



OPEN

Spatial correlation assessment of multiple earthquake intensity measures using physics-based simulated ground motions

Mohammad R. Zolfaghari & Mahboubeh Forghani

Predictive models for spatial correlation play an effective role in the assessment of seismic risk associated with distributed infrastructure and building portfolios. However, existing models often rely on simplified approaches, assuming isotropy and stationarity. This paper verifies these assumptions by presenting a comprehensive study using a database of 3D physics-based simulated broadband ground motions for Istanbul, generated by the SPEED software. The results reveal significant event-to-event variability and nonstationary and anisotropic characteristics of spatial correlation influenced by source, path, and site effects. The development of nonstationary correlation models requires exploring influential metrics beyond spatial proximity and gaining a deep understanding of their impact, which is the focus of this study. Analysis of the spatial correlations of peak ground displacement, peak ground velocity, peak ground acceleration, and response spectral accelerations at different periods, employing both stationary and nonstationary correlation modelling methods and considering the finite fault model, indicates that the slip distribution pattern, direction and distance of station pairs relative to earthquake rupture, soil softness, and homogeneity of soil properties significantly influence the spatial correlations of near-field earthquake ground motions. Implementation of the introduced parameters in predictive spatial correlation models enhances the precision of regional seismic hazard assessments.

Keywords Spatial correlation, 3D physics-based numerical simulation, Near-source ground motion, Regional seismic hazard

Probabilistic seismic hazard analysis (PSHA) is the first important step towards any seismic risk assessment study that provides a probabilistic estimate of ground motion intensity measures (IMs) for a specific site. In this regard, many ground motion prediction equations (GMPE) have been developed to predict the mean and standard deviation of IMs according to earthquake and site characteristics like magnitude, distance to rupture, soil type, fault mechanism, etc. The differences between recorded IMs and estimated IMs using GMPEs are referred to as residuals and are traditionally considered spatially independent. However, previous studies have revealed that the residuals are spatially correlated as a result of the similar source properties, wave propagation path, and site effects, which have not been considered by GMPEs due to their simple functional forms or the insufficiency of records of strong ground motion (SGM) data. The spatial correlation assessment of multiple earthquake intensity measures is essential for accurately evaluating seismic risks, particularly for spatially distributed infrastructure like bridges, lifelines, and building portfolios. When such structures are subjected to the same earthquake, it is unrealistic to treat ground-motion intensity measures at different sites as independent, especially when these sites are in close proximity. Ignoring this spatial correlation can result in inaccurate seismic risk assessments, potentially underestimating or overestimating the actual risks¹⁻⁷. This highlights the importance of considering accurate spatial correlation in understanding the potential impacts of earthquakes on infrastructure.

Many studies have been conducted to develop spatial correlation models for a variety of IMs using ground motion recordings of past earthquakes and geostatistical methods. Due to the shortage of recorded ground motion data, such models mostly assume the stationary (i.e., the site-to-site distance is the only impressive factor in the correlation of the station pairs) and isotropy (i.e., the correlation is not related to the direction of the station pairs) hypotheses⁸⁻¹⁴. These models have significant variability in parameters such as correlation range

Civil Engineering Department, K. N. Toosi University of Technology, Tehran, Iran. email: m.forghani@email.kntu.ac.ir

(the distance beyond which the spatial correlation is negligible). In recent years, a number of studies have been carried out to evaluate the reasons behind the significant heterogeneity in spatial correlation models. Schiappapietra and Douglas¹⁵ provided a literature review of the major potential factors influencing spatial correlation, such as estimation approach, fitting method, applied GMPE, magnitude (Mw), period, and geology. Baker and Chen¹⁶ proposed a procedure to quantify estimation uncertainty and the true variability of spatial correlation. Schiappapietra and Douglas¹⁷ presented some guidelines for considering the uncertainty of spatial correlation in seismic hazard studies. Furthermore, Heresi and Miranda¹¹ underlined the importance of taking into account the notable earthquake-to-earthquake variability of spatial correlation model parameters. Also, some studies have shown that stationarity and isotropy assumptions may need further reconsideration. For example, Sokolov et al.¹⁸ showed the dependency of correlation on geology and propagation path using Taiwan strong ground motion data. Du and Wang¹⁹ indicate the dependency of the correlation of IMs on the regional spatial correlation of V_{S30} (30-m time-averaged shear wave velocity). Garakaninezhad and Bastami²⁰ reported the anisotropy of the PGA residual's correlation in nine past earthquakes related to the focal mechanism, magnitude, and anisotropy of V_{S30} . Abbasnejadfar et al.²¹ showed that the anisotropy direction of residuals of multiple IMs and V_{S30} is similar and Stafford et al.²² denoted that different rupture characteristics lead to variation in the correlation range. Moreover, the recent progress on three-dimensional physics-based numerical simulations of strong ground motions provides the possibility of exploring the dependency of spatial correlation on different potential factors, taking advantage of the richness of simulated data and the exact knowledge of earthquake rupture, path, and site characteristics. Using ground motion data simulated by CyberShake, Chen and Baker²³ detected the nonstationary and anisotropic patterns of the spatial correlation related to the source effect, path effect, and relative location to rupture. Likewise, Infantino et al.²⁴ verified the spatial correlation structure of the simulated broadband ground motions generated by SPEED software in order to explore the variation of SGM spatial correlation, taking into account period, magnitude, directivity effects, and fault normal or fault parallel components. Schiappapietra and Smerzini²⁵ used scenario-based ground motion data simulated by the SPEED code for Norcia (central Italy) to analyze the spatial correlation of ground-motion components (fault-normal, fault-parallel, vertical) with emphasis on the source effect. Furthermore, Chen et al.²⁶ attempted to quantify the non-stationarity of the spatial correlation effect by calculating the site-specific correlation deviation relative to the global models utilizing recent densely recorded ground motions in New Zealand earthquakes. Also, Chen and Baker²⁷ proposed a community detection method to find areas with spatial correlations higher than those estimated by the global correlation models, which could be helpful to gain a much better understanding of influencing parameters such as relative position to rupture and geological conditions.

As a summary, these studies criticize the assumptions of isotropy and stationarity in spatial correlation models. These studies highlight the significant variability and influence of various factors on spatial correlation. Limited recent studies have started to address this by developing nonstationary and anisotropic spatial correlation models. For example, Bodenmann et al.²⁸ have proposed a correlation model that incorporates path and site effects using the NGA-West2 database. Similarly, Liu et al.²⁹ have modeled the spatial correlation structures of systematic source, path, and site effects for the Ridgecrest area. However, more extensive studies utilizing diverse and dense SGM databases are still needed to identify and quantify explanatory variables and their impacts. This is a crucial step towards the development of more precise and sophisticated correlation models. This paper, therefore, presents a comprehensive study on the spatial correlation properties of PGA (peak ground acceleration), PGV (peak ground velocity), PGD (peak ground displacement), and SA (spectrum acceleration) for a broadband range of periods in a more detailed and quantitative manner. The main objective of this research is to survey the nonstationary and anisotropic patterns of spatial correlations to explore the most impressive factors and their resultant variation in spatial correlation. The outcomes would be helpful in determining whether the nonstationary and anisotropic modelling of spatial correlation is essential and, if so, how this could be implemented in studies of seismic hazard and risk.

Simulated ground motions are widely used in seismic risk assessment^{30,31} and can be achieved through various approaches. Empirical GMPEs are computationally efficient but have limitations in capturing correlation structures, site-specific effects, and near-fault ground motion characteristics. Stochastic methods offer greater flexibility but rely on assumptions about source, path, and site parameters. Physics-based simulations deliver the highest level of detail but are computationally intensive. Hybrid approaches aim to balance accuracy and efficiency by integrating elements of both empirical and physics-based methods. Physics-based simulations (PBS) are particularly powerful for assessing spatial correlations, as they provide detailed insights into how seismic source characteristics, wave propagation paths, and local geological conditions influence ground motions. Unlike traditional recordings, which are limited in number and spatial distribution, PBS generates extensive and systematically spatially distributed ground motion data, allowing for the quantification of spatial correlations while relaxing the assumptions of stationarity and isotropy. This approach enables the calculation of correlation coefficients for every pair of stations in a region and facilitates the investigation of how source, path, and site geology impact spatial correlations. In this study, a 3D physics-based simulated broadband SGM database for Istanbul, generated by SPEED software, is employed. The richness of this catalog, with 65 earthquake scenarios and simulated SGM records for 7343 stations, provides a unique opportunity to access the anisotropy and non-stationary properties of SGM spatial correlation. The characteristics of this database are explained in more detail in "Data" of this paper. This paper employs two different approaches to investigate the spatial correlations of SGM. The first method is discussed in sections "Calculation of SGM residuals" and "Stationary spatial correlation model", wherein traditional geostatistical techniques are used to estimate and compare the SGM stationary correlations. This method focuses on examining the stationary assumption and event-to-event correlation variability with respect to the response period. The source effect on spatial correlation is assessed in Section "Source effect" by investigating the variation of the correlation range considering the slip distribution and magnitude. On the other hand, the second method, which is presented in Section "Path and site effect", includes computing

nonlinear nonstationary spatial correlations for each station pair individually. The obtained results from this approach help identify significant contributing factors, including the relative distance and direction of stations to the rupture, indicating the path effect, and soil properties representing site effects. In order to understand the impact of these factors on the variability of the correlation range, the coefficient of variation of correlation ranges is computed with respect to the variation of each metric.

Data

In this study, we utilize a comprehensive 3D physics-based simulated broadband Strong Ground Motion (SGM) database for Istanbul, which was generated by SPEED (<https://speed.mox.polimi.it/istanbul-earthquake>). SPEED (Spectral Elements in Elastodynamics with Discontinuous Galerkin) is an advanced open-source code that uses three-dimensional physics-based numerical simulation of the wave propagation, topography models, and kinematic models of seismic fault rupture³². The distinct advantages of SPEED include its ability to incorporate near-source earthquake features, such as directivity effects³³, and its reliability across a broad period range, achieved through an innovative artificial neural network-based approach known as ANN2BB³⁴.

Even though physics-based simulations cannot reproduce exactly the real ground motions, studies have shown that SPEED can create realistic earthquake ground-motion scenarios and that the resulting waveforms have the same peak values, duration, and frequency content as near-field earthquake recordings. The successful validation of SPEED's numerical approach is achieved through verification against an independent numerical solution, statistical assessment of ground motion, comparison with established ground-motion prediction equations, and a direct comparison of SPEED's simulation results with earthquake recordings. The results indicate that the simulated ground motions align well with empirical GMPEs. However, for Mw 7.4 events, PBSS yield higher ground-motion estimates due to directivity effects, which are amplified by the particular geometry of the Fault³⁵. Moreover, Stupazzini et al.³⁶ demonstrate that the number of simulations conducted for Istanbul is sufficient to yield stable estimates of ground motion moments. They also establish that the complexity and variability within kinematic models of slip distribution provide a realistic representation of both between- and within-event variations. Also, the validation of spatial coherency by Smerzini³⁷, along with the studies conducted by Infantino et al.²⁴ and Lin and Smerzini³⁸, collectively highlight SPEED's proficiency in accurately predicting spatial correlation structures across a broad frequency range.

