Phi: Jurnal Pendidikan Fisika dan Terapan
Vol.12 No.2 (2026) 77-90
ISSN: 2460-4348, e-ISSN: 2549-7162
https://jurnal.ar-raniry.ac.id/index.php/jurnalphi/index

Article Type: *original research*

Beyond Site Effects: Azimuthal Dependence of Ground Motion Residuals in the Volcanic Arc and Sedimentary Basins of West Java, Indonesia

Reno Sudibyo[1*], Wijayanto Wijayanto[1], Sandy Tri Gustono[1]
[1]Indonesian Agency for Meteorological, Climatological and Geophysics, Jakarta, Indonesia

Corresponden E-mail: reno.sudibyo@bmkg.go.id

ARTICLE INFO

Article History:
Received: 12 January 2026
Revised: 21 April 2026
Accepted: 23 April 2026
Published: 29 April 2026

Keywords:
Ground motion prediction equation (GMPE); Azimuthal dependence; Path effects; West Java; Volcanic attenuation; Basin effects; Earthquake early warning

ABSTRACT

The complex tectonic setting of West Java, characterized by the interaction between the Sunda volcanic arc and deep sedimentary basins, presents significant challenges for seismic hazard assessment. This study recalibrates a local Ground Motion Prediction Equation (GMPE) using a hybrid regression strategy to isolate non-ergodic path effects. We implemented a constrained regression for Site Class C fixing the saturation parameter d = 5.80 to match the Class D baseline and applied site-term corrections for Class B. Our results reveal a weak correlation between residuals and V_{S30} (R = 0.215), suggesting that shallow soil velocity is a suboptimal proxy for ground motion variability in this region. Instead, a distinct azimuthal anisotropy emerged, characterized by a dipolar residual pattern. Significant high-frequency attenuation (negative residuals) was identified in azimuths crossing the Quaternary Volcanic Arc, likely driven by multiple scattering in fractured magmatic media. Conversely, systematic amplification (positive residuals) was observed towards the Jakarta and Bandung Basins, consistent with deep basin resonance. These findings underscore the necessity of integrating azimuthal path terms into regional hazard models. Moving beyond 1D site-effect proxies is critical for enhancing the accuracy of Earthquake Early Warning (EEW) systems and urban seismic mitigation in West Java's densely populated basins.

1. INTRODUCTION

The geological configuration of the Indonesian archipelago manifests as a complex assemblage of tectonic interactions, primarily driven by the oblique subduction of the Indo-Australian Plate beneath the Eurasian Plate, a process that has shaped the diverse seismogenic features of the Sunda Arc (Gunawan & Widiyantoro, 2019). This dynamic setting is particularly pronounced in West Java, where recent geodetic investigations using Global Positioning System (GPS) observations have delineated significant strain partitioning, revealing active crustal deformation along major fault systems such as the Lembang, Cimandiri, and Baribis-Kendeng faults (Koulali et al., 2017). Given the high density of population and critical infrastructure in this region, the accurate assessment of seismic hazard is imperative; consequently, the 2017 National Seismic Hazard Maps of Indonesia have integrated these

active fault parameters to provide a robust framework for disaster mitigation (Irsyam et al., 2020). However, while source characterization has improved significantly, the prediction of ground shaking intensity remains a challenging frontier due to the region's extreme crustal heterogeneity.

In engineering seismology, Ground Motion Prediction Equations (GMPEs) serve as the fundamental tool for estimating intensity measures, such as Peak Ground Acceleration (PGA), for a given rupture scenario. Ideally, seismic hazard analyses would utilize locally derived equations; however, in the absence of abundant local data, global models like the Next Generation Attenuation (NGA) relations for example, (Boore & Atkinson, 2008) are frequently adopted. Yet, the transferability of these global empirical models to specific tectonic environments is often limited; comparative studies in other complex regions, such as Italy, have demonstrated that region-specific GMPEs significantly outperform global models by reducing the aleatory variability associated with ergodic assumptions (Bindi et al., 2011). Recognizing this limitation in the context of volcanic island arcs, this study adopts the attenuation framework proposed by (Kanno et al., 2006) as a baseline, which explicitly accounts for the anomalous intensity distributions often observed in fore-arc and back-arc regions similar to those in Japan.

A critical limitation in current ground motion modeling lies in the characterization of site effects, which predominantly relies on the time-averaged shear-wave velocity in the top 30 meters (V_{S30}). While this is a standard proxy in global practice, its efficacy in active tectonic regions with complex subsurface structures, such as the Himalayas, has been questioned, as it may not fully capture the deeper resonance effects (Vyas et al., 2016). Furthermore, site classification schemes based solely on often fail to account for the significant variability in spectral amplification caused by impedance contrasts at depth (Zhao & Xu, 2013). This deficiency is particularly evident in West Java; a recent machine learning-based analysis of strong motion data in the region revealed that was, counter-intuitively, the least influential feature for predicting PGA compared to source magnitude and distance (Rachmadan et al., 2025). This finding strongly suggests that the residual variability in ground motion prediction for this region is driven by factors beyond local site conditions, necessitating a shift in focus toward path-dependent attenuation mechanisms.

The hypothesis of path-dependency is substantiated by the distinct geological contrast observed in West Java, which features massive Quaternary volcanic complexes juxtapositioned with deep sedimentary basins. Recent segmentation studies of the Sunda Arc volcanoes have highlighted how the structural discontinuities and thermal anomalies associated with volcanic chains—such as Mount Sinabung and Tangkuban Parahu, can significantly alter wave propagation properties (Rahmanto et al., 2025). Seismic tomography has further illuminated this complexity, revealing low-velocity zones beneath volcanic arcs that likely act as scattering media, thereby attenuating seismic energy (Sari & Fakhrurrozi, 2020). Conversely, deep sedimentary structures like the Jakarta Basin present a different challenge; studies on basin resonance have demonstrated that the basin’s geometry and depth can trap long-period waves, leading to significant amplification of ground motion (Cipta et al., 2018). The geometry of the bedrock interface, which plays a crucial role in these amplification effects, has been detailed through seismic noise tomography, confirming the basin's capacity to modify incoming wavefields (Saygin et al., 2016).

In addition to seismic-based approaches, recent advancements in multi-parameter Earthquake Early Warning (EEW) systems have highlighted the potential of non-seismic precursors. For instance, ionospheric anomalies detected through Total Electron Content (TEC) monitoring have been investigated as potential indicators for seismic events, offering a complementary perspective to traditional ground-motion-based warnings (Sarkar et al., 2026; Shaw et al., 2025). Integrating such diverse observation methods could significantly enhance the resilience and lead time of regional warning frameworks. Recent studies suggest that combining these stress-indicators with time-magnitude dependent models offers a more robust framework for long-term forecasting (Mukherjee et al., 2025). Furthermore, the development of location-dependent earthquake prediction algorithms has paved the way for more spatially precise hazard assessments (Shaw et al., 2026). Incorporating these multi-parameter approaches into traditional Ground Motion Prediction Equations (GMPE) is essential

for developing a truly resilient, next-generation Earthquake Early Warning (EEW) system in complex regions like West Java.

