Journal of Geography and Cartography 2026, 9(1), 11794.
https://doi.org/10.24294/jgc11794

Article

Bimodality, magnitude, and spatial intensity: Revisiting the past Fiji earthquake with modern statistical tools

Subhrajit Saha, Debashis Chatterjee*

Department of Statistics, Visva-Bharati University, Santiniketan 731235, West Bengal, India
* **Corresponding author:** Debashis Chatterjee, debashis.chatterjee@visva-bharati.ac.in

CITATION

Saha S, Chatterjee D. Bimodality, Magnitude, and spatial intensity: Revisiting the past Fiji earthquake with modern statistical tools. Journal of Geography and Cartography. 2026; 9(1): 11794. https://doi.org/10.24294/jgc11794

ARTICLE INFO

Received: 12 June 2025
Accepted: 18 June 2026
Available online: 21 July 2026

Abstract: The Fiji–Tonga–Kermadec convergent margin represents one of the most seismically volatile regions globally, defined by complex subduction dynamics and a high density of deep-focus events. This study performs a high-resolution exploratory investigation of a 1000-event seismic catalog to resolve the intricate interplay between rupture magnitude, focal depth, and spatial distribution. Moving beyond baseline descriptive summaries, the study employs Exploratory Spatial Data Analysis (ESDA), including non-parametric 2D kernel density estimation and hexagonal binning, to delineate the latent tectonic architecture of the region. Our analysis identifies a pronounced bimodal depth signature, with focal concentrations in the shallow crust (<100 km) and the transition zone (>500 km). These findings empirically validate the thermal and phase-transformation mechanics of the Wadati–Benioffzone. Multivariate diagnostics further reveal a robust linear dependency $r = 0.71$ between seismic magnitude and the global monitoring footprint, providing a quantitative metric for energy-dependent detectability across the station network. By bridging classical geostatistical workflows with modern spatial intensity modeling, this research provides a reproducible, data-driven foundation for enhancing seismic hazard zoning. These insights are particularly critical for improving infrastructure resilience and disaster risk reduction (DRR) strategies in Small Island Developing States (SIDS) facing cascading tectonic risks.

Keywords: earthquake; Fiji Region; spatial statistics; kernel density estimation; exploratory spatial data analysis; multivariate analysis; seismology; tectonic hotspots

1. Introduction

The Tonga–Fiji–Kermadec subduction zone stands as a premier natural laboratory for seismology. It is widely recognized as one of the planet's most active convergent margins, notably producing roughly two-thirds of the world's deep-focus earthquakes at depths exceeding 300 KM [1–3]. These high-pressure events are not merely statistical outliers; they provide our best window into the mechanics of the deep mantle, including processes like phase transformations and dehydration embrittlement [3,4].

Past seismological investigations in this region have consistently pointed toward a clear bimodal depth distribution. This pattern separates shallow crustal activity (< 60 km) from the massive volume of deep-focus intraslab events (> 500 km), reflecting the steep and complex geometry of the local Wadati–Benioff zone [2,5]. Furthermore, the occurrence of earthquake multiplets and doublets, essentially near-simultaneous ruptures, suggests that fault reactivation remains a significant factor even at extreme depths [6,7].

Despite the wealth of data available from this region, the classic R quakes dataset remains an underutilized resource in modern spatial-statistical literature. This dataset, a curated 1,000-event subset of the Harvard PRIM-H catalog originally developed by John H. Woodhouse, offers a robust foundation for multivariate modeling [8,9]. However, much of the existing work on this data remains descriptive rather than analytical.

In this study, the study moves beyond simple data exploration to apply a rigorous, reproducible statistical workflow. Our specific objectives are to:

1. Evaluate the joint distribution of event coordinates, depths, and magnitudes to understand the internal structure of the catalog.
2. Investigate the physical drivers behind the observed bimodal depth distribution, specifically how it differentiates crustal versus mantle-level seismic triggers.
3. Map spatial intensity to pinpoint seismic hotspots; notably, these spatial patterns are further visualized, which highlights the tectonic alignment of the Fiji platform.
4. Model the link between event magnitude and the number of detecting stations, treating the station count as a proxy for energy propagation and signal attenuation.

By merging classical statistics with modern spatial intensity modeling, the study aims to clarify the seismic behavior of the Fiji region while championing the use of open-source, reproducible geoscientific methods [10,11]. All computations were performed in R, with the full source code hosted on GitHub to support further community research.

2. Literature review

The Fiji-Tonga-Kermadec region remains a critical focal point for investigating the mechanical heterogeneity of seismic events, particularly deep-focus ruptures that defy conventional tectonic expectations. While foundational research by Isacks established that deep-seated events often occur as "multiplets" rather than typical aftershock sequences, contemporary scholarship has significantly refined this view. Recent investigations into the 2018 Fiji deep earthquake doublet, for instance, demonstrate that such high-magnitude occurrences are driven by nuanced thermal interactions and the specific geometry of the subducting slab, rather than the monolithic "cold lab core" models previously favored [1,12,13].

2.1. Advances in source characterization and multi-sphere coupling

While historical waveform analyses emphasized the uniformity of stress drops in Fiji’s deep focus events, modern seismology has pivoted toward an integrated, multi-sphere approach [14,15]. A significant breakthrough involves the identification of trans-atmospheric signatures; high-resolution GNSS observations from the 2018 Fiji event revealed profound ionospheric anomalies, offering empirical support for the lithosphere-atmosphere-ionosphere coupling hypothesis [14]. This paradigm shift is further evidenced by recent developments in statistical ionospheric sensing, which seek to isolate pre-seismic signals from atmospheric noise using automated detection algorithms [15–17].

2.2. Structural heterogeneity and integrative geohazard mapping

The Fiji Platform is increasingly characterized as an actively deforming tectonic entity, seismically coupled with adjacent back-arc regions [18,19]. On Viti Levu, the presence of a complex fault network generates a continuous spectrum of activity, ranging from micro-seismicity to rare, high-consequence events [18–20].