Since Istanbul is close to the North Anatolian Fault, which has distinct seismic activity and pure strike-slip rupture behavior, it is a great case study for making simulation-based SGM databases. Moreover, seismic gaps along the segments of the fault, which is located south of Istanbul, have raised significant concerns regarding the occurrence of the near-field earthquake in this city. In this regard, SGM of 65 earthquake scenarios with magnitudes of 7.0, 7.2, and 7.4 and various hypocenter locations and rupture characteristics are calculated for a computational mesh of $165 \times 100 \text{ km}^2$ and 30 km depth in 7343 stations by SPEED. The spectral element model uses topography and bathymetry models (taken from <http://www.cgiarcsi.org>) as well as a shallow geology 3D numerical velocity model linked to a horizontally layered deep crustal model. The 3D shallow soil layer, which encompasses six soil classes based on Vs30 as shown in Fig. 1b, is derived from the study by Özgül³⁹. This layer is characterized by a spatially varying distribution of S and P wave velocities, both horizontally and vertically. Additionally, a viscoelastic soil model with a frequency-dependent quality factor is adopted in the simulation³⁵. The slip distribution for each scenario is generated randomly based on two kinematic rupture models: the "HB94" model by Herrero and Bernard⁴⁰ and the "CA15" model proposed by Crempien and Archuleta⁴¹. The "HB94" model generates slip distribution through concentrated asperities and employs the "k-square" model for the slip spectrum. On the other hand, the "CA15" model results in patchily distributed asperities along the rupture. For

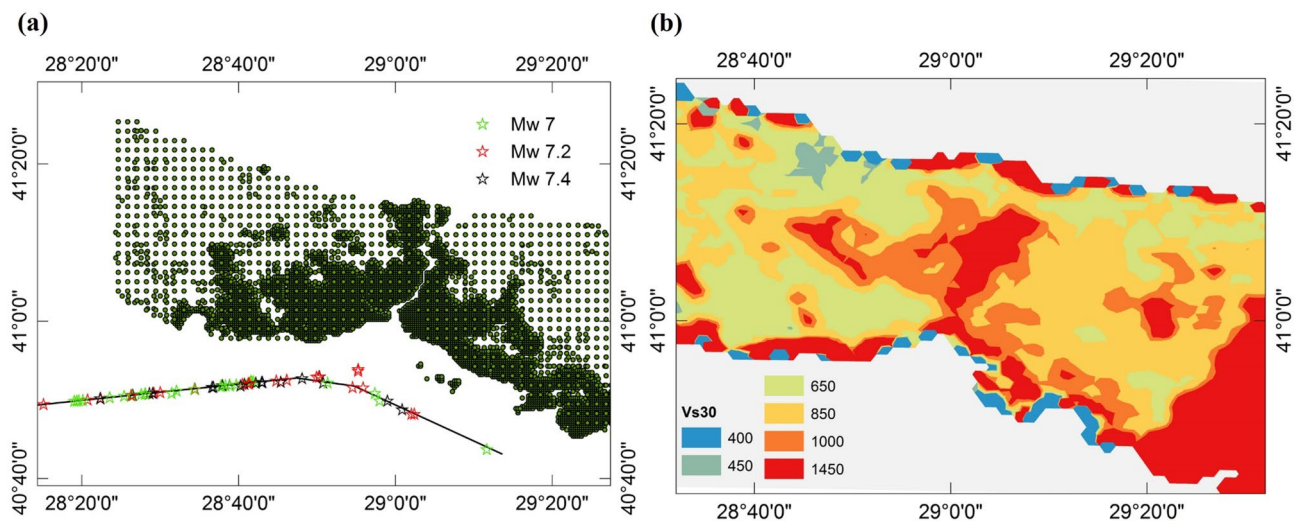


Fig. 1. (a) The stations (shaded area is representing of dense network of close stations) and epicenters (stars) of the simulated earthquake scenarios used in this study; (b) the Vs30 map of the shallow soil of the case study. These maps were created using QGIS software (version 3.12.2) available at <https://qgis.org>.

a deeper understanding and more details, readers are referred to Infantino et al.³⁵. Figure 1a depicts the locations of the strike slip fault with a dip angle of 90°, stations and epicenters of the simulated earthquake scenarios. The richness of the SPEED Istanbul catalog with verified spatial correlation structures, which is rare in a real earthquake database, explains why the authors used it to access the anisotropy and nonstationary properties of SGM spatial correlation. The data set consists of peak ground displacement, peak ground velocity and response spectrum acceleration for a broad period range⁴².

Calculation of SGM residuals

In general, the ground motion parameters caused by the i th earthquake at site j can be estimated as follows:

$$\log(IM_{ij}) = \log \overline{IM}_{(Se_i, Site_j)} + \eta_i + \varepsilon_{ij} \quad (1)$$

in which $\overline{IM}_{(Se_i, Site_j)}$ is the predicted mean value of the intensity measure for the i th scenario, which can be estimated according to existing GMPEs or specific regression analyses on the SGM database^{8,18}. The difference between the mean prediction and the observed or simulated intensity measure is the sum of the between-event residual indicated by η_i for the i th scenario, and the within-event residual indicated by ε_{ij} for site j from the i th scenario. It is assumed that these terms are independently distributed random variables with a normal distribution, zero mean, and standard deviations which are denoted here as τ and ϕ .

In this study, consistent with previous ones^{24,25}, an ad hoc regression relationship is preferred to prevent inconsistencies between simulation amplitudes and empirical ground motion model predictions. To take advantage of this study's huge and dense SGM database, this model is fitted to each scenario individually as follows:

$$\log \overline{IM}_{(Se_i, Site_j)} = c_1 + c_2 \log \sqrt{(R^2 + c_3^2)} + c_4 \sqrt{(R^2 + c_5^2)} + c_6 \log \left(\frac{Vs30}{760} \right) \quad (2)$$

where c_1 , c_2 , c_3 and c_4 are the regression coefficients while R is the shortest distance between the site and the rupture.

The between-event residual can be estimated by Jayaram and Baker⁴³ as:

$$\hat{\eta}_i = \frac{\sum_{j=1}^{n_j} (\eta_i + \varepsilon_{ij})}{\frac{\phi^2}{\tau^2} + n} \quad (3)$$

where n is the total number of stations for each scenario which in our case study is so significant (7343) that it seems rational to ignore the effect of $\frac{\phi^2}{\tau^2}$ in the denominator part. Therefore, the between event residuals can be reasonably defined as:

$$\hat{\eta}_i = \frac{1}{n} \sum_{j=1}^n (\eta_i + \varepsilon_{ij}) = \frac{1}{n} \sum_{j=1}^n (\log(IM_{ij}) - \log \overline{IM}_{(Rup_i, Site_j)}) \quad (4)$$

In addition, the within-event residual can be calculated as:

$$\hat{\varepsilon}_{ij} = \log(IM_{ij}) - \log \overline{IM}_{(Se_i, Site_j)} - \hat{\eta}_i \quad (5)$$

And the standard within-event residuals are defined as:

$$\tilde{\varepsilon}_{ij} = \frac{\hat{\varepsilon}_{ij}}{\phi} \quad (6)$$

Trend of residuals

Understanding the relationship between ground motion residuals and key predictors is crucial for accurate spatial correlation analysis. By examining trends in residuals relative to R (closest distance to rupture), V_{s30} , and M_w , we assess the effectiveness of the ground motion prediction model. Trends are analyzed by plotting the means and standard deviations of normalized within-event residuals, as shown in Fig. 2. The plots reveal that residuals exhibit no significant trends for various IMs. This lack of apparent trend suggests that the ad hoc GMPE of our study adequately accounts for these primary factors. Addressing potential biases in residuals enhances the accuracy of spatial correlation models.

Stationary spatial correlation model

The geostatistical approach based on semivariogram is one of the most common for developing stationary and isotropic spatial correlation models^{10,13,15,44}. The semivariogram is defined by the average dissimilarity of two random variables Z_j and Z_k as follows:

$$\gamma_{(j,k)} = \frac{1}{2} [E\{(Z_j - Z_k)^2\}] \quad (7)$$

Due to the lack of repeated data for a specific station pair, second-order stationary and isotropy assumptions have been taken into account. Therefore, the empirical semivariogram of within-event residuals of the station pair (j, k) can be estimated using the Matheron method as described in:

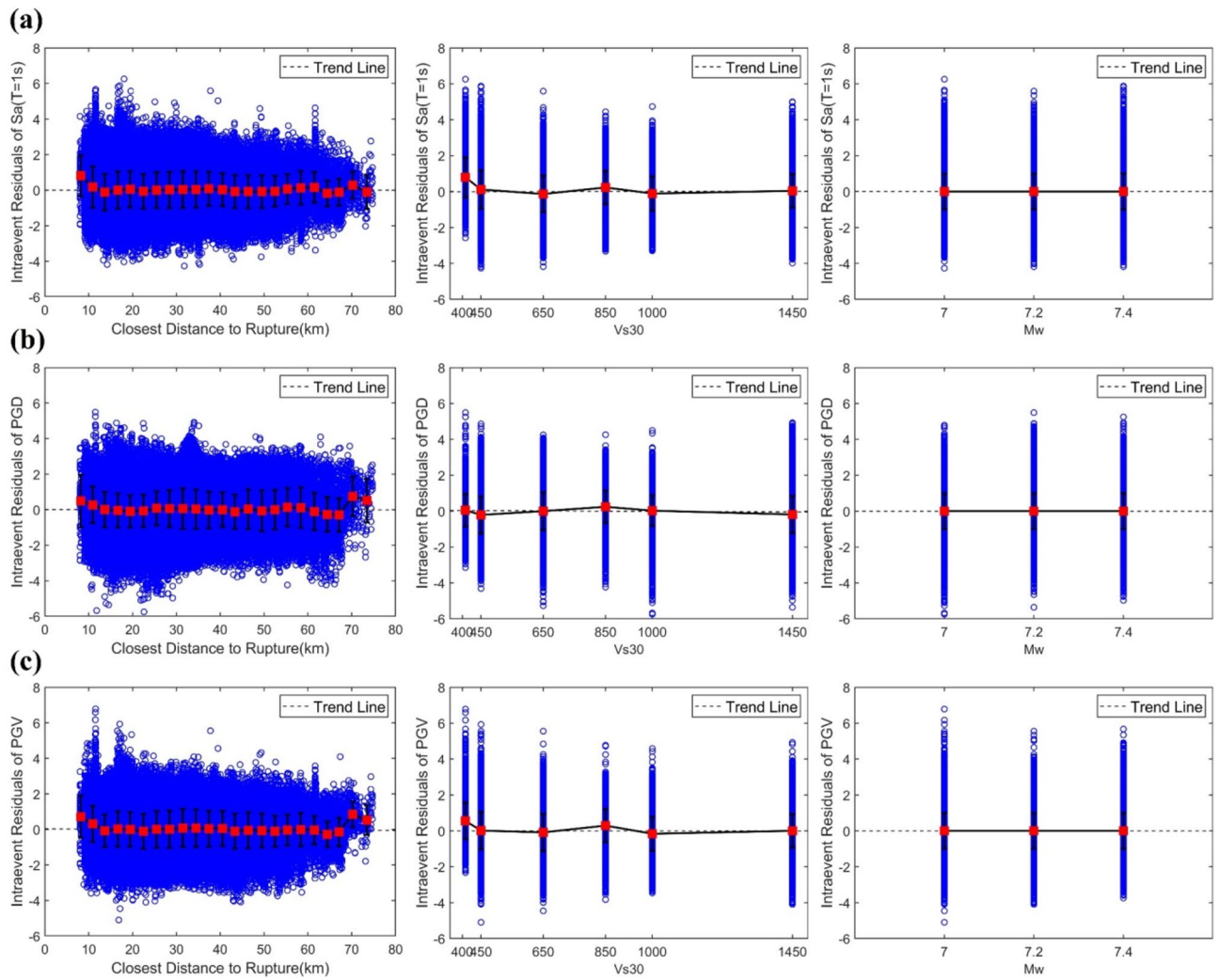


Fig. 2. The mean and standard deviation of normalized within-event residuals versus distance to rupture, V_{s30} and M_w for: (a) Sa($T = 1$ s); (b) PGD; (c) PGV.

