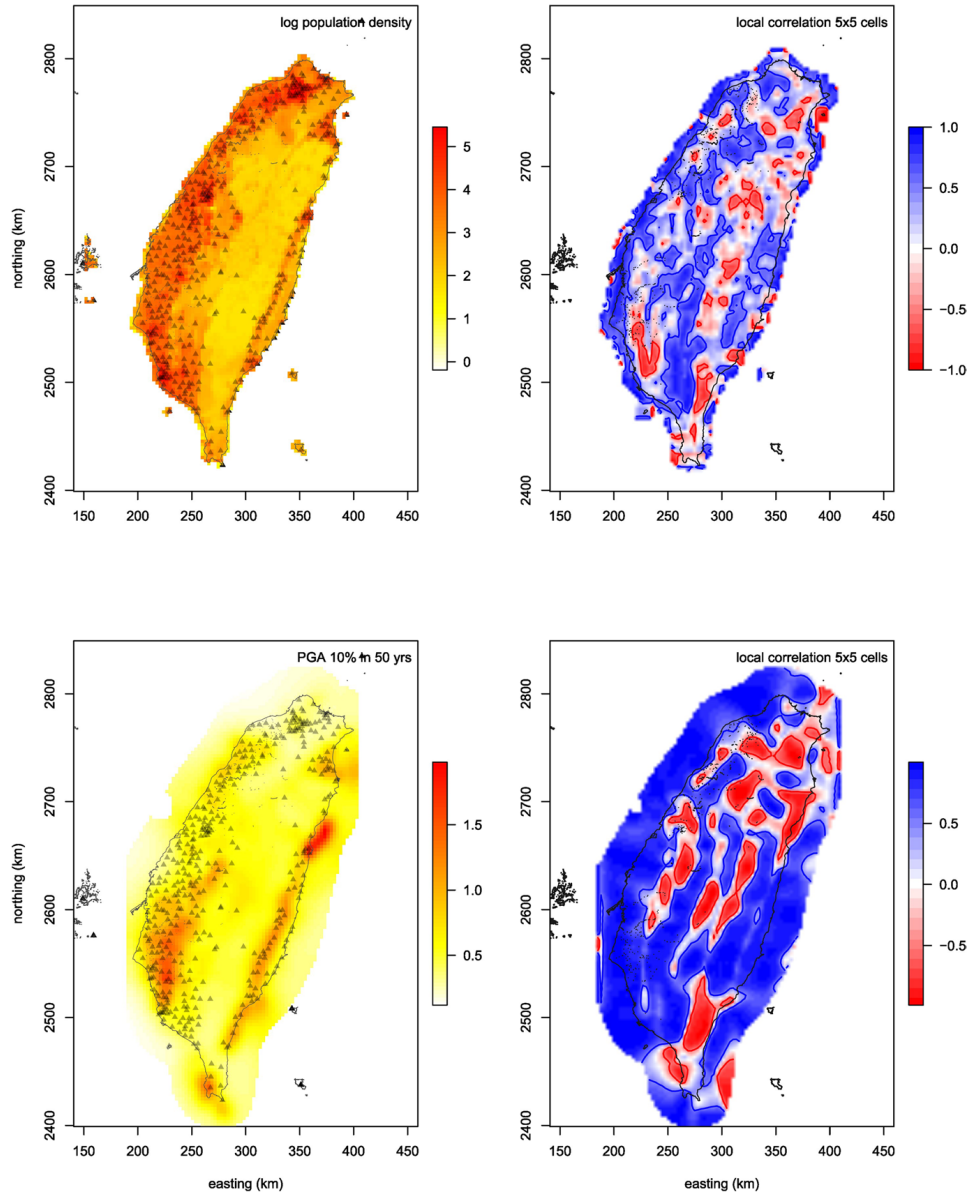

compared with the distributions of the earthquakes indicate that the best-placed stations (i.e., the highest number of events at a distance of 5 km) have a good match with the intensity of instrumental seismicity (Fig. 3). The less coherent stations (i.e., the lowest number of events at a distance of 100 km) are located in the peripheral areas and some of them are the OBS stations. In particular, the latter sensors are essential to better locate and constrain the high seismic area located in the eastern offshore of the main island 3. Considering the cumulative number of earthquake at increasing distance, only a few stations have a low gradient whereas most of them have a notable, even though variable, gradient (Fig. 3). For some of them,

more than 150,000 events have been located at a distance shorter than 100 km.

