Supplemental Material

Vitor A. Monteiro[1], Savvinos Aristeidou[1], and Gerard J. O'Reilly[1]

[1] *Centre for Training and Research on Reduction of Seismic Risk (ROSE Centre), Scuola Universitaria Superiore IUSS di Pavia, Pavia 27100, Italy*

In this study, spatial and cross-intensity measure (IM) correlations can be modelled using a pairwise-principal component analysis (PCA) framework. For any given pair of two IMs, the model provides a valid correlation structure across sites. However, when extending this pairwise formulation to simulate multiple IMs simultaneously, a fundamental consistency issue arises: the resulting global covariance matrix is not guaranteed to be positive semidefinite (PSD).

Let N denote the number of IMs and M the number of sites. The within-event residual vector can be written as:

$$\delta_W = \begin{bmatrix} [\delta_{W,1}] \\ [\delta_{W,2}] \\ \vdots \\ [\delta_{W,N}] \end{bmatrix} \in \mathbb{R}^{NM} \tag{1}$$

where each $\delta_{W,i} \in \mathbb{R}^M$ represents the spatial residual field of the i^{th} IM. The correlation structure between all IMs and sites is assembled in block form as:

$$R = \begin{bmatrix} R_{1,1} & R_{1,2} & \cdots & R_{1,N} \\ R_{2,1} & R_{2,2} & \cdots & R_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ R_{N,1} & R_{N,2} & \cdots & R_{N,N} \end{bmatrix} \tag{2}$$

where each block $R_{i,j} \in R^{M\times M}$ represents the cross-correlation between IMs i and j across all sites. For each pair i and j, the entries are given as:

$$R_{i,j}(m,n) = \rho_{i,j}(m,n) \tag{3}$$

where $\rho_{i,j}(m,n)$ is the cross-spatial correlation between IMs i and j, and m and n denote site locations. Although each pairwise correlation block $R_{i,j}$ is valid by construction, the full matrix R assembled from independently derived pairwise models is not guaranteed to satisfy the global consistency condition required by correlation matrices, namely positive semi-definiteness:

$$R \succeq 0 \tag{4}$$

This issue arises because pairwise consistency does not imply multivariate consistency. Correlation models calibrated independently for different IM pairs may not be mutually compatible when combined into a single multivariate structure. As a result, the assembled matrix may exhibit negative eigenvalues, making it indefinite. From a simulation perspective, this leads to numerical instability and a breakdown of standard decomposition techniques (e.g., Cholesky factorisation). To address this limitation, the assembled correlation matrix R is corrected using a matrix nearness approach following Higham

[2002]. The nearest valid correlation matrix $\tilde{R}$ is obtained by solving:

$$\min_X \|R - X\|_F \quad \text{subject to } X \succeq 0,\ \text{diag}(X) = \mathbf{1}, \tag{5}$$

where X denotes the corrected correlation matrix, i.e., $X \equiv \tilde{R}$, and $\|\cdot\|_F$ denotes the Frobenius norm, which measures the overall difference between matrices as the square root of the sum of squared element-wise differences. This formulation, therefore, seeks the closest matrix to R that is both positive semi-definite and has unit diagonal. The solution is obtained using the alternating projections algorithm with Dykstra's correction proposed by Higham [2002]. This iterative procedure alternates between the projection onto the set of positive semi-definite matrices and the projection onto the set of matrices with unit diagonal until convergence to the nearest feasible matrix is achieved.

Once the corrected correlation matrix $\tilde{R}$ is obtained, it is substituted in place of R and transformed to covariance form using the within-event standard deviations:

$$\tilde{\Sigma}_{i,j}(m,n) = \phi_i(m)\,\phi_j(n)\,\tilde{R}_{i,j}(m,n), \tag{6}$$

or, equivalently, in matrix form:

$$\tilde{\Sigma} = D^{1/2}\tilde{R}D^{1/2}, \tag{7}$$

where $D = \text{diag}(\phi_i^2(m))$ is a diagonal matrix of size $NM \times NM$, with $\boldsymbol{\phi} \in \mathbb{R}^{NM}$ denoting the vector of within-event standard deviations stacked over all IM-site combinations. The resulting matrix $\tilde{\Sigma}$ is symmetric positive semi-definite and preserves, in a least-squares (Frobenius norm) sense, the original correlation structure implied by the pairwise model. A spectral factorisation of $\tilde{\Sigma}$ is then used for simulation. Specifically, the matrix is decomposed as:

$$\tilde{\Sigma} = Q\Lambda Q^T \tag{8}$$

where Q contains the eigenvectors and Λ is a diagonal matrix of non-negative eigenvalues. The factor matrix is then defines as:

$$F = Q\Lambda^{1/2}, \tag{9}$$

such that:

$$\tilde{\Sigma} = FF^T \tag{10}$$

allowing correlated residuals to be generated as:

$$\delta_W = FZ, \quad Z \sim \mathcal{N}(\mathbf{0}, \mathbf{I}). \tag{11}$$

The need for this correction highlights an inherent limitation of pairwise-based correlation models. While they accurately describe bivariate relationships, they do not inherently guarantee a globally consistent multivariate structure. The application of the Higham [2002] procedure offers a solution that ensures mathematical validity of the covariance matrix while minimally modifying the original model, thereby enabling stable and consistent multi-IM simulations.

References

N. J. Higham. Computing the nearest correlation matrix–a problem from finance. *IMA Journal of Numerical Analysis*, 22:329–343, 7 2002. ISSN 0272-4979. doi: 10.1093/imanum/22.3.329.