2019 SECED Conference
Earthquake risk and engineering towards a resilient world
9-10 September 2019, Greenwich, London

CORRELATION MODELS FOR SIGNIFICANT DURATION OF ITALIAN STRONG-MOTION RECORDS

Chen HUANG[1] & Carmine GALASSO[2]

Abstract: *Ground-motion models (GMMs) are widely used in probabilistic seismic hazard analysis (PSHA) to estimate the probability distributions of earthquake-induced ground-motion intensity measures (IMs). Accounting for spatial and cross-IM correlation in ground motions has important implications on the seismic hazard and risk assessment outputs. This study first develops a new Italian GMM with spatial correlation for 5%-95% significant duration ($D_{S5\text{-}95}$). 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. In fact, current state-of-practice approaches estimate the spatial correlations separately from the GMM estimation, resulting in inconsistent and inefficient estimators of the parameters in the spatial correlation models and GMM. This can, in turn, affect the subsequent cross-IM correlation analysis. Based on the newly-developed GMM, the empirical correlation coefficients from inter- and intra-event residuals are investigated. Finally, this study proposes an analytical correlation model for $D_{S5\text{-}95}$ and spectral ordinates and evaluates the empirical correlation coefficients between $D_{S5\text{-}95}$ and peak ground acceleration (PGA) and peak ground velocity (PGV) for Italy. This is of special interest as 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.*

Introduction

The duration of strong ground motions, as well as their peak amplitude and frequency content, are key characteristics of earthquake-induced ground shaking. This study focuses on the empirical analysis of ground-motion durations observed in Italian strong-motion records. One possible definition of ground-motion duration is that proposed by Trifunac and Brady (1975), known as significant duration, i.e., the time interval across which a specified amount of energy is dissipated (as measured by the integral of the square of the ground acceleration or velocity). In particular, a common measure of significant duration is the time interval between 5% to 95% of the Arias intensity (I_A, defined in Equation [1]), referred to as $D_{S5\text{-}95}$. This is one of the commonly used ground-motion intensity measures (IMs) in recent engineering practice (e.g., Kempton and Stewart 2006).

$$I_A = \frac{\pi}{2g}\int_0^{\infty}\left[\ddot{u}_g(t)\right]^2 dt \ . \tag{1}$$

Hancock and Bommer (2007) and Chandramohan *et al.* (2016), among others, have demonstrated that $D_{S5\text{-}95}$ is well correlated to cumulative structural damage measures (e.g., dissipated hysteretic energy and fatigue damage of structural component), although it has little correlation with earthquake-induced peak structural demands (e.g., interstorey drifts and floor accelerations). Moreover, $D_{S5\text{-}95}$ has been used in geotechnical earthquake engineering applications, showing correlation with the earthquake-induced displacement of landslide masses/slope displacements (e.g., Bray and Rathje 1998; Saygili and Rathje 2008) and with lateral spread displacements resulting from soil liquefaction.

Empirical ground-motion models (GMMs) for $D_{S5\text{-}95}$ include Abrahamson and Silva (1996); Kempton and Stewart (2006); Bommer *et al.* (2009); Lee and Green (2014); Afshari and Stewart (2016) (hereafter referred to as AS1996, KJ2006, BSA2009, LG2014, and AS2016, respectively), which are developed based on worldwide database, such as the Next Generation of Ground-Motion Attenuation Models for the Western United States (NGA-West) project database (Chiou *et al.* 2008) and the Enhancement of Next Generation Attenuation Relationships for Western US (NGA-West2) project database (Ancheta *et al.* 2014). A review of these global GMMs is available in Afshari and Stewart

[1] PhD candidate, University College London, London, UK, chen.huang.14@ucl.ac.uk
[2] Associate Professor, University College London, London, UK, c.galasso@ucl.ac.uk

(2016). Recently, Sandıkkaya and Akkar (2017) developed a GMM for $D_{S5\text{-}95}$ (hereafter SA2017) based on the Reference Database for Seismic Ground-Motion in Europe (RESORCE) (Akkar *et al.* 2014).

These GMMs can also be used to develop correlation models between significant duration and other IMs (e.g., peak ground acceleration or PGA or spectral ordinates) and the resultant correlation model can be used to improve ground-motion selection and modification for structural analyses. For instance, Bradley (2011) has demonstrated that $D_{S5\text{-}95}$ and other integral-based IMs can be used as secondary IMs which should be coupled with primary IMs (e.g., amplitude-based IMs) in the state-of-practice ground-motion selection methods (e.g., the generalized conditional intensity measure or GCIM approach (Bradley 2010)).

However, none of the existing GMMs for significant duration accounts for the spatial correlation in ground-motion data. In fact, due to a common source and wave-traveling paths and a similar distance to fault asperities, ground-motion IMs are spatially correlated (e.g., Park *et al.* 2007). It is important to account for this dependence between various IMs from a single event at multiple sites for probabilistic seismic hazard assessment (PSHA) of spatially-distributed systems (e.g., portfolios of structures and lifelines) (e.g., Park *et al.* 2007; Jayaram and Baker 2009; Weatherill *et al.* 2015).