The correlation between the station intensity and the M_c indicates a general positive correlation in the inner and northernmost parts of the island, and a general negative correlation in the outer ones (Fig. 4). The density of the station and M_c is expected to be uncorrelated because the distance from stations is among the controlling factors in reducing the M_c (c.f. Mignan et al. (2011)). For this reason, the areas with positive correlation would represent anomalies (Fig. 4, right).

We found a good coherence also checking the distribution of the seismic network nodes against the

Fig. 10 Top: cumulative population at increasing distance from each strong-motion station and corresponding intensity. Bottom: Top 5% of all the strong-motion stations with the highest cumulative population at a distance of 5 km (green triangles) and top 5% of the stations with the lowest cumulative population at a distance of 100 km (red triangles); strong-motion stations classified according to the cumulative population at a distance of 50 km
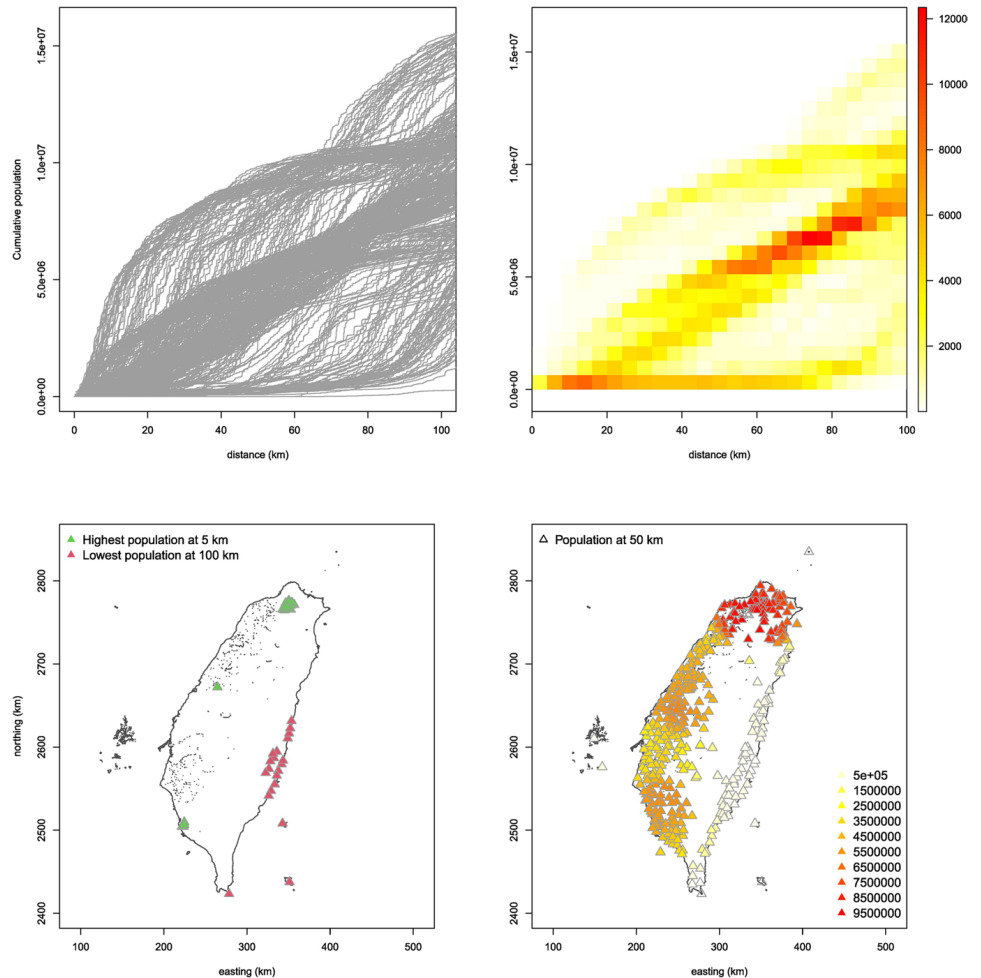
distribution of active faults. This is first confirmed by the cumulative density functions (Fig. 5), while the results of the geographically weighted regression would indicate an apparent disagreement. In fact, in some areas, the number of stations increases at increasing distance from faults (Fig. 6). This can be framed within the almost uniform station distribution (Fig. 2) so that many stations are located in areas where no faults are mapped. In a future perspective, this configuration can be considered a strong point because the fault database cannot be considered definitively complete, and other faults will be mapped in the future also in the areas where they are not mapped at present.

