Yu, B., and Baker, J. W. (2026). "Cross-period ground motion spatial correlation with path and site effects: Bayesian inference and risk implications." *Earthquake Engineering & Structural Dynamics*. `https://doi.org/10.1002/eqe.70269`

# Cross-period ground motion spatial correlation with path and site effects: Bayesian inference and risk implications

Bofan Yu[1] and Jack W. Baker[2]

[1]Department of Civil and Environmental Engineering, Stanford University, Stanford, California, U.S.A. Email: ybfq999@stanford.edu

[2]Department of Civil and Environmental Engineering, Stanford University, Stanford, California, U.S.A.

## ABSTRACT

Accurately characterizing cross-period spatial correlation of ground motion intensity measures is essential for regional earthquake risk assessment. We propose a model that represents multi-period within-event residuals via a principal-component decomposition and, for each latent component, specifies a spatial kernel that separates path and site influences through epicenter-to-station azimuth differences and $V_{s,30}$ contrasts. Parameters are inferred in a Bayesian framework, yielding full posterior distributions, and thus uncertainty quantification. Posterior predictive checks and log pointwise predictive density (LPPD) demonstrate consistent improvements over isotropic baselines. Two synthetic regional case studies illustrate risk implications: in a near-epicenter setting with spatially heterogeneous $V_{s,30}$, traditional isotropic models over-aggregate local extremes and produce heavier-tailed scenario loss-exceedance curves, whereas the proposed model reduces tail inflation; in a far-epicenter setting with relatively uniform $V_{s,30}$, the exceedance curves are nearly indistinguishable, with minor corrections to the baseline's slight underestimation. By jointly capturing cross-period structure and distinct path/site effects, the formulation mitigates these biases and delivers improved portfolio-loss estimates while remaining compatible with standard scenario-based and probabilistic regional portfolio-loss workflows.

**Keywords:** Spatial correlation; Cross-period correlation; Principal component analysis; Bayesian inference; Seismic risk assessment

## INTRODUCTION

Regional seismic risk assessment relies on realistic simulation of spatially distributed ground motion fields across an area of interest, in which spatial correlation plays a central role [1–4].

Most existing spatial correlation models characterize dependence primarily as a function of inter-station separation distance, typically under assumptions of stationarity and isotropy e.g.,[5–11]. Owing to their simplicity and computational efficiency, these distance-based isotropic formulations have been widely adopted in large-scale applications and provide a practical approximation. However, recent studies have highlighted limitations of purely distance-based representations [12–16]. In particular, path effects associated with wave propagation from the rupture to sites, together with site-related effects, can introduce spatial patterns that are not fully captured by inter-station distance alone. When these mechanisms are neglected, the resulting correlation structure may be biased, with implications for regional hazard and risk calculations. Recent modeling work [16] has shown that explicitly accounting for path- and site-related components can improve the representation of spatial dependence. These results motivate the development of spatial correlation models that move beyond isotropic, distance-only assumptions.

A further challenge arises when regional analyses require joint simulation of ground motions across multiple spectral periods. Cross-period correlation is critical when different assets or response measures are governed by different periods. Several approaches have been proposed to address this problem e.g.,[6;17–20]. Goda and Hong [6] explicitly consider cross-period spatial correlation of pseudo spectral acceleration (PSA) residuals among the early studies, proposing a simple separable approximation that combines period-to-period dependence with distance-based spatial correlation. In contrast, multivariate geostatistical formulations such as the linear model of coregionalization (LMC) can provide admissible cross-period and cross-site covariance structures [17]. However, these models can be computationally demanding when constructing and factorizing full covariance matrices for a large number of sites and periods.

Markhvida et al. [20] proposed an alternative representation of cross-period dependence based on principal components (PCs), in which cross-correlated spectral ordinates are expressed through uncorrelated components that can be spatially modeled independently and subsequently reconstructed in the period domain. This framework provides a scalable and practical approximation for modeling cross-period correlation in large regional simulations. Building on this concept, the present study adopts a principal component analysis (PCA)-based decomposition and integrates it with the spatial correlation kernel with path- and site-related components proposed by Bodenmann et al.[16] By formulating the model in this manner, the proposed framework retains a clear and parsimonious representation of spatial dependence while remaining computationally efficient for large regional applications. In addition, a Bayesian inference framework is used to quantify posterior uncertainty in the correlation parameters, which is then propagated into multi-period spatial simulations and regional risk analyses.

To illustrate the implications of the proposed modeling framework, the calibrated correlation models are applied in regional simulations for two representative earthquake scenarios. These simulations compare risk estimates obtained from correlation models that explicitly account for path- and site-related effects with those based on conventional isotropic formulations, focusing on resulting differences in regional loss metrics. Furthermore, posterior distributions of the correlation parameters are propagated through the simulation framework to examine the variability of risk estimates associated with parameter uncertainty in a consistent manner.

## DATA & PREPROCESSING

This study analyzes spatial correlations of ground motion residuals across multiple periods. Records are drawn from the NGA–West2 database[21;22]. The intensity measure is 5%–damped spectral acceleration $SA(T)$ using the rotation-dependent 50th percentile (RotD50) horizontal component. We consider nine oscillator periods $T$ (s) $= \{0.01, 0.03, 0.06, 0.1, 0.3, 0.6, 1, 3, 6\}$. To ensure data quality across all periods, we apply a series of event and station filters following the criteria in[16]:

(1) site shear-wave velocity $V_{s,30}$ (m/s) $\in [180, 760]$;
(2) closest rupture distance $R_{\text{rup}} < 300$ km;
(3) longest-usable period $T_{\text{LUP}} \geq 6$ s so that all target periods are usable;
(4) each event retains at least 40 recordings satisfying the above conditions.

After filtering, we retain 5,555 recordings from 41 earthquakes.

Median and standard deviation predictions for each station are obtained from the Chiou and Youngs[23] model (CY14) , and within–event residuals are computed using a mixed-effects formulation, as below[24]:

$$\ln SA_{ij}(T) \; = \; \mu_{\ln SA}\big(T; rup_{ij},\, site_i\big) \; + \; \delta B_j(T) \; + \; \delta W_{ij}(T), \tag{1}$$

where $i$ and $j$ index recording stations and earthquake events, respectively. $\mu_{\ln SA}\big(T; rup_{ij},\, site_i\big)$ denotes the predicted mean of the logarithmic $SA$ at period $T$, as given by the CY14 ground motion model, conditional on rupture ($rup$) parameters (e.g., magnitude and distance) and site ($site$) parameters (e.g., $V_{S,30}$). The terms $\delta B_j(T)$ and $\delta W_{ij}(T)$ represent the between-event and within-event residuals, respectively. The ground motion model also specifies the standard deviations of $\delta B_j(T)$ and $\delta W_{ij}(T)$, denoted by $\tau_j(T)$ and $\phi_{ij}(T)$ respectively. Based on Jayaram and Baker[25], the joint distribution of $\delta W_{ij}(T)$ for recordings within the same event $j$ is assumed to follow a zero-mean multivariate normal distribution.

To enable cross-period comparison and simulation within a unit-variance multivariate normal framework, the within-event residuals are standardized by scaling by the within-event standard deviation.

$$\delta \bar{W}_{ij}(T) \; \equiv \; \frac{\delta W_{ij}(T)}{\phi_{ij}(T)}, \tag{2}$$

Here, $phi_{ij}(T)$ is the within-event residuals standard deviation predicted by the CY14 variability model. As a consistency check, we computed the sample standard deviations of the scaled within-event residuals for the selected records at the nine oscillator periods. These values range from 1.01 to 1.05, indicating that the empirical variability of the selected subset is broadly consistent with the CY14 within-event sigma model.

To construct the spatial features, for each event, we form all station pairs and compute three distance metrics: (1) the Euclidean inter–station distance $d_E$ (km); (2) the angular separation $d_A$ (degrees) between the epicenter–to–station rays representing path effects; and (3) the site–condition contrast $d_S$, representing differences in near–surface site conditions, as shown in Figure 1.

## METHODOLOGY

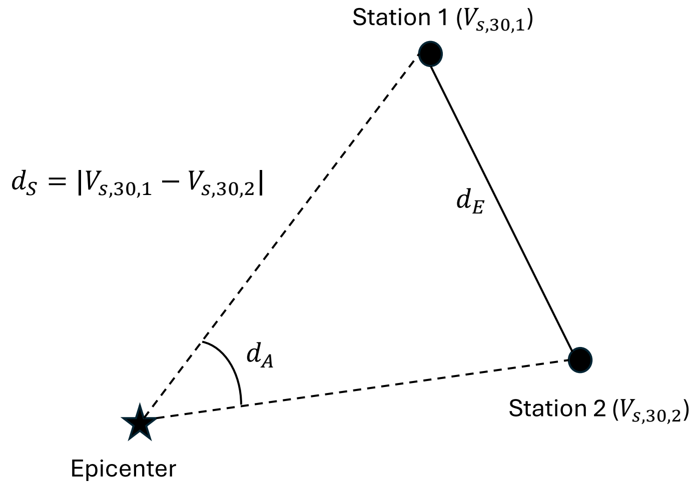


**Fig. 1.** Schematic illustration of pairwise features used in the Euclidean–Azimuth–Soil (EAS) correlation model. The epicenter is marked by a star and two recording sites are shown as circles. Euclidean inter–station distance $d_E$, epicentral–azimuth difference $d_A$, and site–condition contrast $d_S = |V_{s,30,1} - V_{s,30,2}|$ are indicated.