Despite the wealth of global and regional GMPE studies, a significant research gap remains in quantifying the non-ergodic path effects within West Java's unique geological transect. Most existing models adopt isotropic attenuation assumptions, which overlook the systematic variations in ground shaking intensity driven by the Quaternary volcanic arc and deep sedimentary basins. The novelty of this research lies in its specific focus on the azimuthal dependence of GMPE residuals to effectively isolate path effects from source and site terms. Enhancing the precision of these predictions through such high-resolution mapping is directly relevant to the operational needs of Earthquake Early Warning (EEW) systems, where rapid and accurate intensity estimates are crucial for minimizing false alerts and maximizing lead time (Allen & Wald, 2009). The integration of these novel, path-specific attenuation factors is expected to improve the reliability of the pilot EEW system currently being tested in Western Java (Rudyanto et al., 2024), ultimately contributing to more resilient disaster risk reduction strategies for Indonesia.

2. METHODS

2.1. Data Acquisition and Curation

The primary dataset for this study consists of strong-motion waveforms recorded by the accelerograph network of the Agency for Meteorology, Climatology, and Geophysics (BMKG). To ensure the analysis captures the specific attenuation characteristics of the West Java region, we applied a strict geographical filter, selecting only events and stations bounded within latitudes $8.0°S\ to\ 5.5°S$ and longitudes $105.0°E\ to\ 109.0°E$. Furthermore, the events in this study were strictly selected based on hypocenters relocated by the International Seismological Centre (ISC) and the BMKG's internal relocation catalog to ensure maximum precision in source-to-site distance (R_{hypo}) calculations.

The dataset was further refined to include only shallow crustal earthquakes with focal depths $h \leq 30\ km$, thereby isolating the upper crustal attenuation regime from deeper intraslab effects. The data is collected using accelerographs and intensitymeters permanently installed within standardized BMKG seismograph shelters, ensuring consistent instrument response and minimal environmental noise compared to temporary deployments.

Raw waveforms, accelorograms and metadata were retrieved directly from the BMKG's internal Earthquake Early Warning System (EEWS) repository via VSAT connectivity. Data validation involved cross-referencing raw.txt outputs and waveforms with official seismic reports to ensure record integrity. The processing followed standard procedures described by (Rudyanto et al., 2024), including baseline correction and bandpass filtering (0.1–20 Hz). The final dataset comprises 496 records from events with magnitudes M_w 4.34 to 5.6 and hypocentral distances of 5.5 to 427 km (Figure 1). To maintain data independence and mitigate path-effect bias, we focused on relocated mainshocks, cross-referenced with active fault zones identified by (Irsyam et al., 2020) and deformation vectors by (Koulali et al., 2017). In terms of site classification, the dataset is dominated by Site Class D (Medium Soil, N=289) and Site Class C (Stiff Soil, N=206). Data for Site Class B (Rock) was extremely limited (N=1) and thus excluded from the regression analysis, serving only as an indicative reference.

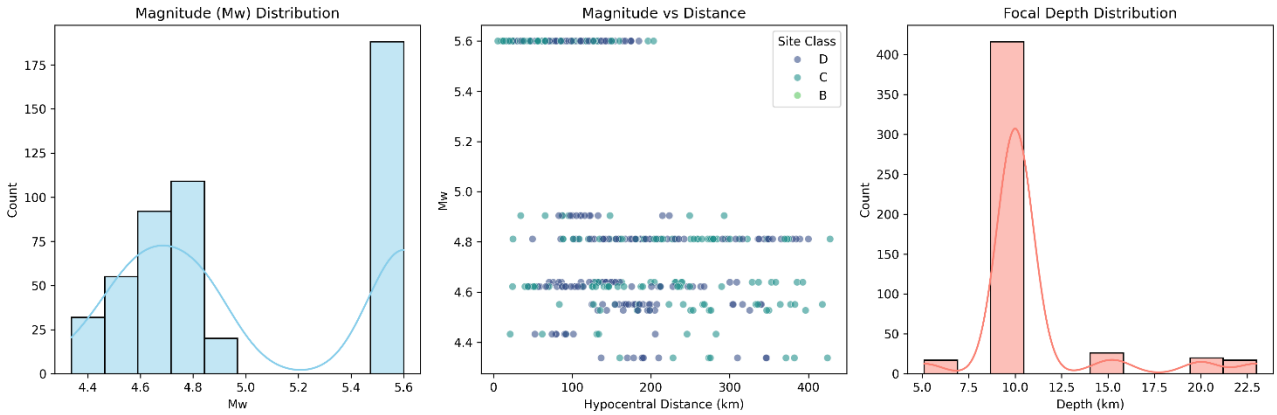

Figure 1. Data Distribution; Statistical distribution of the strong-motion dataset used in this study (N=496). (a) Histogram of Moment Magnitude (M_w), showing a concentration of events between M_w 4.5 and 5.5. (b) Scatter plot of Magnitude versus Hypocentral Distance (R_{hypo}), color-coded by Site Class (C and D), demonstrating adequate sampling across the distance range relevant for hazard assessment. (c) Distribution of Focal Depths, strictly limited to shallow crustal events ($h \leq 30\ km$) to exclude intraslab subduction earthquakes.

2.2. Site Characterization and Stratification

Unlike standard approaches that treat site conditions as a continuous variable via V_{S30} (time-averaged shear-wave velocity in the top 30 m), this study stratifies the dataset into discrete site classes based on the NEHRP/SNI 1726:2019 standard. This stratification allows for the detection of non-linear site responses that a simple linear V_{S30} correction might miss. The dataset was divided into:

- **Site Class B (Rock):** $750 < V_{S30} < 1500$ m/s
- **Site Class C (Stiff Soil):** $350 < V_{S30} \leq 750$ m/s
- **Site Class D (Soft/Medium Soil):** $175 < V_{S30} \leq 350$ m/s

Preliminary analysis indicated that Site Class D is the dominant soil type in the study area. Consequently, the attenuation model derived for Site Class D serves as the **baseline reference** for this study.

2.3. The Hybrid Predictive Modeling Framework

To determine the path-dependent residuals, we first established a local Ground Motion Prediction Equation (GMPE) based on the functional form of (Kanno et al., 2006). The predictive equation is defined as:

$$log_{10}\, PGA = a\, M_w - b\, x - log_{10}(x + d \cdot 10^{0.5\, M_w}) + c + \epsilon \quad (1)$$

Where M_w is moment magnitude, x is hypocentral distance, and a, b, c, d are regression coefficients.