Moreover, most of the existing GMM are developed based on global databases, which may not well address features observed in Italian data. As discussed by Scasserra *et al.* (2009), Italian data is of special interest because (1) it is principally from earthquakes in extensional regions that are poorly represented in global databases, and (2) past practice in Italy has used local GMMs based on limited datasets that cannot resolve many significant source, path, and site effects. The most recent Italian GMMs (e.g., Lanzano *et al.* 2019a) are based on much larger datasets and are quite sophisticated; yet, the spatial correlation and cross-IM correlation are not considered in these GMMs.

To examine the correlation between $D_{S5\text{-}95}$ and several amplitude-based IM, this study first develops a GMM with spatial correlation for $D_{S5\text{-}95}$, based on the Pan-European Engineering Strong Motion (ESM) flatfile (Lanzano *et al.* 2019b). The same database was previously used by the authors to develop a new Italian GMM for PGA, peak ground velocity (PGV), and 5% damped elastic pseudo-spectral acceleration (PSA) (Huang and Galasso 2019). Given the estimated GMM with spatial correlation, this study evaluates the empirical correlation coefficient between $D_{S5\text{-}95}$ and amplitude-based IM. The proposed GMM is validated by comparison with existing GMMs and residual analysis. Finally, this study develops an analytical $D_{S5\text{-}95}$-PSA correlation model for Italy.

Database

As briefly mentioned in the previous section, the selected dataset is extracted from the ESM flatfile (Lanzano *et al.* 2019b) considering:

- events occurred within Italy;
- events with moment magnitude $M_w \geq 4$ and Joyner-Boore distance (i.e., the closest distance to the surface projection of the rupture plane) $R_{JB} \leq 250$ km;
- recording station with at least two recording sites;
- recording stations classified as free-field;

and discarding:

- 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);
- stations with redundant site information (e.g., co-located sites).

The final dataset includes 7,843 records from 233 earthquakes in the magnitude range $4 \leq M_w \leq 6.9$ 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 of the selected data according to Eurocode 8 (CEN, 2004). 66% of the selected data are caused by the rupture of normal faults, 23% strong motion records are caused by reverse faults and 11% records are caused by strike-slip faults. Most data are collected from stations of site class B/stiff soil and the median V_{S30} across stations is about 637 m/s.

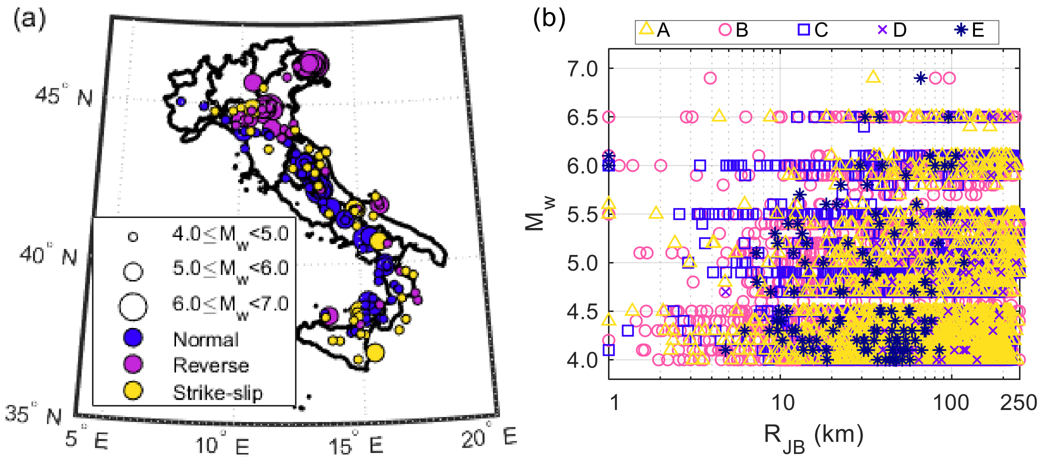

Figure 1. (a) Geographical distribution of considered earthquakes, classified according to focal mechanisms. (b) M_w - R_{JB} distribution with the Eurocode 8 site classification.

Methodology

Model specification

The median prediction for strong-motion duration generally consists of three components: the source duration, the path duration, and the site-effect adjustment (e.g., Abrahamson and Silva 1996). From basic seismological theory, the source duration may be approximated by the reciprocal of the corner frequency $f_c[\Delta\sigma(M), M_0(M)]$, which is assumed be a function of the stress drop $\Delta\sigma(M)$ and seismic moment $M_0(M)$, which in turn, is a function of earthquake magnitude *M*. This results in the following equation, according to Bommer *et al.* (2009):

$$D_S(R_{rup}=0)=\frac{1}{f_c[\Delta\sigma(M),M_0(M)]}=\frac{1}{4.9\times10^6\beta}\left\{\frac{\exp[b_1+b_2(M-M^*)]}{10^{1.5M+16.05}}\right\}^{-1/3} \tag{2}$$

where D_s is the significant duration; R_{rup} is the rupture distance in km; β is the shear-wave velocity of the crust in the vicinity of the source, M^* is a reference magnitude that is simply selected by the analyst. Equation (2) is rearranged by Bommer *et al.* (2009) into the following form,

$$\ln D_S(R_{rup}=0)=c_0+m_1M \tag{3}$$

where *M* is the moment magnitude; c_0 and m_1 are the model coefficients. In addition to the source duration, the increases in significant duration associated with wave propagation and site effects are accounted as follows in Kempton and Stewart (2006),

$$\ln D_S=f(M)+f(R)+f(S) \tag{4}$$

where *R* is the source-to-site distance in km; *S* is the site parameters. Other effects/parameters, such as style-of-faulting, depth-to-top of the rupture fault plane, and soil depth, are also considered in various studies (e.g., Bommer *et al.* 2009; Lee and Green 2014; Afshari and Stewart 2016).