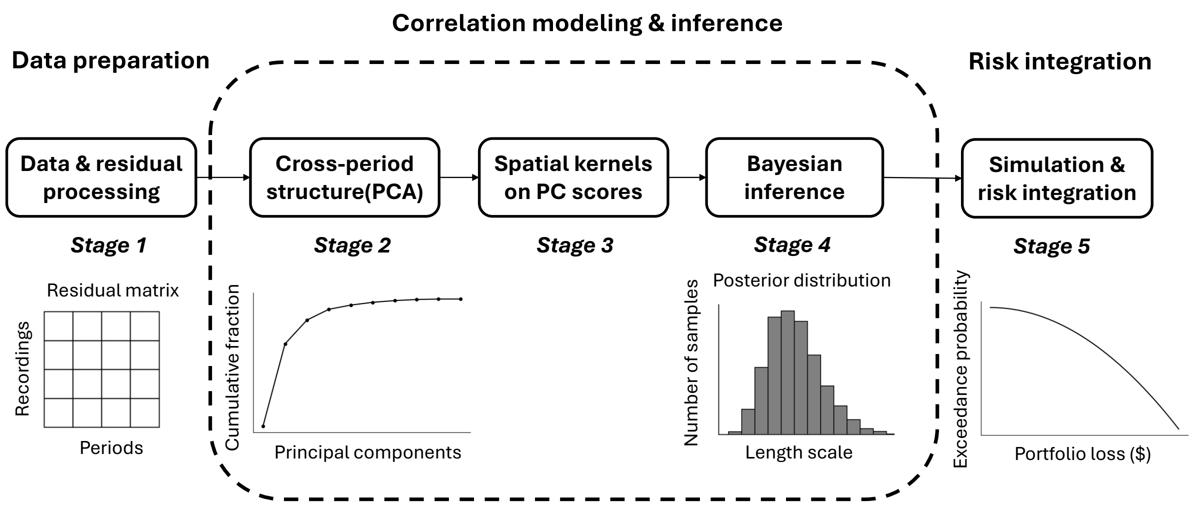


**Fig. 2.** Workflow for cross-period spatial correlation modeling and its integration into ground-motion simulation and seismic risk analysis.

Figure 2 illustrates the overall methodological framework of the study. The core workflow consists of five stages: (1) Data and residual processing; (2) Cross-period structure decomposition via PCA; (3) spatial correlation kernels modeling on PC scores; (4) Bayesian parameter inference; and (5) Simulation and regional risk integration.

Data and residual processing (Stage 1) are described in the Data and preprocessing section. Model performance is evaluated separately after fitting parameters (Stage 4) and not included as an explicit stage in the workflow. The following subsections describe each component of the framework in detail.

### Cross-period structure via PCA

PCA transforms the residual matrix into a set of orthogonal PCs [26]. The transformation can be expressed as

$$\mathbf{P}\bar{\mathbf{W}} = \mathbf{S}, \tag{3}$$

$$\begin{bmatrix} p_{1,T_1} & \cdots & p_{1,T_m} \\ \vdots & \ddots & \vdots \\ p_{m,T_1} & \cdots & p_{m,T_m} \end{bmatrix} \begin{bmatrix} \delta\bar{W}_{T_1}(x_1) & \cdots & \delta\bar{W}_{T_1}(x_n) \\ \vdots & \ddots & \vdots \\ \delta\bar{W}_{T_m}(x_1) & \cdots & \delta\bar{W}_{T_m}(x_n) \end{bmatrix} = \begin{bmatrix} s_1(x_1) & \cdots & s_1(x_n) \\ \vdots & \ddots & \vdots \\ s_m(x_1) & \cdots & s_m(x_n) \end{bmatrix}, \tag{4}$$

where $\mathbf{P}$ is the transformation (loading) matrix, $\bar{\mathbf{W}}$ is the normalized within-event residual matrix that aggregates residuals across all $m$ vibration periods and $n$ recordings, $x_t$ denotes the location of the $t$-th recording, and $\mathbf{S}$ denotes the PC (score) matrix representing the uncorrelated components. Each row of $\mathbf{S}$ corresponds to one PC that can be modeled independently. The loading matrix $\mathbf{P}$ can be expressed as column vectors $[\,\mathbf{p}_1, \ldots, \mathbf{p}_m\,]$, where $\mathbf{p}_k$ is the $k$-th eigenvector of the covariance matrix $\mathbf{C}$ of $\bar{\mathbf{W}}$, arranged in descending order of the corresponding eigenvalues. Principal axes are identifiable only up to sign; we orient each $\mathbf{p}_k$ by requiring that the largest-magnitude entry of pooled within-event residuals matrix be positive. The covariance matrix is defined as

$$\mathbf{C} \;=\; \frac{\bar{\mathbf{W}}\bar{\mathbf{W}}^{\top}}{n-1}. \tag{5}$$

In this framework, we assume that covariance structures evaluated at different separation distances $h$ share a common set of principal axes. This assumption is related to, but less restrictive than, the intrinsic correlation assumption in multivariate geostatistics, where the multivariate correlation structure is assumed to be independent of spatial scale [27]. The covariance matrix $\mathbf{C}$ defined in Equation 5 corresponds to the zero-separation case ($d_E$). That is, the family of covariance matrices can be approximately diagonalized by the same loading matrix $\mathbf{P}$. This assumption allows each component to be modeled independently.

Based on the orthogonality property of the loading matrix, the within-event residual matrix can be reconstructed from the PC matrix as

$$\bar{\mathbf{W}} \;=\; \mathbf{P}^{-1}\mathbf{S} \;=\; \mathbf{P}^{\top}\mathbf{S}, \tag{6}$$

### Spatial correlation kernels on PC scores

We will construct two spatial correlation models for these PCs. The first is Model E (the Euclidean distance-only model), which depends only on $d_E$. The corresponding kernel is given by

$$\rho_{\mathrm{E}}(d_E;\, \ell_E,\, \gamma_E) = \exp\Big[-\big(d_E/\ell_E\big)^{\gamma_E}\Big]. \tag{7}$$

It follows a powered-exponential form with range parameter $\ell_E > 0$ and smoothness exponent $\gamma_E \in (0, 2]$ [28]. This is the typical functional form used by most previous spatial correlation models, and neglects the potential influence of path and site effects.

Model EAS (the Euclidean distance–Azimuth–Soil dissimilarity model) additionally depends on epicenter-to-station azimuth differences and $V_{s,30}$ differences

$$\rho_{\mathrm{EAS}}\big(d_E,\, d_A,\, d_S;\, \boldsymbol{\psi}\big) \;=\; \rho_{\mathrm{E}} \times \big[w\,\rho_A + (1-w)\,\rho_S\big]. \tag{8}$$

where $\boldsymbol{\psi} = \{\ell_E,\, \gamma_E,\, \ell_A,\, \ell_S,\, w\}$ is a vector of the five model parameters. This functional form, proposed by Bodenmann et al. [16], additionally accounts for path and site effects. The first term, $\rho_E$, comes from Equation 7 and controls the spatial decay of correlation with Euclidean inter-station distance. The second term, $\rho_A$, characterizes the directional similarity between sites. It is governed by the angular range parameter $\ell_A > 0$ [29]:

$$\rho_A(d_A;\, \ell_A) = \big(1 + d_A/\ell_A\big)\Big(1 - d_A/180\Big)^{180/\ell_A}. \tag{9}$$

The third term, $\rho_S$, accounts for differences in near-surface conditions. The range parameter $\ell_S > 0$ controls how rapidly correlation decreases with increasing site condition contrast:

$$\rho_S(d_S;\, \ell_S) = \exp\Big[-\,d_S/\ell_S\Big]. \tag{10}$$

**TABLE 1.** Priors for the EAS correlation parameters (per principal component and event). Means and quantiles are reported for the stated parameterization; $q_{0.05}$ and $q_{0.95}$ denote the 5th and 95th percentiles of the corresponding prior distributions. Gamma$(\alpha, \beta)$ uses the rate parametrization (mean $\alpha/\beta$); InvGamma$(\alpha, \beta)$ uses the scale parametrization (mean $\beta/(\alpha - 1)$ for $\alpha > 1$).

| Parameter | Units | Domain | Distribution | Mean | $q_{0.05}$ | $q_{0.95}$ |
|---|---|---|---|---|---|---|
| $\gamma_E$ | – | $(0, 2)$ | $p(\gamma_E/2)$ = Beta(2, 2) | 1.0 | 0.3 | 1.7 |
| $\ell_E$ | km | $\mathbb{R}^+$ | InvGamma(2, 30) | 30.0 | 6.3 | 84.0 |
| $\ell_A$ | $^\circ$ | $(0, 45)$ | $p\big(180/\ell_A - 4\big)$ = Gamma(2, 0.25) | 18.2 | 7.8 | 33.3 |
| $\ell_S$ | $\mathrm{m\,s^{-1}}$ | $\mathbb{R}^+$ | InvGamma(2, 100) | 100.0 | 21.0 | 284.6 |
| $w$ | – | $(0, 1)$ | Beta(2, 2) | 0.5 | 0.1 | 0.9 |

These kernel functions, $\rho_{\mathrm{E}}$ and $\rho_{\mathrm{EAS}}$, ensure that the resulting covariance matrix is positive semi-definite, with values bounded between 0 and 1. Accordingly, the PC scores are normalized so that their covariance structure can be directly interpreted as a correlation matrix, eliminating the need to introduce additional variance parameters for individual PCs during the modeling process.

Each PC is a linear combination of the normalized within-event residuals and therefore follows a zero-mean multivariate normal distribution. The components are further scaled to unit variance to obtain standard normal variables:

$$\bar{s}_k(x) \equiv \frac{s_k(x)}{\sigma_{s_k}}, \tag{11}$$

where $s_k(x)$ denotes the $k$-th PC at location $x$, and $\sigma_{s_k}$ denotes its standard deviation, given by the square root of the corresponding $k$-th eigenvalue of covariance matrix $\mathbf{C}$.

**Bayesian parameter inference**

Using the kernel functions from Equations 7 and 8, we estimate the parameter vector $\boldsymbol{\psi} = \{\ell_E, \gamma_E, \ell_A, \ell_S, w\}$ from the data using a Bayesian inference framework[30;31].

Bayesian inference yields the posterior distribution of unknown parameters by combining prior knowledge and observed data via Bayes' rule. Taking Model EAS as an example, for the $k$-th PC, the posterior distribution of parameters $\boldsymbol{\psi}_k$ given all normalized PC score fields $\bar{\mathbf{S}}_k$ is

$$p(\boldsymbol{\psi}_k \mid \bar{\mathbf{S}}_k) \;\propto\; p(\boldsymbol{\psi}_k)\, p(\bar{\mathbf{S}}_k \mid \boldsymbol{\psi}_k) = p(\boldsymbol{\psi}_k) \prod_{j=1}^{N_{\text{event}}} \mathcal{N}\Big(\bar{\mathbf{S}}_{k,j};\, \mathbf{0},\, \boldsymbol{\rho}(d_{E,j},\, d_{A,j},\, d_{S,j};\, \boldsymbol{\psi}_k)\Big), \tag{12}$$

where $p(\boldsymbol{\psi}_k)$ denotes prior distribution of parameters, $\mathcal{N}$ denotes the multivariate normal distribution, $\bar{\mathbf{S}}_{k,j}$ denotes the normalized score field of the $k$-th PC for event $j$, and $\boldsymbol{\rho}$ denotes the covariance structure defined by Equation 8 with parameter set $\boldsymbol{\psi}_k$. To minimize the influence of prior assumptions, weakly informative priors $p(\boldsymbol{\psi}_k)$ are adopted for all PCs[32;33] (Table 1).

