



Analysis of depth and magnitude heterogeneity in cluster-based segmentation of Andaman and Nicobar earthquake data

DEBASRI SAMANTA^{1*} and RAJIB KUMAR DOLAI²

¹Department of Physics, Vidyasagar University, Midnapore, West Bengal, India

²Department of Economics, Tamralipta Mahavidyalaya, Tamluk, West Bengal, India

(Received 8 March 2024, Accepted 07 July 2026)

*Corresponding author's email: debasri94s@gmail.com

सार – यह अध्ययन भूकंप की गहराई और परिमाण के पावर स्पेक्ट्रल घनत्व (Power Spectral Density–PSD) तथा प्रायिकता घनत्व फलन (Probability Density Function–PDF) विश्लेषण के साथ सिल्हूट स्कोर (silhouette scores) द्वारा अनुकूलित के-माध्य क्लस्टरिंग का उपयोग करते हुए अंडमान और निकोबार क्षेत्र की भूकंपीय विशेषताओं का अध्ययन करता है। अध्ययन में पाँच भूकंपीय क्लस्टर (समूह) की पहचान की गई, जो क्षेत्रीय विवर्तनिक संरचना से निकटता से संबद्ध हैं। PSD विश्लेषण से विभिन्न पैमानों पर समय-आधारित परिवर्तनशीलता का संकेत मिलता है। क्लस्टर 1, 2 और 4 में निम्न-आवृत्ति घटकों का प्रभुत्व पाया गया, जबकि क्लस्टर 3 और 5 में व्यापक स्पेक्ट्रमी वितरण तथा उच्च-आवृत्ति घटकों का अपेक्षाकृत अधिक योगदान देखा गया। स्पेक्ट्रल एज फ्रीक्वेंसी (Spectral Edge Frequency–SEF) के मानों से संकेत मिलता है कि गहराई और परिमाण की समय-श्रृंखलाओं में 95% विचरण इस आवृत्ति-परास से नीचे निहित है। PDF विश्लेषण से ज्ञात हुआ कि क्लस्टर 1, 2 और 4 में गहराई का वितरण द्विविधारी (bimodal) है, जिसमें तथा के आसपास अधिकतम सांद्रता पाई गई, जबकि क्लस्टर 3 और 5 में के निकट गहराई की स्थिर सांद्रता प्रमुख रही। परिमाण की सीमाओं में वृद्धि के साथ परिमाण का वितरण क्रमशः उच्चतर मानों की ओर स्थानांतरित होता है तथा उसका विस्तार बढ़ता जाता है। परिणामों से संकेत मिलता है कि भूकंप की गहराई मुख्यतः स्थायी विवर्तनिक संरचनाओं द्वारा नियंत्रित होती है, जबकि भूकंप के परिमाण में परिवर्तन प्रतिबल के संचय तथा भंश-विफलन (rupture) की प्रक्रियाओं के प्रति प्रतिक्रिया के रूप में होता है। अध्ययन की सीमाओं में कैटलॉग-आधारित आँकड़ों का उपयोग, परिमाण की नियत सीमाएँ तथा प्रत्यक्ष भू-आकृतिक अथवा संरचनात्मक बाधयताओं के बिना सांख्यिकीय व्याख्याओं पर निर्भरता शामिल हैं।

ABSTRACT. This study investigates the seismic characteristics of the Andaman and Nicobar region using Power Spectral Density (PSD) and Probability Density Function (PDF) analyses of earthquake depth and magnitude, combined with K-means clustering optimized by silhouette scores. Five seismic clusters were identified that correspond closely to the regional tectonic framework. PSD analysis indicates multi-scale temporal variability, with Clusters 1, 2, and 4 dominated by low-frequency components and Clusters 3 and 5 showing broader spectral distributions with stronger high-frequency contributions. Spectral Edge Frequency (SEF) values of 0.90–0.96 day⁻¹ indicate that 95% of the variance in the depth and magnitude time series is contained below this frequency range. PDF analysis shows that Clusters 1, 2, and 4 have bimodal depth distributions with concentrations at ~8–12 km and ~28–35 km, whereas Clusters 3 and 5 are dominated by a stable depth concentration near ~30–35 km. Magnitude distributions progressively shift toward higher values and broaden with increasing magnitude thresholds. The results indicate that earthquake depths are controlled by persistent tectonic structures, while earthquake magnitudes vary in response to stress accumulation and rupture processes. Limitations include the use of catalog-based data, fixed magnitude thresholds, and statistical interpretations without direct geodetic or structural constraints.

Key words – Cluster, Probability Density Function, Power Spectral Density, Dominant Frequencies.

1. Introduction

The movement of tectonic plates beneath the surface of the Earth causes abrupt and powerful shaking of the

ground, known as earthquakes. This movement releases energy in the form of seismic waves, which propagate outward from the source. A fault line, typically a crack or weak spot in the Earth's crust where movement has

occurred or may occur in the future, is associated with plate tectonics and is usually found at boundaries between tectonic plates. When stress builds up along a fault due to the movement of tectonic plates, it can eventually overcome the frictional forces holding the rocks together, causing earthquakes. The analysis of earthquake data plays a crucial role in understanding the seismic behaviour. In the Andaman and Nicobar Islands, where seismic events are frequent, segmenting and studying seismic data can shed light on crucial insights. The Andaman-Nicobar Seismic Region formed as a result of the northward movement of the island plate and the subsequent reverse rotation of the ridge thrust beneath the Asian plate. A depression or forearc basin emerged between the Outer High Arc and the volcanic arc as subduction advanced and the Outer High Arc complex rose sharply (see fig-1). The Andaman Nicobar Ridge emerged as an island chain throughout time as a result of constant pressure. Naturally, the Andaman and Nicobar Islands Seismic Region is characterized by numerous major and minor faults, which may exhibit asymmetry in energy release. This phenomenon leads to the division of the Seismic Region into several clusters, highlighting the necessity for cluster-based analysis wherever it can be observed. Cluster analysis is a statistical technique that groups observations in a dataset based on similarities. It helps identify patterns by organizing seismic events according to their proximity and unique features. Asymmetry can also be observed between magnitude and depth, which depends on various factors such as the release of energy, characteristics of the Earth's crust, geographical region, presence of faults, and more. Each seismic event may be considered as a discrete random variable with its own probability distribution. The PDF of magnitude and depth of an earthquake represents the likelihood of observing particular values, describing the distribution of earthquake magnitude and depth within a certain region over a period of time. Therefore, the need for probability density function (PDF) analysis becomes apparent in order to properly understand seismic events in this context. PSD is a tool that assesses the performance of earthquake monitoring systems. It investigates how power is distributed across different frequencies in seismic signals relative to their magnitude and depth. By exploring the PSD plot, the study aims to gain valuable insights into the nature and behaviour of earthquake events, particularly with regard to their frequency content and characteristics. For estimating PSD, Peter (1967) introduced a method named Welch's method, which involves transforming data from a time domain to frequency domain. However, MacQueen (1967) introduced another method called K-means clustering, which is more suitable for dividing data into cluster segments. Later, it was modified by Hartigan and Wong (1979). However, identifying the optimal number of

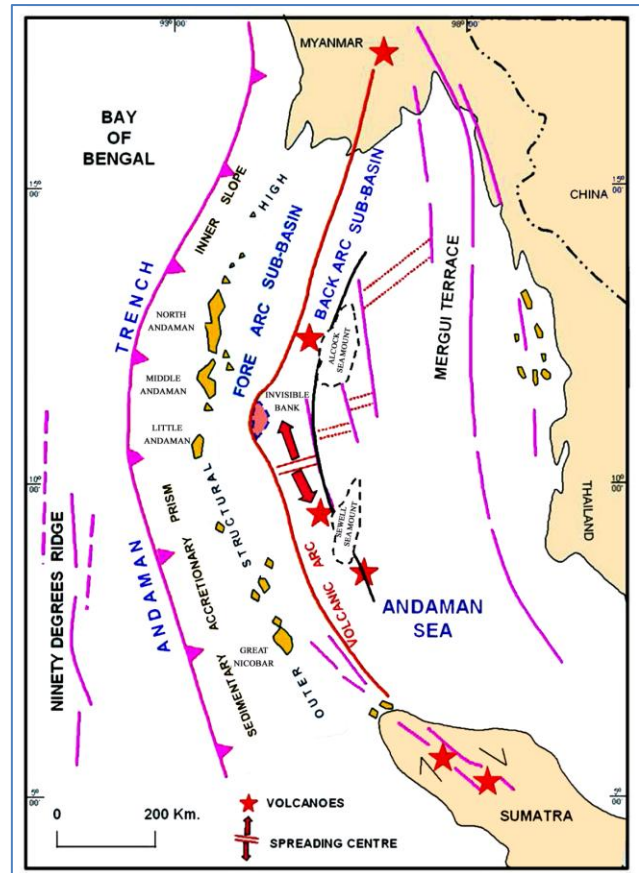


Fig. 1. Seismic Region of Andaman and Nicobar Islands

clusters based on their characteristics is challenging. Rousseeuw (1987) attempted to address this challenge by developing an algorithm called silhouette score. Christian (1989) emphasizes the importance of the power spectral density function in earthquake motion analysis for engineering design, addressing challenges in its computation from response spectra and introducing integration schemes within the SPEED program. Similarly, Manohar and Sarkar (1995) propose innovative methods for identifying critical earthquake excitations in linear systems using linear programming and calculus of variations. While successful in setting upper response limits, these methods lack the ability to capture the inherent unpredictability of actual earthquake data, leading them to incorporate entropy rates for quantifying uncertainty and adjusting critical excitation definitions accordingly. Although Rao and Rao. (1979) use Gumbel's extreme value theory to estimate the probability of occurrence and return periods in the Andaman and Nicobar Islands. Additionally, Anthony *et al.* (2020) discuss the obstacles in obtaining accurate power spectral density estimates for seismic background noise, highlighting the influence of processing techniques and the impact of teleseismic earthquakes on ambient-noise