The optimization of the regression coefficients was executed using a non-linear programming approach to minimize the Sum of Squared Errors (SSE) between the observed and predicted ground motions. Specifically, we employed the Sequential Least SQuares Programming (SLSQP) optimization algorithm. For the baseline model (Site Class D), all primary coefficients (a, b, c, d) were estimated simultaneously to accurately capture the physical trade-offs between magnitude scaling and geometric attenuation.

Due to the heterogeneous spatial distribution of the station network and the uneven number of records per event, a full mixed-effects formulation (separating inter-event and intra-event residual terms) was not implemented, as it would introduce overfitting and convergence issues. Instead, our hybrid regression strategy—supported by the SLSQP optimizer—serves as a pragmatic alternative to

stabilize the path-dependent residuals without violating the region's dominant site characteristics. To ensure the robustness of the derived residuals, a Leave-One-Out Cross-Validation (LOOCV) procedure was subsequently performed during the spatial interpolation analysis (detailed in Section 3.4). The employed **hybrid regression strategy** to address the disparity in data density across site classes:

1. **Baseline Model (Class D):** A full non-linear regression was performed on the abundant Class D dataset to derive robust local coefficients (a_D, b_D, c_D, d_D). Site Class D was selected as the reference baseline because it represents the most geographically ubiquitous and statistically abundant soil type in the dataset (N=289). This empirical dominance is supported by local microzonation studies; for instance, Sari et al. (2019) identified Site Class D (medium/stiff soil) as widely distributed across the western parts of the Bandung Basin, while Widjayanti et al. (2025) confirmed the prevalence of Site Class D in the Serang region, Banten. Therefore, a baseline derived from Class D provides the most representative regional attenuation trend.
2. **Constrained Regression (Site Class C):** Due to the scarcity of near-field records (R < 20 km) for Site Class C, an initial unconstrained optimization using the SLSQP algorithm yielded physically implausible values for the saturation term (d ≈ 3326 km). This numerical artifact indicates a severe parameter identifiability problem, where the lack of near-source data prevents the algorithm from reliably constraining the attenuation curve's flatness near the epicenter. To resolve this, we implemented a constrained regression approach, fixing the saturation parameter d to match the robust baseline derived from Site Class D (d = 5.80). This constraint is strongly supported by theoretical and empirical grounds: the saturation term d is primarily governed by the finite dimensions of the fault rupture and the focal depth, rather than the shallow local site conditions (Kanno et al., 2006). Since the events recorded by both Class C and D stations share the same regional seismotectonic sources, assuming a uniform d for shallow crustal earthquakes is physically consistent.
3. **Site-Term Correction (Class B):** For Site Class B, where data points were insufficient for stable regression (N < 30), we applied a simple site-term correction. We calculated the mean residual of Class B data relative to the Class D baseline and applied this bias as a scalar correction factor to the constant c.

This rigorous approach ensures that the resulting residuals (ϵ) are free from the gross bias of site conditions, as noted in the machine learning validation by Rachmadan et al., (2025).

To further bolster confidence in this assumption, a sensitivity analysis was performed using the Akaike Information Criterion (AIC). The unconstrained model, suffering from the aforementioned identifiability issues, resulted in an unstable fit (AIC = -119.96). In contrast, the constrained model (d=5.80) successfully stabilized the remaining coefficients (source a, attenuation b, and site c) and achieved a superior statistical parsimony (AIC = -121.83). Therefore, fixing d is both a necessary mathematical constraint and a physically sound approach for this dataset.

2.4. Azimuthal and Spatial Decomposition

The total residual for each recording is defined as:

$$\epsilon_{total} = log_{10}(PGA_{obs}) - log_{10}(PGA_{pred}) \quad (2)$$

where PGA_{pred} is calculated using the specific coefficients for the site class of the recording station.

To investigate azimuthal anisotropy, we calculated the source-to-site azimuth (ϕ) for every record and projected the residuals onto a polar coordinate system. This follows the framework of Anggono et al. (2020), who identified crustal shear-wave splitting orientations in West Java. Finally, to isolate path effects such as scattering and basin resonance, the mean residuals were mapped spatially

and overlaid with geological features, including the volcanic arc segmentation defined by Rahmanto et al. (2025) and basin depth models by Saygin et al. (2016). This spatial correlation allows for the physical interpretation of the "Red" (amplification) and "Blue" (attenuation) zones observed in the results.

3. RESULT AND DISCUSSION

3.1. Site-Specific Attenuation Models and the Constrained Regression Approach

The preliminary analysis of the West Java strong-motion dataset demonstrated that a single ergodic attenuation model was insufficient to capture the region's complex site response variability. When the data were processed without site stratification, the standard deviation of residuals remained notably high, suggesting the presence of systematic bias driven by local geological conditions. To address this, we adopted a hybrid regression strategy, stratifying the dataset according to the SNI 1726:2019 classification standards. The calibrated coefficients for each site class are presented in Table 1.

Table 1. Calibrated Regression Coefficients for West Java

Site Class (SNI 1726)	VS30 Range (m/s)	Regression Method	a (Magnitude)	b (Dist. Attenuation)	c (Site/Const)	d (Saturation, km)
Class D (Medium Soil)	175<VS30≤350	Full Non-Linear Regression	0.62	-0.005	0.006	5.8
Class C (Stiff Soil)	350<VS30≤750	Constrained Regression (d fixed)	1.304	0.0046	-1.6188	5.80*
Class B (Rock)	750<VS30<1500	Site-Term Correction	-	-	+0.2341**	-

Note: Parameter fixed to match Class D baseline. ** Value indicates correction factor added to Class D constant.*

A significant methodological challenge was encountered during the analysis of **Site Class C (Stiff Soil)**. Due to the paucity of near-field records (R < 20 km) captured by stations on stiff soil, the initial unconstrained regression yielded physically implausible values for the saturation term ($d \gg 30\ km$), a numerical artifact implying an unrealistic flatness of the attenuation curve in the near-source region. To rectify this, we implemented a **constrained regression approach**, fixing the saturation parameter d to match the robust baseline derived from Site Class D (d=5.80). The use of Site Class D as the primary baseline is empirically supported by recent microzonation studies in the region. Sari et al. (2019) identified Site Class D as a widespread soil class in the western Bandung Basin, while Widjayanti et al. (2025) confirmed its prevalence in Serang, Banten. Given its statistical abundance in our dataset (N=289), Class D provides the most stable reference for regional attenuation. This decision rests on the physical assumption that depth-saturation characteristics are primarily governed by the source geometry rather than local site conditions; therefore, the parameter should remain consistent across site classes for shallow crustal events. By constraining d, the regression successfully isolated and stabilized the magnitude scaling (a) and anelastic attenuation (b) terms.