**Model validation**

*Likelihood-Based Model Evaluation*

We then use the posterior distributions of the parameters to compare different correlation models. The idea is to evaluate the likelihood of the observed data under the covariance matrix constructed from the kernel function, input metrics, and the inferred parameters. Taking Model EAS as an example, this likelihood for each PC can be expressed as

$$p(\bar{\mathbf{S}}_k \mid \boldsymbol{\psi}_k) = \prod_{j=1}^{N_{\text{event}}} \mathcal{N}\Big(\bar{\mathbf{S}}_{k,j};\, \mathbf{0},\, \boldsymbol{\rho}(d_{E,j},\, d_{A,j},\, d_{S,j};\, \boldsymbol{\psi}_k)\Big). \tag{13}$$

Similarly, the likelihood can also be formulated in the single-period domain as

$$p(\bar{\mathbf{W}}_{T_l} \mid \boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_m) = \prod_{j=1}^{N_{\text{event}}} \mathcal{N}\Big(\bar{\mathbf{W}}_{T_l,\, j};\, \mathbf{0},\, \boldsymbol{\rho}_{T_l,\, j}\Big), \tag{14}$$

where $\bar{\mathbf{W}}_{T_l,\,j}$ denotes the vector of observed normalized within-event residuals at the $l$-th vibration period for event $j$, and $\boldsymbol{\rho}_{T_l,\,j}$ is the corresponding covariance matrix, reconstructed from the PCs through the PCA loading matrix as

$$\boldsymbol{\rho}_{T_l,\,j} = \sum_{k=1}^{m} p_{k,T_l}^2\, \sigma_{s_k}^2\, \boldsymbol{\rho}(d_{E,j},\, d_{A,j},\, d_{S,j};\, \boldsymbol{\psi}_k), \tag{15}$$

where $p_{k,T_l}$ denotes the loading of the $k$-th PC at period $T_l$, $\sigma_{s_k}$ denotes $k$-th PC's standard deviation as defined in Equation 11. This formulation follows from Equation 3, in which the residual field at each period is expressed as a linear combination of mutually uncorrelated PCs. Consequently, the period-specific spatial covariance is obtained by summing the covariance contributions from individual components, with each contribution scaled by the squared loading and the variance associated with that component.

*Log Pointwise Predictive Density (LPPD)*

Once the likelihood of the observed data is computed for each posterior parameter sample (either for each normalized PC or for the normalized single-period within-event residual), we obtain a distribution of likelihood values representing how well the model explains the data. Based on this distribution, the log pointwise predictive density (LPPD) can be calculated for two models to enable a more direct comparison, which is defined as (take $k$-th PC as an example):

$$\text{LPPD for the } k\text{-th PC} = \log\left(\frac{1}{N_{\text{sample}}} \sum_{q=1}^{N_{\text{sample}}} p(\bar{\mathbf{S}}_k \mid \boldsymbol{\psi}_k^{(q)})\right), \tag{16}$$

where $N_{\text{sample}}$ is the total number of posterior samples, and $\boldsymbol{\psi}^{(q)}$ represents the $q$-th posterior sample of the parameter vector. A higher LPPD indicates a better model fit to the observed data. The reported LPPD values are interpreted as in-sample predictive fit.

**Simulation and regional risk integration**

This subsection illustrates the workflow of how Model E and Model EAS can be applied in earthquake risk analysis. We employ these two models when forming the $SA$ fields using Equation 1. The first term, representing the median prediction, and the second term, corresponding to the between-event residuals, are not directly relevant to this study. Our focus is on the third term—the within-event residual simulation. Given the posterior distributions of parameters, kernel functions, and input metrics (e.g., $d_E$, $d_A$, $d_S$), we first compute the covariance matrix for each PC as

$$\boldsymbol{\Sigma}_k = \sigma_{s_k}^2\, \boldsymbol{\rho}_k(d_E, d_A, d_S;\, \boldsymbol{\psi}_k). \tag{17}$$

where $\boldsymbol{\Sigma}_k$ is the covariance matrix of the $k$-th PC field, and $\boldsymbol{\rho}_k$ denotes the spatial correlation matrix derived from the corresponding kernel function (Equation 7 or 8) and parameter set $\boldsymbol{\psi}_k$. We then generate realizations of the PC field by drawing samples $\mathbf{s}_k^{(r)}$ from $\boldsymbol{\Sigma}_k$, where the superscript $r$ indexes independent Monte Carlo realizations. Since the PCs are uncorrelated, they can be sampled independently. Using Equation 3, the residuals are reconstructed in the period space, yielding the within-event residual matrix for nine periods. In addition to Model E and Model EAS described above, this study also employs three existing models for comparison—Goda and Hong[6], Loth and Baker[17], and Markhvida et al.[20]—referred to as GH08, LB13, and M18, respectively.

Then, for asset $b$ with fundamental period $T_b$, we use period-matched fragility functions to obtain exceedance probabilities in lognormal form[34]:

$$\mathbb{P}\{\text{DS} \geq ds_i \mid SA = a\} = \Phi\left(\frac{\ln\left(a/\theta_i\right)}{\beta_i}\right), \qquad i = 1, \ldots, 4, \tag{18}$$

where $ds_i$ denotes the $i$-th damage state (*slight*, *moderate*, *extensive*, *complete*), corresponding to $i = 1$–$4$; $\theta_i$ and $\beta_i$ are, respectively, the median and logarithmic standard deviation of the fragility function for asset $b$ at period $T_b$; and $a$ is the $SA$ value at the location of asset $b$[35].

Damage-state category probabilities are then

$$\mathbb{P}\{\text{DS} = ds_i \mid SA = a\} = \mathbb{P}\{\text{DS} \geq ds_i \mid SA = a\} - \mathbb{P}\{\text{DS} \geq ds_{i+1} \mid SA = a\}, \tag{19}$$

where $\mathbb{P}\{\text{DS} \geq ds_5 \mid SA = a\} = 0$ for the complete damage state.

**TABLE 2.** Loading matrix from principal component analysis (PCA) across periods

| **Period (s)** | $\boldsymbol{p}_1$ | $\boldsymbol{p}_2$ | $\boldsymbol{p}_3$ | $\boldsymbol{p}_4$ | $\boldsymbol{p}_5$ | $\boldsymbol{p}_6$ | $\boldsymbol{p}_7$ | $\boldsymbol{p}_8$ | $\boldsymbol{p}_9$ |
|---|---|---|---|---|---|---|---|---|---|
| 0.01 | 0.407 | -0.167 | 0.057 | -0.023 | -0.015 | 0.012 | -0.337 | -0.490 | -0.670 |
| 0.03 | 0.401 | -0.181 | 0.094 | -0.065 | -0.041 | 0.024 | -0.372 | -0.337 | 0.734 |
| 0.06 | 0.380 | -0.255 | 0.233 | -0.211 | -0.122 | 0.053 | -0.261 | 0.771 | -0.107 |
| 0.10 | 0.362 | -0.292 | 0.282 | -0.167 | -0.067 | -0.007 | 0.807 | -0.140 | 0.020 |
| 0.30 | 0.354 | -0.100 | -0.222 | 0.635 | 0.481 | -0.383 | 0.078 | 0.166 | 0.021 |
| 0.60 | 0.332 | 0.169 | -0.515 | 0.183 | -0.164 | 0.716 | 0.134 | 0.069 | 0.009 |
| 1.00 | 0.287 | 0.356 | -0.479 | -0.372 | -0.330 | -0.556 | 0.063 | 0.027 | 0.003 |
| 3.00 | 0.222 | 0.562 | 0.209 | -0.370 | 0.656 | 0.156 | 0.013 | 0.009 | 0.000 |
| 6.00 | 0.172 | 0.555 | 0.517 | 0.460 | -0.426 | -0.055 | 0.003 | 0.004 | -0.002 |

In each Monte Carlo realization $r$, the damage state of asset $b$ is sampled by first drawing a random number $u_b^{(r)} \sim U(0,1)$. This value is then compared with the damage-state exceedance probabilities to assign a discrete damage state $ds_b^{(r)}$. Equivalently, $ds_b^{(r)}$ is sampled from the categorical distribution defined by Eq. (19). The corresponding loss is then computed as

$$loss_b^{(r)} = R_b \, c_{b,ds_b^{(r)}}, \qquad loss^{(r)} = \sum_b loss_b^{(r)}, \tag{20}$$

where $R_b$ is the replacement cost of asset $b$, and $c_{b,ds_b^{(r)}}$ is the fixed loss ratio assigned to asset $b$ for the sampled damage state $ds_b^{(r)}$. Thus, damage-state uncertainty is propagated through Monte Carlo sampling, while the loss ratio conditional on the sampled damage state is treated as deterministic. Aggregating results over all realizations yields scenario losses, loss exceedance curves, and the expected scenario loss.

For Models E and EAS, epistemic uncertainty in the spatial correlation parameters is propagated by sampling from their posterior distributions. For each posterior parameter sample, a full set of Monte Carlo loss realizations is performed following Equation 20, producing a corresponding loss exceedance curve. The collection of curves obtained from all posterior samples constitutes an ensemble that represents parameter-driven (epistemic) uncertainty.

## RESULTS

This section presents and discusses the results obtained by applying the methodology described in the previous section to the dataset introduced in the Data and preprocessing section.

### Cross-period structure via PCA

Figure 3 summarizes the variance captured by successive PCs for individual events (thin grey) and for the pooled (all events) dataset (thick black). The first three PCs together explain approximately 90% of the total variance, indicating that they capture the dominant modes of cross-period variability. Consistent patterns are observed across individual events, in agreement with previous findings by Markhvida et al.[20]

Based on this variance contribution, subsequent analyses focus on the first three PCs. Figure 4 shows the period-dependent loadings $p_1(T)$–$p_3(T)$ for each event (grey), their across-event mean and mean plus standard deviation ($\mu \pm \sigma$) envelope (black solid and red dashed), and the pooled loadings (black triangles). The first component $p_1(T)$ remains positive across all periods and decays slightly with increasing $T$, reflecting an overall amplitude factor. It shows $p_1(T)$ primarily controls the general strength of ground motions across all periods without changing the spectral shape. The second component $p_2(T)$ changes sign from negative to positive with period, representing a short–to–long spectral tilt. The third component $p_3(T)$ exhibits a negative trough near $T \approx 0.1$ s and a positive peak around $T \approx 0.6$–1 s, capturing second-order effect (curvature) in the response spectrum. Inter-event variability is small for the first two components but becomes more pronounced for the third component, as higher-order modes typically explain less variance and are therefore more sensitive to event-specific effects and noise. At a broader spectral level, the full loading matrix (Table 2) confirms that a few leading PCs control the overall spectral form, whereas remaining components capture small-scale deviations.

