



Preliminary results of spatial correlation analysis for acceleration spectral ordinates from Italian data

Simona Esposito, Iunio Iervolino

Dipartimento di Ingegneria Strutturale, Università degli Studi di Napoli Federico II, Naples, Italy.

Keywords: spectral acceleration, semivariogram, lifelines, portfolio, seismic risk.

ABSTRACT

Spatial correlation of peak ground motion amplitudes is required in modeling hazard for risk assessment of spatially distributed systems. In particular when a portfolio of buildings or a transportation/distribution network is of concern, spectral acceleration (S_a) correlation models may be considered in order to evaluate the expected loss in case of seismic events. The estimation of an appropriate spatial correlation model is still a research task in earthquake engineering, because of several issues related to: data, statistical approaches and estimation tools, and tests to evaluate the estimated models. In the presented paper an analysis of the spatial correlation of S_a is carried out using the Italian ACcelerometric Archive (ITACA) dataset. Correlation is estimated on the residuals with respect to a ground motion prediction equation (GMPE) calibrated on the same data considered. Results show that the decay rate of correlation, as a function of inter-site distance, tends to increase with structural period. Based on that, a simple linear formula, although preliminary, is provided to model spatial correlation of S_a as a function of frequency.

1 INTRODUCTION

Assessment of intraevent spatial correlation of ground motion intensity measures (IMs) has become a relevant topic in seismic risk analysis. The importance of modeling such phenomenon is due to the requirement to extend seismic risk analysis, usually related to site-specific structures, to spatially distributed systems and lifelines. In particular, on the hazard side, probabilistic seismic hazard analysis (PSHA; McGuire, 2004) refers to ground motion prediction equations (GMPEs) to model ground motion which provide probabilistic distribution of the chosen IM conditional on earthquake magnitude, source-to-site distance, and other parameters such as local geological conditions. Since it was demonstrated that GMPEs' residuals of IMs (e.g., peak ground acceleration, peak ground velocity, spectral acceleration) are spatially correlated (e.g., Boore et al. 2003, Goda and Hong 2008, Esposito and Iervolino 2011), it is important to have

correlation models for hazard assessment of a region.

Spatial correlation models available in literature have been empirically estimated mainly on earthquakes outside Europe, such as Northridge (1994) or Chi-Chi (1999). Most of the studies are based on dense observations of single events (e.g., Boore et al. 2003, Wang and Takada 2005, Jayaram and Baker 2009); a few works have, instead, combined data from multiple events to obtain a unique estimate of correlation (e.g., Goda and Hong 2008, Goda and Atkinson 2009, Goda and Atkinson 2010, Sokolov et al. 2010).

Those models depend uniquely on inter-site separation distance and provide the distance limit at which correlation may technically considered to be lost (i.e., distance beyond which IMs may be considered uncorrelated). Moreover, if spectral acceleration (S_a) is of concern, the correlation may depend also on the period S_a refers to.

In Figure 1, several models for spectral acceleration, considering as period 1 sec are shown. The correlation coefficient is expressed by Equation 1, where a , b , and c are the model parameters, T is the period and h is the inter-site separation distance (in km).

$$\rho(h, T) = \max \left\{ (1 - c(T)) + c(T) \cdot e^{-a(T) \cdot h^b}, 0 \right\} \quad (1)$$

The black dotted line in Figure 1 represents the value of distance at which the correlation may technically considered to be lost, i.e. equal to 0.05, or the *practical range* as discussed in the following.

It is clear from the figure that distances at which the correlation is conventionally considered lost is very different for each model considered. In fact, models from different authors provide different results even if estimating correlation for the same IM. This is supposed to depend on several factors such as the dataset used, the GMPE chosen to compute intraevent residuals, and the working assumptions of the estimation.

For example, Sokolov *et al.* (2010), starting from the strong-motion database of TSMIP network in Taiwan, investigated the dependency of spatial correlation on site classes and geological structures, asserting that a single generalized spatial model may not be adequate for all of Taiwan territory. In some cases (e.g., Wang and Takada 2005, Jayaram and Baker 2009) existing GMPEs are used, while, in others (e.g., Goda and Hong 2008, Goda and Atkinson 2009, Sokolov *et al.* 2010), *ad-hoc* fit on the chosen dataset is adopted. Generally, regressions analysis used to develop prediction equations does not incorporate the correlation structure of residuals as a hypothesis. Hong *et al.* (2009) and Jayaram and Baker (2010), evaluated the influence of considering the correlation in fitting a GMPE, finding a minor influence on regression coefficients and a more significant effect on the variance components. Goda and Atkinson (2010), investigated the influence of the estimation approach, emphasizing its importance when residuals are strongly correlated.