$$\hat{\gamma}(h) = \frac{1}{|2N(h)|} \sum_{h-\frac{\Delta}{2} \ll d(j,k) \leq h+\frac{\Delta}{2}} (\tilde{\varepsilon}_{ij} - \tilde{\varepsilon}_{ik})^2 \tag{8}$$

$N(h)$ denotes the number of station pairs with the separation distance h by a user-specified tolerance (distance bin Δ) and $\tilde{\varepsilon}$ implies the normalized within-event residual. It is worthy of note that $\hat{\gamma}(h)$ only measures semi-variograms at discrete values of h . To obtain a predictive model for all values of h , a regression analysis should be conducted to fit a function (e.g., exponential, Gaussian, or Spherical) on the empirical semi-variogram data. In the current study, the common exponential function with the following form is employed:

$$\gamma(h) = a \left[1 - \exp\left(-\frac{3h}{b}\right) \right] \tag{9}$$

where a is described as the semi-variogram’s sill (variance of data), which here is $a = 1$, due to the standardization of residuals by Eq. (6) for pooling the data of all records and better comparison between different earthquakes. The semi-variogram’s range, the separation beyond which the residuals can be regarded as independent (where $\hat{\gamma}(h)$ reaches $0.95a$), is denoted by b .

Baker and Chen¹⁶ reviewed different common fitting approaches and concluded that the weighted least squares fitting method (WLS) performs better with limited bias and lower estimation uncertainty, similar to our results. Therefore, the WLS method is selected as the main fitting technique in this paper, taking into account weights assigned to each distance as:

$$w = n_h \times e^{-\frac{h}{c}} \tag{10}$$

in which n_h is the number of station pairs at distance h and $c = 5$ which means the weight of the error at a distance of 0 km is 3 times larger than the weight at a distance of 5 km.

Moreover, the relationship between the semivariogram and spatial correlation could be simply delineated by¹³:

$$\gamma(h) = a[1 - \rho(h)] \quad (11)$$

To implement spatial correlation in seismic hazard assessment, we first calculate the correlation matrix for intra-event residuals at the desired points using the predictive correlation models outlined in Eq. (10). We then simulate the correlated random field of intra-event residuals through methods such as Cholesky decomposition by a multivariate normal distribution with a zero mean and the derived correlation matrix. As described in Eq. (2), by multiplying the simulated intra-event residuals by their standard deviation and then adding the inter-event residuals and the predicted mean values of the intensity measure from the GMPEs, we achieve a more accurate and realistic assessment of seismic hazards. This approach leads to better-informed decisions in earthquake preparedness and risk mitigation.

Modeling

As mentioned in previous sections, the initial step to developing correlation models for this study involves calculating the normalized residuals for the geometric mean of the horizontal spectrum acceleration components at 18 different periods, ranging from 0 to 5 s. For each scenario, experimental semivariograms are estimated using Eq. (8). These semivariograms are then fitted with the exponential functions described in Eq. (9) using the WLS method. The weights for the WLS method are estimated from Eq. (11). The correlation ranges of all scenarios are calculated as the unique explanatory variables of the fitted models. It is important to note that the sill parameter equals 1 due to the previous normalization of residuals. Figure 3a presents a graphical representation of the geometric mean and standard deviation of the spatial correlation ranges for all scenarios plotted against different periods. Consistent with previous research by Schiappapietra and Douglas¹⁵. A positive overall trend is observed between the correlation ranges and the response spectrum periods. However, no specific trend can be identified for periods less than 1 s, and the standard deviation of the correlation range decreases for longer periods. This can be explained by the influence of local soil conditions and path characteristics on short-period waves. The notable gap between the upper and lower bounds of the spatial correlation range in Fig. 3a, indicating significant event-to-event variability in the estimated correlation range at all IMs, suggests that spatial correlation is influenced not only by model and estimation uncertainty but also by source, path, and site effects. However, identifying these effects using empirical data is challenging due to data scarcity. To address this, we use simulated SGM to explore the source, site, and path impacts on the spatial correlation, aiming to explain the observed variability in correlation ranges. Figure 3b presents the coefficient of variation (CV) of correlation ranges over all scenarios. The CV is a statistical measure that expresses the degree of variation or dispersion in a set of data relative to its mean. It is calculated by dividing the standard deviation of the data by its mean, as shown in Eq. (12). The resulting value is generally expressed as a percentage and provides a standardized measure of variability that can be used to compare the variability of different data sets with different means. A higher CV indicates a greater degree of variability relative to the mean, while a lower CV indicates less variability.

$$CV = \frac{\mu}{\sigma} \quad (12)$$

Based on Fig. 3b, the CV is approximately 40% for peak ground displacement (PGD), peak ground velocity (PGV), and long-period spectrum acceleration, while it is around 60% for short-period accelerations. Neglecting this considerable variability can implement biases in regional seismic hazard and risk assessments, as highlighted

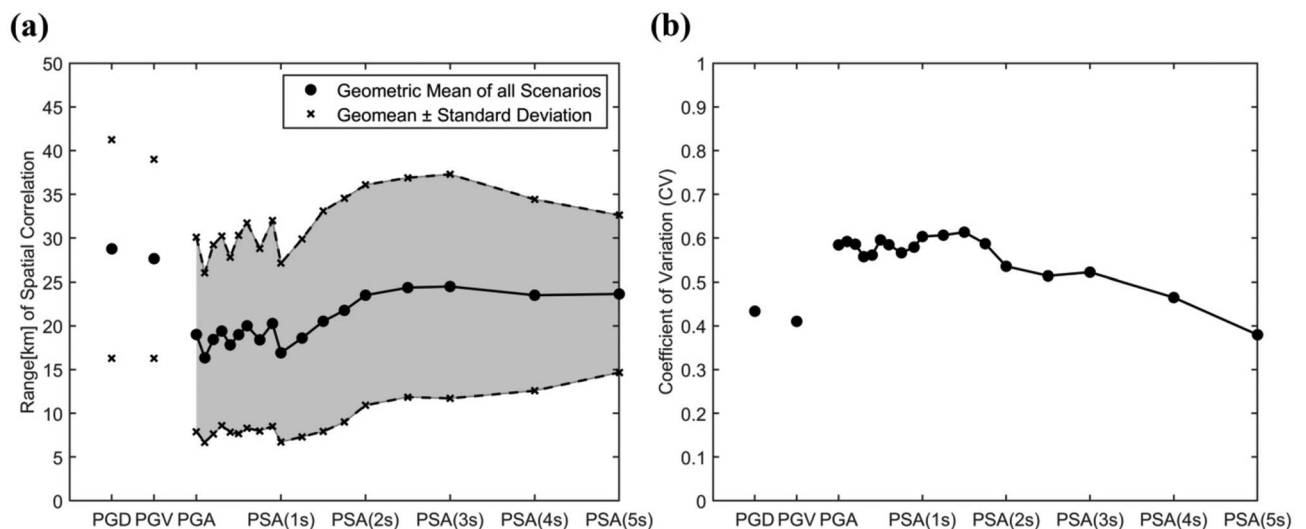


Fig. 3. (a) The mean and standard deviation of correlation ranges for PGD, PGV, PGA and PSA with periods of 0.1 s to 5 s. (b) The coefficient of variation of correlation ranges over 65 earthquake scenarios.

by Heresi and Miranda¹¹, who noted the pronounced event-to-event variability in spatial correlation model parameters for 39 well-recorded earthquakes.

In order to investigate the anisotropy and nonstationarity of the spatial correlations within our random field, we conducted variogram calculations at various orientations using a 10-degree angular bin as well as at different separation distances using a 2-km distance bin. These results were then utilized to create smoothed variogram maps. The variogram surface allows us to gain insights into the nonstationarity and anisotropy of the random field by examining how the variogram changes across different distances and directions. Ideally, in cases where stationarity and isotropy are valid, the variogram surface should display a smooth and symmetric pattern, with variogram values increasing gradually as the lag distance increases until reaching a plateau or upper limit. Moreover, a symmetric structure around the origin signifies that variogram values will be similar in all directions. On the other hand, nonstationarity refers to variations in the statistical properties, such as mean and variance, of the random field across space. These variations can manifest as alterations in the shape or range of the variogram in different areas. As depicted in Fig. 4, the presence of a non-concentric, non-symmetric variogram surface with multiple distinct peaks suggests that different subregions have various spatial structures. Consequently, this indicates a varying spatial correlation throughout the space. Furthermore, when the variogram varies across different directions, it implies that stations exhibit a higher level of correlation along certain orientations. Our findings for all records clearly illustrate the nonstationary and anisotropic correlation of our random field. Figure 4 illustrates three examples of them. Consequently, our next aim is to investigate the underlying reasons for the considerable variations between records as well as the nonstationary and anisotropic pattern of spatial correlations in different aspects of the three categories, including source, path, and site effects. This research is a crucial step in the development of more complex nonstationary and anisotropic spatial correlation models.

Source effect

The magnitude and source model are two main characteristics of an earthquake. However, it is challenging to make generalizations about the relationship between spatial correlation and magnitude because other factors, such as slip distribution and the relative position of the rupture area, may have more deterministic roles and counteract the effect of magnitude. Our results depicted in Fig. 5c reveal no specific trend between correlation range and magnitude, which is consistent with findings from previous studies^{11,13,25}. On the other hand, the source model can significantly influence the propagation of seismic waves in the Earth's subsurface. Variations in slip distribution and rupture characteristics can result in different wave propagation paths, scattering, and interactions with geological structures, leading to variations in ground motion characteristics and correlation ranges. Previous studies by Stafford et al.²², Infantino et al.²⁴, and Schiappapietra and Smerzini²⁵ have also highlighted the impact of source characteristics.