Managing these risks has necessitated a transition toward GIS-integrated Disaster Risk Reduction (DRR). [11] highlighted that the spatial delineation of seismic hotspots is a prerequisite for mitigating cascading hazards such as landslides and tsunamis. Furthermore, recent probabilistic frameworks by [18] and updated tsunami propagation models by [19] advocate for multi-hazard resilience strategies to safeguard coastal infrastructure against the volatile nature of Pacific plate boundaries. These approaches increasingly emphasize the integration of spatial analytics, hazard forecasting, and resilience planning within vulnerable island environments [17,21].

2.3. Statistical innovation and machine learning integration

The methodological evolution from standard Exploratory Data Analysis (EDA) to Exploratory Spatial Data Analysis (ESDA) has allowed for a deeper resolution of latent patterns in seismic catalogs [8,13]. Recent advancements have utilized non-parametric intensity estimation and graph-theory-based network modeling to address the perennial issue of data incompleteness [8,13]. Moreover, the integration of machine learning for seismotectonic regionalization has allowed for more precise cluster identification in the Fiji region, moving beyond simple linear regression to capture non-linear geophysical dependencies [12,22].

The broader socio-political implications of these seismic events have also come to the fore. Major ruptures are now recognized as "focusing events" that catalyze radical shifts in national policy and building standards [20]. By synthesizing advanced statistical workflows, such as the spatial kernel density models proposed by [8], with conceptual frameworks for community resilience, researchers can deliver the reproducible, data-driven platforms required to protect island nations from catastrophic energy releases [17,21,22].

3. Motivation and objectives

The immense power of earthquakes to reshape landscapes and disrupt human infrastructure makes seismic risk assessment a critical pillar of modern geosciences [18,19,21]. Because seismic events are inherently unpredictable, the study must rely on historical instrumental data to build robust disaster preparedness frameworks [17,20]. This study is driven by the necessity of translating raw seismic records into a coherent understanding of tectonic behavior. By utilizing the quakes dataset—a high-quality sample of 1,000 recorded events in the Fiji region—the study provides a rigorous, data-centric exploration of one of the world's most complex seduction zones. Our goal is to extract meaningful patterns from these measurements that can inform both theoretical geophysical research and practical emergency management.

To achieve this, the study focuses on four primary objectives:

(i) Statistical Profiling: The study systematically summarizes the distributions of earthquake magnitude, depth, and spatial coordinates to establish the baseline characteristics of the Fiji-Tonga catalog.
(ii) Visual Evidence: The study develops a suite of analytical graphics, ranging from density estimations and histograms to high-resolution geographical maps, to visualize the physical structure of regional seismicity.
(iii) Variable Interaction: The study quantifies the interplay between earthquake variables, with a specific focus on how magnitude and focal depth influence the sensitivity of the global station network.
(iv) Spatial Intensity and Risk: The study identifies specific clusters and spatial anomalies within the dataset. By pinpointing these seismic "hotspots," the study contributes data-driven insights that can refine current regional risk models and improve infrastructure resilience.

4. The dataset

Our analysis focuses on the quakes dataset, a foundational geophysical record included in R's base datasets package. This collection comprises 1,000 instrumental observations of seismic activity near Fiji, spanning events from 1964 to the present. The dataset is a curated subset of the Harvard PRIM-H catalog, originally managed by Dr. John H. Woodhouse [9]. While the full catalog is extensive, this 1,000-event sample was developed to illustrate the complexities of spatial and multivariate geophysical modeling.

4.1. Data provenance and spatial scope

The geographic focus of the data is a rectangular corridor roughly bounded by latitudes $[-39°, -10°]$ and longitudes $[165°, 190°]$. This region is geophysically significant as it encompasses the high-energy convergence zone of the Indo-Australian and Pacific Plates. Unlike many seismic datasets that focus exclusively on shallow crustal activity, the quake data includes deep-focus events reaching depths beyond 600 km, offering a rare look at mantle-level seismicity within the Fiji-Tonga-Kermadec subduction system.

4.2. Formalization of variables

The study treats the dataset as a random sample $\mathcal{X}$ of $n = 1000$ observations. Each seismic event X_i is represented as a vector in R^5:

$$\mathcal{X} = \{X_1, \dots, X_n\}, \quad X_i = (\lambda_i, \phi_i, d_i, m_i, s_i)^\top \in R^5,$$

where the variables are defined as:
(i) $\lambda_i(lat)$:Epicentral latitude in decimal degrees south.
(ii) $\phi_i(long)$: Epicentral longitude in decimal degrees east.
(iii) $d_i(depth)$: Focal depth measured in kilometers from the surface.
(iv) $m_i(mag)$: Richter magnitude, reflecting the logarithmic scale of seismic energy.
(v) $s_i(stations)$: The count of reporting stations, used here as an indicator of observational coverage and signal detectability.

All recorded events maintain a magnitude $m > 4.0$. While the data are static and preprocessed for statistical use, they represent real-world instrumental recordings that facilitate a deep dive into the seismotectonic regime of the South Pacific. Further documentation is maintained by the R development team at: https://stat.ethz.ch/R-manual/R-devel/library/datasets/html/quakes.html.

4.3. Scientific objectives

The core of our inquiry is to characterize the joint probability distribution P (X) through three specific lenses:

1. Marginal Properties: Estimating the distribution for each parameter to identify unique features like the bimodal depth profile.
2. Spatial Point Processes: Mapping the density of epicenters to distinguish between background seismicity and high-intensity "hotspots."
3. Dependency Structures: Quantifying the link between the physical energy of an event (M) and its resulting observational footprint (S).

This workflow demonstrates how classical statistical formalisms can be applied to clarify complex Earth science phenomena.

5. Methodology

In this study, the study formalizes the *quakes dataset* as a random sample $\mathcal{X} = \{X_1, \ldots, X_n\}$containing $n = 1000$ seismic observations. Each individual event $X_i \in R^5$is characterized by a feature vector comprising spatial, physical, and observational dimensions:

$$X_i = \begin{bmatrix} \lambda_i \\ \phi_i \\ d_i \\ m_i \\ s_i \end{bmatrix}$$

where λ_i and ϕ_i denote the latitude and longitude, d_i represents focal depth in kilometers, m_i is the Richter magnitude, and s_i is the discrete count of detecting seismic stations.

