Ground-motion intensity measure correlations observed in Italian data

Chen Huang
PhD candidate, Dept. of Civil, Environmental and Geomatic Engineering, University College London, London, UK

Carmine Galasso
Associate Professor, Dept. of Civil, Environmental and Geomatic Engineering, University College London, London, UK

ABSTRACT: Ground-motion models (GMMs) are used in probabilistic seismic hazard analysis (PSHA) to estimate the probability distribution of earthquake-induced ground-motion intensity measures (IMs). Accounting for spatial correlation and cross-IM correlation in ground-motion data has important implications on seismic hazard and risk assessment outputs. The current practice estimates the spatial correlation separately from the GMM estimation process, which may result in inconsistent and inefficient estimators of parameters in the spatial correlation models and GMMs. Moreover, several correlation models between different IMs have been calibrated and validated based on the NGA-West and NGA-West2 databases and advanced GMMs. However, modeling the correlation between different IM types has not been adequately addressed by current, state-of-the-art GMMs for Italy. To address those issues, this study first develops a series of new Italian GMMs with spatial correlation for 31 amplitude-related IMs, including peak ground acceleration (PGA) and peak ground velocity (PGV) and 5% damped elastic pseudo-spectral accelerations (PSA) at 29 periods ranging from 0.01 s to 4 s. The model estimation is performed through a recently-developed one-stage non-linear regression algorithm proposed by the authors, known as the Scoring estimation approach. Based on the newly-developed GMMs, this study finally proposes a set of analytical correlation models between the selected IMs for the considered Italian dataset.

1. INTRODUCTION

Ground-motion models (GMMs) are empirical models describing the probability distributions of intensity measures (IMs) at a site, given an earthquake of a certain magnitude occurred at a nearby location. GMMs are widely used in probabilistic seismic hazard analysis (PSHA). The dependence between various IMs from a single event at multiple sites plays a crucial role in PSHA of spatially-distributed systems (e.g., portfolios of structures and lifelines). Such a dependence is due to a common source and a wave traveling paths and due to the similar distance to fault asperities (Park et al., 2007). Jayaram and Baker (2009) and Weatherill et al. (2015), among others, have shown that the spatial correlation in ground motions has important implications for seismic hazard/risk assess-ment. The cross-IM correlation is often required for the performance-based seismic design and assessment of structures, for instance, in the definition of target IMs to be used for ground-motion simulation, selection, and modification (e.g., the generalized conditional intensity measure or GCIM (Bradley, 2010), the conditional spectrum or CS (Lin and Baker, 2015)). These correlations can be addressed in the GMM estimation stage or in a subsequent stage after the GMMs are estimated.

A typical GMM is presented as a mixed-effect nonlinear model with a certain spatial correlation structure (Jayaram and Baker, 2010) and can be written in a vector form as in Eq.(1),

$$\mathbf{Y}_i = \mathbf{f}(\mathbf{X}_i, \mathbf{b}) + \boldsymbol{\eta}_i + \boldsymbol{\varepsilon}_i, \quad i = 1, \dots, N, \quad (1)$$

where $\mathbf{Y}_i = \log_{10} \mathbf{IM}_i = (\log_{10} \mathrm{IM}_{i1}, \dots, \log_{10} \mathrm{IM}_{in_i})^\top$ is

an $n_i \times 1$ logarithmic IM vector at sites $j \in \{1, ..., n_i\}$ of earthquake i; $\mathbf{f}(\mathbf{X}_i, \mathbf{b}) = [f(\mathbf{X}_{i1}, \mathbf{b}), ..., f(\mathbf{X}_{in_i}, \mathbf{b})]^\top$ is an $n_i \times 1$ vector of ground-motion prediction functions $f(\mathbf{X}_{ij}, \mathbf{b})$ of coefficients $\mathbf{b}$ and predictors $\mathbf{X}_{ij}$ (e.g., magnitude and distance); $\boldsymbol{\eta}_i = \eta_i \mathbf{1}_{n_i}$ and $(\eta_i)_{i=1,...,N}$ are independent and identically distributed inter-event errors with $\mathbb{E}(\eta_i) = 0$ and $\text{var}(\eta_i) = \tau^2$ for all $i \in \{1, ..., N\}$ with $\mathbf{1}_{n_i}$ is an $n_i \times 1$ vector of ones; $(\boldsymbol{\varepsilon}_i)_{i=1,...,N}$ are $n_i \times 1$ independent intra-event error vectors with $\mathbb{E}(\boldsymbol{\varepsilon}_i) = \mathbf{0}$ and $\text{cov}(\boldsymbol{\varepsilon}_i) = \sigma^2 \boldsymbol{\Omega}_i(\boldsymbol{\omega})$, where $\boldsymbol{\Omega}_i(\boldsymbol{\omega})$ is the correlation matrix of earthquake i with unknown parameters $\boldsymbol{\omega}$; $(\eta_i)_{i=1,...,N}$ and $(\boldsymbol{\varepsilon}_i)_{i=1,...,N}$ are assumed to be mutually independent; N is the number of earthquakes; n_i is the number of stations for earthquake i.

