17th World Conference on Earthquake Engineering, 17WCEE

Sendai, Japan - September 13th to 18th 2020

INSIGHTS INTO SPATIALLY CORRELATED GROUND MOTION INTENSITY MEASURES USING ITALIAN EARTHQUAKES

E. Schiappapietra[(1)], J. Douglas[(2)]

[(1)] *Ph.D. Student, Department of Civil and Environmental Engineering, University of Strathclyde, Glasgow (UK), erika.schiappapietra@strath.ac.uk*

[(2)] *Senior Lecturer, Department of Civil and Environmental Engineering, University of Strathclyde, Glasgow (UK), john.douglas@strath.ac.uk*

Abstract

Modelling the spatial correlation of ground motion intensity measures (IMs) has become a keystone in seismic hazard and risk analysis of portfolios of buildings, spatially distributed infrastructures and earthquake-induced phenomena. The quantification of the seismic performance of such systems over a region requires knowledge of the joint probability of occurrence of different ground motion IMs at multiple locations. Therefore, the classical Probabilistic Seismic Hazard Assessment (PSHA) tools, which are based on the hypothesis of independency between IMs at closely spaced sites, are not appropriate.

Over the past decade, the spatial correlation of peak ground acceleration (PGA) and spectral acceleration (SA) has been widely studied. Although common findings suggest that the correlation of intra-event residuals decreases quite rapidly with increasing separation distances, these models feature different rates of decay. Among the causes that may lead to inconsistencies between models, with significant impact on hazard and loss estimates, are the multiple techniques used to estimate the correlation structure, the region and local site conditions, as well as the choice of the databases.

Furthermore, little effort has been directed towards other IMs suitable to characterize the resulting damage to structures and predict ground failure: peak ground velocity (PGV), peak ground displacement (PGD) and spectral displacement (SD) as well as Arias intensity (I_a) and cumulative absolute velocity (CAV), to name but a few. A proper definition of the seismic action in terms of spectral displacement ordinates has progressively gained importance in performance-based seismic design, and I_a and CAV have been found to be adequate for many other earthquake engineering applications, such as evaluating the susceptibility to liquefaction and earthquake-induced landslides.

In this study, we use geostatistical tools in order to compute the spatial correlations of such ground motion parameters. We perform comparisons with other existing models with the aim of: (1) identifying factors that most affect the correlation structure, and (2) quantifying the variability of correlation lengths between different events and regions. Moreover, spatial correlation models are usually calibrated on the within-event component of residuals, obtained based on ergodic ground motion prediction equations (GMPEs). Therefore, we also analyse the spatial correlation of event- and site- corrected residuals, retrieved relaxing the ergodic assumption, to further investigate the factors that determine the spatial dependency of IMs. In order to address these issues, we use the 2016-2017 Central Italy seismic sequence database, which includes nine Mw $\geq$ 5.0 earthquakes that occurred over a time period of five months. These data allow some uncertainties to be removed and an evaluation of the event-to-event variability of the spatial correlation because the same seismic region is considered. Our preliminary results will provide a more accurate picture of ground motions, and thus improve the modelling of earthquake losses for risk model development.

Keywords: spatial correlation; ground motion intensity measures; regional probabilistic seismic hazard analysis.

1. Introduction

Modelling the spatial correlation of ground motion intensity measures (IMs) has become a keystone in seismic hazard and risk analysis of portfolios of buildings, spatially distributed infrastructures and earthquake-induced phenomena. The quantification of the seismic performance of such systems over a region requires not only the estimation of independent IMs values at different sites, but also knowledge of the joint probability of occurrence of such IMs at multiple locations during the same earthquake. Therefore, the classical Probabilistic Seismic Hazard Assessment (PSHA) tools, which are grounded on the hypothesis of independency between IMs at closely spaced sites, are not appropriate.

Over the past decade, the spatial correlation of peak ground acceleration (PGA) and spectral acceleration (SA) has been widely studied. A comprehensive summary of such models can be found in Schiappapietra and Douglas [1], in which a comparison among the various models is provided, highlighting both common findings and differences. The authors demonstrate that multiple techniques used to estimate the correlation structure, region and local site conditions, as well as the choice of the databases are the primary causes of inconsistencies among the models, which can have a significant impact on hazard and loss estimates. Furthermore, little effort has been directed towards other IMs suitable to characterize the resulting damage to structures and predict ground failure: peak ground velocity (PGV) and peak ground displacement (PGD) as well as Arias intensity (I_a) and cumulative absolute velocity (CAV), to name but a few. A proper definition of the seismic action in terms of spectral displacement ordinates has progressively gained importance in performance-based seismic design [2], and I_a and CAV have been found to be adequate for many other earthquake engineering applications, such as evaluating the susceptibility to liquefaction and earthquake-induced landslides as well as predicting the structural damage [3,4].

I_a is defined as the integral of the square of the acceleration time history $a(t)$ over the entire duration adjusted by a constant factor [5]:

$$I_a = \frac{\pi}{2g}\int_0^{t_{max}} a(t)^2\,dt \tag{1}$$

where g is the acceleration of gravity. CAV is the area under the absolute acceleration time history:

$$CAV = \int_o^{t_{max}} |a(t)|dt \tag{2}$$

Differently from the other IMs, which are primarily related to either the amplitude or the frequency content of the ground motion, I_a and CAV implicitly reflect multiple characteristics of the time histories, including the cumulative effects of the ground motion duration [3,6]. As a matter of fact, they turn out to be more efficient to represent the cumulative potential damage due to the ground shaking [4].

Moreover, spatial correlation models are usually calibrated on the within-event component of residuals, obtained based on ergodic ground motion prediction equations (GMPEs). The ergodic assumption implies that the distribution of ground motions over time at given site is the same as their spatial distribution over all site [7]. In the last decade many efforts have been made in order to relax the ergodic assumption, thereby leading towards a non-ergodic approach, in which the systematic and repeatable characteristics of the ground motion (e.g. site-specific and path-specific component) are specified to adjust the median prediction of a GMPE and hence significantly reduce the aleatory uncertainties. Despite the importance of such approaches in case of site-specific PSHA, the assessment of spatial correlation under the non-ergodic assumption has not been widely studied yet. The major contributions in this context are by Kuehn and Abrahamson [8] and Sgobba et al. [9].

