

Earthquake Analysis of The Hindukush Range by Using Statistical Methods



By

Aiman Zaka

(Registration No: 00000365130)

Department of Statistics

School of Natural Sciences

National University of Sciences and Technology (NUST)

Islamabad, Pakistan

(2024)

Earthquake Analysis of The Hindukush Range by Using Statistical Methods



By

Aiman Zaka

(Registration No: 00000365130)

A thesis submitted to the National University of Sciences and Technology, Islamabad,

in partial fulfillment of the requirements for the degree of

Masters of Science in

Statistics

Supervisor: Dr.Tahir Mehmood

School of Natural Sciences

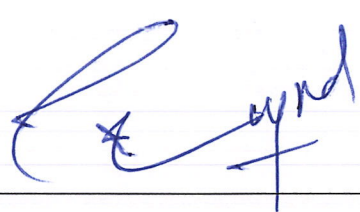
National University of Sciences and Technology (NUST)

Islamabad, Pakistan

(2024)

THESIS ACCEPTANCE CERTIFICATE

Certified that final copy of MS thesis written by Aiman Zaka (Registration No. 00000365130), of School of Natural Sciences has been vetted by undersigned, found complete in all respects as per NUST statutes/regulations, is free of plagiarism, errors, and mistakes and is accepted as partial fulfillment for award of MS/M.Phil degree. It is further certified that necessary amendments as pointed out by GEC members and external examiner of the scholar have also been incorporated in the said thesis.

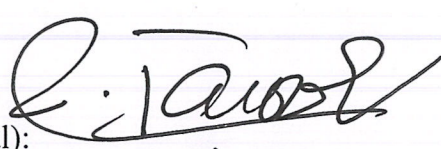
Signature:  _____

Name of Supervisor: Prof. Tahir Mehmood

Date: 05/09/2024

Signature (HoD):  _____

Date: 5/9/2024

Signature (Dean/Principal):  _____

Date: 05.09.2024

National University of Sciences & Technology

MS THESIS WORK


We hereby recommend that the dissertation prepared under our supervision by: **“Aiman Zaka”** Regn No. **00000365130** Titled: **“Earthquake Analysis of The Hindukush Range by Using Statistical Methods”** accepted in partial fulfillment of the requirements for the award of **MS** degree.

Examination Committee Members

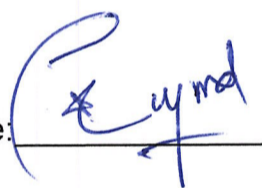
1. Name: DR. FIRDOS KHAN


Signature: 

2. Name: DR. ZAMIR HUSSAIN

Signature: 

Supervisor's Name: PROF. TAHIR MAHMOOD

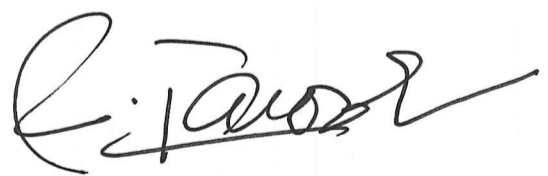
Signature: 


Head of Department

5/9/2024
Date

COUNTERSIGNED

Date: 05.09.2024


Dean/Principal

Dedication

I dedicate this work to my beloved parents, whose unwavering support, endless sacrifices, and constant encouragement have been the foundation of all my achievements

Acknowledgments

Glory be to Allah (S.W.A), the Creator, the Sustainer of the Universe, whose boundless blessings and guidance have led me through every step of this journey. I would like to express my deepest gratitude to my Supervisor, Dr. Tahir Mehmood, whose guidance, encouragement, and invaluable support have been instrumental in the completion of this thesis. His expertise and patience have profoundly shaped my research journey. I also owe a heartfelt thanks to my parents, whose endless love, sacrifices, and unwavering belief in me have provided the foundation for all my achievements. Without their constant support and prayers, this accomplishment would not have been possible.

Contents

1	Introduction	1
1.1	Hindukush Region Tectonics	1
1.2	The Hindukush Region: A Seismic Hotspot	2
1.3	Dependent and Independent Earthquakes	3
1.4	Seismicity Declustering	3
1.5	Declustering Techniques	4
1.5.1	Deterministic Models	4
1.5.2	Stochastic Model	5
1.6	Time and Distance Window	6
1.7	Seismic Database	7
1.8	Properties of the Earthquake Catalog	7
1.8.1	Relationship between Magnitude and Depth	7
1.8.2	Correlation of Magnitude, Depth, and Period Duration	8
1.8.3	Cumulative Frequency of earthquakes	9
1.9	Global Research on b-value	9
1.10	Problem Statement	11
1.11	Aim & Objective	11
1.12	Research Question	11
1.13	Covered Topics	11
2	Literature Review	12

3	Materials and Methods	16
3.1	Data Acquisition	16
3.2	Window-Based Declustering Method	17
3.2.1	Gardner and Knopoff Method	18
3.3	Gutenberg-Richter Relationship:Linking Earthquake Frequency and Magnitude	18
3.4	Magnitude of Completeness (Mc)	19
3.5	The ETAS Model	19
3.5.1	Marked-Point Temporal Processes	19
3.6	The Space-Time ETAS Model	19
3.6.1	Total Spatial Intensity	21
3.6.2	Clustering (Triggering) Coefficient	22
3.6.3	Stochastic Declustering	22
3.6.4	Parameter Estimation and Model Fitting	22
4	Results and Discussions	24
5	Conclusions	44
6	Limitations	46

List of Tables

1.1	Representation of Earthquake Catalog	7
4.1	Earthquake Events categorized by Depth Distribution. This table classifies the overall number of earthquake occurrences into two depth ranges: shallow earthquakes (0-70 km) and intermediate earthquakes (70-280 km), with 6274 and 3258 occurrences respectively.	24
4.2	Aftershock identification windows	29
4.3	ML estimates of model parameters are presented. β equals to 1.4878 indicates a rapid decay in the occurrence of larger magnitude target events in dataset. A μ background seismicity rate of 1.01 suggests steady background earthquake occurrences. A small earthquake (at threshold magnitude) might trigger roughly 0.0912 aftershocks. α equals to 1.7511 indicates that for each unit increase in earthquake magnitude, aftershocks increase by approximately 5.75 times. c and p are temporal decay parameters. With $c = 10.4336$ and $p = 1.1559$ the model suggests that aftershocks will keep happening for a while after the main shock, but they'll start to become less frequent at a moderate pace. D , γ and q are spatial decay parameters. With $D = 0.7144$, aftershocks are relatively close to the main event. With $\gamma=5.5429$, even a small increase in the magnitude will significantly expand the area where aftershocks might happen. With $q = 6.9006$, the probability of aftershocks decreases rapidly as move away from main shock.	38
4.4	Declustering probabilities from the ETAS model fit indicates that the 75.94% earthquakes are classified as background events.	39

List of Figures

1.1	Algorithmic representation of stochastic declustering depicts the random classification of earthquake clusters, showing the computation of ϕ_j and ρ_{ij} for every event j , creating a random variable U_j , and deciding the classification of event j depending on the total impact of previous events. This procedure determines if an event is part of the background or a child event in the cluster.	6
1.2	Classification of earthquakes based on their focal depth: shallow (0-70 km), intermediate (70-300 km), and deep (300-700 km).	8
3.1	The visual representation of all seismic events in Hindukush and the surrounding area of different magnitudes, from 2008 to 2010.	17
4.1	The plot represents the correlation between magnitude and depth of earthquake catalog. Shallow depth are represented in blue dots and Intermediate depth are represented in grey dots.	25
4.2	(a) Relationship between seismic events and earthquake magnitudes, (b) Relationship between Focus depth and seismic events, (c) Relationship between number of seismic events and years, (d) Relationship between focal depth, magnitude, and year of occurrence of seismic events represented by a 3D graph. . .	27
4.3	The cumulative number of earthquakes is shown in relation to the time period (year).	28
4.4	By using the window approach, the remaining events (mainshocks) are shown as deep sky blue-colored triangles. The total number of events are shown in grey-colored points.	30

4.5	Plot of the cumulative number of earthquake events (on a logarithmic scale) against magnitude with “ <i>a</i> ” value of 5.21 and “ <i>b</i> ” value of 0.98, obtaining magnitude of completeness 3.03 through maximum curvature method.	31
4.6	The regional and temporal mapping of the <i>a</i> and <i>b</i> -values for the rate of change in seismicity and region’s slope respectively.	33
4.7	The epicenters’ locations in the top left panel, the logarithm of frequency by magnitude in the bottom left panel, the cumulative frequency over time in the bottom middle panel, and the latitude, longitude, and magnitude against time in the right panels of 9532 earthquakes with a magnitude greater than or equal to 4 occurred between 2008-08-01 to 2010-01-01 in the Hindukush region and its surrounding area (36°–37°N and 70°–72°E).	35
4.8	Plots of the fitted ETAS model’s estimates of the background seismicity rate $\mu(x,y)$, total spatial intensity $\Lambda(x,y)$, clustering or triggering coefficient $\omega(x,y)$, and conditional intensity function $\lambda_{\theta}(x,y,t H_t)$ to the Hindukush Catalog	41
4.9	The Q-Q plot compares the observed quantiles of earthquake occurrences with a hypothesized uniform distribution ranging from 0 to 1. Points that closely align with the red line indicate a strong general fit of the model, while deviations reflect minor deviations in prediction accuracy.	42
4.10	A Q-Q plot is generated for the transformed times using the ETAS model. It compares the ordered transformed times to the expected quantiles of a standard uniform distribution. The close alignment between the black line and the red 45-degree line suggests that the model fits the earthquake data well.	43

LIST OF ABBREVIATIONS

ETAS	Epidemic-Type Aftershock Sequence
M_c	Magnitude of Completeness
PDF	Probability Density Function
MLE	Maximum Likelihood Estimation
CGS	California Geological Survey
ISC	International Seismological Centre
USGS	United States Geological Survey

Abstract

The Hindukush region is renowned for its high seismic activity, making it a crucial area for earthquake analysis. This research examines the earthquake data from 2008 to 2010 utilizing sophisticated statistical techniques. The earthquake catalog contains both dependent and independent events, required declustering to eliminate dependent events. The study focuses declustering methods with the Gardner-Knopoff approach to identify mainshocks. The declustering found 3164 clusters of earthquakes with a total of 3574 (37.5 %) events out of 9532 with only the mainshock shown on a color-coded seismicity map. The Gutenberg-Richter relationship was used to assess the distribution of magnitudes, indicating a completeness magnitude of $M_c = 3.03$. Temporal and spatial analyses revealed a high productivity rate “ a ” of 5.21, indicating significant seismic activity, and “ b ” value of 0.98, reflecting notable higher frequency of smaller earthquakes. Variations in b -values were linked to the region’s distribution of earthquake magnitudes, while the high “ a ” value was due to elevated seismic activity. Moreover, the research employed the ETAS model to distinguish between background and triggered seismic activity in a defined area, identifying 146 target events and classifying 75% as background events. Using an iterative stochastic declustering method, the ETAS model yielded Maximum Likelihood estimates for important parameters, emphasizing the connection between mainshocks and aftershocks. The plot of estimates of ETAS model and spatial and temporal plots were utilized providing valuable information for assessing seismic hazard and evaluate the model’s accuracy. This research provides crucial insights on earthquake risks in the Hindukush area, enhancing earthquake assessment and risk analysis.

CHAPTER 1

Introduction

Among all the natural calamities, earthquakes are the most powerful. These seismic occurrences have the potential to cause massive destruction and fatalities because to their sheer magnitude. Reducing the effects of earthquakes requires an understanding of their dynamics. Accurate forecasting of earthquakes is crucial in order to implement preventive measures and mitigate potential damage, particularly in terms of human casualties [42]. Because of its high seismic activity and potential for catastrophic occurrences, the Hindukush region in particular has been a focus of seismic research.

1.1 Hindukush Region Tectonics

The Hindukush region is very prone to seismic activity, particularly for earthquakes that occur at depths between 70 and 300 kilometers [27]. The geological formation of the region is a consequence of the collision between the Eurasian and Indian plates during the Eocene epoch [16]. Several models have been suggested regarding the earthquake activity pattern in the area. These models are primarily classified into two primary groups [28]. Earlier studies [9] proposed a concept suggesting that an isolated, heavily twisted slab causes Pamir-Hindukush earthquake events. The model shows two subducting plates: the eastern Indian plate is migrating northward and subducting under the Hindukush region, while the Asian lithosphere under the Pamir region is also subducting [11].

1.2 The Hindukush Region: A Seismic Hotspot

This Region has been the focus of much research on earthquakes because of its high level seismic activity, and potential for large-scale, catastrophic calamities. It is, hence, one of the most seismically active regions globally. There has been ongoing discussion on the seismic activity of this area. Prior research refers to this area as the affected 'Benioff zone' [9] due to the occurrence of seismic events in a deep region that is 30 kilometers wide and has a slab-like structure, resembling oceanic subduction. According to the geological analysis of Pamir, Hindukush, and the adjacent areas, it is indicated that there is an absence of oceanic crust throughout entire region, including Karakoram and Tibet [23]. The subducting slab undergoes significant deformation between longitudes 71°E and 73°E gradually weakens deeply around 250 km. This region is likely experiencing ongoing tectonic activity, resulting in frequent seismic events. This makes it an unique seismic area globally. The region of hindukush is characterized by earthquake activity at intermediate depths, although it also has deep seismic events, including those occurring at depths greater than 300 km. Researchers have presented various conflicting explanations for the deep seismic activity in the area. Nevertheless, this deep seismicity remains interesting in that particular region, for which no acceptable or trustworthy theory can now explain [32].

Gaining insight into the earthquake occurrence in the area of Hindukush it is essential for understanding the larger tectonic movements occurring in Central Asia. It is also important to reduce effects of seismicity on the local and nearby populations. Peak ground acceleration estimation is thus an important task in geophysics, seismology, and civil engineering whenever seismic hazard assessment is required. The intensity of the movement of ground can be used to estimate crustal deformation [45]. It is essential to comprehend and analyze these seismic events for evaluating the region's seismic risks and enhance preparedness for earthquakes. A collection of related seismic occurrences grouped together according to their magnitude within a given time frame is known as an earthquake cluster [47]. The stress-strain dynamics in the neighboring fault systems are the driving forces behind the formation of these seismic clusters. The main directory was a thorough earthquake catalogue, which included extremely helpful data like waveform duration and epicentral positions [38]. Regardless of the intensity of the subsequent earthquake, this event leads to a surface rupture that serves as its source. Every earthquake event includes time, depth, magnitude, and epicenter location on a time scale data. It is vital to examine the ground motion and its corresponding factors to understand the previous seismicity and its features [49]. The seismicity shows the process of an earthquake occurring due to sev-

eral geological and environmental events within a specific time and space window [56]. Active tectonic processes cause the relatively high seismicity rate. Numerous seismic occurrences of varying magnitudes have been noticed in northern regions of HinduKush, which is covered by the Pakistan-Afghan border [58].

1.3 Dependent and Independent Earthquakes

There exists a variety of words to describe these two classes. Independent earthquakes are alternatively referred to as background earthquakes, mainshocks, or parent earthquakes. On the other hand, earthquakes that are dependent on the mainshocks are also known as aftershocks, foreshocks, triggered earthquakes, or offspring. Identifying background earthquakes is crucial in seismology for several purposes such as assessing seismic danger, developing models for clustered seismic activity, researching earthquake prediction, and estimating changes in seismicity rate. Nevertheless, this problem is ambiguous as it lacks a distinct solution.[46]. In a recent study, two distinct declustering methods were used on the Italian Instrumental Catalogue analyzed how declustering affected the calculation of b-values and the assessment of seismic danger. There was an observation that the declustered catalogue yielded a lower hazard than the full catalogue [62]. Several research have demonstrated that excluding the aftershocks or foreshocks from the major catalogues results in underestimating occurrences per year and, subsequently, associating with no significant earthquake risk. Plate tectonic movements-related permanent structural stresses are the primary cause of background or independent earthquakes, also known as the parent earthquake or the mainshock. Dependent earthquakes include foreshocks, aftershocks, and triggered earthquakes resulting from temporary stress shifts either before or after mainshock-related fault activity. There may be duplicate occurrences, aftershocks, major shocks, and foreshocks in the earthquake data repository. Declustering is used to correct this catalogue and extract the mainshock without any duplicates [22]. Terminating duplicate or dependent events in the earthquake catalogue is a challenging procedure [10][44].