5.1. Multivariate moments and correlation structure

To uncover latent dependencies within the Fiji seismic catalog, the study first evaluates the empirical mean vector $\widehat{\boldsymbol{\mu}}$ and the sample covariance matrix $\widehat{\boldsymbol{\Sigma}}$:

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i, \tag{1}$$

$$\hat{\Sigma} = \frac{1}{n-1} \sum_{i=1}^{n} (\boldsymbol{X_i} - \widehat{\boldsymbol{\mu}})(\boldsymbol{X_i} - \widehat{\boldsymbol{\mu}})^{\top}. \tag{2}$$

Standardizing these results allows us to construct the correlation matrix $R = [R_{jk}]_{1 \le j,k \le 5}$, where each element is defined as:

$$, \quad R_{jk} = \frac{\widehat{\Sigma_{jk}}}{\sqrt{\widehat{\Sigma_{jj}\Sigma_{kk}}}}.$$

This matrix is vital for isolating the linear relationship between magnitude (m_i) and the resulting detection signature (s_i). A high R_{ms} coefficient would validate the use of station counts as a reliable proxy for the energy released during an event.

5.2. Nonparametric marginal estimation

Rather than assuming a specific distribution for earthquake depth or magnitude, the study utilizes Kernel Density Estimators (KDE) to approximate the underlying density $f_X(x)$:

$$\widehat{f_h}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right), \quad K(\cdot) \in \mathcal{K},$$

Here, K represents a symmetric Gaussian kernel, and h is a smoothing bandwidth determined via cross-validation. This approach is particularly useful for identifying the bimodal structure often observed in focal depths. To quantify the deviation from symmetry and confirm the presence of distinct seismic regimes (crustal vs. deep-focus), the study calculates the empirical skewness:

$$\text{Skewness}(D) = \frac{\frac{1}{n}\sum_{i=1}^{n}\left(d_i - \bar{d}\right)^3}{\left(\frac{1}{n}\sum_{i=1}^{n}\left(d_i - \bar{d}\right)^2\right)^{3/2}}.$$

5.3. Spatial intensity modeling

The epicentral coordinates $\{(\lambda_i, \phi_i)\}_{i=1}^{n} \subset R^2$ represent a spatial point process $\mathcal{P}$ distributed over a geographic domain $\mathcal{D}$. The study models the spatial intensity function $\lambda(s)$, the expected event density per unit area, using a bivariate Gaussian kernel K_2:

$$\widehat{\lambda_h}(s) = \frac{1}{nh^2}\sum_{i=1}^{n} K_2\left(\frac{s - s_i}{h}\right),$$

This allows us to visually and mathematically delineate seismic "hotspots," providing a data-driven basis for tectonic interpretation in the Fiji region.

5.4. Linear dependence and detection sensitivity

To test the hypothesis that seismic energy propagation follows a predictable attenuation pattern, the study models the station count S_i as a function of magnitude M_i:

$$S_i = \beta_0 + \beta_1 M_i + \varepsilon_i, \qquad \varepsilon_i \overset{iid}{\sim} \mathcal{N}(0, \sigma^2),$$

The coefficient β_1 indicates the sensitivity of the global monitoring network to changes in event magnitude. The study estimatê $\hat{\beta}$ via the least square method:

$$\hat{\beta} = \left(\mathbf{X}^{\mathsf{T}}\mathbf{X}\right)^{-1}\mathbf{X}^{\top}\mathbf{S}$$

Model validity is subsequently checked using Cook's distance and residual diagnostic plots to ensure the results are not biased by influential outliers.

5.5. Analytical graphics suite

Our visualization strategy, implemented via the *ggplot2* framework in R, uses a multi-layered approach to explore the data:

(i) Distributions: Histograms paired with KDE overlays to analyze marginal behavior.

(ii) Spatial Patterns: Geospatial heatmaps and hexbin plots to resolve localized epicentral density.

(iii) Associations: Correlograms for pairwise R_{jk} values and faceted plots to examine the(d_i, m_i) relationship across different magnitude classes.

6. Results

The analysis of the 1,000 seismic events reveals a highly structured environment defined by intense tectonic stress and distinct depth regimes. This section transitions from a baseline numerical summary to high-dimensional spatial and multivariate visualizations that characterize the Fiji-Tonga-Kermadec subduction system.

6.1. Quantitative baseline and variable dependencies

The study first establishes the baseline properties of the catalog to understand the scale of regional seismicity. **Table 1** provides the central tendencies and dispersions for the five core variables. The median magnitude of 4.60 suggests a regime dominated by moderate-intensity events, while the depth statistics (mean = 311.4 km) reflect a concentration of deep-focus activity that is rare in most other global subduction zones.

The initial statistical overview and the correlation matrix provide a foundational understanding of the dataset.

Table 1. Summary statistics for the *quakes* dataset variables.

Statistic	lat	long	depth	mag	stations
Min.	−38.59	165.7	40.0	4.00	10.0
1st Qu.	−23.47	179.6	99.0	4.30	22.0
Median	−20.30	181.4	181.4	4.60	33.0
Mean	−20.64	179.5	179.5	4.62	39.2
3rd Qu.	−17.67	183.2	183.2	4.90	53.0
Max.	. −10.72	188.1	188.1	6.40	132.0

The Pearson correlation matrix (**Table 2**) identifies the magnitude-station link (r = 0.71) as the most robust dependency. This is physically significant, as it validates station count as an observational proxy for seismic energy release. Interestingly, the near-zero correlation between depth and magnitude indicates that the physical mechanism of energy release in this region is relatively constant, whether the rupture occurs in the crust or the deep mantle.

Table 2. Pearson correlation matrix of the dataset variables.

Variable	lat	long	depth	mag	stations
lat	1.00	0.22	0.03	−0.15	−0.21
long	0.22	1.00	0.07	−0.13	−0.12
depth	0.03	0.07	1.00	−0.02	−0.14
mag	−0.15	−0.13	−0.02	1.00	0.71
stations	−0.21	−0.12	−0.14	0.71	1.00

6.2. Distributions and seismogenic modality

The marginal distributions for magnitude (**Figure 1**) and depth (**Figure 2**) exhibit contrasting behaviors. While magnitudes follow a standard exponential decay, the depth profile is explicitly bimodal. The study observes a primary peak at shallow crustal levels (0–100 km) and a secondary surge in the transition zone (500–650 km). This bimodality marks the transition from brittle crustal fracture to phase-transformation-induced stresses within the subducting lithosphere.