In this study, we use geostatistical tools in order to compute the spatial correlations of different ground motion parameters, such as PGA, PGV, PGD, spectral acceleration, CAV and I_a. We perform comparison with other existing models and we also analyse the spatial correlation of event-and-site- corrected residuals, retrieved relaxing the ergodic assumption, to further investigate the factors that determine the spatial dependency of IMs. In order to address these issues, we use the 2016-2017 Central Italy seismic sequence

database, which includes nearly 6900 records from 63 Mw ≥ 3.7 earthquakes that occurred over a time period of five months (August 2016 – January 2017).

2. Database

Starting from 24th August, one of the most important sequences ever recorded in Italy struck the central Apennines, one of the highest hazard zones in Italy. The sequence caused widespread damage to the built environment, thousands of homeless people and invaluable losses for historical heritage as well as earthquake-induced landslides and ground failure. The first mainshock (M_w 6.0) occurred on 24th August near the municipality of Amatrice and was followed by an M_w 5.3 aftershock within less than one hour. After two months, on 26th October, two other large earthquakes (M_w 5.4 and M_w 5.9) struck further north, near the village of Visso and close to the aftershock area of the 1997 Umbria-Marche sequence [10]. On 30th October, the largest event of the sequence (M_w 6.5) occurred close to the municipality of Norcia. Few months later, on 18th January, four other $M_w \geq 5.0$ earthquakes happened to the south of the Norcia mainshock, near the village of Campotosto and Montereale, adjacent to the aftershock zone of the 2009 L'Aquila sequence [10] (Fig. 1, Table 1). The sequence includes 63 $M_w \geq 3.7$ events in the period from August 2016 to January 2017, which were recorded by nearly 400 ground-motion stations, mainly belonging to two major permanent networks: the Italian strong-motion network (RAN, Rete Accelerometrica Nazionale) and the Italian National Seismometric Network (INSN). In addition, several temporary networks were installed in the aftermath of the first mainshock to retrieve more accurate observations of the ground shaking in the epicentral region and further investigate site effects [10].

In this study, we select 6048 records from 44 $M_w \geq 4.0$ well-recorded events with more than 100 recordings, recorded by 367 strong-motion stations within an epicentral distance of 200 km (Fig. 1). The distributions of the selected data with respect to distance, magnitude and magnitude-distance are summarized in Fig. 2. The site conditions at each recording site are expressed according to the EC8 soil categories [11], which are based either on the average shear-wave velocity averaged over the upper-most 30 m ($V_{s,30}$) or on the available geological information. Only about 25% of the considered stations are characterised by shear-wave velocity profiles, so that most of them are classified using geology. The majority of the selected stations are classified as site class B, which includes very dense sand or gravel and very stiff clay deposits.

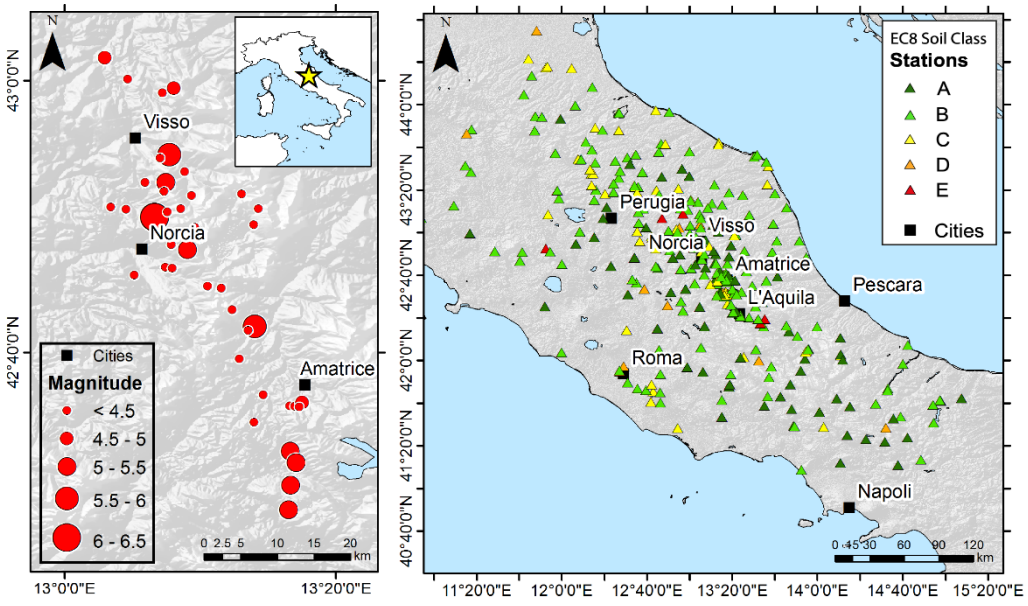

Fig. 1 – Strong ground-motion stations considered in the analysis. Stations are color-coded based on the EC8 site classification. In the zoom-view, epicentres of the events considered in the analysis.