1.4 Seismicity Declustering

The primary objective of declustering is to separate the category of independent earthquakes, that are not influenced by any previous earthquakes. Independent earthquakes are often explained by long-term tectonic forces or by transient changes in stress, neither of which for

seismic swarms appears to be induced by previous earthquakes. The seismic events that are dependent caused by stress changes, fluid movements activated by seismic activity, after-slip, and other mechanical processes that are influenced to some extent by prior earthquakes. Seismicity declustering refers to the process of categorizing earthquakes into two distinct classifications[46]. It the process of eliminating the shocks that are dependent from the seismic data, is crucial in seismicity research [30] [46]. The overall aim here is the attainment of a pure background seismicity rate that excludes dependent earthquakes, leaving only earthquakes independent of time and space. To estimate the annual seismicity rate in any given location by reviewing the earthquake catalogue and then using the frequency-magnitude relation of the Richter law with the specific a and b values for that area [1]. Declustering can impact b-value estimation for the same seismotectonic zone [63].

1.5 Declustering Techniques

Various methods for declustering earthquake catalogues have been offered by different researchers.

1.5.1 Deterministic Models

Previously, the majority of users have been utilizing modified versions of the techniques suggested by Gardner and Knopoff (1974) or Reasenberg (1985) mostly due to their easy accessibility and relative simplicity in implementation. Gardner and Knopoff method, developed in the 1970s, established fundamental methodologies for differentiating aftershocks from main shocks. These methods utilized statistical criteria to assess the temporal and spatial proximity between events [5][8]. Researchers [8] devised a method to detect aftershocks in seismicity catalogs by analyzing the time and spatial intervals between events. In addition, they furnished precise measurements of space-time distances based on the size of the mainshock in order to detect aftershocks. This technique is referred to as a window method and is considered one of the most straightforward approaches to identifying aftershocks. They ignored subsequent and more significant aftershocks (known as aftershocks of aftershocks). If an earthquake C occurred within the time window that could cause two background events A and B , only the greater shock was designated the mainshock of C . The proximity of C to other shock, regardless of how close it may be in space and time, was not taken into account. The Reasenberg technique in 1985 proposed the use of space-time windows to decluster seismic events, whereas in

1986 the Uhrhammer approach focused on analyzing regional seismic activity using clustering algorithms [17][18]. In 2001 Wiemer utilized sophisticated statistical methods and geographic clustering to adaptively modify declustering criteria [29]. Grunthal in 2009 further improved upon these techniques by taking into account the seismic activity rates in specific areas [39]. An earthquake catalogue that is dependable is essential for the analysis of seismic trends and the enhancement of hazard assessments.

1.5.2 Stochastic Model

The probabilistic separation of the background component and clustering component was introduced as an alternative to the deterministic declustering approaches [30] [21]. Another seismicity declustering approach was introduced by Zhuang et al. [30], that makes use of a random model with origins in a process that branches across space and time. This model explains how every earthquake leads to additional smaller earthquakes. There are two ways in which this method improves upon and builds upon earlier approaches:

1. Because of the constraints imposed by the ETAS model, the chosen space-time distance is fine-tuned so that it accurately depicts the earthquake data. Thus, even if the exact formula for the space-time distance is known in advance, it is not necessary to simply approximate values for the factors that impact it. Constant monitoring and waiting around during the optimization process is a potential downside.
2. As an alternative to attributing each aftershock to a single mainshock, this technique takes into account all past earthquakes as potential mainshocks and determines the probability that each earthquake is an aftershock of any prior earthquake. This approach is more complex than the previous one. It assigns both A and B as primary earthquakes of C with identical possibilities when the space-time distance between them is about the same, instead of only choosing one.

This reflects the challenge of making a definitive decision in such a scenario. A complete catalogue is essential for examining earthquake activity rates and studying the occurrence and reoccurrence through time series data [43]. The Epidemic-Type Aftershock Sequence (ETAS) model is widely employed in earthquake science to simulate aftershock sequences. This method improves traditional declustering methods by incorporating the impact of past earthquakes on future seismic events [20]. The subjective selection of parameter values for window sizes or link

distances in the windowing and link-based declustering algorithms. Varying parameter values yield distinct declustered catalogs and diverse estimations of background seismicity. Typically, these criteria are selected depending on the researchers' expertise in handling the particular data. Alternatively, in some instances, the determination can be made by employing a method of trial and error, which considers both the declustering outcomes and the temporal coherence of the declustered catalog.

Choosing the right algorithm is crucial for creating an earthquake catalogue usable for declustering. In this dissertation Gardner and Knoff method and ETAS model is used for declustering and seismic analysis. Therefore, all declustering approaches must depend on a conceptual model that defines the characteristics of a mainshock. The distinguishing factor among declustering approaches is in their underlying model, which is also the reason why seismologists find their comparison intriguing. Since all methods rely on models to some extent, no universally superior way that can be determined beforehand is there. [46].



Figure 1.1: Algorithmic representation of stochastic declustering depicts the random classification of earthquake clusters, showing the computation of ϕ_j and ρ_{ij} for every event j , creating a random variable U_j , and deciding the classification of event j depending on the total impact of previous events. This procedure determines if an event is part of the background or a child event in the cluster.

[46]

1.6 Time and Distance Window

Windowing techniques provide a straightforward method for distinguishing mainshocks from aftershocks. An earthquake is classed as an aftershock in the database if its magnitude, denoted as M , falls within the time interval $T(M)$ and distance interval $L(M)$. Foreshocks are considered and handled similarly to aftershocks. This indicates that if the largest earthquake happens after other smaller earthquakes, the foreshock is considered a type of aftershock. Therefore, the time-space frames are modified according to the size of the biggest earthquake in a series. The time periods during which aftershocks are identified can differ significantly between studies and are typically not determined through an optimization process.

1.7 Seismic Database

A seismic database listed in chronological order is compilation of historical or instrumental recordings of earthquakes within a given geographical region S and during a defined time frame. The usual information consists of the time, epicenter coordinates, magnitude, and focal depth of earthquakes with magnitudes equal to or above a set threshold (m_0) that happened in or around area S during the given time frame. An earthquake catalog, from a statistical perspective, refers to the collection of data about earthquakes within a specific space-time window [25]. Subsequently, statistical analysis can be employed to identify an appropriate model for the fundamental earthquake mechanism. This model provides a conceptual comprehension of seismic activities in the examined area and enables the anticipation of the likelihood of future events [22] [31]. Therefore, it is essential to choose an appropriate statistical model for a certain earthquake database in order to accurately evaluate the probability of seismic risk. In order to analyze the seismic risks of a certain region, it is necessary to have an earthquake catalog specifically for that location. The earthquake catalog include both dependent earthquakes, which rely on each other in terms of time and space, and independent or background earthquakes [47]. The Table 1.1 displays the earthquake catalog's portrayal containing columns for seismic events, date, time, longitude, latitude, magnitude, and depth.

Seismic Event	Date	Time	Longitude	Latitude	Magnitude	Depth
1	D_1	t_1	x_1	y_1	m_1	d_1
2	D_2	t_2	x_2	y_2	m_2	d_2
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
N	D_N	t_N	x_N	y_N	m_N	d_N

Table 1.1: Representation of Earthquake Catalog

1.8 Properties of the Earthquake Catalog

1.8.1 Relationship between Magnitude and Depth

A crucial problem examined in earthquake catalogs is the correlation between earthquake magnitude and depth as it helps researchers in understanding the characteristics of seismic events. Large seismic occurrences typically occur at shallow depths inside the Earth's crust, but minor

earthquakes may occur at various depths, up to approximately 700 km (400 mi). Earthquakes are seismic events that take place in the crust of earth in the uppermost layer, with an average thickness ranging from 7 to 30 km (4 to 18 mi). The crust of earth described by its low temperature and high brittleness, making it prone to the occurrence of earthquakes along numerous fault networks. The occurrence of these earthquakes is attributed to the accumulation of tectonic stress, which leads to the occurrence of frictional sliding along fault lines. Earthquakes can also happen in regions where tectonic plates are being forced downwards into the Earth, known as subducting slabs. These slabs have the capability to reach depths of approximately 700 km (400 mi). At this depths, higher temperatures cause the rock to become softer. The mechanism behind frictional sliding in that location remains unknown, yet earthquakes occur. Another possible explanation is that these deep earthquakes are triggered by a "phase transition", in which the rocks undergo a quick and significant shift in their physical structure, transitioning to a different state. Earthquakes with magnitudes below approximately 7 can happen in both crust of earth and in the subducting slabs. Earthquakes with larger magnitudes, such as the 9.5 magnitude earthquake measured in 1960 in Chile, usually happen within the Earth's crust. These significant and severe earthquakes are linked to the sliding of faults due to friction, a phenomenon that can only happen at lower temperatures. The amount of strain energy stored within the rock caused variations in hypo central depth. Fault rupture begins toward the higher surface at this stage. The hypo central depth and magnitude relation provides an indication of crustal behavior and describes ground variability of earthquakes [45]. The representation in 1.2 shows the earthquake classification based on focal depth.

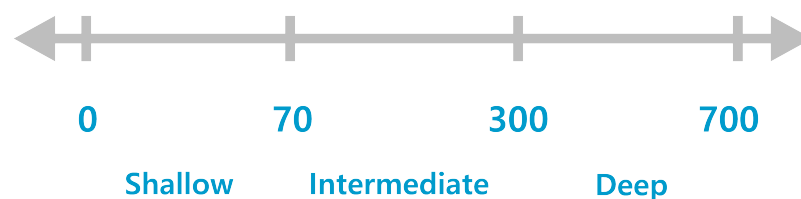


Figure 1.2: Classification of earthquakes based on their focal depth: shallow (0-70 km), intermediate (70-300 km), and deep (300-700 km).

1.8.2 Correlation of Magnitude, Depth, and Period Duration

Statistical methods and suitable tools can easily be used to interpret vast quantities of data. Studying the relationship between magnitude, depth, and period duration in earthquake seis-

mology is crucial for understanding the dynamics of seismic occurrences. Each of these criteria has a crucial role in characterizing earthquakes and evaluating their potential effects. Depth versus number of events represents the distribution of earthquake occurrences at various depths within the Earth. Depth of earthquakes, indicating their distance below surface of earth, meanwhile the number of earthquake events, showing the frequency of earthquakes at each specific depth. The correlation between earthquake size and frequency is often used to demonstrate the distribution of earthquake magnitudes. The magnitude of earthquakes as an indicator of their strength is represented on horizontal axis and the occurrence rate of earthquakes of various magnitudes as another measure represented on y-axis. The distribution of earthquake magnitudes in this plot often follows to a logarithmic pattern, indicating that smaller earthquakes occur with more frequency compared to bigger ones. The correlation between the duration of a period and the frequency of events often reflects the temporal pattern of earthquake events. Period duration refers to the time periods during which earthquakes are monitored, whereas the number of occurrences represents the frequency of earthquakes within each time interval.

1.8.3 Cumulative Frequency of earthquakes

The total number of earthquakes that have happened up to a specific moment in time is known as the cumulative frequency of earthquakes. The total number of earthquakes within a given time span (year) reflects the overall frequency of earthquakes occurring in that particular timeframe. A steeper slope indicates a higher rate of new events being added to the total. The relationship between them helps the identification of seismic activity trends across time, allowing for the determination of whether earthquake occurrences are experiencing an upward, downward, or remaining constant over time. Potential reasons for the rise in seismic activity rate can be attributed to factors such as climate change, a dynamic tectonic system, and the temperature beneath the Earth's surface [41].

1.9 Global Research on b-value

The value of b is an essential seismic parameter that characterizes the earthquake activity of a specific area during a specified time frame. The variable b is often known as b -slope, usually indicates the slope of the trend line indicating the correlation between the magnitude and the log of total number of earthquake events. The value of b mostly represents the relative distribution

of earthquake magnitudes. The value typically ranges from 0.5 to 1.5. In seismic active region the value of b is close to 1. Extensive research has been conducted globally on the b -value at regional and local levels [13] [48][54]. The value of b is a significant measure in seismology that characterizes the comparative occurrence rate of minor and major earthquakes within a specific area. The estimation is commonly conducted using the Gutenberg-Richter relationship, which establishes a connection between the magnitude and occurrence rate of earthquakes. Seismic asperity model and b -value relies on the Gutenberg-Richter relationship [51] [57]. Some researchers determine the b -value using either the least squares method or the maximum likelihood technique [3] [14] [26]. The relationship between frequency and magnitude of earthquakes is influenced by three key elements, which contribute to the large fluctuation in the annual occurrence of earthquakes.

1. The lack of uniformity in the material results in a higher b -value. [2]
2. Greater the lithospheric pressure, the lower the b -value [7].
3. The increase in the b -value is a result of a rise in the temperature gradient [4].

Broader Contexts for Studying b -Value

1. volume of crustal rock globally: Pertains to the total quantity and spatial arrangement of crustal rock on a global scale, which has an impact on earthquake occurrences and the b -value.
2. Tectonic environment characterized by a wide range of faulting patterns: The b -value is also influenced by different tectonic environments that have diverse fault processes and stress conditions.

The b -value and stress accumulation within the crust have an inverse relationship, meaning higher stress may be indicated by a lower b -value [6][7]. Additionally, there is a relationship between mainshock magnitude and effective stress [15]. Accurate identification of the b -value is crucial for assessing seismic hazards and understanding the underlying tectonic processes. Gaining a comprehension of these relationships can offer valuable understanding of the fundamental geophysical processes and contribute to enhancing the evaluation of seismic hazards. Collectively, the worldwide investigation into b -value and estimating techniques has enhanced comprehension of seismic activity and its consequences for evaluating and reducing risks.

1.10 Problem Statement

Earthquake Analysis Of The Hindukush Range By Utilizing Statistical Methods.

1.11 Aim & Objective

This dissertation aims to comprehensively analyze the seismic behavior of the Hindukush region, characterized by high seismic activity. The objective of this thesis is to conduct a thorough analysis of the seismic activity in the Hindukush region. This will involve studying the characteristics of the earthquake catalogue. In order to improve accuracy of seismic hazard declustering approaches such as the Gardner and Knopoff and the ETAS model are used. Additionally, it seeks to evaluate seismic risk by establishing the parameters of the Gutenberg-Richter relationship a for assessing the seismicity of the area and b for analyzing the tectonic stress and environment in which seismic activity occurs.

1.12 Research Question

- How can the application of the ETAS model and Gardner and Knopoff approach to declustering improve the accuracy of seismic hazard assessments in the Hindukush region.
- What insights can be gained about the region's seismicity?

1.13 Covered Topics

The layout of this thesis is outlined as follows: Initially, we present a summary of the tectonic features of the hindukush region, emphasizing its intricate seismic behavior and the merging of the Eurasian and Indian tectonic plates. Subsequently, we emphasize the significance of precise earthquake catalogs and declustering techniques including deterministic and stochastic models in enhancing assessments of seismic hazards. Chapter 2 of this thesis is devoted mostly to reviewing the existing literature. The dataset is presented and the methods used are outlined in the third chapter of our study. The presentation of results and discussions is provided in chapter 4 and is followed by the conclusion and Limitations in Chapter 5 and 6 respectively. The list of references is provided in the bibliography section.

CHAPTER 2

Literature Review

The exact date, location, and magnitude of an earthquake cannot be predicted with any degree of certainty using the current models since earthquakes are thought to be extremely unpredictable or stochastic events. Researchers have utilized diverse statistical methods to assess the existing earthquake data to gain a more thorough comprehension of the seismic characteristics associated with this region. The studies utilize many methodologies, such as statistical analysis, historical earthquake cataloging, machine learning algorithms, time series analysis and seismic hazard evaluations.

Verma thoroughly analyzed the seismotectonics of the Hindu kush region its earthquake activity and geological features of the Hindu Kush and Baluchistan arc identified the intricate interaction between the plates of Eurasian and Indian in this area that signifies understanding of regional tectonics to more accurately forecast and mitigate the consequences of earthquakes in these areas. Researchers found that there is a notable level of seismic activity in region of Hindu kush [12].