The PCA transformation assumes that the spatial covariance matrices at all separation distances $d_E$ can be diagonalized by a common loading matrix $\mathbf{P}$. Under this assumption, each PC can be modeled with an independent

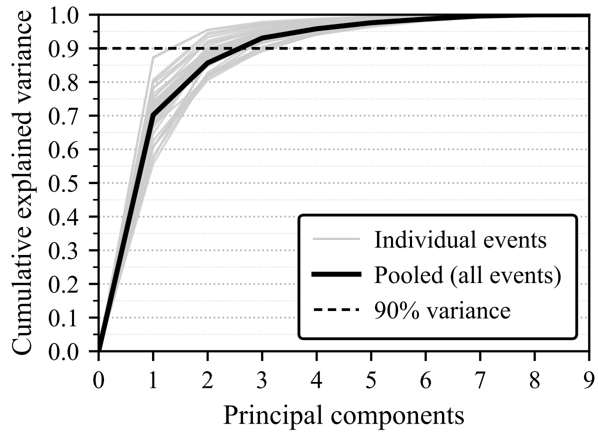


**Fig. 3.** Cumulative variance explained by successive principal components for individual events (thin grey) and pooled data (thick black). The dashed line marks the 90% threshold.

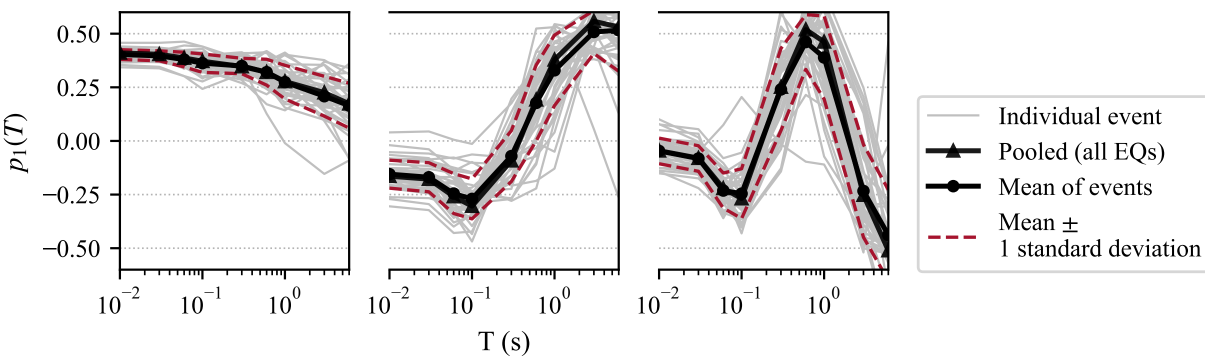


**Fig. 4.** Period-dependent loadings for the first three principal components. Grey lines: individual events; black triangles: pooled loadings; black solid with markers and red dashed: across-event mean and mean ± 1 standard deviation.

spatial kernel, and the period-specific spatial correlations recovered through Equation 15. To assess this assumption, Figure 5 compares per-period spatial correlation curves reconstructed from the PCA-based approach using Equation 15 with those obtained from direct single-period fits to within-event residuals using the same Model E functional form (Equation 7). The direct-fit models show $SA$ at longer periods having higher correlations at a given distance, consistent with the physical expectation that longer-wavelength motions are spatially coherent over larger distances[7]. The PCA reconstruction, by contrast, compresses this range variation, with all nine periods having similar correlations, suggesting that an approximation has been introduced by the transformation.

However, the PCA transformation remains useful for several reasons. First, the magnitude of the PCA-induced correlation compression is comparable to inter-model variability among published spatial correlation studies[6;17;20], as illustrated in Figure S1 of the Supplementary Material. Second, the portfolio-level loss results presented below, using models with and without PCA transformations, show that mean losses are broadly consistent across correlation models tested, suggesting that the per-period approximation does not materially affect risk estimates for this application. Third, the PCA framework's computational efficiency, guaranteed positive-definiteness, and compact representation of cross-period dependence are practically valuable.

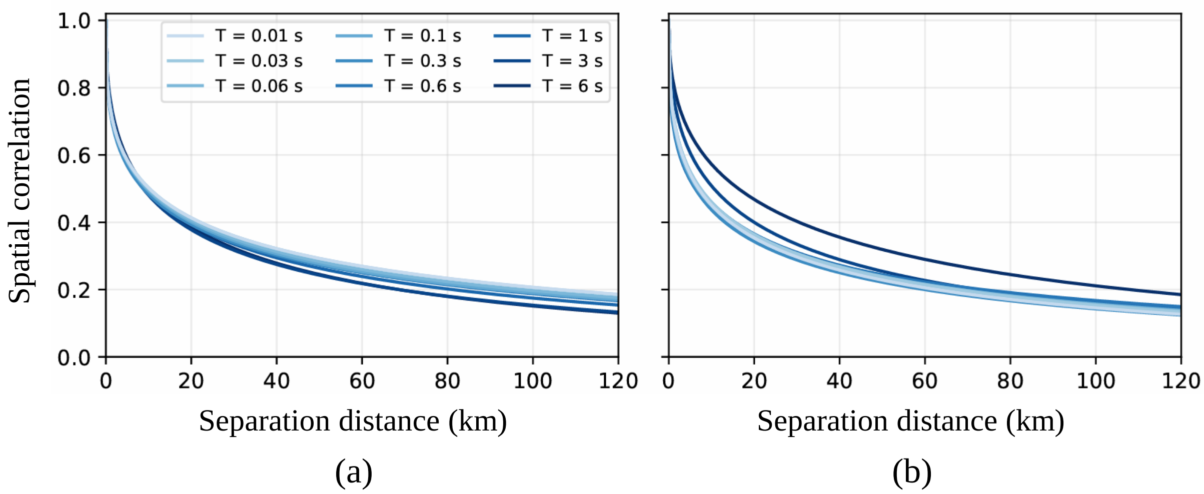


**Fig. 5.** Comparison of per-period spatial correlation curves obtained via (a) reconstruction from the PCA-based approach using Equation 15 with fitted PC parameters and the Model E (Equation 7), and (b) direct single-period maximum-likelihood fits using Model E (Equation 7) to each period's within-event residuals. Line darkness increases with oscillator period.

**Posterior behavior across PCs and spatial interpretation**

Posterior distributions of Model E and Model EAS for each normalized PC are summarized in Figure 6. The large dots indicate the posterior medians, the thick lines show the 50% credible intervals (CIs), and the thin lines show the 95% CIs. Compared to the isotropic Model E, Model EAS exhibits longer path ranges (larger $\ell_E$), while the corresponding smoothness exponents $\gamma_E$ remain broadly comparable between the two models, although some differences are present. This pattern is particularly evident for the first three PCs.

The angular scales ($\ell_A$) and site–similarity scales ($\ell_S$) are non-negligible. For the leading PCs, which determine the mainly practical effect, $\ell_A$ is generally on the order of several tens of degrees; an azimuth difference of $d_A = \ell_A$ reduces $\rho_A$ to roughly 0.65–0.70, while $d_A = 2\ell_A$ reduces it to about 0.25–0.35. Similarly, $d_S = \ell_S$ gives $\rho_S = e^{-1} \approx 0.37$, and $d_S = 2\ell_S$ gives $\rho_S \approx 0.14$. These values therefore indicate when angular separation and $V_{s,30}$ contrast are large enough for Model EAS to noticeably lower the correlation relative to a distance-only model. The mixing weight ($w$) remains away from the degenerate 0/1 limits, indicating that both path and site terms are active in shaping the correlation structure.

Across PCs, the CI of the angular length scale ($\ell_A$), site-similarity length scale ($\ell_S$), and mixing weight ($w$) are generally broader than those of the distance-related parameters $\ell_E$ and $\gamma_E$, indicating that the model is less confident in the directionality and site-similarity effects than in distance-based correlations.

In the posterior summaries, the distance scale parameter ($\ell_E$) is largest for PC1 and decreases sharply with PC order. For higher PCs, $\ell_E$ approaches zero, implying that the spatial correlation nearly collapses to a nugget effect and carries little spatial information. This difference further highlights the dominant role of inter-station distance in controlling spatial variability. The absence of systematic variation in other parameters does not imply that directionality or site similarity has no influence on spatial correlation. Rather, these effects are secondary and less stable across PCs, as the dominant spatial structure is already captured by inter-station distance in the leading components. Detailed numerical summaries of these posterior distributions are provided in the supplementary material (Tables S1 and S2).

To further investigate the dependencies among the inferred spatial parameters, Figures 7 and 8 present the joint posterior distributions (pair plots) of the key parameters $\{\gamma_E, \ell_E, \ell_A, \ell_S, w\}$ for the first and second PCs, respectively. The diagonal panels show marginal posterior densities, while the off-diagonal panels illustrate pairwise relationships among parameters for Model E and Model EAS.