To investigate the correlation dependency on the rupture model, we categorized earthquake scenarios based on the rupture model and plotted the average correlation range for each category as a function of period in Fig. 5d. As mentioned earlier, two kinematic rupture models referred to as "HB94" and "CA15" were considered by SPEED to simulate the SGM database. The "HB94" model produces the slip distribution by modeling concentrated asperity and the "k-square" model for the slip spectrum, while the "CA15" model disperses asperities unevenly throughout the rupture (as shown in Fig. 5a,b). The scenarios simulated by the "HB94" model show higher correlation ranges of response spectrum acceleration, as well as PGV and PGD, compared to the "CA15" model. This suggests that a slip distribution that is more heterogeneous can result in a lower residual correlation. In fact, when slip occurs in multiple patches along the fault, it results in greater variation in ground motion. This variation is influenced by the contributions of slip from different patches, leading to spatially varying shaking and impacting spatial correlation. This effect is particularly pronounced in near-field earthquakes, where the distribution of slip has a greater impact on ground motion. Even slight alterations in slip can cause significant variations in ground motion at nearby locations, ultimately affecting the correlation of ground motion.

Path and site effect

As mentioned earlier, the semivariogram method is used to measure spatial correlation under the assumptions of isotropy and stationarity. However, in this study, considering the size of our SGM database, we have the opportunity to relax these assumptions and explore factors beyond spatial proximity, such as path and site characteristics.

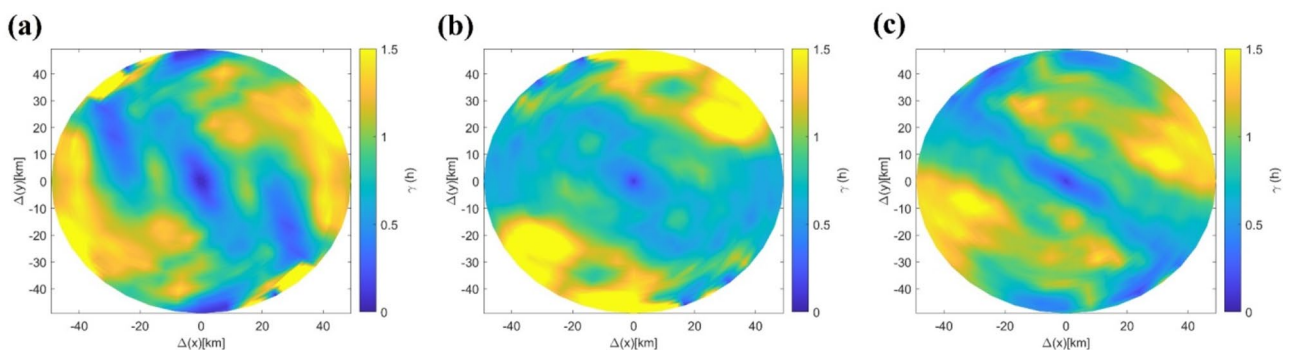


Fig. 4. Smoothed variogram surface for within-event residuals of (a) PGD, Mw=7, Scenario No. 3 (b) PGV, Mw=7.2, Scenario No. 2. (c) Sa(T=1 s), Mw=7, Scenario No. 1

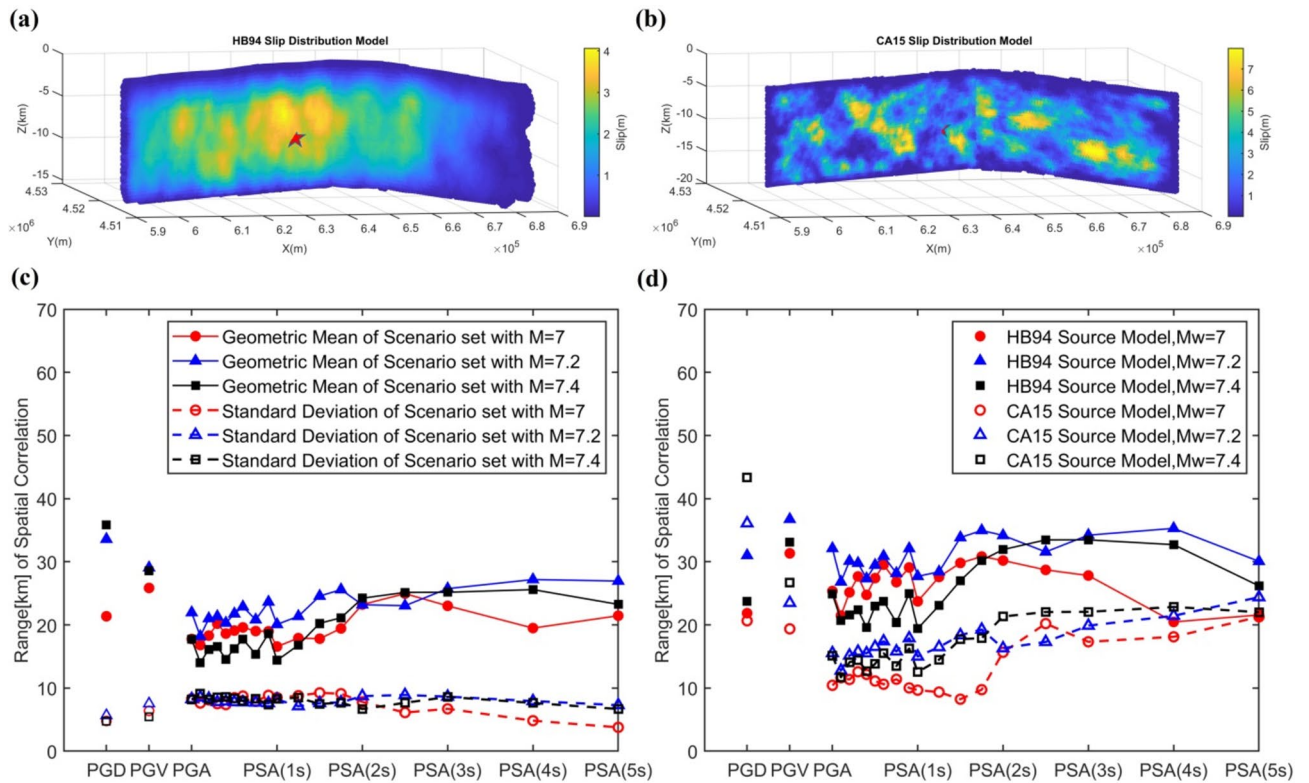


Fig. 5. (a) HB94 slip distribution model. (b) CA15 slip distribution model (c) the mean and deviation of correlation range for PGD, PGV, PGA, and PSA of earthquake scenarios classified by magnitude (d) the mean and deviation of the correlation range for PGD, PGV, PGA, and PSA of earthquake scenarios classified by source model and magnitude.

To achieve this, it is useful to compute site-specific correlations that quantify correlation in a non-stationary and anisotropic manner for each individual station pair. In previous studies^{24,26,27,45}, the common practice was to use Pearson correlation to measure the nonstationary correlation. The Pearson correlation coefficient measures the strength and direction of the relationship between two variables, assuming a normal data distribution and linear dependence. Although these assumptions may not hold in spatial data, utilizing the coefficient can still provide valuable insights by revealing influential factors and how they affect the spatial correlation of within-event residuals. This coefficient quantifies the correlation between the residuals of each station pair (j, k) as follows:

$$P(\epsilon_j, \epsilon_k) = \frac{\sum_{i=1}^n \tilde{\epsilon}_{ij} \tilde{\epsilon}_{ik}}{\sqrt{\sum_{i=1}^n \tilde{\epsilon}_{ij}^2} \sqrt{\sum_{i=1}^n \tilde{\epsilon}_{ik}^2}} \tag{13}$$

where $\tilde{\epsilon}_{ij}$ and $\tilde{\epsilon}_{ik}$ indicate the within-event residuals of the i th scenario at sites j and k , and n is the number of earthquake scenarios that have generated records at sites j and k .

However, we opted to utilize the Spearman correlation method in this study. The Spearman correlation determines the correlation between two variables by calculating the Pearson correlation of their rank values. Unlike Pearson’s correlation, which assesses linear relationships, Spearman’s correlation evaluates monotonic (nonlinear) relationships, regardless of whether they are linear or not. This method is non-parametric, meaning it does not assume a specific data distribution, and it is robust to outliers, allowing for reliable measurements even with extreme observations. By considering potential nonlinear relationships between the within-event residuals of IMs at different stations, the Spearman correlation, formulated as the following equation, is more suitable for our analysis.

$$r_s = P(R(\epsilon_j), R(\epsilon_k)) = \frac{cov(R(\epsilon_j), R(\epsilon_k))}{\sigma_{R(\epsilon_j)} \sigma_{R(\epsilon_k)}} \tag{14}$$

where $cov(R(\epsilon_j), R(\epsilon_k))$ is the covariance of the rank variables and $\sigma_{R(\epsilon_j)}$ and $\sigma_{R(\epsilon_k)}$ are the standard deviations of them. In our study, we aimed to identify the spatial correlation between all pairs of stations by computing correlation coefficients. This was made possible due to the availability of multiple simulated ground motions at each station. In order to predict spatial correlation based solely on separation distance (h), we fit commonly used exponential models as follows:

$$\rho(h) = \exp\left(-\frac{3h}{b}\right) \quad (15)$$

The resulting models are presented in Fig. 6, where they are compared to other well-known models. The spatial correlation model of simulated ground motions generated by SPEED code had the same decay pattern as the empirical models^{8,9,11}, particularly the ones proposed by Jayaram and Baker¹³ and Wagener et al.¹⁰. It is worth noting that Wagener et al.¹⁰ developed a spatial correlation model for Istanbul (similar to our study area) using SGM data and the dense network of the Istanbul Rapid Response and Early Warning System.

In this particular study, the relative distance and direction to the rupture and the soil properties of two sites have been identified as potential factors that could affect the spatial correlation. By assessing these factors, it may be possible to better understand how they contribute to the overall correlation observed in the IMs.

Distance to rupture dependency

The shortest distance from the site to the rupture is a common measure of source-to-site distance in all recent GMPE models. To assess the dependency of the spatial correlation of the station pair, we need to consider the distance of both stations from the rupture. There are several methods to have an average distance⁴⁶. In this study, considering the finite fault model, the harmonic mean of the shortest distances between the station pair and the rupture plane is selected as the distance-to-rupture measure of each station pair as follows:

$$\bar{R}(j,k) = \frac{2}{\frac{1}{R_j} + \frac{1}{R_k}} \quad (16)$$

We categorized station pairs according to their mean distance to rupture (8 km bins) in order to examine how the correlation varies with the distance of stations from the rupture. For each class, we then develop a correlation model by fitting the exponential function of Eq. (15). The findings for SA(3 s) is presented as an examples of results in Fig. 7a. Our results imply that in general, station pairs located in close proximity to the earthquake rupture exhibit lower spatial correlation when compared to those situated at greater distances. This holds true for various types of intensity measures, including PGD, PGV, and SA. This phenomenon can be explained by the significant influence of near-source earthquake characteristics and source effects on ground motion variability at stations that are nearer to the rupture. Vyas et al.⁴⁷ reveal that the variability of within-event residuals of PGV in near-field earthquakes is higher at a closer distance to the strike-slip rupture. Also, the rupture complexity is more effective at nearer stations to the rupture. For instance, Ripperger et al.⁴⁸ found that stress heterogeneity-induced variations in SGM are more pronounced near the rupture. Additionally, considering the finite fault model in this study, waves propagating from different parts of the fault have different radiation paths and incident angles, which lead to variations in the ground motion of the stations near the rupture. However, farther away from the earthquake source, ground motions become less influenced by the characteristics of the finite fault model of rupture, and waves propagating from the fault tend to reach nearby stations almost simultaneously, resulting in increased spatial correlation among them. Moreover, the spatial correlation range is influenced by the distance between station pairs and the rupture. Plotting the correlation range's CV for each IM in Fig. 7b reveals that the range's CV is roughly 30% regarding the variation of station pairs distance from the rupture.