estimates. Meanwhile, Yoshimitsu *et al.* (2023) Utilized the F-distribution as a probability density function for seismic source parameter estimation, improving estimate quality with increased iterations and reducing variation in stress drop estimates within a cluster. Sadeghian (2018) also uses PDF analysis and observes higher accuracy than the gamma, beta, and other more well-known methods. Kolathayar and Sitharam (2012) employed a probabilistic seismic hazard analysis, using a homogenized earthquake catalogue and identifying 11 seismic source zones to address local variability. Kataria *et al.* (2013) performed a deterministic seismic hazard assessment, dividing the region into seven seismogenic zones and providing insights into ground motion potential through peak ground acceleration (PGA) and pseudo-spectral acceleration (PSA) estimates. Mishra *et al.* (2007) analyzed aftershocks following the 2004 Sumatra-Andaman earthquake, highlighting the region's complex seismotectonic patterns. Dewangan *et al.* (2021) focused on the ambient noise characteristics in the Andaman region, while Rahayu *et al.* (2020) tested new seismic sensors for waveform detection. D'Alessandro *et al.* (2021) assessed seismic noise variability across the Italian Seismic Network, contributing to a new reference model for seismic noise. In contrast, Samanta and Dolai (2025) applied advanced waveform analysis techniques, including FFT, PSD, CWT, and PDF, and found that the Mw 7.7 Myanmar earthquake generated stronger, longer-lasting low-frequency ground motion and posed a greater structural hazard than the Mw 6.7 event.

In this paper, the study region is logically divided into clusters. Subsequently, depth and magnitude power spectral density analyses are conducted. Finally, probability density function analysis is performed. The characteristics of each cluster are identified and analyzed to gain insights into seismic activity.

2. Data and methodology

2.1. Silhouette score

In determining the optimal number of clusters for our study dataset, we employed the silhouette score method. The silhouette score (1987), calculated using the formula $s = (b - a) / \max(a, b)$, is a metric used to evaluate the efficacy of clustering algorithms in unsupervised machine learning. In the formula, 'a' represents the mean distance between a sample and all other points in the same cluster, while 'b' represents the mean distance between a sample and all others. The denominator, $\max(a, b)$, ensures that the silhouette score is normalized between -1 and 1. A high silhouette score close to 1 indicates that the data point is well-matched to its own cluster and poorly matched to samples in other clusters. Conversely, a low

score near -1 indicated that the sample may have been assigned to the wrong cluster, as it is more similar to samples in neighbouring clusters. A score around 0 implies that the sample lies on or near the decision boundary between clusters.

2.2. K-means clustering

Here, we employed K-means clustering to partition the dataset into the optimal number of clusters determined using the silhouette score method. The K-means clustering algorithm is based on a mathematical formulation that aims to minimize the within-cluster variance and maximize the between-cluster variance. Given a dataset with N data points represented as vectors $X = \{x_1, x_2, \dots, x_N\}$, where each data point x_i is a d -dimensional vector, the algorithm begins by randomly initializing K cluster centroids $\mu = \{\mu_1, \mu_2, \dots, \mu_K\}$, where each centroid μ_j is also a d -dimensional vector. The algorithm iteratively performs two steps: the assignment step and the update step. In the assignment step, each data point x_i is assigned to the cluster with the nearest centroid based on the Euclidean distance or any other chosen distance metric. The assignment process determines the minimum distance between x_i and the centroids. In the update step, the centroids of the clusters are recalculated by taking the mean of all the data points assigned to each cluster. The new centroid for cluster j , denoted as μ_j' , is computed as the average of the data points x_i that belong to cluster j . This process of assignment and update is repeated until convergence, which is typically defined by a maximum number of iterations or when there is minimal change in the assignments or centroid positions. The final result of the algorithm is the assignment of data points to clusters and the positions of the centroids, representing the K clusters that best capture the structure of the data. The algorithm's effectiveness lies in its ability to find optimal centroids that reduce the sum of squared distances between the allocated centroids and the data points, thereby grouping similar data points together and separating different clusters.

2.3. Power spectral density (PSD)

Fundamentally the power spectral density (PSD) describes how a signal's power is distributed over its frequency range. The Welch Method is a widely used technique to estimate the PSD of a signal based on its discrete samples. The original earthquake catalog consists of irregularly spaced events; therefore, prior to spectral analysis, the data were converted into a uniformly sampled time series using a 12-hour resampling interval. Within each interval, the mean depth and magnitude were computed. This results in a discrete-time signal $x(n)$ with a constant sampling interval of 12 hours, corresponding to

a sampling frequency $f_s = 2$ samples per day. Consequently, the frequency axis in the PSD is expressed in cycles per day (day^{-1}). Mathematically, Welch's method (1967) can be expressed as follows. Let $x(n)$ be a discrete-time signal, where n denotes the sample index. The signal $x(n)$ is divided into M overlapping segments of length N , denoted as $x_k(n)$, where $k = 0, 1, 2, \dots, M - 1$. Each segment is multiplied by a window function $w(n)$ to reduce spectral leakage. The discrete Fourier transform (DFT) is then computed for each windowed segment to obtain $X_k(f)$, which represents the frequency-domain representation of the k -th segment. The power spectral density at frequency f , denoted as $PSD(f)$, is estimated by averaging the squared magnitudes of the Fourier transforms across all segments. It can be expressed as:

$$PSD(f) = \left(\frac{1}{(M \cdot f_s)} \right) * \sum_{k=0}^{M-1} |X_k(f)|^2 \quad (1)$$

where, $PSD(f)$ represents the power spectral density at frequency f , $X_k(f)$ is the Fourier transform of the k -th windowed segment, M is the number of segments, and f_s is the sampling frequency.

2.4. Spectral edge frequency (SEF)

The Spectral Edge Frequency (SEF) is a parameter derived from the power spectral density (PSD) that represents the frequency below which a specified percentage of the total spectral power is contained. It is widely used to characterize the distribution of power across frequencies in a signal. Let $PSD(f)$ denote the power spectral density obtained using the Welch Method. The cumulative distribution function (CDF) of the PSD is computed by cumulatively summing the PSD values and normalizing by the total power. In discrete form, this can be written as:

$$CDF(f_i) = \frac{\sum_{j=0}^i PSD(f_j)}{\sum_{j=0}^N PSD(f_j)} \quad (2)$$

where f_i and f_j denote the i -th and j -th discrete frequency bins, respectively, and N is the total number of frequency bins. The CDF represents the fraction of total spectral power accumulated up to frequency f_i .

The SEF at a specified percentile P (e.g., 95%) is defined as the frequency at which the CDF reaches that percentile. It is obtained by interpolating the frequency corresponding to the desired cumulative power level:

$$SEF(P) = \text{interp}(P, CDF(f), f) \quad (3)$$

where interp denotes a linear interpolation function, $CDF(f)$ is the cumulative distribution, and f is the

frequency vector. By determining the SEF, we can identify the frequency at which a specified percentage of the total power is reached.

2.5. Probability density function (PDF)

In probability theory and statistics, the Probability Density Function (PDF) describes how the probability of a continuous random variable is distributed over its range of values. In this study, the PDF is estimated using a histogram-based approach, which provides a simple and non-parametric representation of the underlying distribution derived directly from the observed data. Let X be a continuous random variable with observed values $x_1, x_2, \dots, x_n$. The range of X is partitioned into a set of contiguous intervals (bins). Unlike fixed binning schemes, the bin width Δx is determined using the Freedman–Diaconis rule, which adapts to the variability and size of the dataset, ensuring a more robust and data-driven representation of the distribution. The bin width is given by:

$$\Delta x = \frac{2IQR}{n^{1/3}} \quad (4)$$

where IQR denotes the interquartile range of the data and n is the total number of observations.

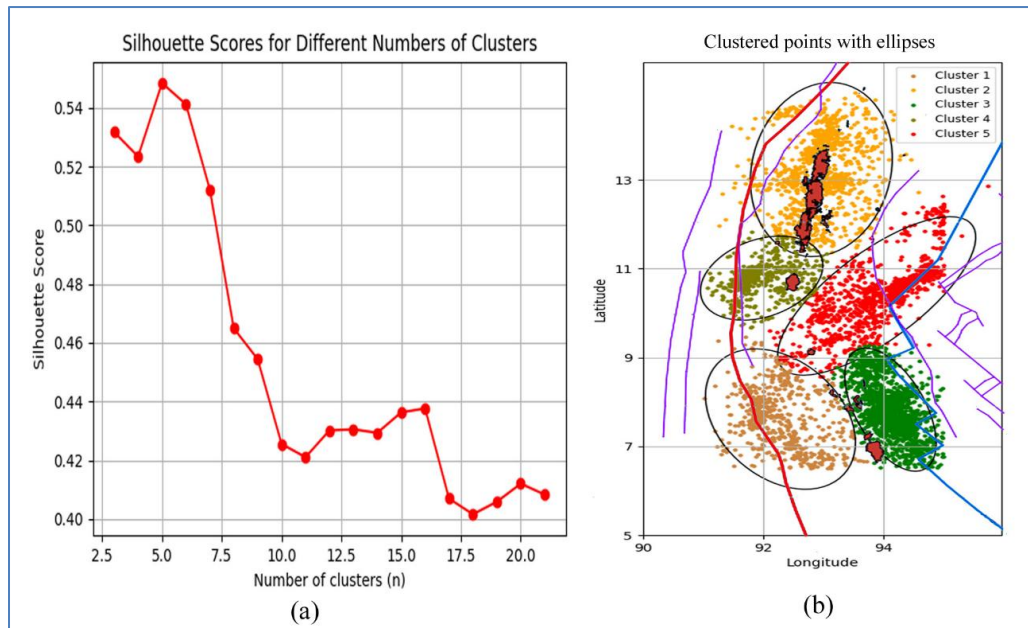
For each bin, the number of observations n_i falling within that interval is counted. The PDF is then estimated by normalizing these counts by both the total number of observations and the bin width. The estimated PDF at the i -th bin is expressed as:

$$f(x_i) = \frac{n_i}{n \cdot \Delta x} \quad (5)$$

where n_i is the number of observations in the i -th bin, n is the total number of observations, and Δx is the bin width. The quantity $f(x_i)$ represents the probability density associated with that bin. By calculating the PDF, we obtain a representation of the distribution of the variable, highlighting the regions of higher and lower likelihood.