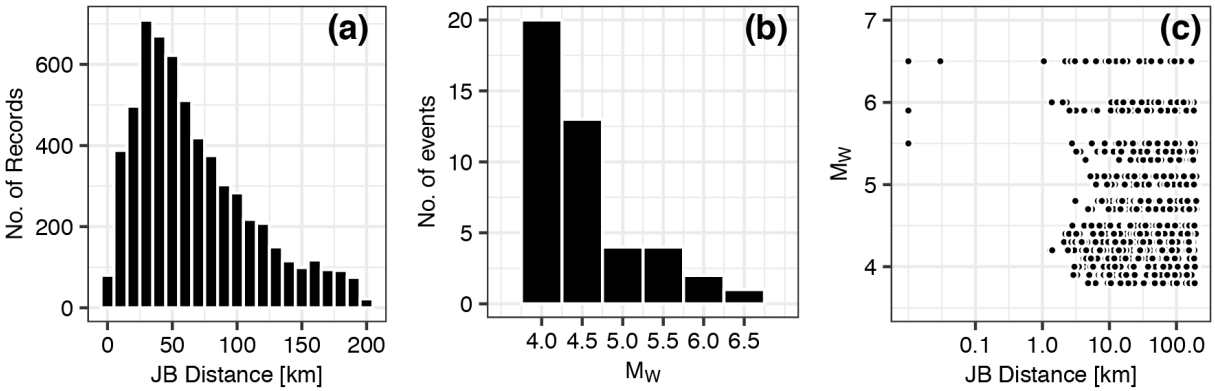

Fig. 2 – (a) Records-distance distribution; (b) Events-magnitude distribution; (c) Magnitude-distance distribution. Distances are in terms of Joyner and Boore distance (R_{JB}, i.e. the distance to the surface projection of the rupture).

Table 1 – Main characteristic of the $M_w \geq 5.0$ earthquakes of the Central Italy seismic sequence.

Event	Event Time	Latitude [°]	Longitude [°]	Depth [km]	Mw	# Records
Amatrice	24/08/2016 01:36	42.70	13.23	8.1	6.0	173
Amatrice Aftershock	24/08/2016 02:33	42.79	13.15	8.0	5.3	173
Visso I	26/10/2016 17:10	42.88	13.13	8.7	5.4	176
Visso II	26/10/2016 19:18	42.91	13.13	7.5	5.9	183
Norcia	30/10/2016 06:40	42.83	13.11	9.2	6.5	175
Montereale I	18/01/2017 09:25	42.55	13.26	9.2	5.1	128
Montereale II	18/01/2017 10:14	42.53	13.28	9.1	5.5	142
Pizzoli I	18/01/2017 10:25	42.49	13.31	8.9	5.4	129
Pizzoli II	18/01/2017 10:33	42.48	13.28	10.0	5.0	127

3. Spatial Correlation Modelling

GMPEs commonly model IMs as lognormally-distributed random variables and predict such intensities at an individual site *i* during an earthquake *j* as a function of source-, path- and site-related parameters. GMPEs take the following form:

$$\log_{10} Y_{ij} = \log_{10} \bar{Y}_{ij}(M, R, S, \theta) + \eta_j + \varepsilon_{ij} \quad (3)$$

where Y_{ij} is the IM of interest at site i^{th} due to the j^{th} event and $\bar{Y}_{ij}$ is the predicted median function of magnitude (M), distance from the source (R), local-site conditions (S) and other explanatory variables (θ). η_j denotes the between-event residual term and represents the systematic deviation of observed IMs with respect to the median prediction of an event. ε_{ij} is the within-event residual term, which represents the misfit between an individual observation at site i from the event-specific average model due to the path and local-site effects. Therefore, the spatial dependency of IMs is explained by the within-event term, as η_j is constant across all sites when a single event is considered. Furthermore, it is assumed that residuals η_j and ε_{ij} are independent and normally distributed random variables with zero mean and standard deviation τ and ϕ, respectively. The joint probability of occurrence of IMs at multiple locations (i,k) during the same earthquake j is usually

modelled as a multivariate gaussian distribution, which is fully described by its mean function, $E[\varepsilon_{ij}]$, and covariance function defined as in Eq. (4):

$$C_j(i,k) = E[\varepsilon_{ij} \cdot \varepsilon_{kj}] - E[\varepsilon_{ij}] \cdot E[\varepsilon_{kj}] \quad (4)$$

For each location, only one realization of the random variable is available (e.g. one observation from a given earthquake) and the mean function can vary depending on the site, so that it is impossible to draw any inference from it [12]. Hence, further simplifications are needed, and the hypotheses of second-order stationarity and isotropy are commonly assumed. This implies that: (1) the mean function is constant for all sites and (2) the covariance is location-independent and depends only on the separation distance (h) between two sites.

In geostatistical analysis, it is common practice to adopt the experimental semivariogram to represent the spatial dependency of IMs values with varying separation distance, which measures the average dissimilarity between spatially distributed data [9, 13, 14, 15, 16]. The semivariogram is defined as:

$$\gamma_j(h) = \frac{1}{2}\mathrm{Var}[\varepsilon_{ij} - \varepsilon_{kj}] \quad (5)$$

In the stationary case, the semivariogram and the covariance function are equivalent, so that the following relation holds [17]:

$$\gamma_j(h) = \mathrm{Var}[\varepsilon_j] - C_j(h) = \mathrm{Var}[\varepsilon_j] \cdot [1 - \rho_j(h)] \quad (6)$$

where $\rho_j(h)$ is the correlation function, defined as $\rho_j(h) = C_j(h)/\mathrm{Var}[\varepsilon_j]$.

3.1 GMPE regression

Generally, an existing GMPE or an *ad hoc* GMPE is adopted to compute the within-event residuals at each site in order to compute the experimental semivariogram. In this study, we develop a data-driven ground motion model based on the data of the Central Italy earthquake sequence (Section 2) to avoid any dependency on the selected GMPE, without including any site-response component. We use RotD50 values of PGA, PGV, PGD, CAV, I_a and 5%-damped SA at 15 periods of vibration between 0.1 and 5 s; hence, the equation is:

$$\log_{10} Y_{ij} = b_1 + b_2 M + b_3 M^2 + (b_4 + b_5 M)\log_{10}\sqrt{R_{JB}^2 + b_7^2} + b_6\sqrt{R_{JB}^2 + b_7^2} + \eta_j + \varepsilon_{ij} \quad (7)$$

where M is either the moment magnitude or the local magnitude and R_{JB} is the Joyner-Boore distance (i.e. the closest distance to the surface projection of the rupture). $b_1 \ldots b_7$ are the model coefficients inferred through a non-linear mixed-effect regression approach, computed using the NLMER algorithm of [18] implemented in [19]. The advantage of using such approach is twofold. Namely, it allows: (1) quantifying the between- and within-event components and further partitioning the within-event group into systematic and non-systematic components; and (2) obtaining unbiased regression for each group (event or station) which has a different number of ground motion records [20].