To account for the spatial correlation, the jj'-th entry of $\boldsymbol{\Omega}_i(\boldsymbol{\omega})$, $\boldsymbol{\Omega}_{i,jj'}(\boldsymbol{\omega})$, is

$$\boldsymbol{\Omega}_{i,jj'}(\boldsymbol{\omega}) = \rho\left(\varepsilon_{ij}(\mathbf{s}_{ij}), \varepsilon_{ij'}(\mathbf{s}_{ij'})\right), \qquad (2)$$

where $\rho\left(\varepsilon_{ij}(\mathbf{s}_{ij}), \varepsilon_{ij'}(\mathbf{s}_{ij'})\right)$ represents the spatial correlation between ε_{ij} and $\varepsilon_{ij'}$ at locations $\mathbf{s}_{ij}$ and $\mathbf{s}_{ij'}$ of stations j and j' during earthquake i. Assuming the field of $\boldsymbol{\varepsilon}_i$ is stationary and isotropic, the spatial correlation only depends on the distance between stations j and j', $d_{i,jj'}$, such that,

$$\rho\left(\varepsilon_{ij}(\mathbf{s}_{ij}), \varepsilon_{ij'}(\mathbf{s}_{ij'})\right) = \rho(d_{i,jj'}). \qquad (3)$$

There are many correlation functions available in the literature (e.g., Rasmussen and Williams, 2006), one example is the exponential model,

$$\rho(d_{i,jj'}) = \exp\left(-d_{i,jj'}/h\right), \qquad (4)$$

where h is the range parameter in km, at which the spatial correlation is around 0.37.

Numerous GMMs are developed for Europe and Italy, including Akkar and Bommer (2010); Bindi et al. (2014) for Europe and Bindi et al. (2011) for Italy (hereafter, AB10, BMLA14 and ITA10). However, due to the complexity in the estimation process, these GMMs do not account for the spatial correlation in their analysis (i.e., $\boldsymbol{\Omega}_i(\boldsymbol{\omega}) = \boldsymbol{I}_i$, where $\boldsymbol{I}_i$ is an $n_i \times n_i$ identity matrix for earthquake i).

To evaluate the spatial correlation in European and Italian ground motions, Esposito and Iervolino (2012) have studied the semivariograms computed from the residuals of AB10 and ITA10, respectively. The spatial correlation function they used is

similar to Eq. (4) but with a factor three in the numerator, which estimates the effective range $\tilde{h}$, the distance at which the spatial correlation is around 0.05. The relation between the effective range $\tilde{h}$ and the range parameter h is $\tilde{h} = 3h$ (Zimmerman and Michael, 2010). Moreover, Esposito and Iervolino (2012) developed a predictive model for $\tilde{h}$ as a function of the structural period T.

To incorporate the spatial correlation into the GMM estimation, Jayaram and Baker (2010) adopted the framework of the classical geostatistical method (Zimmerman and Michael, 2010) and proposed a multi-stage algorithm, which, however, may result in an inconsistent estimator of $\boldsymbol{\omega}$, thus, resulting in estimators of $\mathbf{b}$ (although consistent) statistically inefficient and estimators of τ^2 and σ^2 both inconsistent and statistically inefficient. Also, the multi-stage algorithm may suffer from slow convergence and it is sensitive to the initial parameter values. Furthermore, the multi-stage algorithm can not account for more advanced (e.g., non-stationary) spatial correlation functions. Recently, Ming et al. (2019) introduced a one-stage nonlinear algorithm for GMMs with spatial correlation, also known as the Scoring estimation approach. This method is proved to be statistically rigorous, numerically stable, and capable of estimations of various spatial correlation models.

Once the GMMs with spatial correlation are estimated, they can be used to develop analytical cross-IM correlation models. Several cross-IM correlation models have been calibrated and validated based on the NGA-West and NGA-West2 databases (Baker and Cornell, 2006; Bradley, 2011, 2012; Baker and Bradley, 2017) and advanced GMMs. Furthermore, Cimellaro (2013) studied the correlation in spectral ordinates for Europe based on 595 records of events from 1973 to 2003 in Europe and Middle East with $M_w \geq 5$ and $R_{JB} \leq 100$ km, using the GMMs without spatial correlation proposed by Ambraseys et al. (2005). However, modeling the correlation between different IM types has not been adequately addressed by current, state-of-the-art ground-motion models for Italy.

Based on the discussion above, this study first develops a series of new Italian GMMs with spatial

correlation for various amplitude-based IMs. The model estimation is performed through the Scoring estimation approach recently proposed by the authors. Given these newly-developed GMMs, this study finally proposes a set of analytical correlation models between the selected IMs for the Italian dataset. It is worth noting that the developed one-stage estimation approach can also account for the cross-IM correlation and this feature is currently under investigation by the authors, especially in terms of implications on the GMM estimates.

The Italian case is of special interest in this study because the Italian data is principally from earthquakes in extensional regions that are poorly represented in global databases (Scasserra et al., 2009) and past practice in Italy has been to use local GMMs based on limited datasets that cannot resolve many significant source, path, and site effects.

2. DATASET

The dataset is extracted from the European Strong-Motion (ESM) flatfile (Lanzano et al., 2018) and the following selection criteria are applied,

- events occurred within Italy;
- events with moment magnitude $M_w \geq 4$
- events with at least two recording sites.
- recording station with Joyner-Boore distance[1] (i.e., the closest distance to the surface projection of the rupture plane) $R_{JB} \leq 250$ km;
- recording stations are free-field;
- removing records without information of M_w, fault types, or V_{S30} (i.e., the average shear-wave velocity in the upper 30 m of the soil);
- remove stations with redundant site information (e.g., co-located sites).

