STATISTICAL ANALYSIS OF SEISMICITY IN THE FIJI-TONGA REGION: INTEGRATING CLASSICAL AND CIRCULAR METHODS

SUBHRAJIT SAHA

Department of Statistics, Visva-Bharati University, Santiniketan - 731235, India

Email: subhrajitsaha553@gmail.com

DEBASHIS CHATTERJEE*

Department of Statistics, Visva-Bharati University, Santiniketan - 731235, India

Email: debashis.chatterjee@visva-bharati.ac.in

SUMMARY

This study presents a comprehensive exploratory and directional analysis of 1,000 geo-referenced earthquake events from the seismically active Fiji–Tonga region, based on the classical `quakes` dataset in R. Integrating both multivariate and circular statistical methods, the analysis investigates the spatial, depth-dependent, and directional structure of seismicity along a complex subduction zone. Classical techniques—including kernel-based density estimation, spatial point process modeling, and correlation analysis—reveal a distinctly bimodal depth distribution and a strong linear relationship between magnitude and the number of reporting stations. The estimated spatial intensity surface highlights seismic hotspots consistent with subduction interfaces and tectonic boundaries. To capture directional behavior, a unified circular statistical framework is introduced, incorporating bearing computation, uniformity tests (Rayleigh, Kuiper, Watson), circular-linear regression, and finite mixture modeling using mixed von Mises distributions. This enables decomposition of complex directional patterns into interpretable fault-related clusters and identification of depth-direction coupling. Applied to the Fiji trench, the method detects SW–NE bimodality, a dominant orientation near $127°$, a secondary cluster near $292°$, and significant regression effects ($\hat{\beta} = 0.036$, $p < 0.001$). The findings underscore the utility of classical datasets for modern geostatistical workflows and highlight the value of open-access seismic data in understanding global tectonic processes. All analyses were performed using reproducible code.

Keywords and phrases: Seismicity Analysis, Circular Statistics, Spatial Point Process, von Mises Mixture Model, Subduction Zone Tectonics, Earthquake Directionality.

AMS Classification: 62H11, 62P12, 86A15

---

1 Introduction

Earthquakes are complex geophysical phenomena governed by interactions between stress accumulation, fault geometry, and deep tectonic processes. In highly active convergent margins such as the Tonga–Fiji–Kermadec subduction zone—one of the world's most seismically dynamic regions, generating nearly two-thirds of all deep-focus earthquakes—understanding the spatial and directional organization of seismicity provides fundamental insight into subduction dynamics, stress-field orientation, and slab morphology (Green and Houston, 1995; Okal and Kirby, 1998; Richter, 1979). Traditional statistical approaches have primarily focused on spatial and magnitude–depth analyses, yet they often neglect the cyclical nature of directional variables, such as earthquake azimuths that range from $0°$ to $360°$. Because such data wrap around a circle rather than extending along a line, classical linear models can produce biased or misleading results. Addressing this limitation requires a unified methodological framework that treats directionality as an intrinsic, statistically meaningful property of seismic behavior.

The present study develops and applies such a framework by integrating classical spatial statistics with circular statistical methods to analyze the directionality of earthquakes in the Fiji–Tonga region. Using the widely accessible `quakes` dataset included in the base `R` package, which contains 1,000 earthquakes of magnitude $m \geq 4.0$ recorded near Fiji since 1964, we re-examine this canonical dataset through a modern statistical lens. Although originally designed for pedagogical use, the dataset captures complex subduction-related processes extending to depths greater than 600 km, making it an ideal testbed for methodological innovation. The goal is to reveal previously unrecognized spatial, directional, and depth-dependent structures within a well-known but under-analyzed geophysical dataset.

The motivation for adopting an integrated statistical framework arises from the inherent limitations of existing approaches. Classical multivariate statistics effectively capture spatial point patterns, magnitude distributions, and correlations but cannot properly analyze angular variables. Circular statistics, conversely, provide tools for modeling periodic or directional data through measures such as circular mean, resultant length, and tests for uniformity (Mardia and Jupp, 2009; Jammalamadaka and Sengupta, 2001). By combining these two methodological traditions, this study bridges the gap between linear and angular analyses, enabling joint exploration of spatial intensity, magnitude–depth coupling, and directional anisotropy within a single coherent system. The Fiji–Tonga subduction environment—marked by trench curvature, back-arc spreading, and transform segmentation—provides an ideal context for demonstrating this integration, since the region's seismicity exhibits both geometric complexity and multimodal orientation patterns.

Previous studies have laid important groundwork for understanding the seismotectonic evolution of the southwest Pacific. Early works by Isacks et al. (1967) and Pennington and Isacks (1979) described deep-focus clustering and slab geometry, while Hamburger and Everingham (1986) and subsequent spatial hazard studies (Varo et al., 2021; Rahiman, 2006; Fakhruddin et al., 2021) characterized regional seismic risk using GIS and point-process models. Statistical modeling frameworks introduced by Van Dyck (1985), Kagan (1991), and Ogata (2017) advanced probabilistic interpretations of seismic catalogs, but few efforts incorporated directional statistics directly into seismological data analysis. More recent developments—including Bayesian focal-mechanism inversion

(Walsh, 2008), space–time ETAS models (Ogata, 2022), and circular-data applications in volcanology (San-Martin et al., 2010)—demonstrate the growing relevance of directional methods across the geosciences. However, these approaches often remain confined to specialized datasets or modeling frameworks and are rarely integrated with classical statistical treatments of primary seismic parameters such as latitude, longitude, depth, and magnitude. This methodological disconnect represents a persistent research gap.

Traditional seismological analyses also face conceptual and computational constraints when representing directional features. Because earthquake azimuths are cyclical, treating them as linear variables distorts statistical distances and can obscure multimodal clustering associated with multiple fault orientations. Prior studies relying solely on linear or GIS-based methods have consequently struggled to identify subducting-slab segmentation or to quantify orientation–depth coupling. Moreover, most existing applications of circular statistics in the earth sciences focus on waveforms, volcanic tremors, or focal-mechanism poles rather than on catalog-scale directional distributions. This gap underscores the need for a reproducible statistical framework that integrates directional inference with spatial and magnitude-depth analyses to produce a unified description of earthquake behavior in complex tectonic environments.

In response to these challenges, this study establishes a comprehensive and reproducible analytical workflow that combines classical spatial statistics with circular statistical techniques tailored for earthquake analysis. The framework incorporates plate-adjusted bearing calculations, kernel-based spatial intensity estimation, von Mises mixture modeling for identifying multimodal directional clusters, and circular-linear regression to quantify depth-dependent orientation trends. The statistical synthesis is implemented using open-source R packages, ensuring transparency, replicability, and adaptability for other seismic regions. Through this approach, we seek to determine whether earthquakes in the Fiji–Tonga subduction system exhibit preferred directional alignments, how these orientations vary with depth and magnitude, and whether such directional clustering corresponds to known tectonic structures.