In this paper, the evaluation of the spatial correlation of S_a residuals is carried out using the Italian ACcelerometric Archive (ITACA). The analysis of correlation was performed through geostatistical tools pooling data from multiple events to fit a unique model following the same approach of Esposito *et al.* (2010) and Esposito and Iervolino (2011).

The GMPE with respect to which residuals are computed is those of Bindi *et al.* (2011) and only

records used to estimate the considered GMPE are employed to estimate spatial correlation.

In the paper, the first part describes the framework adopted to estimate the correlation. The second part provides the working assumptions and a description of the dataset considered. Then, results of the estimation of correlation lengths for spectral acceleration for the eight periods ranging from 0 s to 2 s are given.

Finally, an approximated preliminary relationship of the correlation range as a function of period is provided and compared with previous research on the same topic.

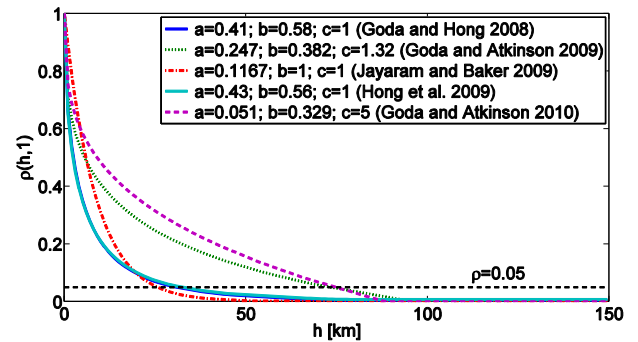


Figure 1. Some correlation models available in literature for $S_a(1s)$: the black dashed line intersects the curves at the distance at which the correlation is conventionally considered as almost lost (i.e., the correlation coefficient is equal to 0.05).

2 SEMI-EMPIRICAL MODELING OF SPATIAL CORRELATION

2.1 Geostatistical analysis

GMPEs model the logs of spectral acceleration for a specific period T , and related heterogeneity, at a site p due to earthquake j as in Equation (2).

$$\log Sa(T)_{pj} = \overline{\log Sa(T)_{pj}}(M, R, \underline{\theta}) + \eta_j + \varepsilon_{pj} \quad (2)$$

$\overline{\log Sa(T)_{pj}}(M, R, \underline{\theta})$ is the mean of the logs conditional on parameters such as magnitude (M), source-to-site distance (R), and others ($\underline{\theta}$); η_j denotes the inter-event residual, which is a constant term for all sites in a given earthquake and represents a systematic deviation from the mean of the specific seismic event; and ε_{pj} is the intra-event variability of ground motion. ε_{pj} and η_j are usually assumed to be independent random variables, normally distributed with zero mean and standard deviation σ_{intra} and σ_{inter} , respectively. Then, $\log Sa(T)_{pj}$ is modeled as a normal random variable with mean

$\overline{\log Sa(T)}_{pj}(M, R, \underline{\theta})$ and standard deviation, σ_T where $\sigma_T^2 = \sigma_{intra}^2 + \sigma_{inter}^2$.

If the hazard assessment at two or more sites is of concern, the joint probability density function (PDF) for Sa(T) at all locations is required, and it can be modeled with a multivariate normal distribution (Jayaram and Baker 2008).

It is assumed that the logs of Sa(T) form a Gaussian random field (GRF), defined as a set of random variables, one for each site \mathbf{u} in the study area $S \in \mathbb{R}^2$.

To any set of n sites $\mathbf{u}_p, p=1, \dots, n$, corresponds to a vector of n random variables that is characterized by the covariance matrix, Σ , as in Equation 3 where the first term produces perfectly correlated inter-event residuals (Malhotra 2008), while the second term (symmetrical) produces partially correlated intra-event residuals.

$$\Sigma = \sigma_{inter}^2 \cdot \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 1 \end{bmatrix} + \sigma_{intra}^2 \cdot \begin{bmatrix} 1 & \rho_{12} & \dots & \rho_{1n} \\ \rho_{21} & 1 & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n1} & \rho_{n2} & \dots & 1 \end{bmatrix} \quad (3)$$

In Equation 3, the correlation is heterogeneous as it depends on the pairs of sites considered, and the intra-event variance is homoscedastic as it is constant for all sites (this assumed in most of GMPEs, although some studies have found dependence of intra-event variability on distance, magnitude and non-linear site effects; Strasser et al. 2009).