The selected dataset includes 7,843 records from 233 earthquakes of $4 \leq M_w \leq 6.9$ in Italy from 1976 to 2016. The geographical distribution of the selected dataset is shown in Figure 1, together with the $M_w - R_{JB}$ distribution and the site classifications according to Eurocode 8 (CEN, 2004). 66% of the selected data are caused by the rupture of normal

---

[1] If the finite-fault model is available, R_{JB} is computed based on the fault geometry from ESM; if not, for $M_w > 5.5$, R_{JB} is estimated from epicenter distance R_{epi} by empirical relationship (Stucchi et al., 2011), otherwise the earthquake source is assumed to be a point source and $R_{JB} = R_{epi}$.
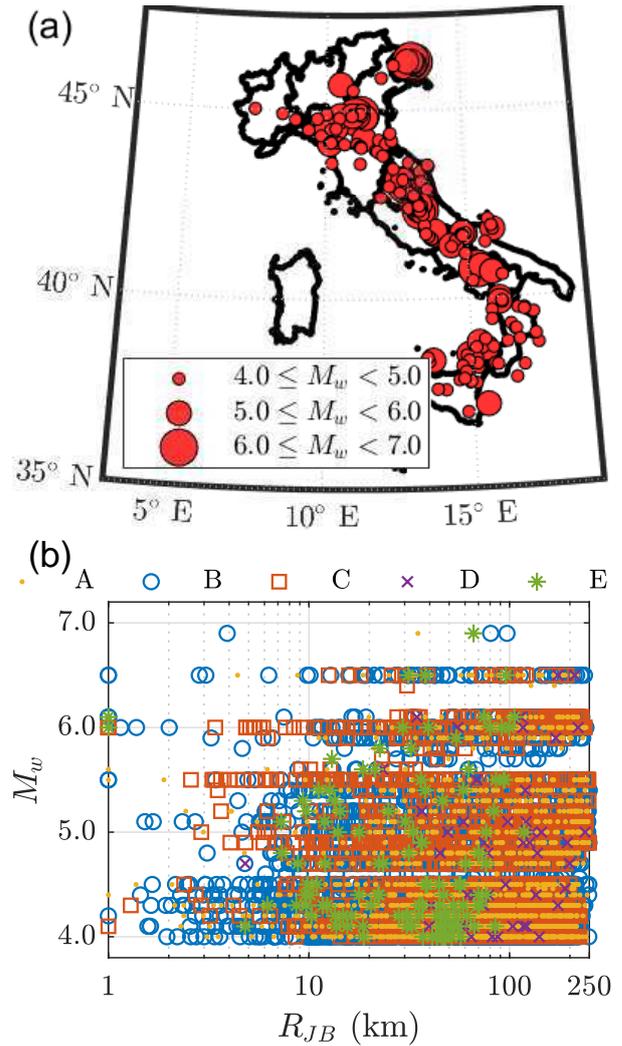
Figure 1: (a) Geographical distribution of 233 events $4 \leq M_w \leq 6.9$ in Italy from 1976 to 2016. The epicenters are in circles, whose size is scaled by event M_w. (b) M_w-R_{JB} distribution. The EC8 site classification is shown: '·' for class A ; '∘' for class B; '□' for class C; '×' for class D; '∗' for class E.

faults, 23% caused by the reverse faults and 11% by strike-slip faults; most data is collected from stations of site class B/stiff soil.

3. METHODOLOGY

3.1. Model specification

The GMMs are developed for 31 amplitude-based IMs: peak ground acceleration (PGA) in cm/s², peak ground velocity (PGV) in cm/s and 5% damped elastic pseudo-spectral accelerations (PSA) in cm/s² at 29 periods ranging from 0.01 s to 4 s. The RotD50 IM definition proposed by Boore (2010) is used, consisting of the median single-component horizontal ground motion across

all non-redundant azimuths. It is noted that AB10, ITA10 and BMLA14 used the geometric mean IM, however, the difference between the two definitions is limited (Boore, 2010).

For consistency with Ming et al. (2019), the functional form, $f(\mathbf{X}_{ij}, \mathbf{b})$, is the same as AB10 and is similar to ITA10, as follows,

$$
\begin{aligned}
f(\mathbf{X}_{ij}, \mathbf{b}) = {} & b_1 + b_2 M_{w,i} + b_3 M_{w,i}^2 \\
& + (b_4 + b_5 M_{w,i}) \log_{10} \sqrt{R_{JB,ij}^2 + b_6^2} \\
& + b_7 S_{S,ij} + b_8 S_{A,ij} + b_9 F_{N,i} + b_{10} F_{R,i},
\end{aligned} \tag{5}
$$

where M_w is the moment magnitude; R_{JB} is the Joyner-Boore distance in km; S_S and S_A are the dummy variables for soil type, such that $(S_S, S_A) = (1,0)$ for soft soil, $(S_S, S_A) = (0,1)$ for stiff soil and $(S_S, S_A) = (0,0)$ for rock; F_N and F_R are the dummy variables for style-of-faulting, such that $(F_N, F_R) = (1,0)$ for normal fault, $(F_N, F_R) = (0,1)$ for reverse fault and $(F_N, F_R) = (0,0)$ for strike-slip fault. The spatial correlation function is set as Eq.(4).