The characteristics of the strong motion network have been assessed taking into account the distribution of the strongest events ($M >= 5.5$) and their associated seismic release (M_0), the population, and the seismic hazard model. The results of the local correlation between the intensities of stations and of the $M >=$ 5.5 events highlight large portions with negative correlation (red zones in Fig. 7, right). In particular, the large red area in the inner part of the island is likely the result of a relative lack of strong motion stations in an area experiencing a relatively abundant strong seismicity; as a counterpart, this area corresponds to a low-populated area (Fig. 8). Conversely, the smaller area along the upper-eastern coast is likely the result of the high density of $M >= 5.5$ earthquakes.

The stations with the higher gradient and greater cumulative M_0 at a 50 km distance (Fig. 8) are placed coherently with the expected values of ground motion (Fig. 9). The worst-placed stations according to the lowest M_0 at a distance of 100 km are exactly located where the expected ground motion is lower and where few people live 9.

The distribution of the population is highly uneven, being characterized by small and highly populated zones, and large areas with few inhabitants (Fig. 9 top-left). The resulting local correlation shows

small-scale variations whose interpretation is not straightforward. The gradient of the cumulative population at increasing distance highlights three well-defined groups of stations (Fig. 10 top) clearly recognizable in the map (Fig. 10 bottom). The best-located strong motion stations with respect to the population are clustered in a very restricted area in correspondence with the capital city (green triangles in Fig. 10), while the worst-placed stations are located along the lower eastern coast and in some islands (red triangles in Fig. 10). The local correlation between the strong motion stations and the seismic hazard model (TEM) results in general positive values (Fig. 9), with the negative values grouped in small portions clustered mainly in the northern part of the study area. This is also the area in which the seismic hazard is generally lower (Fig. 9) and uncorrelated areas are the effect of a relatively greater number of stations.

6 Conclusive remarks

The geometry of a network plays a role in its performances. In particular, earthquake monitoring networks should be validated according to their final objectives in consideration of associated information like the earthquakes' and faults' distributions, as well as seismic hazard model and populations.

The planned design of the geometry of a given network is crucial in its future performance; however, its assessment should be validated afterwards through recorded data because local-scale factors could play a role (e.g., the hypocentral depth and crustal velocity structure) (Scudero et al. 2021). Although these factors have not been included in our analysis, they do not present limitations on the interpretation of the results at a larger scale (Siino et al. 2020). Indeed, this approach represents a very useful tool to assess the degree of coverage of monitoring networks as a function of the spatial distributions of correlated parameters (including complex and irregular distributions), and it has been applied for the characterization of the earthquake monitoring networks in Taiwan.

For both of the Taiwan networks (i.e., seismic and strong motion), the adopted approach revealed useful to assess the present-day quality of the earthquake monitoring and to suggest directions in planning their future optimization. The main considerations achieved with this work can be summarized as follows:

- The shape of the island inevitably influences the geometry of the earthquake networks and this, in turn, clearly influences many results showing NNE-SSE elongated features. However, at a large scale, does not emerge any evidence of critical weak areas for both the seismic and strong-motion network in Taiwan. They both are coherent with their own objectives of seismic surveillance and intensity/hazard assessment, respectively.