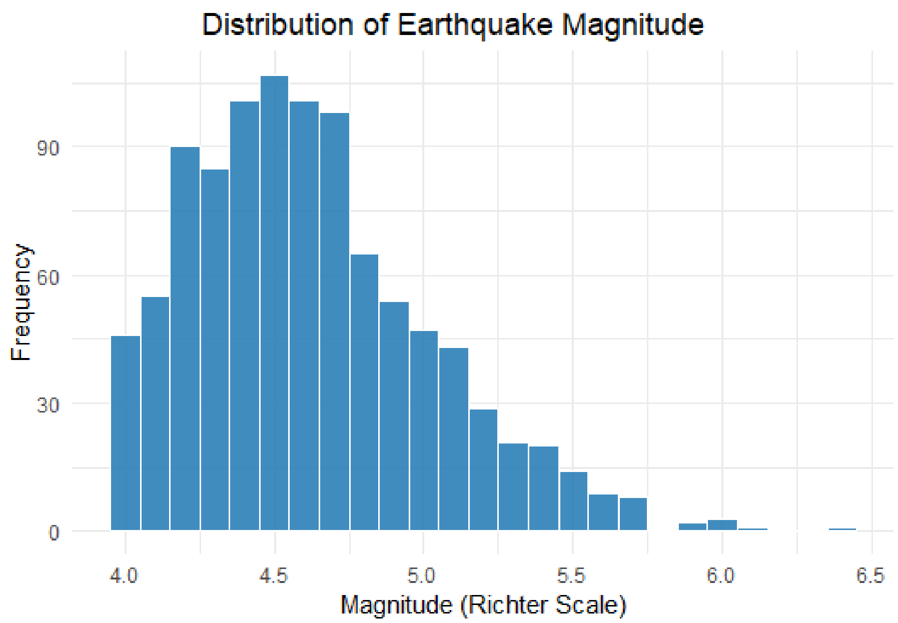

Figure 1. Frequency distribution of earthquake magnitudes.

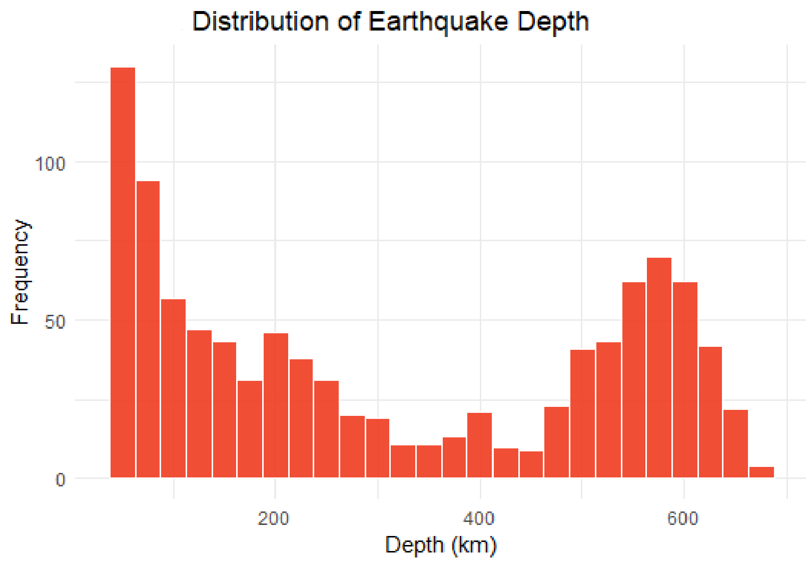

Figure 2. Focal depth distribution illustrating the bimodal seismic signature.

6.3. Spatial intensity and tectonic mapping

To understand the geographic drivers of this activity, the study refers to the epicentral coordinates plotted in **Figure 3**. The clustering of blue points in **Figure 3** clearly traces the subduction interface where the Pacific Plate descends beneath the Indo-Australian Plate. This alignment highlights a well-defined seismic arc extending from the Tonga Trench toward the Fiji platform.

Figure 4 reinforces this, showing that seismicity is not random but strictly constrained by regional structural geometry. Further resolution is provided by density heatmaps (**Figure 5**) and hexbin mapping (**Figure 6**). Both visualizations identify a major seismic "hotspot" near the 180° meridian. The high density in this corridor suggests a zone of maximum deformation.

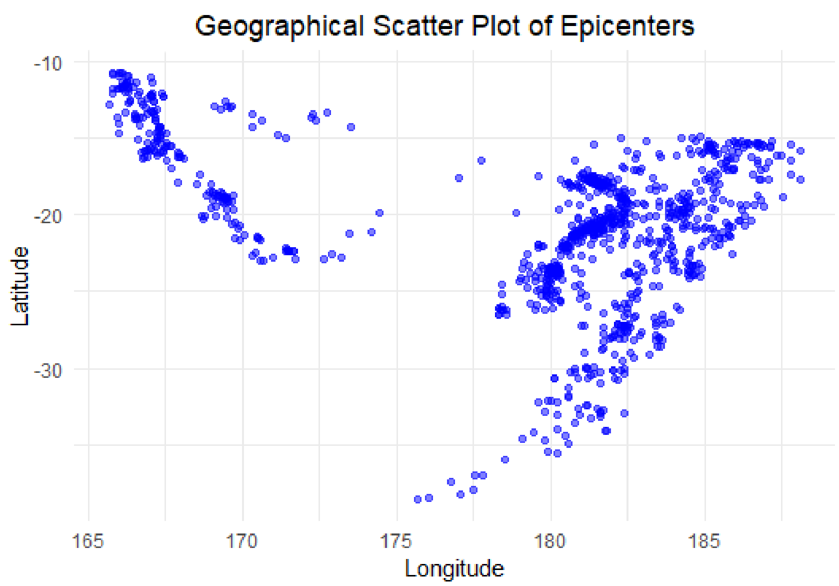

Figure 3. Geographical scatter plot of epicenters illustrating tectonic alignment.