3.2. Estimation algorithm

The Scoring estimation approach, which is a modified Newton–Raphson algorithm, is briefly introduced here. Details can be found in Ming et al. (2019). The unknown parameters in the GMM, denoted by $\boldsymbol{\alpha} = (\mathbf{b}^\top, \tau^2, \sigma^2, \boldsymbol{\omega}^\top)^\top$ as the complete vector of model parameters are estimated by maximizing the log-likelihood function, $l(\boldsymbol{\alpha})$, as follows,

$$
\begin{aligned}
l(\boldsymbol{\alpha}) = {} & -\frac{\sum_{i=1}^N n_i}{2} \ln(2\pi) - \frac{1}{2} \ln |\mathbf{C}(\boldsymbol{\theta})| \\
& - \frac{1}{2} [\mathbf{Y} - \mathbf{f}(\mathbf{X}, \mathbf{b})]^\top \mathbf{C}^{-1}(\boldsymbol{\theta}) [\mathbf{Y} - \mathbf{f}(\mathbf{X}, \mathbf{b})]
\end{aligned} \tag{6}
$$

where $\mathbf{Y} = (\mathbf{Y}_1^\top, ..., \mathbf{Y}_N^\top)^\top$; covariance matrix $\mathbf{C}(\boldsymbol{\theta})$ is a block diagonal matrix of $\mathbf{C}_i = \tau^2 \mathbf{1}_{n_i \times n_i} + \sigma^2 \boldsymbol{\Omega}_i(\boldsymbol{\omega})$; $\mathbf{f}(\mathbf{X}, \mathbf{b}) = [\mathbf{f}(\mathbf{X}_1, \mathbf{b})^\top, ..., \mathbf{f}(\mathbf{X}_N, \mathbf{b})^\top]^\top$.

The Scoring estimation approach finds the estimate of $\boldsymbol{\alpha}$ that maximizes $l(\boldsymbol{\alpha})$ in equation (6) via the general updating equation in equation (7):

$$
\widehat{\boldsymbol{\alpha}}^{(k+1)} = \widehat{\boldsymbol{\alpha}}^{(k)} + \mathbf{I}^{-1}(\widehat{\boldsymbol{\alpha}}^{(k)}) \mathbf{S}(\widehat{\boldsymbol{\alpha}}^{(k)}) \tag{7}
$$

where $\widehat{\boldsymbol{\alpha}}^{(k)}$ is the estimate of $\boldsymbol{\alpha}$ at iteration k, and

$$
\mathbf{S}(\boldsymbol{\alpha}) = \frac{\partial l(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}}, \quad \mathbf{I}(\boldsymbol{\alpha}) = \mathbb{E} \left[\frac{\partial l(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} \frac{\partial l(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}^\top} \right]. \tag{8}
$$

The updating equation for the Scoring estima-

tion approach are obtained by replacing the negative Hessian matrix in the Newton–Raphson algorithm, $-\mathbf{H}(\boldsymbol{\alpha})$, by the Fisher information matrix, $\mathbf{I}(\boldsymbol{\alpha})$ (Fisher, 1925). In summary, the steps of the Scoring estimation approach are as follows,

1. Set initial values $\boldsymbol{\alpha}^{(1)}$;
2. Given $\widehat{\boldsymbol{\alpha}}^{(k)}$, update parameters by Eq. (7);
3. Repeat step 2 until $l(\boldsymbol{\alpha})$ in Eq. (6) is maximized and the estimates converge.

3.3. Computation of cross-IM correlation

Once the GMMs with spatial correlation have been estimated, the cross-IM correlation can be estimated by the empirical Pearson correlation coefficients. Following Baker and Cornell (2006), this study have applied the following steps to compute the empirical correlation coefficients:

1. Compute the inter- and intra-event residuals for each IM,

$$
\widehat{\eta}_i = \frac{\mathbf{1}_{n_i,1}^\top \boldsymbol{\Omega}_i^{-1} [\mathbf{Y}_i - \mathbf{f}(\mathbf{X}_i, \widehat{\mathbf{b}})]}{\frac{1}{\widehat{\tau}^2} + \mathbf{1}_{n_i,1}^\top \boldsymbol{\Omega}_i^{-1} \mathbf{1}_{n_i,1}} \tag{9}
$$

$$
\widehat{\boldsymbol{\varepsilon}}_i = \mathbf{Y}_i - \mathbf{f}(\mathbf{X}_i, \widehat{\mathbf{b}}) - \widehat{\eta}_i
$$

2. Scale the residuals by the estimated standard deviations,

$$
\widetilde{\eta}_i = \widehat{\eta}_i / \widehat{\tau}, \quad \widetilde{\boldsymbol{\varepsilon}}_i = \widehat{\boldsymbol{\varepsilon}}_i / \widehat{\sigma} \tag{10}
$$

3. Compute the empirical correlation coefficient,

$$
\begin{aligned}
& \widehat{\rho}(\mathrm{IM}_1, \mathrm{IM}_2) = \\
& \frac{\widehat{\rho}(\widetilde{\boldsymbol{\eta}}^{(1)}, \widetilde{\boldsymbol{\eta}}^{(2)}) \widehat{\tau}^{(1)} \widehat{\tau}^{(2)} + \widehat{\rho}(\widetilde{\boldsymbol{\varepsilon}}^{(1)}, \widetilde{\boldsymbol{\varepsilon}}^{(2)}) \widehat{\sigma}^{(1)} \widehat{\sigma}^{(2)}}{\sigma_{\mathrm{total}}^{(1)} \sigma_{\mathrm{total}}^{(2)}}
\end{aligned} \tag{11}
$$

where $\widehat{\rho}(\widetilde{\boldsymbol{\eta}}^{(1)}, \widetilde{\boldsymbol{\eta}}^{(2)})$ and $\widehat{\rho}(\widetilde{\boldsymbol{\varepsilon}}^{(1)}, \widetilde{\boldsymbol{\varepsilon}}^{(2)})$ are the correlation coefficients of the inter- and intra-event residuals of a pair of IMs of interest, respectively; $\sigma_{total} = \sqrt{\widehat{\sigma}^2 + \widehat{\tau}^2}$.

Although the current practice estimates the GMMs for each IM individually and then assess the cross-IM correlation separately, several studies (e.g., Goda and Hong, 2008; Arroyo and Ordaz, 2010) have shown that it is possible but challenging to incorporate the spatial cross-IM (i.e., considering both spatial correlation and cross-IM correlation at the same time) in the GMM estimation process, requiring often stricter assumptions. This aspect is currently under investigation by the authors.