2.6. Data sources

Our study involved collecting comprehensive data from the databases of the United States Geological Survey (USGS) and the India Meteorological Department (IMD). We focused on the Andaman and Nicobar Islands Seismic Region, spanning latitudes of 6.5N to 15N and longitudes of 91E to 96E. We specifically analyzed earthquakes with magnitudes greater than 3.5, 4, 4.5, and 5, resulting in a dataset of 5,155 earthquakes. The temporal span of our dataset extended from 1921 to 2022, allowing us to examine earthquake patterns over a significant timeframe



Figs. 2(a&b). (a) Silhouette Scores & (b) K-means clustering for earthquakes in Andaman and Nicobar Islands region.

3. Results and discussion

By considering the geographical context, including fault lines, minor and major plate boundaries, and the distribution of islands, we recognized that the region's seismic behaviour could not be adequately represented by fewer than three clusters due to its unique and complex characteristics. However, based on the geological and tectonic segmentation of the Andaman–Nicobar region, the study area can be more accurately divided into at least five distinct seismic clusters. Cluster 1 covers regions along the Mentawai Fault, the Sunda-Banda Trench, and parts of the Andaman Trench, and is associated with large-scale subduction processes that produce deep, high-energy seismicity. Cluster 2 includes Central to South Andaman and incorporates a complex network of geological structures such as the Irawaddy Fan, Arakan Trench, Homfray's Strait–Button Fault, Outram Fault, Havelock Fault, Neil Fault, and Narcondam and Barren Islands, which are indicative of active arc volcanism and strike-slip faulting. The presence of thick sediments and volcanic centers makes this cluster seismically and geodynamically distinct. Cluster 3 lies within the backarc Andaman Sea, covering features like the Sewell Rise, Andaman Basin, Sumatra Fault, and several volcanic seamounts, where extensional tectonics and complex crustal deformation result in intermediate-depth earthquakes with a broad range of magnitudes. Cluster 4 encompasses parts of the Andaman Trench and Ninetyeast Ridge, extending into the outer forearc and Nicobar

region. It is characterized by accretionary prism activity and shallow seismicity distributed across Car Nicobar, Little Andaman, and Great Nicobar islands. Cluster 5 is associated with the southern subduction front and incorporates key tectonic elements including the Alcock Rise (AR), Eastern Margin Fault (EMF), West Andaman Fault (WAF), South Andaman Fault, Diligent Fault (DF), Andaman Basin, and Mergui Basin. This cluster reflects highly complex fault dynamics and deep subduction-related seismicity. Although Mukhopadhyay *et al.* (2010) identified up to ten clusters in the broader Andaman–Nicobar region, our objective was to balance geological meaningfulness with statistical robustness. Therefore, to refine our cluster selection, we applied silhouette score analysis for a range of cluster numbers (3 to 21). This iterative process revealed that five clusters, corresponding to the highest silhouette score of 0.54, provided the most meaningful segmentation, aligning closely with known fault structures and regional tectonic boundaries. In contrast, a cluster count of 21 yielded the lowest silhouette score of 0.39, indicating poor separation and likely overlapping or ambiguous seismic groupings (see Fig. 2a).

Now, in Fig. 2(b), clusters are depicted with distinct colours: light brown represents cluster-1, yellow for cluster-2, green for cluster-3, yellow-green for cluster-4, and red for cluster-5. Additionally, Table 1 provides a summary of basic information about each cluster, including the number of earthquakes, centroid coordinates, average magnitude, and average depth.

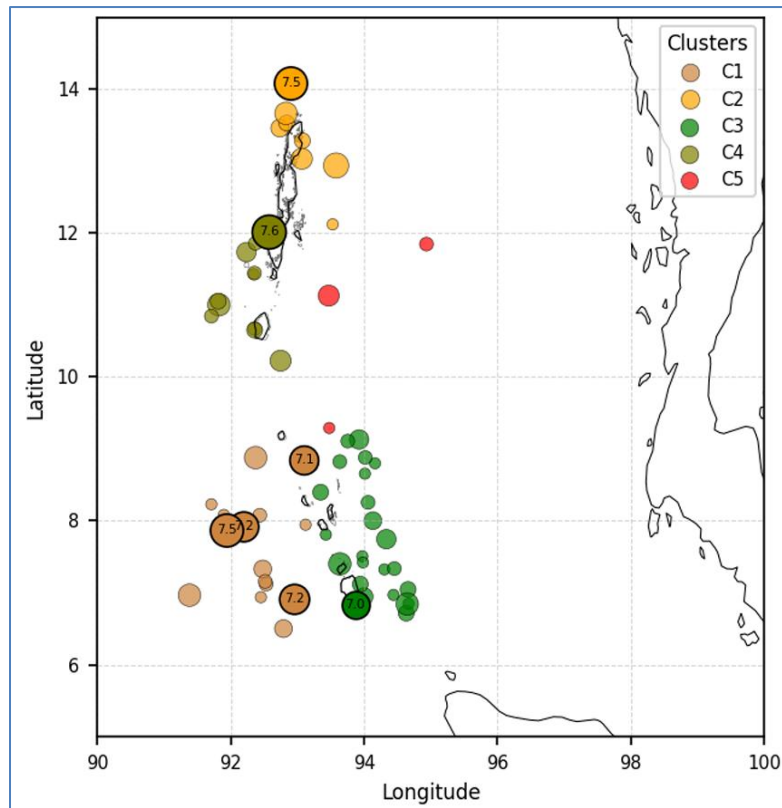


Fig. 3. Major earthquakes (magnitude ≥ 6) in the Andaman and Nicobar Islands region from 1921 to 2022

TABLE 1

Summary of the Clusters (Earthquake magnitude > 3.5)

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Number of Earth Quake	592	929	1967	722	945
Centroid	7.72, 92.25	13.38, 92.98	7.80, 94.09	10.80, 92.12	10.50, 94.02
Average Magnitude	4.61	4.52	4.61	4.51	4.58
Average Depth(KM)	29.93	35.63	31.08	27.31	37.83

The spatial distribution of major earthquakes ($M \geq 6$) during the study period in the Andaman and Nicobar Islands region. A total of 62 earthquake events are identified, distributed as follows: Cluster 1 contains 15 events, Cluster 2 includes 8 events, Cluster 3 has the highest number with 25 events, Cluster 4 comprises 11 events, and Cluster 5 contains only 3 events. Cluster 3 contains a relatively higher number of earthquake events; however, most of them are of moderate magnitude, with only a few events approaching magnitude 7. In contrast, Cluster 1 includes several high-magnitude events, with

approximately four major earthquakes exceeding magnitude 7 ($M > 7$). Cluster 2 and Cluster 4 also show the occurrence of strong earthquakes, including events above magnitude 7. Notably, Cluster 5 does not exhibit any earthquake events with magnitude greater than 7 (see Fig 3).

3.1. Power spectral density analysis

The power spectral density (PSD), computed from the uniformly resampled time series, represents the