In PC1 (Figure 7), two discernible dependencies are observed: a negative correlation between the distance scale $\ell_E$ and the smoothness parameter $\gamma_E$, and a negative relationship between the site-similarity scale $\ell_S$ and the weight $w$. These patterns indicate coupled effects among certain parameter pairs, where a decrease in one parameter

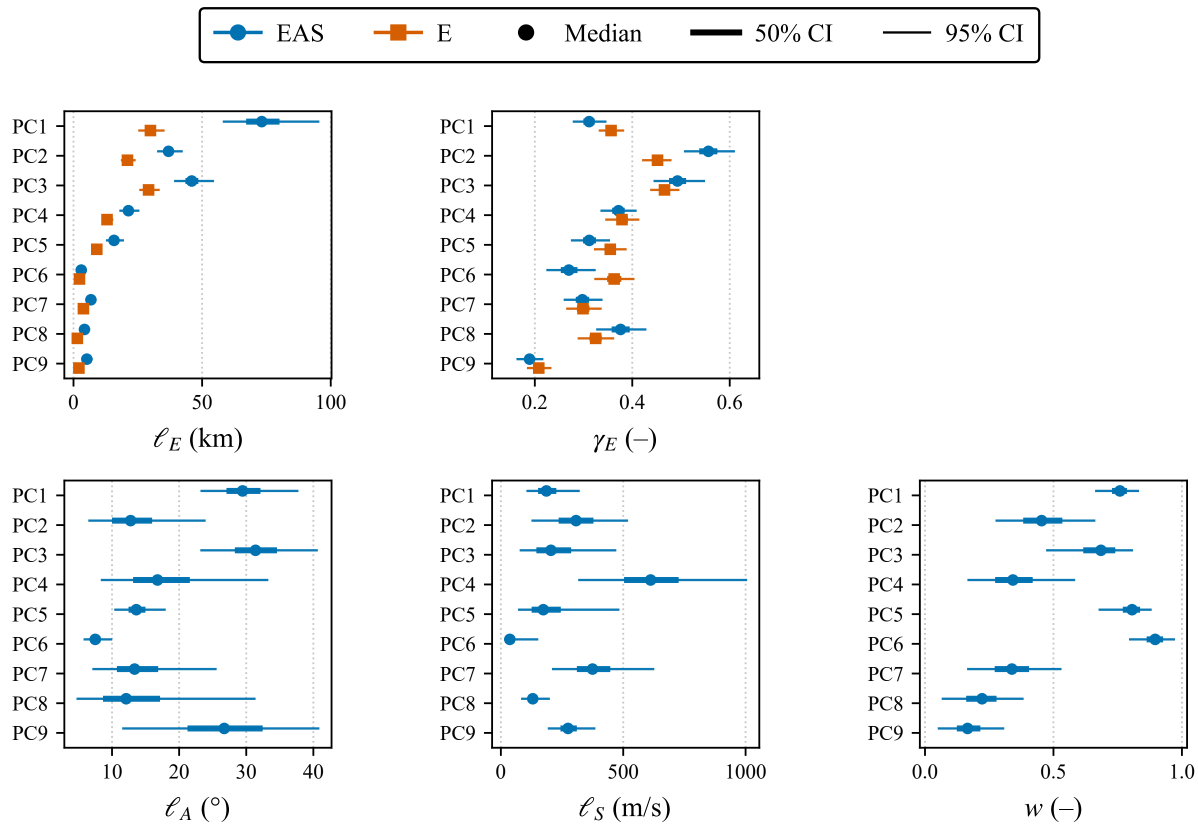


**Fig. 6.** Posterior parameter distributions of Model E (orange) and Model EAS (blue) for each normalized principal component (PC). Large dots mark posterior medians, thick lines show 50% credible intervals (CIs), and thin lines show 95% CIs.

is compensated by an increase in the other. This behavior reflects compensating roles among parameters within the kernel formulation, rather than distinct physical relationships imposed by the data.

In contrast, PC2 (Figure 8) exhibits stronger dependencies involving the angular length scale $\ell_A$, indicating that its spatial structure is governed by a combination of directional effects and site-similarity clustering. The observed parameter interactions suggest that these two mechanisms can provide alternative explanations for the spatial variability captured by PC2, in contrast to the predominantly isotropic structure observed for PC1.

**Predictive fit**

We next assess predictive performance using the methods described in the Model validation section. Figure 9 compares the posterior distributions of the log probability for Model E and the proposed Model EAS, computed using posterior samples of the correlation parameters for each PC following Equation 13 and 14. Across all nine components (PC1–PC9), Model EAS consistently yields higher log probabilities than Model E, indicating improved predictive performance under the path- and site-aware formulation. Notably, even the lower tail of the posterior log-probability distribution for Model EAS exceeds the upper tail of the corresponding distribution for Model E, suggesting that the improvement is robust to parameter uncertainty.

At the same time, the log-probability distributions for Model EAS exhibit larger variance than those for Model E for most components. This behavior is expected, as the additional structure introduced in Model EAS allows for a wider range of plausible correlation realizations. An exception is observed for PC8 and PC9, where both models exhibit larger variance relative to the other components, consistent with a greater influence of noise and the weaker, less coherent spatial structure typically associated with higher-order components.

While Figure 9 illustrates the posterior distributions of log probability and highlights the consistent shift toward higher likelihood values for Model EAS, a complementary summary of predictive performance can be obtained

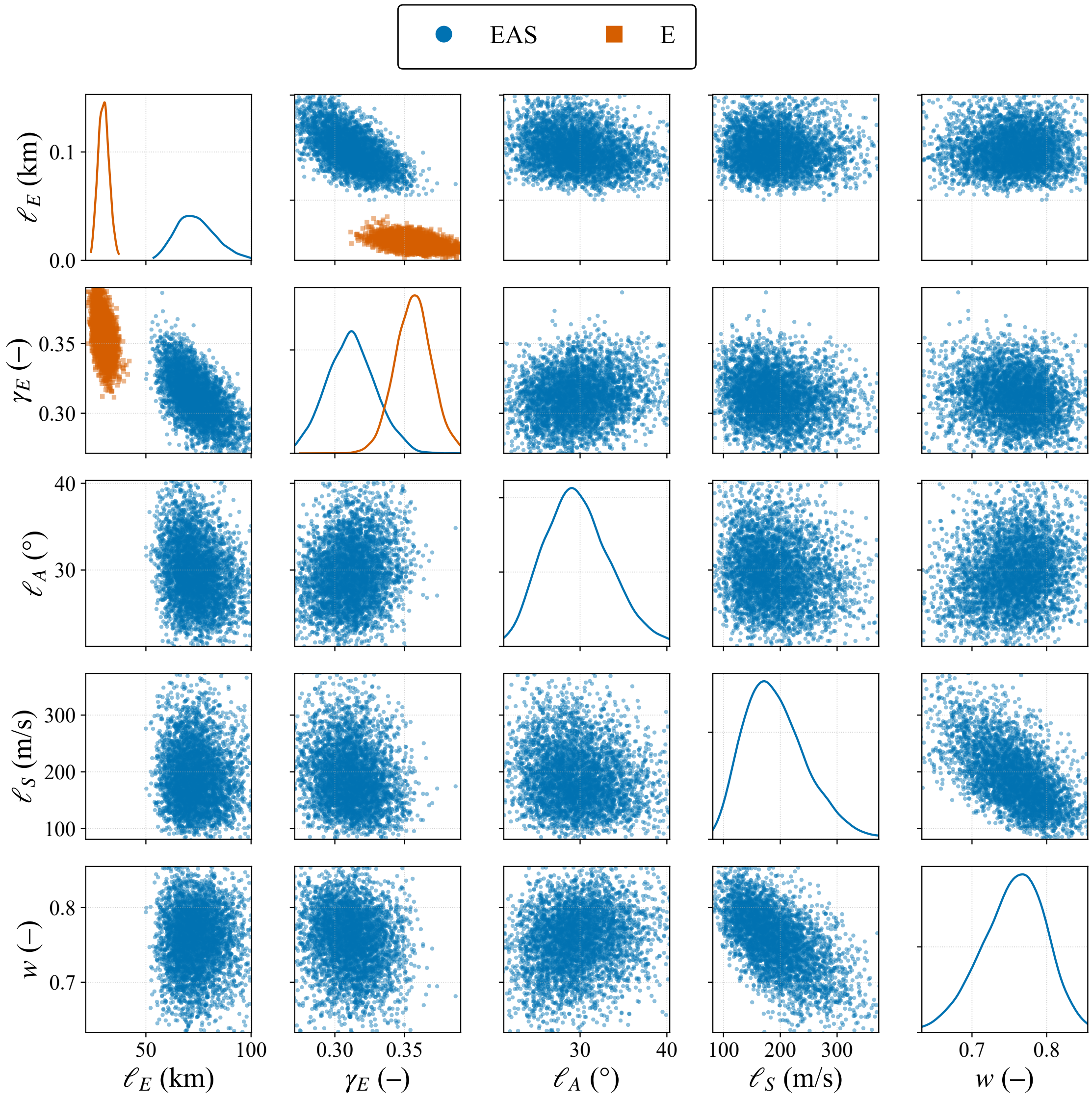


**Fig. 7.** Joint posterior distributions of key parameters $\{\gamma_E, \ell_E, \ell_A, \ell_S, w\}$ for the first principal component (PC1). The diagonal panels show marginal posterior densities, while the off-diagonal panels illustrate pairwise relationships between parameters for Model E (orange) and Model EAS (blue).

using the LPPD. The LPPD integrates over posterior uncertainty and provides a concise and interpretable metric for comparing predictive performance across PCs and oscillator periods.

Figure 10 summarizes the LPPD comparison results, taking Model E as the reference baseline to evaluate improvements from the updated anisotropic Model EAS. Positive values indicate a better predictive fit for model EAS. As shown in the left panel of Figure 10, the relative improvement in LPPD is concentrated in the leading PCs. Specifically, Model EAS yields an increase of approximately 1.9% for both PC1 and PC2, followed by a still notable improvement of about 1.5% for PC3. Beyond PC3, the relative gain decreases rapidly, dropping to around 1.0% for PC4 and below 0.5–0.6% by PC6–PC7. A small rebound occasionally appears for the highest-order PCs, but any aggregate effect is negligible because high-order components explain little variance. These 0.5–2% gains are modest in absolute magnitude but systematic across PCs and periods.

