Spatial Cross-correlation Models for Vector Intensity Measures (PGA, Ia, PGV and Sa’s) Considering Regional Site Conditions

Gang Wang and Wenqi Du

Department of Civil and Environmental Engineering
Hong Kong University of Science and Technology

March 15, 2013
PEER-GMSM Working Group Meeting
Outlines

• Introduction
• Spatial Correlation of Scalar Intensity Measures
• Strong Motion Database and Regional Site Conditions
• Spatial Cross-correlation of Vector Intensity Measures
• A Site-dependent LMC Model for [PGA, Ia, PGV]
• A Site-dependent LMC Model for Sa(T)
• Applications and Conclusions

The presentation is based on
Introduction

- Modeling spatial variability of ground-motion intensity measures (IMs) is essential for the seismic hazard analysis and risk assessment of spatially distributed infrastructure, such as lifelines, transportation systems and structure portfolios;

- Spatial correlation is not accounted for by GMPEs;

- It is necessary to consider the simultaneous occurrence of multiple intensity measure (Vector IM);

- It is important to consider the influence of regional geological features on the correlation structures.
Introduction

Epicenter and distribution of record stations for the Chi-Chi earthquake

Ground motion prediction equation:

$$\ln Y_{ij} = \ln Y_{ij}(M, R, \theta) + \eta_i + \varepsilon_{ij}$$

- Measured data
- Median prediction
- Inter-event residual
- Intra-event residual

Distribution of intra-event residuals (from Baker)
A Missing Link – Spatial Correlation

The joint probability of occurrence of ground motion residuals in space.

Empirical semi-variograms for intra-event residuals can be developed to measure the dissimilarity of data separated by separation distance \( h \).

\[
\hat{\gamma}(h) = \frac{1}{2|N(h)|} \sum_{i=1}^{N(h)} [z(u_i + h) - z(u_i)]^2
\]

The closer the separation distance, the higher probability they are similar.

An exponential model can be used to fit the semivariograms.

\[
\gamma(h) = a(1 - \exp(-3h / b))
\]

\( a \): the sill of correlation
\( b \): the range of correlation
\( h \): the separation distance

\( b = 24 \text{ km} \)
Spatial Correlation of Scalar Intensity Measures

- An exponential model can be used to fit the semivariograms.
  \[ \gamma(h) = a(1 - \exp(-3h / b)) \]

- Spatial correlation coefficient
  \[ \rho(h) = \exp(-3h / b) \]

- The valid spatial correlation matrix is always (symmetric) positive semi-definite
  \[
  \begin{bmatrix}
  1 & \rho(h_{12}) & \cdots & \rho(h_{1J}) \\
  \rho(h_{12}) & 1 & \cdots & \\
  \vdots & \ddots & \ddots & \\
  \rho(h_{1J}) & \rho(h_{2J}) & \cdots & 1
  \end{bmatrix} \geq 0
  \]

Valid correlation matrix

- \( a \): the sill of correlation
- \( b \): the range of correlation
- \( h \): the separation distance

\( b = 24 \text{ km} \)
### Strong Motion Database

<table>
<thead>
<tr>
<th>Earthquake name</th>
<th>Date (dd/mm/yyyy)</th>
<th>Moment magnitude</th>
<th>Location</th>
<th>Fault mechanism</th>
<th>Num. of recordings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northridge</td>
<td>01/17/1994</td>
<td>6.69</td>
<td>California</td>
<td>Reverse</td>
<td>152</td>
</tr>
<tr>
<td>Chi-Chi</td>
<td>09/20/1999</td>
<td>7.62</td>
<td>Taiwan</td>
<td>Reverse-oblique</td>
<td>401</td>
</tr>
<tr>
<td>Tottori</td>
<td>10/06/2000</td>
<td>6.61</td>
<td>Japan</td>
<td>Strike-slip</td>
<td>235</td>
</tr>
<tr>
<td>Parkfield</td>
<td>09/28/2004</td>
<td>6</td>
<td>California</td>
<td>Strike-slip</td>
<td>90</td>
</tr>
<tr>
<td>Anza</td>
<td>06/12/2005</td>
<td>5.2</td>
<td>California</td>
<td>Reverse-oblique</td>
<td>111</td>
</tr>
<tr>
<td>Chuetsu</td>
<td>07/16/2007</td>
<td>6.8</td>
<td>Japan</td>
<td>Reverse</td>
<td>401</td>
</tr>
<tr>
<td>Alum Rock</td>
<td>10/30/2007</td>
<td>5.4</td>
<td>California</td>
<td>Strike-slip</td>
<td>161</td>
</tr>
<tr>
<td>Iwate</td>
<td>06/13/2008</td>
<td>6.9</td>
<td>Japan</td>
<td>Reverse</td>
<td>279</td>
</tr>
<tr>
<td>Chino Hills</td>
<td>07/29/2008</td>
<td>5.4</td>
<td>California</td>
<td>Reverse-oblique</td>
<td>337</td>
</tr>
<tr>
<td>El Mayor</td>
<td>04/04/2010</td>
<td>7.2</td>
<td>Mexico</td>
<td>Strike-slip</td>
<td>154</td>
</tr>
</tbody>
</table>