The results reveal a striking disparity in the magnitude scaling coefficient (a) between site classes. For Site Class C, the coefficient is **1.304**, which is more than double the baseline value for Site Class D (a=0.620). This suggests that ground motions on stiff soil scale more aggressively with magnitude. This phenomenon is likely attributed to the **absence of non-linear soil damping** on stiff sites; unlike soft soil deposits (Class D) which tend to dissipate high-frequency energy during large magnitude events due to hysteretic damping, stiff soils transmit this energy more efficiently. This observation aligns with high-frequency attenuation behaviors reported in other subduction environments, such as those analyzed by Zhao & Xu (2013).

Furthermore, the analysis of **Site Class B (Rock)** uncovered an unexpected positive site correction factor ($c_{corr} = +0.2341$) relative to the soft soil baseline. Theoretical models typically predict de-amplification on rock sites; however, this anomaly can be rationalized by considering the geomorphological context of the station network. As noted by Rudyanto et al. (2024), a significant portion of the BMKG rock stations in West Java are installed on hilltops or mountain ridges to minimize cultural noise. Such locations are highly susceptible to **topographic amplification (ridge effects)**, where seismic waves are focused at the crest, resulting in PGA values that exceed those expected for flat rock sites. Additionally, the thin weathering layer characteristic of these sites minimizes the high-frequency absorption parameter (κ_0), allowing short-period energy to dominate the record, a factor often underestimated in standard global GMPEs (Boore & Atkinson, 2008)

3.2. The Explanatory Power of V_{S30} for Remaining Residual Variability

Following the application of the site-class specific corrections detailed in Table 1, the logical progression is to evaluate whether the remaining residual variability (σ) can be attributed to finer variations in V_{S30}. In modern GMPE development, the time-averaged shear-wave velocity in the top 30 meters is frequently regarded as the "gold standard" for parameterizing site amplification. However, our residual analysis uncovers a limitation that challenges the dominance of this parameter within the complex tectonic setting of West Java.

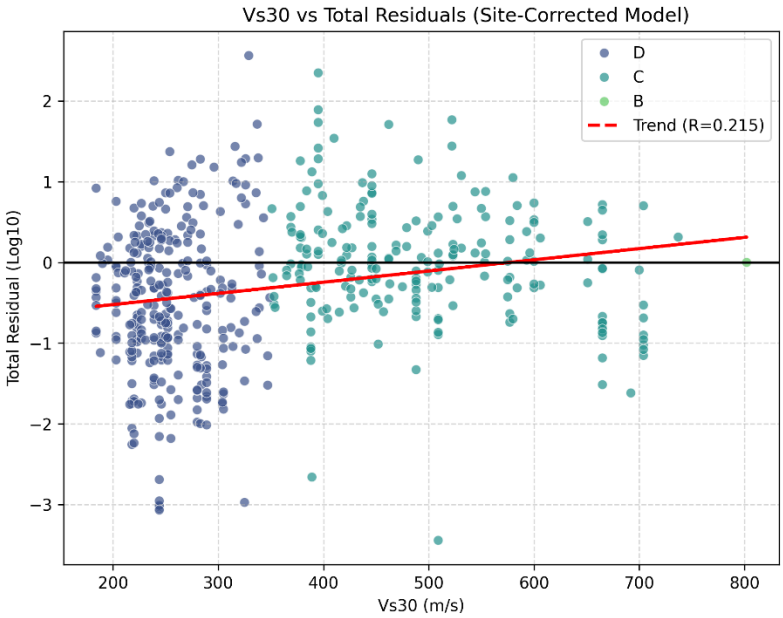

Figure 2. Scatter plot showing the correlation between total residuals (ϵ) and measured V_{S30} after applying site-class specific corrections. The weak correlation (R = 0.215) indicates that intra-class velocity variations explain only a fraction of the remaining ground motion variability, suggesting that path effects are the more dominant source of uncertainty.

As illustrated in **Figure 2**, the correlation between the total residuals (ϵ) and the measured V_{S30} is weak (R = 0.215). While a slight trend is observable, the coefficient of determination ($R^2 \approx 0.046$) indicates that intra-class velocity variations explain less than 5% of the remaining ground motion variability. This finding demonstrates that once the data are stratified into broad geological categories using the SNI 1726:2019 framework, intra-class variations in V_{S30} exhibit low explanatory power for the remaining ground motion variability. We do not suggest that V_{S30} is an inherently flawed proxy; rather, within this specific residual framework, its utility is overshadowed by non-ergodic path effects. This corroborates the recent machine learning analysis by Rachmadan et al. (2025), where V_{S30} was ranked

significantly lower in feature importance compared to geometric path parameters. The limited capacity of a simple 1D V_{S30} profile to capture the total residual variance likely stems from the deep subsurface architecture of West Java. In profound sedimentary environments like the Jakarta Basin, ground motion amplification is often controlled by basin depth and low-frequency resonance that extend hundreds of meters—parameters that remain largely "invisible" to a shallow 30-meter metric.

In deep sedimentary environments like the Jakarta Basin, ground motion amplification is often controlled by basin depth and resonance frequencies that extend hundreds of meters into the subsurface parameters that remain "invisible" to the V_{S30} metric. A similar decoupling of V_{S30} from site response was observed by Vyas et al. (2016) in the Himalayan region, where extreme topography and crustal heterogeneity rendered simple 1D proxies ineffective.

The implication of these findings is critical: since the substantial residual spread is not primarily driven by local site conditions (V_{S30}), the dominant source of uncertainty must lie within the **wave propagation path**. Consequently, the persistence of these residuals compels us to shift our analytical focus from the vertical domain (local soil profile) to the horizontal domain (azimuthal and spatial heterogeneity), which we investigate in the following section.

3.3. Azimuthal Anisotropy and Path-Dependent Residuals

Since the site-specific calibration (V_{S30} and Site Class) failed to account for the majority of the ground motion variability ($\sigma_{total} \approx 0.30$), we shifted our investigation to the horizontal domain: the wave propagation path. By projecting the total residuals (ϵ) onto a polar coordinate system relative to the source-to-site azimuth (Az), a systematic non-ergodic pattern emerged.