Consistent with the PC-based results discussed above, Model EAS provides improved predictive performance over Model E across all oscillator periods considered. As shown in the right panel of Figure 10, the improvement is positive

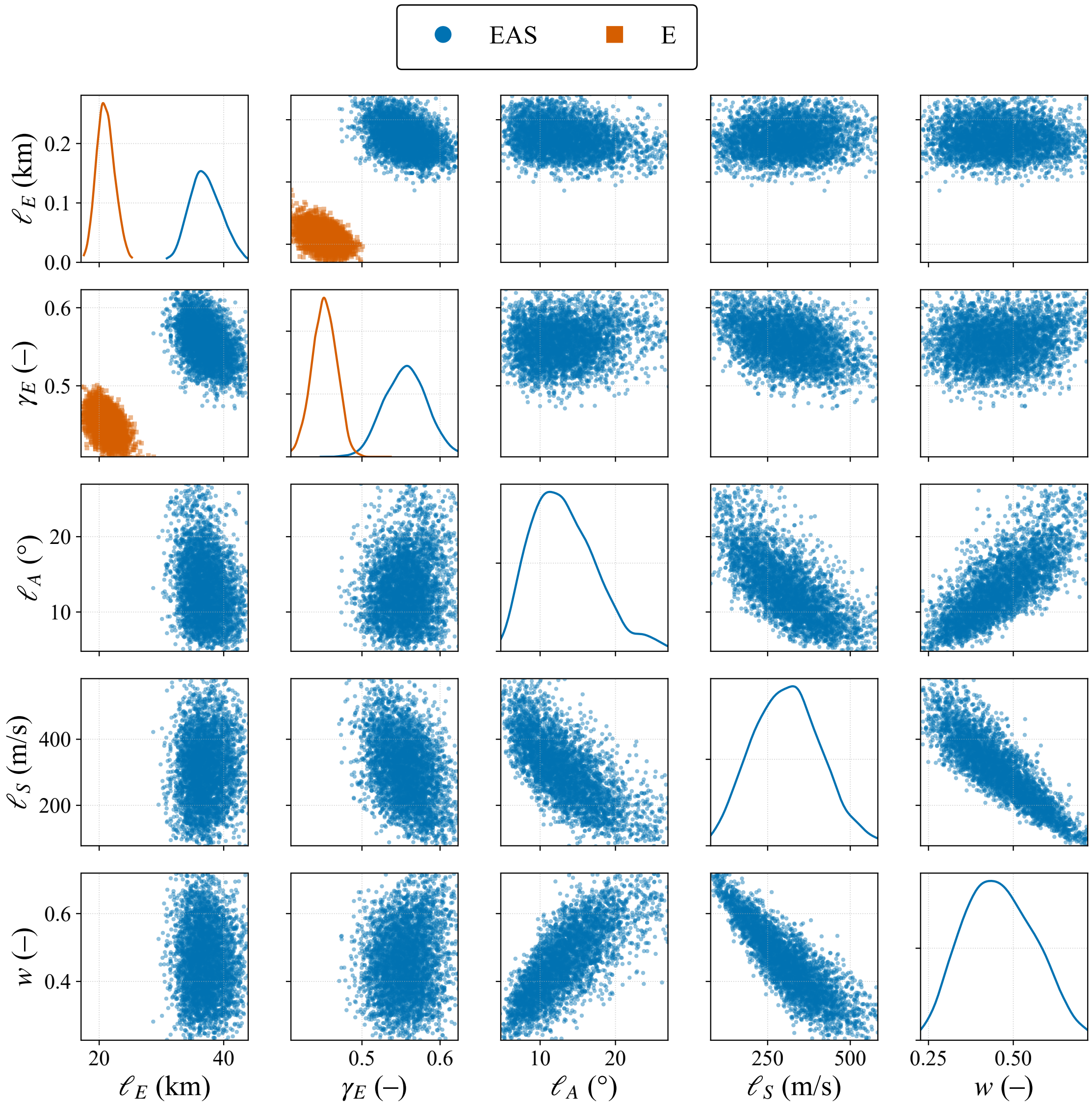


**Fig. 8.** Joint posterior distributions of key parameters $\{\gamma_E, \ell_E, \ell_A, \ell_S, w\}$ for the second principal component (PC2). The diagonal panels show marginal posterior densities, while the off-diagonal panels illustrate pairwise relationships between parameters for Model E (orange) and Model EAS (blue).

throughout the period range, indicating that incorporating anisotropic path effects and site similarity leads to systematic gains relative to the isotropic baseline. At the same time, the improvement exhibits a clear overall period-dependent trend. Relative gains at shorter periods are modest, on the order of about 1– 1.5%, while improvements become more pronounced at longer periods. Although fluctuations are observed across intermediate periods, the overall enhancement tends to increase toward longer periods, reaching values approaching 2% at periods on the order of 1–3 s and above.

This behavior is physically consistent. Longer-period ground motions are influenced by longer-wavelength propagation processes and regional-scale geologic heterogeneity, which produce stronger spatial coherence across stations. These mechanisms introduce systematic dependence on propagation path and regional site contrasts that are explicitly captured by the EAS formulation, leading to larger predictive gains at long periods. In comparison, shorter-period motions tend to be more localized and therefore provide less opportunity for improvement through spatial correlation modeling.

## CASE STUDY AND SCENARIO-BASED PORTFOLIO LOSS RESULTS

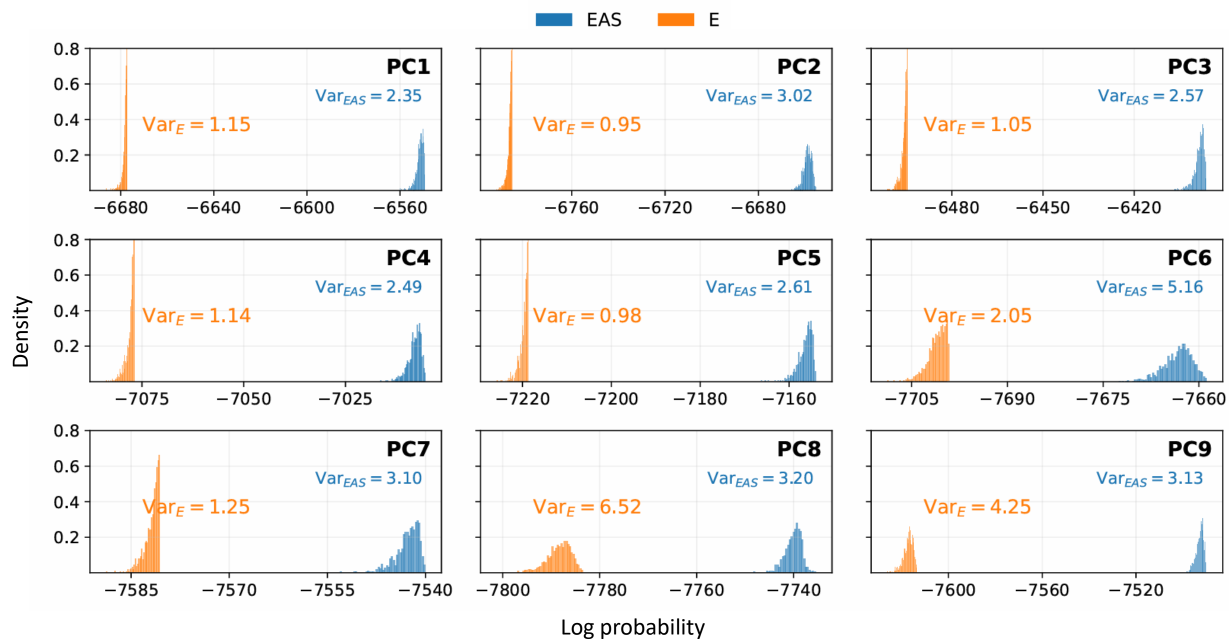


**Fig. 9.** Comparison of log-likelihood distributions for Model E (orange) and Model EAS (blue) across nine principal components (PC1–PC9). For each panel, histograms show the log-likelihood values of the observed principal-component residual field under the spatial correlation kernel of each model, evaluated using posterior samples of the model parameters. Annotated values indicate the variance of the log-likelihood distributions for each model.

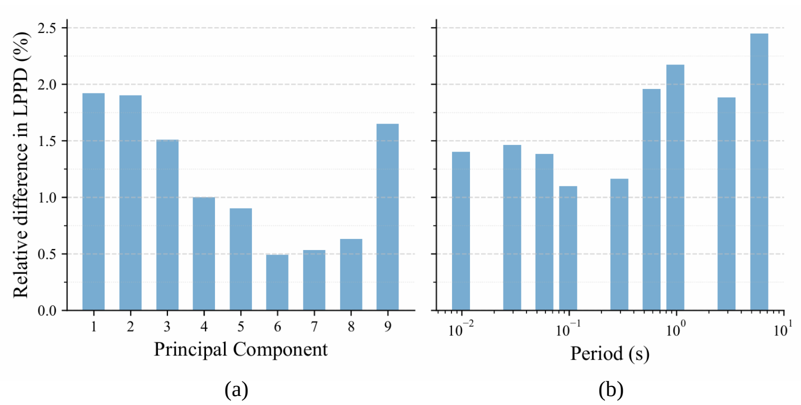


**Fig. 10.** Predictive performance of Model EAS relative to Model E in terms of log pointwise predictive density (LPPD). (a) LPPD gains by principal component (PC). (b) LPPD gains by oscillator period.

We implement the risk integration following the procedure described in Simulation and regional risk integration section. The target region is a $0.05° \times 0.05°$ area in the San Francisco Bay Area (37.80°–37.85° N, 122.23°–122.18° W; Figure 11a), subject to a magnitude 7.05 Hayward–Rodgers Creek rupture scenario. Median *SA*s are computed using the CY14 ground motion model. Residuals consist of (i) cross-period correlated between-event terms sampled following Baker and Bradley[36], and (ii) a within-event residual field generated using one of five cross-period spatial correlation models: Model E, and Model EAS, M18, LB13, GH08. The *SA* fields are generated using the SimCenter R2D framework, with the source code modified to accommodate the present study's parameter sampling and spatial correlation simulation requirements[37;38].