We derive two GMPEs with identical functional forms, but with different decomposition of residuals: (1) between-event and within-event residuals and, (2) between-event, between-site ($\delta S2S$) and event-and-site corrected residuals (δWS_{es}). The second model allows partially relaxing the ergodic assumption through the definition of the site-to-site term ($\delta S2S$), which includes all the local-site specific effects [7, 21]. This partitioning lets us compute not only the spatial correlation of the within-event residuals, but also of both the site-to-site and event-and-site corrected residuals to further investigate the factors that determine the spatial dependency of IMs. In both case, visual inspection of the residuals (Fig. 3) suggests that the adopted functional form (Eq.7) is performing well, as no significant trends with respect to the predictor variables are detected. It is noted that, for sake of brevity, we show only the plot pertinent to the partially ergodic GMPE. However, we obtain similar results for the ergodic GMPE, where η_j and ε_{ij} are plotted against M_w and R_{JB}, respectively.

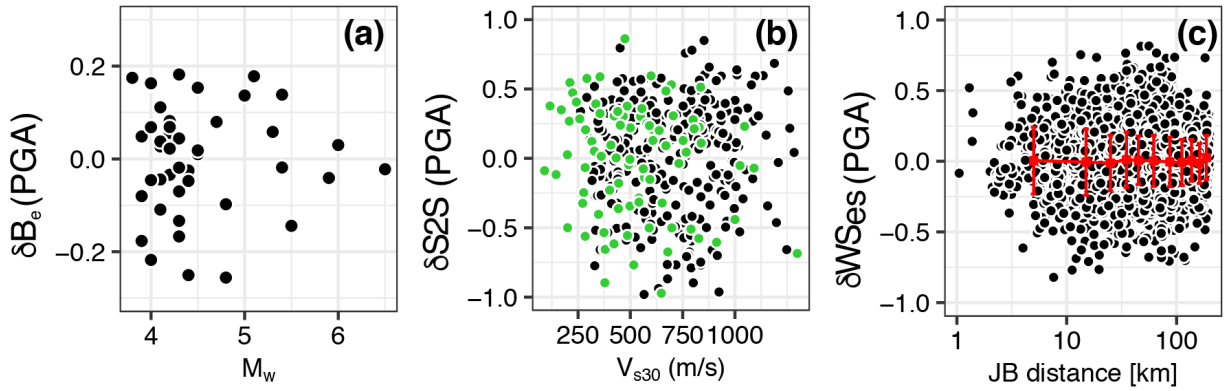

Fig. 3 – Residual plots to check if residuals show any systematic trend with predictor variables: a) η_j (δB_e) against M_w; b) $\delta S2S$ against $V_{s,30}$; c) δWS_{es} against R_{JB}. Green dots in b) highlight stations with a measured $V_{s,30}$ value, whereas black dots stations for which the $V_{s,30}$ is inferred from the slope, following [22].

3.2 Estimation of empirical semivariograms

We use the robust estimator proposed by Cressie [24] to compute the experimental semivariogram, defined as:

$$\hat{\gamma}(h) = \frac{1}{2}\left\{\frac{\left[\frac{1}{|N(h)|}\sum_{N(h)}|\varepsilon_{ij} - \varepsilon_{kj}|^{0.5}\right]^4}{0.457 + \frac{0.494}{|N(h)|}}\right\} \quad (8)$$

where $\hat{\gamma}(h)$ represents the empirical semivariogram and *N(h)* is the number of pairs of observations separated by *h*. Alternatively, one can use the classic estimator based on the method of moments proposed by Matheron [25], which, however, is more sensitive to outliers [3, 26]. The semivariogram is computed for each pair of stations (s_i,s_k) whose inter-site spacing falls in a distance bin set as $h - \Delta h/2 \leq |s_i - s_k| \leq h + \Delta h/2$. Du and Wang and Esposito and Iervolino [3,14] suggest fixing the bin width in such a way that there are at least 30 pairs in each bin. Other studies recommend having at least 100 pairs per bin, so that the estimations result to be more reliable and representative. In our case, we select either a bin size of 5 km if each single event is individually studied or of 1 km if residuals from all earthquakes are pooled to develop a combined model.

3.3 Parametric functions

Several parametric functions have been developed to fit the experimental values obtained through Eq.8 in order to retrieve the spatial dependency of the different residual components (e.g. ε_{ij}, $\delta S2S$, δWS_{es}) for any *h*. Basic second-order stationary and isotropic models are: (1) Gaussian, (2) Exponential, and (3) Spherical, to name but a few [12]. We select an exponential model, as it usually provides the best performance and we implement the following functional form:

$$\gamma(h) = a_0 + a\left[1 - \exp\left(-\frac{ch}{b}\right)\right] \quad (9)$$

where *h* is the separation distance, *a* and *b* are the sill and the range of the semivariogram and *c* is a positive constant set to 3. a_0 is the so called nugget effect, which represents the semivariogram variability at the origin due to either measurements errors or microscale spatial variations. We set a_0 equal to zero throughout the analysis. The sill equals the variance of the data, whereas the range represents the distance beyond which the spatial dependency between site vanishes. For the exponential model, the range is conventionally defined as

the distance at which $\gamma(h)$ equals 0.95 times the sill, since the semivariogram only approaches the sill asymptotically.

Different approaches have been proposed to retrieve the model parameters, such as the weighted least-squares and manual-fitting techniques. We opt for a weighted least-squares approach in which the weights are computed based on both the number of pairs in each bin and the separation distance, so that shorter lags are weighted higher.

3.4 Trend surface

Practically, the hypothesis of second-order stationarity may not hold. The expected value of the random variable may not be constant across all sites, but indeed varying depending on the location. In such cases, the semivariogram increases with separation distance, without reaching a stable sill [17]. These long-range spatial trends should be removed so that small-range correlation structures can be detected. It is common practice to model spatial trends through trend surface models, namely the mean function is described by either first- or second-order polynomial functions of the geographic coordinates. Alternatively, the spatial trends may be defined as a function of any other "scientifically" pertinent parameter of the site, rather than coordinates, in order to better capture the spatial variation of the mean function [17].