Anisotropy

The objective of this section is to investigate the anisotropic properties of spatial correlation. To achieve this, our research delves into the relationship between spatial correlation and the orientation of the separation distance vector. We compute and plot the average correlation coefficients of station pairs that have similar directions

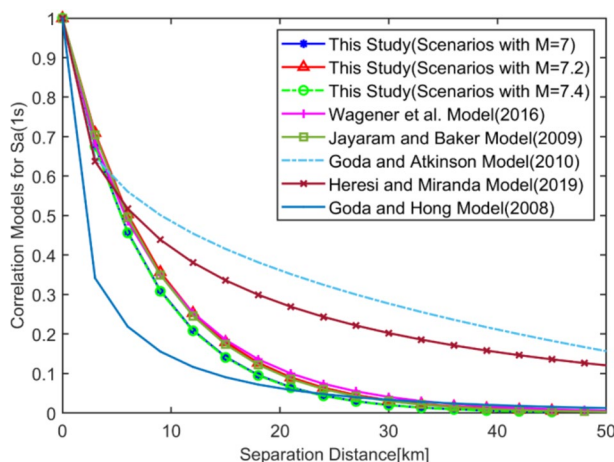


Fig. 6. Comparison of spatial correlation models from this study for Sa(T = 1 s) with existing empirical models.

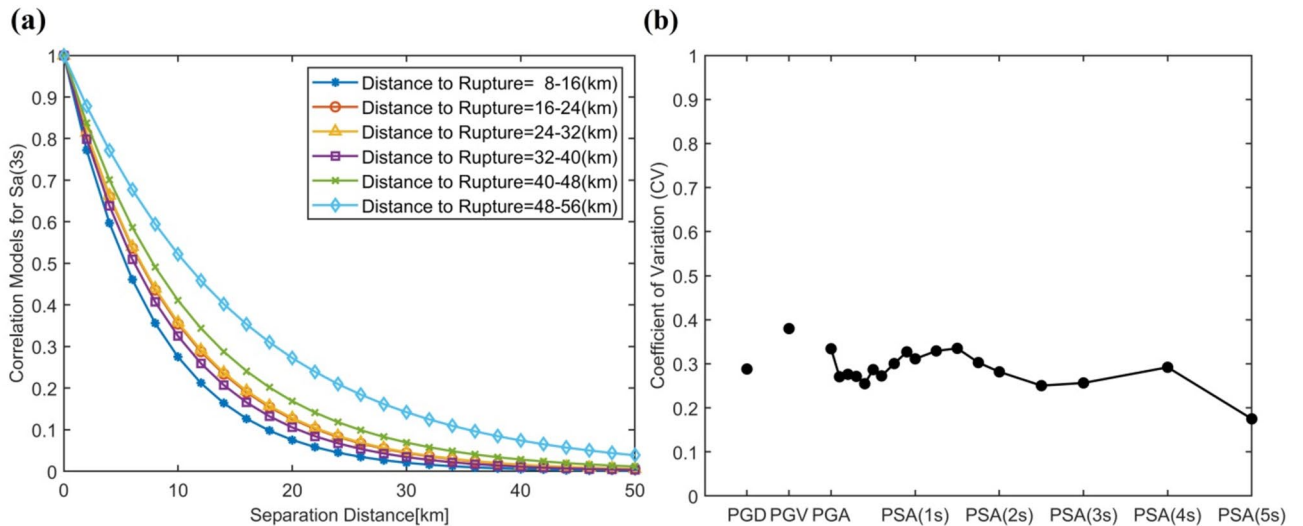


Fig. 7. (a) Spatial correlation models of station pairs with various mean distances to the rupture using Mw = 7.4 scenarios for within-event residuals of Sa(T = 3 s); (b) the coefficient of variation of correlation ranges regarding the variation of station pairs distance from the rupture.

against the azimuth of station pairs, which is the angle formed between their separation distance vector and the north direction. For example, Fig. 8 illustrates polar plots displaying the average correlation coefficients of station pairs classified based on direction bins ($\Delta\theta = 10^\circ$ intervals) and separation distance (10 km intervals) for PGD, PGV, PGA, Sa(1 s), and Sa(3 s). Furthermore, Fig. 8f displays the mean separation distances of station pairs within each angular bin. The consistent separation distance across all angular bins serves as confirmation that the variation in correlation is not due to differences in the scalar values of separation distances. Our findings indicate systematic variations in average correlation coefficients based on the direction of station pairs. The most possible explanation for this observed anisotropy lies in the dependence of spatial correlation on the relative positioning of stations in relation to the earthquake rupture. This finding aligns with Sheng et al.⁴⁹'s research, which underscores the impact of seismic wave direction on spatial correlation patterns. Their work highlights the strong influence of seismic energy incidence angles on waveform similarity. Moreover, they have also noted

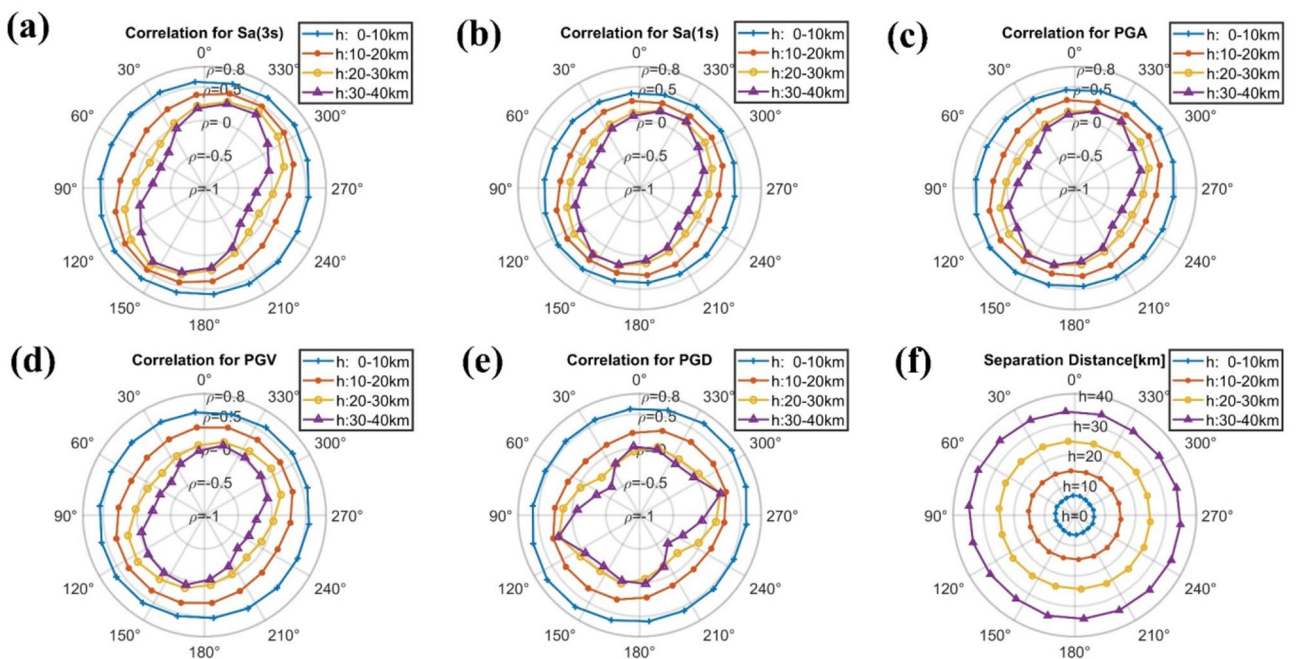


Fig. 8. The mean correlation coefficients versus the azimuth of separation distance vector using Mw = 7.4 scenarios for within-event residuals of: (a) Sa(T = 3 s); (b) Sa(T = 1 s); (c) PGA; (d) PGV; (e) PGD; (f) the mean separation distances of station pairs associated with different direction bins.

that the interaction between seismic waves and local geology, often referred to as site effects, varies with the direction of propagated seismic waves.

In this study, based on the location of the earthquake rupture, we can conclude that station pairs positioned perpendicular to the rupture direction exhibit stronger correlations. When station pairs are aligned perpendicularly to the rupture direction, they effectively occupy the same portion of the rupture area, resulting in more similar wave propagation directions. In the context of near-field earthquakes, this alignment enables them to capture more direct, coherent, and synchronized seismic waves originating from the seismic source. Consequently, the spatial correlation of intra-event residuals for such station pairs tends to be higher. Alternatively, when station pairs align parallel to the rupture direction, they are situated along distinct segments of the rupture. This alignment can lead to variations in wave propagation paths and directivity effects, resulting in greater variability in recorded ground motions. This observation is consistent with the findings of Vyas et al.⁴⁷ in their study of near-field earthquakes with strike-slip ruptures, where they noted higher variations in ground motion along the rupture-propagation direction and lower variations perpendicular to it. Additionally, studies by Chen and Baker²³ and Infantino et al.²⁴ reported higher spatial correlations of within-event residuals for stations located perpendicular to the fault.

Additionally, we categorized station pairs based on their azimuth angles to compare the correlation among station pairs with different azimuths. For each angular bin, we developed a correlation model by fitting the exponential function from Eq. (15). The correlation models for $Sa(T = 3\text{ s})$ are presented in Fig. 9a. Our findings indicate that correlation vary across different directions, and station pairs with specified separation distances exhibit diverse spatial correlations in relation to their azimuth angles. Consequently, the correlation range is also influenced by the orientation of station pairs. This observation remains consistent across all ground motion parameters (IMs). Figure 9b illustrates the CV of correlation ranges concerning variations in station pair azimuths for all IMs. It is evident that all IMs display significant variation in correlation range, typically falling within the CV range of 20% to 35%.

Site effect dependency

Soil characteristics are key factors in determining the spatial correlation of ground motions during an earthquake. Variations in soil properties, including amplification and attenuation effects, local site conditions, wave propagation speeds, damping characteristics, and frequency content, contribute to differences in ground motion intensity across locations^{50–53}. Softer soils may amplify ground motions, while stiffer soils may attenuate them, leading to variability that reduces spatial correlation. Additionally, differences in wave propagation and frequency content due to varying soil conditions further influence the correlation between sites. This section aims to investigate how soil properties impact spatial correlation by examining the relationship between site characteristics and the spatial correlation of intensity measures (IMs) in two key aspects: (1) the effect of soil dissimilarity between two stations and (2) the soil properties of each station. We consider $Vs30$ as the most well-known and accessible information about soil conditions to account for site effects.