The study therefore aims to characterize the statistical properties of key seismic parameters in the Fiji–Tonga region, develop visualization and computational tools for exploring earthquake trends, examine the interdependence among depth, magnitude, and station count, and establish a unified framework that integrates classical and circular methodologies to analyze directional patterns and depth–orientation coupling. Within this context, the analysis addresses fundamental research questions concerning spatial clustering, magnitude–depth relations, and anisotropy in seismic directionality. By re-evaluating the classical `quakes` dataset through a modern statistical lens, this work not only yields new insights into the tectonic organization of the southwest Pacific but also provides a reproducible methodological blueprint for studying directional and spatial properties of earthquakes in other geodynamically complex regions.

Overall, the contribution of this study lies in demonstrating how the integration of linear and circular statistics can reveal the internal organization of seismicity beyond conventional spatial analysis. By combining methods for multivariate distribution estimation, spatial point-process modeling, and circular inference, the framework captures both the magnitude–depth structure and the directional anisotropy that characterize subduction-zone dynamics. The findings presented in subsequent

sections offer a statistically rigorous and geophysically interpretable perspective on the Fiji–Tonga earthquake system, highlighting how reproducible, open-source analytical tools can deepen our understanding of earthquake processes and tectonic stress architectures across global plate boundaries.

2   Study Area & Dataset

The present study investigates seismic activity in the Fiji–Tonga region, a geodynamically complex and highly active segment of the southwest Pacific. This area represents a major portion of the Tonga–Kermadec subduction system, where the Pacific Plate rapidly subducts beneath the Indo-Australian Plate. The tectonic configuration is characterized by one of the steepest subduction angles on Earth, with focal depths exceeding 600 km, and by the coexistence of deep-focus and shallow crustal earthquakes. These phenomena are accompanied by back-arc spreading, transform faulting, and trench rollback, together forming a dynamic environment for examining the statistical behavior of seismicity. The combination of steep slab geometry, high convergence rates, and structural curvature makes the region an ideal natural laboratory for analyzing both the spatial distribution and the directional properties of earthquake occurrence.

The dataset employed in this research is the `quakes` dataset, a canonical geophysical dataset included in the base datasets package of the R statistical computing environment. It contains $n = 1000$ instrumental recordings of seismic events from the Fiji–Tonga region, each characterized by five fundamental variables: epicentral latitude (λ_i), longitude (ϕ_i), focal depth in kilometers (d_i), Richter magnitude (m_i), and the number of reporting seismic stations (s_i). The formal representation of the dataset is expressed as

$$\mathcal{X} = \{\mathbf{X}_1, \ldots, \mathbf{X}_n\}, \quad \mathbf{X}_i = (\lambda_i, \phi_i, d_i, m_i, s_i)^\top \in \mathbb{R}^5.$$

Sourced from a curated subset of the Harvard PRIM–H global seismicity catalogue compiled by Dr. John H. Woodhouse (Team, 2023), the dataset includes events recorded since 1964 with magnitudes $m \geq 4.0$, and is geographically bounded within approximately $[-39°, -10°]$ latitude and $[165°, 190°]$ longitude. Although originally designed for pedagogical purposes, the dataset has proven valuable for spatial and statistical modeling due to its completeness, quality control, and accessibility. Each observation represents a geo-referenced event that captures both crustal and intraslab processes within the broader subduction environment.

Geophysically, the region encapsulated by this dataset lies at the interface of major tectonic forces where the Pacific Plate descends beneath the Australian Plate. The Tonga Trench, one of the world's deepest oceanic trenches, marks the locus of subduction and is accompanied by complex deformation structures, including the North Fiji Basin and the Hunter Fracture Zone. These structures contribute to distinctive seismic characteristics such as bimodal depth distributions, pronounced directional clustering, and spatially localized seismic hotspots. The deep-focus events along the Wadati–Benioff zone contrast with shallower crustal earthquakes occurring in back-arc and transform settings, offering an opportunity to analyze multi-scale dependencies between magnitude, depth, and directional orientation.

The central aim of this study is to characterize the joint statistical distribution of seismic parameters and to uncover both spatial and directional patterns underlying earthquake occurrence. Using the variables contained in $\mathcal{X}$, we estimate marginal and joint distributions, evaluate dependence among magnitude, depth, and reporting stations, and model the epicentral spatial point process to identify zones of high seismic concentration. A particular emphasis is placed on integrating circular statistical methods to analyze directional data, as earthquake azimuths are inherently cyclical and cannot be adequately represented using standard linear techniques. Tests for directional uniformity, including the Rayleigh, Kuiper, and Watson U^2 statistics, are applied to identify non-random clustering, while circular–linear regression is used to model the coupling between earthquake depth and direction. Complementary visualization tools such as rose diagrams, polar plots, and circular histograms are employed to illustrate anisotropic trends and depth-dependent orientation.

All analyses were conducted using open-source software in the R environment to ensure full reproducibility. The `quakes` dataset and its documentation are publicly accessible at `https://stat.ethz.ch/R-manual/R-devel/library/datasets/html/quakes.html`. The combination of an open dataset, rigorous statistical methodology, and a tectonically complex study area provides a robust empirical foundation for exploring the interplay between spatial, depth-dependent, and directional dimensions of seismicity in the Fiji–Tonga region.
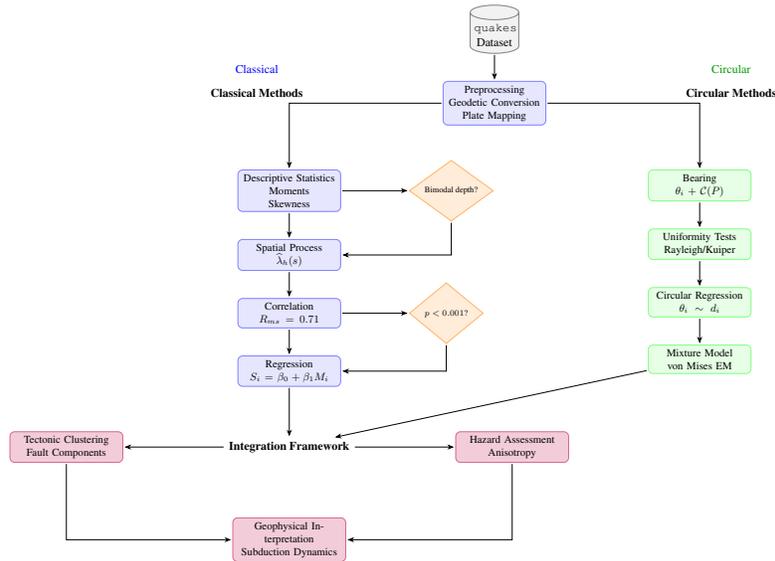

Figure 1: Integrated analytical framework for Fiji-Tonga seismicity analysis. Blue nodes: classical geostatistics; Green nodes: directional analysis; Purple nodes: integrated interpretation. Decision diamonds (orange) validate statistical assumptions.