In our preliminary study, we simply correct variables through a detrending processing using trend surfaces based on geographical coordinates. Fig. 4 compares the original data (e.g. within-event residuals), which clearly show a pattern from north to south, with residuals obtained after fitting a second-order surface model. In the latter case, the spatial variation is taken up by the trend surface, allowing identification of the small-scale correlation structure. The corresponding semivariograms are plotted in Fig. 5. The raw data feature an increasing experimental semivariogram and hence an unstable sill. Conversely, detrended data show a behaviour more typical of a stationary spatially-correlated process. It is noted that the modelled semivariograms and the correlation parameters presented in Fig. 5 are similar, despite the different behaviour of the experimental values. This is a consequence of the adopted weighting system, which gives more weight to shorter separation distances.

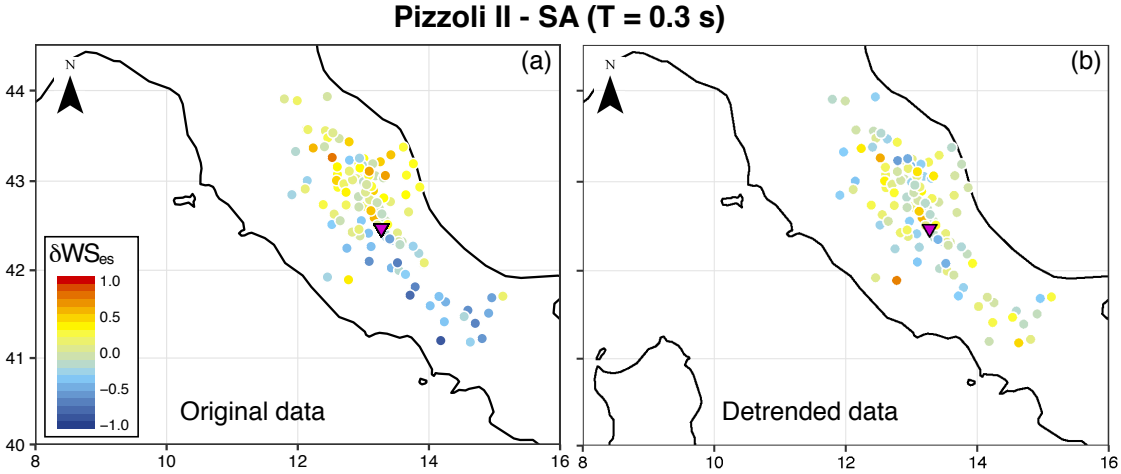

Fig. 4 – Distribution of within-event residuals obtained for SA at 0.3 s: a) original data; b) residuals from the second-order polynomial trend surface. The purple triangle represents the epicentre of the Pizzoli II event.

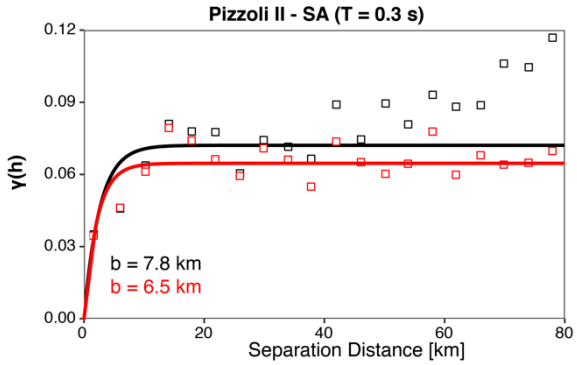

Fig. 5 – Experimental semivariograms (squares) and modelled semivariograms (solid line): raw data are plotted in black; detrended data in red.

4. Results of the geostatistical analyses

4.1 Spatial correlation of different residual components

Generally, spatial correlation studies are performed on the within-event component of residuals, obtained based on ergodic GMPEs. Therefore, the correlation structure of such terms includes spatial variation of both local-site and travel-path effects. In this study, we decide to analyse the spatial correlation not only of the within-event term, but also of the site-to-site term and event- and site-corrected residuals, estimated partially relaxing the ergodic assumption. It is recalled that Eq. (7) does not involve any site-response parameter, hence $\delta S2S$ absorbs all site-specific features and it can be considered as a proxy of the amplification function of each station [27]. Conversely, δWS_{es} is the remaining aleatory variability, which mostly accounts for path effects. Fig. 6 compares the semivariograms obtained for the different residual components of PGV, before and after the detrending processing. Indeed, ε_{ij} and δWS_{es} do not clearly comply with the hypothesis of second-order stationarity, as the semivariogram tends to increase without reaching a stable plateau. On the contrary, $\delta S2S$ does not reveal any spatial trend and thus we do not apply any detrending correction. Similar outcomes were found by Sgobba et al. [28], in which the authors model the spatial correlation under the hypothesis of stationarity for the site term and non-stationarity for the path term, while analysing data from Central Italy. Kuehn and Abrahamson [8] employed a non-stationary covariance function, which includes a dependency on both the inter-site distance and source-to-site distance, to investigate the correlation structure of the path term. According to Kuehn and Abrahamson [8], fitting both the spatial dependency of IMs from near and distant stations is not appropriate; indeed, closely spaced sites located near the source may show different travel paths due to small-scale heterogeneities in the rupture process, whereas seismic waves to distant sites will be along almost identical travel paths.

Fig. 7 shows the spatial correlation results for different IMs by considering within-event and event- and site-corrected residuals as well as the site-to-site term. In agreement with Sgobba et al. [28], the latter term does not show any correlation, with a very small range for all investigated IMs. On the contrary, the correlation distance of ε_{ij} tends to increase with increasing period, as previously observed in the literature. The spatial correlation structure is found to be affected by the response-spectral period considered: in particular, range and period are directly proportioned. δWS_{es} shows a very similar behaviour to ε_{ij}, despite lower correlation values at longer periods. This suggests that spatial dependency of IMs is not due only to site-specific features, but also to travel-path effects and other unexplained effects, not fully captured by our GMPE. A comparison of these residual components for CAV and I_a was also investigated by Foulser-Piggott and Goda [4], who concluded that the two terms have in general different correlation structures, and thus this variability should be accounted for in seismic hazard and risk analysis.