The first outlook involves analyzing the effect of the differences between the $Vs30$ values of station pairs on spatial correlation. We categorize station pairs into five groups based on their dissimilarity of $Vs30$, ranging from 0 to more than 800 m/s, in order to compare the correlation among each group of station pairs. For each group, we develop a correlation model by fitting the exponential function from Eq. (15). The correlation models for $Sa(T = 3\text{ s})$, PGV, and PGA are presented in Fig. 10a–c. The findings indicate that as the disparity in soil's shear wave velocity between station pairs increases, there is a decrease in correlation for all IMs. Generally, station

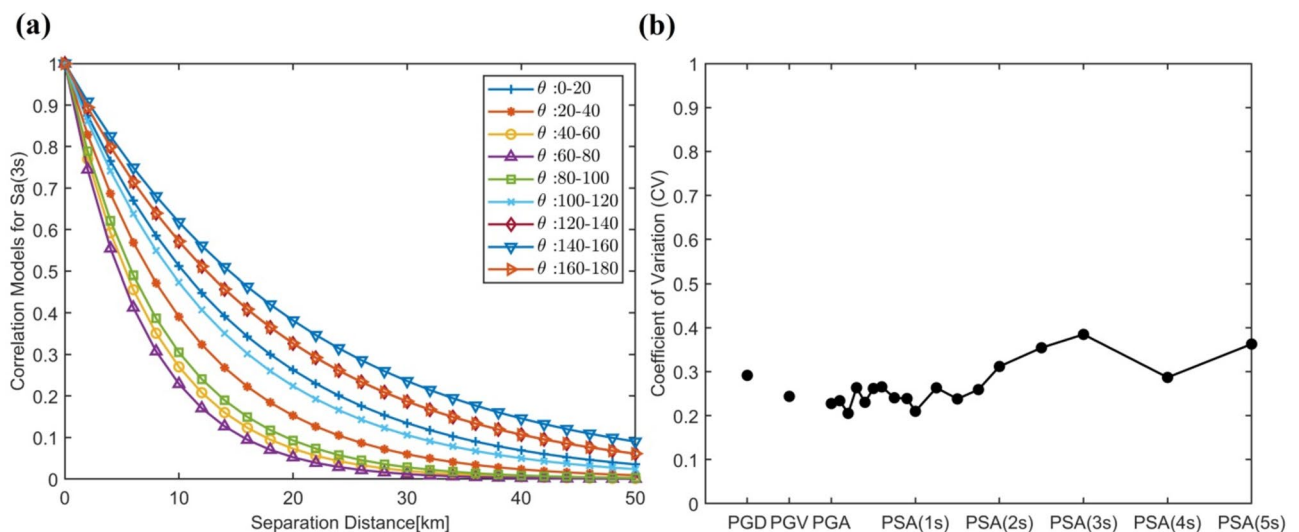


Fig. 9. (a) The spatial correlation models for the within-event residuals of $Sa(T = 1\text{ s})$ computed for station pairs in various directions as a function of separation distance using $M_w = 7.4$ scenarios; (b) the coefficient of variation of correlation ranges for different IMs using $M_w = 7.4$ scenarios.

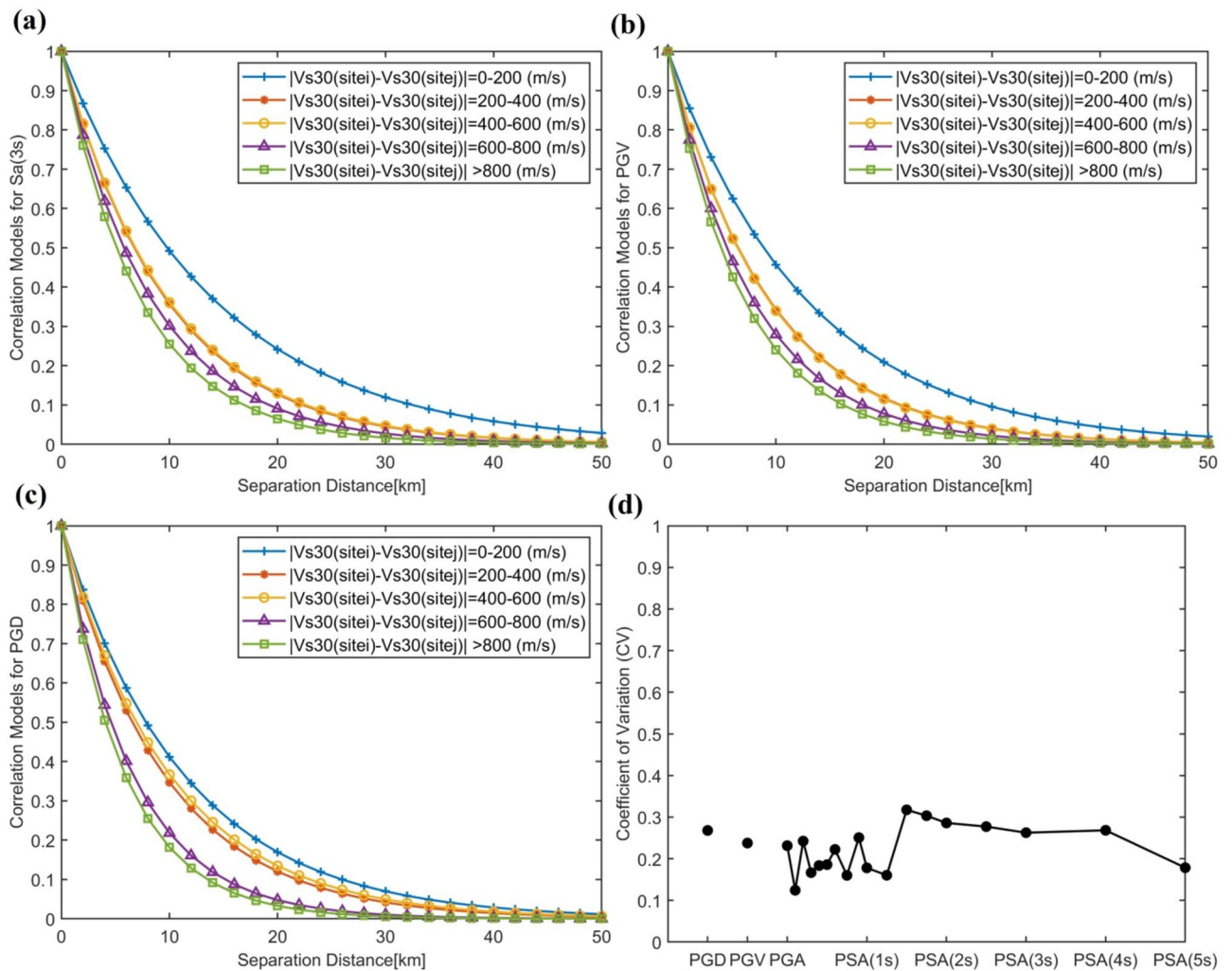


Fig. 10. Spatial correlation as a function of the differences of the soil properties of station pairs.

pairs with similar soil conditions exhibit higher spatial correlations compared to stations with different soils. Therefore, the correlation range varies depending on the heterogeneity of the soil. Figure 10d illustrates the CV of correlation ranges concerning the soil dissimilarity for all IMs, typically falling within the range of 20% to 30%. This variation in correlation can be attributed to the differences in surface and subsurface geological features at each site, resulting in variations in soil properties, such as V_{s30} , which affect how seismic waves propagate through the ground. It is important to note that locations with higher V_{s30} values generally have denser and more compact soils, which allows seismic waves to propagate faster. Conversely, lower V_{s30} values indicate less dense soil, resulting in slower wave velocities. So, when there is a greater contrast in shear wave velocity between two stations, the seismic waves velocities differ more for them. This contrast in wave velocities can lead to significant differences in wave behavior as they travel through the subsurface leading to a lower correlation between recordings. These observations highlight the importance of considering the variability in soil properties when characterizing spatial correlation in seismic hazard analysis.

The second feature explored in the study is the impact of site-specific soil properties on spatial correlation. To achieve this, station pairs were categorized into five groups based on their V_{s30} values to examine how correlation varies across different soil classes, including soft and rock soil conditions. Each group was then individually analyzed, and correlation models were developed using the exponential function (Eq. 15) for $S_a(T = 3 \text{ s})$, PGV, and PGA, as illustrated in Fig. 11a–c. The results reveal that stations located on rock exhibit stronger correlations compared to those situated on soft soil. This could be attributed to the fact that soft soils exhibit more pronounced site effects and tend to amplify and scatter seismic waves, thereby reducing the overall similarity of the recorded data. This trend holds true across all IMs. This observation is consistent with findings from Smerzini³⁷ studies on the spatial coherency of earthquakes. It is also worth noting that the range of correlation varies depending on the soil class, as depicted in Fig. 11d, with a CV typically falling within the range of 30% to 40%.

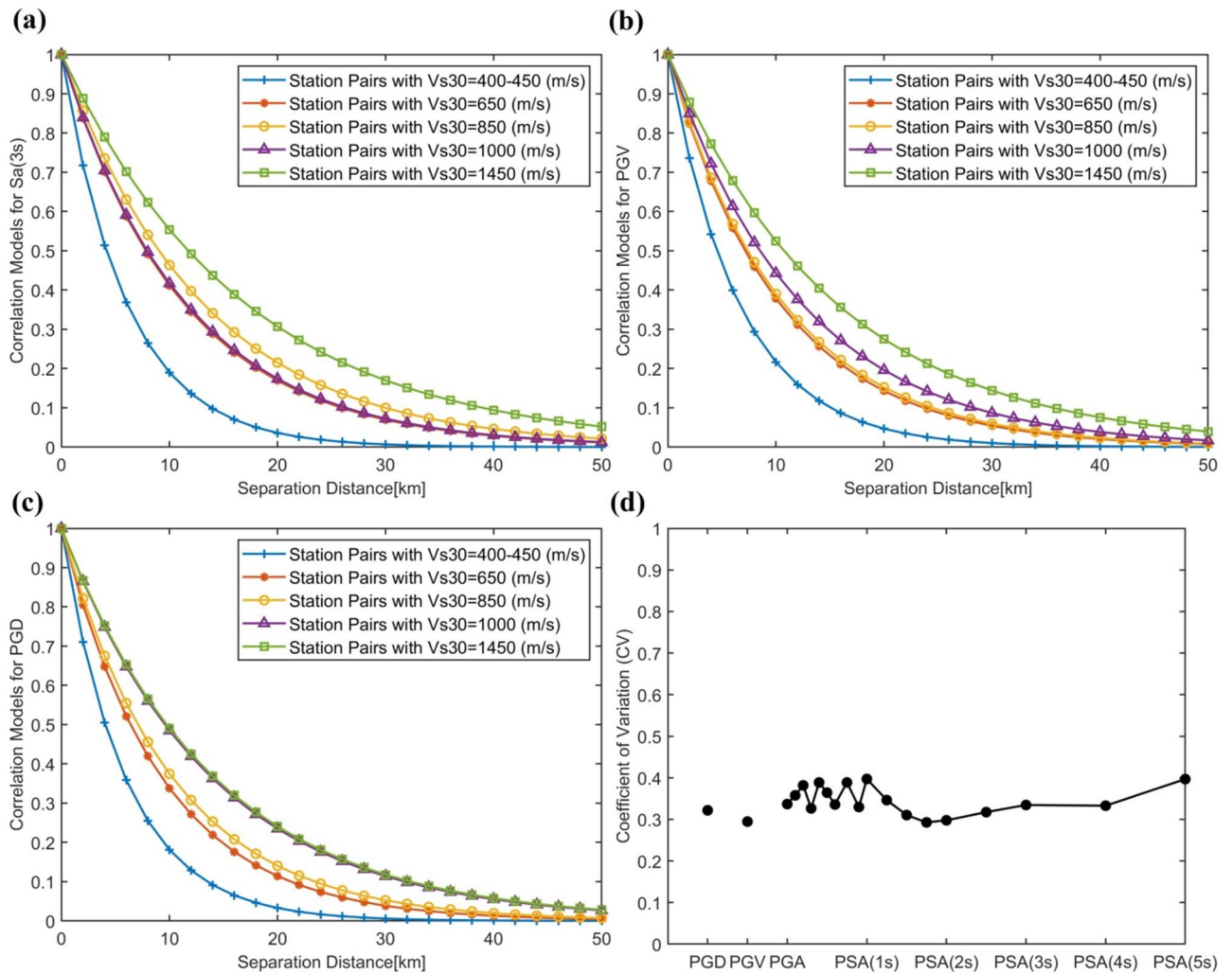


Fig. 11. Spatial correlation as a function of shear wave velocity of the soil of station pairs.