The functional form used in this study is similar to Equation (4), as follows,

$$\begin{aligned}\log D_{s5-95,ij}&=f(M)+f(R)+f(S)+f(Mech)+\eta_i+\varepsilon_{ij}\\ f(M)&=b_1+b_2M_i+b_3M_i^2\\ f(R)&=(b_4+b_5M_i)\log\left(\sqrt{R_{JB,ij}^2+b_6^2}\right)\\ f(S)&=b_7S_{S,j}+b_8S_{A,j}\\ f(Mech)&=b_9F_{N,i}+b_{10}F_{R,i}\end{aligned} \tag{5}$$

where:

- $D_{S5\text{-}95}$ is combined from the two as-recorded horizontal components to produce the RotD50 values, which is the median amplitude among all possible azimuths (Boore 2010).
- M_i is the moment magnitude M_w of event *i*;
- $R_{JB,ij}$ is the Joyner-Boore distance in km at station *j* during event *i*;
- $S_{S,j}$ and $S_{A,j}$ are dummy variables determining the soil type at station *j* according to

$$(S_{S,j}, S_{A,j}) = \begin{cases} (1 \;,\; 0), & soft\ \ soil(V_{S30} < 360\text{m/s}) \\ (0 \;,\; 1), & stiff\ \ soil(360\text{m/s} \le V_{S30} \le 750\text{m/s}) \\ (0 \;,\; 0), & rock(V_{S30} > 750\text{m/s}) \end{cases}; \tag{6}$$

- $F_{N,i}$ and $F_{R,i}$ are dummy variables for the focal mechanism (*Mech*) indicating the style-of-faulting of earthquake *i* as

$$(F_{N,i}, F_{R,i}) = \begin{cases} (1 \;,\; 0) & normal\ \ fault \\ (0 \;,\; 1), & reverse\ \ fault \\ (0 \;,\; 0), & strike-slip\ \ fault \end{cases}; \tag{7}$$

- $(\eta_i)_{i=1,\dots,N}$ are independent and identically distributed inter-event errors with $E(\eta_i) = 0$ and $\text{var}(\eta_i) = \tau^2$ for all $i \in \{1,\dots,N\}$;
- *N* is the number of earthquakes and n_i is the number of stations during earthquake *i*;
- $(\boldsymbol{\varepsilon}_i)_{i=1,\dots,N}$ are independent intra-event error vectors of size $n_i \times 1$ with $E(\boldsymbol{\varepsilon}_i) = \mathbf{0}$ and $\text{cov}(\boldsymbol{\varepsilon}_i) = \phi^2 \boldsymbol{\Omega}_i(\boldsymbol{\omega})$ where $\boldsymbol{\Omega}_i(\boldsymbol{\omega})$ is the correlation matrix corresponding to earthquake *i* with $\boldsymbol{\omega}$, a vector of unknown parameters. To take the spatial correlation into account, the *jj'*-th entry of $\boldsymbol{\Omega}_i(\boldsymbol{\omega})$, $\boldsymbol{\Omega}_{i,jj'}(\boldsymbol{\omega})$, is specified as

$$\boldsymbol{\Omega}_{i,jj'}(\boldsymbol{\omega}) = k(\boldsymbol{s}_{ij}, \boldsymbol{s}_{ij'}) = \rho(\varepsilon_{ij}, \varepsilon_{ij'}) \tag{8}$$

where $k(\boldsymbol{s}_{ij}, \boldsymbol{s}_{ij'})$ gives the correlation $\rho(\varepsilon_{ij}, \varepsilon_{ij'})$ between ε_{ij} and $\varepsilon_{ij'}$ at locations $\boldsymbol{s}_{ij}$ and $\boldsymbol{s}_{ij'}$ of sites *j* and *j'* during earthquake *i*. If it is assumed there are no spatial correlations between intra-event errors at station $j \neq j'$, we have

$$k(\boldsymbol{s}_{ij}, \boldsymbol{s}_{ij'}) = 0. \tag{9}$$

Assuming the spatial field of intra-event errors is stationary and isotropic, the spatial correlations depends on the inter-station distance *d.* An exponential correlation function is used in this study,

$$k(d) = \exp(-d/h) \tag{10}$$

where *h* is a positive range parameter in km at which the spatial correlation is around 0.37 and the effective range parameter corresponding to *ρ*=0.05 spatial correlation is $\tilde{h} = 3h$ (Zimmerman and Michael 2010), which is considered in Esposito and Iervolino (2011, 2012) and Jayaram and Baker (2009).

It is worth pointing out that, the focus of this study is on investigating the spatial correlation model for $D_{S5\text{-}95}$ and its correlations with amplitude-based ground motion IMs observed in Italian data. Hence, a fairly simple ground-motion prediction function (Equation [5]) has been selected for this purpose. As discussed in Baker and Cornell (2006) and Baker and Jayaram (2008), the choice of a particular GMM functional form has an almost negligible effect on the correlation estimates. Moreover, there are many options for correlation functions (e.g., non-stationary and anisotropic) available in the literature (Rasmussen and Williams 2006).