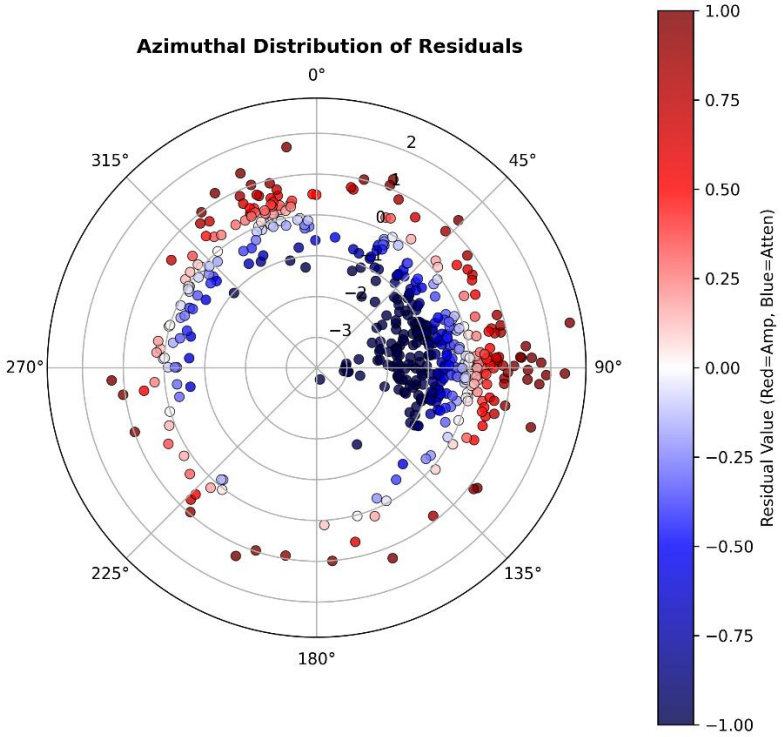

Figure 3. Polar plot of ground motion residuals (ϵ) derived from the site-corrected model. The radial axis represents the residual magnitude, while the angular axis represents the source-to-site azimuth ($0° = North$). The plot reveals a strong dipolar anisotropy: significant attenuation (blue) is observed in the SW azimuths facing the volcanic arc, while amplification (red) occurs in the NE azimuths towards the sedimentary basins.

As visualized in Figure 3, the distribution of residuals is not random but exhibits a distinct **dipolar anisotropy**. A clear dichotomy exists: negative residuals (indicating strong attenuation/blue zones) are predominantly clustered in the Western and Southwestern sectors ($225° \leq Az \leq 315°$), while positive residuals (indicating amplification/red zones) dominate the Northern and Eastern sectors ($45° \leq Az \leq 135°$). This azimuthal dependence is statistically significant. The boxplot analysis in Figure 3 confirms that the median residual shift between these opposing sectors exceeds 0.4 log units, a variance far greater than what is typically attributed to random source effects (Supendi et al., 2025).

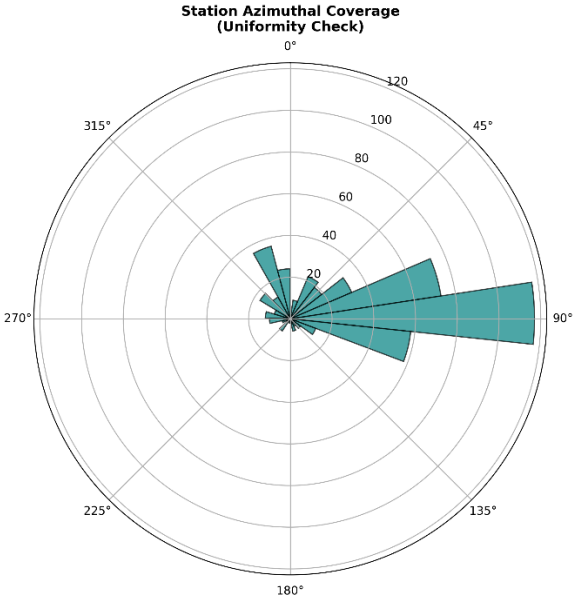

Figure 4. Azimuthal coverage of the recording stations relative to the earthquake epicenters. The polar histogram displays the number of records within each 15° azimuthal bin. The distribution covers all azimuthal sectors ($0° - 360°$) without significant gaps, with a count standard deviation of $\sigma \approx 26.6$. This uniformity confirms that the dipolar residual pattern observed in this study is a genuine physical feature of the wavefield and not an artifact of biased network geometry.

To verify that the observed dipolar anisotropy is a statistically significant physical feature and not an artifact of random data distribution, we performed a formal hypothesis test on the residuals grouped by 45° azimuthal sectors. First, a Shapiro-Wilk test indicated that the residuals depart from a perfect normal distribution (p-value = 0.003), prompting the use of the non-parametric Kruskal-Wallis H-test.The Kruskal-Wallis test yielded an H-statistic of 69.71 and a highly significant p-value of 1.68×10^{-12}. This allows us to definitively reject the null hypothesis, confirming that the ground motion variability in West Java is systematically and strongly dependent on the propagation path. This statistical robustness provides a firm foundation for the physical claim that path effects—specifically volcanic scattering and basin resonance—are the primary drivers of non-ergodic variability in the region, rather than localized source or site effects.

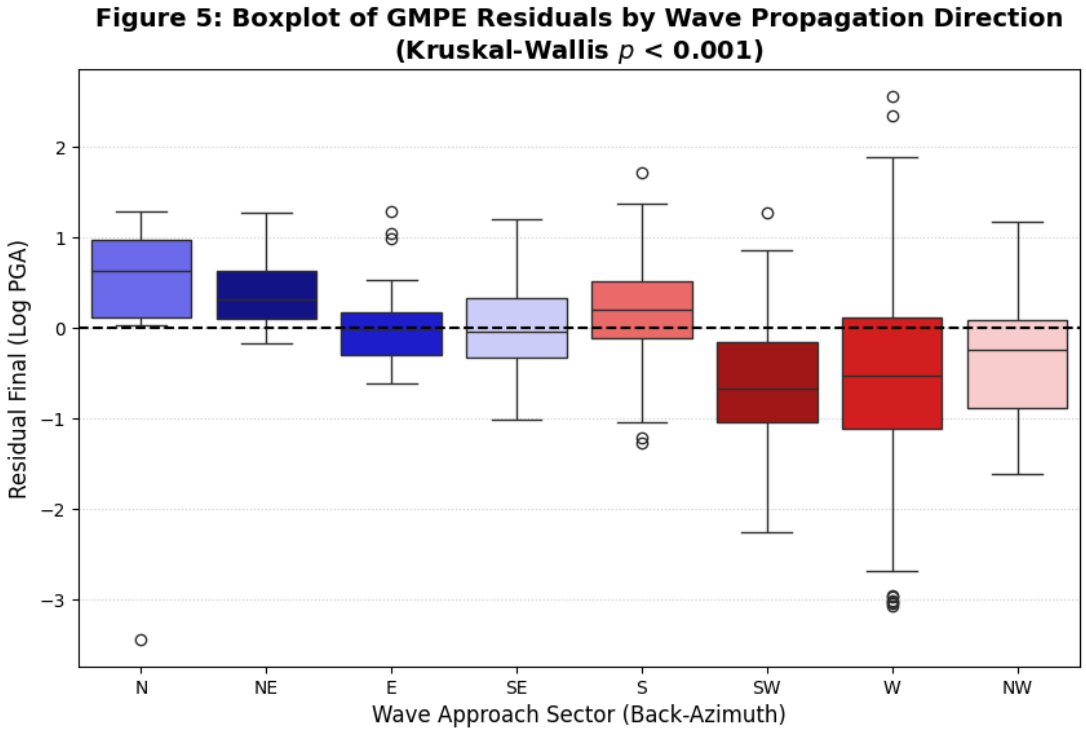

Figure 5. Statistical boxplot of residuals grouped by wave approach sector (Back-Azimuth, 45° bins). The Kruskal-Wallis test confirms a highly significant variance (p < 0.001). The separation between the median residuals of the West-Southwest sectors (strong negative bias / attenuation) and the North-East sectors (positive bias / amplification) indicates that the azimuthal dependence is a geologically driven feature of the wavefield.