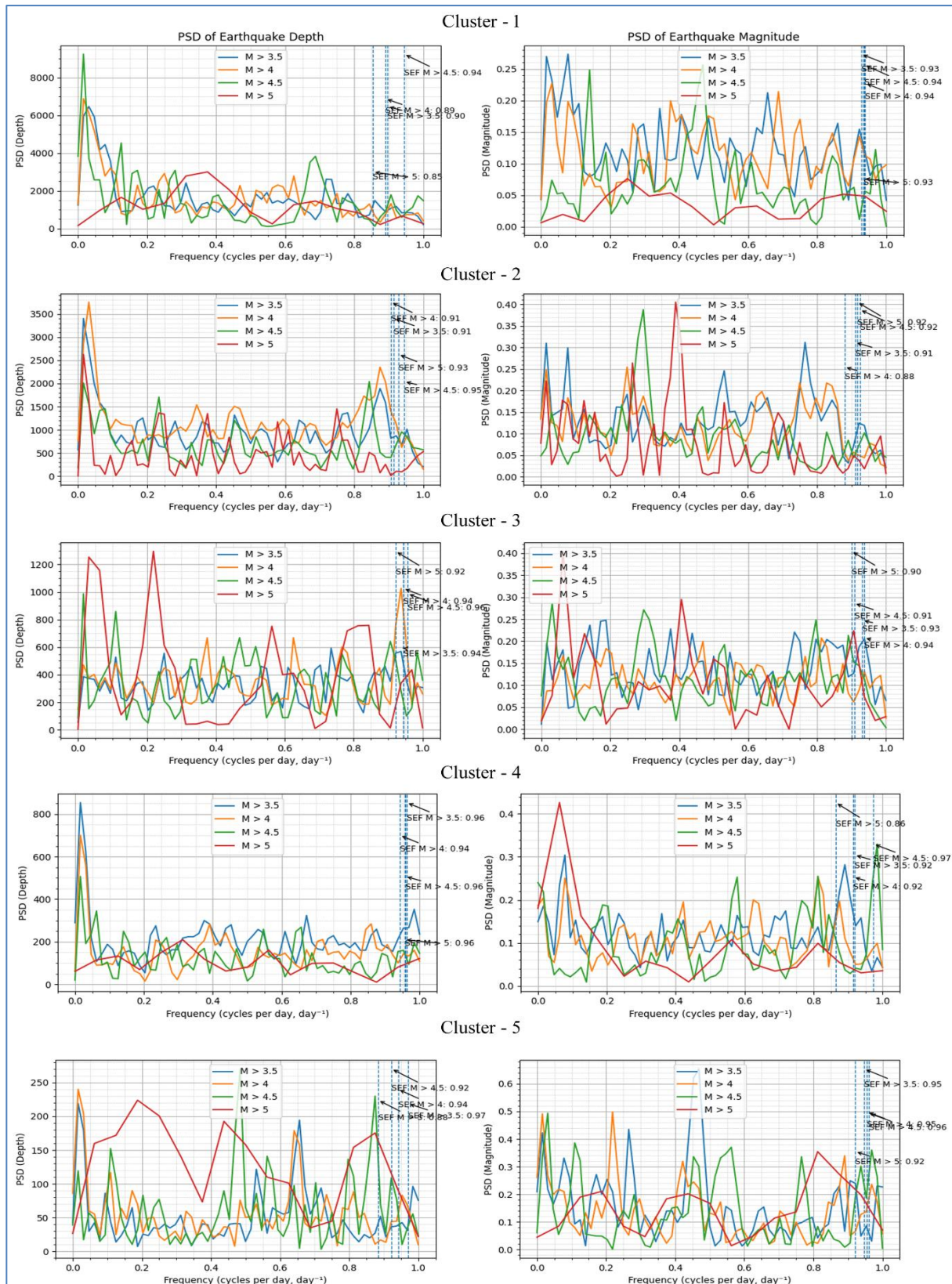


Fig. 4. Cluster-based PSD analysis of depth and magnitude for earthquakes with magnitudes greater than 3.5, 4, 4.5, and 5 in the Andaman and Nicobar region

distribution of variance (power) of earthquake depth and magnitude over frequency. Cluster-1, centered at 7.7° N and 92.25° E, is located southwest of Great Nicobar within the Andaman Trench, a major subduction zone characterized by relatively infrequent but significant seismic activity. Analysis of the PSD of earthquake depth indicates that for magnitudes greater than 3.5 and 4, the dominant frequencies are concentrated in the low-frequency range ($\sim 0.016\text{--}0.062\text{ day}^{-1}$), where the highest PSD peaks occur (*e.g.*, ~ 6462 and ~ 6876). The variance of the depth time series is strongly concentrated toward lower frequencies, indicating dominant long-period variability (see fig-4). For magnitudes greater than 4.5, the strongest contribution remains at 0.016 day^{-1} ($\text{PSD} \approx 9238$), reinforcing the persistence of very low-frequency variability, although secondary peaks at 0.125 and 0.688 day^{-1} indicate the emergence of shorter-period components. For magnitudes greater than 5, the dominant depth frequencies shift toward intermediate to higher frequencies ($\sim 0.312\text{--}0.375\text{ day}^{-1}$), reflecting increased short-term variability in focal depth for larger events (see fig-4). It may also involve variations in stress accumulation and relief across the segmented plate boundary between the pre- and post-seismic rupture phases (Khan *et al.*, 2012).

In contrast, the PSD of earthquake magnitude shows that for $M > 3.5$ and $M > 4$, the variance is primarily concentrated at low frequencies ($\sim 0.016\text{--}0.078\text{ day}^{-1}$), although isolated higher-frequency components (*e.g.*, ~ 0.656 and $\sim 0.688\text{ day}^{-1}$) are present but not dominant. For $M > 4.5$, the dominant frequencies shift toward an intermediate range ($\sim 0.141\text{--}0.469\text{ day}^{-1}$), indicating a transition from long-period to shorter-period variability in magnitude. This trend becomes more pronounced for $M > 5$, where the dominant frequencies extend into the higher-frequency range ($\sim 0.25\text{--}0.94\text{ day}^{-1}$), reflecting stronger short-term fluctuations in the magnitude time series.

Subduction zones are commonly associated with deeper seismic processes that may produce long-period temporal variability, reflected here by strong low-frequency PSD peaks (especially near $\sim 0.016\text{ day}^{-1}$). The shift toward higher frequencies for larger magnitudes may indicate increasing temporal variability related to more complex rupture processes and fault interactions (Khan *et al.*, 2012). Low-frequency components may be linked to large-scale, slow tectonic deformation and deeper seismic processes, whereas higher-frequency components may reflect more localized and rapid changes in seismic activity (Khan *et al.*, 2012). Thick sedimentary layers in the Andaman region may also contribute to enhanced long-period variability. High-resolution multichannel seismic reflection data and receiver function

profiles reveal that the subducting crust in this province carries exceptionally thick continental margin and sediment fan sequences from the Bengal Fan system (Mishra and Prajapati, 2019). These deposits expand the wide accretionary wedge up to 250 km in width (Singh and Moeremans, 2017), where the shallow crust consists of 4 to 7 km of soft, fluid-saturated accretionary sediments that attenuate high frequencies and emphasize long-period, low-frequency variations (Mishra and Prajapati, 2019). The increasing contribution of higher frequencies for larger earthquakes may indicate more multi-scale temporal behavior.

On the other hand, Cluster-2, centered at 13.38° N and 92.98° E, includes the Irawaddy Fan, the northern Andaman Trench, and parts of the Arakan Trench. This region is tectonically complex, marked by multiple active fault systems (*e.g.*, Homfray's Strait-Button Fault, Outram Fault, Havelock Fault, and Neil Fault) and nearby volcanic centers such as Barren and Narcondam Islands, indicating a highly dynamic seismotectonic regime governed by highly oblique plate convergence (Yadav and Tiwari, 2020).

The power spectral density (PSD) analysis of earthquake depth demonstrates a consistent dominance of low-frequency components ($\sim 0.016\text{--}0.031\text{ day}^{-1}$) across magnitudes $M > 3.5$ and $M > 4$, where the highest PSD peaks are observed (*e.g.*, ~ 3402 and ~ 3755 , respectively). The variance (power) of the depth time series is concentrated toward the lower-frequency range, despite the presence of secondary peaks at higher frequencies ($\sim 0.875\text{--}0.891\text{ day}^{-1}$). For $M > 4.5$, low-frequency dominance persists (notably at 0.016 day^{-1}), but the spectrum broadens with contributions from intermediate ($\sim 0.234\text{ day}^{-1}$) and high frequencies ($\sim 0.844\text{ day}^{-1}$), indicating increased temporal variability. For $M > 5$, the strongest peak again occurs at low frequency (0.016 day^{-1}), although the spectrum shows a noticeable spread into intermediate and higher frequencies ($0.234\text{--}0.750\text{ day}^{-1}$), suggesting the coexistence of long-term and short-term depth fluctuations (see Fig.-4).

In contrast, the PSD of earthquake magnitude exhibits a more complex and distributed spectral pattern. For $M > 3.5$ and $M > 4$, the variance is distributed across both low-frequency ($\sim 0.016\text{--}0.078\text{ day}^{-1}$) and high-frequency ranges ($\sim 0.75\text{--}0.84\text{ day}^{-1}$), indicating mixed temporal behavior. As magnitude increases ($M > 4.5$), the dominant peaks shift toward the intermediate-frequency range ($\sim 0.234\text{--}0.312\text{ day}^{-1}$), where the highest PSD values (up to ~ 0.39) are observed. This transition is clearly visible in the Fig. 4 as a concentration of peaks in the mid-frequency band. For $M > 5$, the

magnitude PSD remains dominated by intermediate to moderately high frequencies ($\sim 0.266\text{--}0.406\text{ day}^{-1}$), although a persistent low-frequency peak at 0.016 day^{-1} indicates that long-period variability still plays a role.

Depth variations are dominated by low-frequency components, which may be associated with subduction-related tectonic processes. Magnitude variations shift toward intermediate and higher frequencies with increasing magnitude, which may reflect localized temporal fluctuations related to fault segmentation, secondary fault interactions, and volcanic influences. This unique spectral structure corresponds to distinct subslab conditions in the northern Andaman sector, which is structurally characterized by a significantly steeper slab dip angle, high intrinsic attenuation due to slab dehydration fluids, and a distinct trench-parallel mantle anisotropy compared to the southern clusters (Srijayanthi and Kumar, 2020).