Estimation algorithm

The nonlinear mixed-effects GMM with spatial correlation is estimated by the Scoring estimation approach proposed by the authors (Ming *et al.* 2019). The Scoring estimation approach finds the estimate of the complete vector of model parameters $\boldsymbol{\alpha} = (\mathbf{b}^{\mathrm{T}}, \boldsymbol{\theta}^{\mathrm{T}})^{\mathrm{T}}$ where $\boldsymbol{\theta} = (\tau^2, \phi^2, \boldsymbol{\omega}^{\mathrm{T}})$ that maximizes $l(\boldsymbol{\alpha})$ in Equation (11)

$$l(\boldsymbol{\alpha}) = \frac{\sum_{i=1}^{N} n_i}{2}\ln(2\pi) - \frac{1}{2}\ln|\mathbf{C}(\boldsymbol{\theta})| - \frac{1}{2}[\mathbf{Y} - \mathbf{f}(\mathbf{X}, \boldsymbol{b})]^T \mathbf{C}^{-1}(\boldsymbol{\theta})[\mathbf{Y} - \mathbf{f}(\mathbf{X}, \boldsymbol{b})], \tag{11}$$

via the general updating equation in Equation (12)

$$\hat{\boldsymbol{\alpha}}^{(k+1)} = \hat{\boldsymbol{\alpha}}^{(k)} + \mathbf{I}^{-1}(\hat{\boldsymbol{\alpha}}^{(k)})\mathbf{S}(\hat{\boldsymbol{\alpha}}^{(k)}) \tag{12}$$

where $\hat{\boldsymbol{\alpha}}^{(k)}$ denotes the estimate of $\boldsymbol{\alpha}$ at iteration step *k* and

$$\mathbf{S}(\boldsymbol{\alpha}) = \frac{\partial l(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} \text{ and } \mathbf{I}(\boldsymbol{\alpha}) = E\left[\frac{\partial l(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} \frac{\partial l(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}^{\mathrm{T}}}\right] \quad (13)$$

The updating equation for the Scoring estimation 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. Update the estimates of $\boldsymbol{\alpha}$ by Equation (12);
3. Repeat step 2 until the log-likelihood function in Equation (11) is maximized and the estimates for the parameters converge.

It is worth noting that the GMM for $D_{S5\text{-}95}$ with and without spatial correlation is estimated for the comparison of model performance.

Computation of empirical cross-IM correlation

Once the GMM with spatial correlation has been estimated, the cross-IM correlation can be estimated by the empirical Pearson correlation coefficients, which is also used in Baker and Cornell (2006) and Bradley (2011). This study has applied the following steps to compute the empirical correlation coefficients:

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

$$\hat{\eta}_i = \frac{\frac{1}{\hat{\phi}^2}\mathbf{1}_{n_i,1}^T \Omega_i^{-1}(\hat{\omega})\left[\mathbf{Y}_i - \mathbf{f}(\mathbf{X}_i, \hat{\mathbf{b}})\right]}{\frac{1}{\hat{\tau}^2} + \frac{1}{\hat{\phi}^2}\mathbf{1}_{n_i,1}^T \Omega_i^{-1}(\hat{\omega})\mathbf{1}_{n_i,1}} \text{ and } \hat{\boldsymbol{\varepsilon}}_i = \mathbf{Y}_i - \mathbf{f}(\mathbf{X}_i, \hat{\mathbf{b}}) - \hat{\eta}_i \mathbf{1}_{n_i,1} \quad (14)$$

2. Scale the residuals by the estimated standard deviations from the proposed GMM with spatial correlation;

$$\tilde{\eta}_i = \frac{\hat{\eta}_i}{\hat{\tau}} \text{ and } \tilde{\boldsymbol{\varepsilon}}_i = \frac{\hat{\boldsymbol{\varepsilon}}_i}{\hat{\phi}} \quad (15)$$

3. Compute the empirical correlation coefficient, as follows,

$$\rho(IM_1, IM_2) = \frac{\rho(\tilde{\eta}^{(1)}, \tilde{\eta}^{(2)})\hat{\tau}^{(1)}\hat{\tau}^{(2)} + \rho(\tilde{\varepsilon}^{(1)}, \tilde{\varepsilon}^{(2)})\hat{\phi}^{(1)}\hat{\phi}^{(2)}}{\hat{\sigma}^{(1)}\hat{\sigma}^{(2)}} \quad (16)$$

where $\rho(\tilde{\eta}^{(1)}, \tilde{\eta}^{(2)})$ and $\rho(\tilde{\varepsilon}^{(1)}, \tilde{\varepsilon}^{(2)})$ are the correlation coefficients of the inter- and intra-event residuals of a pair of IMs of interest (i.e., $D_{S5\text{-}95}$ and PGA, PGV, and PSA), respectively; $\hat{\sigma} = \sqrt{\hat{\tau}^2 + \hat{\phi}^2}$.

Modeling of cross-IM correlation

Following Baker and Cornell (2006) and Bradley (2011), the analytical correlation model between various IMs is developed through the following steps:

1. Apply the Fisher *z* transformation to the empirical correlation coefficients, as follows,

$$z = \frac{1}{2}\ln\left(\frac{1+\rho}{1-\rho}\right) \quad (17)$$

where z is the transformed data with a constant variance, $Var(z) = 1/\sqrt{\sum_{i=1}^{N} n_i - 3}$;

2. Propose a parametric correlation model $\rho(\varphi)$;
3. Estimate the parameters φ by nonlinear least squares and the objective function is

$$\min_{\varphi} \sum_{i=1}^{K}\sum_{j=1}^{K}\left[z_{ij} - \frac{1}{2}\ln\left(\frac{1+\rho_{ij}(\varphi)}{1-\rho_{ij}(\varphi)}\right)\right]^2 \quad (18)$$

where *K* is the number of IMs considered.