If the spatial correlation for intra-event residuals is a function of the relative location of sites, it becomes as in Equation 4, where p and q are two locations at the end of \mathbf{h}_{pq} (the separation vector between the two sites).

$$\rho_{pq} = \rho(\mathbf{h}_{pq}) \quad (4)$$

Under the hypothesis of second-order stationarity and isotropy of the GRF, correlation depends only on the separation distance $h = \|\mathbf{h}\|$. Therefore if intra-event residuals may be modeled as a stationary and isotropic GRF, all data available from different earthquakes and regions,

therefore deemed homogeneous, are used to fit a unique model (see Esposito and Iervolino 2011, for more details).

A common tool to quantify spatial variability of georeferenced data is the semivariogram $\gamma_j(h)$

It is used to model the covariance structure of GRF through suitable functions. Under the hypothesis of second-order stationarity and isotropy it is defined as in Equation 5.

$$\gamma_j(h) = Var(\varepsilon_j) \cdot [1 - \rho_j(h)] \quad (5)$$

where $\rho_j(h)$ denotes the spatial correlation coefficient between intra-event residuals separated by the distance h .

The estimation of correlation usually develops in three steps:

1. computing the empirical semivariogram¹;
2. choosing a functional form;
3. estimating the correlation parameters by fitting the empirical data with the functional model.

Empirical semivariograms are computed as a function of site-to-site separation distance, with different possible estimators. The classical estimator is the method-of-moments (Matheron 1962) which is defined for an isotropic random field in Equation 6, where $N(h)$ is the set of pairs of sites separated by the same distance h , and $|N(h)|$ is the cardinal of $N(h)$.

$$\hat{\gamma}(h) = \frac{1}{2 \cdot |N(h)|} \cdot \sum_{N(h)} [\varepsilon(\mathbf{u} + \mathbf{h}) - \varepsilon(\mathbf{u})]^2 \quad (6)$$

Since this estimator can be badly affected by atypical observations (Cressie 1993), Cressie and Hawkins (1980) proposed a more robust estimator (less sensitive to outliers). Both estimators will be used in the evaluation of intra-event spatial correlation of Sa(T).

To compute the semivariogram it may be useful, when dealing with earthquake records, to define tolerance bins around each possible h value. The selection of distance bins has important effects: if its size is too large, correlation at short distances may be masked; conversely if it is too small, empty bins, or bins with samples small in size, may impair the

¹ Assuming a common semivariogram for different events, that is, invariant through earthquakes, allows to neglect the subscript j in the following equations.

estimate. A rule of thumb is to choose the maximum bin size as a half of the maximum distance between sites in the dataset, and to set the number of bins so that there are at least thirty pairs per bin (Journal and Huijbregts 1978).

The interpretation of experimental semivariograms consists in the identification of a model among the family of functions able to capture and emulate its trend. The three basic stationary and isotropic models are: exponential, spherical, and Gaussian. In particular, the exponential model which is the most common one, is described in Equation 7:

$$\gamma(h) = c_0 + c_e \cdot (1 - e^{-3h/b}) \quad (7)$$

where c_0 is defined *nugget*, i.e. the limit value of the semivariogram when h is zero, c_e is the *sill*, or the population variance of the random field (Barnes 1991) and b is the practical range defined as the inter-site distance at which $\gamma(h)$ equals 95% of the sill.

Note that the parameter b defined in Equation 7, which will be used in the following, does not correspond to the parameter defined in Equation 1.

Several goodness-of-fit criteria for finding the best parametric model have been proposed in geostatistical literature. Studies dealing with earthquake data sometimes use visual or trial and error approaches in order to appropriately model the semivariogram structure at short site-to-site distances (Jayaram and Baker 2009), where it is significant.

In this work experimental semivariograms are fitted visually, although using the least squares estimation as a starting point.

2.2 Estimating correlation on multievent data

Empirical semivariograms are computed starting from normalized intra-event residuals obtained for a single earthquake j and a generic site p as $\varepsilon_{pj}^* = \varepsilon_{pj} / \sigma_p$ where σ_p is the standard deviation of the intra-event residual at the site p (in the study the intra-event standard deviation is common for all sites consistent with GMPEs used to compute residuals, to follow).

The standardization enables to not estimate the sill, as it should be equal to one; therefore Equation 5 becomes Equation 8, where the superscript represents an empirical estimate.

$$\hat{\gamma}(h) = 1 - \hat{\rho}(h) \quad (8)$$

With earthquake data, standardization can be carried out with the standard deviation provided by the GMPE². Another option is to use the sample variance as an estimate of the true variance (e.g., Jayaram and Baker 2010). Moreover Goda and Atkinson (2010) used the intra-event standard deviation inferred from the large-separation-distance plateau of the semivariogram, assuming that at those distances residuals are not correlated.

