27 May 2010, 4:30 pm - 6:20 pm

Incorporation of the Spatial Correlation of Arias Intensity Within Earthquake Loss Estimation

Roxane Foulser-Piggott
Imperial College London, U.K.

Peter J. Stafford
Imperial College London, U.K.

Fifth International Conference on
Recent Advances in Geotechnical Earthquake Engineering and Soil Dynamics
and Symposium in Honor of Professor I.M. Idriss
May 24-29, 2010 • San Diego, California

INCORPORATION OF THE SPATIAL CORRELATION OF ARIAS INTENSITY WITHIN EARTHQUAKE LOSS ESTIMATION

Roxane Foulser-Piggott
Imperial College London
London SW7 2AZ, UK

Peter J. Stafford
Imperial College London
London SW7 2AZ, UK

ABSTRACT

Arias Intensity (I_a) has been identified as an efficient intensity measure for the purpose of estimating the likelihood and extent of landslides. This efficiency implies that Arias intensity may logically be used within earthquake loss estimation applications in order to ultimately estimate the damage to spatially-distributed systems or portfolios. In order to estimate the effects of ground motions on such spatially-distributed systems it is important to take into account the spatial correlation of the intensity measure. However, existing landslide loss-estimation models, which use I_a as an input, do not take this aspect of the ground motion into account. Due to the areal nature of landslides, accounting for the spatial distribution of I_a is important if one wishes to accurately predict the probability of landslides occurring, and their subsequent displacements. In this paper, a model for the spatial correlation of Arias intensity is proposed. In order to obtain this model, a new empirical prediction equation for Arias intensity is first developed. The empirical predictive model is developed using recordings from the PEER NGA database while the model for spatial correlation makes use of the well-recorded events from this database, i.e. the Northridge and Chi-Chi earthquakes.

INTRODUCTION

Earthquake-induced ground-motion cannot be adequately characterized by a single scalar measure for all conceivable applications. In order to comprehensively characterize a ground motion, and its associated damage potential, one must be able to quantify features of the ground motion that are associated with its energy and frequency content as well as the variation of these characteristics in time. However, for certain applications, it has been found that some of these characteristics are not as influential as others and in these cases a scalar representation of some characteristic of the ground shaking can be efficiently used to infer the likelihood of the motion to cause damage. Of the scalar intensity measures that have been proposed in the literature, Arias Intensity (Arias, 1970) is a measure that has been found to be well-suited to application in a number of problems in earthquake engineering. This utility results from the ability of Arias intensity to reflect multiple characteristics of the ground motion, despite being a scalar measure.

Travasarou *et al*. (2003) discuss the effectiveness of employing Arias intensity for the assessment of the seismic performance of structures whose response is dominated by the high-frequency content of a ground motion. Arias Intensity is also useful for predicting ground failure resulting from earthquakes as discussed in Kramer (1996). Egan and Rosidi (1991), Kayen and Mitchell (1997) and Travasarou *et al*. (2003), among others, have all discussed the utility of Arias intensity for estimating the propensity of a soil deposit to liquefy. However, of greatest relevance to the present article are the works, such as those of Harp *et al*. (1995), Keefer (2002) and Jibson (2007), that discuss the strong correlation that has been observed between Arias intensity and the distribution of earthquake-induced landslides.

In order to conduct earthquake loss analyses in terms of Arias intensity, a stable empirical ground-motion model must be available for use. However, very few models for this purpose have been derived (Stafford *et al*., 2009a). The most robust, and generally applicable, model that has been developed to date is that of Travasarou *et al*. (2003). This model was developed using recordings from the strong-motion database of the Pacific Earthquake Engineering Research (PEER) Center which are associated with a large number of shallow crustal earthquakes that have occurred throughout the globe. However, despite being the most robust of those currently available, the model of Travasarou *et al*. (2003) also has some shortcomings, primarily associated with the modeling of near-surface site response. For this reason, and to facilitate the development of new model for the spatial correlation of Arias intensity among sites, a new empirical model is developed and presented herein for the prediction of Arias intensity.

This paper firstly describes the development of the new empirical relationship for the prediction of Arias Intensity just mentioned. The final functional form gives an expression for Arias intensity in terms of the common predictor variables of magnitude, distance, style-of-faulting and the average shear-wave velocity over the upper 30 m. The most significant enhancement of the new model over existing models is the inclusion of the continuous variable, average shear-wave velocity, as a predictor variable representing local site conditions and the inclusion of terms to account for nonlinear site response.