The exposure model consists of a synthetic portfolio of 900 buildings distributed evenly within a target region

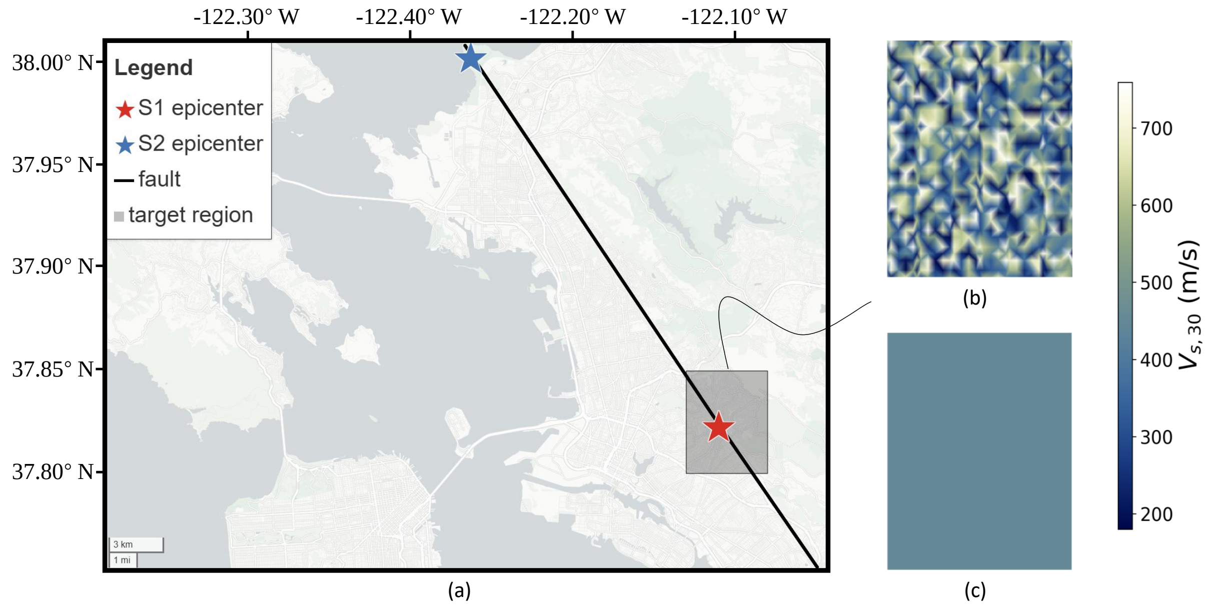


**Fig. 11.** Illustration of the case study setup. (a) Map of the target region showing the locations of two earthquake scenarios (S1 and S2) and the fault trace. (b) Spatial distribution of $V_{S,30}$ $(m/s)$ for Scenario S1. (c) Spatial distribution of $V_{S,30}$ for Scenario S2.

discretized by a $30 \times 30$ station lattice. The $SA$ simulations are performed for the same 900 locations. The resulting analysis domain spans approximately $5.4 \times 4.3$ km, with an average spacing of about 240 m between adjacent lattice points.

Table 3 summarizes the building material type, height category, and design level of the buildings. The synthetic building classes are derived from the Global Earthquake Model (GEM) taxonomy and simplified into a reduced set of representative categories for this study[39;40]. For simplicity, we model steel buildings as steel moment-resisting frames and masonry buildings as confined masonry structures. We further assume that the building code level corresponds directly to the structural ductility level assigned to each building class. Low-rise, mid-rise, and high-rise categories are assumed to represent buildings with 1, 3, and 10 stories, respectively.

**TABLE 3.** Summary of the synthetic building exposure model. Count indicates the number of buildings of this type in the portfolio. Period indicates the spectral acceleration period used for the corresponding GEM fragility function.

| **Building Type** | **Height** | **Design Level** | **Count** | **Period** |
|---|---|---|---:|---:|
| Masonry | Low-rise | Low | 100 | 0.01 |
| Steel | Low-rise | High | 200 | 0.3 |
| Steel | Low-rise | Moderate | 200 | 0.3 |
| Steel | Low-rise | Low | 100 | 0.3 |
| Steel | Mid-rise | High | 100 | 0.6 |
| Steel | Mid-rise | Moderate | 100 | 0.6 |
| Steel | Mid-rise | Low | 40 | 0.6 |
| Steel | High-rise | High | 30 | 1.0 |
| Steel | High-rise | Moderate | 20 | 1.0 |
| Steel | High-rise | Low | 10 | 1.0 |

Fragility functions for each building class are adopted from the GEM vulnerability model library[39;40]. In this study, peak ground acceleration (PGA) is approximated by the $SA$ at $T = 0.01s$. These fragility models utilize $SA$s at

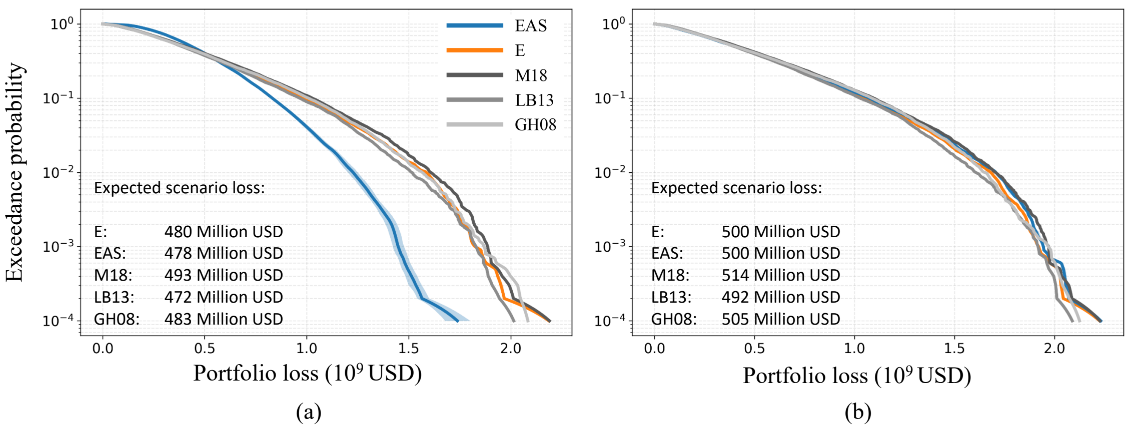


**Fig. 12.** Portfolio-level loss exceedance curves for the two earthquake scenarios, (a) S1 and (b) S2. Curves compare the anisotropic Model EAS (blue) with the isotropic distance-only Model E (orange) and existing spatial correlation models (M18, LB13, and GH08). Each curve represents the probability of exceeding a given portfolio ground-up loss, expressed in billion USD. The probabilities are not annualized and do not represent a full regional probabilistic risk assessment. Shaded bands indicate the 10th–90th percentile uncertainty range for Models E and Model EAS only.

multiple vibration periods, which necessitates modeling both cross-period and spatial correlation in the ground motion simulations.

To isolate the role of anisotropy and site heterogeneity, we analyze two scenarios with identical magnitude and mechanism but different geometry and site conditions (Figure 11): S1 (near-epicenter; heterogeneous $V_{s,30}$) places the epicenter inside the target box and uses a randomized heterogeneous $V_{s,30}$ field; S2 (far-epicenter; uniform $V_{s,30}$) places the epicenter far outside the target box and fixes $V_{s,30} = 450$ m/s (uniform NEHRP C-like conditions).

For each scenario and correlation model, we perform 10,000 Monte Carlo realizations using fixed random seeds to ensure result reproducibility. The loss of each building is computed following Equation 20, and total portfolio losses are obtained by aggregating individual building losses. Loss exceedance curves are then derived for both earthquake scenarios.

For Models E and EAS, epistemic uncertainty associated with the parameters in Equation 7 or 8 is additionally quantified by sampling 1,000 realizations from their posterior distributions. For each posterior parameter sample, a full set of 10,000 Monte Carlo loss realizations is performed following Equation 20, yielding a corresponding loss exceedance curve. This procedure results in an ensemble of 1,000 loss exceedance curves (i.e., a total of $1{,}000 \times 10{,}000$ loss simulations), which represents parameter-driven uncertainty. In Figure 12, this uncertainty is summarized by the 10th–90th percentile envelope of exceedance probabilities across the ensemble at each loss level.

Figure 12 shows portfolio loss exceedance curves. For S1, because the epicenter is located near the center of the target region and the $V_{s,30}$ field is randomly generated, sites exhibit large azimuthal differences and strong soil dissimilarity. These factors lead to weaker spatial correlation, so Model EAS predicts that large losses are less likely to occur simultaneously across many sites, shifting the tail of the exceedance curve leftward (lower high-loss probability). For S2, the epicenter is distant and all sites lie in a similar azimuthal direction, while the soil condition is uniform ($V_{s,30} = 450$ m/s). As a result, both $\rho_A$ and $\rho_S$ approach 1, making the overall $\rho_{\mathrm{EAS}}$ nearly identical to the isotropic $\rho_E$. Because $\ell_E$ values inferred for Model EAS are generally larger than those for Model E, the EAS tail is slightly higher, though the difference is minor. Other published models (M18, LB13, GH08) depend solely on inter-site distance and thus behave similarly to Model E under these uniform conditions.

Across both scenarios, mean portfolio losses are broadly comparable across models. Meaningful differences are observed mainly in the tail of the exceedance curves and are largely limited to Model EAS in S1, consistent with its distinct anisotropic correlation mechanism relative to the isotropic models. The crossing of the curves in S1 indicates that Model EAS does not reduce exceedance probabilities uniformly. Instead, it lowers the probability of very large losses while slightly increasing exceedance probabilities at more moderate loss levels. This redistribution of probability mass explains why the high-loss tail is lighter under Model EAS, while the mean scenario losses remain nearly unchanged.

Additional sensitivity cases in the Supplementary Material further compare independent and fully correlated residual bounds and introduce S3, which combines the S1 epicenter with the uniform $V_{s,30}$ condition of S2, as shown in Supplementary Figure S2.

Epistemic uncertainty associated with Models E and Model EAS is small for both scenarios, as reflected by the narrow spread of exceedance curves across posterior parameter samples. In S2 in particular, epistemic uncertainty is nearly negligible for both models, indicating that parameter uncertainty plays a minimal role and that differences among models are dominated by their treatment of spatial correlation rather than uncertainty in the inferred parameters.

## CONCLUSION

We present a cross-period spatial correlation model for earthquake ground motion within-event residuals built on a principal component representation. The model incorporates path and site effects through a structured covariance framework and quantifies parameter uncertainty using Bayesian inference. Path effects are introduced through epicentral–azimuth differences and site effects through contrasts in $V_{s,30}$, enabling the model to capture directionality and site-similarity.

The principal component representation provides a compact framework for modeling within-event residuals that are correlated across oscillator periods of spectral acceleration. By decomposing the residual field into a small number of orthogonal components, cross-period dependence is absorbed into a few uncorrelated principal components, allowing each principal component to be modeled independently in space. Combined with a physically motivated spatial covariance kernel, this structure guarantees positive-definite covariance matrices, ensuring statistical consistency and numerical stability in regional seismic risk simulations.

Bayesian inference provides posterior distributions for all model parameters, allowing both the quantification of parameter uncertainty and the examination of joint dependencies among them. The results indicate that spatial correlation is primarily controlled by Euclidean inter-station distance, while directional differences, site similarity, and their relative contributions play secondary roles and are associated with greater uncertainty in their estimated effects. Moreover, the leading principal components exhibit spatially coherent structures, whereas higher-order principal components are weakly structured and primarily represent local or uncorrelated variability. Joint posterior analyses reveal the first principal component primarily captures isotropic distance-based effects, while the second principal component reflects directionality and site-similarity interactions, reinforcing the conclusion that inter-station distance remains the dominant driver of spatial correlation.