In this work, the variance provided by the GMPEs was preferred. In fact, evaluation of possible alternatives for standardization leads to results which seem to be not significantly affected by a choice with respect to another (e.g., in Esposito and Iervolino 2011).

Normalized intra-event residuals from multiple events (and regions) are then pooled to fit a unique correlation model. This because geostatistical estimation needs a relatively large number of data to model the semivariogram (i.e., many records to have more than thirty pairs in each h bin), which are not available for individual events in the chosen dataset.

Assuming the same isotropic semivariogram with the same parameters for all earthquakes the experimental semivariogram becomes that of Equation 9 where n_j is the number of records for the j th event and $|N(h)|$ is the number of pairs in the specific h bin.

Equation 10 shows how individual events are kept separated in computing the empirical semivariogram. In fact, the differences of residuals of Equation (9) are computed only between pairs of residuals (standardized with the common standard deviation from the GMPE) from the same earthquake, then differences from different earthquakes are pooled.

This process is visually sketched in Figure 3 from which it is possible to note that the empirical semivariogram point at h_1 is not the average of experimental semivariograms from different earthquakes, as $N(h)$ is the number of pairs in the specific h bin from all earthquakes.

$$\hat{\gamma}(h) = \frac{1}{2 \cdot |N(h)|} \sum_{N(h)} [\varepsilon_{pj}^* - \varepsilon_{qj}^*]^2 \quad (9)$$

² It is usual to use the sample variance as an estimator of the sill for the experimental semivariogram, but this may be improper in some circumstances; see Barnes (1991) for a discussion.

$$N(h) = \left\{ \left(j, \varepsilon_{pj}^*, \varepsilon_{qj}^* \right) : \left\| \varepsilon_{pj}^* - \varepsilon_{qj}^* \right\| = h; \right. \\ \left. p, q = 1, \dots, n_j; j = 1, \dots, k \right\} \quad (10)$$

3 DATASET

The number of considered recordings from ITACA corresponds exactly to data used to fit Bindi *et al.* 2011 GMPE (except that only earthquakes for which more than one record was available were considered) and it is equal to 763 ground motions from 97 events over the 4-6.9 magnitude range (moment magnitude, M_w).

Source-to-site distance is the closest horizontal distance to the vertical projection of the rupture (i.e., Joyner-Boore distance, R_{jb}) and is up to 196 km in the data.

Characteristics of the datasets, with respect to explanatory variables of the considered prediction equation (magnitude, distance and local site conditions) are shown in Figure 2.

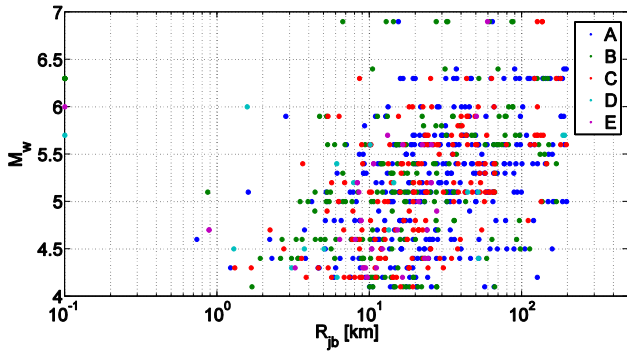


Figure 2. ITACA strong-motion subsets with respect M_w , R_{jb} and local site conditions according to Eurocode 8 (CEN, 2003).

4 SPATIAL CORRELATIONS OF SPECTRAL ORDINATES FROM ITACA

In order to have a reasonable number of data pairs in the bins (at least 30) and a stable trend of correlation, the experimental semivariograms were obtained using a width of 2 km at seven periods ranging between 0.1 seconds and 2 seconds.

In Figure 4 the distribution of data pairs as a function of separation distance bins (2 km) is shown.

Because the GMPE used to obtain intraevent residuals refers to geometric mean of horizontal components, the correlation was estimated for this IM.

Both estimators (classical and robust) were used; no significant difference was found in the shape of the fitted semivariogram.

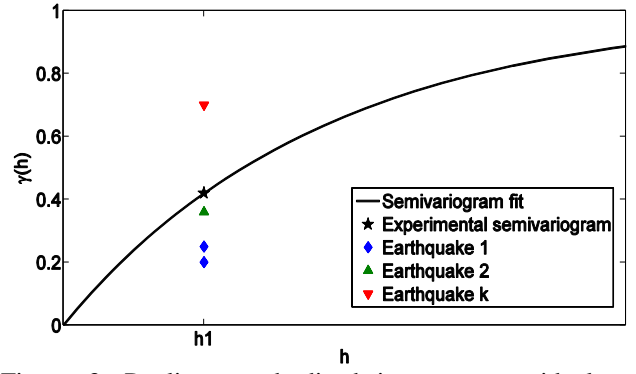


Figure 3. Pooling standardized intra-event residuals of multiple events ($j=1,2,\dots,k$) to compute experimental semivariogram.