In addition, Cluster-3, encompassing the Sewell Rise and the Andaman Basin, includes important geological features such as the northern extension of the Sumatra Fault, a volcanic seamount, the southern segment of the West Andaman Fault, and the Eastern Margin Fault. This combination of active fault systems and volcanic structures contributes to a highly complex seismotectonic environment. The Fig. 4 clearly shows that the PSD of earthquake depth exhibits a broad spectral distribution across frequencies. For magnitudes greater than 3.5, dominant peaks occur in both the intermediate ($\sim 0.109\text{--}0.250\text{ day}^{-1}$) and high-frequency ranges ($\sim 0.734\text{--}0.938\text{ day}^{-1}$), with slightly higher PSD values in the high-frequency band, indicating significant short-term variability. For $M > 4$, this pattern becomes more pronounced, with strong peaks near $\sim 0.922\text{--}0.938\text{ day}^{-1}$, while intermediate frequencies ($\sim 0.375\text{--}0.625\text{ day}^{-1}$) remain important. This frequency seismic distribution can be physically linked to active shallow seafloor spreading, crustal rifting, and elevated thermal conditions near back-arc spreading segments and submarine volcanic structures (Kamesh *et al.*, 2020).

For magnitudes greater than 4.5, the depth PSD shows a shift toward low-frequency dominance, with the strongest peak at $\sim 0.016\text{ day}^{-1}$ ($PSD \approx 987$), while intermediate and high-frequency components remain visible in the spectrum. For $M > 5$, the dominance of low-frequency components ($\sim 0.031\text{--}0.062\text{ day}^{-1}$) becomes more evident (PSD up to ~ 1253), although intermediate ($\sim 0.219\text{ day}^{-1}$) and high-frequency ($\sim 0.812\text{--}0.844\text{ day}^{-1}$) contributions persist, confirming a wide and multi-scale variability in depth.

In contrast, the PSD of earthquake magnitude shows a consistently broad and distributed spectral pattern in the Fig. 4. For $M > 3.5$, dominant frequencies are mainly in the intermediate range ($\sim 0.141\text{--}0.188\text{ day}^{-1}$), with additional contributions at higher frequencies ($\sim 0.734\text{ day}^{-1}$), indicating mixed temporal variability. For $M > 4$, the magnitude PSD remains spread across intermediate ($\sim 0.188\text{--}0.203\text{ day}^{-1}$) & higher frequencies ($\sim 0.812\text{--}0.828\text{ day}^{-1}$), without a single dominant peak. For $M > 4.5$, the dominant frequencies shift slightly toward intermediate values ($\sim 0.281\text{--}0.312\text{ day}^{-1}$), while a low-frequency contribution ($\sim 0.031\text{ day}^{-1}$) also appears. For $M > 5$, a multi-band distribution, with peaks at low ($\sim 0.062\text{--}0.125\text{ day}^{-1}$), intermediate ($\sim 0.406\text{--}0.500\text{ day}^{-1}$), and high frequencies ($\sim 0.906\text{ day}^{-1}$), indicating strong multi-scale temporal variability in magnitude.

Cluster-3 contains both low- and high-frequency components, indicating the presence of long-term and short-term temporal variability. The strong high-frequency components, particularly in depth for lower magnitudes, may reflect rapid and localized changes associated with active faulting and volcanic influences. The increasing contribution of low-frequency components for larger magnitudes may indicate broader-scale tectonic processes. The interaction of multiple fault systems, basin structures, and volcanic activity may produce this multi-scale seismic variability, resulting in more diverse temporal behavior than observed in Clusters-1 and 2 (Kamesh *et al.*, 2020).

Similarly, Cluster-4, centered at 10.8° N and 92.12° E , is intersected by the Andaman Trench and represents an active seismotectonic region influenced by subduction processes and associated fault systems. Analysis of the PSD of earthquake depth shows a clear and consistent dominance of low-frequency components ($\sim 0.016\text{--}0.031\text{ day}^{-1}$) for magnitudes $M > 3.5$ and $M > 4$, where the highest PSD values occur (e.g., ~ 854 and ~ 700 , respectively). The depth spectra are strongly concentrated toward the lower-frequency end, indicating dominant long-period variability.

For magnitudes greater than 4.5, low-frequency dominance persists ($peak\ at\ \sim 0.016\text{ day}^{-1}$, $PSD \approx 507$), although intermediate frequencies ($\sim 0.062\text{--}0.141\text{ day}^{-1}$) and higher frequencies ($\sim 0.547\text{--}0.656\text{ day}^{-1}$) also contribute, suggesting increased temporal variability. However, for $M > 5$, the spectral behavior changes noticeably: the dominant peaks shift toward intermediate frequencies ($\sim 0.312\text{ day}^{-1}$), while very low-frequency dominance weakens (see fig-4). However, larger events show a transition from predominantly long-period behavior to more mixed temporal variability in depth.

In contrast, the PSD of earthquake magnitude exhibits a broad and mixed spectral distribution across all magnitude thresholds. For $M > 3.5$, dominant frequencies occur in both low-frequency ($\sim 0.062\text{--}0.078\text{ day}^{-1}$) and high-frequency ranges ($\sim 0.875\text{--}0.906\text{ day}^{-1}$), indicating mixed temporal characteristics. For $M > 4$, the spectrum remains widely distributed across low ($\sim 0.016\text{--}0.078\text{ day}^{-1}$), intermediate ($\sim 0.625\text{ day}^{-1}$), and high frequencies ($\sim 0.812\text{--}0.828\text{ day}^{-1}$), without a single dominant band.

For magnitudes greater than 4.5, the magnitude PSD shows increasing contribution from higher frequencies ($\sim 0.812\text{--}0.984\text{ day}^{-1}$), along with a notable low-frequency (near 0 day^{-1}) component, reflecting both long-term and rapid variability. For $M > 5$, the strongest peak shifts back to the low-frequency range ($\sim 0.062\text{ day}^{-1}$, $PSD \approx 0.426$), while intermediate ($\sim 0.125\text{--}0.562\text{ day}^{-1}$) and high-frequency ($\sim 0.812\text{ day}^{-1}$) components remain present.

The tectonic setting of the Andaman Trench, where subduction-related processes contribute to long-period variability, while fault interactions introduce shorter-term fluctuations. The strong low-frequency dominance in depth indicates the influence of large-scale, slow tectonic processes, particularly for smaller magnitudes. In contrast, the broader spectral distribution observed in magnitude indicates more complex temporal behavior.

The coexistence of low-, intermediate-, and high-frequency components indicates multi-scale seismic variability in Cluster-4. Low frequencies may reflect large-scale tectonic deformation, whereas intermediate and high frequencies may represent localized and more rapid changes associated with subduction and crustal fault interactions across the outer arc system (Haripriya *et al.*, 2020; Mishra and Prajapati, 2019).

Cluster-5, centered at 10.5°N and 94.028°E , represents a structurally intricate region influenced by multiple interacting tectonic elements, including the Eastern Margin Fault, West Andaman Fault, South Andaman Fault, Diligent Fault, Alcock Rise, and the Andaman Basin. This structural complexity is reflected in the spectral characteristics of both depth and magnitude time series. The nested nature of these active strands across the Alcock Rise produces intense strain partitioning, shearing, and crustal block rotations (Kamesh *et al.*, 2020; Yadav and Tiwari, 2020).

The PSD of earthquake depth reveals a clear evolution in temporal variability with increasing magnitude. For lower magnitudes ($M > 3.5$ and $M > 4$), the figure shows that the variance is primarily

concentrated at low frequencies ($\sim 0.016\text{--}0.031\text{ day}^{-1}$), corresponding to the highest PSD peaks, while secondary contributions appear in the intermediate-to-high frequency range ($\sim 0.53\text{--}0.66\text{ day}^{-1}$). As magnitude increases to $M > 4.5$, the spectral pattern visibly shifts in the figure, with stronger contributions emerging at intermediate and higher frequencies ($\sim 0.48\text{--}0.88\text{ day}^{-1}$), indicating enhanced short-term variability. For the largest events ($M > 5$), the depth spectrum becomes more evenly distributed across intermediate frequencies ($\sim 0.125\text{--}0.438\text{ day}^{-1}$), with no single dominant peak, reflecting a transition toward mixed temporal behavior. A high-resolution velocity framework derived from ambient noise cross-correlations constrains the upper plate crust beneath Central Andaman, where a 2–6 km thick low-velocity Andaman Flysch overlies an active, intermediate continental-type crystalline basement (Gupta *et al.*, 2016). For $M > 3.5$, dominant peaks are observed in the intermediate frequency range ($\sim 0.26\text{--}0.47\text{ day}^{-1}$), with additional contributions from low ($\sim 0.016\text{ day}^{-1}$) and higher frequencies. For $M > 4$, the spectrum remains widely spread, with notable peaks at both low ($\sim 0.016\text{ day}^{-1}$) and intermediate frequencies ($\sim 0.22\text{--}0.42\text{ day}^{-1}$), along with high-frequency components ($\sim 0.89\text{ day}^{-1}$). As magnitude increases further ($M > 4.5$ and $M > 5$), the figure shows a gradual strengthening of higher-frequency contributions ($\sim 0.81\text{--}0.94\text{ day}^{-1}$), although intermediate-frequency variability remains significant, indicating persistent multi-scale fluctuations in magnitude.

Taken together, Cluster-5 does not exhibit a single dominant frequency band, but instead shows a continuously shifting and distributed spectral structure. Depth variability transitions from low-frequency dominance to intermediate-frequency control, while magnitude variability remains broadband across low to high frequencies for all magnitude thresholds. This spectral behavior may reflect the interaction of multiple fault systems and basin structures, which may produce both long-period (low-frequency) and short-term (high-frequency) temporal variations (Kamesh *et al.*, 2020). The absence of a stable dominant frequency band, particularly at higher magnitudes, may indicate that seismicity in Cluster-5 is influenced by multi-scale processes involving large-scale tectonic deformation and localized fault interactions (Khan *et al.*, 2012).