This anisotropic behavior aligns remarkably well with the crustal stress field of West Java. Anggono et al. (2020) identified a consistent shear-wave splitting parameter with a fast-polarization direction generally parallel to the trench (East-West). Seismic waves propagating along this strike-parallel direction (towards the East) travel through relatively intact lithospheric blocks, preserving their high-frequency energy. In contrast, waves propagating orthogonal to the structure (towards the volcanic arc in the South/West) encounter high fracture density and thermal anomalies. This interpretation is supported by the strain partitioning models of Koulali et al. (2017), which delineate the localized shear zones that act as waveguides or barriers depending on the propagation angle.

The "Blue Zone" of strong attenuation in the Southwest correlates spatially with the segmentation of the Quaternary Volcanic Arc. As detailed by Rahmanto et al. (2025), this region is characterized by intense magmatic intrusion and crustal heterogeneity. Theoretical models of multiple scattering by Bracale et al. (2025) suggest that such heterogeneous media act as low-pass filters, rapidly scattering high-frequency PGA (f > 5 Hz). Consequently, the negative residuals here represent a "scattering loss" that standard GMPEs which assume a homogeneous Q-structure fail to predict.

Conversely, the "Red Zone" of amplification in the North cannot be explained solely by source radiation patterns. This cluster spatially coincides with the deep alluvial deposits of the Jakarta Basin. Saygin et al. (2016) mapped the basement geometry of this basin using seismic noise tomography, revealing sediment thicknesses that facilitate trapped-wave resonance. Although our site classification (Table 1) attempted to correct for soil softness, the *resonance effect* of the deep basin introduces prolonged shaking duration and constructive interference that a simple V_{S30} term cannot capture. This aligns with the hazard warnings by Cipta et al. (2018) and Irsyam et al. (2020) regarding the spectral

amplification potential in the capital region. Thus, the positive residuals here are likely a composite signature of both path-guided waves and basin resonance.

3.4. Spatial Distribution of Path Effects: Volcanic Scattering and Basin Resonance

To determine the optimal interpolation technique for mapping the azimuthal residuals, we conducted a Leave-One-Out Cross-Validation (LOOCV) comparison between Inverse Distance Weighting (IDW, power=2) and Radial Basis Function (RBF) interpolation. The analysis revealed that the RBF method suffered from severe numerical instability (ill-conditioned matrices) due to the dense clustering of stations in the metropolitan areas, yielding divergent error metrics. In contrast, the IDW method proved to be numerically robust, producing a stable interpolation with a Root Mean Square Error (RMSE) of 0.67 log units. While this variance reflects the inherent heterogeneity of the path effects, the IDW approach successfully delineates the regional attenuation and amplification patterns without introducing mathematical artifacts. Consequently, IDW was selected as the most representative method for this dataset.

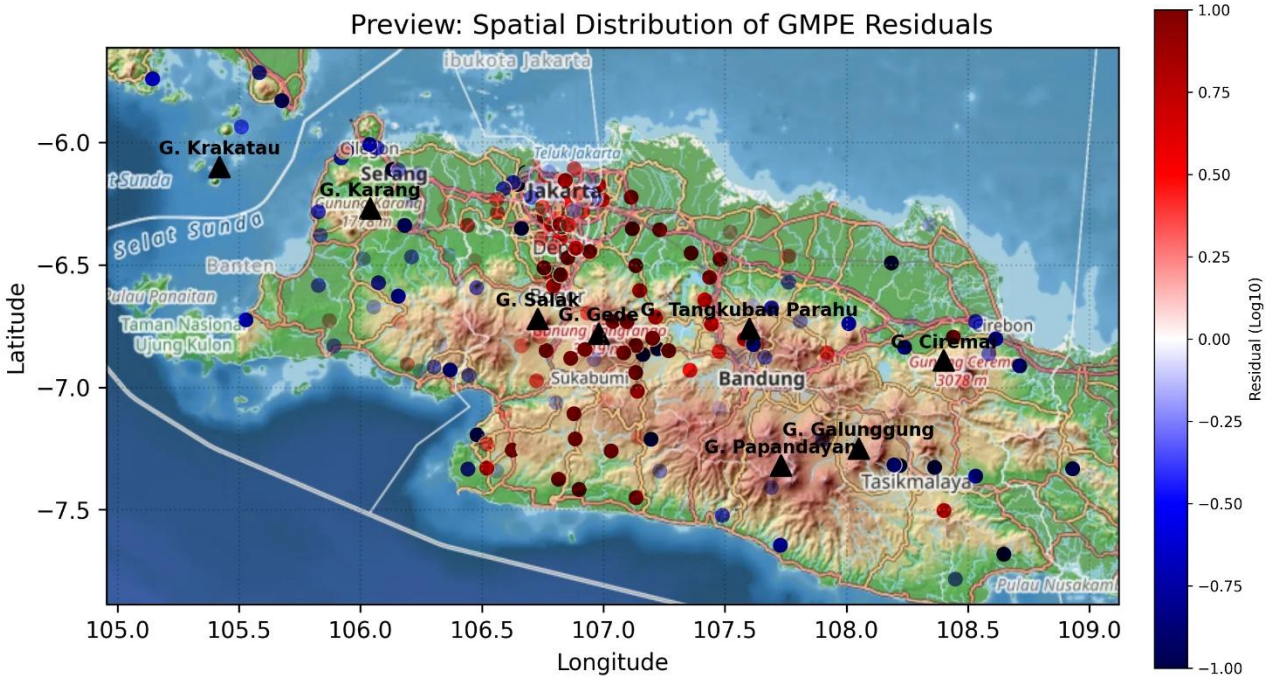

Figure 6. Spatial distribution map of mean GMPE residuals (ϵ) overlaid with major geological features, generated using QGIS 3.28. Blue contours indicate zones of strong attenuation (negative residuals) coinciding with the Quaternary Volcanic Arc (black triangles), attributed to scattering effects in fractured media. Red contours indicate zones of amplification (positive residuals) aligning with the deep sedimentary deposits of the Jakarta and Bandung Basins. The distinct spatial separation confirms that path effects are controlled by the regional geological architecture.