- The seismic network extends even over zones apparently not justifiable at present because of the lack of mapped faults. This might turn out in a strong point in a future perspective if further studies will reveal previously unmapped faults, as revealed in the experiment by Wu et al. (2016). The OBSs deployed in the eastern offshore area reveal essential to locate the relevant seismicity of this sector. However, all the stations are essential to better locate and constrain all those earthquakes occurring in the more external parts of the network. The hazardous area along the eastern coast is well covered, and the weaker parts of the networks in the inner part of the island, all things considered, are low seismic zones. Another advantage of such a high-density network is enabling the scientific research like seismic tomographies and other seismological studies with great coverage and resolution (Wu et al. 2007, 2008; Sun et al. 2015; Kuo et al. 2016). Similarly, the strong-motion network well covers zones where the current seismic hazard model indicates moderate hazard, without overlooking the most hazardous zones. Moreover, especially for the eastern coast, there is a relevant simultaneous coherence among all the information: high expected ground motion, low population, and appropriately placed strong motion stations.

- For both networks, only restricted portions of the country show a limited coherence between the networks with the associated information, but not with multiple parameters simultaneously.

were accessed from the "Seismological Centre of the Central Weather Bureau," at https://gdmsn.cwb.gov.tw/network_cwbsn.php and at https://gdmsn.cwb.gov.tw/catalogDownload.php respectively (Central Weather Bureau (CWB, Taiwan) 2012) (last accessed on May 2023). Data about population density have been retrieved from the United Nations population estimate at https://population.un.org/wpp/ (last accessed on March 2021).

Declarations

Ethics approval   This research has been conducted in compliance with ethical standards.

Competing interests   The authors declare no competing interests.

References

Amorese D (2007) Applying a change-point detection method on frequency-magnitude distributions. Bull Seismol Soc Am 97:1742–1749. https://doi.org/10.1785/0120060181

Baddeley A (2018) Spatstat.local: extension to 'spatstat' for local composite likelihood. r package version 3.5-7. https://CRAN.R-project.org/package=spatstat.local. Accessed May 2023

Baddeley A, Rubak E, Turner R (2015) Spatial point patterns: methodology and applications with R. Chapman and Hall/CRC Press, London

Baddeley A, Turner R (2005) Spatstat: an r package for analyzing spatial point patterns. J Stat Softw 12(i06)

Barrier E, Angelier J (1986) Active collision in eastern Taiwan: the coastal range. Tectonophysics 125(1-3):39–72

Cao A, Gao S (2002) Temporal variation of seismic b-values beneath northeastern Japan island arc. Geophys Res Lett 29. https://doi.org/10.1029/2001GL013775

Central Weather Bureau (CWB, Taiwan) (2012) Central weather bureau seismographic network. International Federation of Digital Seismograph Networks. https://www.fdsn.org/networks/detail/T5/, https://doi.org/10.7914/SN/T5. Accessed May 2023

Chan C-H, Ma K-F, Shyu JBH, Lee Y-T, Wang Y-J, Gao J-C, Yen Y-T, Rau R-J (2020) Probabilistic seismic hazard assessment for Taiwan: Tem psha2020. Earthq Spectra 36(1_suppl):137–159

Chang C-H, Yih-Min W, Da-Yi C, Shin T-C, Tai-Lin C, Wen-Yen C (2012) An examination of telemetry delay in the central weather bureau seismic network. TAO. Terr Atmospheric Ocean Sci 23(3):261

Chen K-P, Tsai Y-B (2008) A catalog of Taiwan earthquakes (1900–2006) with homogenized m w magnitudes. Bull Seismol Soc Am 98(1):483–489

Cheng SN, Yeh YT (1989) Catalog of the earthquakes in Taiwan from 1604 to 1988. Institute of Earth Sciences, Academia Sinica, Institute of Earth Scienze, p 255

Chin S-J, Lin J-Y, Yeh Y-C, Kuo-Chen H, Liang C-W (2019) Seismotectonic characteristics of the Taiwan collision-manila subduction transition: the effect of pre-existing structures. J Asian Earth Sci 173:113–120

D'Alessandro A, Badal J, D'Anna G, Papanastassiou D, Baskoutas I, Özel NM (2013) Location performance and detection threshold of the Spanish national seismic network. Pure Appl Geophys 170(11):1859–1880. https://doi.org/10.1007/s00024-012-0625-y

D'Alessandro A, Luzio D, D'Anna G, Mangano G (2011a) Seismic network evaluation through simulation: an application to the Italian national seismic network. Bull Seismol Soc Am 101(3):1213–1232. https://doi.org/10.1785/0120100066