Across all five clusters, the Spectral Edge Frequency (SEF), computed for different magnitude thresholds ($M > 3.5$ to $M > 5$), consistently falls within the range of approximately 0.90 to 0.96 day^{-1} . Since the power spectral density (PSD) in this study is derived from uniformly resampled time series of earthquake depth and magnitude, it represents the distribution of variance across

frequencies rather than true physical seismic energy. Accordingly, the SEF indicates that 95% of the variance in the temporal fluctuations of depth and magnitude is contained below these frequency levels.

Although the SEF values are broadly consistent among the clusters, some variations are observed. In particular, the depth SEF does not remain uniform across all magnitude thresholds in Clusters-1 and 5, while the magnitude SEF shows greater variability in Clusters-4 and 5. Although SEF does not directly quantify physical source parameters such as stress drop or rupture velocity, it provides a robust measure of the temporal variability and spectral bandwidth of the observed seismic parameters.

3.2. Probability density function analysis

The probability density function (PDF) analysis reveals a well-defined and consistent pattern in the distribution of earthquake depth across all magnitude thresholds. In Cluster-1, for events with $M > 3.5$, the highest PDF values occur at depths of approximately 12 km ($PDF \approx 0.058$) and ~ 35 km ($PDF \approx 0.050$), indicating that earthquake occurrence is strongly concentrated within these shallow to intermediate depth ranges. As the magnitude threshold increases to $M > 4$ and $M > 4.5$, this bimodal depth structure remains stable, with dominant peaks consistently observed around ~ 10 – 12 km and ~ 30 – 35 km, and the maximum PDF reaching ~ 0.067 and ~ 0.080 , respectively. For larger events ($M > 5$), the dominant depth peaks shift slightly to ~ 7 – 8 km and ~ 33 – 34 km, but the overall pattern persists. Beyond these ranges, the PDF values decrease rapidly, confirming that deeper events (> 50 km) are comparatively rare in this cluster.

The magnitude PDF exhibits a similarly structured but magnitude-dependent distribution. For $M > 3.5$, the highest probabilities are concentrated in the range ~ 4.3 – 4.7 , with a peak near 4.55 ($PDF \approx 1.16$). As the threshold increases to $M > 4$, the dominant range remains similar but becomes more pronounced, with higher PDF values (up to ~ 1.49) centered around ~ 4.4 – 4.7 . For $M > 4.5$, the distribution narrows further, with strong clustering around ~ 4.6 – 4.9 , where the PDF reaches its maximum (~ 2.82). For the largest events ($M > 5$), the distribution shifts upward, with a sharp and dominant peak at ~ 5.16 ($PDF \approx 4.41$), followed by a rapid decline beyond ~ 5.5 – 5.7 , indicating that very high-magnitude events are increasingly rare. Fig. 5 supports these observations by showing pronounced peaks in the depth PDF at two primary zones, approximately 10–15 km and 30–35 km. The magnitude distributions are tightly clustered and shift toward higher values as the magnitude

threshold increases. The presence of these two preferred seismogenic layers may be associated with shallow crustal faulting and deeper subduction-related processes (Lay *et al.*, 1995; Scholz, 2019; Špičák *et al.*, 2013).

The strongest concentration of events consistently occurs around ~ 30 km, accompanied by a secondary but stable peak near ~ 9 – 11 km in cluster-2. For $M > 3.5$, the dominant peak at ~ 30 km ($PDF \approx 0.224$) clearly exceeds the shallow peak (~ 9 km, $PDF \approx 0.090$), and this contrast becomes even more pronounced for $M > 4$, where the maximum PDF increases to ~ 0.265 at ~ 30 km. As the magnitude threshold increases ($M > 4.5$ and $M > 5$), both peaks remain present with only minor shifts (≈ 29 – 31 km and ≈ 10 km), demonstrating that the depth distribution is structurally stable. Outside the ~ 5 – 40 km range, the PDF rapidly approaches zero, confirming that earthquake occurrence is strongly confined within these two depth intervals.

A similarly systematic pattern is observed in the magnitude PDF, although here the distribution evolves progressively with increasing magnitude threshold. For $M > 3.5$, the highest probabilities are concentrated within ~ 4.3 – 4.7 , with a peak near ~ 4.5 – 4.6 ($PDF \approx 1.47$). This concentration becomes sharper and more pronounced for $M > 4$ (PDF up to ~ 1.90), indicating tighter clustering of moderate magnitudes. As the threshold increases to $M > 4.5$, the dominant range narrows further and shifts slightly upward (~ 4.6 – 4.8), where the PDF reaches its maximum (~ 2.82). For $M > 5$, the distribution clearly transitions to higher magnitudes, with strong peaks around ~ 5.1 – 5.4 ($PDF \approx 2.4$ – 2.5), followed by a rapid decay beyond ~ 5.8 – 6.0 , reflecting the reduced likelihood of larger events. Taken together, the seismicity appears to be highly structured rather than randomly distributed. The two dominant depth levels may represent distinct seismogenic layers, while the upward shift in magnitude distribution may reflect scaling in earthquake occurrence (Kagan, 2010; Kolathayar and Sitharam, 2012). Both shallow and intermediate-depth processes may influence where earthquakes occur and how large they become.

Additionally, For $M > 3.5$, the highest probability occurs at ~ 29 – 30 km ($PDF \approx 0.261$), followed by closely spaced peaks at ~ 32 – 35 km, showing that most events are tightly clustered within this depth interval in Cluster-3 (see Fig. 5). A much weaker secondary contribution appears near ~ 10 km, but its PDF value is significantly lower, indicating that shallow events are comparatively less frequent.

As the magnitude threshold increases ($M > 4$ and $M > 4.5$), this pattern remains stable, with

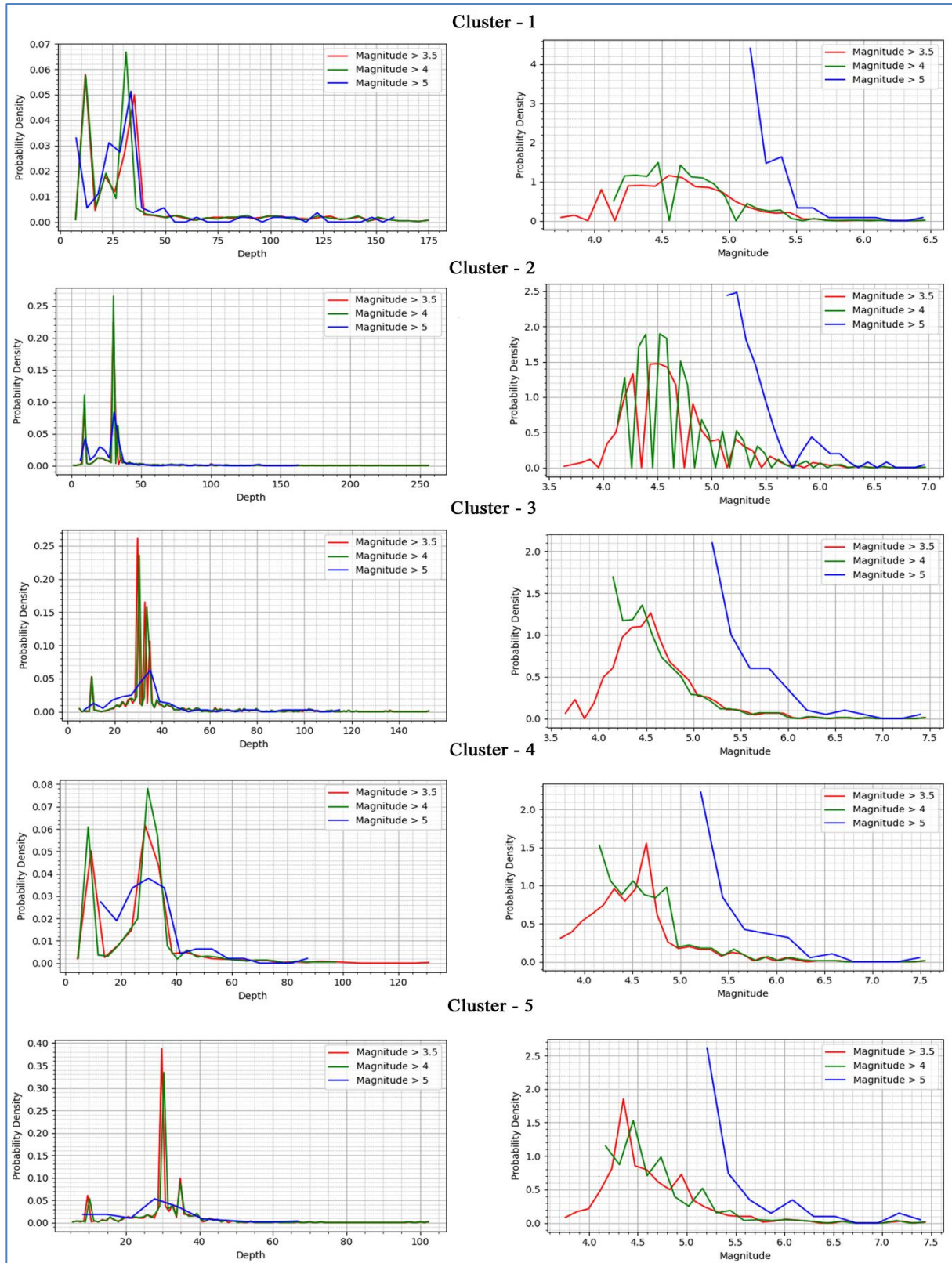


Fig. 5. Cluster-based probability density function (PDF) of earthquake depth and magnitude for events with $M > 3.5, 4.0, 4.5$, and 5.0 in the Andaman and Nicobar region, estimated using a histogram-based approach with Freedman–Diaconis binning

dominant peaks consistently centered around $\sim 30\text{--}34\text{ km}$ (e.g., $\sim 30.3\text{ km}$ and $\sim 33.5\text{ km}$ for $M > 4$; $\sim 34.0\text{ km}$ and $\sim 29.9\text{ km}$ for $M > 4.5$). For $M > 5$, the concentration persists within the same range, with peaks near $\sim 31\text{--}35\text{ km}$, while additional minor contributions appear at shallower ($\sim 10\text{--}20\text{ km}$) and slightly deeper ($\sim 38\text{--}42\text{ km}$) levels.