A distinct zone of strong negative residuals ($-0.5 \leq \epsilon \leq -1.2$), depicted as the "Blue Zone" in Figure 6, stretches across the southern and western highlands. This area spatially coincides with the segmented volcanic arc identified by Rahmanto et al. (2025), a region characterized by high thermal gradients and intense crustal fracturing. The significant attenuation of Peak Ground Acceleration (PGA) in this zone is consistent with the physics of multiple scattering in heterogeneous media. This 'volcanic shadow' effect is likely driven by multiple scattering mechanisms in the highly fractured and thermally anomalous magmatic media of the Sunda Arc. As suggested by Bracale et al. (2025), short-wavelength seismic energy (high-frequency PGA) undergoes rapid scattering when propagating through

heterogeneous volcanic structures, which act as a natural low-pass filter. This interpretation is consistent with the volcanic segmentation model by Rahmanto et al. (2025).

In sharp contrast, the Northern part of the study area encompassing the Jakarta and Bandung metropolitan areas exhibits a cluster of positive residuals ($+0.4 \leq \epsilon \leq +0.9$), marked as the "Red Zone.". This amplification cannot be attributed to soft soil conditions alone, as our site-specific model (Table 1) has already corrected for V_{S30} effects, as our model has already corrected for site terms. The elevated hazard in these sectors is further corroborated by independent microtremor studies, such as the mapping of high seismic vulnerability indices in West Java by Erdison et al. (2024) using the HVSR method. This persistent amplification aligns perfectly with the deep basin geometry mapped by Saygin et al. (2016) using seismic noise tomography. The sediment thickness in the Jakarta Basin, reaching depths of over 1000 meters, traps seismic energy and generates basin-edge induced surface waves. Cipta et al. (2018) warned that such basin resonance effects could prolong shaking duration and amplify spectral accelerations beyond the design levels of the current Indonesian building codes. Furthermore, the proximity of this amplified zone to the active Baribis Fault, as delineated by Irsyam et al. (2020) and Koulali et al. (2017), suggests a compounded seismic hazard scenario where path-dependent amplification could exacerbate the impact of near-field crustal rupture. Furthermore, the observed positive residuals in the northern sectors are not only a product of basin resonance but are also closely linked to the regional tectonic architecture. Aribowo et al. (2022) highlighted the intricate relationship between active fault propagation and heavy foreland sedimentation in North West Java. This structural setting creates thick sedimentary traps that potentially enhance the 'Red Zone' amplification, where seismic energy becomes trapped and amplified within the North Java basin's complex layers.

Discussion on Residual Variance: Regarding the LOOCV result, the RMSE of 0.67 log units, while seemingly high, is reflective of the extreme crustal heterogeneity of the West Java-Banten region. This 'unresolved variability' is typical for regional non-ergodic studies where 3D path effects are projected onto a 2D map. However, our Kruskal-Wallis test ($p < 0.001$) confirms that despite the local variance, the large-scale azimuthal dipolar pattern is a statistically robust physical feature that significantly outperforms standard isotropic models. Therefore, the spatial segregation of residuals into "volcanic attenuators" and "basin amplifiers" serves as critical evidence that ground motion in West Java is governed by a **non-ergodic path process**. Relying solely on distance-based attenuation models (R) without accounting for these specific geological path effects may lead to a dangerous underestimation of hazard in the basins and an overestimation in the volcanic highlands.

4. CONCLUSION

This study has successfully recalibrated a local Ground Motion Prediction Equation (GMPE) for West Java by implementing a hybrid regression strategy that accounts for site-specific disparities. The adoption of a constrained regression approach for Site Class C, where the saturation parameter d was fixed to match the robust Class D baseline has proven effective in stabilizing coefficients despite the paucity of near-field data. Our findings demonstrate that while site-class stratification is essential for reducing broad biases, the traditional reliance on V_{S30}as a primary proxy for site response is highly suboptimal in this region. The weak correlation between V_{S30} and residuals ($R = 0.215$) confirms that intra-class velocity variations explain less than 5% of the remaining ground motion variability, suggesting that vertical soil profiles are secondary to deeper crustal architecture.

The core novelty of this research lies in the identification of a distinct azimuthal anisotropy in ground motion residuals, which follows a systematic dipolar pattern. We have identified two geologically controlled path effects: (1) a "Blue Zone" of strong high-frequency attenuation in the Southwest, likely driven by multiple scattering within the fractured and thermally anomalous Quaternary Volcanic Arc, and (2) a "Red Zone" of significant amplification in the North, attributed to basin resonance and trapped-wave effects within the deep Jakarta and Bandung sedimentary basins. These non-ergodic path effects introduce systematic errors that standard GMPEs which assume isotropic attenuation consistently fail to capture.

Ultimately, these results underscore the urgent necessity of moving beyond 1D site-effect models in West Java's seismic hazard assessments. Integrating azimuthal-dependent path terms into regional hazard maps and Earthquake Early Warning (EEW) algorithms is not merely a theoretical refinement; it is a critical requirement for mitigating risks in high-density metropolitan areas situated over deep basins. Future research should prioritize high-resolution Q-tomography to further quantify the frequency-dependent nature of volcanic scattering identified in this study.

REFERENCES

Allen, T. I., & Wald, D. J. (2009). On the use of high-resolution topographic data as a proxy for seismic site conditions (VS30). Bulletin of the Seismological Society of America, 99(2 A), 935–943. https://doi.org/10.1785/0120080255

Aribowo, S., Husson, L., Natawidjaja, D. H., Authemayou, C., Daryono, M. R., Puji, A. R., ... & Djarwadi, D. (2022). Active back-arc thrust in North West Java, Indonesia. Tectonics, 41(10), e2021TC007120. https://doi.org/10.1029/2021TC007120

Bindi, D., Pacor, F., Luzi, L., Puglia, R., Massa, M., Ameri, G., & Paolucci, R. (2011). Ground motion prediction equations derived from the Italian strong motion database. Bulletin of Earthquake Engineering, 9(6), 1899–1920. https://doi.org/10.1007/s10518-011-9313-z

Boore, D. M., & Atkinson, G. M. (2008). Ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods between 0.01 s and 10.0 s. Earthquake Spectra, 24(1), 99–138. https://doi.org/10.1193/1.2830434

Bracale, M., Campillo, M., Shapiro, N., Brossier, R., Melnik, O., & Shapiro, N. M. (2025). Multiple scattering of seismic waves in a heterogeneous magmatic system and spectral characteristics of long period volcanic earthquakes. Journal of Geophysical Research : Solid Earth, 2025(12), 2025–031850. https://doi.org/10.1029/2025JB031850