Expanding upon this fundamental research, Khattri [19] examined the possibility of significant earthquakes occurring along the Himalayan plate border, which is geologically linked to the Hindukush region. Khattri's research focused on seismic gaps, which are areas that have not seen severe earthquakes in recent times. This indicates that stress is accumulating in these locations, potentially leading to major seismic occurrences in the future. This study emphasized the significance of ongoing surveillance and evaluation of risks in the area to minimize the possible impact of earthquake catastrophes.

In 1998, a famous mathematical model created by Dahmen et al. [24] demonstrated a correlation between the magnitude of earthquakes and their frequency. The earthquake probability distribu-

tion model is crucial for the construction of structures. Petersen et al. (2007) [36] worked on a model for the California Geological Survey (CGS) that does not depend on time and shows the likelihood of an earthquake occurring according to a Poisson distribution.

Kagan et al. (2007) [35] analyzed abnormalities in earthquake occurrences in Southern California to make predictions about future seismic events with a magnitude of 5.0 and above for the next five years. Shen et al. (2007) [37] proposed a model for earthquake prediction using the analysis of tectonic plate strain. According to their theory, higher levels of strain increase the chances of an earthquake happening. Ebel et al. (2007) [34] implemented a forecasting model that projects future seismic events by extrapolating data from past earthquakes of 5.0 magnitude or higher.

This research conducted in southern California (2010) presents a statistical approach for identifying the background seismic activity and the induced seismicity. The approach utilizes the ETAS model and is used for monitoring seismic events. That model is employed for declustering seismic data, evaluate seismic background, and estimate parameters such as productivity coefficient, triggering distance, spatial distribution exponent, and magnitude productivity law [40].

The study [50] in 2017 employed the Maximum Likelihood approach to examine seismic activity in the Southern sea of D.I. Yogyakarta by utilizing the Gutenberg-Richter connection. The study concentrated on calculating the values of b , which are essential for comprehending the relationship between the frequency and magnitude of earthquakes. The results suggested a strong likelihood of earthquakes with a magnitude exceeding 6.0 happening in certain places within the Southern sea of D.I. Yogyakarta in the near future. These findings emphasize the importance of continuous seismic monitoring and the creation of methods to minimize future earthquake damage in the area.

The research conducted in the Hindukush region has employed this principle to calculate value of b . As an example, a study conducted by researchers [52] utilized the Gutenberg-Richter law to examine earthquake activity in region of Hindukush. The results showed value of b is similar to other areas with active tectonic activity. The discovery is crucial in grasping the seismic hazard and likelihood of major earthquakes in the region.

The writers of this paper [53] in 2018 used the ETAS model to decluster the seismic dataset of the Iranian plateau, covering the period from 1983 to 2017. The authors conducted a comparative analysis of the efficiency of the ETAS model in relation to other well-established method-

ologies, including Gardner and Knopoff and the Reasenberg method. The results suggested that the ETAS model shown superior performance in declustering seismic events, as it exhibited enhanced ability to distinguish between primary shocks, aftershocks, and background seismic activity. This paper offers a helpful assessment of several techniques employed in seismic analysis, with a particular focus on the usefulness of the ETAS model in comprehending and controlling earthquake catalogs.

The authors of this research in 2019 [55] derived the Gutenberg-Richter recurrence relationship by analyzing a 197 year earthquake catalogue from Gujarat, India. The study calculated hazard parameter values and emphasized the importance of the b -value's reliance on magnitude intervals. The research offers valuable insights on the seismic hazard potential of the region, which can be crucial for the development of safety measures and infrastructure resilience. This study highlights the need of utilizing long-term seismic data to precisely predict the occurrence of earthquakes and evaluate the hazards associated with earthquakes in a certain region.

Researchers in 2021 utilizes the Gardner and Knopoff approach to distinguish between mainshocks and aftershocks in earthquake data, with a specific emphasis on a magnitude of completeness of 3.8. The researchers utilized the Gutenberg-Richter law to compute the seismicity parameters, resulting in an a value of 7.34 and b values spanning from 0.2 to 2.0. The study emphasized that northern Pakistan displayed notably high values of b , indicating a higher occurrence of lower magnitude earthquakes in this area. The findings indicate that seismic activities in northern Pakistan are likely to occur more frequently than in the southwestern region. This has significant consequences for assessing and preparing for seismic hazards in the area [59].

The study conducted by Rahman et al. (2021) [61] centers on the Pamir-Hindu Kush region susceptible to earthquakes. Its main objective is to fill the gap in knowledge by creating a thorough earthquake catalog. They compiled comprehensive databases spanning from 1900 to 2017, utilizing data from the *ISC* and *USGS*. The analysis indicated that catalogs prior to 1960 are incomplete, whereas catalogs after 1960 provide accurate documentation of events with a moment magnitude (M_w) exceeding 5, with a precision of ± 0.1 . The revised catalogs offer an accurate basis for evaluating seismic danger and can be customized for forthcoming seismic risk analyses. The significance of precise historical data in assessing and reducing earthquake hazards in the area is emphasized by the insights.

Mirzrahi et al. (2021) [60] found that the frequency-magnitude distribution of mainshocks and the b -value of declusters are affected by the selection of declustering method and algorithm-

specific parameters. Declustering has a notable impact on the mainshock frequency and magnitude distribution, leading to a decrease in the b -value by as much as 30%.

This study [65] examined Pakistani seismic data using Gardner and Knopoff and Uhrhammer declustering methods. We wanted to find out which strategy best differentiates mainshocks from aftershocks to better understand the seismicity in the region. The Gardner and Knopoff approach was 4% more accurate than the Uhrhammer method, making it the better declustering method. Improved seismic hazard assessments and better understanding of Pakistan's seismic processes require this accuracy.

The primary objective of the work done by researchers in 2022, is to utilize the ETAS model for the purpose of forecasting aftershock sequences that occur after major seismic occurrences in the Kermanshah region. The authors employed the ETAS model to detect and monitor the series of aftershocks that followed two significant earthquakes in 2017 and 2018. The study's findings suggest that the ETAS model is successful in predicting the sequence of aftershocks, offering significant understanding of the particular dynamics of aftershocks in the Kermanshah region. This study highlights the model's capacity to accommodate various parameterizations, indicating that ETAS can serve as a flexible tool for predicting aftershocks in places with diverse seismic characteristics [66].

In 2022 [64] researchers analyzed how the declustering method affects the hazard level in the Kathmandu Valley scenario. The Gardner and Knopoff declustering method is deemed better suited for the study area compared to Reasenbergs' method after a thorough evaluation. The comparison highlights how these declustering methods impact the assessment of seismic hazard. Furthermore, The study conducted in 2023 used the Gutenberg-Richter relation to examine seismic activity in the Banda Arc region, with a particular focus on computing seismic parameters such as the a value and b value. These parameters play a vital role in comprehending the occurrence and intensity of earthquakes. The analysis uncovers a significant amount of seismic activity, characterized by a magnitude of completeness (M_c) of 4.5, an a value of 9.07, and a b value of 1.33. The research findings suggest that the earthquake data is comprehensive up to a magnitude of 4.5, which offers a strong dataset for examining seismic trends and assessing possible hazards in the Banda Arc region. These findings emphasize the significance of comprehending regional seismicity factors in order to improve earthquake preparedness and risk mitigation measures [67].

CHAPTER 3

Materials and Methods

In this study, we present a summary of the dataset utilized for our analysis, along with two different statistical methods employed to analyze the earthquakes in the Hindukush region: Gardner and Knopoff method and ETAS Model. The methodology outlines the steps involved in applying these declustering methods to identify mainshocks, aftershocks, and seismic sequences, while the selected dataset serves as the foundational component for our research endeavors.

3.1 Data Acquisition

The study used the earthquake data sourced from National Seismic Monitoring Center Islamabad, Pakistan. The dataset spans from August 1, 2008, to July 13, 2010, capturing seismic events over a significant time frame. The seismic data pertains to the Hindukush region, bounded by latitudes 35.72°N to 43.66°N and longitudes 68.59°E to 77.90°E . The dataset includes essential information on seismic events, such as date, time, magnitude, depth, latitude, and longitude of over 9532 earthquakes. The seismicity map of hindukush is shown in [3.1](#)

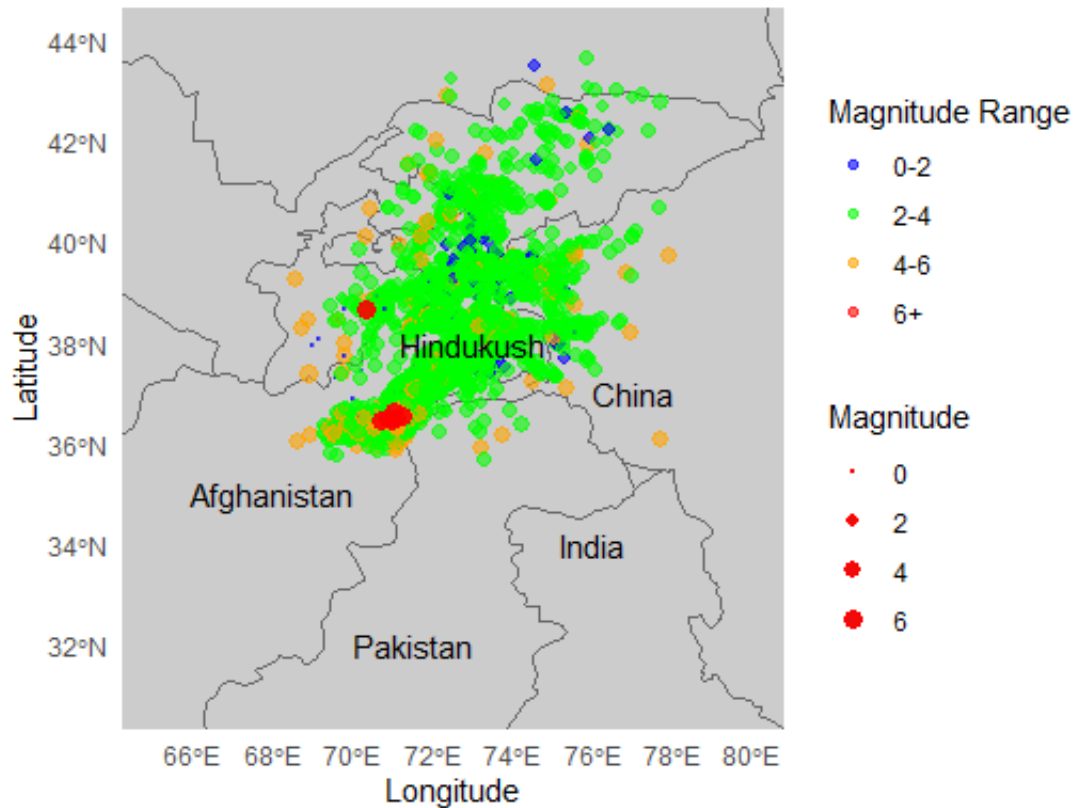


Figure 3.1: The visual representation of all seismic events in Hindukush and the surrounding area of different magnitudes, from 2008 to 2010.

3.2 Window-Based Declustering Method

The process of declustering the earthquake catalogue is essential to isolate mainshocks by removing dependent events, such as foreshocks and aftershocks, which can distort the analysis of seismic hazard. This study employs the Gardner and Knopoff (1974) window method, a widely recognized technique in seismology for identifying and removing these dependent events.[8]

3.2.1 Gardner and Knopoff Method

The Gardner and Knopoff method is based on the concept of spatial and temporal windows. By defining specific windows around each earthquake event, the method identifies clusters of events that are likely to be foreshocks or aftershocks associated with a mainshock. The parameters for these windows are typically based on empirical relationships derived from historical earthquake data.

1. **Parameter Selection** To apply the Gardner and Knopoff method effectively, it is crucial to select appropriate spatial and temporal windows.

- **Spatial Window (d):** The spatial window defines the distance in kilometers within which events are considered related. This study used the following empirical formula:

$$d = 10^{0.1238M+0.983} \quad [\text{km}]$$

Where M is the magnitude of the earthquake.

- **Temporal Window (t):** The temporal window defines the time period within which events are considered related. This study used the following empirical formula:

$$t = \begin{cases} 10^{0.032M+2.7389} & \text{if } M \geq 6.5 \\ 10^{0.5409M-0.547} & \text{else} \end{cases} \quad [\text{days}]$$

Where the magnitude of the earthquake is represented by M .

3.3 Gutenberg-Richter Relationship: Linking Earthquake Frequency and Magnitude

Gutenberg-Richter (1935) developed a relationship between earthquake occurrence and magnitude frequency, which is used to calculate annual earthquake occurrence rates.

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

$$N = 10^{a-bM}$$

The magnitude of the earthquake is denoted by M , while the total number of earthquakes with a magnitude greater than M is represented by N . Two constants, a and b are used here. The

overall seismicity of the region is represented by a value. Different magnitudes of earthquakes are described by their relative likelihood by the b value.

3.4 Magnitude of Completeness (M_c)

The magnitude of completeness (M_c) is a vital metric in seismic hazard assessments as it denotes the minimal magnitude required to accurately identify all earthquakes in a specific location. Rigid seismic hazard assessment is made possible by an accurate M_c determination, which ensures that the earthquake database is complete. For the purpose of evaluating M_c , the maximum curvature method was employed due to its simplicity and reliability. The maximum curvature approach can be employed to determine the magnitude of completeness (M_c), by finding the point on the earthquake magnitude-frequency distribution curve where the curvature is maximum, or where the slope changes the most sharply.

3.5 The ETAS Model

3.5.1 Marked-Point Temporal Processes

In an earthquake catalog with N occurrences, let t denote the time, i the longitude, w the latitude, and m the magnitude of the i th earthquake event, respectively. A point pattern $\{(t_i, x_i, y_i, m_i) : i = 1, \dots, N\}$ on $\mathbb{R}^+ \times \mathbb{R}^2 \times [m_0, \infty)$ can be used to represent the catalogue data. This pattern of points is generated by the temporal marked point process X , which regulates the spatial and temporal frequency of earthquakes. The conditional intensity function $\lambda(t, x, y, m | H_t)$ uniquely determines the distribution of the temporal marked point process X . This is the case for the conditional intensity function:

$$\lambda(t, x, y, m | H_t) dt dx dy dm \approx \mathbb{P}\{X \cap [t, t + dt) \times [x, x + dx) \times [y, y + dy) \times [m, m + dm) | H_t\}.$$

Until time t , the variable H_t represents the history of earthquake events.

3.6 The Space-Time ETAS Model

For the purpose of declustering the Hindukush seismic catalogue, the Epidemic-Type Aftershock Sequence (ETAS) model is presented in this study. As part of the procedure, we estimate

the likelihood of an earthquake triggering and establish a threshold value for declustering. The ETAS model can be represented by the conditional intensity function:

$$\lambda(t, x, y) = \mu(x, y) + \sum_{k=1}^{\infty} \kappa(m_k) g(t - t_k) f(x - x_k, y - y_k | m_k).$$

The $\mu(x, y)$ intensity function, which is believed to be independent of time, represents the background seismicity. The $g(t)$ and $f(x, y | m_k)$ as the normalized response functions. Given an ancestor event of magnitude m_k , these functions show the probability distributions of its timing and location as offspring events. The intensity function $\kappa(m_k) g(t - t_k) f(x - x_k, y - y_k | m_k)$ represents the non-stationary Poisson process that is initiated by the k th event. The expected number of descendants from a parent with a magnitude of m_k is denoted by $\kappa(m_k)$ in this intensity function.

The ETAS model encompasses various essential elements to depict the occurrence of aftershocks, such as the anticipated quantity of aftershocks, the temporal distribution of their events, and the spatial distribution of their locations.

Expected Number of Triggered Events The expected number of aftershocks produced by an occurrence with a magnitude of m_i can be mathematically represented as:

$$\kappa_{A, \alpha}(m) = A \exp(\alpha(m - m_0))$$

The equation represents a mathematical relationship, where κ is a variable that depends on A , α , and m . The parameters A and α are unknown and need to be calculated. It is important to note that α must be greater than zero and A must be more than zero as well. The parameter A governs the fundamental rate of aftershocks, while α defines the degree to which aftershock productivity is influenced by the magnitude of the original event.