Of the three basic models (Gaussian, spherical and exponential), the exponential model has been chosen to fit empirical points since this model is widely adopted in the literature. Moreover, the choice of using the same model for all periods allows to compare results and to investigate possible dependency of the model's parameters on the fundamental period.

Assuming that there is no nugget effect (as this study does not investigate variations at a smaller scale with respect to that of the tolerance); the only parameter to estimate is the range, b .

Least square method (LSM) was used as a reference to manually fit the model in the empirical semivariogram in order to give more importance to the small separation distances. In particular, correlation lengths evaluated for spectral acceleration for the seven periods, 0.1 s, 0.2 s, 0.3 s, 0.5 s, 1 s, 1.5 s, 2 s, resulted equal to 11.4 km, 9 km, 13.2 km, 11.9 km, 17.8 km, 25.7 km, and 33.7 km, respectively.

In Figures 5 to Figure 11 all estimated exponential models are shown together with data points referring to both classical and robust estimators.

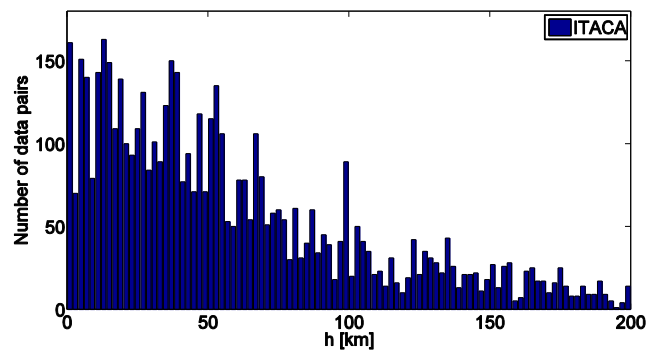


Figure 4. Histograms of the number of data pairs as a function of site-to-site separation distance.

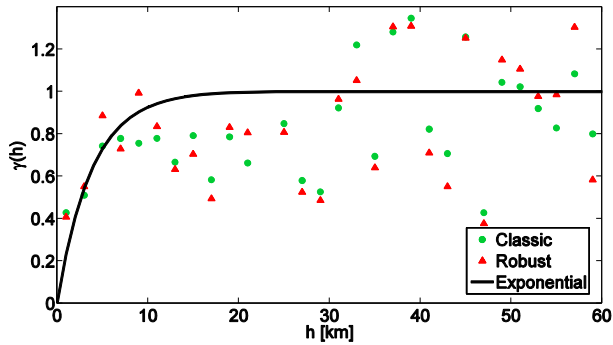


Figure 5. Empirical semivariogram and fitted exponential model for Sa(0.1 s).

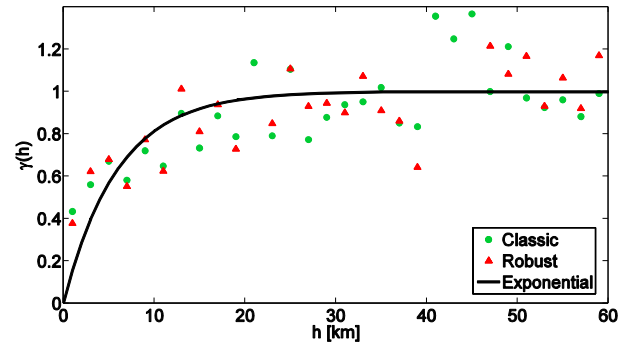


Figure 9. Empirical semivariogram and fitted exponential model for Sa(1 s).

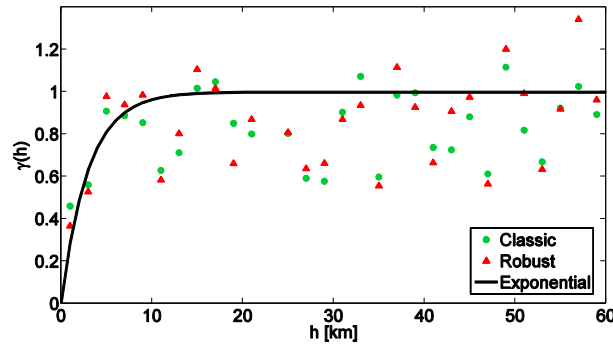


Figure 6. Empirical semivariogram and fitted exponential model for Sa(0.2 s).