3.4. *Modeling of cross-IM correlation*

The analytical correlation model between various IMs and structural period is developed through the following steps (Baker and Cornell, 2006),

1. Apply the Fisher z transformation to the empirical correlation coefficients

$$z = \frac{1}{2} \ln \left(\frac{1 + \widehat{\rho}}{1 - \widehat{\rho}} \right) \qquad (12)$$

where z is the transformed data with a constant standard deviation $\text{var}(z) = 1/\sqrt{\sum_{i=1}^{N} n_i - 3}$

2. Propose a parametric correlation model $\widetilde{\rho}(\boldsymbol{\phi})$;
3. Estimate the parameters $\boldsymbol{\phi}$ by nonlinear least squares and the objective function is

$$\min_{\boldsymbol{\phi}} \sum_{i=1}^{K} \sum_{j=1}^{K} \left(z_{ij} - \frac{1}{2} \ln \left(\frac{1 + \widetilde{\rho}_{ij}(\boldsymbol{\phi})}{1 - \widetilde{\rho}_{ij}(\boldsymbol{\phi})} \right) \right)^2 \qquad (13)$$

where K is the number of IMs.

4. RESULTS AND DISCUSSIONS

4.1. *GMMs with spatial correlation*

The estimated GMM parameters are presented in Table 1.The median predictions for PGA, PGV and PSA(T=1.0) for stiff soil assuming V_{S30}=580 m/s for a normal fault event are visually compared to AB10, ITA10 and BMLA14 in Figure 2. The median predictions for PGA and PSA(T=1.0) are very consistent with the reference GMMs (i.e., generally lie within $\pm 1 \sigma_{total}$ of the derived models). The median predictions for PGV are similar to that of ITA10. Generally, the newly-developed GMMs seem fairly consistent with the literature.

The distribution of residuals from the derived model against distance and magnitude are investigated, although they are not reported here due to space limitations, showing there is no distance- and magnitude-dependency in either inter- or intra-event residuals and implying an overall good fitting.

The parameters in the spatial correlation are estimated by the Scoring estimation approach developed by the authors as a by-product of the GMM estimation. To compare the results with existing studies, the effective range $\widetilde{h}$ computed from h (i.e., $\widetilde{h} = 3h$) is compared to the predictive models of Jayaram and Baker (2009) for cluster sites scenario and the model of Esposito and Iervolino (2012) for

Italy, as shown in Figure 3. It is shown that the overall trend of $\widetilde{h}$ obtained in this study is consistent with the two reference models. However, the effective range $\widetilde{h}$ derived in this study is generally smaller than those of the existing models. This may be due to the use of the classical geostatistical method, which generally tends to overestimate the parameters in spatial correlation (Ming et al., 2019). Moreover, Ming et al. (2019) demonstrated that, although both the multi-stage algorithm and the Scoring estimation approach produce consistent estimators of **b**, the estimators of τ and σ produced by the multi-stage algorithm are inconsistent and it seems to underestimate τ and overestimate σ compared to the Scoring estimation approach.

4.2. *The empirical correlation coefficient*

The empirical correlation coefficients for PSA for multiple period pairs, T_1 and T_2 between 0.01 s and 4 s, are shown in Figure 4, which are compared to three existing cross-IM correlation models: Cimellaro (2013) for Europe, applicable for $0.05 \text{ s} \leq T \leq 2.5$ s; Akkar et al. (2014) for Europe and the Middle East, applicable for $0.01 \text{ s} \leq T \leq 4$ s; Baker and Jayaram (2008) for worldwide shallow crustal regions, applicable for $0.01 \text{ s} \leq T \leq 10$ s. The comparison is not reported here due to space limitations.

The results show that the empirical PSA-PSA correlation coefficients derived in this study follow a similar trend as the existing models except for the Cimellaro (2013) model. Baker and Bradley (2017) have reported that the model of Cimellaro (2013) may have a numerical error since it produces negative correlations for some period pairs. In general, the observed PSA correlation coefficients in the Italian data are slightly higher than the considered studies, particularly, when the separation between T_1 and T_2 is large (e.g., 0.01 s and 4 s).

The empirical PGA-PSA and PGV-PSA correlation coefficients are shown in Figure 5. The results are compared to several studies (Bradley, 2011, 2012; Baker and Bradley, 2017). As shown in Figure 5, the PGA-PSA correlations are similar to that in the NGA-West2 dataset (Baker and Bradley, 2017) and the models of Bradley (2011). Regarding the PGV-PSA correlation, the trend obtained in this study are similar to the NGA-West2

Table 1: Estimated parameters for the ground-motion models proposed in this study