In many situations, such as when estimating the impacts of landslides upon spatially distributed networks, it is not appropriate to only predict independent values of Arias intensity at a series of locations. Instead, knowledge of the joint probability of occurrence of Arias intensity values at multiple locations is required. In order to estimate losses to spatially distributed systems for a particular earthquake event, a model for the spatial correlation of an intensity measure is required. Recent focus has been placed upon the development of such models for a range of common intensity measures, *e.g.*, Wang and Takeda (2005), Goda and Hong (2008) and Jayaram and Baker (2009). However, no efforts have been directed towards the development of a spatial correlation model for Arias intensity.

In light of the above, the second part of the article is concerned with the development of a model to represent the correlation of Arias intensity values at spatially separated locations. The model is derived using recordings from well-recorded events within the PEER Next Generation of Attenuation (NGA) database (Chiou *et al*., 2008), the Chi-Chi and Northridge events. The new empirical model for the prediction of Arias intensity presented in this paper is used to obtain the intra-event residuals which are then used in the development of the spatial correlation model.

STRONG-MOTION DATASET

The dataset used for the derivation of the predictive equation for Arias intensity, as well as for the development of the spatial correlation model, is a subset of the PEER NGA database. The complete NGA database consists of 3551 accelerograms from 173 earthquakes (Chiou *et al*., 2008). However, not all of these records have associated metadata and for this reason the total dataset was restricted using certain criteria. The actual subset that was used is the same as that which has recently been used for the development of other empirical models for numbers of cycles (Stafford and Bommer, 2009), duration measures (Bommer *et al*., 2009), and envelope parameters (Stafford *et al*., 2009b).

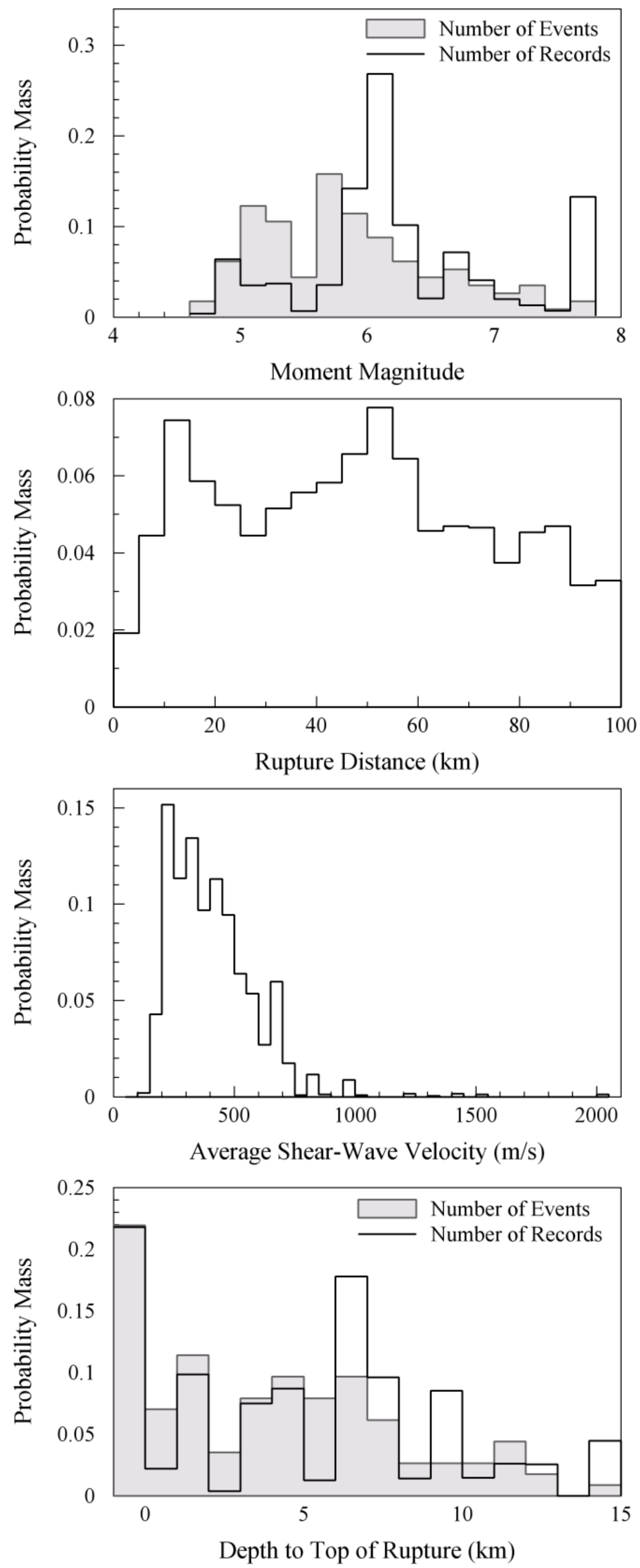

Fig. 1. Distribution of independent variables used for the development of the predictive model for Arias intensity. Modified from Stafford and Bommer (2009).