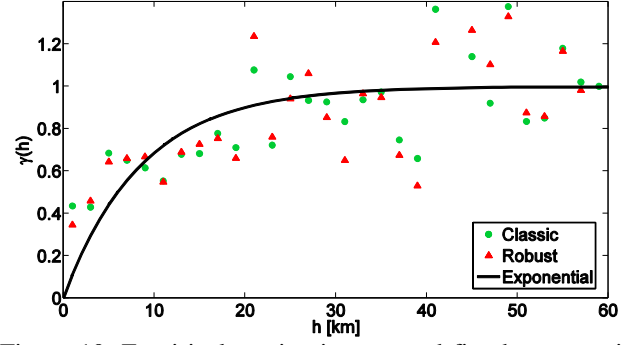


Figure 10. Empirical semivariogram and fitted exponential model for Sa(1.5 s).

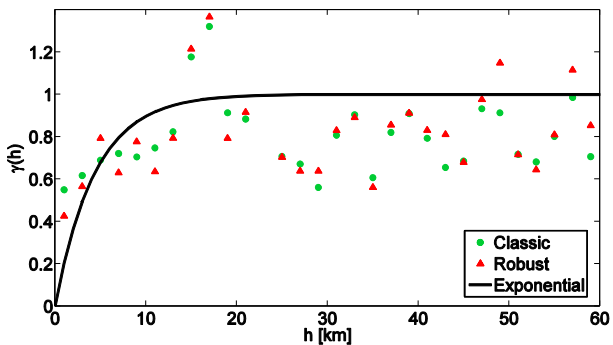


Figure 7. Empirical semivariogram and fitted exponential model for Sa(0.3 s).

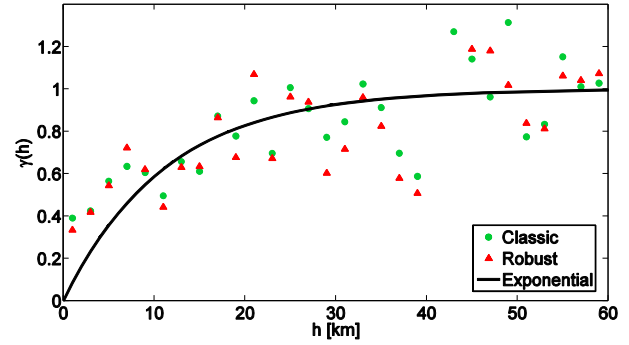


Figure 11. Empirical semivariogram and fitted exponential model for Sa(2 s).

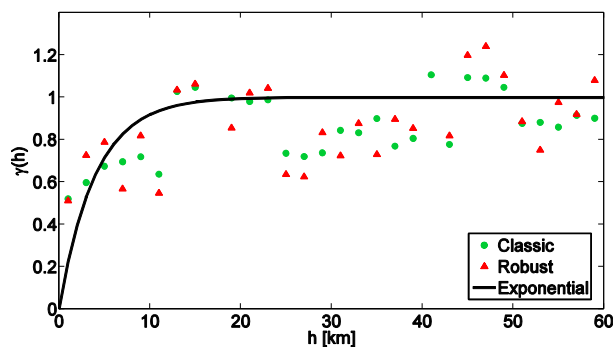


Figure 8. Empirical semivariogram and fitted exponential model for Sa(0.5 s).

Results indicate that correlation range tends to increase with period reaching the value of about 40 km for $T = 2$ s.

It should be noted that in Esposito and Iervolino (2011) the proposed methodology was used to estimate the horizontal peak ground acceleration (PGA) intra-event residuals' correlation starting from a less recent GMPE, the Bindi *et al.* (2010), and a larger dataset that includes the one used herein. For completeness, their spatial correlation model has been also estimated for horizontal PGA. The resulting range was similar to that of the mentioned study (Figure 12); i.e., 10.8 km with respect to 11.5 km, as expected.

The slight difference may also be related to the bin width used in the estimation of empirical

semivariograms (1 km instead of 2 km considered herein).

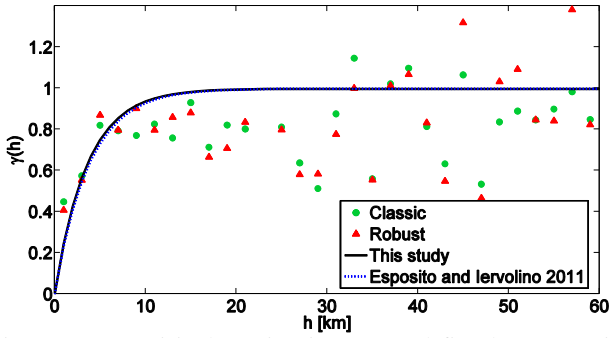


Figure 12. Empirical semivariogram and fitted exponential model for horizontal PGA compared with Esposito and Iervolino (2011).