Cipta, A., Cummins, P., Irsyam, M., & Hidayati, S. (2018). Basin resonance and seismic hazard in jakarta, Indonesia. Geosciences (Switzerland), 8(4), 1–25. https://doi.org/10.3390/geosciences8040128

Erdison, T. A., Pranata, B., & Hafidz, A. (2024). Analisis Indeks Kerentanan Guncangan Gempabumi di Provinsi Jawa Barat Menggunakan Data Mikrotremor Berdasarkan Metode Horizontal to Vertical Spectral Ratio (HVSR). (HAGI)

Gunawan, E., & Widiyantoro, S. (2019). Active tectonic deformation in Java, Indonesia inferred from a GPS-derived strain rate. Journal of Geodynamics, 123, 49–54. https://doi.org/10.1016/j.jog.2019.01.004

Irsyam, M., Cummins, P. R., Asrurifak, M., Faizal, L., Natawidjaja, D. H., Widiyantoro, S., Meilano, I., Triyoso, W., Rudiyanto, A., Hidayati, S., Ridwan, M., Hanifa, N. R., & Syahbana, A. J. (2020). Development of the 2017 national seismic hazard maps of Indonesia. Earthquake Spectra, 36(1_suppl), 112–136. https://doi.org/10.1177/8755293020951206

Kanno, T., Narita, A., Morikawa, N., Fujiwara, H., & Fukushima, Y. (2006). A new attenuation relation for strong ground motion in Japan based on recorded data. Bulletin of the Seismological Society of America, 96(3), 879–897. https://doi.org/10.1785/0120050138

Koulali, A., McClusky, S., Susilo, S., Leonard, Y., Cummins, P., Tregoning, P., Meilano, I., Efendi, J., & Wijanarto, A. B. (2017). The kinematics of crustal deformation in Java from GPS observations: Implications for fault slip partitioning. Earth and Planetary Science Letters, 458, 69–79. https://doi.org/10.1016/j.epsl.2016.10.039

Mukherjee, B., Shaw, R. L., Sharma, M. L., & Sain, K. (2025). Earthquake prediction using machine learning perspectives in Himalayan seismic belt and its surroundings. Journal of Asian Earth Sciences, 293, Article 106764. https://doi.org/10.1016/j.jseaes.2025.106764

Rachmadan, A., Koeshidayatullah, A., & Kaka, S. L. I. (2025). Developing ground motion prediction models for West Java: A machine learning approach to support Indonesia's earthquake early warning system. Applied Computing and Geosciences, 25. https://doi.org/10.1016/j.acags.2024.100212

Rahmanto, R., Saepuloh, A., Kriswati, E., & Purnamasari, H. D. (2025). Volcanoes Segmentation at the Western Sunda Arc based on Satellite-derived Geological Lineaments and Land Surface Temperatures. Journal of Engineering and Technological Sciences, 57(3), 327–342. https://doi.org/10.5614/j.eng.technol.sci.2025.57.3.4

Rudyanto, A., Wijaya, A., Widiyantoro, S., Sahara, D. P., Rosalia, S., Wibowo, A., Pramono, S., & Putra, A. S. (2024). Performance test of pilot Earthquake Early Warning system in western Java, Indonesia. International Journal of Disaster Risk Reduction, 115. https://doi.org/10.1016/j.ijdrr.2024.105010

Sari, A. M., & Fakhrurrozi, A. (2020). SEISMIC HAZARD MICROZONATION BASED ON PROBABILITY SEISMIC HAZARD ANALYSIS IN BANDUNG BASIN. RISET Geologi Dan Pertambangan, 30(2), 215. https://doi.org/10.14203/risetgeotam2020.v30.1138

Sari, A., Soebowo, E., Fakhrurrozi, A., Syahbana, A., & Tohari, A. (2019). Microzonation of Soil Amplification Based on Microtremor, Spt and Cptu Data in Bandung Basin. Riset Geologi dan Pertambangan, 29(1), 53-64. https://dx.doi.org/10.14203/risetgeotam2019.v29.978

Sarkar, P., Shaw, R. L., Mukherjee, B., Dutta, B., Tiwari, A., Roy, P. N. S., Prajapati, S. K., & Sharma, M. L. (2026). TEC variation as earthquake precursor: A statistical and SARIMA-based study from Northeast India. Advances in Space Research, 77(5), 6184–6212. https://doi.org/10.1016/j.asr.2025.12.083

Saygin, E., Cummins, P. R., Cipta, A., Hawkins, R., Pandhu, R., Murjaya, J., Masturyono, Irsyam, M., Widiyantoro, S., & Kennett, B. L. N. (2016). Imaging architecture of the Jakarta Basin, Indonesia with transdimensional inversion of seismic noise. Geophysical Journal International, 204(2), 918–931. https://doi.org/10.1093/gji/ggv466

Shaw, R. L., Mukherjee, B., Sharma, M. L., & Kar, S. (2026). Toward Location Reliant Early Earthquake Detection: A Paradigm Deployed Deep Learning Algorithms and Clustered Seismic Indicators. Journal of Earthquake Engineering, 30(2), 488–516. https://doi.org/10.1080/13632469.2025.2565617

Shaw, R. L., Mukherjee, B., Tiwari, A., & Sharma, M. L. (2025). b-value and fractal dimension assisted spatiotemporal seismicity pattern assessment along Himalayan seismic belt. Journal of Seismology, 29, 1337–1361. https://doi.org/10.1007/s10950-025-10325-9

Supendi, P., Widiyantoro, S., Rawlinson, N., Daryono, M. R., Ardianto, A., Baskara, A. W., Damanik, R., & Husni, Y. M. (2025). Evidence of the West Java back-arc thrust from earthquake activity. Tectonophysics, 911. https://doi.org/10.1016/j.tecto.2025.230853

Vyas, J. C., Mai, P. M., & Galis, M. (2016). Distance and azimuthal dependence of ground-motion variability for unilateral strike-slip ruptures. Bulletin of the Seismological Society of America, 106(4), 1584–1599. https://doi.org/10.1785/0120150298

Widjayanti, R. Y., AL HS, U. A., Rahman, A. S., Permana, D., & Alfiandy, S. (2025). Analisis Nilai Percepatan Tanah Maksimum dan Klasifikasi Kelas Situs di Kota Serang Provinsi Banten. Buletin GAW Bariri (BGB), 6(1), 1-7. https://doi.org/10.31172/bgb.v6i1.128

Zhao, J. X., & Xu, H. (2013). A Comparison of VS30 and Site Period as Site-Effect Parameters in Response Spectral Ground-Motion Prediction Equations. Bulletin of the Seismological Society of America, 103(1), 1–18. https://doi.org/10.1785/0120110251