- Eleven well-recorded earthquakes (2686 records) are used to investigate the spatial correlation of PGA, Ia, PGV and Sa(T).

Magnitude and rupture distance distribution of records in the database.

Note: only recorded data within rupture distance of 200 km are included for Japan earthquakes.
• The trend of residuals versus rupture distance and $V_{S30}$ should be corrected to avoid artificial correlation.

$$
\varepsilon_{corr} = \ln Y_{ij} - \ln Y_{ij}(M, R, \theta) - \left( \varphi_1 + \varphi_2 \ln(R_{ij}) + \varphi_3 \ln(V_{S30}) \right)
$$
Regional Site Conditions

Vs category
- B
- BC
- C
- CD
- D
- DE
- E
- WATER

Northridge

Heterogeneous

Chi-Chi

Homogeneous
The correlation range of normalized $V_{s30}$ values ($R_{Vs30}$) are used to quantify the regional site conditions.

$R_{Vs30} = 0$ km

Heterogeneous

$R_{Vs30} = 26$ km

Homogeneous
Regional Site Conditions

- Influence of inferred Vs30 data.

- A redistributed procedure is applied considering the uncertainty of Vs30 data to reduce artificial correlation induced by inferred data.
Regional Site Conditions

Anza earthquake

Chino Hills earthquake

Randomized data
### Regional Site Conditions

<table>
<thead>
<tr>
<th>Earthquake name</th>
<th>Median $V_{s30}$ (m/s)</th>
<th>Std. dev. $V_{s30}$ (m/s)</th>
<th>Correlation range $R_{Vs30}$ (km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northridge</td>
<td>422</td>
<td>218</td>
<td>0</td>
</tr>
<tr>
<td>Chi-Chi</td>
<td>384</td>
<td>178</td>
<td>26</td>
</tr>
<tr>
<td>Tottori</td>
<td>425</td>
<td>174</td>
<td>18.8</td>
</tr>
<tr>
<td>Parkfield</td>
<td>395</td>
<td>131</td>
<td>3.5</td>
</tr>
<tr>
<td>Niigata</td>
<td>404</td>
<td>168</td>
<td>21.8</td>
</tr>
<tr>
<td>Anza</td>
<td>348</td>
<td>118</td>
<td>20.3</td>
</tr>
<tr>
<td>Chuetsu</td>
<td>415</td>
<td>167</td>
<td>20.8</td>
</tr>
<tr>
<td>Alum Rock</td>
<td>386</td>
<td>144</td>
<td>14.2</td>
</tr>
<tr>
<td>Iwate</td>
<td>407</td>
<td>176</td>
<td>8.7</td>
</tr>
<tr>
<td>Chino Hills</td>
<td>342</td>
<td>101</td>
<td>14.5</td>
</tr>
<tr>
<td>El Mayor Cucapah</td>
<td>422</td>
<td>180</td>
<td>20.3</td>
</tr>
</tbody>
</table>

$R_{Vs30}$ is closely related to the correlation range of scalar IMs.

(Du and Wang, BSSA 2013, in press)
Spatial Cross-correlation of Vector Intensity Measures