Probability Distribution of Occurrence Time The PDF of the time until a triggered event occurs, following a mainshock of magnitude m_i at time t_i , is represented by $g_{c,p}(t - t_i)$. The assumption is that the time difference $(t - t_i)$ between the direct mainshock and a specific event (t_i) is not influenced by the magnitude of that event (m_i). The PDF (Probability Density Function) can be determined using the inverse power law, which is a modification of Omori's law.

$$g_{c,p}(t-t_i) = \begin{cases} \frac{p-1}{c} \left(1 + \frac{t-t_i}{c}\right)^{-p}, & t-t_i > 0, \\ 0, & t-t_i \leq 0, \end{cases}$$

Given that c is more than 0 and p is greater than 1, where c and p are unknown factors. The parameter c regulates the initial rate of decay of aftershocks immediately after the mainshock, while p defines the long-term decay rate.

Probability Distribution of Occurrence Location The spatial distribution of aftershocks is represented by the probability density function (PDF) $f_{D,\gamma,q}(x-x_i, y-y_i; m_i)$ which describes the likelihood of a triggered event occurring at a specific location (x, y) , given a mainshock of magnitude m_i at a different location (x_i, y_i) presented as:

$$f_{D,\gamma,q}(x-x_i, y-y_i; m_i) = \frac{q-1}{\pi D \exp(\gamma(m_i - m_0))} \left(1 + \frac{(x-x_i)^2 + (y-y_i)^2}{D \exp(\gamma(m_i - m_0))}\right)^{-q}$$

Let D , γ , and q be unknown parameters such that $D > 0$, $\gamma > 0$, and $q > 1$. The parameter D determines the size of the aftershocks, γ indicates how the magnitude of the mainshock affects their spatial distribution, and q determines how the density of aftershocks diminishes as the distance from the mainshock rises.

3.6.1 Total Spatial Intensity

This metric approximates the overall seismic activity in a certain area (x, y) and is called the total spatial intensity function $\Lambda(x, y)$. It takes into account both the baseline seismicity, which is the average frequency of earthquakes, and the aftershocks, which are triggered occurrences that are affected by earlier earthquakes. It can be expressed mathematically as:

$$\Lambda(x, y) \approx \mu u(x, y) + \frac{T_1}{T} \sum_{i:t_i < T} \kappa(m_i) f(x-x_i, y-y_i | m_i)$$

$\mu u(x, y)$ stands for the seismic activity in the background. $\frac{1}{T} \sum_{i:t_i < T} \kappa(m_i) f(x-x_i, y-y_i | m_i)$ accounts for aftershocks caused by previous earthquakes are taken into consideration by adding up the contributions from all events i that happened before time T .

3.6.2 Clustering (Triggering) Coefficient

The clustering or triggering coefficient $\omega(x,y)$ measures how much the measured seismic activity in a region (x,y) is affected by aftershocks, or triggered events, rather than the background seismicity. It is defined as

$$\omega(x,y) = 1 - \frac{\Lambda(x,y)}{u(x,y)}$$

Aftershocks, rather than background events, are responsible for the majority of seismic activity in the area.

3.6.3 Stochastic Declustering

To determine if a particular earthquake (Event j) was induced by a preceding earthquake (Event i) or if it occurred randomly as part of the background seismic activity is calculated by comparing the background earthquake rate at the location of event j to the overall rate of earthquakes.

$$\phi_j = \frac{\mu(x_j, y_j)}{\lambda(t_j, x_j, y_j)}$$

and to calculate triggering probability, meaning if it's not a background event, it's a triggered one is calculated as:

$$\rho_j = 1 - \phi_j = \sum_{i=1}^{j-1} \rho_{ij}.$$

ρ_j is close to 1, the event is likely a background event otherwise triggered. ϕ_j is close to 1, the event is likely a triggered event, otherwise background.

3.6.4 Parameter Estimation and Model Fitting

The ETAS model, as described by the above equations, is a semi-parametric model that has an infinite-dimensional parameter $u(x,y)$, $(x,y) \in S$, and Euclidean parameters β and $\theta = (\mu, A, \alpha, c, p, D, \gamma, q)$ [33].

- **Log-Likelihood Function**

To assess the goodness-of-fit of the ETAS model to the observed earthquake data, a log-likelihood function is created:

$$l(\beta, \theta | H_T) = \sum_{i=1}^N \delta_i \log \lambda_{(\beta, \theta)}(t_i, x_i, y_i, m_i | H_T),$$

while δ_i is equal to 1 when event i is a target event and equal to 0 when it is not. Because β and θ are independent variables in the conditional intensity function, we may break down the log-likelihood into its component parts accordingly:

$$l(\beta, \theta | H_T) = l_1(\beta | H_T) + l_2(\theta | H_T),$$

where for β the log-likelihood function (l_1) is defined as,

$$l_1(\beta | H_T) = \sum_{i=1}^N \delta_i \log V_\beta(m_i) = \log \beta \sum_{i=1}^N \delta_i - \beta \sum_{i=1}^N \delta_i (m_i - m_0),$$

and for θ the log-likelihood function (l_2) is defined as,

$$l_2(\theta | H_T) = \sum_{i=1}^N \delta_i \log \lambda_\theta(t_i, x_i, y_i | H_{t_i}) - \int_{t_{start}}^{t_{start}+T} \int_S \lambda_\theta(t, x, y | H_t) dx dy dt.$$

• Parameter Estimation

The Maximum Likelihood Estimation (MLE) approach is used for estimating parameters. The parameter $\hat{\beta}$ is estimated by maximizing the portion of the log-likelihood function that is associated with β .

$$\hat{\beta} = \frac{N'}{\sum_{i=1}^N \delta_i (m_i - m_0)},$$

The number of target events is represented by N' .

Maximizing $l_2(\theta | H_T)$ is the MLE for the parameter vector θ . Usually, methods that use gradients, such the Davidon-Fletcher-Powell (DFP) approach, are used to carry out this optimization.

CHAPTER 4

Results and Discussions

In this thesis the properties of Earthquake Catalogs are representing in detail by presenting the relationship between magnitude and depth and correlation of magnitude, depth and time. After that declustering methods are used to remove dependent earthquakes, which is essential for analyzing earthquake patterns and assessing seismic hazards. A detailed analysis is given below.

The scatter plot illustrates the relationship between magnitude and depth that is shown in 4.1. Shallow earthquakes predominantly have magnitudes between 1 and 5. Intermediate earthquakes have wider range of magnitudes from 1.5 to 7. The magnitude exceeds 7 at intermediate depth in their relationship. Saturation is evident at various depth levels in plots, ranging from magnitude 2 to 3.5 at shallow depths and from 2.5 to 4 at intermediate depths, highlighting a frequent incidence within these ranges at their specific depths. No deep earthquakes can be detected beyond 280 km depth as there are no seismic events observed. The table 4.1 demonstrates that shallow-depth earthquakes occur more frequently than intermediate-depth earthquakes.

	Earthquake Events
Shallow	6274
Intermediate	3258

Table 4.1: Earthquake Events categorized by Depth Distribution. This table classifies the overall number of earthquake occurrences into two depth ranges: shallow earthquakes (0-70 km) and intermediate earthquakes (70-280 km), with 6274 and 3258 occurrences respectively.

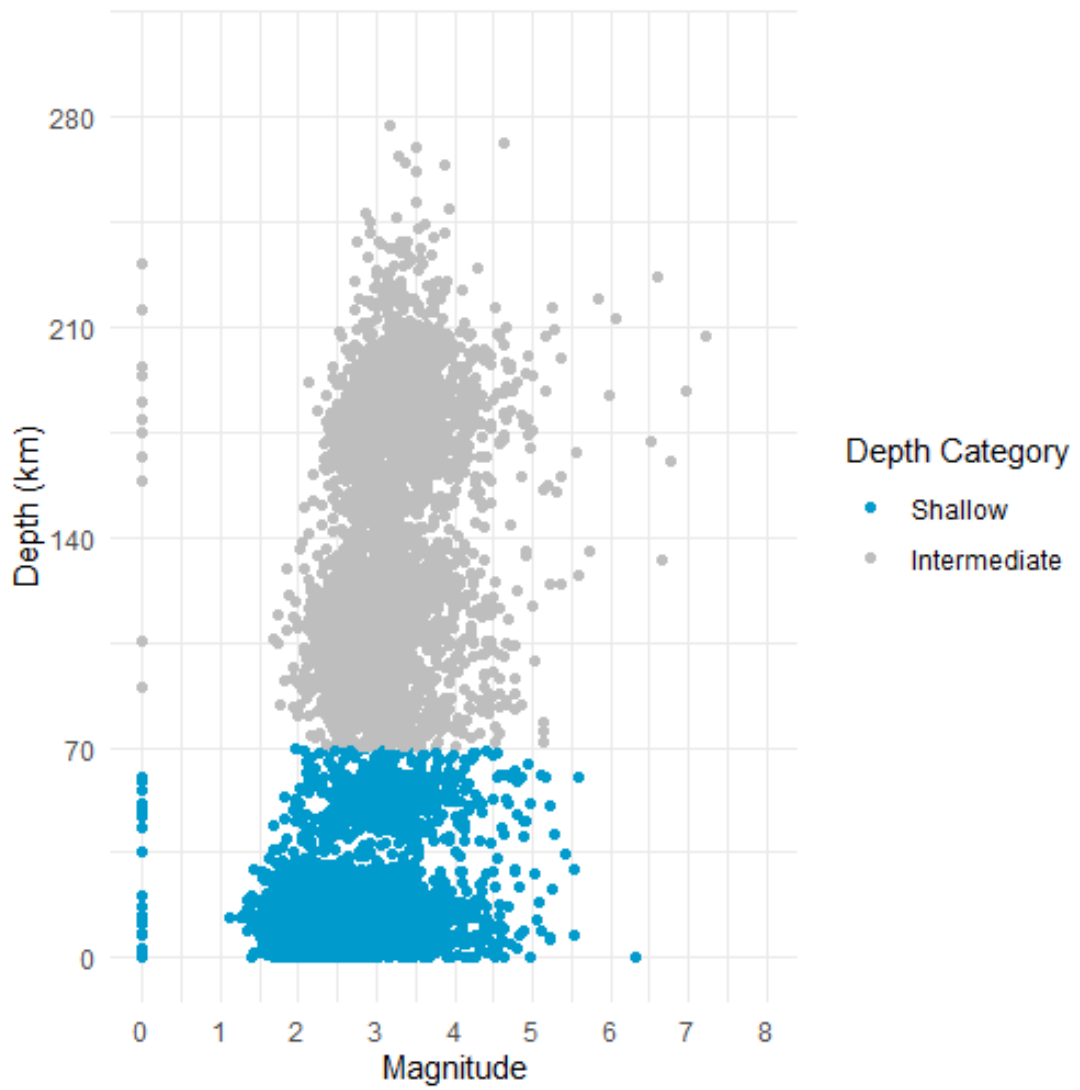


Figure 4.1: The plot represents the correlation between magnitude and depth of earthquake catalog. Shallow depth are represented in blue dots and Intermediate depth are represented in grey dots.

In figure 4.2 graph displaying the relationship between the depth of earthquakes, their magnitude, and the frequency of occurrence is presented. The bar chart in subfigure 4.2a illustrates the frequency distribution of earthquakes with magnitude ranges from 0 to 8. The earthquake's magnitude is represented by the height of the bar, which has 500 intervals per bar. The histogram indicates about a thousand instances of earthquakes with magnitudes from 1.8 to 3.4. Seismic activity is occurring rapidly within this magnitude range. The highest amount of events reaches a magnitude of 2.5. The seismic distribution relation to depth was described using the bar chart in 4.2b. Shallow earthquakes are observed between 0 and 70 kilometers and have more earthquakes than intermediate events. The subfigure 4.2c illustrates the relationship between the frequency of earthquakes and the number of years they occur. There has been a noticeable increase in earthquake frequency in 2009. Overall it indicates the high frequency of earthquake events throughout the time period. The 3D subfigure 4.2d displays the frequency distribution of shallow to intermediate depth seismic events based on year, magnitude, and depth.

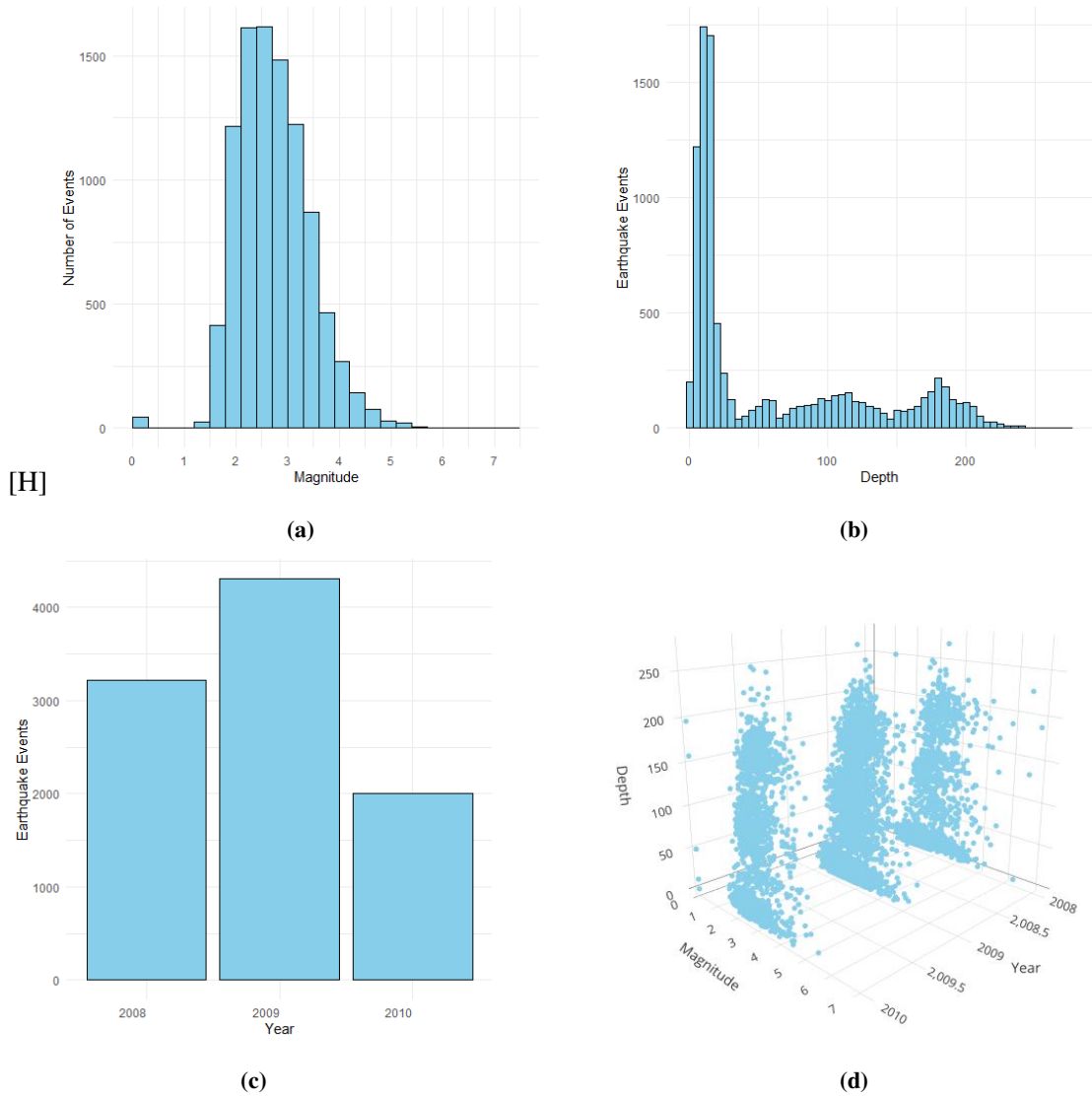


Figure 4.2: (a) Relationship between seismic events and earthquake magnitudes, (b) Relationship between Focus depth and seismic events, (c) Relationship between number of seismic events and years, (d) Relationship between focal depth, magnitude, and year of occurrence of seismic events represented by a 3D graph.

The cumulative number of earthquakes is shown in figure 4.3. The graph begins at a low point in mid-2008, indicating a small accumulation of events. There is a noticeable and rapid increase in the frequency of events from late 2008, indicating a significant rise in earthquake activity. The slope of the curve remains steep, indicating a persistent and elevated frequency of earthquake events. During this period, there was a significant rise in the overall quantity of earthquakes, providing evidence that 2009 had the highest occurrence of earthquake events. The slope of the curve starts to flatten out considerably after mid-2009, suggesting a minor decrease in the frequency of earthquake events. Nevertheless, the total number of events is increasing consistently but at slower pace, reaching above 9000 events by mid 2010. The general upward trend throughout the entire study period signifies ongoing seismic activity, with no extended period of inactivity.