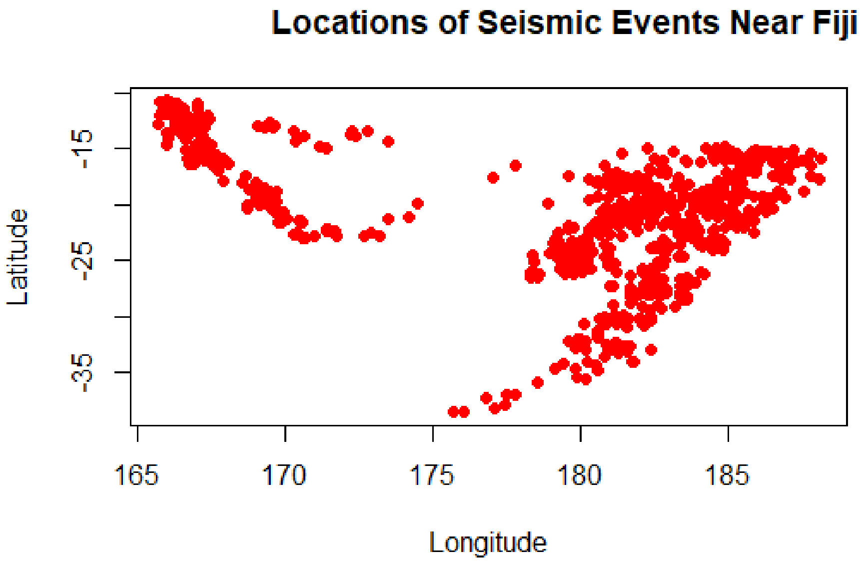

Figure 4. Locations of Seismic Events Near Fiji.

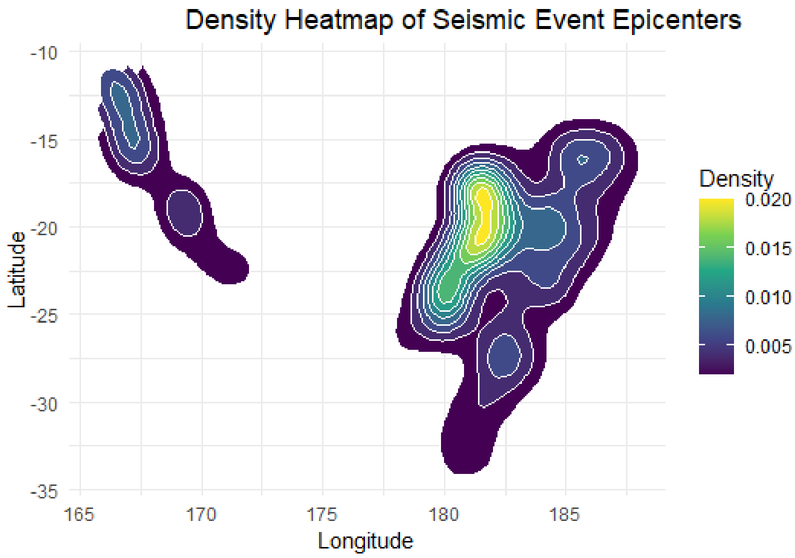

Figure 5. 2D Kernel Density Estimation of seismic event epicenters.

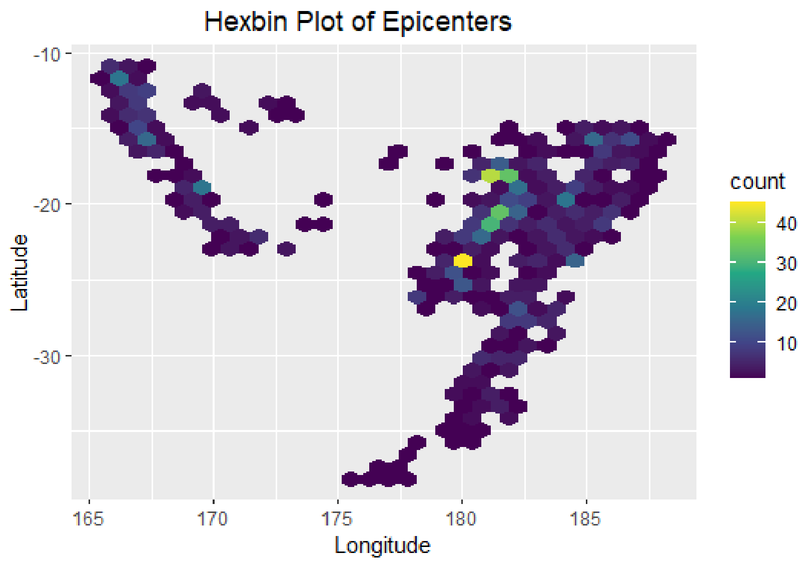

Figure 6. Hexbin plot providing localized intensity estimation.

6.4. Energy propagation and reporting sensitivity

The relationship between physical energy and detectability is explored in **Figure 7**. The strong linear increase in stations for higher-magnitude events reflects the broader propagation of seismic waves. This link is further quantified in the correlogram (**Figure 8**), where r = 0.85 represents the strongest statistical relationship in the study.

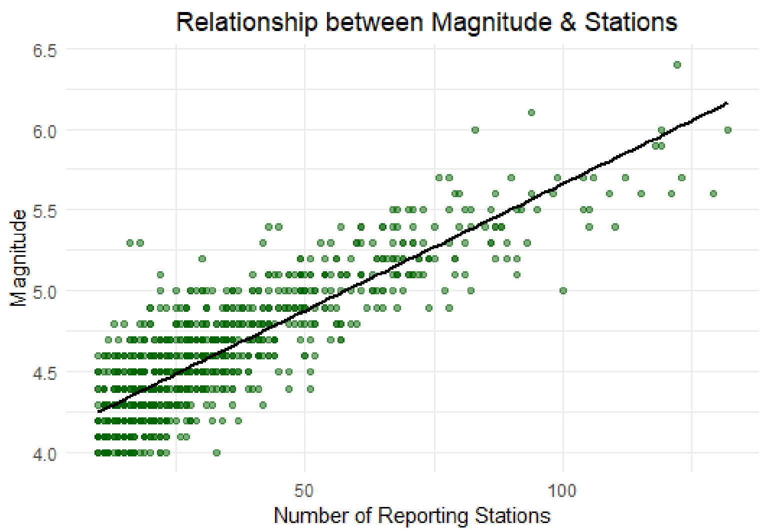

Figure 7. Regression analysis of magnitude vs. reporting stations.