The article of Stafford and Bommer (2009) should be consulted for specific details regarding the actual dataset. However, for now it suffices to say that the general philosophy adopted by Abrahamson and Silva (2008) was applied, but with the main difference being that the distance range was limited to 100 km. The final dataset used in this study contains 2406 recordings from 114 earthquakes, each of which have two orthogonal horizontal components. The distributions of these records with respect to the primary predictor variables that were used in the development of the empirical model for Arias intensity are shown in Fig. 1. It should also be noted that, of these 114 earthquakes, 23 events have normal or normal-oblique mechanisms, 35 have reverse or reverse-oblique mechanisms and the remaining 56 earthquakes are strike-slip events. The records have been used in their processed form and are freely available from the following website: http://peer.berkeley.edu/products/nga_project.html.

MOTIVATION FOR DEVELOPING A NEW MODEL FOR ARIAS INTENSITY

Before commencing development of a new model for Arias intensity it is instructive to investigate the performance of the existing model of Travasarou *et al*. (2003) – hereafter referred to as TBA. The datasets used in the present study and that used by TBA share many similarities, with the main difference being the availability of shear-wave velocity in the latter case. However, despite these similarities, it is found that the performance of the TBA model when applied to the dataset of this study is not particularly good. An example of the performance can be illustrated in two ways. Table 1 shows the coefficients of the TBA model that are obtained when the functional form used by these authors is used with the dataset of the present study. As can be seen, the values of many of the coefficients differ markedly. Importantly, the new calibration of these parameters indicates that only 4 out of the 11 coefficients are significant at the 95% confidence level. In particular, the terms that are used to account for the site response and style-of-faulting are far from being significant.

The practical impact of these differences is demonstrated in Fig. 2 in which median predictions from these two models are compared. In this particular case one can observe differences that are of the order of a factor of 2 or more. No effort has been made to replicate the variance structure used by TBA, but inspection of Fig. 3 indicates that the homoskedastic variance model differs quite significantly from that employed by Travasarou *et al*. (2003) (this will account for some of the difference in the determined coefficients, but will not be a strong contributor).

A plot of the residuals from the model derived using the functional form of TBA and the dataset of this study against the shear-wave velocity (Fig. 4) indicates that the model for site response employed by TBA does a very poor job of representing the scaling of ground motions with respect to the shear-wave velocity.

Table 1. Comparison of coefficients using the functional form of Travasarou et al. (2003) and the dataset of this study (NGA). Variance components of the TBA model are intensity, magnitude, and site class dependent (hence the ranges).

Coef.	TBA (2003)	NGA	p-value	Sig.
c_1	2.8	3.9626	0.0000	Y
c_2	-1.981	-1.7523	0.2329	N
c_3	20.72	20.7610	0.0195	Y
c_4	-1.703	-1.9344	0.0000	Y
h	8.78	9.0455	0.0000	Y
s_{11}	0.454	-0.1299	0.0544	N
s_{12}	0.101	-0.1476	0.0759	N
s_{21}	0.479	0.0185	0.7435	N
s_{22}	0.334	-0.0289	0.6721	N
f_1	-0.166	-0.0945	0.6544	N
f_2	0.512	0.1756	0.2863	N
σ_E	0.475-0.611	0.6126		
σ_A	0.730-1.180	0.9984		
σ_T	0.871-1.329	1.1714		