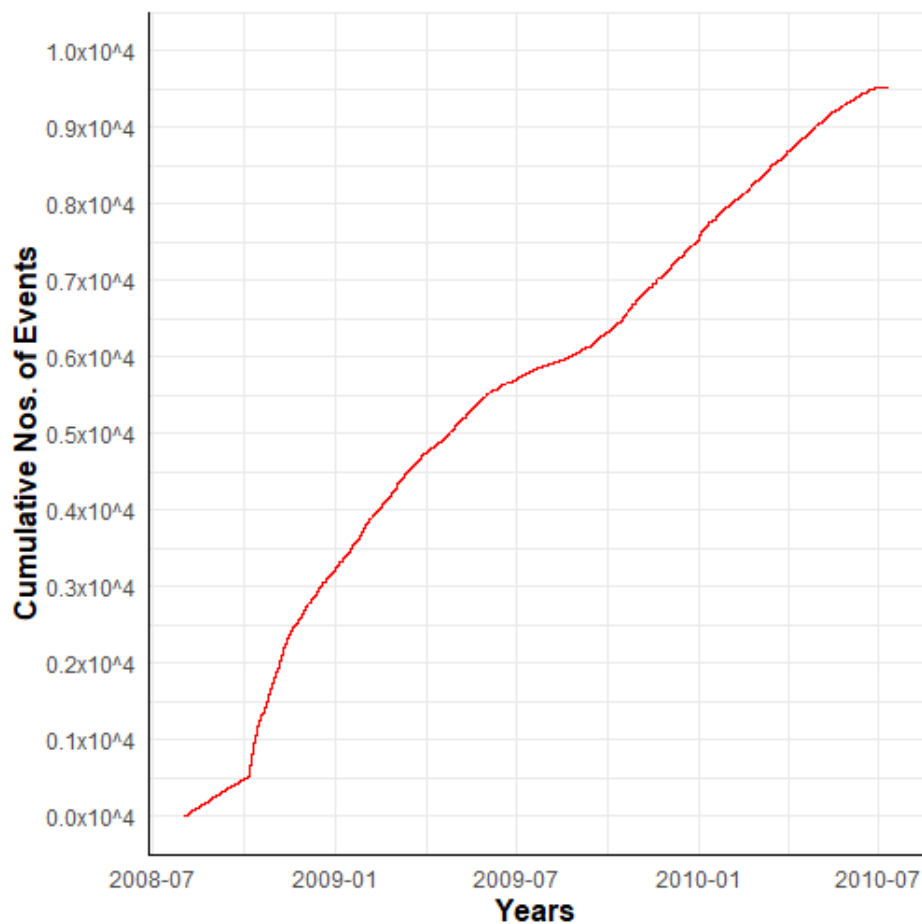


Figure 4.3: The cumulative number of earthquakes is shown in relation to the time period (year).

The elimination procedure of foreshocks and aftershocks from an earthquake catalogue is a fundamental step in verifying the consistency and authenticity of the catalogue. When the data is precisely recorded and best fitted, the result becomes comprehensible and readily interpretable. This thesis uses the ETAS model and the Gardner and Knopoff approach to decluster earthquakes. The ETAS model is utilized to enhance understanding of the fundamental patterns in earthquake data, for precise declustering and advanced analysis of seismic activity.

The Gardner and Knoff declustering approach relies on the use of both distance and time window criteria. The time and distance window by Gardner and Knopoff (1974) is used for aftershock identification. Any seismic event occurring within the two specified time intervals will be classified as either foreshocks or aftershocks. Table 4.2 depicts the drawing of window that is used to quantify the temporal and spatial correlation between the aftershock intensity and the time and distance data.

M	L (km)	T (days)
2.5	19.5	6
3.0	22.5	11.5
3.5	26	22
4.0	30	42
4.5	35	83
5.0	40	155
5.5	47	290
6.0	54	510
6.5	61	790
7.0	70	915
7.5	81	960
8.0	94.0	985

Table 4.2: Aftershock identification windows

[8]

According to this method, if an earthquake of magnitude 4.5 occurs and another of the same magnitude happens within the next 83 days at an epicentral distance of 35 km, the latter is considered an aftershock. This process is repeated until the entire catalogue is reviewed. The declustering found 3164 clusters of earthquakes with a total of 3574 (37.5 %) events out of

9532. The declustering approach detected 3164 clusters of earthquakes, which represents a substantial proportion of the overall seismic occurrences. Among the initial 9532 events, 3574 were categorized as independent mainshocks. Approximately (62.5 %) of the events were eliminated as aftershocks or foreshocks, which are dependent on the mainshocks. After the declustered events were identified, a map was drawn for these events 4.4. These are referred to as the mainshocks in the remaining earthquake catalogue.

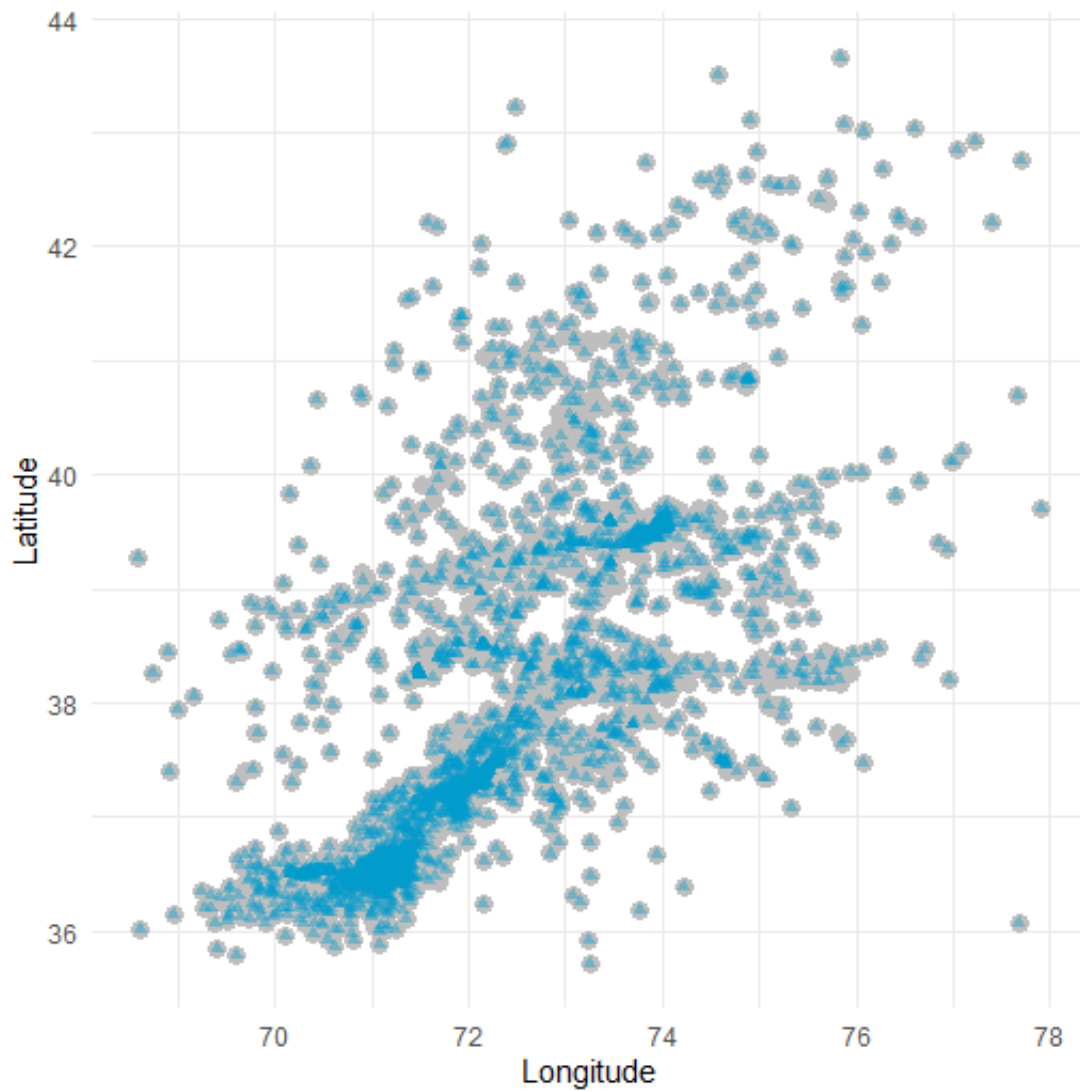


Figure 4.4: By using the window approach, the remaining events (mainshocks) are shown as deep sky blue-colored triangles. The total number of events are shown in grey-colored points.

The magnitude of Completeness was obtained by using maximum curvature method. A very low M_c value (nearing negative) implies problems in seismic parameters and possible errors in data collection, whereas a greater M_c value indicates an undersampled catalogue with duplicate data [45]. In this study, the samples were properly balanced, yielding a significant M_c value. Using the maximum curvature method for the period 2008-2010, M_c was determined to be 3.03 as shown in Figure 4.5. This means that earthquakes with a magnitude equal to or greater than 3.03 were reliably detected and recorded during that time frame.

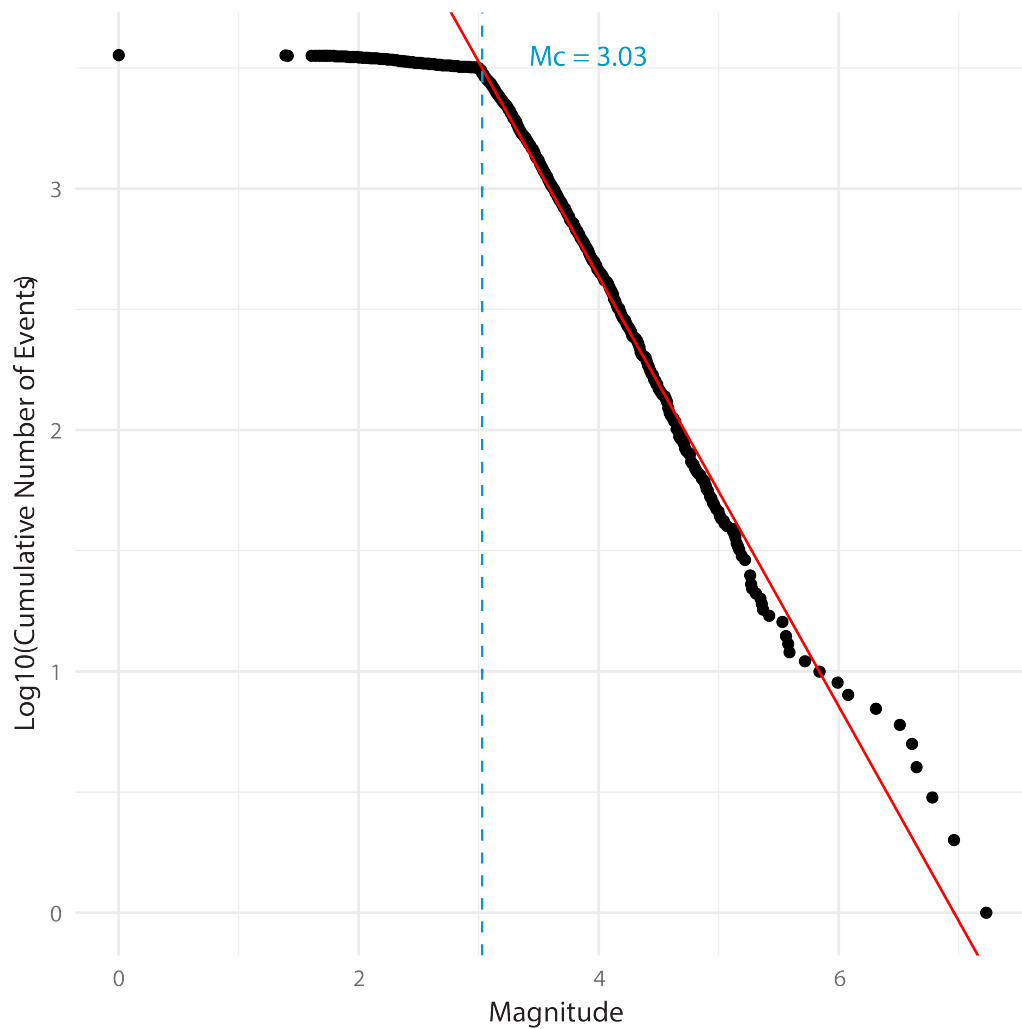


Figure 4.5: Plot of the cumulative number of earthquake events (on a logarithmic scale) against magnitude with “ a ” value of 5.21 and “ b ” value of 0.98, obtaining magnitude of completeness 3.03 through maximum curvature method.

Using log-linear regression of the earthquake frequency-magnitude distribution, the parameters were determined to have an a -value of 5.21 and a b -value of 0.98. The high value of a indicates that the region is highly seismic active. The behavior of the b -value is a critical factor to take into account. As the b value is close to 1.0 in seismically active regions. The b -value is utilized to characterize the regime of stresses and tectonic environment. The parameter b indicates that the region has moderate tectonic stress and it increases as the value of b decreases. The region's tectonically activeness determines by the a and b -values of the various zones. In 4.6 representation of the spatiotemporal values of a and b for Hindukush Region is shown. The map shows that the a -value changes in different parts of the Hindukush area, which means that there are different levels of seismic activity. The moderate value of seismicity rate " a " ranges from 3 to 3.5. Regions with higher frequencies of earthquake occurrence are indicated by higher a values (in red and orange areas). Regions with lower frequencies of earthquake occurrence are indicated by lower a values (seen in blue and green areas). A larger value of parameter a suggests a region with greater seismic activity. The value of b indicates that small earthquakes are more frequent and the region has moderate tectonic stress and it increases as the value of b decreases. Regions with values close to 1.0 have comparatively less tectonic stress. The high stress is indicated by yellow orange area. Geographically, along the Pakistan-Afghanistan and Afghanistan-China borders a value is high. The frequency of earthquake events is increased as a result of the intense seismicity that is observed in this area. The tectonic stress is comparatively high along Afghanistan-China border as compared to the Afghanistan-Pakistan. These regions are highly prone to earthquakes, as evidenced by a and b -values. There are numerous fault systems that have been generated by the convergence of plate boundaries at this location. They are predominantly reverse faults that are produced as a result of over thrusting. Different zone exhibit distinct a and b -values, which are contingent upon the level of tectonic activity in each zone.