Cross-correlation matrix

<table>
<thead>
<tr>
<th></th>
<th>$IM_1(x_1)$</th>
<th>$IM_2(x_1)$</th>
<th>$IM_1(x_2)$</th>
<th>$IM_2(x_2)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$IM_1(x_1)$</td>
<td>$\rho_{11}(0)$</td>
<td>$\rho_{12}(0)$</td>
<td>$\rho_{11}(h_{12})$</td>
<td>$\rho_{12}(h_{12})$</td>
</tr>
<tr>
<td>$IM_2(x_1)$</td>
<td>$\rho_{12}(0)$</td>
<td>$\rho_{22}(0)$</td>
<td>$\rho_{12}(h_{12})$</td>
<td>$\rho_{22}(h_{12})$</td>
</tr>
<tr>
<td>$IM_1(x_2)$</td>
<td>$\rho_{11}(h_{12})$</td>
<td>$\rho_{12}(h_{12})$</td>
<td>$\rho_{11}(0)$</td>
<td>$\rho_{12}(0)$</td>
</tr>
<tr>
<td>$IM_2(x_2)$</td>
<td>$\rho_{12}(h_{12})$</td>
<td>$\rho_{22}(h_{12})$</td>
<td>$\rho_{12}(0)$</td>
<td>$\rho_{22}(0)$</td>
</tr>
</tbody>
</table>

- Cross spatial correlation between $IM_1$ and $IM_2$ at separation distance $h_{12}$
- Given an $n$-component vector $IM$ distributed at $J$ sites, the total correlation matrix is $[Jxn, Jxn]$ in dimension

$$
\begin{bmatrix}
R(0) & \cdots & R(h_{1,J}) \\
\vdots & \ddots & \vdots \\
R(h_{J1}) & \cdots & R(0)
\end{bmatrix}
$$
Linear Model of Coregionalization for Vector IMs

- A combination of a short range and a long range exponential basic function are selected to fit empirical data (LMC)

\[
R(h) = P^1 \left( \exp \left( \frac{-3h}{r_1} \right) \right) + P^2 \left( \exp \left( \frac{-3h}{r_2} \right) \right)
\]

- As long as \( P^1 \) and \( P^2 \) are positive semi-definite, the total correlation matrix is guaranteed to be positive semi-definite regardless of sites considered (a permissible/valid LMC model).

\[
P^i = \begin{bmatrix}
p_{11} & p_{12} & p_{13} \\
p_{21} & p_{22} & p_{23} \\
p_{31} & p_{32} & p_{33}
\end{bmatrix}
\]

\[
R(0) \quad \cdots \quad R(h_{1J}) \\
\vdots \quad \ddots \quad \vdots \\
R(h_{J1}) \quad \cdots \quad R(0)
\]

Valid correlation matrix

- The \( P^1 \) and \( P^2 \) matrices can be obtained from each earthquake
Cross-semivariograms and fitted LMC curves obtained for the Niigata earthquake

\[ R(h) = P_1 \exp \left( \frac{-3h}{10} \right) + P_2 \exp \left( \frac{-3h}{60} \right) \]
Influence of Site Condition on LMC Matrices

\[ R(h) = P^1 \left( \exp \left( \frac{-3h}{10} \right) \right) + P^2 \left( \exp \left( \frac{-3h}{60} \right) \right) \]

Short range \quad Long range

\[
P^i = \begin{bmatrix}
p_{11}^i & p_{12}^i & p_{13}^i \\
p_{21}^i & p_{22}^i & p_{23}^i \\
p_{31}^i & p_{32}^i & p_{33}^i \\
\end{bmatrix}
\]
Influence of Site Condition on LMC Matrices

\[
P^1 = P^0 - K \cdot R_{Vs30}
\]
\[
P^2 = K \cdot R_{Vs30}
\]

\[
K = \begin{bmatrix}
  0.28 & 0.24 & 0.17 \\
  0.24 & 0.22 & 0.16 \\
  0.17 & 0.16 & 0.31 
\end{bmatrix}
\]

\[
P^0 = \begin{bmatrix}
  1 & 0.91 & 0.65 \\
  0.91 & 1 & 0.71 \\
  0.65 & 0.71 & 1 
\end{bmatrix}
\]
A Site-dependent LMC Model for [PGA, Ia, PGV]

- Site-dependent LMC model

\[
R(h, R_{v_s30}) = P^1(R_{v_s30}) \left( \exp\left(\frac{-3h}{10}\right) \right) + P^2(R_{v_s30}) \left( \exp\left(\frac{-3h}{60}\right) \right)
\]

\[
P^1 = P^0 - K \left( \frac{R_{v_s30}}{10} \right)
\]

\[
P^2 = K \left( \frac{R_{v_s30}}{10} \right)
\]

A permissible LMC model

Positive definite for \( R_{vs30} \leq 25 \text{ km} \)