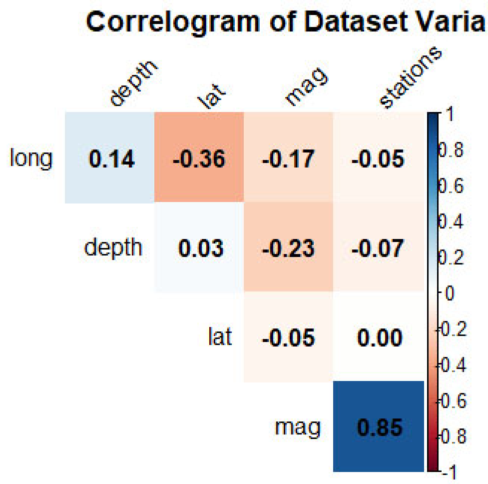

Figure 8. Variable correlogram displaying Pearson coefficients.

Finally, the faceted analysis **(Figure 9)** and magnitude-depth scatter **(Figure 10)** examine whether energy release is depth-dependent. The consistency of magnitude ranges across all depths implies that deep-focus events are as capable of generating significant seismic moments as shallow ones, posing a unique challenge for regional hazard models.

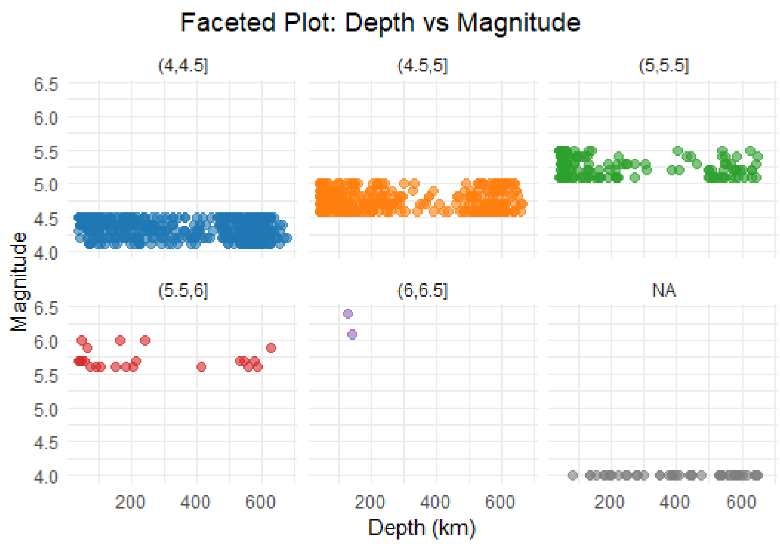

Figure 9. Faceted plot of earthquake depth vs. magnitude bins.

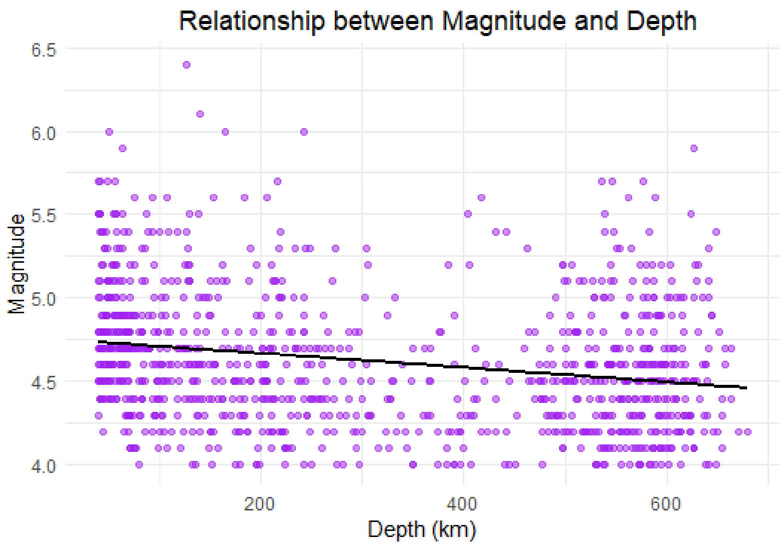

Figure 10. Bivariate scatter plot of magnitude vs. depth.

7. Discussion

The results of this study, derived from a systematic analysis of the quakes dataset, offer a window into the complex seismic architecture of the Fiji-Tonga region. Rather than treating the 1,000 events as isolated data points, our multivariate approach allows us to interpret mathematical patterns as direct signatures of underlying geophysical mechanisms.

A primary feature of our analysis is the pronounced bimodality in the estimated marginal density for earthquake depth, $\hat{f}(d)$. Statistically, this distribution represents the superposition of two distinct seismogenic regimes. The initial peak, clustered at depths $(d < 100\ \text{km})$, captures the brittle crustal activity typical of the overriding plate and the shallow subduction interface. However, the secondary peak at $d > 500\ \text{km}$ is more diagnostic; it serves as a clear statistical signature of a Wadati–Benioff zone. This suggests that the Pacific Plate's subduction beneath the Indo-Australian Plate is not a continuous, uniform descent but one where stress release is concentrated in specific thermal or phase-change windows within the subducting slab [2].

The spatial analysis, particularly the 2D kernel density estimates and the epicentral mapping in **Figure 3**, reinforces this tectonic narrative. The clustering of seismic events is not random; the points align along curvilinear features that trace the active plate boundary and deep-seated fault systems of the Fiji platform. These "hotspots," further resolved in the density heatmap (**Figure 5**) and hexbin plot (**Figure 6**), represent regions where accumulated tectonic stress is preferentially released. Identifying these zones is a critical step for regional hazard assessment and disaster risk reduction (DRR) strategies [11].

One of the most robust statistical results is the strong positive Pearson correlation $\left(R_{m,s} \approx 0.71\right)$ between earthquake magnitude M and the number of reporting stations S. This relationship transcends mere correlation and has a clear physical basis in wave propagation. As magnitude is a logarithmic scale of energy release, higher-magnitude events generate high-amplitude seismic waves that maintain a signal-to-noise ratio above the detection threshold over much larger distances. Consequently, the linear relationship observed between M and S provides an empirical validation of seismic attenuation physics; larger events simply have a wider "detection footprint" across the global monitoring network.