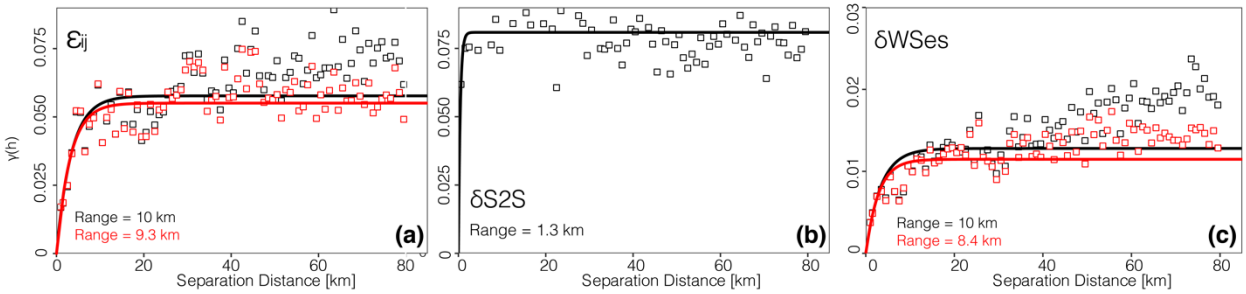

Fig. 6 – Semivariograms of different residual components for PGV obtained pooling all the events considered in the analysis: a) Within-event residuals; b) Site-to-site term; c) Event- and site-corrected residuals. Black squares represent the original data, whereas red dots refer to the detrended data.

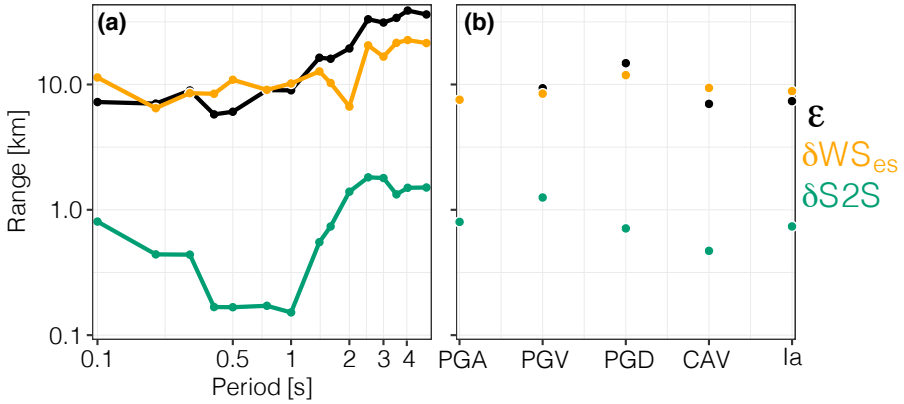

Fig. 7 – Range of within-event, site-to-site and event- and site-corrected term for different IMs: a) Spectral acceleration at period between 0.1 and 5 s; b) PGA, PGV, PGD, CAV and I_a.

4.2 Relationship between magnitude and range

Foulser-Piggott and Goda and Sokolov et al. [4, 29] provided evidence that the range tends to increase with increasing magnitude, since moderate-to-large earthquake are characterized by a lower frequency content. Conversely, Jayaram and Baker [16] did not find any clear relationship between correlation distance and magnitude, at least for the M_w interval analysed ($5 \leq M_w \leq 6.5$). In this study, we expand the M_w interval including all $M_w \geq 4.0$ events. In general, there is not an evident trend for all the IMs considered, as illustrated in Fig. 8 and as already found in Schiappapietra and Douglas [1] for spectral acceleration. We are aware that our database is biased towards small-magnitude earthquakes and this could influence the results. Besides, the large variability in terms of range at lower M_w values may be attributed to other factors, such as the rupture process and stress-drop. Stafford et al. [30] showed that the variability of the correlation distance of events of equal magnitude is strongly influenced by the rupture process. In addition, central Italy earthquake stress-drops were found to have a large variability [10].

4.3 Comparison with other intra-event correlation models

In Fig. 9, we compare our correlation models for PGV, CAV and I_a with some of the studies reported in literature. In particular, we select the models of: Esposito and Iervolino [13], based on Italian data; Sokolov and Wenzel [31], calibrated on Japanese events with $4.2 \leq M_w \leq 7.2$; Wang and Takada [32], which selected earthquakes recorded in Japan and Taiwan with $6.2 \leq M_w \leq 8$; and Du and Wang [3], which developed a prediction equation for the range based on the correlation structure of $V_{s,30}$ values, using data from events occurred in California and Japan. Clearly, the models calibrated on different regions show a much larger

correlation distance than our results. This suggests that regional and local site effects play a first -order role in defining the correlation structure. The range estimated by Esposito and Iervolino [13] is slightly larger than that found for Central Italy. A possible explanation can lie with the extended database considered in Esposito and Iervolino [13], which includes events that occurred in all of the Italian mainland. Similar outcomes were found in Schiappapietra and Douglas [1] for spectral acceleration, suggesting that a unique correlation model based on a large database is not appropriate to represent small regions. Finally, ranges estimated for the Amatrice and Norcia earthquakes are compared to those obtained by Costanzo [6]. CAV and I_a show very similar outcomes, in agreement also with Foulser-Piggott and Goda [4]. However, our results differ to those by Costanzo [6] and the main causes can be attributable to the different estimation process.