Positive definite

Positive definite

K = \[
\begin{bmatrix}
0.28 & 0.24 & 0.17 \\
0.24 & 0.22 & 0.16 \\
0.17 & 0.16 & 0.31 \\
\end{bmatrix}
\]

\[
P^0 = \[
\begin{bmatrix}
1 & 0.91 & 0.65 \\
0.91 & 1 & 0.71 \\
0.65 & 0.71 & 1 \\
\end{bmatrix}
\]
A Site-dependent LMC Model for [PGA, Ia, PGV]

- **Site-dependent LMC model**

\[
R(h, R_{Vs30}) = \left[ P^0 - K \left( \frac{R_{Vs30}}{10} \right) \right] \exp\left( \frac{-3h}{10} \right) + K \left( \frac{R_{Vs30}}{10} \right) \exp\left( \frac{-3h}{60} \right)
\]

- **Examples**

\[
R(h = 5, R_{Vs30} = 20) = \begin{bmatrix}
0.53 & 0.47 & 0.33 \\
0.47 & 0.47 & 0.34 \\
0.33 & 0.34 & 0.57 \\
\end{bmatrix}
\]

\[
R(h = 5, R_{Vs30} = 10) = \begin{bmatrix}
0.38 & 0.34 & 0.24 \\
0.34 & 0.35 & 0.25 \\
0.24 & 0.25 & 0.40 \\
\end{bmatrix}
\]

**Influence of site conditions**

\[
K = \begin{bmatrix}
0.28 & 0.24 & 0.17 \\
0.24 & 0.22 & 0.16 \\
0.17 & 0.16 & 0.31 \\
\end{bmatrix}
\]

Weaker

Stronger
A Site-dependent LMC Model for [PGA, Ia, PGV]

- **Site-dependent LMC model**

\[
R(h, R_{Vs30}) = \begin{bmatrix} P^0 - K \left( \frac{R_{Vs30}}{10} \right) \exp \left( \frac{-3h}{10} \right) \\
K \left( \frac{R_{Vs30}}{10} \right) \exp \left( \frac{-3h}{60} \right) \end{bmatrix}
\]

- **Reduce to local correlation matrix (h=0)**

\[
R(h = 0) \overset{\Delta}{=} R(0) = P^0 = \begin{bmatrix} 1 & 0.91 & 0.65 \\
0.91 & 1 & 0.71 \\
0.65 & 0.71 & 1 \end{bmatrix}
\]

Compare with Campbell and Bozorgnia (2012):

\[
\begin{bmatrix} \rho_{PGA,PGA} & \rho_{PGA,Ia} & \rho_{PGA,PGV} \\
\rho_{PGA,Ia} & \rho_{Ia,Ia} & \rho_{Ia,PGV} \\
\rho_{PGA,PGV} & \rho_{Ia,PGV} & \rho_{PGV,PGV} \end{bmatrix} = \begin{bmatrix} 1 & 0.88 & 0.69 \\
0.88 & 1 & 0.74 \\
0.69 & 0.74 & 1 \end{bmatrix}
\]
A Site-dependent LMC Model for [PGA, Ia, PGV]

- Site-dependent LMC model

\[
R(h, R_{Vs30}) = \left[ P^0 - K \left( \frac{R_{Vs30}}{10} \right) \right] \exp \left( \frac{-3h}{10} \right) + K \left( \frac{R_{Vs30}}{10} \right) \exp \left( \frac{-3h}{60} \right)
\]

- Reduce to heterogeneous site conditions \((R_{Vs30}=0)\)

\[
R(h) = R(0) \exp \left( \frac{-3h}{10} \right)
\]

- Averaged LMC model for [PGA, Ia, PGV]

\[
R(h) = P^1_{avg} \left( \exp \left( \frac{-3h}{10} \right) \right) + P^2_{avg} \left( \exp \left( \frac{-3h}{60} \right) \right)
\]

\[
P^1_{avg} = \begin{bmatrix} 0.61 & 0.57 & 0.38 \\ 0.57 & 0.67 & 0.45 \\ 0.38 & 0.45 & 0.50 \end{bmatrix} \quad P^2_{avg} = \begin{bmatrix} 0.39 & 0.34 & 0.24 \\ 0.34 & 0.33 & 0.24 \\ 0.24 & 0.24 & 0.50 \end{bmatrix}
\]

\[\text{avg} \]
Model Predictions -- Northridge Earthquake

\[
R(h) = P^1 \left( \exp\left(\frac{-3h}{10}\right) \right) + P^2 \left( \exp\left(\frac{-3h}{60}\right) \right)
\]