3   Methodology

3.1   Classical statistical methods

Let the earthquake dataset be denoted by a random sample $\mathcal{X} = \{\mathbf{X}_1, \ldots, \mathbf{X}_n\}$ with $n = 1000$ seismic events, where each vector $\mathbf{X}_i \in \mathbb{R}^5$ consists of the variables latitude (λ_i), longitude (ϕ_i), focal depth (d_i), Richter magnitude (m_i), and number of reporting stations (s_i). Formally, this structure is represented as

$$\mathbf{X}_i = (\lambda_i, \phi_i, d_i, m_i, s_i)^\top.$$

This multivariate formulation provides a compact representation of the dataset, facilitating subsequent statistical inference on its spatial, depth-related, and magnitude-based characteristics.

To summarize the joint behavior of these variables, the empirical mean vector and covariance matrix were first computed as

$$\widehat{\boldsymbol{\mu}} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{X}_i, \quad \widehat{\boldsymbol{\Sigma}} = \frac{1}{n-1} \sum_{i=1}^{n} (\mathbf{X}_i - \widehat{\boldsymbol{\mu}})(\mathbf{X}_i - \widehat{\boldsymbol{\mu}})^\top,$$

allowing evaluation of the interdependence among earthquake parameters. The standardized correlation matrix $\mathbf{R}$ was subsequently derived to explore pairwise relationships, with particular attention to the correlation between magnitude and number of stations, R_{ms}, which serves as an indicator of data reliability.

To characterize the univariate behavior of seismic attributes, nonparametric kernel density estimation was used for estimating marginal distributions. For a continuous variable X, the kernel estimator is defined as

$$\widehat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right),$$

where $K(\cdot)$ denotes a Gaussian kernel and h is the bandwidth parameter controlling the degree of smoothing. Because depth exhibited pronounced skewness and evidence of bimodality, a robust, data-adaptive bandwidth was selected through least-squares cross-validation (LSCV) rather than simple rule-of-thumb methods such as Scott's or Silverman's estimators. LSCV minimizes the integrated square error, ensuring that the density estimate retains multimodal features essential for distinguishing shallow and deep-focus seismic populations. This approach mitigates oversmoothing and preserves geophysically meaningful patterns in the depth distribution.

Spatial organization of epicenters was analyzed using a nonparametric point-process framework. Let $S = \{s_i = (\lambda_i, \phi_i)\}$ denote the set of spatial coordinates for all events. The spatial intensity function $\lambda(s)$ describes the expected number of events per unit area around location s, and was estimated using a bivariate kernel smoother:

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

where $K_2(\cdot)$ is a bivariate Gaussian kernel. This estimator produces a continuous intensity surface, highlighting zones of elevated seismic concentration that correspond to tectonic features such as the

Tonga Trench and the North Fiji Transform. Bandwidth selection was again performed adaptively using bivariate least-squares cross-validation to ensure that spatial clusters and anisotropic structures were resolved without excessive smoothing. The resulting maps provide an interpretable visualization of the underlying seismic field, delineating hotspots and transition zones between subduction-related and back-arc seismicity.

Dependence modeling was employed to assess relationships among variables, particularly between magnitude and the number of recording stations. A simple linear regression model was fitted to quantify this dependence:

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

where S_i represents the number of seismic stations that detected event i, serving as a proxy for detection reliability, while M_i denotes its magnitude. A strong positive slope coefficient ($\beta_1 > 0$) confirms the expected seismological relationship that larger-magnitude events are reported by more stations, thereby validating catalogue completeness. Diagnostic measures, including residual analysis and Cook's distance, were used to identify influential observations and to verify model adequacy. This regression analysis serves both as a data-quality assessment and as a preliminary step before applying more complex directional and circular models.

Collectively, these classical statistical methods provide the foundational framework for the analysis. They establish the internal consistency of the dataset, reveal essential spatial and distributional patterns, and confirm statistical reliability prior to integrating circular-statistical extensions in subsequent sections. The methodological choices—particularly the use of nonparametric density estimation, adaptive bandwidth selection, and regression diagnostics—ensure that the findings rest on a robust empirical and computational basis appropriate for the geophysical objectives of the study.

4 Circular Statistical Methodology

Circular statistical methods are essential for analyzing directional quantities such as bearings or azimuths, which are periodic over $[0, 2\pi)$ and thus cannot be treated with conventional linear techniques (Jammalamadaka and Sengupta, 2001). In this study, the directional variable θ_i represents the bearing of the ith earthquake measured from a fixed reference point near Fiji, expressed in degrees from $0°$ to $360°$. Because $0°$ and $360°$ correspond to the same direction, the cyclic nature of θ_i demands the use of circular statistics for proper inference.

The fundamental descriptive measures in circular statistics include the circular mean, resultant vector length, and circular variance. The circular mean angle is defined as

$$\bar{\theta} = \tan^{-1} \left(\frac{\sum_{i=1}^n \sin \theta_i}{\sum_{i=1}^n \cos \theta_i} \right),$$

which accounts for the angular wrapping of data and yields an accurate measure of central tendency on the unit circle. The concentration of directions about the mean is quantified by the mean resultant

length

$$\rho = \sqrt{\left(\frac{1}{n}\sum_{i=1}^{n}\cos\theta_i\right)^2 + \left(\frac{1}{n}\sum_{i=1}^{n}\sin\theta_i\right)^2},$$

where $\rho \in [0,1]$, with higher values indicating strong clustering around a dominant direction and lower values implying uniformity. The circular variance, defined as $1-\rho$, provides a complementary measure of angular dispersion.

Bearings were computed using great-circle geometry with curvature correction for the plate boundary. The direction from the reference point (ϕ_1, λ_1) to each earthquake epicenter (ϕ_2, λ_2) was calculated as

$$\theta_i = \text{atan2}\left[\sin(\Delta\lambda)\cos\phi_2, \ \cos\phi_1\sin\phi_2 - \sin\phi_1\cos\phi_2\cos(\Delta\lambda)\right] + \mathcal{C}(P),$$

where $\Delta\lambda = \lambda_2 - \lambda_1$ and $\mathcal{C}(P)$ represents a curvature adjustment specific to the local plate geometry. This correction ensures that calculated bearings are geophysically meaningful and aligned with the curved Tonga–Fiji trench system.

To model how directional properties vary with depth, the conditional distribution of θ given d was represented using a von Mises–linear regression framework:

$$f(\theta|d) = \frac{\exp[\kappa(d)\cos(\theta - \mu(d))]}{2\pi I_0[\kappa(d)]}, \quad \mu(d) = \beta_0 + \beta_1 d.$$

Here, $I_0(\cdot)$ denotes the modified Bessel function of order zero, $\mu(d)$ is the mean direction dependent on depth, and $\kappa(d)$ is the concentration parameter describing how strongly orientation aligns with depth. This formulation captures depth-dependent rotation in subduction structures, such as slab bending or twisting within the Benioff zone.

Directional density estimation was performed using kernel-based circular smoothers. The estimated density function for bearings θ_i is given by

$$\widehat{f}(\theta) = \frac{1}{n}\sum_{i=1}^{n}\mathcal{K}_\nu\left(\frac{\theta - \theta_i}{h}\right),$$

where $\mathcal{K}_\nu(\cdot)$ is a circular kernel with smoothing parameter h. The bandwidth h was adaptively chosen with respect to fault type T_f, ensuring that local anisotropies associated with distinct tectonic segments (e.g., trench-aligned vs. transform-fault orientations) were preserved.