Results and discussion

GMM with spatial correlation

The estimated parameters and the corresponding 95% confidence interval (CI) of GMM with and without spatial correlation for $D_{S5\text{-}95}$ are presented in Table 1. For illustrative purposes, the median prediction for $D_{S5\text{-}95}$ and its 95% confidence limits are shown in Figure 2, in comparison with existing GMMs, for stiff soil assuming V_{S30}=580 m/s for an M_w 5.5 normal fault event, following the practice of Douglas (2007). It is worth noting that M_w is set to 5.5, which is the median of the applicable magnitude range of this study; no aftershock/hanging-wall effect/basin effect is considered.

α	S	CI	NS	CI
b_1	-0.918	±0.722	-0.771	±0.760
b_2	0.051	±0.283	0.000	±0.302
b_3	0.020	±0.028	0.022	±0.030
b_4	1.399	±0.125	1.354	±0.114
b_5	-0.117	±0.024	-0.102	±0.018
b_6	9.483	±1.224	10.107	±0.960
b_7	0.093	±0.013	0.108	±0.013
b_8	0.015	±0.007	0.015	±0.009
b_9	0.015	±0.032	0.004	±0.034
b_{10}	-0.010	±0.037	-0.024	±0.040
τ^2	$0.070^2(0.162^2)$	±0.001	$0.082^2(0.188^2)$	±0.002
ϕ^2	$0.181^2(0.416^2)$	±0.001	$0.173^2(0.398^2)$	±0.001
h	7.421	±0.549	–	–
AIC	-5602	–	-4853	–
BIC	-5511	–	-4770	–

Table 1. Estimated parameters for $D_{S5\text{-}95}$ GMM with and without spatial correlation (denoted by S and NS, respectively). The values in brackets refer to the results in natural log unit.

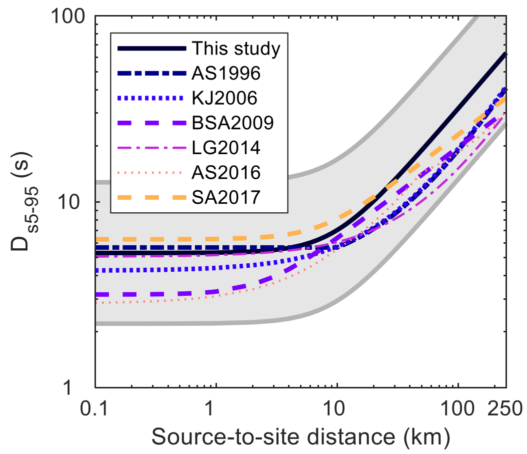

Figure 2. Median predictions for $D_{S5\text{-}95}$ and its 95% confidence limits in comparison with existing GMMs for Italy, for stiff soil V_{S30} = 580 m/s for an M_w 5.5 normal fault event. AS1996, KJ2006, BSA2009, LG2014, and AS2016 refer to Abrahamson and Silva (1996); Kempton and Stewart (2006); Bommer et al. (2009); Lee and Green (2014); Afshari and Stewart (2016), respectively.

It is shown that the proposed GMM for $D_{S5\text{-}95}$ is generally consistent with the reference GMMs, as the reference GMMs generally lie within 95% confident limits of the derived models. However, the $D_{S5\text{-}95}$ predicted in this study by the proposed GMM increases faster than those predicted by the considered studies, particularly at large source-to-site distance, implying that the Italian data may not be well represented by the existing GMMs.

The estimated inter- and intra-event standard deviations are within the ranges of existing GMMs (i.e., τ~[0.20, 0.33] and ϕ~[0.28,0.43] in natural logarithm unit). In comparison with the model estimated without spatial correlation in Table 1, the incorporation of spatial correlation has resulted in a reduction of the inter-event variance and an increase of the intra-event variance, which is consistent with the findings of Jayaram and Baker (2010) and Ming *et al.* (2019) for amplitude-based IMs.

To further compare the performances of the GMM with and without spatial correlation, the Akaike Information Criteria (AIC) (Akaike 1974) and the Bayesian Information Criteria (BIC) (Schwarz 1978), which deal with the trade-off between the goodness of fit of the model and the simplicity of the model (i.e., whether or not to include the spatial correlation), are also reported in Table 1. The model with lower AIC or BIC value would be the preferred one. It is shown that the GMM with spatial correlation has about 15% lower AIC and BIC than the GMM without spatial correlation, implying that the GMM with spatial correlation model provides a better representation of the considered dataset over those without spatial correlation.