D'Alessandro A, Papanastassiou D, Baskoutas I (2011b) Hellenic unified seismological network: an evaluation of its performance through SNES method. Geophys J Int 185(3):1417–1430

D'Alessandro A et al (2019) Urban seismic networks, structural health and cultural heritage monitoring: the national earthquakes observatory (ingv, italy) experience. Front Built Environ 5:127

D'Alessandro A, Dănet A, Grecu B (2012) Location performance and detection magnitude threshold of the romanian national seismic network. Pure Appl Geophys 169:2149–2164

Fotheringham AS, Charlton ME, Brunsdon C (1998) Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis. Environment and Planning A 30(11):1905–1927

GEBCO Bathymetric Compilation Group (2020) The gebco 2020 grid - a continuous terrain model of the global oceans and land. British Oceanographic Data Centre, National Oceanography Centre, NERC, UK. https://doi.org/10.5285/a29c5465-b138-234d-e053-6c86abc040b9

Gee LS, Leith WS (2011) The global seismographic network. Tech. rep., US Geological Survey

Godano C, Convertito V, Pino NA, Tramelli A (2022) An automated method for mapping independent spatial b values. Earth Space Sci 9(6):e02205. https://doi.org/10.1029/2021EA002205

Guan Z-K, Kuo-Chen H, Sun W-F (2020) Re-calculation of the attenuation functions for local magnitude from the upgraded central weather bureau seismic network in Taiwan. Terr Atmos Ocean Sci 31:479–486

Gutenberg B, Richter CF (1944) Frequency of earthquakes in California*. Bull Seismol Soc Am 34(4):185–188. https://doi.org/10.1785/BSSA0340040185

Hanks TC, Kanamori H (1979) A moment magnitude scale. J Geophys Res Solid Earth 84(B5):2348–2350

Hijmans, RJ (2018) Raster: geographic data analysis and modeling. r package version 2.8-4. https://CRAN.R-project.org/package=raster. Accessed May 2023

Hsiao N-C, Wu Y-M, Shin T-C, Zhao L, Teng T-L (2009) Development of earthquake early warning system in Taiwan. Geophys Res Lett 36(5):L00B02. https://doi.org/10.1029/2008GL036596

Kim K-H, Chen K-C, Wang J-H, Chiu J-M (2010) Seismogenic structures of the 1999 mw 7.6 Chi-Chi, Taiwan, earthquake and its aftershocks. Tectonophysics 489(1-4):119–127

Kuo C-H, Chen C-T, Lin C-M, Wen K-L, Huang J-Y, Chang S-C (2016) S-wave velocity structure and site effect parameters derived from microtremor arrays in the western plain of Taiwan. J Asian Earth Sci 128:27–41

Mezcua J (1995) Fundamentos de la red sísmica deEspaña. En: Mezcua J (ed) Redes sísmicas regionales. Instituto Geográfico Nacional Monografía nº. 11. Madrid, pp 63–86

Michelini A et al (2016) The Italian national seismic network and the earthquake and tsunami monitoring and surveillance systems. Adv Geosci

Mignan A, Werner M, Wiemer S, Chen C-C, Wu Y-M (2011) Bayesian estimation of the spatially varying completeness magnitude of earthquake catalogs. Bull Seismol Soc Am 101(3):1371–1385

Popa M, Radulian M, Ghica D, Neagoe C, Nastase E (2015) Romanian seismic network since 1980 to the present. Springer Proceedings in Physics book series (SPPHY, volume 163, pp 117–131)

R Development Core Team (2005) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org. Accessed May 2023

Rau RJ, Wu FT (1998) Active tectonics of Taiwan orogeny from focal mechanisms of small-to-moderate-sized earthquakes. Terr Atmospheric Ocean Sci 9(4):755–778

Scott DW (1992) Density estimation: theory, practice and visualization. Wiley Series in Probability and Statistics, John Wiley & Sons, Inc. https://doi.org/10.1002/9780470316849

Scudero S, Marcocci C, D'Alessandro A (2021) Insights on the Italian seismic network from location uncertainties. J Seismol 25:1061–1076