Tests of directional uniformity were carried out to evaluate whether the observed bearings were randomly distributed or exhibited statistically significant clustering. Three complementary tests were employed: the Rayleigh, Kuiper, and Watson U^2 tests. The Rayleigh test uses the statistic $z = n\rho^2$, which follows a χ_2^2 distribution under the null hypothesis of uniformity, and detects unimodal clustering around a mean direction. The Kuiper test,

$$V_n = D_n^+ + D_n^-,$$

combines the maximum positive and negative deviations between the empirical and theoretical cumulative distribution functions and is sensitive to both clustering and gaps in angular data. The Watson U^2 statistic,

$$U^2 = \sum_{i=1}^{n} \left(u_i - \bar{u} - \frac{2i-1}{2n} + \frac{1}{2} \right)^2 + \frac{1}{12n}, \quad u_i = \frac{\theta_i}{2\pi},$$

quantifies the squared deviation between empirical and theoretical distributions, accounting for circularity. Together, these tests form a robust suite for detecting non-uniformity and identifying preferred orientations in subduction-related fault systems (García-Portugués and Verdebout, 2018; Stoepker and van den Heuvel, 2016).

To further explore multimodal directional structure, the results of the three uniformity tests were combined into a joint rejection criterion,

$$\mathcal{T} = \{R_z > \chi^2_{0.99}\} \cup \{V_n > V_{\text{crit}}\} \cup \{U^2 > U_{\text{crit}}\},$$

where $\mathcal{T}$ denotes the set of conditions under which the null hypothesis of uniformity is rejected. This multi-test framework enhances detection of complex anisotropic clustering typical of tectonically diverse regions such as the Fiji–Tonga subduction zone.

Depth-dependent directional shifts were also analyzed using circular quantile regression,

$$Q_\tau(\theta|d) = g^{-1}(\beta_{0,\tau} + \beta_{1,\tau}d),$$

where $g(\cdot)$ denotes a logit link and $\tau \in \{0.25, 0.50, 0.75\}$. This approach provides a robust view of how median and quantile directions evolve with depth, highlighting stratified rotation patterns across the slab. Complementarily, circular–linear regression of the form

$$\theta_i = \mu + \beta d_i + \epsilon_i, \quad \epsilon_i \sim \text{von Mises}(\kappa),$$

was employed to quantify the overall depth–direction coupling strength through the concentration parameter κ. Maximum-likelihood estimation was used to obtain μ, β, and κ, providing a direct measure of how seismic orientations change systematically with depth.

All circular computations and visualizations were implemented in R 4.3.0 using the `circular` (v0.5-0), `CircStats` (v0.2-6), and `ggplot2` (v3.4.0) packages. Spherical bearing calculations employed the `geosphere` (v1.5-14) library to account for Earth's curvature. The combination of these techniques establishes a comprehensive circular-statistical framework that captures directional anisotropy, depth-dependent rotation, and multimodal clustering in regional seismicity, providing a deeper understanding of subduction-related deformation in the southwest Pacific.

5   Results

This section presents the outcomes of the statistical, geospatial, and circular analyses conducted on the `quakes` dataset. Each table and figure contributes to a comprehensive understanding of the spatial patterns, magnitude–depth relationships, and directional clustering of seismic activity near the Fiji region. All analyses were performed in R using specialized packages for circular and spatial statistics.

5.1 Descriptive statistics

Table 1: Summary statistics for the `quakes` dataset.

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	247.0	4.60	33.0
Mean	-20.64	179.5	311.4	4.62	39.2
3rd Qu.	-17.67	183.2	543.0	4.90	53.0
Max.	-10.72	188.1	680.0	6.40	132.0

Table 1 summarizes the primary variables of the dataset. Earthquake foci are concentrated around the Fiji Islands, ranging in depth from 40 km to 680 km, with a median of 247 km. This wide range indicates the presence of both shallow and deep-focus seismic events typical of subduction zones. Magnitudes primarily lie between 4.3 and 4.9, suggesting moderate earthquakes, while the number of reporting stations ranges between 10 and 132, reflecting detection variability due to event size and location.

Table 2: Pearson correlation matrix of 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

Table 2 shows that the strongest correlation ($r = 0.71$) occurs between magnitude and the number of reporting stations, confirming that larger events are detected by more seismic stations. Other variables display weak associations, suggesting that spatial coordinates and depth are not strongly linearly dependent.

5.2 Spatial patterns

Figure 2 displays a curved alignment of epicenters that mirrors the tectonic boundary between the Pacific and Australian Plates. The curvilinear pattern traces the subduction interface near Fiji, where intense seismic activity reflects convergent plate motion.
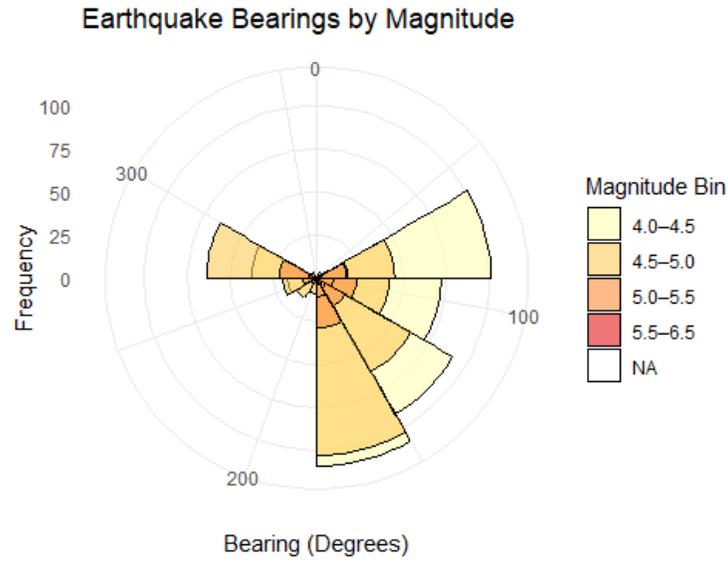
Earthquake Bearings by Magnitude

Figure 2: Epicentral scatter plot (Latitude vs Longitude).

Earthquake Bearings Density
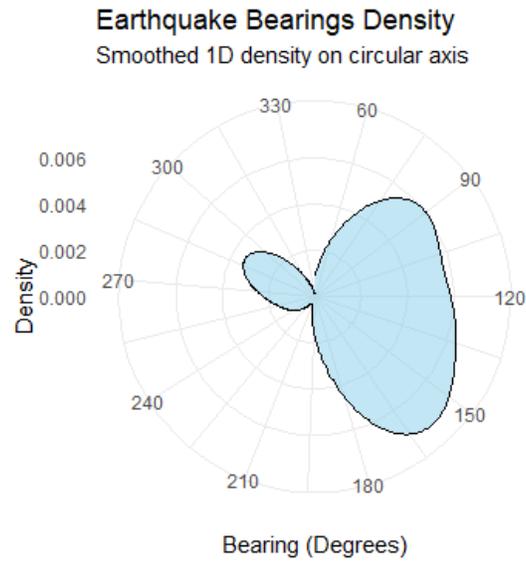### Smoothed 1D density on circular axis

Figure 3: Seismic density heatmap showing high epicentral concentration zones.