The magnitude PDF exhibits a progressive shift toward higher values with increasing magnitude thresholds. For $M > 3.5$, the dominant range lies between ~ 4.3 and 4.6 , with a peak near ~ 4.55 ($\text{PDF} \approx 1.26$). For $M > 4$, the distribution remains within a similar range but becomes more concentrated, with stronger peaks around $\sim 4.1\text{--}4.5$ ($\text{PDF up to} \sim 1.69$).

For $M > 4.5$, the distribution shifts upward and becomes more sharply peaked, with a dominant maximum at ~ 4.65 ($\text{PDF} \approx 3.76$), followed by decreasing probabilities toward higher magnitudes. For $M > 5$, the distribution clearly transitions to higher values, with a strong peak at ~ 5.2 ($\text{PDF} \approx 2.10$) and additional contributions extending toward $\sim 5.4\text{--}6.0$, while probabilities beyond ~ 6.5 decrease rapidly. Taken together, Cluster-3 shows a stable depth distribution but a progressively changing magnitude distribution. The consistent depth concentration may reflect persistent structural control, whereas the variation in magnitude may be related to differences in stress accumulation, release processes, and fault interactions (Scholz, 2019).

In contrast, Cluster-4 presents a well-defined but slightly broader depth distribution compared to the previous clusters, where seismicity is concentrated within two preferential depth bands rather than a single confined layer. The PDF curves indicate consistent peaks around a shallow level ($\sim 8\text{--}12\text{ km}$) and a deeper band ($\sim 28\text{--}33\text{ km}$) for all magnitude thresholds. For $M > 3.5$ and $M > 4$, the strongest concentrations occur near $\sim 9\text{ km}$ and $\sim 29\text{--}30\text{ km}$, while for $M > 4.5$ and $M > 5$, the deeper peak becomes slightly more prominent, extending toward $\sim 32\text{--}36\text{ km}$. Beyond $\sim 40\text{--}50\text{ km}$, the probability rapidly diminishes, with only minor contributions at greater depths.

The magnitude distribution evolves more noticeably with increasing thresholds. At lower magnitudes ($M > 3.5$ and $M > 4$), the PDF is concentrated within a compact range of $\sim 4.1\text{--}4.7$, forming a relatively symmetric peak. As the threshold increases to $M > 4.5$, the distribution broadens and shifts toward higher magnitudes, with a strong concentration around $\sim 4.6\text{--}5.3$ and a pronounced peak near ~ 4.66 . For $M > 5$, the dominant range moves

further upward ($\sim 5.2\text{--}5.7$), accompanied by a long tail extending beyond magnitude 6, although with progressively lower probabilities. A key characteristic of Cluster-4 exhibits two persistent depth bands across all magnitude thresholds. Earthquakes of different sizes may originate from both shallow and deeper sources. The stronger contribution of the deeper band at higher magnitudes may indicate that larger events are associated with deeper structural interfaces, possibly related to subduction processes (Lay *et al*, 1995; Turcotte and Schubert, 2014).

Finally, the depth PDF reveals a strong and consistent peak centered near $\sim 29\text{--}31\text{ km}$ (e.g., 29.78 km , 30.37 km , 30.14 km), accompanied by a secondary but weaker concentration around $\sim 34\text{--}35\text{ km}$ in cluster-5. A shallow component near $\sim 8\text{--}10\text{ km}$ is present but comparatively subdued, and probabilities decrease rapidly beyond $\sim 40\text{ km}$, with only sparse deep events. As the magnitude threshold increases ($M > 4.5$ and $M > 5$), the overall spread slightly broadens, yet the primary concentration remains anchored within the $\sim 28\text{--}35\text{ km}$ range.

In contrast to the tightly constrained depth behavior, the magnitude PDF exhibits a progressive shift and expansion toward higher values. For $M > 3.5$, the distribution is concentrated within $\sim 4.2\text{--}4.6$, with a prominent peak around ~ 4.35 . As the threshold increases to $M > 4$, the dominant range extends to $\sim 4.2\text{--}4.7$, maintaining a relatively compact structure but with increased spread. A more noticeable transition occurs for $M > 4.5$, where the distribution shifts upward ($\sim 4.6\text{--}5.2$) and displays a strong peak near ~ 4.66 . For $M > 5$, the dominant magnitudes move further into the $\sim 5.2\text{--}5.6$ range, accompanied by a long tail extending beyond magnitude 6 and even approaching ~ 7 , though with low probability. Cluster-5 shows a persistent depth concentration near $\sim 30\text{ km}$ across all magnitude thresholds. This stable depth band may reflect a major tectonic interface or a mechanically favorable rupture zone. At the same time, the magnitude PDF becomes broader as magnitude increases, indicating that earthquakes may repeatedly originate within the same depth range but vary considerably in size depending on stress accumulation and release processes (Scholz, 2019; Kolathayar and Sitharam, 2012).

4. Conclusions

This study investigates the seismic characteristics of the Andaman and Nicobar region using Power Spectral Density (PSD) and Probability Density Function (PDF) analyses of earthquake depth and magnitude, combined with cluster-based segmentation through K-means

optimized by silhouette scores. The results identify five distinct seismic clusters that correspond well with the region's complex tectonic framework.

The PSD analysis highlights clear multi-scale temporal behavior in seismicity across all clusters. The PSD was estimated using the Welch method from earthquake data resampled at 12-hour intervals (2 samples per day). Clusters-1, 2, and 4 are characterized by stronger low-frequency components, indicating relatively stable and gradual variations in earthquake depth. In contrast, Clusters-3 and 5 exhibit broader spectral distributions with significant high-frequency contributions, suggesting more rapid and heterogeneous temporal fluctuations. As the magnitude threshold increases, low-frequency components become more prominent in the depth PSD, particularly for larger events, reflecting the influence of large-scale tectonic controls. The magnitude PSD shows a broad frequency distribution across all clusters, indicating persistent multi-scale variability. Higher magnitude thresholds are associated with enhanced high-frequency components, suggesting increased short-term complexity in the temporal behavior of earthquake magnitudes. Spectral Edge Frequency (SEF) values of approximately $0.90\text{--}0.96\text{ day}^{-1}$ indicate that 95% of the variance in the depth and magnitude time series is contained below this frequency range.

The PDF analysis indicates distinct structural patterns in earthquake distributions across all clusters. Clusters-1, 2, and 4 show bimodal depth distributions, with seismicity concentrated in shallow ($\sim 8\text{--}12\text{ km}$) and intermediate ($\sim 28\text{--}35\text{ km}$) zones, indicating multiple active seismogenic layers. In contrast, Clusters-3 and 5 show a more focused and stable depth concentration near $\sim 30\text{--}35\text{ km}$, indicating a dominant structural interface controlling earthquake occurrence. Depth distributions remain largely unchanged as the magnitude threshold increases, indicating that earthquakes repeatedly nucleate within the same tectonically favorable depth ranges. The magnitude PDFs progressively shift toward higher values and become broader as the magnitude threshold increases. Smaller events are concentrated within a narrow range ($\sim 4.1\text{--}4.7$). Larger events ($M > 5$) extend to higher magnitudes ($\sim 5.2\text{--}6+$) with lower probabilities. Together, these results indicate that earthquake depths are governed by persistent tectonic structures, whereas earthquake magnitudes vary in response to stress accumulation, rupture processes, and fault interactions.

Despite these insights, some limitations remain. The analysis is based on cataloged earthquake depth and magnitude rather than raw waveform data. Fixed magnitude thresholds reduce the number of events at higher magnitudes and may affect the stability of the PSD

and PDF estimates. Resampling at 12-hour intervals may smooth short-term variations and influence the spectral characteristics. The interpretations are primarily statistical and are not constrained by detailed geodetic, geological, or three-dimensional structural models.

Acknowledgements

We extend my heartfelt gratitude to Dr. Anurag Sahay, Assistant Professor at NIT Patna, and Dr. Bappasona Kar, Assistant Professor at Panskura Banamali College, for their invaluable insights and guidance that greatly enriched the quality and depth of this research. Their expertise has played a pivotal role in shaping the direction of our work and has contributed significantly to our success. I would also like to express appreciation to Soumen Khatua, a research scholar at NIT Patna, for his unwavering support and encouragement throughout this journey.

Furthermore, we sincerely thank the anonymous reviewers for their thorough and constructive feedback, which has been instrumental in enhancing the clarity, scientific rigour, and overall relevance of this manuscript. Their thoughtful suggestions prompted deeper analysis and refinement of our interpretations, significantly improving the quality and impact of our study.

Funding and Conflict of Interest:

No funding was received for this research project. The authors declare no conflicts of interest.