5 DISCUSSION

Empirical results demonstrate that correlation length tends to increase with period. Except for high frequencies at which there is no a significant increment. This seems to be consistent with past studies of ground motion *coherency* (Zerva and Zervas 2002). In fact, the coherency describes the degree of correlation between amplitudes and phases angles of two time histories at each of their component frequencies. Considering that coherency decreases with increasing distance between measuring points and with increasing frequency, it may be reasonable to expect more coherent ground motion, as spectral acceleration evaluated at high periods exhibits more correlated peak amplitudes.

This same issue aspect was also discussed in Jayaram and Baker (2009) where in all earthquakes analyzed, the estimated ranges increased with period except for some cases.

In Figure 13 estimated ranges have been compared with some correlation lengths available in literature; those models have been chosen considering: Californian dataset from Goda and Hong (2008), “all earthquakes” model from Hong et al. (2009), and the “predictive model” based on all earthquakes from Jayaram and Baker (2009). The range b has been obtained evaluating the distance at which correlation is equal to 0.05.

Results provided by this study seem to be comparable with ranges estimated in literature in terms of both trend as a function of T , and the estimated value of correlation lengths.

This holds generally, except with respect to Goda and Atkinson (2009, 2010) models, in which ranges are larger (never below about 60

km) and the dependence of the correlation on period is not significant.

A simple linear predictive model (red dashed line in Figure 13) has been estimated using LSM method in order to obtain ranges based on the period of interest. The resulting linear model is expressed by Equation (11).

$$b(T) = 8.6 + 11.6 \cdot T \quad (11)$$

Based on this model, the correlation between normalized intraevent residuals separated by h is obtained as follows:

$$\rho(h, T) = e^{\left(-3 \cdot h / b(T)\right)} \quad (12)$$

derived from Equation 7 and Equation 8 where model parameters c_0 and c_e , the *nugget* and the *sill* respectively, are equal to one.

Starting from this predictive model it is possible to get the joint probability density function for $S_a(T)$ at all locations, for which it is required characterizing the covariance matrix, Σ , expressed in the Equation 3.

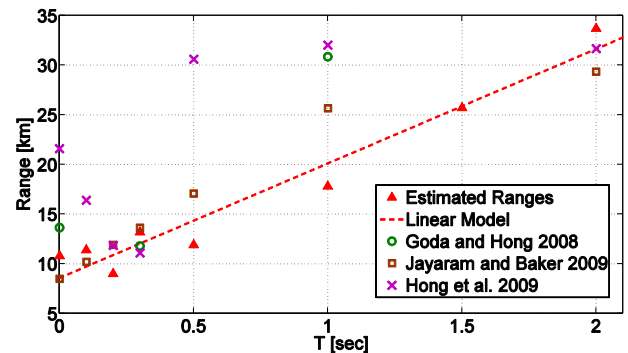


Figure 13 Estimated ranges in this study compared with some correlation lengths available in literature.

6 CONCLUSIONS

The study presented focused on the assessment of intraevent spatial correlation of spectral acceleration at eight periods ranging between 0 s and 2 s.

A subset of the Italian Accelerometric Archive has been used to compute residuals starting from a GMPE calibrated on the same dataset.

Consistent with the available literature on the topic, hypotheses of the stationarity and isotropy of the random fields were retained to compute experimental semivariograms of standardized intraevent residuals (with respect to the standard deviation estimated by the GMPE).

Moreover, because a relatively small number of records for each earthquake was available,

records from multiple events and regions within Italy were pooled to develop a unique model fitted with a large number of observations.

Exponential correlation models were calibrated by finding that practical ranges tends to increase with the period. The choice of using the same model (exponential) for all periods allows to compare results and to investigate possible dependency of the parameters on the period the spectral ordinate refers to.

Results have been also compared with previous researches finding, generally, a similar trend.

Finally a simple linear predictive model has been estimated in order to provide the correlation coefficient between spectral accelerations as a function of structural period.

ACKNOWLEDGMENTS

This study was partially developed in the activities of the Rete dei Laboratori Universitari di Ingegneria Sismica - ReLUIS for the research program funded by the Dipartimento della Protezione Civile 2010-2013 in a joint task with the European Community's Seventh Framework Programme [FP7/2007-2013] under the frame of the SYNERG project (project contract no. 244061). Authors want to thank Dr. Dario Bindi (*Deutsches GeoForschungsZentrum, Germany*) and Dr. Francesca Pacor (*Istituto Nazionale di Geofisica e Vulcanologia, Italy*), for kindly providing us with datasets used in this study.