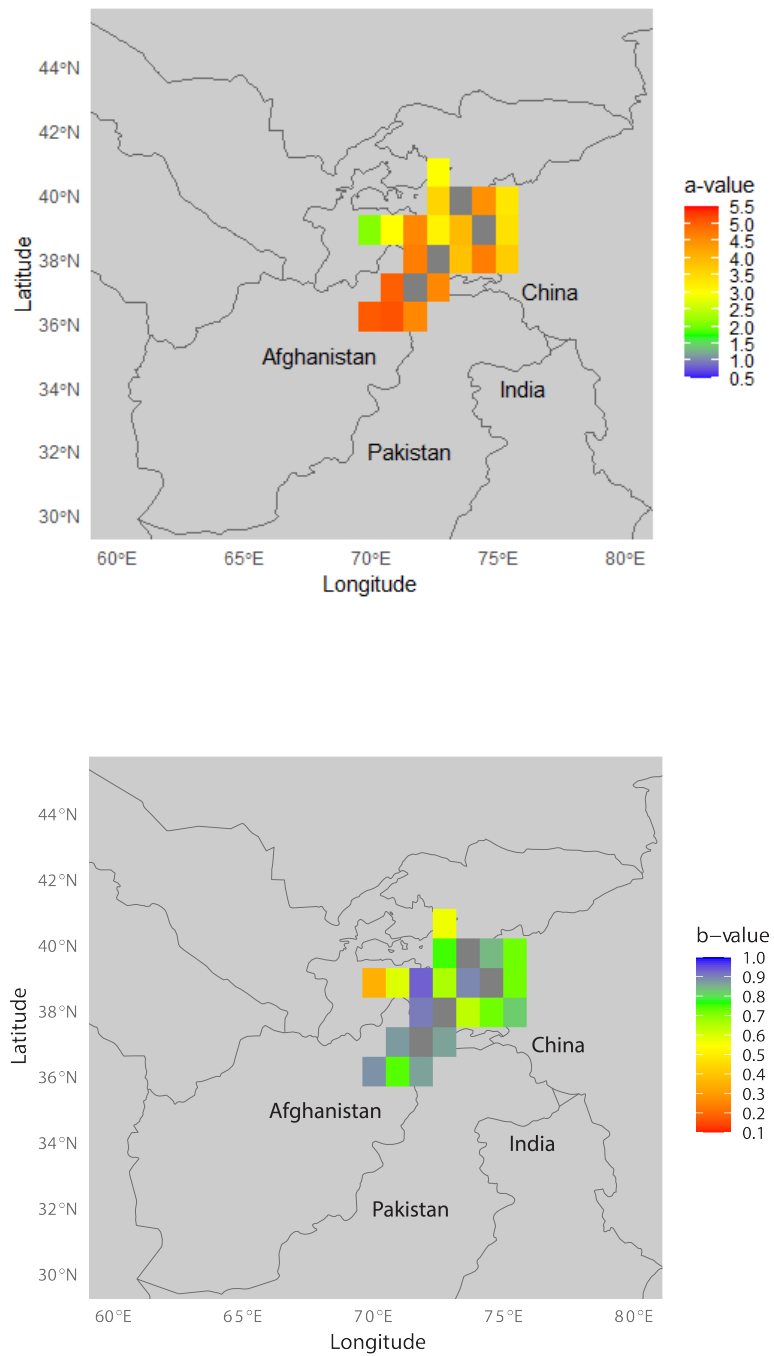


Figure 4.6: The regional and temporal mapping of the *a* and *b*-values for the rate of change in seismicity and region's slope respectively.

The quality and quantity of data used in statistical approaches like the ETAS model significantly affect the final results. A collection of point coordinates within a rectangular region of two-dimensional research zone is referred as a geographical region in ETAS. The latitude and longitude ranges for a rectangular geographical region are defined. The rectangular geographical study area is bounded by 36° – 37° N and 70° – 72° E, and the earthquake catalog for fitting ETAS model started on August 1, 2008, and it covered a study period of 518 days, from August 1, 2008, to January 1, 2010. The magnitude threshold was set to 4. The duration of the study period is $T = 518$ days. The total events are 324, among them, 146 took place as target events inside the designated geographic area, and 178 took place as complementary events outside the area. Nothing was recorded outside of the study period; all events occurred during that time. Figure 4.7 shows a plot depicting the geographical distribution of events within a particular area. Three graphs illustrate the temporal variations in latitude, longitude, and magnitude of occurrences. The two charts visually evaluate the extent to which the catalog is comprehensive and stationary. The relationship between $\log_{10}(N_m)$ and m exhibits a linear pattern that aligns with the Gutenberg-Richter law. The relationship between N_t and t exhibits a linear pattern throughout the research period, suggesting that the catalogue is both complete and stationary over time.

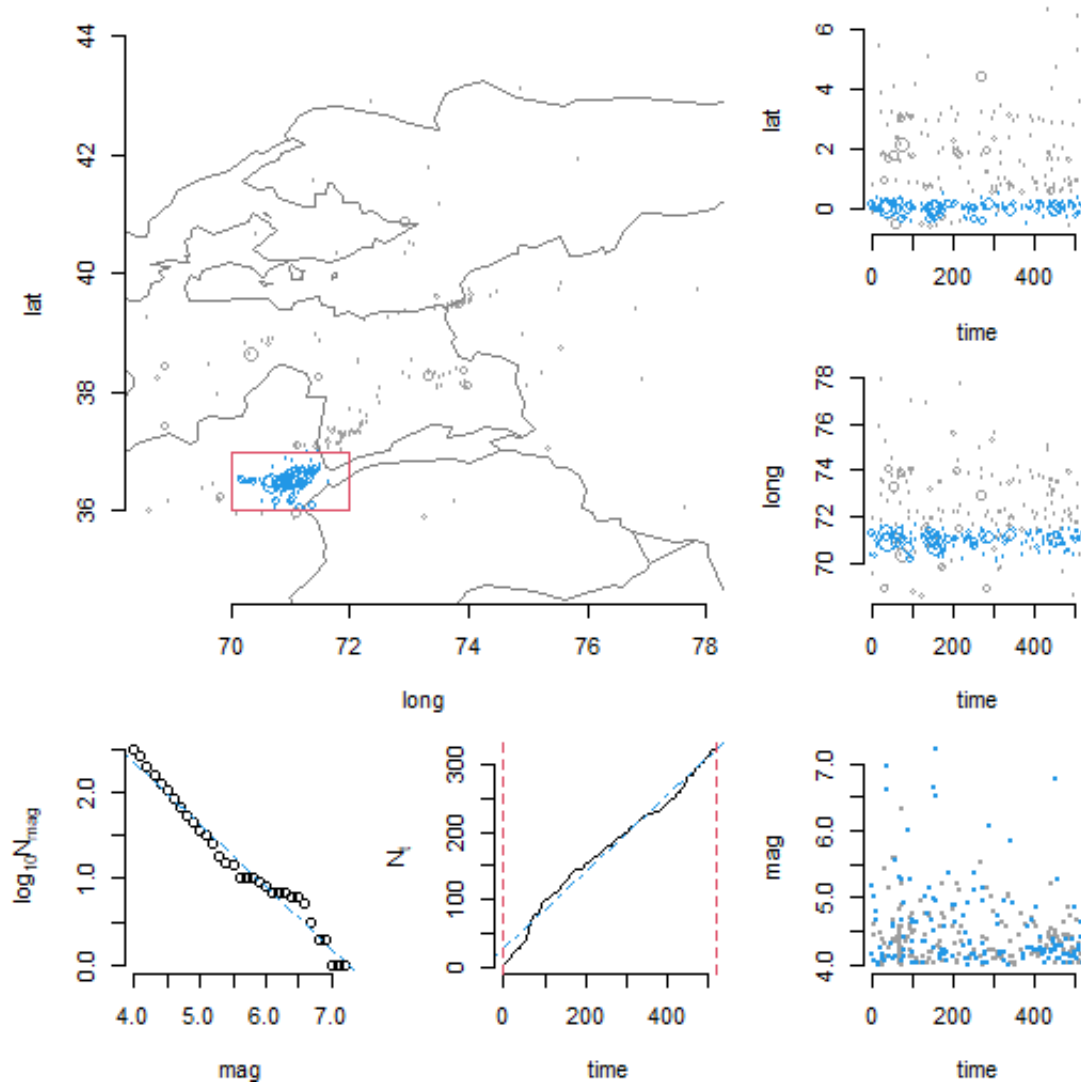


Figure 4.7: The epicenters' locations in the top left panel, the logarithm of frequency by magnitude in the bottom left panel, the cumulative frequency over time in the bottom middle panel, and the latitude, longitude, and magnitude against time in the right panels of 9532 earthquakes with a magnitude greater than or equal to 4 occurred between 2008-08-01 to 2010-01-01 in the Hindukush region and its surrounding area (36° – 37° N and 70° – 72° E).

The ETAS model was calibrated using the iterative stochastic declustering approach and reached convergence after 7 rounds, with a total execution time of 6.44 minutes. The Maximum Likelihood (ML) estimates of the model parameters and their standard errors are presented in 4.3. The value of log-likelihood and AIC obtained was -135.8734 and 287.7468 respectively. In the ETAS model, every parameter has a specified role. The model integrates both parametric and non-parametric components, enabling a flexible representation of seismic activity. Infinite-dimensional parameter $u(x,y)$ is form of non-parametric component which depends on the location (x,y) within the region S where the earthquakes are being studied. Euclidean Parameters β and a vector $\theta=(\mu,A,\alpha,c,p,D,\gamma,q)$. Additionally, the role of each parameter in the model provides further insight into the mechanics behind earthquake formation and propagation. The Maximum Likelihood estimates of the model parameters, along with their corresponding standard errors, reveal the intricate dynamics that regulate the incidence of earthquakes and the following tremors that follow. The distribution of earthquake magnitudes within the dataset is governed by the β parameter. A greater β value indicates a more rapid decline of higher magnitude events frequencies. In this instance, a value of 1.4878 for β indicates that in the Hindukush area, larger earthquakes occur less often, and smaller events are more common in the dataset. This aligns with the fact that the majority of earthquakes are minor, while major ones are infrequent. It helps in evaluating the overall seismic activity within a specific region. In simple terms value of β indicates small earthquakes occur frequently, while major earthquakes are significantly less frequent. This information is crucial for comprehending earthquake hazards and making necessary preparations. This refers to the estimation of the predicted frequency of an earthquake within a particular time frame in a specific region, without taking into account any aftershocks resulting from past earthquakes. The average number of earthquakes happening that are not caused by other earthquakes is known as the background seismicity rate, μ . A value of 1.01 suggests the occurrence of background earthquake with a magnitude of 4.0 or more, indicating a steady amount of seismic activity regardless of aftershock sequences. This is the underlying level of seismic activity in the area, excluding any secondary earthquakes. The productivity parameter A describes that a small earthquake (at threshold magnitude) might trigger roughly 0.0912 aftershocks. This suggests that aftershocks are quite infrequent in this location. Nevertheless, the standard error of 0.3464 suggests a substantial degree of uncertainty in this estimation, implying that the true output of aftershocks could differ considerably and may not be accurately forecasted. The value of parameter c , which is equal to 10.4336, signifies that there will be a higher occurrence of aftershocks for roughly 10.43 units of time after the mainshock.

After this period, the frequency of aftershocks will gradually decrease. The parameter p determines the rate at which the occurrence of aftershocks decreases over time after a mainshock. A result of p equals to 1.1559 implies that the aftershock activity decays at a moderate rate, indicating a balanced and gradual decrease, it indicates that after the duration of 10.43 days, the frequency of aftershocks will gradually decrease at a moderate rate, neither too fast nor too slow. Within the framework of the ETAS model, this parameter holds significant importance as it determines the duration of the aftershock sequence, namely the length of time in which aftershocks persist. A p value near to 1, such as 1.1559, is commonly observed in numerous seismic zones, suggesting a typical rate of decay for aftershocks. When examining both values c and p collectively, they characterize the temporal decline pattern of aftershocks following a mainshock. The values of $c=10.4336$ and $p=1.1559$ suggest that following a large seismic event, the rate of aftershocks remains elevated for around 10.43 units of time before gradually starting to decline. The decrease in aftershock activity exhibits a modest and controlled decline, which is determined by the p value. These factors are crucial for predicting the succession of aftershocks and comprehending the temporal patterns of seismic activity that occur after a mainshock in the Hindukush region. The temporal decay parameter is essential for evaluating the possible duration of aftershock sequences that occur after a mainshock. A higher value indicates a greater likelihood of aftershocks persisting for an extended duration, which is crucial for disaster preparedness and measures to reduce risks. Furthermore, it emphasizes the necessity for ongoing surveillance and analysis of aftershock sequences in order to precisely predict the progression of future aftershocks and inform response strategies. The parameter α indicates that the expected amount of aftershocks rises exponentially in relation to the magnitude of the mainshock. More precisely, for every one unit increase in magnitude, the productivity of aftershocks increases by a factor of around 5.75, as determined by the exponential function, $\exp(1.7511) \approx 5.75$. The standard error of 0.1716 suggests that there is possibility of a certain amount of variation, although the connection between magnitude and aftershock productivity is relatively well-established. D specifies the distribution's scale. It scales the spatial component of the distribution and acts as a distance parameter. A small value of D , shows that the events triggered are more confined to the vicinity of the main shock. In other words, the dispersion of events is more concentrated near the triggering event. In this scenario, the value of D is 0.7144 indicates that the likelihood of seeing triggered events diminishes at a faster rate as the distance from the primary shock increases. The estimate is linked with a standard error of 0.9431. A standard error signifies an uncertainty regarding the estimation of D . This implies that the true value of

D can exhibit some variation, and it is important to take caution when interpreting this outcome. The standard error indicates that acquiring more data or conducting further research is required to provide a more accurate estimation of aftershock triggering behavior. The variable q defines the form of the distribution and the rate at which the density declines as distance increases. It governs the characteristics of the distribution's tail. The higher parameter q indicates the possibility of a significant decrease in aftershock as the distance from the main shock increases. With a value of q equals 6.9006, the distribution exhibits a significant reduction, indicating that the occurrence of triggered events is less probable at a considerable distance from the primary shock. γ controls the impact of the primary shock's magnitude on the spatial distribution of triggered events. As the value of γ is higher it amplifies the impact of the magnitude on the distribution, The value of γ equals to 5.5429 indicates that the stronger primary earthquakes have a greater spatial impact, meaning that secondary earthquakes are more likely to happen at greater distances from the original earthquake as its magnitude grows.

	β	μ	A	c	α	p	D	q	γ
Estimate	1.4878	1.0114	0.0912	10.4336	1.7511	1.1559	0.7144	6.9006	5.5429
StdErr	0.0426	0.0506	0.3464	0.0737	0.1716	0.0068	0.9431	0.0963	0.0558

Table 4.3: ML estimates of model parameters are presented. β equals to 1.4878 indicates a rapid decay in the occurrence of larger magnitude target events in dataset. A μ background seismicity rate of 1.01 suggests steady background earthquake occurrences. A small earthquake (at threshold magnitude) might trigger roughly 0.0912 aftershocks. α equals to 1.7511 indicates that for each unit increase in earthquake magnitude, aftershocks increase by approximately 5.75 times. c and p are temporal decay parameters. With $c = 10.4336$ and $p = 1.1559$ the model suggests that aftershocks will keep happening for a while after the main shock, but they'll start to become less frequent at a moderate pace. D , γ and q are spatial decay parameters. With $D = 0.7144$, aftershocks are relatively close to the main event. With $\gamma=5.5429$, even a small increase in the magnitude will significantly expand the area where aftershocks might happen. With $q = 6.9006$, the probability of aftershocks decreases rapidly as move away from main shock.

The declustering probabilities derived from the ETAS (Epidemic-Type Aftershock Sequence) model applied to an earthquake dataset provides useful insights into the likelihood of each earthquake being classified as either an independent event or an aftershock. Probabilities approaching 1 (such as 0.9 or above) indicate a high likelihood that the earthquake will be independent. These seismic events have a significant probability of generating aftershocks. Probabilities approach-

ing 0 (such as 0.1 or less) indicate a high likelihood that the earthquake is an aftershock or triggered event. These earthquakes have a little probability of causing aftershocks and are more prone to being triggered by other seismic events. Probabilities within the range of 0 to 1 indicate that the earthquake exhibits a combination of independent and dependent characteristics. The precise interpretation relies on the particular probability value. The probability distribution suggests that a substantial proportion of the earthquakes are likely to be Independent events. More precisely, a minimum likelihood of 0.0113 implies that certain earthquakes are highly likely to be aftershocks, whilst a maximum probability of 1.0000 indicates that other earthquakes are highly likely to be separate events. The first quartile value of 0.6089 and the median value of 0.8705 indicate that more than half of the earthquakes are likely to be independent. Based on the data, earthquakes in this dataset typically have a 75.94% probability of being classified as background events. This indicates that the seismic activity in the region is greatly influenced by the baseline seismicity. Accurately differentiating between background events and aftershocks is crucial for exact evaluation of seismic risk and understanding the patterns of seismic activity in the Hindukush region. As shown in Table 4.4, the declustering probabilities vary between 0.0113 and 1.0000, with an average of 0.7594.

	Min	1st Qu.	Median	Mean	3rd Qu.	Max
Declustering Probabilities	0.0113	0.6089	0.8705	0.7594	0.9524	1.0000

Table 4.4: Declustering probabilities from the ETAS model fit indicates that the 75.94% earthquakes are classified as background events.