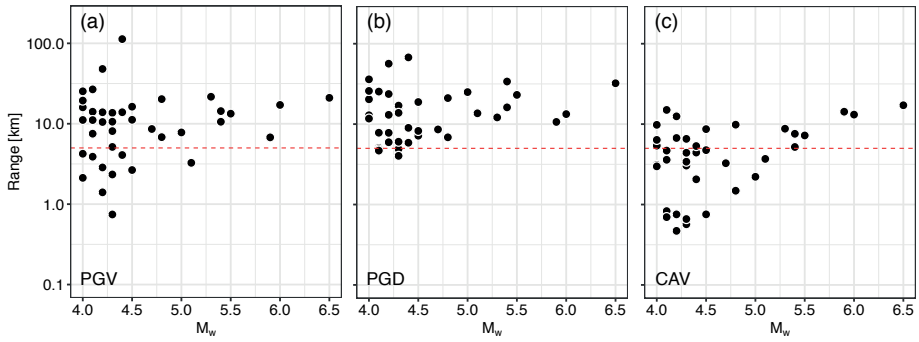

Fig. 8 – Range a as function of magnitude M_w: a) PGV; b) PGD; c) CAV. The dashed red line represents the bin size of 5 km. Any range value smaller than the bin size is indicative of non-correlation and should not be taken as an appropriate estimate.

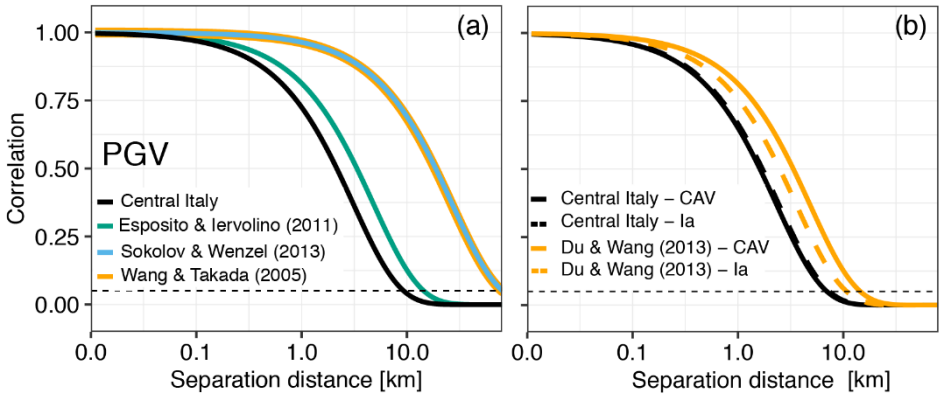

Fig. 9 – Comparison among different correlation models: a) PGV; b) CAV and Ia. Black dashed line points out a level of correlation equal to 0.05.

5. Conclusions

This study focuses on the 2016-2017 Central Italy earthquake sequence to analyse the spatial correlation of different IMs and different residual components. We derive a GMPE through the mixed-effects regression approach for use in computing the correlation distance. We develop both a correlation model for each $M_w \geq 4.0$ event and a global model obtained by pooling data from all considered earthquakes. In agreement with Kuehn and Abrahamson [8] and Sgobba et al. [28], we found that event- and site-corrected residuals, which are mostly influenced by path effects, do not comply with the hypothesis of stationarity. We get around the problem by modelling spatial trends through trend-surface models, calibrated on geographic coordinates. However, the spatial variation of the mean function may be better calibrated using more appropriate parameters

of the site and this aspect requires further analysis to draw more firm conclusions. Alternatively, the approach proposed by Kuehn and Abrahamson [8], which directly includes the computation of a non-stationary covariance function, can be used. Moreover, as suggested by Foulser-Piggott and Goda [4], we found that the analysed residual components have different degrees of spatial correlation, which may be related to the varying underlying physical processes. Therefore, seismic hazard and risk analysis should account for this variability. Finally, the range appears to be regionally-dependent, and thus a region-specific spatial correlation model should be developed.

6. Acknowledgements

We thank the research group of Dr Francesca Pacor at INGV (Istituto di Geofisica e Vulcanologia) for providing the dataset related to the $M_w \geq 4.0$ events that occurred during the Central Italy earthquake sequence. We also thank Dr Peter Stafford for his helpful suggestions about implementing mixed-effects regressions. Finally, we thank the University of Strathclyde for funding the PhD during which the work was undertaken.

7. References

[1] Schiappapietra, E. and Douglas, J. (2020): Modelling the spatial correlation of earthquake ground motion: Insights from the literature, data from the 2016–2017 Central Italy earthquake sequence and ground-motion simulations. *Earth-Science Reviews*, https://doi.org/10.1016/j.earscirev.2020.

[2] Paolucci, R., Rovelli, A., Faccioli, E., Cauzzi, C., Finazzi, D., Vanini, M., Di Alessandro, C. and Calderoni, G. (2008): On the reliability of long-period response spectral ordinates from digital accelerograms. *Earthquake Engineering and Structural Dynamics.* https://doi.org/10.1002/eqe.781.

[3] Du, W., and Wang, G. (2013): Intra-event spatial correlations for cumulative absolute velocity, arias intensity, and spectral accelerations based on regional site conditions. *Bulletin of the Seismological Society of America,* 103(2 A), 1117–1129. https://doi.org/10.1785/0120120185.

[4] Foulser-Piggott, R., and Goda, K. (2015): Ground-motion prediction models for arias intensity and cumulative absolute velocity for Japanese earthquakes considering single- station sigma and within-event spatial correlation. *Bulletin of the Seismological Society of America*, 105(4), 1903–1918. https://doi.org/10.1785/0120140316.

[5] Kramer, S. L. (1996): Geotechnical Earthquake Engineering. *William J. Hall, Editor.*

[6] Costanzo, A. (2018): Shaking maps based on cumulative absolute velocity and arias intensity: The cases of the two strongest earthquakes of the 2016-2017 central Italy seismic sequence. *ISPRS International Journal of Geo-Information*, 7(7). https://doi.org/10.3390/ijgi7070244.

[7] Lanzano, G., Pacor, F., Luzi, L., D'amico, M., Puglia, R. and Felicetta, C. (2017): Systematic source, path and site effects on ground motion variability: The case study of Northern Italy. *Bulletin of Earthquake Engineering*, 15(11), 4563-4583.