REFERENCES

- Barnes, R. J., 1991. The Variogram Sill and the Sample Variance. *Mathematical Geology* **23** (4): 673-678.
- Bindi D., Pacor, F., Luzi, L., Puglia, R., Massa, M., Ameri, G., Lucci, R., 2011. Ground Motion Prediction Equations Derived from the Italian Strong Motion Data Base. *Bulletin of Earthquake Engineering* (Submitted).
- Bindi, D., Luzi, L., Rovelli, A., 2010. Ground Motion Prediction Equations (GMPEs) derived from ITACA. Deliverable No.14. Project S4: Italian Strong Motion Data Base, <http://esse4.mi.ingv.it>.
- Boore, D. M., Gibbs, J. F., Joyner, W.B., Tinsley, J. C., Ponti, D. J., 2003. Estimated ground motion from the 1994 Northridge, California, earthquake at the site of the Interstate 10 and La Cienega Boulevard bridge collapse, West Los Angeles, California. *Bulletin of the Seismological Society of America* **93** (6): 2737-2751.
- CEN, European Committee for Standardization, 2003. Eurocode 8: design provisions for earthquake resistance of structures, Part 1.1: general rules, seismic actions and rules for buildings, prEN 1998-1.
- Cressie, N., 1993. *Statistics for Spatial Data*. Revised edition. Wiley, New York, USA, 900 pp.
- Cressie, N., Hawkins, D.M., 1980. Robust estimation of variogram. *Mathematical Geology* **12** (2): 115-125.
- Esposito, S., Iervolino, I., Manfredi, G., 2010. Pga semi-empirical correlation models based on European data. *14th European Conference on Earthquake Engineering*. August 30 – September 3. Ohrid, Macedonia
- Esposito, S., Iervolino, I., 2011. Pga and Pgv Spatial Correlation Models Based On European Multievent Datasets. *Bulletin of the Seismological Society of America* (In press).
- Goda, K., Hong, H.P., 2008. Spatial correlation of peak ground motions and response spectra. *Bulletin of the Seismological Society of America* **98** (1): 354-365.
- Goda, K., Atkinson, G. M., 2010. Intraevent spatial correlation of ground-motion parameters using SK-net data. *Bulletin of the Seismological Society of America* **100** (6): 3055-3067.
- Hong, H.P., Zhang, Y., Goda, K., 2009. Effect of spatial correlation on estimated ground motion prediction equations. *Bulletin of the Seismological Society of America* **99**: 928-934.
- Jayaram, N., Baker, J.W., 2008. Statistical tests of the joint distribution of spectral acceleration values. *Bulletin of the Seismological Society of America* **98**: 2231-2243.
- Jayaram, N., Baker, J.W., 2009. Correlation model for spatially distributed ground-motion intensities. *Earthquake Engineering and Structural Dynamics* **38** (15): 1687-1708.
- Jayaram, N., Baker, J.W., 2010. Considering spatial correlation in mixed-effects regression, and impact on ground-motion models. *Bulletin of the Seismological Society of America* **100** (6): 3295-3303.
- Journel, A.G., Huijbregts, Ch. J., 1978. *Mining Geostatistics*. Academic Press, London, 600 pp.
- Malhotra, P., 2008. Seismic Design Loads from Site-Specific and Aggregate Hazard Analyses. *Bulletin of the Seismological Society of America* **98** (4): 1849-1862.
- Matheron, G., 1962. *Traité de géostatistique appliquée*. Editions Technip, Paris, France, 333 pp. (in French)
- McGuire, R.K., 2004. *Seismic Hazard and Risk Analysis*. Earthquake Engineering Research Institute, Oakland, MNO-10, 178 pp.
- Sokolov, V., Wenzel, F., Jean, WY., Wen, KL., 2010. Uncertainty and spatial correlation of earthquake ground motion in Taiwan. *Terrestrial, Atmospheric and Oceanic Sciences* **21** (6): 905-921.
- Strasser, F.O., Abrahamson, N. A., Bommer, J.J., 2009. Sigma: issues, insights, and challenges. *Seismological Research Letters* **80** (1): 40-56.
- Wang, M., Takada, T., 2005. Macrospatial correlation model of seismic ground motions. *Earthquake Spectra* **21** (4): 1137-1156.
- Zerva, A., Zervas, V., 2002. Spatial variation of seismic ground motion, *Applied Mechanics Reviewers* **55** (3): 271-297.