○ Empirical semivariogram
- Predicted curve using site–dependent matrices
- Predicted curve using averaged matrices
Model Predictions – Chi-Chi Earthquake

\[ R(h) = P^1 \left( \exp \left( \frac{-3h}{10} \right) \right) + P^2 \left( \exp \left( \frac{-3h}{60} \right) \right) \]

Empirical semivariogram
- Predicted curve using site-dependent matrices
- Predicted curve using averaged matrices

PGAs

PGVs

IAs
A Site-dependent LMC Model for Sa(T)

- Site-dependent LMC model for Sa(T):
  \[
  R(h, R_{Vs30}) = P^1 \left( \exp \left( \frac{-3h}{10} \right) \right) + P^2 \left( \exp \left( \frac{-3h}{70} \right) \right)
  \]

  \[
  P^1 = P_{Sa}^0 - K_{Sa} \left( \frac{R_{Vs30}}{10} \right)
  \]

  \[
  P^2 = P_{Sa}^0 + K_{Sa} \left( \frac{R_{Vs30}}{10} \right)
  \]

- Averaged LMC model (Loth and Baker, 2013)
  \[
  R(h) = P^1 \left( \exp \left( \frac{-3h}{20} \right) \right) + P^2 \left( \exp \left( \frac{-3h}{70} \right) \right) + P^3 I_{h=0}
  \]

A Site-dependent LMC Model for Sa(T)

<table>
<thead>
<tr>
<th>Period (s)</th>
<th>0.01</th>
<th>0.1</th>
<th>0.2</th>
<th>0.5</th>
<th>1</th>
<th>2</th>
<th>5</th>
<th>7.5</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>0.96</td>
<td>0.96</td>
<td>0.95</td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>0.9</td>
<td>0.8</td>
<td>0.81</td>
<td>0.76</td>
<td>0.62</td>
<td>0.54</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>0.8</td>
<td>0.81</td>
<td>0.83</td>
<td>0.75</td>
<td>0.45</td>
<td>0.54</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.3</td>
<td>0.7</td>
<td>0.65</td>
<td>0.75</td>
<td>0.71</td>
<td>0.34</td>
<td>0.42</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.4</td>
<td>0.6</td>
<td>0.55</td>
<td>0.66</td>
<td>0.67</td>
<td>0.32</td>
<td>0.43</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>0.5</td>
<td>0.46</td>
<td>0.56</td>
<td>0.62</td>
<td>0.31</td>
<td>0.42</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Positive definite for \( R_{vs30} \leq 25 \) km

\[
P^1 = P_{Sa}^{01} - K_{Sa} \left( \frac{R_{Vs30}}{10} \right)
\]

\[
P^2 = P_{Sa}^{02} + K_{Sa} \left( \frac{R_{Vs30}}{10} \right)
\]

The model can be interpolated for other periods and still remains valid.

\[
R = P^1 \left( \exp \left( \frac{-3h}{10} \right) \right) + P^2 \left( \exp \left( \frac{-3h}{70} \right) \right)
\]

Permissible LMC model
Model Prediction – Northridge Earthquake

\[ R(h) = P_1 \left( \exp \left( \frac{-3h}{10} \right) \right) + P_2 \left( \exp \left( \frac{-3h}{60} \right) \right) \]

**Empirical semivariogram**
- Blue line: Predicted curve by this study
- Red dashed line: Predicted curve by Loth and Baker (2013)
Model Prediction – Chi-Chi Earthquake

\[ R(h) = P^1 \left( \exp \left( \frac{-3h}{10} \right) \right) + P^2 \left( \exp \left( \frac{-3h}{60} \right) \right) \]

- Empirical semivariogram
- Predicted curve by this study
- Predicted curve by Loth and Baker (2013)
Model Applications -- Random Vector IM Fields

$R_{Vs30} = 0 \text{ km}$

$R_{Vs30} = 25 \text{ km}$

(PGA, la and PGV are in the natural log scale, in the unit of g, m/s and cm/s, respectively; $Vs30$ is in the unit of m/s)
Model Applications -- Fully Probabilistic Approach using Spatially-correlated Vector IMs

Conclusions

• Simple permissible spatial correlation models are developed for vector IMs (PGA, Ia, PGV, and Sa) using eleven recent earthquakes.

• The correlation range of $V_{s30}$, $R_{V_{s30}}$, is found to be a good indicator to characterize the regional geological conditions. In general, the spatial correlations of IMs becomes stronger for a homogeneous regional site condition.

• The spatial correlation models can be conveniently used in regional-specific seismic hazard analysis and loss estimation of spatially-distributed infrastructure.
Questions?