Discussion

Predictive models for spatial ground motion correlation play a crucial role in assessing seismic risk to distributed infrastructure and building portfolios, as well as in rapid post-earthquake ground motion estimation. However, due to insufficient recordings, existing spatial correlation models of ground motion rely on simplified approaches based on the assumptions of isotropy and stationarity. As highlighted by the mentioned studies in the introduction and demonstrated by this study, these assumptions may not accurately predict the spatial distribution of ground motions in space, particularly in near-source conditions. This issue is particularly important for seismic hazard assessments of regional-scale infrastructure or urban areas, where spatial correlation structures play a crucial role. Therefore, there is a growing emphasis on developing nonstationary and anisotropic correlation models to generate random fields of residuals. Recent studies, such as those by Bodenmann et al.²⁸ and Liu et al.²⁹, have shed light on this subject. To make further progress in this field, it is essential to recognize the impactful factors and gain a deeper understanding of their influence on the spatial correlation of intra-event residuals. This can be achieved through additional research conducted on dense databases of SGM in different regions.

The findings of our study offer valuable insights into the spatial correlation of ground motions. Through the utilization of a large catalog of simulated ground motions, this paper provides a comprehensive assessment of the nonstationary and anisotropic characteristics of spatial correlation. We aim to detect the underlying causes influencing a wide range of intensity measures in near-field earthquakes. Using the semivariogram approach, we developed spatial correlation models for the simulated SGMs of 65 earthquake scenarios. The results revealed that the CVs of correlation ranges for various earthquake records, which reflect the record-to-record variability, exhibited significant values of 40–60% for different IMs. Neglecting this remarkable uncertainty when performing regional seismic hazard and risk assessments may introduce bias. We have observed that the slip distribution, which serves as an indicator of the source effect, has a significant impact on the spatial correlation range. Importantly, the variogram analysis demonstrated that the residuals displayed non-stationarity and anisotropic patterns attributed to the effects of the source, path, and site. Consequently, the authors moved toward nonstationary correlation analysis to investigate the nonstationary and anisotropic properties of correlation. The Pearson correlation coefficient, which calculates the linear correlation between variables, was used in previous research to

quantify the nonstationary site-specific correlation. However, we employed the Spearman correlation approach to determine the spatial correlation coefficient for every pair of individual stations, taking into account the possibility of nonlinear correlations in spatial data. This method takes into account the nonlinear relationships between the intra-event residuals of the two stations. Through this approach, we were able to identify the major contributing factors, including the relative distance and direction of the stations to the rupture (an indicator of the path effect), as well as the soil properties that represent the site effects. This study has demonstrated that the variability of spatial correlation ranges for different intensity measures (IMs), considering various factors such as source and site characteristics, the relative location, and the direction of station pairs to the rupture, is significant and should not be neglected. The findings question the widely applied approach in seismic hazard studies, which models the spatial correlation of IMs by employing only the distance between station pairs as the explanatory variable. Taking into account additional significant parameters, as shown by this study, enables the development of more complex and precise spatial correlation models. We may greatly improve our hazard analysis's precision by adding these factors to empirical correlation prediction models.

The spatial correlation model of the simulated ground motions made by the SPEED code has the same decay pattern as the empirical models, especially the one developed for Istanbul, which is similar to the area we studied. However, it is important to acknowledge that the current ground motion simulation methods do not fully capture the variability that naturally exists in the complex physics of earthquake sources, propagation paths, and local site effects. For example, the variation of source parameters is limited in our database due to the high computational costs and difficulty in assigning a realistic probability of occurrence to various conditions. Additionally, considering such an extended computational domain, the physics-based simulations may not rely on a sufficiently detailed 3D model. Successful validation tests, however, showed that in the Istanbul case study, combining a crustal model with a thorough urban microzonation proved sufficient to provide a realistic modeling of the area. Despite being aware of these limitations, we opted to utilize the simulated database in our study due to the distinctive richness and quantity of our database. Our database, comprising a network of 7,343 stations and a minimum station separation distance of 15 m, offers a level of richness that is rarely found in real earthquake databases. This is particularly advantageous in near-field earthquakes, where nonstationary and anisotropic spatial correlation patterns are more pronounced, but obtaining real earthquake data is more challenging due to the scarcity of ground motion recordings available in close proximity to earthquake ruptures.

Conclusion

This study analyzes the spatial correlation properties of PGV, PGD, and response spectral acceleration over a broadband range of periods using 3D physics-based simulated ground motions generated by SPEED for Istanbul through both stationary semivariogram and nonstationary Spearman correlation methods. Our results show that the correlation range, a key parameter in the stationary spatial correlation model, has a positive relationship with the period of the response spectrum. However, no specific trend can be identified for periods shorter than 1 s. Moreover, the standard deviation of the correlation range decreases for longer periods due to the enhanced influence of local soil conditions on short-period waves. The coefficient of variation (CV) of correlation ranges is approximately 40% for PGD, PGV, and long-period spectrum acceleration and around 60% for short-period accelerations, highlighting significant event-to-event variability. Furthermore, the non-concentric, non-symmetric variogram maps demonstrate nonstationarity and anisotropic patterns of spatial correlation.

We examined source, path, and site effects to explain the significant variability and the nonstationary, anisotropic nature of spatial correlations. Exploring the source effect, we observed that the correlation range is not specifically correlated with magnitude but rather strongly varies depending on the rupture model. A more heterogeneous slip distribution leads to a lower spatial correlation for all intensity measures compared to a rupture model with a concentrated asperity.

Examination of path effects shows that spatial correlation decreases for station pairs closer to the rupture due to the enhanced influence of near-field earthquake characteristics in the vicinity of the rupture. The CVs for correlation ranges are around 30% based on the distances of station pairs from the rupture. Furthermore, the azimuth of station pairs relative to the rupture contributes to the anisotropy in spatial correlation. Station pairs oriented perpendicularly to the rupture exhibit a higher correlation, primarily due to their exposure to more similar wave propagation directions, with a CV of 20% to 35% for correlation ranges. The soil effect also plays a significant role in spatial correlation. Station pairs with similar soil conditions demonstrate higher correlations compared to pairs with differing soil properties, with CVs ranging from 20 to 30%. Also, stations located on rock exhibit stronger correlations compared to those on soft soil, with CVs generally between 30 and 40%.

Overall, this study enhances seismic hazard assessment by providing insights into spatial ground motion correlation, which is crucial for developing nonstationary and anisotropic models. Incorporating these factors can improve predictive accuracy, leading to a more realistic simulation of random fields of residuals. Future research should continue to explore empirical and simulated earthquake data from various regions to deepen our knowledge in this field.

Data availability

Simulated ground motion data utilized in this study was graciously provided by SPEED, through personal communication. BB-SPEEDset data is publicly accessible on their website at <https://speed.mox.polimi.it/bb-speedset>.