IM	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9	b_{10}	τ	σ	h
PGA	3.524	0.247	-0.020	-3.936	0.351	12.417	0.228	0.160	-0.060	0.080	0.247	0.370	8.476
PGV	0.742	0.188	0.015	-3.089	0.286	8.529	0.308	0.144	-0.021	0.037	0.261	0.301	3.788
0.010	3.544	0.244	-0.019	-3.943	0.352	12.438	0.228	0.160	-0.060	0.080	0.247	0.370	8.333
0.025	3.770	0.191	-0.016	-3.995	0.359	12.220	0.224	0.156	-0.059	0.082	0.248	0.372	7.730
0.040	4.340	0.099	-0.014	-4.198	0.387	11.956	0.212	0.148	-0.054	0.091	0.249	0.385	7.596
0.050	4.668	0.048	-0.013	-4.303	0.399	11.931	0.211	0.155	-0.055	0.096	0.243	0.401	9.919
0.070	4.975	0.034	-0.013	-4.401	0.404	12.404	0.215	0.157	-0.070	0.092	0.237	0.420	12.964
0.100	4.941	0.099	-0.015	-4.345	0.379	14.067	0.212	0.163	-0.088	0.090	0.244	0.430	12.816
0.150	3.667	0.445	-0.032	-3.867	0.290	15.633	0.192	0.160	-0.087	0.099	0.248	0.416	9.761
0.200	2.584	0.687	-0.042	-3.454	0.225	16.378	0.190	0.162	-0.088	0.094	0.251	0.394	6.343
0.250	1.710	0.793	-0.039	-3.011	0.165	15.061	0.195	0.132	-0.076	0.083	0.260	0.366	2.080
0.300	1.214	0.808	-0.034	-2.748	0.137	13.969	0.220	0.140	-0.076	0.059	0.257	0.357	2.396
0.350	0.867	0.802	-0.026	-2.538	0.109	13.637	0.246	0.141	-0.071	0.048	0.255	0.346	1.927
0.400	0.573	0.786	-0.019	-2.387	0.096	12.917	0.260	0.138	-0.068	0.042	0.258	0.337	1.360
0.450	0.170	0.834	-0.021	-2.274	0.090	12.086	0.280	0.146	-0.068	0.031	0.261	0.333	1.375
0.500	-0.131	0.861	-0.020	-2.174	0.081	11.509	0.293	0.149	-0.069	0.025	0.265	0.329	1.405
0.600	-0.481	0.838	-0.012	-2.020	0.068	10.626	0.312	0.151	-0.053	0.015	0.269	0.324	2.227
0.700	-0.648	0.764	-0.002	-1.913	0.066	9.487	0.319	0.153	-0.038	0.010	0.276	0.316	2.922
0.750	-0.844	0.785	-0.002	-1.869	0.063	9.292	0.323	0.152	-0.032	0.006	0.278	0.314	3.375
0.800	-0.884	0.753	0.002	-1.850	0.066	8.990	0.326	0.151	-0.031	-0.001	0.281	0.312	3.823
0.900	-1.235	0.798	0.000	-1.786	0.064	8.238	0.331	0.145	-0.024	-0.007	0.286	0.310	3.682
1.000	-1.329	0.754	0.006	-1.753	0.068	7.660	0.343	0.144	-0.013	-0.006	0.291	0.307	3.877
1.200	-1.602	0.744	0.008	-1.720	0.076	7.043	0.355	0.143	-0.002	-0.017	0.298	0.305	4.463
1.400	-1.827	0.726	0.013	-1.670	0.077	6.393	0.356	0.137	0.004	-0.020	0.301	0.303	5.485
1.600	-1.869	0.684	0.016	-1.714	0.091	6.070	0.365	0.133	0.008	-0.022	0.304	0.300	5.599
1.800	-1.782	0.580	0.029	-1.692	0.089	5.903	0.358	0.129	0.011	-0.026	0.306	0.300	6.547
2.000	-1.887	0.572	0.030	-1.689	0.091	5.858	0.345	0.127	0.021	-0.020	0.308	0.299	7.921
2.500	-2.114	0.596	0.026	-1.785	0.114	5.873	0.324	0.115	0.042	-0.014	0.320	0.298	9.095
3.000	-2.113	0.531	0.032	-1.822	0.122	6.108	0.314	0.112	0.061	-0.017	0.330	0.298	8.906
3.500	-2.166	0.500	0.035	-1.843	0.126	6.275	0.304	0.101	0.081	-0.010	0.337	0.298	9.585
4.000	-2.088	0.438	0.039	-1.914	0.141	6.361	0.305	0.101	0.094	0.000	0.340	0.301	9.688
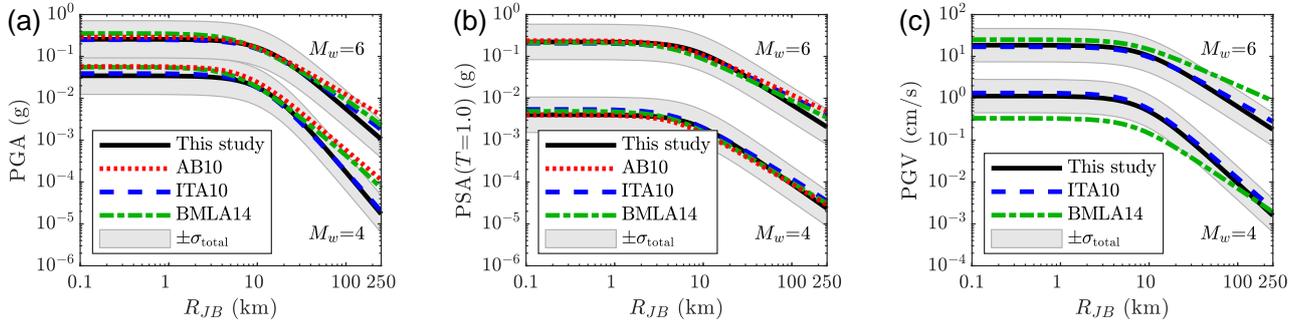
Figure 2: Median predictions for (a) PGA, (b) PSA(T=1.0), and (c) PGV for stiff soil assuming $V_{S30} = 580$ m/s for a normal fault event.
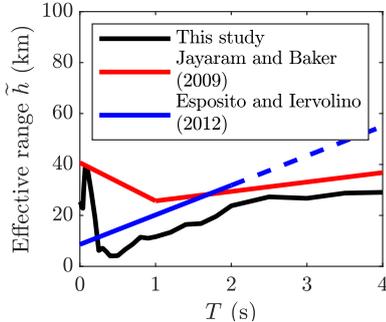
Figure 3: Comparison of effective range $\tilde{h}$ for different models.
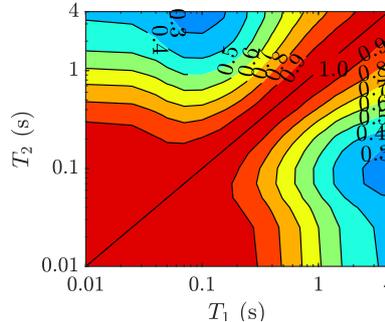
Figure 4: Contours of empirical PSA-PSA correlation.
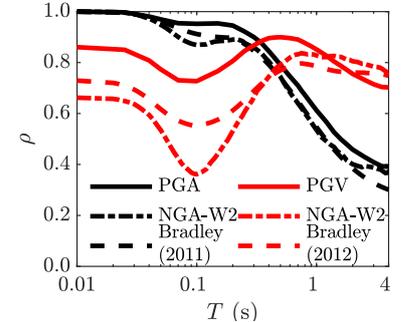
Figure 5: Empirical PGA-PSA and PGV-PSA correlation.