Furthermore, it is also shown in Table 1 whether the estimated parameters are significantly different from zeros at a 5% significance level (i.e., whether zero is included within the 95% confidence interval or CI). The range parameter h is significantly different from zero as its 95% CI does not include zero, which implies that the spatial correlation is a non-negligible feature of ground motions. However, the parameters b_2 and b_3 for magnitude scaling and b_9 and b_{10} for style-of-faulting scaling may be zeros (i.e., these terms may not be significant in capturing the ground-motion features), since it cannot reject the null hypothesis that these parameters equal to zeros at a 5% significance level. These findings are consistent with the observations in Bommer *et al*. (2009) for $D_{S5\text{-}95}$ and Bommer *et al.* (2003) and Lanzano *et al.* (2019a) for amplitude-based IMs. However, these results do not mean that these physical parameters are not important in explaining significant duration but, rather, it implies that the functional form involving these parameters may not be a good representation of the specific feature. Lanzano *et al.* (2019a) have suggested that the failure to reject null hypothesis regarding the magnitude scaling may be because of the large variability in magnitude scaling and uncertainty in the estimation of some predefined parameters (e.g., M_h hinge magnitude in their models). Regarding the style-of-faulting scaling, the failure to reject the null hypothesis may be because of the limited difference between amplitudes of motions from normal faulting earthquakes, with respect to those from strike-slip events (Bommer *et al.* 2003). However, it is decided here to keep the functional form as in Equation (5), although some parameters may have a limited impact on model performance.

An illustrative example of residual analysis is shown in Figure 3, in which the inter-event residuals are presented with respect to magnitude and the intra-event residuals with respect to distance and V_{S30}. It is shown that there is no major bias in the residuals with respect to distance, magnitude or V_{S30}, which implies an overall good fitting of the derived models to the Italian data.

The parameter h in the spatial correlation is estimated by the Scoring estimation approach as a by-product of the GMM estimation, as shown in Table 1. However, the existing spatial correlation studies on ground motions focused on the amplitude-based IM and a very few studies are available for $D_{S5\text{-}95}$. Thus, this study provides an estimate of spatial correlation for $D_{S5\text{-}95}$, that is, the range parameter h of the exponential model in Equation (10) is 7.421 km corresponding to 0.37 spatial correlation and the effective range parameter $\tilde{h} = 3h = 22.263$ km corresponding to 0.05 spatial correlation.

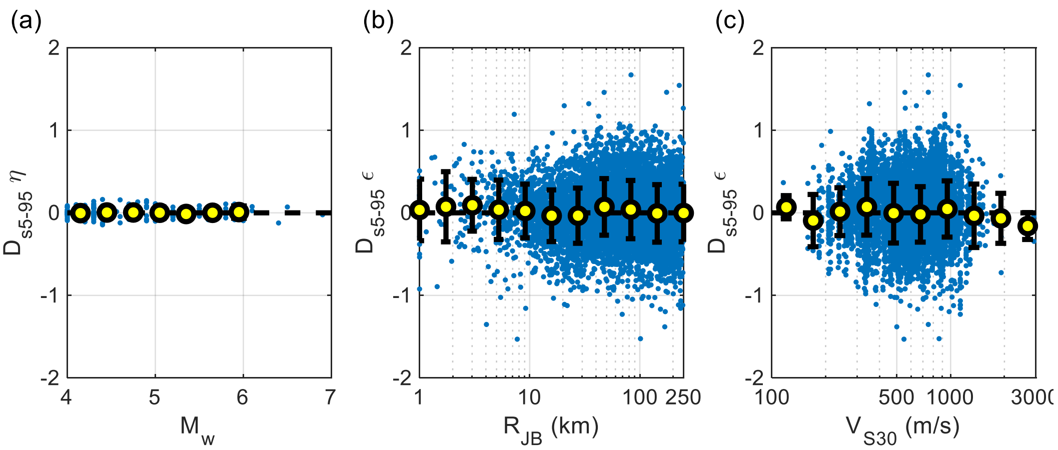

Figure 3. (a) Inter-event residuals versus magnitude; (b) intra-event residuals versus distance; (c) intra-event residuals versus VS30.

The empirical correlation coefficient

The empirical correlation coefficients of $D_{S5\text{-}95}$ and PSA are shown in Figure 4, in comparison to the cross-IM correlation models of Bradley (2011) and Baker and Bradley (2017) developed for global databases (hereafter, referred to B2011 and BB2017) and Sandıkkaya and Akkar (2017) for Europe (hereafter SA2017). It is shown that $D_{S5\text{-}95}$ seems to be negatively correlated with short-period PSA, weakly negatively correlated with moderate-period PSA, and weakly positively correlated with long-period PSA, which is consistent with the findings in Bradley (2011), Baker and Bradley (2017) and Sandıkkaya and Akkar (2017).

Bradley (2011) and Baker and Bradley (2017) have shown that the negative correlation is because ground motions with longer-than-predicted durations tend to have ground motion energy arriving over a longer period of time, and thus less likely to cause large peak responses in a damped oscillator (while long-period oscillators require a longer duration of shaking to build up resonance and so have little or no negative correlation). It is also shown in Figure 4 that the correlations between $D_{S5\text{-}95}$ and short-period PSA observed in Italian data are lower than that implied by the global and European models, while the correlations between $D_{S5\text{-}95}$ and moderate- and long-period PSA are similar to the global models while being slightly lower than the European model.

Furthermore, the empirical correlation coefficients of $D_{S5\text{-}95}$ and PGA and PGV are presented in Table 2. In comparison with Bradley (2011) and Baker and Bradley (2017), the correlations observed in Italian data are lower than the considered studies.

In summary, the correlations observed in this study are generally lower than the considered models at short periods while are consistent with the existing models at longer periods, which implies different features in the Italian data from that in the global dataset and may be possibly due to the poor representation of normal fault events in NGA-West1 and NGA-West2 dataset.