Authors' Contributions

Debasri Samanta: Conceptualization, Methodology, Validation, Formal Analysis, Investigation, Resources, Data Curation, Writing – Original Draft, and Review & Editing.

Rajib Kumar Dolai: Conceptualization, Methodology, Software, Validation, Formal Analysis, Investigation, Resources, Data Curation, Writing – Original Draft, Review & Editing, Visualization, Coding, and Computational Analyses. ([e-mail:rajib@tmv.ac.in](mailto:rajib@tmv.ac.in))

Disclaimer: The opinions and conclusions expressed in this study are those of the authors and do not necessarily reflect the views of their affiliated organizations.

Reference

- Anthony, R.E., Ringler, A.T., Wilson, D.C., Bahavar, M. and Koper, K.D., 2020, "How processing methodologies can distort and bias power spectral density estimates of seismic background noise", *Seismological Research Letters*, **91**, 3, 1694–1706, <https://doi.org/10.1785/0220190212>

- Christian, J.T., 1989, "Generating seismic design power spectral density functions", *Earthquake Spectra*, **5**, 2, 351–368, <https://doi.org/10.1193/1.1585526>
- D'Alessandro, A., Greco, L., Scudero, S. and Lauciani, V., 2021, "Spectral characterization and spatiotemporal variability of the background seismic noise in Italy", *Earth and Space Science*, **8**, 10, e2020EA001579, <https://doi.org/10.1029/2020EA001579>
- Dewangan, P., Reddy, R., Raju, K.A.K., Singha, P., Aswini, K.K., Yatheesh, V., Samudrala, K. and Shuhail, M., 2018, "Nature of the ambient noise, site response, and orientation of ocean-bottom seismometers (OBSs): Scientific results of a passive seismic experiment in the Andaman Sea", *Bulletin of the Seismological Society of America*, **108**, 1, 248–259, <https://doi.org/10.1785/0120170163>
- Gupta, S., Borah, K. and Saha, G., 2016, "Continental-like crust beneath the Andaman Island through joint inversion of receiver function and surface wave from ambient seismic noise", *Tectonophysics*, **690**, 328–341, <https://doi.org/10.1016/j.tecto.2016.09.013>
- HariPriya, K., Radhakrishna, M. and Mukhopadhyay, M., 2020, "Crustal model for the Andaman Outer Arc: Constraints from earthquake, gravity and receiver function data", In: Ray, J.S. and Radhakrishna, M. (eds.), *The Andaman Islands and Adjoining Offshore, Society of Earth Scientists Series, Springer, Cham*, 161–181, http://doi.org/10.1007/978-3-030-39843-9_8
- Hartigan, J.A. and Wong, M.A., 1979, "Algorithm AS 136: A k-means clustering algorithm", *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, **28**, 1, 100–108, <https://doi.org/10.2307/2346830>
- Kagan, Y.Y., 2010, "Statistical distributions of earthquake numbers: Consequence of branching process", *Geophysical Journal International*, **180**, 3, 1313–1328, <https://doi.org/10.1111/j.1365-246X.2009.04487.x>
- Kamesh Raju, K.A., Aswini, K.K. and Yatheesh, V., 2020, "Tectonics of the Andaman Backarc Basin—Present understanding and some outstanding questions", In: Ray, J.S. and Radhakrishna, M. (eds.), *The Andaman Islands and Adjoining Offshore, Society of Earth Scientists Series, Springer, Cham*, 237–259, http://doi.org/10.1007/978-3-030-39843-9_12
- Kataria, N.P., Shrikhande, M. and Das, J.D., 2013, "Deterministic seismic hazard analysis of Andaman and Nicobar Islands", *Journal of Earthquake and Tsunami*, **7**, 4, Article 1350035, <https://doi.org/10.1142/S1793431113500358>
- Khan, P.K., Mohan, A. and Chowdhury, S., 2012, "Pre- and post-seismic activities along the Myanmar-Andaman-Sumatra Subduction Margin: Insights for tectonic segmentation", *Journal of the Indian Geophysical Union*, **16**, 3, 87–96. <https://iguonline.in/journal/Archives/16-3/1prasanta%20kumar.pdf>
- Kolathayar, S. and Sitharam, T.G., 2012, "Comprehensive probabilistic seismic hazard analysis of the Andaman-Nicobar regions", *Bulletin of the Seismological Society of America*, **102**, 5, 2063–2076, <https://doi.org/10.1785/0120110219>
- Lay, T., and Wallace, T. C., 1995, *Modern global seismology*, Academic Press.
- MacQueen, J., 1967, "Some methods for classification and analysis of multivariate observations", In: *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, **1**, 281–297. https://digicoll.lib.berkeley.edu/record/113015/files/math_s5_v1_article-17.pdf
- Manohar, C.S. and Sarkar, A., 1995, "Critical earthquake input power spectral density function models for engineering structures", *Earthquake Engineering & Structural Dynamics*, **24**, 12, 1549–1566, <https://doi.org/10.1002/eqe.4290241202>
- Mishra, O.P., Kayal, J.R., Chakraborty, G.K., Singh, O.P. and Ghosh, D., 2007, "Aftershock investigation in the Andaman-Nicobar Islands of India and its seismotectonic implications", *Bulletin of the Seismological Society of America*, **97**, 1A, S71–S85, <https://doi.org/10.1785/0120050629>
- Mishra, S. and Prajapati, S., 2019, "Crustal structure of Andaman Islands using joint inversion of receiver functions and surface wave dispersion measurements", *Annals of Geophysics*, **62**, 6, Article GT675, <https://doi.org/10.4401/ag-8063>
- Mukhopadhyay, B., Fnais, M., Mukhopadhyay, M. and Dasgupta, S., 2010, "Seismic cluster analysis for the Burmese-Andaman and West Sunda Arc: Insight into subduction kinematics and seismic potentiality", *Geomatics, Natural Hazards and Risk*, **1**, 4, 283–314, <https://doi.org/10.1080/19475705.2010.494014>
- Rahayu, T., Sipayung, R., Ainun, A.R., Sinambela, M. and Tarigan, K., 2020, "An application of power spectral density and probability density function for seismic mini region stations performance: Padang Lawas earthquake", In: 2020 3rd International Conference on Mechanical, Electronics, Computer, and Industrial Technology (MECnIT), 202–206, IEEE, <https://doi.org/10.1109/MECnIT48290.2020.9166654>
- Rao, P.S. and Rao, B.R., 1979, "Estimated earthquake probabilities in the northeast India and Andaman-Nicobar Islands", *Mausam*, **30**, 2–3, 267–273, <https://doi.org/10.54302/mausam.v30i2.3021>
- Rousseeuw, P.J., 1987, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis", *Journal of Computational and Applied Mathematics*, **20**, 53–65, [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7)
- Sadeghian, R., 2018, "Finding a probability density function for time intervals between earthquake occurrences", *Arabian Journal of Geosciences*, **11**, 1–10, <https://doi.org/10.1007/s12517-018-3777-y>
- Samanta, D., and Dolai, R. K., 2025, "Wave-Based Seismic Characterization of the 2025 Myanmar Earthquakes (Mw 7.7 and Mw 6.7)", *Journal of Earthquake Science and Soil Dynamics Engineering*, **8**, 3, <https://doi.org/10.5281/zenodo.17205067>
- Scholz, C. H., 2019, *The Mechanics of Earthquakes and Faulting*, 3rd ed., Cambridge University Press.
- Singh, S.C. and Moeremans, R., 2017, "Anatomy of the Andaman-Nicobar subduction system from seismic reflection data", *Geological Society, London, Memoirs*, **47**, 1, 193–204, <https://doi.org/10.1144/m47.13>
- Špičák, A., and Vaněk, J., 2013, "Earthquake clustering in the tectonic pattern and volcanism of the Andaman Sea region", *Tectonophysics*, **608**, 728–736. <https://doi.org/10.1016/j.tecto.2013.08.007>
- Srijayanthi, G. and Kumar, M.R., 2020, "Seismicity, lithospheric structure and mantle deformation in the Andaman Nicobar Subduction Zone", In: Ray, J.S. and Radhakrishna, M. (eds.), *The Andaman Islands and Adjoining Offshore, Society of Earth Scientists Series, Springer, Cham*, 107–136, https://doi.org/10.1007/978-3-030-39843-9_6
- Turcotte, D.L. and Schubert, G., 2014, *Geodynamics*, 3rd ed., Cambridge University Press. [https://www.eps.mcgill.ca/~courses/c510/%5BTurcotte_D.L.,_Schubert_G.%5D_Geodynamics\(Bookos.org\).pdf](https://www.eps.mcgill.ca/~courses/c510/%5BTurcotte_D.L.,_Schubert_G.%5D_Geodynamics(Bookos.org).pdf)

Welch, P., 1967, "The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms", *IEEE Transactions on Audio and Electroacoustics*, **15**, 2, 70–73, <https://doi.org/10.1109/TAU.1967.1161901>

Yadav, R. and Tiwari, V.M., 2020, "Lithospheric framework of Sumatra-Andaman Subduction Zone—A review", In: Ray, J.S. and

Radhakrishna, M. (eds.), *The Andaman Islands and Adjoining Offshore*, *Society of Earth Scientists Series*, Springer, Cham, 137–159, https://doi.org/10.1007/978-3-030-39843-9_7

Yoshimitsu, N., Maeda, T. and Sei, T., 2023, "Estimation of source parameters using a non-Gaussian probability density function in a Bayesian framework", *Earth, Planets and Space*, **75**, 1, Article 33, <https://doi.org/10.1186/s40623-023-01770-2>