Finally, the weak correlations observed between geographic coordinates and physical parameters like depth and magnitude suggest that energy release is not a simple function of latitude or longitude. Seismicity here is likely governed by local heterogeneities, variations in slab temperature, rock mineralogy, and complex fault geometries that are not captured by surface coordinates alone. This underscores the necessity of high-dimensional datasets and modern statistical tools, such as those implemented in our R-based workflow, to uncover the latent structures of the Earth's interior that traditional linear models might overlook.

By merging classical exploratory data analysis with spatial intensity modeling, this work contributes to a quantitative understanding of Fiji's seismicity while promoting a transparent, reproducible framework for geoscientific research.

8. Conclusion

This exploratory interrogation of the *quakes* dataset has successfully characterized the seismic architecture of the Fiji region through 1,000 instrumental observations. By merging classical statistical moments with spatial intensity modeling, the study has extracted a multi-layered understanding of regional seismicity that moves beyond simple descriptive summaries.

The study identifies several defining seismogenic signatures that clarify the nature of tectonic activity in the South Pacific. First, the magnitude profile confirms a high frequency of moderate-intensity events (magnitude 4.0–5.0), establishing a baseline for regional background seismicity. More significantly, the depth distribution exhibits a stark bimodality, with focal points concentrated in the shallow crust (< 100 km), and again within the deep-mantle transition zone (> 500 km),. This statistical duality provides clear evidence of the complex tectonic plumbing of the Fiji-Tonga subduction system, where stress is released at both the brittle plate interface and within the dehydrating subducting slab.

Geospatially, our analysis confirms that activity is not stochastic but is instead governed by structural constraints. The identified seismic "hotspots" and curvilinear clusters align with known tectonic boundaries, providing essential data for local hazard zonation. Furthermore, the robust linear association ($r = 0.71$) between magnitude and station count offers a quantitative validation of wave propagation physics; the larger the energy release, the wider the observational footprint across the global station network. This confirms that station count is a reliable, albeit indirect, proxy for the seismic significance of an event in data-limited scenarios.

Ultimately, this work demonstrates the power of the R environment for open-source, reproducible geoscientific research. By integrating density estimations, spatial heatmaps, and regression diagnostics, the study has turned raw instrumental data into actionable geophysical insights. These findings not only enhance our understanding of the seismic behavior in the Fiji region but also underscore the importance of data-driven approaches in seismology, disaster preparedness, and the mitigation of seismic risks in one of the world's most active convergent margins.

9. Limitations

Although the *quakes* dataset provides a useful foundation for examining seismic patterns in the Fiji–Tonga region, several limitations should be considered when interpreting the results.

9.1. Geospatial and temporal constraints

One important limitation is the geographic focus of the dataset. Since all recorded events originate from the Fiji–Tonga–Kermadec region, the findings primarily reflect the tectonic characteristics of this particular subduction system and may not be directly applicable to other tectonic environments.

Another important limitation is the lack of high-resolution temporal information. The dataset records the locations and characteristics of earthquake events but does not provide detailed timestamps that would allow the evolution of seismicity to be examined through time. As a result, the present study is limited to a spatial and static assessment of earthquake patterns. Analyses of aftershock sequences, seismic recurrence intervals, and temporal clustering of events could not be performed. These temporal characteristics play an important role in understanding earthquake occurrence and are frequently used in seismic hazard forecasting. Consequently, the findings should be interpreted as describing the spatial distribution and statistical properties of the available earthquake records rather than the temporal dynamics of regional seismic activity.

9.2. Data completeness and variable scope

The *quakes* dataset contains 1,000 earthquake events selected from the larger Harvard PRIM-H catalogue and distributed as a built-in dataset in R. Because the dataset was prepared primarily for statistical analysis and teaching purposes, it is not possible to determine whether the selected sample fully represents the long-term seismic behavior of the Fiji–Tonga region. Consequently, some of the patterns observed in this study, including the distributions of earthquake magnitudes and focal depths, may be influenced by the composition of the sample itself. Therefore, the results should be interpreted as an exploratory assessment of the available data rather than a complete characterization of regional seismicity. Another limitation relates to the magnitude range represented in the dataset. The recorded earthquakes have magnitudes between 4.0 and 6.4, meaning that very small micro-seismic events as well as rare great earthquakes are not included. As a result, the dataset does not capture the full spectrum of seismic activity in the region. While it is suitable for exploring general seismic patterns, it may not fully reflect the behavior of low-magnitude background seismicity or the extreme events that often contribute most significantly to seismic hazard assessments. The dataset is also limited by the availability of seismological variables. While it includes earthquake location, focal depth, magnitude, and station coverage, it does not provide focal mechanism parameters such as strike, dip, and rake. These parameters are important for understanding fault motion, stress orientation, and earthquake source processes. In addition, information on fault geometry, regional stress fields, and lithological characteristics is unavailable. The absence of these variables limits our ability to interpret the observed seismic patterns from a seismotectonic perspective and to relate them directly to the physical mechanisms driving earthquake activity in the Fiji–Tonga subduction zone. Future studies incorporating more comprehensive earthquake catalogues and geophysical datasets would enable a more detailed understanding of regional tectonic processes.

9.3. Analytical simplification

The statistical methods used in this study are primarily exploratory and rely on correlation analysis and linear regression. While these approaches are useful for identifying broad patterns and relationships, earthquake processes are often governed by complex and non-linear interactions that may not be fully captured by linear models. More advanced statistical and machine-learning techniques may therefore reveal additional structures and dependencies within the data.

These limitations should be kept in mind when interpreting the findings. At the same time, they highlight several opportunities for future research using larger datasets, richer geophysical information, and more advanced analytical approaches.

10. Future directions

To advance the study of Fiji's seismic patterns and overcome the limitations of the current dataset, the following research avenues are proposed:

- Spatiotemporal Integration: Future studies should incorporate high-resolution timestamps to move beyond static spatial mapping. This will enable the modeling of aftershock sequences and the calculation of seismic recurrence intervals within the Tonga-Kermadec subduction zone.
- Advanced Predictive Modeling: Transitioning from linear correlations to machine learning architectures—such as Random Forests or Neural Networks—could better resolve the non-linear relationships between focal depth and energy release (M_w).
- Data Fusion and Geospatial Context: Combining the *quakes* dataset with GPS-derived tectonic plate velocities and bathymetric data would provide a deeper understanding of the slab-pull mechanics driving deep-focus events.
- Uncertainty and Risk Communication: Future work should prioritize Bayesian uncertainty quantification to refine seismic "hotspot" identification, facilitating more reliable evidence-based decision making for disaster management in South Pacific island nations.