[8] Kuehn, N. M., and Abrahamson, N. A. (2020): Spatial correlations of ground motion for non-ergodic seismic hazard analysis. *Earthquake Engineering & Structural Dynamics*, 49(1), 4-23.

[9] Sgobba, S., Lanzano, G., Pacor, F., Puglia, R., D'Amico, M., Felicetta, and., & Luzi, L. (2019): Spatial Correlation Model of Systematic Site and Path Effects for Ground-Motion Fields in Northern Italy. *Bulletin of the Seismological Society of America.* https://doi.org/10.1785/0120180209.

[10] Luzi, L., Pacor, F., Lanzano, G., Felicetta, C., Puglia, R., and D'Amico, M. (2019): 2016–2017 Central Italy seismic sequence: strong-motion data analysis and design earthquake selection for seismic microzonation purposes. *Bulletin of Earthquake Engineering*, 1-19. https://doi.org/10.1007/s10518-019-00676-3

[11] CEN (2004) EN 1998-1:2004 Eurocode 8: design of structures for earthquake resistance—part 1: general rules, seismic actions and rules for buildings. European Committee for Standardization, Brussels.

[12] Webster, R., and Oliver, M. A. (2007): Geostatistics for environmental scientists. *John Wiley & Sons*.

[13] 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), pp. 2532–2541. doi: 10.1785/0120110117.

[14] Esposito, S. and Iervolino, I. (2012): Spatial correlation of spectral acceleration in European data. *Bulletin of the Seismological Society of America*, 102(6), pp. 2781–2788. doi: 10.1785/0120120068.

[15] Heresi, P. and Miranda, E. (2019): Uncertainty in intraevent spatial correlation of elastic pseudo-acceleration spectral ordinates. *Bulletin of Earthquake Engineering*. Springer Netherlands, 17(3), pp. 1099–1115. doi: 10.1007/s10518-018-0506-6.

[16] Jayaram, N. and Baker, J. W. (2009): Correlation model for spatially distributed ground-motion intensities. *Earthquake Engineering and Structural Dynamics*, 38. doi: 10.1002/eqe.

[17] Diggle, P., and Paulo Justiniano, R. (2007): Model-based Geostatistics. *1st Ed. 2007. Springer Ser*. in Statistics. Web.

[18] Bates, D., Machler, M., Bolker, B. and Walker, S. (2014): Fitting linear mixed-effects models using lme4. arXiv preprint arXiv:1406.5823

[19] R Development Core Team. R: a language and environment for statistical computing. Vienna, Austria: R *Foundation for Statistical Computing*; 2010.

[20] Lee, J. (2009): Engineering characterization of earthquake ground motions. *PhD Thesis*. University of Michigan.

[21] Kotha, S. R., Bindi, D., and Cotton, F. (2017): From ergodic to region-and site-specific probabilistic seismic hazard assessment: Method development and application at European and Middle Eastern sites. *Earthquake spectra*, 33(4), 1433-1453.

[22] Wald, D.J. and Allen, T.I., (2007): Topographic slope as a proxy for seismic site conditions and amplification. *Bulletin of the Seismological Society of America*, 97(5), pp. 1379.

[23] Lanzano G, Puglia R, Russo E, Luzi L, Bindi D, Cotton F, D'Amico M, Felicetta C, Pacor F & ORFEUS WG5 (2018). ESM strong-motion flat-file 2018. Istituto Nazionale di Geofisica e Vulcanologia (INGV), Helmholtz-Zentrum Potsdam Deutsches GeoForschungsZentrum (GFZ), Observatories & Research Facilities for European Seismology (ORFEUS). PID: 11099/ESM_flatfile_2018.

[24] Cressie, N. (1985): Fitting variogram models by weighted least squares. *Journal of the international Association for mathematical Geology*, *17*(5), 563-586.

[25] Matheron, G., (1962): Traité de géostatistique appliquée. 1 (1962). *Editions Technip*. Cited in Esposito and Iervolino (2011).

[26] Oliver, M. A. and Webster, R. (2014): A tutorial guide to geostatistics: Computing and modelling variograms and kriging. *Catena. Elsevier B.V.*, 113, pp. 56–69. doi: 10.1016/j.catena.2013.09.006.

[27] Kotha, S. R., Cotton, F., and Bindi, D. (2018): A new approach to site classification: Mixed-effects Ground Motion Prediction Equation with spectral clustering of site amplification functions. *Soil Dynamics and Earthquake Engineering*, 110, 318–329. https://doi.org/10.1016/j.soildyn.2018.01.051.

[28] Sgobba, S., Lanzano, G., Pacor, F., Felicetta, C. (2019): A scenario-based approach to generate empirical shaking maps in Central Italy from non-ergodic ground motion models. *38° GNGTS conference, 12-14 November, 2019* Rome.

[29] Sokolov, V., Wenzel, F., Wen, K. L., and Jean, W. Y. (2012): On the influence of site conditions and earthquake magnitude on ground-motion within-earthquake correlation: analysis of PGA data from TSMIP (Taiwan) network. *Bulletin of Earthquake Engineering*, *10*, 1401–1429. https://doi.org/DOI 10.1007/s10518-012-9368-5.

[30] Stafford, P. J., Zurek, B. D., Ntinalexis, M., and Bommer, J. J. (2018): Extensions to the Groningen ground-motion model for seismic risk calculations: component-to-component variability and spatial correlation. *Bulletin of Earthquake Engineering*, *17*(8), 4417–4439. https://doi.org/10.1007/s10518-018-0425-6.

[31] Sokolov, V., and Wenzel, F. (2013): Further analysis of the influence of site conditions and earthquake magnitude on ground-motion within-earthquake correlation: analysis of PGA and PGV data from the K-NET and the KiK-net (Japan) networks. *Bulletin of Earthquake Engineering*, 11(6), 1909-1926.

[32] Wang, M., and Takada, T. (2005): Macrospatial correlation model of seismic ground motions. *Earthquake spectra*, 21(4), 1137-1156.