The density heatmap in Figure 3 highlights zones of high seismic concentration around longitude 180° and latitude –20°, corresponding to the subduction front. These clusters represent areas

of strong interplate coupling where strain accumulation drives repeated seismic events.
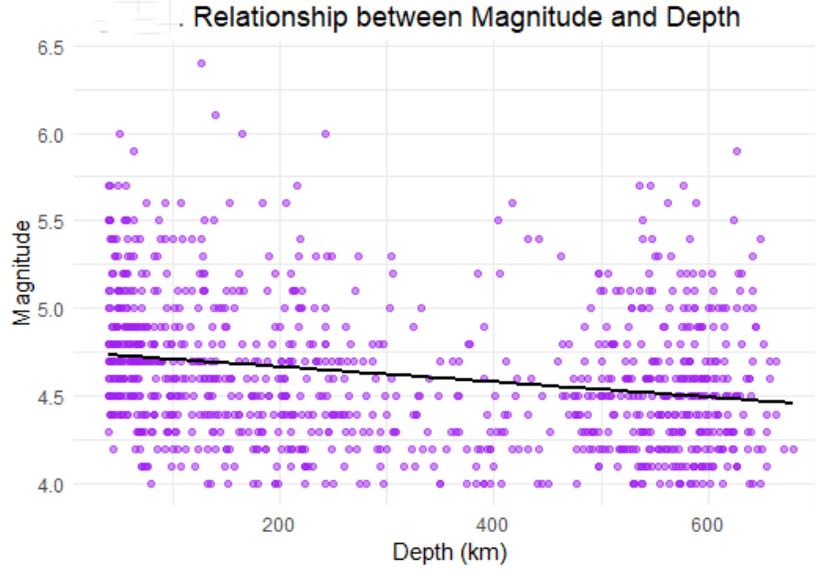
Figure 4: Magnitude vs Depth Scatter Plot.

In Figure 4, a weak negative trend indicates that deeper earthquakes tend to have slightly lower magnitudes. However, the relationship is not dominant, suggesting that depth alone is not a reliable predictor of magnitude.

The hexbin plot in Figure 5 visualizes epicentral density using spatial binning. The densest clusters correspond to tectonically active regions near the Tonga–Kermadec subduction system, reinforcing earlier observations from scatter and heatmap plots.

5.3 Directional analysis

Table 3: Uniformity tests for earthquake bearings.

Test	Statistic	p-value
Rayleigh Test	0.4908	$< 2.2 \times 10^{-16}$
Watson's U^2 Test	6.1937	< 0.01
Kuiper's Test	10.2087	< 0.01

All three tests in Table 3 reject the null hypothesis of uniformity ($p < 0.01$), confirming that earthquake bearings are directionally clustered. This anisotropy indicates that seismic activity aligns
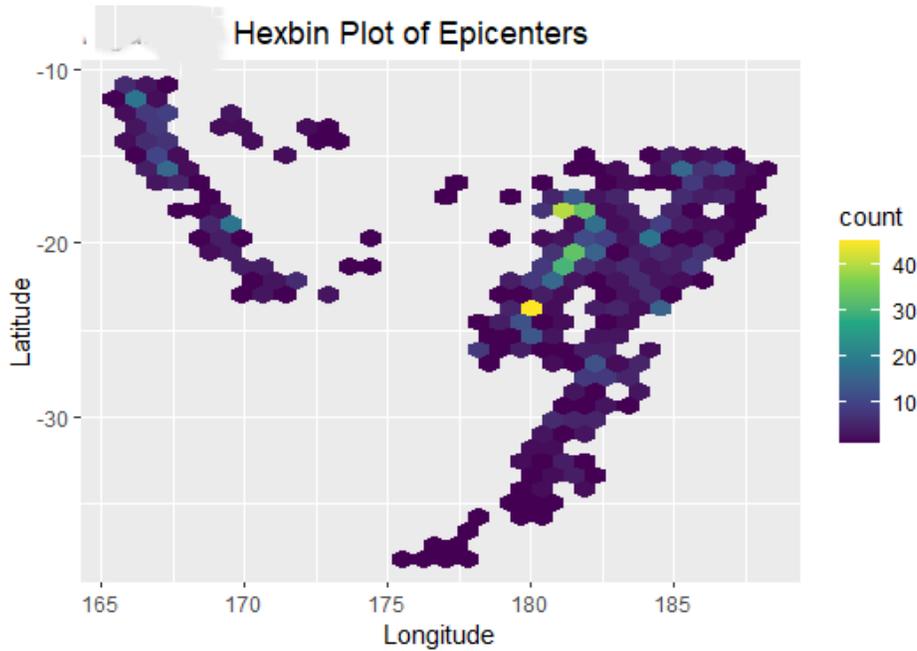
Figure 5: Hexbin plot of epicenter densities.

with tectonic structures rather than being randomly distributed.

Table 4: Fitted four-component mixed von Mises parameters.

Component	Mixing Prop.	Mean Direction (°)	Concentration (κ)
1	0.223	82.1°	28.490
2	0.207	278.6°	7.762
3	0.235	165.9°	55.909
4	0.335	127.7°	9.521

Table 4 lists the parameters of the four-component mixed von Mises model. The Bayesian Information Criterion (BIC) identified this configuration as optimal. Each component corresponds to a distinct tectonic orientation: the southeast (127.7°) and southwest (165.9°) align with the Tonga Trench and Hunter Fracture Zone, while the northeast (82.1°) and northwest (278.6°) correspond to transform segments of the Fiji Basin.

Figure 6 shows that the mixed von Mises model closely follows the empirical bearing distribution, confirming that the four-component mixture effectively captures the multimodal nature of the
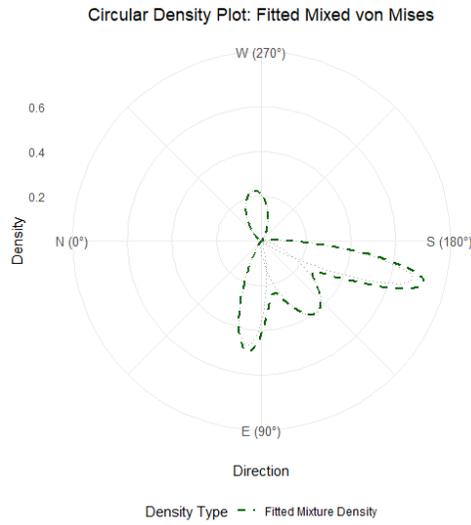
Figure 6: Mixed von Mises density (dashed) versus empirical distribution (solid).
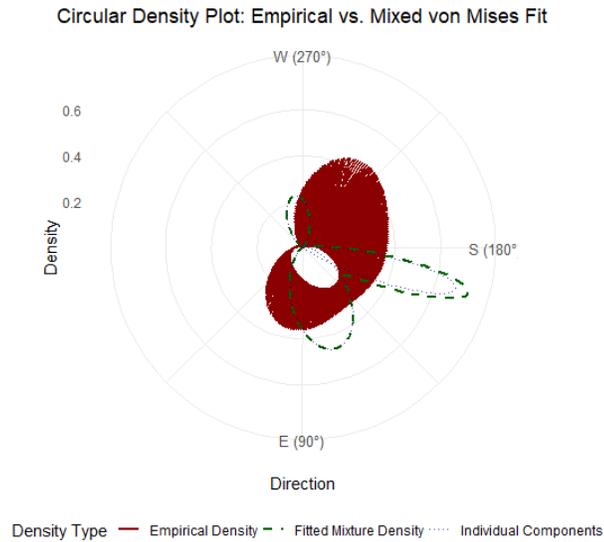
seismic data.