data (Baker and Bradley, 2017) and the models of Bradley (2012), though the derived PGV-PSA correlation is higher than the reference models. Furthermore, the empirical PGA-PGV correlation coefficient is 0.861 in this study, which is slightly higher than 0.733 obtained in Bradley (2012).

These results show that the correlations observed in this study have a higher (in terms of absolute

value) correlation than the reference models, which implies different features in the Italian data from that in the global data. This may be due to the poor representation of normal fault events in global dataset (e.g., 19% normal fault events and 7% of total records NGA-West2 dataset).

4.3. The cross-IM correlation models

The results in the previous sections show there is a need for correlation models calibrated based on the Italian data. In this section, a set of analytical cross-IM correlation models is developed.

Following Baker and Jayaram (2008), the PSA-PSA correlation model is as follows: if $T_{\max} \leq 0.1$, $\widetilde{\rho} = C_2$; else if $T_{\min} > 0.1$, $\widetilde{\rho} = C_1$; else if $T_{\max} \leq 0.2$, $\widetilde{\rho} = \min(C_2, C_3)$; else $\widetilde{\rho} = C_3$, where $T_{\max} = \max(T_1, T_2)$, $T_{\min} = \min(T_1, T_2)$ and

$$C_1 = 1 - \cos\left[\pi/2 - 0.2351 \ln(T_{\max}/T_{\min})\right], \quad (14)$$

$$
\begin{aligned}
C_2 =& 1 - 0.0617\left[1 - \frac{1}{1 + \exp(100T_{\max} - 5)}\right] \\
& \times \left(\frac{T_{\max} - T_{\min}}{T_{\max} - 0.0099}\right),
\end{aligned} \quad (15)
$$

$$C_3 = C_1 + 0.3131(\sqrt{C_1} - C_1)\left[1 + \cos\left(\frac{\pi T_{\min}}{0.1}\right)\right]. \quad (16)$$

Following Bradley (2011, 2012), the analytical correlation models between PGA/PGV and period T is as follows, for $t_{n-1} \leq T < t_n$,

$$\widetilde{\rho} = \frac{(\phi_1 + \phi_2)}{2} - \frac{(\phi_1 - \phi_2)}{2}\tanh\left[\phi_4 \ln\left(\frac{T}{\phi_3}\right)\right] \quad (17)$$

where the parameters ϕ_n are shown in Table 2.

5. CONCLUSIONS

This paper develops a series of GMMs with spatial correlation for 31 amplitude-based IMs by a recently-developed one-stage non-linear regression algorithm proposed by the authors. The estimated model parameters, including the effective range of the spatial correlation function, are consistent with the literature. Based on the newly-developed GMM, the empirical correlation between various IMs observed in the considered dataset are computed and compared to the existing correlation

Table 2: The estimated parameters in Eq. (17)

IM	n	t_n	ϕ_1	ϕ_2	ϕ_3	ϕ_4
PGA	0	0.01	-	-	-	-
	1	0.2	1.000	0.950	0.045	2.225
	2	4	1.000	0.344	0.783	0.824
PGV	0	0.01	-	-	-	-
	1	0.1	0.859	0.722	0.045	2.533
	2	0.5	0.711	0.912	0.203	1.681
	3	4	0.917	0.686	1.450	1.306

models, which implies that the correlation features in the Italian data have not been adequately addressed by literature. Finally, this study proposed a set of analytical correlation models between the selected IMs for the considered Italian ground-motion data. The results of this study can be used to improve hazard/risk assessment exercises in Italy.

6. REFERENCES

Akkar, S. and Bommer, J. J. (2010). "Empirical equations for the prediction of PGA, PGV, and spectral accelerations in Europe, the Mediterranean region, and the Middle East." *Seismological Research Letters*, 81(2), 195–206.

Akkar, S., Sandıkkaya, M. A., and Ay, B. Ö. (2014). "Compatible ground-motion prediction equations for damping scaling factors and vertical-to-horizontal spectral amplitude ratios for the broader Europe region." *Bulletin of Earthquake Engineering*, 12(1), 517–547.

Ambraseys, N. N., Douglas, J., Sarma, S. K. and Smit, P. M. (2014). "Equations for the estimation of strong ground motions from shallow crustal earthquakes using data from Europe and the middle east: Horizontal peak ground acceleration and spectral acceleration." *Bulletin of Earthquake Engineering*, 3(1), 1–53.