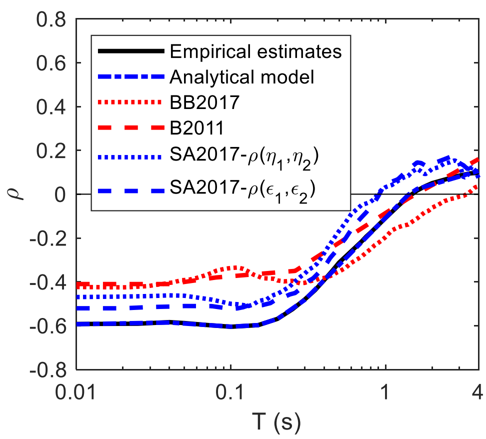

Figure 4 Empirical correlation coefficients $D_{S5\text{-}95}$-PSA versus T ranging from 0.01 s and 4 s and the corresponding analytical model, in comparison of cross-IM correlation models applicable to Italy. B2011, BB2017 and SA2017 refer to Bradley (2011), Baker and Bradley (2017) and Sandıkkaya and Akkar (2017) respectively.

ρ	$D_{S5\text{-}95}$-PGA	$D_{S5\text{-}95}$-PGV
This study	-0.592	-0.359
B2011	-0.405	-0.211
BB2017	-0.424	-0.273

Table 2. Correlation of DS5-95 with PGA and PGV.

The cross-IM correlation models

The results in the previous section show that there is a need for correlation models specifically calibrated based on the Italian data. In this section, a set of analytical correlation models between the selected IMs is developed. Following Bradley (2011), the $D_{S5\text{-}95}$-PSA correlation model is as follows,

$$\rho = a_{n-1} + \left[\ln\left(\frac{T}{t_{n-1}}\right) \Big/ \ln\left(\frac{t_n}{t_{n-1}}\right) \right] (a_n - a_{n-1}) \text{ for } t_{n-1} < T < t_n \quad (19)$$

where a_n is the model coefficients at specific structural periods t_n as listed in Table 3.

n	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
t_n (s)	0.01	0.04	0.10	0.15	0.20	0.30	1.50	4.00
a_n	-0.593	-0.586	-0.604	-0.597	-0.569	-0.482	0.018	0.106

Table 3. The estimated parameters in $D_{S5\text{-}95}$-PSA correlation model in Equation (19)

It is worth noting that there is no physical interpretation of the proposed functional forms in Equations (19), which is only fitting of the observed data and therefore should not be extrapolated. The proposed correlation model is consistent with the empirical correlations observed in the Italian data as shown in Figure 4. Moreover, it accounts for the features observed in the Italian data that is not well captured in the considered studies.

Conclusion

This study has investigated the correlation between significant duration and several amplitude-based IMs observed in Italian data. To this aim, this paper first used Italian ground-motion data to develop a new GMM with spatial correlation for $D_{S5\text{-}95}$ through a recently-developed one-stage non-linear regression algorithm proposed by the authors. The median prediction of the proposed GMM is generally consistent with the existing GMMs for Europe. This study demonstrated that the inclusion of spatial correlation in GMM estimation reduces the inter-event variance and increases the intra-event variance. The residual analysis showed that there is no bias in inter-event residuals with respect to magnitude or in intra-event residuals with respect to distance, implying an overall good fitting of the proposed models to the Italian data. Moreover, this study provided an estimate of the spatial correlation in $D_{S5\text{-}95}$ under the exponential correlation models. The inter- and intra-event standard deviations in this study were similar to the existing models in the literature. Based on the newly-developed GMM, the empirical correlation between $D_{S5\text{-}95}$ and several amplitude-based IMs observed in the considered dataset were computed and compared to the existing correlation models. It is shown that the correlation in the Italian data are lower than those implied by the global and European models, particularly, between $D_{S5\text{-}95}$ and short structural periods, PGA and PGV, which implied that the correlation features observed in Italian data have not been adequately addressed by the literature. Finally, this study proposed an analytical correlation model between $D_{S5\text{-}95}$ and spectral ordinates, capturing well the features of Italian data.

References

Abrahamson NA and Silva WJ (1996) *Empirical ground motion models*. Report to Brookhaven National Laboratory.

Afshari K and Stewart JP (2016), Physically parameterized prediction equations for significant duration in active crustal regions, *Earthquake Spectra*, 32(4): 2057–2081

Akaike H (1974), A new look at the statistical model identification, *IEEE Transactions on Automatic Control*, 19(6): 716–723

Akkar S, Sandıkkaya MA, Şenyurt M, Azari Sisi A., Ay BÖ, Traversa P, Douglas J, Cotton F, Luzi L, Hernandez B and Godey S (2014), Reference database for seismic ground-motion in Europe (RESORCE), *Bulletin of Earthquake Engineering*, 12(1): 311–339

Ancheta TD, Darragh RB, Stewart JP, Seyhan E, Silva WJ, Chiou BSJ, Wooddell KE, Graves RW, Kottke AR, Boore DM, Kishida T and Donahue JL (2014), NGA-West2 Database, *Earthquake Spectra*, 30(3): 989–1005

Baker JW and Bradley BA (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 JW and Cornell CA (2006), Correlation of response spectral values for multicomponent ground motions, *Bulletin of the Seismological Society of America*, 96(1): 215–227

Baker JW and Jayaram N (2008), Correlation of spectral acceleration values from NGA ground motion models, *Earthquake Spectra*, 24(1):299–317

Bommer JJ, Douglas J and Strasser FO (2003), Style-of-faulting in ground-motion prediction equations, *Bulletin of Earthquake Engineering*, 1(2): 171–203

Bommer JJ, Stafford PJ and Alarcón JE (2009), Empirical equations for the prediction of the significant, bracketed, and uniform duration of earthquake ground motion, *Bulletin of the Seismological Society of America*, 99(6): 3217–3233

Boore DM (2010), Orientation-independent, nongeometric-mean measures of seismic intensity from two horizontal components of motion. *Bulletin of the Seismological Society of America* 100(4):1830–35.