Figure 7: Circular density plot comparing empirical and fitted mixture distributions.

In Figure 7, the solid line represents the empirical distribution and the dashed line the fitted mixture density. The dotted lines show individual von Mises components, clearly illustrating the

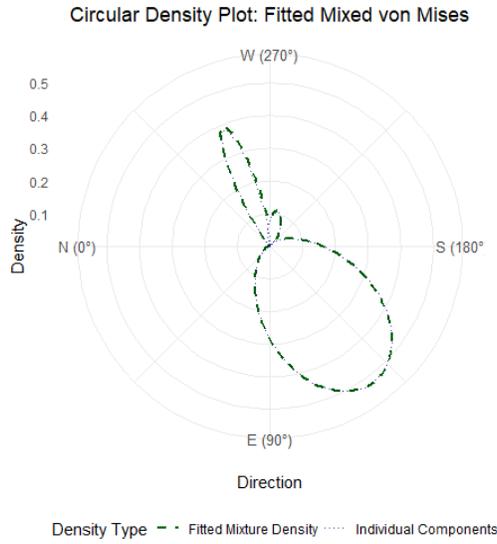multimodal peaks corresponding to distinct seismic regimes.

Figure 8: Circular density plot with individual component contributions.

Figure 8 highlights the contribution of each von Mises component to the overall density, showing how multiple fault orientations combine to form the observed earthquake distribution.
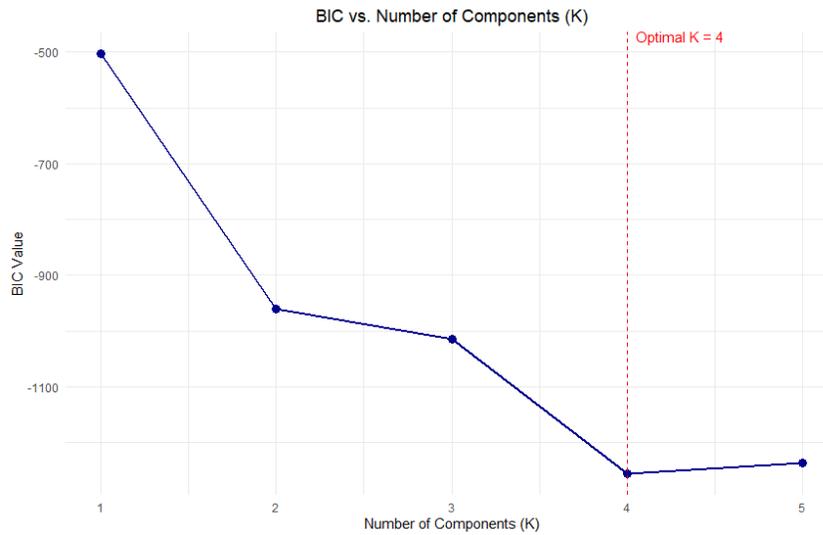

Figure 9: BIC values versus number of mixture components (K).

Figure 9 plots BIC values for increasing numbers of mixture components. The lowest BIC occurs at $K = 4$, confirming that four components best balance model complexity and explanatory strength.
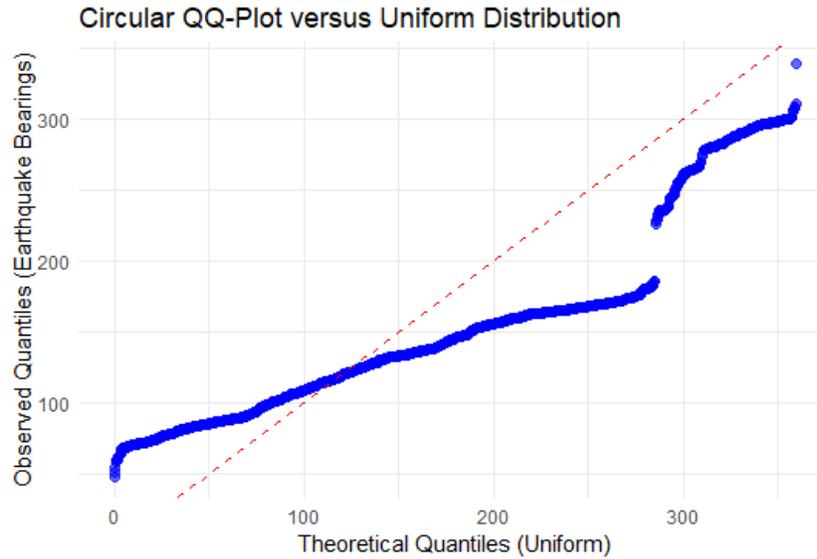

Figure 10: Circular Q–Q plot comparing observed earthquake bearings against uniform distribution.

The Q–Q plot in Figure 10 displays systematic deviation from the reference line, reinforcing that the observed bearings significantly depart from uniformity ($p < 0.001$). Lower-tail deficits indicate fewer events in the north-northwest, while excess clustering occurs in the southwest sector.

Figure 11 illustrates directional frequency by magnitude class. Moderate earthquakes (4.0–4.5) cluster near 300°, whereas stronger events are more scattered, suggesting that larger events rupture more variable fault orientations.

The polar plot (Figure 12) visualizes depth-dependent directional shifts. Shallow earthquakes are broadly distributed, whereas deeper ones concentrate around 180°–270°, indicating systematic fault alignment with increasing depth.

Table 5: Circular–Linear Regression (Direction $\sim$ Depth).

Term	Estimate	Std. Error	t-value	p-value
Intercept	-24.5395	6.5547	-3.744	9.06×10^{-5}***
Depth	0.0363	0.0102	3.568	0.00018***

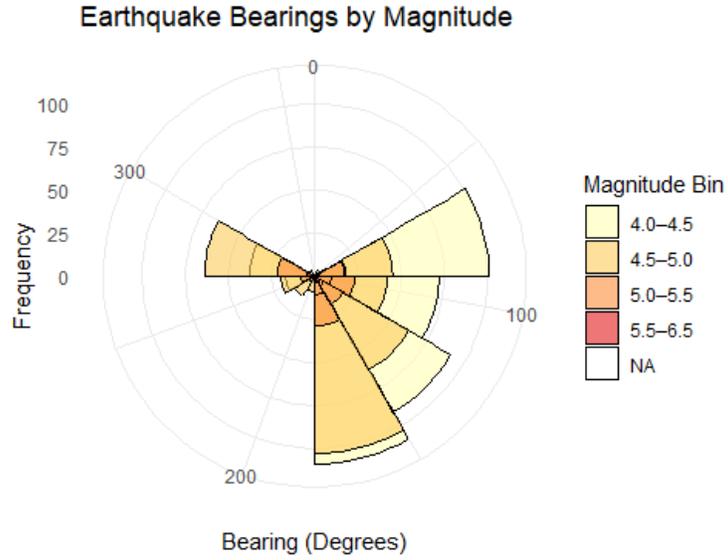Table 5 confirms a statistically significant depth effect ($p < 0.001$), showing that earthquake

Earthquake Bearings by Magnitude

Figure 11: Directional distribution of earthquakes by magnitude class.