Arroyo, D. and Ordaz, M. (2010). "Multivariate Bayesian regression analysis applied to ground-motion prediction equations, part 1 : Theory and synthetic example." *Bulletin of the Seismological Society of America*, 100(4), 1568–1577.

Baker, J. W. and Bradley, B. A. (2017). "Intensity measure correlations observed in the NGA-West2 database, and dependence of correlations on rupture and site parameters." *Earthquake Spectra*, 33(1), 145–156.

Baker, J. W. and Cornell, C. A. (2006). "Correlation of response spectral values for multicomponent ground motions." *Bulletin of the Seismological Society of America*, 96(1), 215–227.

Baker, J. W. and Jayaram, N. (2008). "Correlation of

spectral acceleration values from NGA ground motion models." *Earthquake Spectra*, 24(1), 299–317.

Bindi, D., Massa, M., Luzi, L., Ameri, G., Pacor, F., Puglia, R., and Augliera, P. (2014). "Pan-European ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods up to 3.0 s using the RESORCE dataset." *Bulletin of Earthquake Engineering*, 12(1), 391–430.

Bindi, D., Pacor, F., Luzi, L., Puglia, R., Massa, M., Ameri, G., and Paolucci, R. (2011). "Ground motion prediction equations derived from the Italian strong motion database." *Bulletin of Earthquake Engineering*, 9(6), 1899–1920.

Boore, D. M. (2010). "Orientation-independent, non-geometric-mean measures of seismic intensity from two horizontal components of motion." *Bulletin of the Seismological Society of America*, 100(4), 1830–1835.

Bradley, B. A. (2010). "A generalized conditional intensity measure approach and holistic ground-motion selection." *Earthquake Engineering and Structural Dynamics*, 39, 1321–1342.

Bradley, B. A. (2011). "Empirical correlation of PGA, spectral accelerations and spectrum intensities from active shallow crustal earthquakes." *Earthquake Engineering and Structural Dynamics*, 40, 1707–1721.

Bradley, B. A. (2012). "Empirical correlations between peak ground velocity and spectrum-based intensity measures." *Earthquake Spectra*, 28(1), 17–35.

CEN. (2004). "Eurocode 8: Design of structures for earthquake resistance, Part 1: General rules, seismic actions and rules for buildings, EN 1998-1." European Committee for Standardization (CEN), Brussels.

Cimellaro, G. P. (2013). "Correlation in spectral accelerations for earthquakes in Europe." *Earthquake Engineering and Structural Dynamics*, 42(4), 623–633.

Esposito, S. and Iervolino, I. (2012). "Spatial correlation of spectral acceleration in European data." *Bulletin of the Seismological Society of America*, 102(6), 2781–2788.

Fisher, R. A. (1925). "Theory of statistical estimation." *Mathematical Proceedings of the Cambridge Philosophical Society*, Vol. 22, Cambridge University Press, 700–725.

Goda, K. and Hong, H. P. (2008). "Spatial correlation of peak ground motions and response spectra." *Bulletin of the Seismological Society of America*, 98(1), 354–365.

Jayaram, N. and Baker, J. W. (2009). "Correlation model for spatially distributed ground-motion intensities." *Earthquake Engineering and Structural Dynamics*, 38(15), 1687–1708.

Jayaram, N. and Baker, J. W. (2010). "Considering spatial correlation in mixed-effects regression and the impact on ground-motion models." *Bulletin of the Seismological Society of America*, 100(6), 3295–3303.

Lanzano, G., Sgobba, S., Luzi, L., Puglia, R., Pacor, F., Felicetta, C., D'Amico, M., Cotton, F., and Bindi, D. (2018). "The Pan-European engineering strong motion (ESM) flatfile: compilation criteria and data statistics." *Bulletin of Earthquake Engineering*, 17(2), 561–582.

Lin, T. and Baker, J. W. (2015). "Conditional Spectra." *Encyclopedia of Earthquake Engineering*, 1–11.

Ming, D., Huang, C., Peters, G. W., and Galasso, C. (2019). "An advanced estimation algorithm for ground-motion models with spatial correlation." *Bulletin of the Seismological Society of America*. (In press)

Park, J., Bazzurro, P., and Baker, J. W. (2007). "Modeling spatial correlation of ground motion Intensity Measures for regional seismic hazard and portfolio loss estimation." *Applications of Statistics and Probability in Civil Engineering*, Taylor and Francis Group, London, United Kingdom, 1–8.

Rasmussen, C. E. and Williams, C. K. I. (2006). *Gaussian processes for machine learning*. MIT Press.

Scasserra, G., Stewart, J. P., Bazzurro, P., Lanzo, G., and Mollaioli, F. (2009). "A comparison of NGA ground-motion prediction equations to Italian data." *Bulletin of the Seismological Society of America*, 99(5), 2961–2978.

Stucchi, M., Meletti, C., Montaldo, V., Crowley, H., Calvi, G. M., and Boschi, E. (2011). "Seismic hazard assessment (2003-2009) for the Italian building code." *Bulletin of the Seismological Society of America*, 101(4), 1885–1911.

Weatherill, G. A, Silva, V.. Crowley, H. and Bazzurro, P. (2015). "Exploring the impact of spatial correlations and uncertainties for portfolio analysis in probabilistic seismic loss estimation." *Bulletin of Earthquake Engineering*, 13(4), 957–981.

Zimmerman, D. L. and Michael, S. (2010). "Classical Geostatistical Methods." *Handbook of Spatial Statistics*, A. E. Gelfand, P. J. Diggle, M. Fuentes, and P. Guttorp, eds., CRC Press, 29–44.