These steps are essential for evolving this exploratory study into a robust predictive framework for regional geohazard mitigation.

Funding: The authors confirm that no external funding or financial grants from government, commercial, or non-profit organizations were received for the conduct of this research.

Conflict of interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Author contributions: The authors contributed equally to the conceptualization, methodology, analysis, and writing of this manuscript.

Availability of data: The empirical basis for this study is the quakes dataset, a foundational geophysical record included in the R statistical environment's base datasets package. This collection consists of 1,000 seismic observations from the Fiji region, originally extracted from the Harvard PRIM-H catalogue curated by Dr.~John H. Woodhouse. Researchers can access the data directly within any R session (version 4.0 or later) by executing the command data(quakes). Comprehensive metadata and technical documentation are maintained by the R Development Core Team at: https://stat.ethz.ch/R-manual/R-devel/library/datasets/html/quakes.html.

Further disclosure: To ensure the reproducibility of our statistical and geospatial findings, the complete analytical workflow has been made public. All R scripts used for data cleaning, spatial intensity estimation, and multivariate visualization—utilizing libraries such as ggplot2, MASS, and corrplot—are hosted in the following GitHub repository: https://github.com/debashisdotchatterjee/Earthquake-Fiji-Classical.

References

1. Jia Z, Shen Z, Zhan Z, Li C, Peng Z, Gurnis M. The 2018 Fiji Mw 8.2 and 7.9 deep earthquakes: One doublet in two slabs. Earth and Planetary Science Letters. 2020; 531: 115997.
2. Richter D, et al. When slabs collide: A tectonic assessment of deep earthquakes in the Tonga subduction zone. Geology. 2011; 39(8): 787-790.
3. Yu WC, Wen L. Deep-focus repeating earthquakes in the Tonga–Fiji subduction zone. Bulletin of the Seismological Society of America. 2012; 102(4): 1829-1849.
4. Jia J, Ye W. Deep learning for earthquake disaster assessment: objects, data, models, stages, challenges, and opportunities. Remote Sensing. 2023; 15(16): 4098.
5. Trott M, Pelletier M, Harrap R, Leybourne M, Layton-Matthews D. Open data to visualize subduction zones and improve large scale prospectivity interpretation: A case study examining the Wadati-Benioff zone of Cocos-Caribbean subduction in Central America. Applied Computing and Geosciences. 2026; 100352.
6. Gibowicz SJ, Lasocki S. Earthquake doublets and multiplets in the Fiji–Tonga–Kermadec region. Acta Geophysica Polonica. 2005; 53(3): 239-274.
7. Lorito S, Romano F, Lay T. Tsunamigenic major and great earthquakes (2004–2013): source processes inverted from seismic, geodetic, and sea-level data. In: Complexity in Tsunamis, Volcanoes, and their Hazards. Springer; 2021:247-298.
8. Saha S, Chatterjee D. Statistical analysis of seismicity in the Fiji-Tonga region: Integrating classical and circular methods. Journal of Statistical Research. 2025; 59(2): 145-166.
9. R Core Team. quakes: Locations of earthquakes off Fiji [dataset]. R package datasets; 2023.
10. Ansari A, Zaray AH, Wahidi AW, et al. Post-event investigation of multiscale failure patterns associated with the Mw 6.0 2025 Nurgal-Kunar earthquake, Afghanistan. Natural Hazards. 2026; 122(5): 189.
11. Varo J, Sekac T, Jana SK, Pal I. GIS perspective hazard risk assessment: A study of Fiji Island. In: Disaster Resilience and Sustainability. Elsevier; 2021:197-238.
12. Kumar S, Sengupta A. Application of ML-based approach for co-seismic landslides susceptibility mapping and identification of important controlling factors in eastern Himalayan region. Environmental Earth Sciences. 2024; 83(21): 603.

13. Yates A, Caudron C, Lesage P, Mordret A, Lecocq T, Soubestre J. Assessing similarity in continuous seismic cross-correlation functions using hierarchical clustering: application to Ruapehu and Piton de la Fournaise volcanoes. Geophysical Journal International. 2023; 233(1): 472-489.
14. Kiyani A, Shah M, Ahmed A, et al. Seismo ionospheric anomalies possibly associated with the 2018 Mw 8.2 Fiji earthquake detected with GNSS TEC. Journal of Geodynamics. 2020; 140: 101782.
15. Li W, Jiang Y, Ji H, Wei W. Amplitude scintillation detection with geodetic GNSS receivers leveraging machine learning decision tree. Satellite Navigation. 2024; 5(1): 18.
16. Patel AS. Lightning-Induced NOx Production over Canada: A Multivariate Earth Observation Analysis [dissertation]. Carleton University; 2023.
17. Shah KU, Kelman I. Data management for sustainable development: case analyses of small island developing states (SIDS). Development Policy Review. 2025; 43(4): e70017.
18. Fakhruddin B, Kintada K, Tilley L. Probabilistic tsunami hazard and exposure assessment for the Pacific Islands-Fiji. International Journal of Disaster Risk Reduction. 2021; 64: 102458.
19. Singh A, Mishra O, Pandey AP, Bhatla R. Seismic signature of the super cyclone Amphan in Bay of Bengal using coastal observatories operating under National Seismological Network of India. Earth and Space Science. 2024; 11(4): e2023EA003191.
20. Wanner MS. Change in policy regimes for disaster risk reduction in Fiji and Nepal. International Journal of Disaster Risk Reduction. 2022; 77: 103030.
21. Oh-Seng JR, Klöck C, Deenapanray P. Conceptualising a governance framework for climate change adaptation and disaster risk reduction in small island developing states through a systematic review. Sustainability. 2025; 17(22): 9965.
22. Ansari A, Alluqmani AE, Huang Z, et al. Fuzzy synthetic approach for seismic risk assessment of bridges with insights from the 2023 Kahramanmaras earthquake in Turkiye. Scientific Reports. 2025; 15(1): 13418.