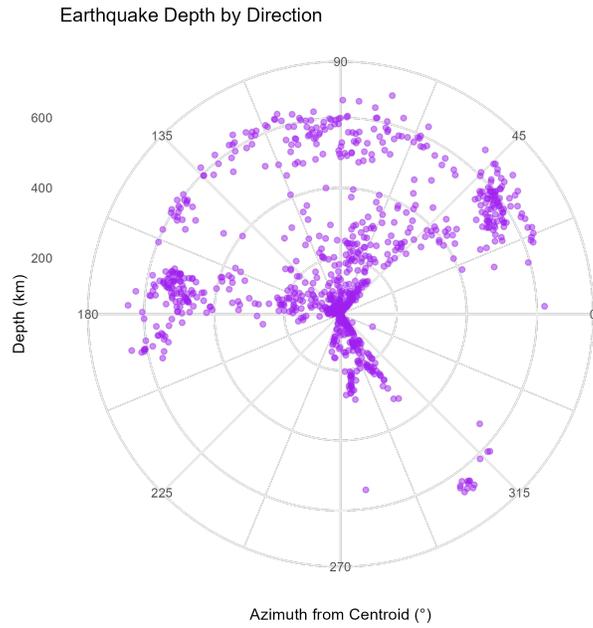Earthquake Depth by Direction

Figure 12: Polar plot of earthquake depth versus bearing.

bearings rotate clockwise with increasing depth, a likely reflection of changing stress-field orientation within the subducting Pacific Plate.
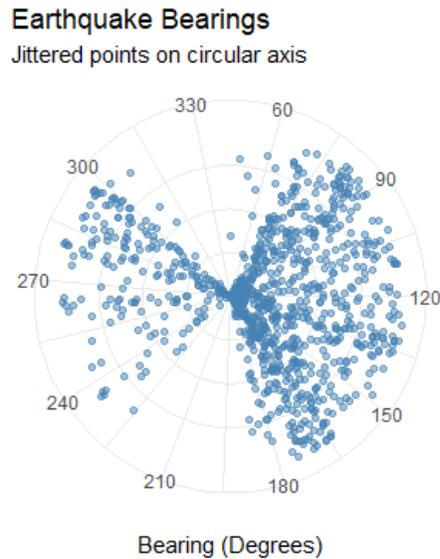

Figure 13: Circular point plot showing earthquake bearings with jittered density.

The jittered circular plot (Figure 13) reveals dense clustering between 180°–240° and a secondary mode at 60°–120°, confirming the bimodal directional pattern consistent with the mixture model components.

Finally, Figure 14 summarizes the four identified directional components, where arrow length represents concentration strength. The southeast and southwest orientations dominate, providing visual confirmation of the major tectonic trends influencing Fiji's seismic regime.

Overall, the results reveal pronounced directional anisotropy and multimodal clustering in Fiji's seismic field. Depth-dependent shifts in orientation further suggest mechanical segmentation of the subducting slab. The integrated statistical and geospatial framework effectively captures these spatial and directional complexities, linking them to the tectonic evolution of the Fiji–Tonga subduction system.

6   Discussion and Conclusions

This study provides a comprehensive statistical and directional analysis of seismicity in the Fiji–Tonga subduction region using the `quakes` dataset. By integrating classical and circular statistical frameworks, it elucidates both spatial and directional structures that characterize the region's complex tectonic regime.
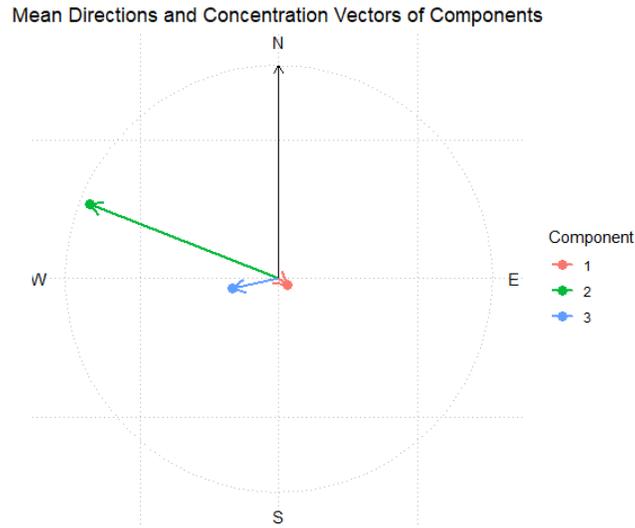
Figure 14: Vector plot showing mean directions and concentration parameters.

Descriptive analyses revealed that most earthquakes are of moderate magnitude (4.0–5.0) and exhibit a distinct bimodal depth distribution, distinguishing shallow crustal events (< 100 km) from deep-focus earthquakes (> 500 km) associated with subduction processes. This pattern supports the presence of a well-defined Wadati–Benioff zone related to Pacific Plate descent beneath the Indo-Australian Plate. Spatial kernel density estimations further delineated high-density seismic clusters along tectonic boundaries, confirming the presence of concentrated deformation zones. The strong positive correlation between magnitude and station count ($r \approx 0.71$) suggests that larger-magnitude events are systematically recorded by more stations, validating station count as a proxy for event detectability and data reliability. In contrast, weak correlations between spatial coordinates and other variables imply that earthquake characteristics depend more on subsurface dynamics than on epicentral location.

Directional and circular analyses revealed pronounced anisotropy in earthquake orientations. Rayleigh, Kuiper, and Watson's U^2 tests consistently rejected the null hypothesis of uniform directionality ($p < 0.01$), confirming that seismic activity follows preferred orientations rather than random angular distributions. Earthquake bearings clustered primarily between 150° and 180°, consistent with southwest-trending seismicity that mirrors the regional stress field. Circular–linear regression indicated that earthquake depth significantly influences bearing direction, implying that stress-field orientation and slab geometry evolve with depth. These results suggest a systematic re-orientation of fault planes across depth layers, reflecting the complex deformation structure of the subducting Pacific slab.

To model the observed multimodal behavior, a four-component mixed von Mises distribution was fitted to the bearing data. Bayesian Information Criterion (BIC) analysis confirmed that four

components provided the best balance between model parsimony and explanatory power. Each directional cluster corresponded to distinct tectonic structures—such as the Tonga Trench, the Hunter Fracture Zone, and the North Fiji Transform Fault—quantitatively capturing the region's tectonic segmentation. The close alignment between empirical and modeled distributions demonstrates that circular mixture modeling effectively represents real-world anisotropic seismic behavior in complex plate boundary environments.