Received: 27 February 2024; Accepted: 5 September 2024

Published online: 11 September 2024

References

- Weatherill, G., Silva, V., Crowley, H. & Bazzurro, P. Exploring the impact of spatial correlations and uncertainties for portfolio analysis in probabilistic seismic loss estimation. *Bull. Earthq. Eng.* **13**, 957–981. <https://doi.org/10.1007/s10518-015-9730-5> (2015).
- Lee, R. & Kiremidjian, A. S. Uncertainty and correlation for loss assessment of spatially distributed systems. *Earthq. Spectra* **23**, 753–770. <https://doi.org/10.1193/1.2791001> (2007).
- Jayaram, N. & Baker, J. W. Efficient sampling and data reduction techniques for probabilistic seismic lifeline risk assessment. *Earthq. Eng. Struct. Dyn.* **39**, 1109–1131. <https://doi.org/10.1002/eqe.988> (2010).
- Sokolov, V. & Wenzel, F. Areal exceedance of ground motion as a characteristic of multiple-site seismic hazard: Sensitivity analysis. *Soil Dyn. Earthq. Eng.* **126**, 105770. <https://doi.org/10.1016/j.soildyn.2019.105770> (2019).
- Dong, Y. & Frangopol, D. M. Probabilistic assessment of an interdependent healthcare–bridge network system under seismic hazard. *Struct. Infrastruct. Eng.* **13**, 160–170. <https://doi.org/10.1080/15732479.2016.1198399> (2017).
- Du, W. & Wang, G. Fully probabilistic seismic displacement analysis of spatially distributed slopes using spatially correlated vector intensity measures. *Earthq. Eng. Struct. Dyn.* **43**, 661–679. <https://doi.org/10.1002/eqe.2365> (2014).
- Schiappapietra, E., Stripajová, S., Pažák, P., Douglas, J. & Trendafiloski, G. Exploring the impact of spatial correlations of earthquake ground motions in the catastrophe modelling process: A case study for Italy. *Bull. Earthq. Eng.* **20**, 5747–5773. <https://doi.org/10.1007/s10518-022-01413-z> (2022).
- Goda, K. & Atkinson, G. M. Intraevent spatial correlation of ground-motion parameters using SK-net data. *Bull. Seismol. Soc. Am.* **100**, 3055–3067. <https://doi.org/10.1785/0120100031> (2010).
- Goda, K. & Hong, H. Spatial correlation of peak ground motions and response spectra. *Bull. Seismol. Soc. Am.* **98**, 354–365. <https://doi.org/10.1785/0120070078> (2008).
- Wagener, T., Goda, K., Erdik, M., Daniell, J. & Wenzel, F. A spatial correlation model of peak ground acceleration and response spectra based on data of the Istanbul earthquake rapid response and early warning system. *Soil Dyn. Earthq. Eng.* **85**, 166–178. <https://doi.org/10.1016/j.soildyn.2016.03.016> (2016).
- Heresi, P. & Miranda, E. Uncertainty in intraevent spatial correlation of elastic pseudo-acceleration spectral ordinates. *Bull. Earthq. Eng.* **17**, 1099–1115. <https://doi.org/10.1007/s10518-018-0506-6> (2019).
- Du, W. & Ning, C.-L. Modeling spatial cross-correlation of multiple ground motion intensity measures (SAs, PGA, PGV, Ia, CAV, and significant durations) based on principal component and geostatistical analyses. *Earthq. Spectra* **37**, 486–504. <https://doi.org/10.1177/8755293020952442> (2021).
- Jayaram, N. & Baker, J. W. Correlation model for spatially distributed ground-motion intensities. *Earthq. Eng. Struct. Dyn.* **38**, 1687–1708. <https://doi.org/10.1002/eqe.922> (2009).
- Loth, C. & Baker, J. W. A spatial cross-correlation model of spectral accelerations at multiple periods. *Earthq. Eng. Struct. Dyn.* **42**, 397–417. <https://doi.org/10.1002/eqe.2212> (2013).
- Schiappapietra, E. & Douglas, J. Modelling the spatial correlation of earthquake ground motion: Insights from the literature, data from the 2016–2017 Central Italy earthquake sequence and ground-motion simulations. *Earth-Sci. Rev.* **203**, 103139. <https://doi.org/10.1016/j.earscirev.2020.103139> (2020).
- Baker, J. W. & Chen, Y. Ground motion spatial correlation fitting methods and estimation uncertainty. *Earthq. Eng. Struct. Dyn.* **49**, 1662–1681. <https://doi.org/10.1002/eqe.3322> (2020).
- Schiappapietra, E. & Douglas, J. Assessment of the uncertainty in spatial-correlation models for earthquake ground motion due to station layout and derivation method. *Bull. Earthq. Eng.* **19**, 5415–5438. <https://doi.org/10.1007/s10518-021-01179-w> (2021).
- Sokolov, V., Wenzel, F. & Kuo-Liang, W. Uncertainty and spatial correlation of earthquake ground motion in Taiwan. *TAO: Terrest. Atmos. Ocean. Sci.* **21**, 9. [https://doi.org/10.3319/TAO.2010.05.03.01\(T\)](https://doi.org/10.3319/TAO.2010.05.03.01(T)) (2010).
- Du, W. & Wang, G. Intra-event spatial correlations for cumulative absolute velocity, Arias intensity, and spectral accelerations based on regional site conditions. *Bull. Seismol. Soc. Am.* **103**, 1117–1129. <https://doi.org/10.1785/0120120185> (2013).
- Garakaninezhad, A. & Bastami, M. A novel spatial correlation model based on anisotropy of earthquake ground-motion intensity. *Bull. Seismol. Soc. Am.* **107**, 2809–2820. <https://doi.org/10.1785/0120160367> (2017).
- Abbasnejadfar, M., Bastami, M. & Fallah, A. Investigation of anisotropic spatial correlations of intra-event residuals of multiple earthquake intensity measures using latent dimensions method. *Geophys. J. Int.* **222**, 1449–1469. <https://doi.org/10.1093/gji/ggaa255> (2020).
- Stafford, P. J., Zurek, B. D., Ntinalexis, M. & Bommer, J. J. Extensions to the Groningen ground-motion model for seismic risk calculations: Component-to-component variability and spatial correlation. *Bull. Earthq. Eng.* **17**, 4417–4439. <https://doi.org/10.1007/s10518-018-0425-6> (2019).
- Chen, Y. & Baker, J. W. Spatial correlations in cybershake physics-based ground-motion simulations. *Bull. Seismol. Soc. Am.* **109**, 2447–2458. <https://doi.org/10.1785/0120190065> (2019).
- Infantino, M., Smerzini, C. & Lin, J. Spatial correlation of broadband ground motions from physics-based numerical simulations. *Earthq. Eng. Struct. Dyn.* **50**, 2575–2594. <https://doi.org/10.1002/eqe.3461> (2021).
- Schiappapietra, E. & Smerzini, C. Spatial correlation of broadband earthquake ground motion in Norcia (Central Italy) from physics-based simulations. *Bull. Earthq. Eng.* **1–25**. <https://doi.org/10.1007/s10518-021-01160-7> (2021).
- Chen, Y., Bradley, B. A. & Baker, J. W. Nonstationary spatial correlation in New Zealand strong ground-motion data. *Earthq. Eng. Struct. Dyn.* <https://doi.org/10.1002/eqe.3516> (2021).
- Chen, Y. & Baker, J. W. Community detection in spatial correlation graphs: Application to non-stationary ground motion modeling. *Comput. Geosci.* **154**, 104779. <https://doi.org/10.1016/j.cageo.2021.104779> (2021).
- Bodenmann, L., Baker, J. W. & Stojadinović, B. Accounting for path and site effects in spatial ground-motion correlation models using Bayesian inference. *Nat. Hazards Earth Syst. Sci. Disc.* **1–23**, 2023. <https://doi.org/10.5194/nhess-23-2387-2023> (2023).
- Liu, C., Macedo, J. & Kuehn, N. Spatial correlation of systematic effects of non-ergodic ground motion models in the Ridgecrest area. *Bull. Earthq. Eng.* **1–27**. <https://doi.org/10.1007/s10518-022-01441-9> (2022).
- Ornthammarath, T., Chua, C. T., Suppasri, A. & Foytong, P. Seismic damage and comparison of fragility functions of public and residential buildings damaged by the 2014 Mae Lao (Northern Thailand) earthquake. *Earthq. Spectra* **39**, 126–147 (2023).
- Naguit, M. et al. From source to building fragility: post-event assessment of the 2013 M7. 1 Bohol, Philippines, Earthquake. *Earthq. Spectra* **33**, 999–1027 (2017).
- Mazzieri, I., Stupazzini, M., Guidotti, R. & Smerzini, C. SPEED: SPectral Elements in Elastodynamics with Discontinuous Galerkin: A non-conforming approach for 3D multi-scale problems. *Int. J. Numer. Methods Eng.* **95**, 991–1010. <https://doi.org/10.1002/nme.4532> (2013).
- Paolucci, R. et al. Earthquake ground motion modeling of induced seismicity in the Groningen gas field. *Earthq. Eng. Struct. Dyn.* **50**, 135–154. <https://doi.org/10.1002/eqe.3367> (2021).
- Paolucci, R. et al. Broadband ground motions from 3D physics-based numerical simulations using artificial neural networks/broadband ground motions from 3D PBSSs using ANNs. *Bull. Seismol. Soc. Am.* **108**, 1272–1286. <https://doi.org/10.1785/0120170293> (2018).
- Infantino, M., Mazzieri, I., Özcebe, A. G., Paolucci, R. & Stupazzini, M. 3D physics-based numerical simulations of ground motion in Istanbul from earthquakes along the Marmara segment of the North Anatolian Fault. *Bull. Seismol. Soc. Am.* **110**, 2559–2576. <https://doi.org/10.1785/0120190235> (2020).

36. Stupazzini, M., Infantino, M., Allmann, A. & Paolucci, R. Physics-based probabilistic seismic hazard and loss assessment in large urban areas: A simplified application to Istanbul. *Earthq. Eng. Struct. Dyn.* <https://doi.org/10.1002/eqe.3365> (2020).
37. Smerzini, C. In *Proceedings of the 16th European Conference on Earthquake Engineering*. 1–12.
38. Lin, J. & Smerzini, C. Variability of physics-based simulated ground motions in Thessaloniki urban area and its implications for seismic risk assessment. *Front. Earth Sci.* **10**, 951781. <https://doi.org/10.3389/feart.2022.951781> (2022).
39. Özgül, N. Geology of Istanbul City Area (executive summary). In *Metropolitan Municipality of Istanbul, Directorate of Earthquake & Geotechnical Investigation (in Turkish)* (2011).
40. Herrero, A. & Bernard, P. A kinematic self-similar rupture process for earthquakes. *Bull. Seismol. Soc. Am.* **84**, 1216–1228. <https://doi.org/10.1785/BSSA0840041216> (1994).
41. Crempien, J. G. & Archuleta, R. J. UCSB method for simulation of broadband ground motion from kinematic earthquake sources. *Seismol. Res. Lett.* **86**, 61–67. <https://doi.org/10.1785/0220140103> (2015).
42. Paolucci, R., Smerzini, C. & Vanini, M. BB-SPEEDset: A validated dataset of broadband near-source earthquake ground motions from 3D physics-based numerical simulations. *Bull. Seismol. Soc. Am.* <https://doi.org/10.1785/0120210089> (2021).
43. Jayaram, N. & Baker, J. W. Considering spatial correlation in mixed-effects regression and the impact on ground-motion models. *Bull. Seismol. Soc. Am.* **100**, 3295–3303. <https://doi.org/10.1785/0120090366> (2010).
44. Esposito, S. & Iervolino, I. Spatial correlation of spectral acceleration in European data. *Bull. Seismol. Soc. Am.* **102**, 2781–2788. <https://doi.org/10.1785/0120120068> (2012).
45. Aldea, S., Heresi, P. & Pastén, C. Within-event spatial correlation of peak ground acceleration and spectral pseudo-acceleration ordinates in the Chilean subduction zone. *Earthq. Eng. Struct. Dyn.* **51**, 2575–2590. <https://doi.org/10.1002/eqe.3674> (2022).
46. Abbasnejadfar, M., Bastami, M. & Fallah, A. Investigating the spatial correlations in univariate random fields of peak ground velocity and peak ground displacement considering anisotropy. *Geoenviro. Disasters* **8**, 1–16. <https://doi.org/10.1186/s40677-021-00196-w> (2021).
47. Vyas, J. C., Mai, P. M. & Galis, M. Distance and azimuthal dependence of ground-motion variability for unilateral strike-slip ruptures. *Bull. Seismol. Soc. Am.* **106**, 1584–1599. <https://doi.org/10.1785/0120150298> (2016).
48. Ripperger, J., Mai, P. & Ampuero, J.-P. Variability of near-field ground motion from dynamic earthquake rupture simulations. *Bull. Seismol. Soc. Am.* **98**, 1207–1228. <https://doi.org/10.1785/0120070076> (2008).
49. Sheng, Y., Kong, Q. & Beroza, G. C. Network analysis of earthquake ground motion spatial correlation: A case study with the San Jacinto seismic nodal array. *Geophys. J. Int.* **225**, 1704–1713. <https://doi.org/10.1093/gji/ggab058> (2021).
50. Mase, L. Z. Local seismic hazard map based on the response spectra of stiff and very dense soils in Bengkulu city Indonesia. *Geodesy. Geodyn.* **13**, 573–584 (2022).
51. Somantri, A. K., Mase, L. Z., Susanto, A., Gunadi, R. & Febriansya, A. Analysis of ground response of Bandung region subsoils due to predicted earthquake triggered by Lembang Fault, West Java Province Indonesia. *Geotech. Geol. Eng.* **41**, 1155–1181 (2023).
52. Mase, L. Z., Somantri, A. K., Chaipayut, S., Febriansya, A. & Syahbana, A. J. Analysis of ground response and potential seismic damage to sites surrounding Cimandiri Fault, West Java Indonesia. *Nat. Hazards* **119**, 1273–1313 (2023).
53. Mase, L. Z. & Keawsawasvong, S. Seismic hazard maps of Bengkulu City, Indonesia, considering probabilistic spectral response for medium and stiff soils. *Open Civ. Eng. J.* **16**, e187414952210210 (2022).

Author contributions

M.Z: supervision of the project, Conceptualization and design the study, review and editing of manuscript. M.F: conceptualization and design the study, data collection and analysis, interpretation and visualization of results, software, drafting and editing the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to M.F.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024