The figure 4.8 shows the plot of estimates of ETAS model. In first plot the rate of earthquakes classified as background events, which occur independently of other seismic activities. The middle part of the region has the highest background seismicity rate, as denoted by the red area. This implies that this particular location is most likely to be experiencing independent seismic occurrences. The rate of background seismicity diminishes as moves farther from the center, with blue areas denoting low rates. The overall spatial intensity encompasses both the underlying seismic activity and the additional impact from aftershocks induced by prior earthquakes. It quantifies the overall frequency of earthquakes at a particular location. The figure shows the combined effect of both the background seismic activity and the subsequent aftershock sequences throughout the whole region. High intensity zones are indicative of places that have high background rates and/or strong aftershock activity. The highest overall spatial seismicity rate is approximately 71 longitude and 36.5 latitude, similar to the background seismicity rate.

However, when aftershocks are included, this plot depicts somewhat larger or more intense areas. The clustering coefficient measures the extent to which earthquakes in a particular region are affected by preceding earthquakes. It provides information about the proportion of earthquakes that are aftershocks resulting from previous seismic events. High clustering coefficient regions in the figure denote locations where a high proportion of earthquakes are aftershocks occurring in the vicinity of earlier seismic events. On the other hand, areas with low clustering coefficients indicate places where the majority of earthquakes occur on their own and are not greatly impacted by past quakes. The conditional intensity function at the conclusion of the study period represents the anticipated frequency of earthquakes at location t , which is defined as $t + t_{start} + T$, where t_{start} is the start of the research period and T is the amount of time that passes after t_{start} . The plot of the conditional intensity function displays the regions with the highest probability of experiencing earthquakes near the conclusion of the study period. This plot, which takes into consideration both independent events and aftershocks, helps in determining the location and frequency of earthquakes anticipated to happen during that specific period. Given the history of all prior earthquakes, the conditional intensity function at the End of the Study period plot shows the instantaneous rate of earthquake occurrence at the end of the study period. High values (highlighted in red and yellow) indicate areas where there is a high probability of seismic activity at that particular time. With aftershocks and independent events taken into account, this plot offers a quick overview of the areas with the highest probability of experiencing earthquakes at the end of the research period.

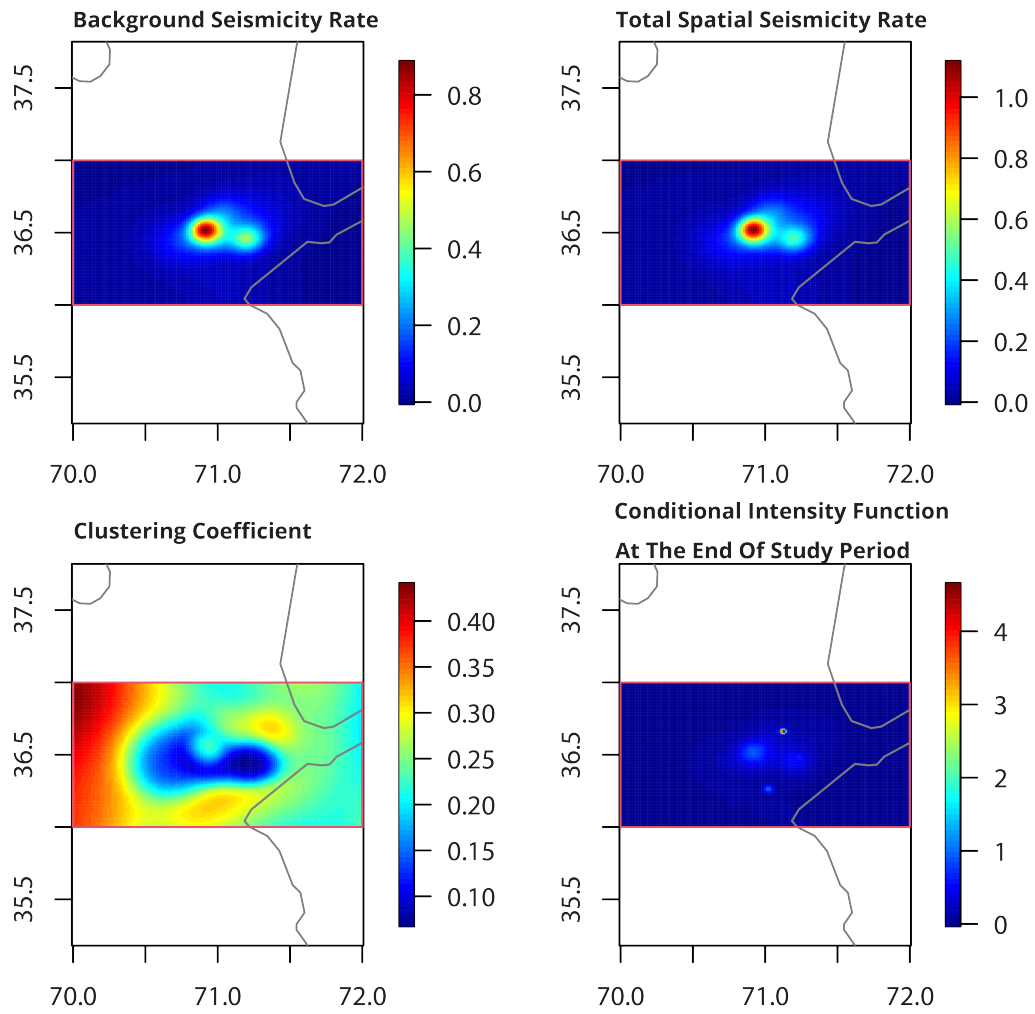


Figure 4.8: Plots of the fitted ETAS model's estimates of the background seismicity rate $\mu(x,y)$, total spatial intensity $\Lambda(x,y)$, clustering or triggering coefficient $\omega(x,y)$, and conditional intensity function $\lambda_\theta(x,y,t | H_t)$ to the Hindukush Catalog

Here 4.8, the data points nearly align with the red line, indicating that the model’s predictions are correct and that it effectively represents the occurrence rates of earthquakes. Nevertheless, there are certain deviations seen from the red line. These variations show discrepancies in the model’s accuracy, suggesting that the model may overestimate or underestimate occurrence rates in some regions or time periods. Although there are some variations, the overall positioning of the points along the red line suggests that the model is reliable in accurately representing the overall pattern of seismic events.

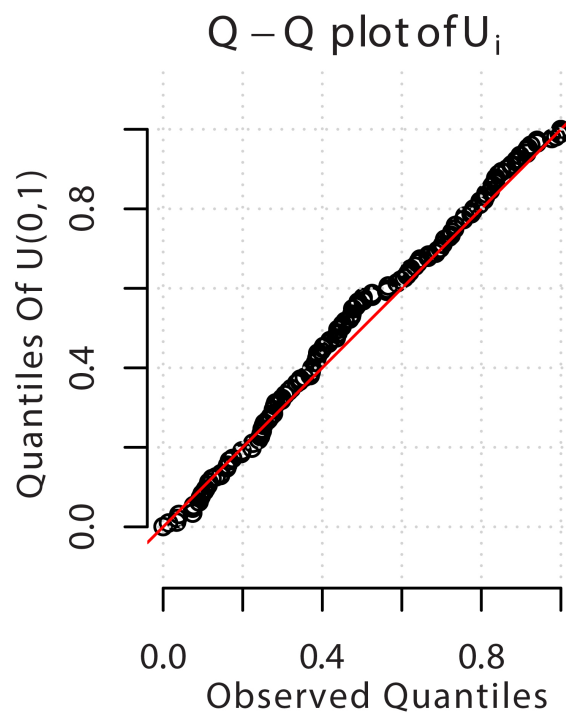


Figure 4.9: The Q-Q plot compares the observed quantiles of earthquake occurrences with a hypothesized uniform distribution ranging from 0 to 1. Points that closely align with the red line indicate a strong general fit of the model, while deviations reflect minor deviations in prediction accuracy.

The Q-Q plot of transformed time in 4.10 is a diagnostic tool used to evaluate the suitability of the ETAS model for the earthquake data. The close alignment between the black line and the red line suggests that the model well reflects the time of earthquakes, whilst deviations indicate places where the model may require more improvement. This analysis facilitates comprehension

of the efficacy of the ETAS model in forecasting earthquake events and emphasizes potential improvements for better precision.

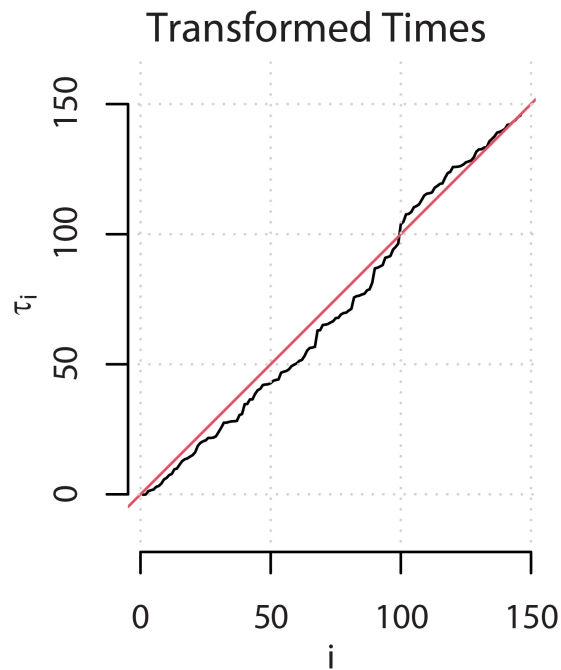


Figure 4.10: A Q-Q plot is generated for the transformed times using the ETAS model. It compares the ordered transformed times to the expected quantiles of a standard uniform distribution. The close alignment between the black line and the red 45-degree line suggests that the model fits the earthquake data well.

CHAPTER 5

Conclusions

This thesis provides in depth analysis of earthquake properties in the Hindukush region, with a specific emphasis on the relationships between magnitude, depth, and time. The study shows that earthquakes occurring at depths of 0-70 km predominantly consist of magnitudes ranging from 1 to 5, whereas earthquakes at depths of 70-280 km have a broader range of magnitudes, from 1.5 to 7. Significantly, no seismic events were detected beyond a depth of 280 km, suggesting a lack of deep earthquakes in the area under analysis. The seismic data shows that there were 6,274 shallow earthquakes and 3,258 intermediate events, showing a greater occurrence of shallow quakes. The research revealed a notable rise in seismic activity in 2009, namely in the magnitude range of 1.8-3.4, with a peak occurrence around 2.5 on the frequency-magnitude scale. For declustering two distinct earthquake catalogues were created for both methods to identify the dependent and independent events. The declustering technique, employing the Gardner and Knopoff approach it was determined if an earthquake of $M=4.5$ occurs within the epicentral distance of 35 km and a subsequent earthquake of the same magnitude occurs within the next 83 days, it will be considered an aftershock of the earlier large earthquake. The Gardner and Knopoff method analyzed 9,532 events (magnitude ≥ 3) from August 1, 2008 to July 13, 2010, identifying 3574 mainshock events accounting for 37.5 % of the mainshocks and the remaining events (62.5 %) were eliminated as aftershocks or foreshocks. The magnitude of completeness (M_c) was found from maximum curvature method value of 3.03, earthquakes of magnitudes equal to or larger than this value can be successfully recognized and recorded. For earthquake hazard assessment, the Gutenberg-Richter relationship is important. Log-linear regression modeling provided productivity rate and slope, $a = 5.21$ and $b = 0.98$ respectively for the region's seismotectonic level. The correlation between a and b values highlighting the region's different seismic risks. The ETAS model studied 36° - 37° N latitudes and 70° - 72° E

longitudes with a magnitude threshold of 4 from August 1, 2008, to January 1, 2010 with 146 target events out of 324. Analysis confirmed the catalog's completeness and stationarity as indicated by the linear correlation revealed in the plots of $\log_{10}(N_m)$ versus m and N_t versus t . The log-likelihood and AIC values confirm that the model accurately represents the observed seismicity. The Euclidean parameters β and vector $\theta = (\mu, A, \alpha, c, p, D, \gamma, q)$ offer earthquake dynamics insights. The β value of 1.4878 indicates fewer larger earthquakes, whereas the background rate $\mu = 1.01$ shows a continuous seismic activity. The parameters $c = 10.4336$ and $p = 1.1559$ indicate a slight drop in aftershock frequency over time at moderate rate. Parameter $A = 0.0912$ indicates that the a small events at the threshold magnitude might trigger roughly 0.0912 aftershocks, whereas α indicates increase in aftershock productivity by 5.75 times per unit increase in magnitude. The parameters $D = 0.7144$ and $q = 6.9006$ show that aftershocks are concentrated and diminish with distance, whereas $\gamma = 5.5429$ enhances the spatial impact of stronger earthquakes. These findings contribute to our comprehension of seismic hazards and assist in evaluating risks. Using the ETAS model's declustering probabilities, over 75% of earthquakes in the dataset are background events and the remaining events are triggered, indicating significant Hindukush seismic activity. Background activity (Independent Events) is most at the region's core. Total spatial seismicity rate indicates mostly background and triggered events are in the center indicated by red spot. While the high clustering coefficients in red zone suggest triggered events and blue area indicates isolated events. The conditional intensity function indicates places with strong seismic activity at the research's end, indicating the complicated interaction between baseline seismicity and aftershocks. The model predicts accurately in areas with residuals approaching zero. In general, the Q-Q plot fits well, although some deviations show probable inaccuracies in specific places or time frames. A transformed temporal Q-Q plot shows that the model accurately predicts earthquake timing, but variances offer opportunity for improvement.

CHAPTER 6

Limitations

The Gardner and Knopoff technique and Epidemic-Type Aftershock Sequence (ETAS) model have limitations for earthquake modeling. Gardner and Knopoff technique excludes higher-order aftershocks, which may exclude critical secondary aftershocks. This method considers a fixed space and time distance window which may not necessarily adapt to aftershock sequence variations across different regions. On the other hand, ETAS model is computationally difficult, time-consuming, and delicate, especially for large earthquake catalogues that require many computational resources. Another problem is the assumption that the seismic catalog is comprehensive and stationary. If the catalog is inadequate or there is non-stationarity like seasonal or growing trends, the model may be unreliable. It's also sensitive to parameter estimates. A and D 's estimated standard errors are bigger than their estimates, indicating that the fitted model cannot adequately capture this region's complicated earthquake clustering. Insufficient data (small catalog) or model inadequacy may be the cause of this. The ongoing collision between Indian Plate and Eurasian Plate is making Hindukush most seismically active region. Thus, the ETAS model with fixed parameters may not be suitable for modeling earthquakes in an active tectonic zone. Enhanced ETAS models with regionally changing parameters may be better for this location.

Bibliography

- [1] B. Gutenberg and C. F. Richter. “Seismicity of the Earth and Associated Phenomena”. In: *Princeton University Press* 2 (1954). DOI: [10.1111/j.2153-3490.1950.tb00313.x](https://doi.org/10.1111/j.2153-3490.1950.tb00313.x).
- [2] K. Mogi. “Study of elastic shocks caused by the fracture of heterogeneous materials and its relations to earthquake phenomena”. In: *Bulletin of the Earthquake Research Institute, University of Tokyo* 40 (1962), pp. 125–173.
- [3] K. Aki. “Maximum likelihood estimate of b in the formula $\log N=a-bM$ and its confidence limits”. In: *Bulletin of the Earthquake Research Institute, University of Tokyo* 43 (1965), pp. 237–239. DOI: [10.1007/BF00881745](https://doi.org/10.1007/BF00881745).
- [4] N.W. Warren and G.V. Latham. “An experimental study of thermally induced microfracturing and its relation to volcanic seismicity”. In: *Journal of Geophysical Research* 75.23 (1970), pp. 4455–4464.
- [5] L. Knopo and J. Gardner. “Higher Seismic Activity During Local Night on the Raw Worldwide Earthquake Catalogue”. In: *Geophysical Journal of the Royal Astronomical Society* 28 (1972), pp. 311–313.
- [6] S.J. Gibowicz. “Variation of the frequency-magnitude relation during earthquake sequences in New Zealand”. In: *Bulletin of the Seismological Society of America* 63 (1973), pp. 517–528. DOI: [10.1785/BSSA0630020517](https://doi.org/10.1785/BSSA0630020517).
- [7] M. Wyss. “Towards a physical understanding of the earthquake frequency distribution”. In: *Geophysical Journal International* 31.4 (1973), pp. 341–359.
- [8] J.K. Gardner and L. Knopoff. “Is the sequence of earthquakes in southern California, with aftershocks removed Poissonian?” In: *Bulletin of the Seismological Society of America* 64 (1974), pp. 1363–1367. DOI: [10.1785/BSSA0640051363](https://doi.org/10.1785/BSSA0640051363).