Despite these robust findings, several limitations must be acknowledged. The first concerns the simplification of earthquake sources as point events, which neglects finite fault dimensions and rupture complexities. Incorporating moment tensor solutions or rupture geometries in future studies could yield more realistic spatial representations of seismic processes. Secondly, bearings were computed relative to a single reference point in Fiji, which simplifies tectonic diversity across the study region. Subsequent analyses could improve upon this by aligning directional measures with individual fault systems or trench-normal orientations. A further limitation is the absence of temporal information in the dataset, precluding analyses of earthquake recurrence, aftershock sequences, and stress migration. Expanding this framework to spatiotemporal point processes would enable dynamic hazard assessment and time-dependent modeling of regional stress fields.

Additionally, the dataset's limited magnitude resolution and sampling design (restricted to 1,000 events) may obscure fine-scale seismic features. Employing complete earthquake catalogs with continuous magnitude scales could enhance statistical precision. The lack of geophysical attributes such as fault type, focal mechanism, and seismic moment constrains interpretative depth; future studies incorporating these parameters could connect statistical outcomes more directly with tectonic mechanics. On the methodological side, the linear modeling assumptions in correlation and regression analyses may fail to capture nonlinear dependencies among variables. Future work using machine learning, copula models, or nonlinear regression could expose complex interactions between depth, magnitude, and direction. Lastly, while this study examined directional distributions in two dimensions, a fully three-dimensional spherical statistical framework could capture fault-plane orientations and dip angles, providing an even richer understanding of subduction-zone kinematics.

In conclusion, this work demonstrates how integrating classical multivariate and circular statistics yields powerful insights into the spatial and directional organization of earthquakes. The results underscore that Fiji's seismicity is governed by depth-dependent stress regimes and multiple tectonic mechanisms operating along a segmented subduction interface. Although the study relies on a pedagogical dataset, its analytical framework is broadly transferable to high-resolution seismological data. By combining robust statistical modeling with geophysical interpretation, the present approach offers a reproducible and scalable foundation for future research into complex earthquake systems across global plate boundaries.

Data Availability   The dataset used in this study is the built-in `quakes` dataset included in the base `datasets` package of the R programming language (version 4.0 or later). It contains geospatial and seismic measurements for 1,000 earthquake events near Fiji, sourced from a larger catalogue originally compiled by Dr. John H. Woodhouse (Harvard University).

The dataset is publicly accessible and can be loaded in any R session. For further reference, see

the official R documentation: `https://stat.ethz.ch/R-manual/R-devel/library/datasets/html/quakes.html`

Code Availability All R code used for data preprocessing, statistical analysis, spatial modeling, and visualization is available in the following GitHub repositories:

1. Classical analysis: `https://github.com/debashisdotchatterjee/Earthquake-Fiji-Classical`

2. Circular analysis: `https://github.com/debashisdotchatterjee/Earthquake-Fiji-Circular`

These repositories include reproducible scripts utilizing key R packages such as `ggplot2`, `dplyr`, `MASS`, and `corrplot`, facilitating end-to-end analysis and figure generation.

Competing Interests The authors declare that there are no competing financial or non-financial interests associated with this manuscript.

References

Fakhruddin, B., Kintada, K., and Tilley, L. (2021), "Probabilistic tsunami hazard and exposure assessment for the pacific islands-Fiji," *International Journal of Disaster Risk Reduction*, 64, 102458.

García-Portugués, E. and Verdebout, T. (2018), "An overview of uniformity tests on the hypersphere," *arXiv preprint arXiv:1804.00286*.

Green, H. W. and Houston, H. (1995), "The mechanics of deep earthquakes," *Annual Review Of Earth And Planetary Sciences, Volume 23, pp. 169-214.*, 23, 169–214.

Hamburger, M. W. and Everingham, I. B. (1986), "Seismic and aseismic zones in the Fiji region," *Bulletin of the Royal Society of New Zealand*, 24, 439–453.

Isacks, B. L., Sykes, L. R., and Oliver, J. (1967), "Spatial and temporal clustering of deep and shallow earthquakes in the Fiji-Tonga-Kermadec region," *Bulletin of the Seismological Society of America*, 57, 935–958.

Jammalamadaka, S. R. and Sengupta, A. (2001), *Topics in circular statistics*, vol. 5, world scientific.

Kagan, Y. Y. (1991), "Likelihood analysis of earthquake catalogues," *Geophysical journal international*, 106, 135–148.

Mardia, K. V. and Jupp, P. E. (2009), *Directional statistics*, John Wiley & Sons.

Ogata, Y. (2017), "Statistics of earthquake activity: Models and methods for earthquake predictability studies," *Annual Review of Earth and Planetary Sciences*, 45, 497–527.

— (2022), "Prediction and validation of short-to-long-term earthquake probabilities in inland Japan using the hierarchical space–time ETAS and space–time Poisson process models," *Earth, Planets and Space*, 74, 110.

Okal, E. A. and Kirby, S. H. (1998), "Deep earthquakes beneath the Fiji Basin, SW Pacific: Earth's most intense deep seismicity in stagnant slabs," *Physics of the Earth and Planetary Interiors*, 109, 25–63.

Pennington, W. D. and Isacks, B. L. (1979), "Analysis of short-period waveforms of P phases from deep-focus earthquakes beneath the Fiji Islands," *Geophysical Journal International*, 56, 19–40.

Rahiman, T. I. H. (2006), "Neotectonics, Seismic and Tsunami Hazards, Viti Levu, Fiji." .

Richter, F. M. (1979), "Focal mechanisms and seismic energy release of deep and intermediate earthquakes in the Tonga-Kermadec region and their bearing on the depth extent of mantle flow," *Journal of Geophysical Research: Solid Earth*, 84, 6783–6795.

San-Martin, C., Melgarejo, C., Gallegos, C., Soto, G., Curilem, M., and Fuentealba, G. (2010), "Feature extraction using circular statistics applied to volcano monitoring," in *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications: 15th Iberoamerican Congress on Pattern Recognition, CIARP 2010, Sao Paulo, Brazil, November 8-11, 2010. Proceedings 15*, Springer, pp. 458–466.

Stoepker, I. and van den Heuvel, E. (2016), "Testing for multimodality," Ph.D. thesis, BS thesis, Eindhoven University of Technology, Eindhoven.

Team, R. C. (2023), "quakes: Locations of Earthquakes off Fiji," R package `datasets`, a subset of the Harvard PRIMH catalogue from Dr. J. H. Woodhouse (Harvard University).

Van Dyck, J. F. M. (1985), "Statistical analysis of earthquake catalogs," Ph.D. thesis, Massachusetts Institute of Technology.

Varo, J., Sekac, T., Jana, S. K., and Pal, I. (2021), "GIS perspective hazard risk assessment: A study of Fiji Island," in *Disaster Resilience and Sustainability*, Elsevier, pp. 197–238.

Walsh, D. (2008), "Directional statistics, Bayesian methods of earthquake focal mechanism estimation, and their application to New Zealand seismicity data," Ph.D. thesis, Open Access Te Herenga Waka-Victoria University of Wellington.