From a predictive perspective, the anisotropic Model EAS, which explicitly incorporates site and path effects, achieves consistently better predictive performance than the isotropic distance-only Model E. Across both principal components and oscillator periods, the log pointwise predictive density increases by approximately 0.5–2%, demonstrating that the added model complexity meaningfully improves predictive skill and confirming the necessity of accounting for directionality and site similarity in spatial correlation modeling.

A synthetic dataset was used to perform a regional seismic risk assessment under different within-event spatial correlation models, including Model E, Model EAS, GH08, LB13, and M18, as described in Section 3. Two scenarios were designed to isolate the effects of source proximity and site heterogeneity: scenario 1 represents a near-epicenter region with heterogeneous $V_{s,30}$, and scenario 2 represents a far-epicenter region with uniform $V_{s,30}$. This setup allows evaluation of the models under conditions with strong versus weak spatial gradients.

Across both scenarios, mean portfolio losses are broadly consistent across all five within-event spatial correlation models, with differences concentrated in the high-loss tail. In scenario 1, Model EAS produces a substantially lighter tail than isotropic models, indicating that accounting for directional effects and site dissimilarity effectively reduces spatial correlation and limits spurious co-aggregation of extreme losses. For both scenario 1 and scenario 2, epistemic uncertainty associated with Model E and Model EAS is small, suggesting that risk estimates are relatively insensitive to parameter uncertainty when parameters are sampled from their posterior distributions.

The proposed framework could be further refined by exploring alternative spatial correlation functions or by introducing new input metrics, such as finite-fault-informed path-orientation metrics. As part of model development, we also tested a simple alternative definition in which the epicenter-to-site azimuth was replaced by an azimuth measured from the closest point on the rupture plane to each station, but this substitution did not improve predictive fit. Future work could explore more flexible functional forms for the path-orientation term or incorporate additional finite-rupture input metrics, which may better capture the effects of extended source geometry on spatial correlation. Nevertheless, the current formulation provides a statistically rigorous, physically interpretable, and computationally efficient basis with strong predictive capability for modeling spatial correlation of within-event residuals. By improving the representation of spatial correlation in ground motion within-event residuals, the proposed model strengthens the reliability of regional seismic risk assessments and provides a robust analytical basis for decision making and resilience planning.

**Code and data availability**

The data and code used in this study are publicly available at `https://github.com/BofanYu/cross-period-correlation-project`. The code used to generate the case-study portfolio loss results was developed based on the SimCenter R2D framework, with additional modifications implemented for this study. Due to dependencies on the SimCenter software ecosystem, the modified R2D workflow is not fully distributed with the repository; however, additional details are available from the authors upon reasonable request.

**Acknowledgments**

The Bayesian inference implementation in this study was informed by the publicly available code of Bodenmann[41]. We acknowledge this contribution to reproducible research. The visualization style adopted in several figures was inspired by the presentation approach in prior work by Markhvida et al.[20] This work was partially supported by the National Science Foundation under Grant No. CMMI-2053014. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

## REFERENCES

[1] Wesson RL, Perkins DM. Spatial correlation of probabilistic earthquake ground motion and loss. *Bull. Seismol. Soc. Am.* 2001; 91(6): 1498–1515.

[2] Lee R, Kiremidjian AS. Uncertainty and correlation for loss assessment of spatially distributed systems. *Earthq Spectra* 2007; 23(4): 753–770.

[3] Shiraki N, Shinozuka M, Moore JE, Chang SE, Kameda H, Tanaka S. System risk curves: Probabilistic performance scenarios for highway networks subject to earthquake damage. *J. Infrastruct. Syst.* 2007; 13(1): 43–54.

[4] Weatherill GA, 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.* 2015; 13(4): 957–981.

[5] Park J, Bazzurro P, Baker JW. Modeling spatial correlation of ground motion intensity measures for regional seismic hazard and portfolio loss estimation. In: Kanda J, Takada T, Furuta H., eds. *Applications of Statistics and Probability in Civil Engineering*London: Taylor & Francis Group. 2007 (pp. 1–8).

[6] Goda K, Hong HP. Spatial correlation of peak ground motions and response spectra. *Bull. Seismol. Soc. Am.* 2008; 98(1): 354–365.

[7] Jayaram N, Baker JW. Correlation model for spatially distributed ground-motion intensities. *Earthq Eng Struct Dyn.* 2009; 38(15): 1687–1708.

[8] Goda K, Atkinson GM. Intraevent spatial correlation of ground-motion parameters using SK-net data. *Bull. Seismol. Soc. Am.* 2010; 100(6): 3055–3067.

[9] Goda K. Interevent variability of spatial correlation of peak ground motions and response spectra. *Bull. Seismol. Soc. Am.* 2011; 101(5): 2522–2531.

[10] Esposito S, Iervolino I. PGA and PGV spatial correlation models based on European multievent datasets. *Bull. Seismol. Soc. Am.* 2011; 101(5): 2532–2541.

[11] Esposito S, Iervolino I. Spatial correlation of spectral acceleration in European data. *Bull. Seismol. Soc. Am.* 2012; 102(6): 2781–2788.

[12] Chen Y, Baker JW. Spatial correlations in CyberShake physics-based ground-motion simulations. *Bull. Seismol. Soc. Am.* 2019; 109(6): 2447–2458.

[13] 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.* 2020; 203: 103139.

[14] Abbasnejadfard 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.* 2020; 222(2): 1449–1469.

[15] Chen Y, Bradley BA, Baker JW. Nonstationary spatial correlation in New Zealand strong ground-motion data. *Earthq Eng Struct Dyn.* 2021; 50(13): 3421–3440.

[16] Bodenmann L, Baker JW, Stojadinović B. Accounting for path and site effects in spatial ground-motion correlation models using Bayesian inference. *Nat. Hazards Earth Syst. Sci.* 2023; 23(7): 2387–2402.

[17] Loth C, Baker JW. A spatial cross-correlation model of spectral accelerations at multiple periods. *Earthq Eng Struct Dyn.* 2013; 42(3): 397–417.
[18] Du W, Wang G. Fully probabilistic seismic displacement analysis of spatially distributed slopes using spatially correlated vector intensity measures. *Earthq Eng Struct Dyn.* 2014; 43(5): 661–679.
[19] Worden CB, Thompson EM, Baker JW, Bradley BA, Luco N, Wald DJ. Spatial and spectral interpolation of ground-motion intensity measure observations. *Bull. Seismol. Soc. Am.* 2018; 108(2): 866–875.
[20] Markhvida M, Ceferino L, Baker JW. Modeling spatially correlated spectral accelerations at multiple periods using principal component analysis and geostatistics. *Earthq Eng Struct Dyn.* 2018; 47(5): 1107–1123.
[21] Ancheta TD, Darragh RB, Stewart JP, et al. NGA-West2 database. *Earthq Spectra* 2014; 30(3): 989–1005.
[22] Bozorgnia Y, Abrahamson NA, Al Atik L, et al. NGA-West2 research project. *Earthq Spectra* 2014; 30(3): 973–987.
[23] Chiou BSJ, Youngs RR. Update of the Chiou and Youngs NGA model for the average horizontal component of peak ground motion and response spectra. *Earthq Spectra* 2014; 30(3): 1117–1153.
[24] Jayaram N, Baker JW. Considering spatial correlation in mixed-effects regression and the impact on ground-motion models. *Bull. Seismol. Soc. Am.* 2010; 100(6): 3295–3303.
[25] Jayaram N, Baker JW. Statistical tests of the joint distribution of spectral acceleration values. *Bull. Seismol. Soc. Am.* 2008; 98(5): 2231–2243.
[26] Jackson JE. *A user's guide to principal components*. John Wiley & Sons . 2005.
[27] Wackernagel H. *Multivariate geostatistics: an introduction with applications*. Springer Science & Business Media . 2003.
[28] Seeger M. Gaussian Processes for Machine Learning. *International Journal of Neural Systems* 2004; 14(2): 69–106.
[29] Padonou E, Roustant O. Polar Gaussian processes and experimental designs in circular domains. *SIAM/ASA J. Uncertain. Quantif.* 2016; 4(1): 1014–1033.
[30] Bernardo JM, Smith AFM, Berliner M. *Bayesian theory*. 586. Wiley . 1994.
[31] Gelman A, Carlin JB, Stern HS, Rubin DB. *Bayesian data analysis*. Chapman and Hall/CRC . 1995.
[32] Kuehn NM, Stafford PJ. Estimating Ground-Motion Models via Bayesian Inference using Stan. engrXiv preprint; 2021. engrXiv preprint
[33] Liu C, Macedo J, Kuehn N. Spatial correlation of systematic effects of non-ergodic ground motion models in the Ridgecrest area. *Bull. Earthq. Eng.* 2023; 21(11): 5319–5345.
[34] Bradley BA. Epistemic uncertainties in component fragility functions. *Earthquake Spectra* 2010; 26(1): 41–62.
[35] Baker JW, Bradley BA, Stafford PJ. *Seismic hazard and risk analysis*. Cambridge University Press . 2021.
[36] Baker JW, Bradley BA. Intensity measure correlations observed in the NGA-West2 database, and dependence of correlations on rupture and site parameters. *Earthq Spectra* 2017; 33(1): 145–156.
[37] McKenna F, Gavrilovic S, Zhao J, et al. NHERI-SimCenter/R2DTool: Version 5.2.0. Zenodo; 2024
[38] Deierlein GG, McKenna F, Zsarnóczay A, et al. A cloud-enabled application framework for simulating regional-scale impacts of natural hazards on the built environment. *Front. Built Environ.* 2020; 6: 558706. doi: 10.3389/fbuil.2020.558706
[39] Martins L, Silva V. Global Vulnerability Model of the GEM Foundation. `https://github.com/gem/global_vulnerability_model/`; 2023. DOI: 10.5281/zenodo.8391742
[40] GEM Foundation . Fragility Damage Thresholds. `https://docs.openquake.org/vulnerability/main/fragility/fragility_damage_thresholds.html`; 2024. OpenQuake vulnerability documentation.
[41] Bodenmann L. Bayesian parameter estimation for ground-motion correlation models (v1.0.1). Zenodo; 2022. Software