Shin T-C, Chang C-H, Pu H-C, Hsiao-Wei L, Leu P-L (2013) The geophysical database management system in Taiwan. Terr Atmospheric Ocean Sci 24(1):11

Shin T-C, Teng TL (2001) An overview of the 1999 Chi-Chi, Taiwan, earthquake. Bull Seismol Soc Am 91(5):895–913

Shyu JBH, Chuang Y-R, Chen Y-L, Lee Y-R, Cheng C-T (2016) A new on-land seismogenic structure source database from the Taiwan earthquake model (tem) project for seismic hazard analysis of Taiwan. Terr Atmospheric Ocean Sci 27(3):311–323. https://doi.org/10.3319/TAO.2015.11.27.02

Shyu JBH, Sieh K, Chen Y-G, Liu C-S (2005) Neotectonic architecture of Taiwan and its implications for future large earthquakes. J Geophys Res Solid Earth 110(B8):B08402. https://doi.org/10.1029/2004JB003251

Sibuet J-C, Hsu S-K (1997) Geodynamics of the Taiwan arc-arc collision. Tectonophysics 274(1-3):221–251

Siino M, Scudero S, Greco L, D'Alessandro A (2020) Spatial analysis for an evaluation of monitoring networks: examples from the Italian seismic and accelerometric network. J Seismol 24(6):1045–1061. https://doi.org/10.1007/s10950-020-09937-0

Styron R, Pagani M (2020) The gem global active faults database. Earthq Spectra 36(1_suppl):160–180

Sun W-F, Peng Z, Lin C-H, Chao K (2015) Detecting deep tectonic tremor in Taiwan with a dense array. Bull Seismol Soc Am 105(3):1349–1358

Taroni M (2021) Back to the future: Old methods for new estimation and test of the Gutenberg-Richter b-value for catalogues with variable completeness. Geophys J Int 224(1):337–339. https://doi.org/10.1093/gji/ggaa464

Theunissen T, Font Y, Lallemand S, Liang W-T (2010) The largest instrumentally recorded earthquake in Taiwan: revised location and magnitude, and tectonic significance of the 1920 event. Geophys J Int 183(3):1119–1133

Tsai Y-B, Lee C-P (2005) Strong motion instrumentation programs in Taiwan. Directions in strong motion instrumentation. Springer, pp 255–278

Tsai Y-B, Wu H-H (1997) A study of the errors in locating earthquakes due to the geometry of the Taiwan seismic network. Terr Atmospheric Ocean Sci 8(3):355–370

Wang JH (1989) The Taiwan telemetered seismographic network. Phys Earth Planet In 58(1):9–18

Wiemer S (2001) Sa software package to analyze seismicity: Zmap. Seismol Res Lett 72:373–382

Wiemer S, Wyss M (2000) Minimum magnitude of completeness in earthquake catalogs: examples from Alaska, the Western United States, and Japan. Bull Seismol Soc Am 90(4):859–869. https://doi.org/10.1785/0119990114

Woessner J, Wiemer S (2005) Assessing the quality of earthquake catalogues: estimating the magnitude of completeness and its uncertainty. Bull Seismol Soc Am 95(2):684–698

Wu FT, Ross ZE, Okaya D, Ben-Zion Y, Wang C-Y, Kuo-Chen H, Liang W-T (2016) Dense network, intense seismicity and tectonics of Taiwan. Tectonophysics 692(B):152–163. https://doi.org/10.1016/j.tecto.2016.04.025

Wu Y-H, Chen C-C, Turcotte DL, Rundle JB (2013) Quantifying the seismicity on Taiwan. Geophys J Int 194(1):465–469

Wu Y-M, Chang C-H, Zhao L, Shyu JBH, Chen Y-G, Sieh K, Avouac J-P (2007) Seismic tomography of Taiwan: improved constraints from a dense network of strong motion stations. J Geophys Res Solid Earth 112(B8)

Wu Y-M, Chang C-H, Zhao L, Teng T-L, Nakamura M (2008) A comprehensive relocation of earthquakes in Taiwan from 1991 to 2005. Bull Seismol Soc Am 98(3):1471–1481