- [9] S. Billington, B. L. Isacks, and M. Barazangi. “Spatial distribution and focal mechanisms of mantle earthquakes in the Hindu Kush–Pamir region: a contorted Benioff zone”. In: *Geology* 5.11 (1977), pp. 699–704.
- [10] H. Kanamori. “The Energy Release in Great Earthquakes”. In: *Journal of Geophysical Research* (1977).
- [11] J.-L. Chatelain et al. “Microearthquake seismicity and fault plane solutions in the Hindu Kush region and their tectonic implications”. In: *Journal of Geophysical Research: Solid Earth* 85.B3 (1980), pp. 1365–1387.
- [12] R.K. Verma, M. Mukhopadhyay, and A.K. Bhanja. “Seismotectonics of the Hindukush and Baluchistan arc”. In: *Tectonophysics* 66.4 (1980), pp. 301–322. ISSN: 0040-1951. DOI: [https://doi.org/10.1016/0040-1951\(80\)90247-4](https://doi.org/10.1016/0040-1951(80)90247-4).
- [13] B. Bender. “Maximum likelihood estimation of b values for magnitude-grouped data”. In: *Bulletin of the Seismological Society of America* 73 (1983), pp. 831–851. DOI: [10.1785/BSSA0730030831](https://doi.org/10.1785/BSSA0730030831).
- [14] B. Bender. “Maximum likelihood estimation of b values for magnitude-grouped data”. In: *Bulletin of the Seismological Society of America* 73 (1983), pp. 831–851. DOI: [10.1785/BSSA0730030831](https://doi.org/10.1785/BSSA0730030831).
- [15] H. Kanamori. “Magnitude scale and quantification of earthquakes”. In: *Tectonophysics* 93 (1983), pp. 185–199. DOI: [10.1016/0040-1951\(83\)90273-1](https://doi.org/10.1016/0040-1951(83)90273-1).
- [16] A. Farah et al. “Evolution of the lithosphere in Pakistan”. In: *Tectonophysics* 105.14 (1984), pp. 207–227.
- [17] P. Reasenber. “Second-Order Moment of Central California Seismicity, 1969-82”. In: *Journal of Geophysical Research: Solid Earth* 90.B7 (1985), pp. 5479–5495. DOI: [10.1029/JB090iB07p05479](https://doi.org/10.1029/JB090iB07p05479).
- [18] R. Uhrhammer. “Characteristics of Northern and Southern California Seismicity”. In: *Earthquake Notes* 57.1 (1986), p. 21.
- [19] K. N. Khattri. “Great earthquakes, seismicity gaps and potential for earthquake disaster along the Himalayan plate boundary”. In: *Tectonophysics* 138.1 (1987), pp. 79–92.
- [20] Yosihiko Ogata. “Statistical models for earthquake occurrences and residual analysis for point processes”. In: *Journal of the American Statistical Association* 83.401 (1988), pp. 9–27. DOI: [10.1080/01621459.1988.10478560](https://doi.org/10.1080/01621459.1988.10478560).

- [21] Y. Kagan and D. Jackson. “Long-term earthquake clustering”. In: *Geophysical Journal International* 104 (1991), pp. 117–133.
- [22] Yury Y. Kagan. “Likelihood Analysis of Earthquake Catalogues”. In: *Geophysical Journal International* 106.1 (1991), pp. 135–148. DOI: [10.1111/j.1365-246X.1991.tb04607.x](https://doi.org/10.1111/j.1365-246X.1991.tb04607.x).
- [23] M. W. Hamburger et al. “Structural and seismic evidence for intracontinental subduction in the Peter the First Range, central Asia”. In: *Geological Society of America Bulletin* 104.4 (1992), pp. 397–408.
- [24] K. Dahmen, D. Erta, and Y. Ben-Zion. “Gutenberg–Richter and characteristic earthquake behavior in simple mean-field models of heterogeneous faults”. In: *Physical Review E* 58.2 (1998), p. 1494.
- [25] Yosihiko Ogata. “Seismicity Analysis through Point-Process Modeling: A Review”. In: *Pure and Applied Geophysics* 155.2-4 (1999), pp. 471–507. DOI: [10.1007/s000240050275](https://doi.org/10.1007/s000240050275).
- [26] I. G. Main. “Apparent breaks in scaling in the earthquake cumulative frequency-magnitude distribution: Factor artifact?” In: *Bulletin of the Seismological Society of America* 90 (2000), pp. 86–97. DOI: [10.1785/0119990086](https://doi.org/10.1785/0119990086).
- [27] G. L. Pavlis and S. Das. “The Pamir–Hindu Kush seismic zone as a strain marker for flow in the upper mantle”. In: *Tectonics* 19.1 (2000), pp. 103–115.
- [28] M. Searle, B. R. Hacker, and R. Bilham. “The Hindu Kush seismic zone as a paradigm for the creation of ultrahigh-pressure diamond-and coesite-bearing continental rocks”. In: *Journal of Geology* 109.2 (2001), pp. 143–153.
- [29] Stefan Wiemer. “A Software Package to Analyze Seismicity: ZMAP”. In: *Seismological Research Letters* 72 (2001), pp. 373–382. DOI: [10.1785/gssr1.72.3.373](https://doi.org/10.1785/gssr1.72.3.373).
- [30] Jiancang Zhuang, Yosihiko Ogata, and David Vere-Jones. “Stochastic Declustering of Space-Time Earthquake Occurrences”. In: *Journal of the American Statistical Association* 97 (Feb. 2002), pp. 369–380. DOI: [10.1198/016214502760046925](https://doi.org/10.1198/016214502760046925).
- [31] David Vere-Jones. “Probabilities and Information Gain for Earthquake Forecasting”. In: *Selected Papers From Volume 30 of Vychislitel'naya Seysmologiya* (2003), pp. 104–114. DOI: [10.1029/cs005p0104](https://doi.org/10.1029/cs005p0104).

- [32] Ivan Koulakov and Stephan V. Sobolev. “A tomographic image of Indian lithosphere break-off beneath the Pamir–Hindukush region”. In: *Geophysical Journal International* 164.2 (2006), pp. 425–440.
- [33] Jiancang Zhuang, Yosihiko Ogata, and David Vere-Jones. “Diagnostic Analysis of Space-Time Branching Processes for Earthquakes”. In: *Case Studies in Spatial Point Process Modeling*. Ed. by Adrian Baddeley et al. Springer-Verlag, 2006, pp. 275–292. DOI: [10.1007/0-387-31144-0_15](https://doi.org/10.1007/0-387-31144-0_15).
- [34] J. E. Ebel et al. “Non-poissonian earthquake clustering and the hidden Markov model as bases for earthquake forecasting in California”. In: *Seismological Research Letters* 78.1 (2007), pp. 57–65.
- [35] Y. Y. Kagan, D. D. Jackson, and Y. Rong. “A testable five-year forecast of moderate and large earthquakes in southern California based on smoothed seismicity”. In: *Seismological Research Letters* 78.1 (2007), pp. 94–98.
- [36] M. D. Petersen et al. “Time-independent and time-dependent seismic hazard assessment for the State of California: uniform California Earthquake Rupture Forecast Model 1.0”. In: *Seismological Research Letters* 78.1 (2007), pp. 99–109.
- [37] Z.-K. Shen, D. D. Jackson, and Y. Y. Kagan. “Implications of geodetic strain rate for future earthquakes, with a five-year forecast of m5 earthquakes in southern California”. In: *Seismological Research Letters* 78.1 (2007), pp. 116–120.
- [38] I. Zaliapin et al. “Clustering analysis of seismicity and aftershock identification”. In: *Phys. Rev. Lett.* 101.1 (2008), pp. 22–23. DOI: [10.1103/PhysRevLett.101.018501](https://doi.org/10.1103/PhysRevLett.101.018501).
- [39] G. Grünthal, R. Wahlström, and D. Stromeyer. “The Unified Catalogue of Earthquakes in Central, Northern, and Northwestern Europe (CENEC) – Updated and Expanded to the Last Millennium”. In: *Journal of Seismology* 13 (2009), pp. 517–541. DOI: [10.1007/s10950-008-9144-9](https://doi.org/10.1007/s10950-008-9144-9).
- [40] R. Console, D. D. Jackson, and Y. Y. Kagan. “Using the ETAS Model for Catalog Declustering and Seismic Background Assessment”. In: *Pure and Applied Geophysics* 167.6 (2010), pp. 819–830.
- [41] M. Usman and M. Zafar. “Effects of temperature increase on earthquake frequency and depth in Northern Pakistan”. In: *International Conference on Biology, Environment and Chemistry*. Vol. 1.1. IPCBEE. Singapore: IACSIT Press, 2010, p. 2011.

- [42] G. Ananthachari Preethi, Assistant Professor, and B. Santhi. “Study on Techniques of Earthquake Prediction”. In: *International Journal of Computer Applications* 29 (2011), pp. 55–58.
- [43] S. Gupta et al. “Statistical analysis of completeness of earthquake data around Dehradun city and its implication of seismicity evaluation”. In: (2012).
- [44] B. Luen and P. B. Stark. “Poisson Tests of Declustered Catalogues”. In: *Geophysical Journal International* 189.1 (2012), pp. 691–700. DOI: [10.1111/j.1365-246X.2012.05400.x](https://doi.org/10.1111/j.1365-246X.2012.05400.x).
- [45] A. Mignan and Jochen Woessner. “Estimating the magnitude of completeness for earthquake catalogs”. In: *Community Online Resource for Statistical Seismicity Analysis* (Apr. 2012). DOI: [10.5078/corssa-00180805](https://doi.org/10.5078/corssa-00180805).
- [46] T. Van Stiphout, J. Zhuang, and D. Marsan. “Seismicity Declustering”. In: *Community Online Resource for Statistical Seismicity Analysis* (2012). February 1–25. DOI: [10.5078/corssa-52382934](https://doi.org/10.5078/corssa-52382934).
- [47] T. Van Stiphout, T. J. Zhuang, and D. Marsan. “Seismicity declustering”. In: *Community Online Resource for Statistical Seismicity Analysis* (2012). DOI: [10.5078/CORSSA-52382934](https://doi.org/10.5078/CORSSA-52382934).
- [48] K. Rehman et al. “Spatio-temporal variation of b value in and around the north Pakistan”. In: *Journal of Earth System Science* 124.7 (2015), pp. 1445–1456. DOI: [10.1007/s12040-015-0625-2](https://doi.org/10.1007/s12040-015-0625-2).
- [49] M.S. Ali. “Statistical analysis of seismicity in Egypt and its surroundings”. In: *Arab. J. Geosci.* 9 (2016), p. 52. DOI: [10.1007/s12517-015-2079-x](https://doi.org/10.1007/s12517-015-2079-x).
- [50] Arum Handini Primandari and Khusnul Khotimah. “Seismic analysis using maximum likelihood of Gutenberg-Richter”. In: *Bulletin of Social Informatics Theory and Application* 1.1 (2017), pp. 34–40.
- [51] P. Senartorski. “Effect of slip-area scaling on the earthquake frequency–magnitude relationship”. In: *Physics of the Earth and Planetary Interiors* 267 (2017), pp. 41–52. DOI: [10.1016/j.pepi.2017.04.004](https://doi.org/10.1016/j.pepi.2017.04.004).
- [52] Khawaja M. Asim et al. “Earthquake prediction model using support vector regressor and hybrid neural networks”. In: *PLOS ONE* 13 (July 2018), pp. 1–22. DOI: [10.1371/journal.pone.0199004](https://doi.org/10.1371/journal.pone.0199004).

- [53] Nader Davoudi et al. “Declustering of Iran earthquake catalog (1983–2017) using the epidemic-type aftershock sequence (ETAS) model”. In: *Acta Geophysica* 66.6 (2018), pp. 1359–1373. ISSN: 1895-7455. DOI: [10.1007/s11600-018-0211-5](https://doi.org/10.1007/s11600-018-0211-5).
- [54] M.M. Khan and G.K. Kumar. “Statistical completeness analysis of seismic data”. In: *Journal of the Geological Society of India* 19 (2018), pp. 749–753. DOI: [10.1007/s12594-018-0934-6](https://doi.org/10.1007/s12594-018-0934-6).
- [55] Shivamant Angadi et al. “Development of Gutenberg-Richter Recurrence Relationship Using Earthquake Data”. In: *Green Buildings and Sustainable Engineering*. 2019, pp. 281–288.
- [56] F.V. Pisarenko and V.M. Rodkin. “Declustering of the seismicity flow. Statistical analysis”. In: *Phys. Solid Earth* 55.5 (2019), pp. 773–745.
- [57] P. Senartorski. “Gutenberg–Richter’s b value and earthquake asperity models”. In: *Pure and Applied Geophysics* 177 (2020), pp. 1891–1905. DOI: [10.1007/s00024-019-02385-z](https://doi.org/10.1007/s00024-019-02385-z).
- [58] S. Khurram et al. “Assessment of seismic hazard of roller compacted concrete dam site in Gilgit-Baltistan of northern Pakistan”. In: *Earthquake Engineering and Engineering Vibration* 20 (2021), pp. 621–630. DOI: [10.1007/s11803-021-2042-7](https://doi.org/10.1007/s11803-021-2042-7).
- [59] Shahzada Khurram and Pervez Khalid. “Statistical Analysis of Earthquake Catalogue and Magnitude of Completeness for Northern and Southwestern Pakistan”. In: *Acta Geodyn. Geomater.* 19.1 (2021), pp. 23–34.
- [60] Leila Mizrahi, Shyam Nandan, and Stefan Wiemer. “The Effect of Declustering on the Size Distribution of Mainshocks”. In: *Seismological Research Letters* 92.4 (Feb. 2021), pp. 2333–2342. ISSN: 0895-0695. DOI: [10.1785/0220200231](https://doi.org/10.1785/0220200231).
- [61] Zahid Rahman et al. “Re-appraisal of earthquake catalog in the Pamir—Hindu Kush region, emphasizing the early and modern instrumental earthquake events”. In: *Journal of Seismology* 25.6 (2021), pp. 1461–1481. ISSN: 1573-157X. DOI: [10.1007/s10950-021-10046-9](https://doi.org/10.1007/s10950-021-10046-9).
- [62] M. Taroni and A. Akinici. “Good practices in PSHA, Declustering, b value estimation, foreshocks and aftershocks. A case study in Italy”. In: *Geophysical Journal International* 224 (2021), pp. 1174–1187. DOI: [10.1093/gji/ggaa462](https://doi.org/10.1093/gji/ggaa462).

- [63] M. Taroni and A. Akinci. “Good practices in PSHA, Declustering, b value estimation, foreshocks and aftershocks. A case study in Italy”. In: *Geophysical Journal International* 224 (2021), pp. 1174–1187. DOI: [10.1093/gji/ggaa462](https://doi.org/10.1093/gji/ggaa462).
- [64] Pawan Chhetri, Deepak Chamlagain, and Ram Chandra Tiwari. “Effect of Declustering Algorithm in Probabilistic Seismic Hazard Analysis for Kathmandu Valley”. In: *Journal of Advanced College of Engineering and Management* 7.01 (2022), pp. 107–119. DOI: [10.3126/jacem.v7i01.47337](https://doi.org/10.3126/jacem.v7i01.47337).
- [65] Shahzada Khurram and Pervez Khalid. “Comparison of Different Declustering Procedures Using Pakistan Earthquake Dataset”. In: *The Geological Bulletin of the Punjab University* 55 (2022).
- [66] Davood Kazemi Lafmejadi and Elham Shabani. “Modeling Earthquake Data Using ETAS Model to Forecast Aftershock Subsequences Applying Different Parameterizations in Kermanshah Region, Iran”. In: *PREPRINT (Version 1) available at Research Square* (2022). DOI: [10.21203/rs.3.rs-1996634/v1](https://doi.org/10.21203/rs.3.rs-1996634/v1).
- [67] Shofiyatul Mahmudah and Zainul Arifin Imam Supardi. “ANALISIS PARAMETER KEGEMPAAN (NILAI-a DAN NILAI-b) DI WILAYAH BUSUR BANDA”. In: *JURNAL RISET RUMPUN MATEMATIKA DAN ILMU PENGETAHUAN ALAM* 2.1 (2023), pp. 254–262.