Bradley BA (2010), A generalized conditional intensity measure approach and holistic ground-motion selection, *Earthquake Engineering and Structural Dynamics*, 39(12): 1321–1342
Bradley BA (2011), Correlation of significant duration with amplitude and cumulative intensity measures and its use in ground motion selection, *Journal of Earthquake Engineering*, 15(6): 809–832
Bray JD and Rathje EM (1998), Earthquake-induced displacements of solid-waste landfills, *Journal of Geotechnical and Geoenvironmental Engineering*, 124(3: 242–253.
Comite Europeen de Normalisation (2004), *Eurocode 8: Design of structures for earthquake resistance - Part 1: General rules, Seismic action and rules for buildings*. Brussels: CEN
Chandramohan R, Baker JW and Deierlein GG (2016), Quantifying the influence of ground motion duration on structural collapse capacity using spectrally equivalent records, *Earthquake Spectra*, 32(2): 927–950
Chiou BSJ, Darragh R, Gregor N and Silva WJ (2008), NGA project strong-motion database, *Earthquake Spectra*, 24(1): 23–44
Douglas J (2007), On the regional dependence of earthquake response spectra. *ISET Journal of Earthquake Technology,* 44(1): 71-99
Esposito S and Iervolino I (2011), PGA and PGV spatial correlation models based on European multievent datasets, *Bulletin of the Seismological Society of America*, 101(5): 2532–2541
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 RA (1925), Theory of statistical estimation, *Mathematical Proceedings of the Cambridge Philosophical Society*, 22(5): 700–725
Hancock J and Bommer JJ (2007), Using spectral matched records to explore the influence of strong-motion duration on inelastic structural response, *Soil Dynamics and Earthquake Engineering*, 27: 291–299
Huang C and Galasso C (2019), Ground-motion intensity measure correlations observed in Italian data, in *13th International Conference on Applications of Statistics and Probability in Civil Engineering(ICASP13)*, Seoul, South Korea, May 26-30.
Jayaram N and Baker JW (2009), Correlation model for spatially distributed ground-motion intensities, *Earthquake Engineering and Structural Dynamics*, 38(15): 1687–1708
Jayaram N and Baker JW (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
Kempton JJ and Stewart JP (2006), Prediction equations for significant duration of earthquake ground motions considering site and near-source effects, *Earthquake Spectra*, 22(4): 985–1013
Lanzano G, Luzi L, Pacor F, Felicetta C, Puglia R, Sgobba S and D'Amico M (2019a), A revised ground-motion prediction model for shallow crustal earthquakes in Italy, *Bulletin of the Seismological Society of America*, doi: 10.1785/0120180210
Lanzano G, Sgobba S, Luzi L, Puglia R, Pacor F, Felicetta C, D'Amico M, Cotton F. and Bindi, D. (2019b), The Pan-European Engineering Strong Motion (ESM) flatfile: Compilation criteria and data statistics, *Bulletin of Earthquake Engineering*, 17(2): 561–582
Lee J and Green RA (2014), An empirical significant duration relationship for stable continental regions, *Bulletin of Earthquake Engineering*, 12(1): 217–235.
Ming D, Huang C, Peters GW and Galasso C (2019), An advanced estimation algorithm for ground-motion models with spatial correlation, *Bulletin of the Seismological Society of America*, 109(2): 541–566
Park J, Bazzurro P and Baker JW (2007), Modeling spatial correlation of ground motion Intensity Measures for regional seismic hazard and portfolio loss estimation, in *Applications of Statistics and Probability in Civil Engineering*. London: Taylor and Francis Group
Rasmussen CE and Williams CKI (2006), *Gaussian processes for machine learning*. Massachusetts: MIT Press.
Sandıkkaya MA and Akkar S (2017), Cumulative absolute velocity, Arias intensity and significant duration predictive models from a Pan-European strong-motion dataset, *Bulletin of Earthquake Engineering*, 15(5): 1881–1898
Saygili G, Rathje EM and Asce M (2008), Empirical predictive models for earthquake-induced sliding displacements of slopes, *Journal of Geotechnical and Geoenvironmental Engineering*, 134(6): 790–803
Scasserra G, Stewart JP, 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
Schwarz G (1978), Estimating the dimension of a model, *The Annals of Statistics*, 6(2): 461–464

Trifunac MD and Brady AG (1975), A study on the duration of strong earthquake ground motion, *Bulletin of the Seismological Society of America*, 65(3): 581–626

Weatherill GA, 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 DL and Michael S (2010), Classical geostatistical methods, in Gelfand, A. E., Diggle, P. J., Fuentes, M., and Guttorp, P. (eds) *Handbook of Spatial Statistics*. CRC Press, 29–44.