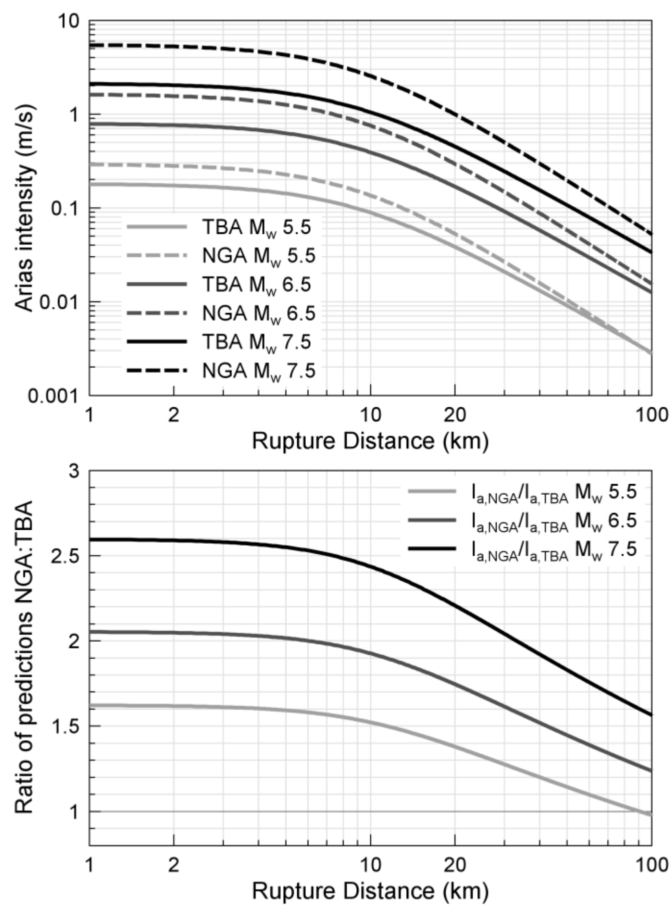

Fig. 2. Comparison of the median predictions of the TBA model and that derived in this study using the TBA functional form. Predictions are for a strike-slip event on rock. The bottom panel shows the ratio of the predictions.

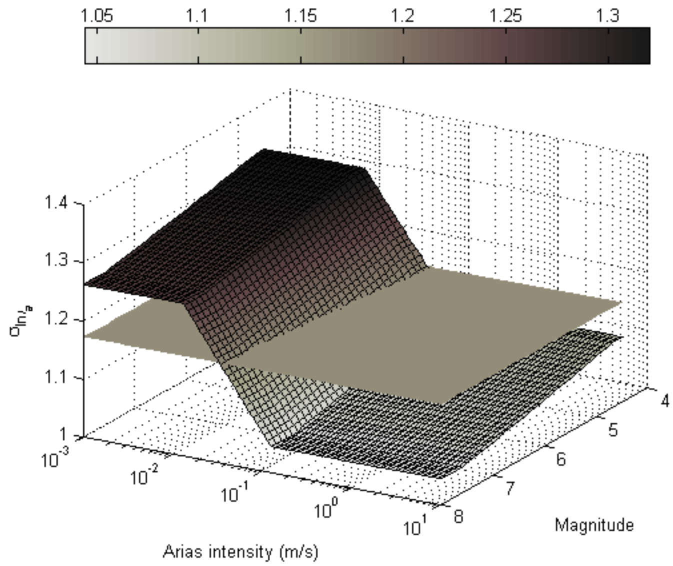

Fig. 3. Comparison between the heteroskedastic variability model TBA with the homoskedastic model using their functional form and the NGA database.

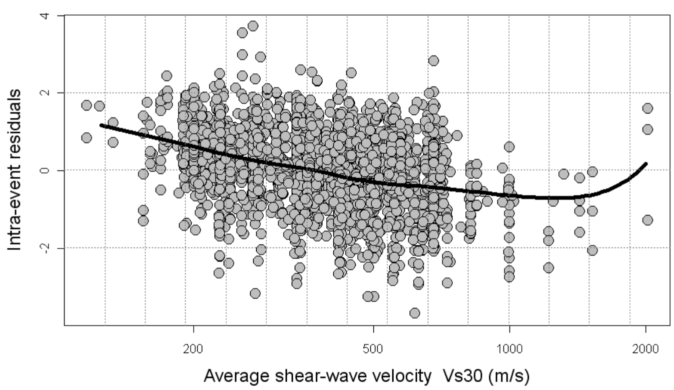

Fig. 4. Intra-event residuals plotted against Vs30 for the model derived using the NGA dataset and the TBA functional form.

The combination of these findings strongly suggests that the development of a new model is warranted prior to the derivation of a correlation model for Arias intensity.

PREDICTIVE MODEL FOR ARIAS INTENSITY

The studies of Travasarou *et al*. (2003) and Stafford *et al*. (2009) have both provided theoretical arguments to support the use of particular functional forms. Through the application of numerical strategies, Travasarou *et al*. (2003) demonstrated that the influence of stress drop on the Arias intensity led to a nonlinear scaling with respect to magnitude. The analytical approach adopted by Stafford *et al*. (2009) does not suggest such scaling and results in only linear terms. Following the theoretical considerations both groups of authors just mentioned then deviate from their base models in order to incorporate known features of earthquake ground-motion scaling such as near-source saturation, site response and style-of-faulting. Therefore, for the purposes of developing a new model for Arias intensity we take note of the theoretically constrained functional forms previously exposed but also adopt functional expressions that have recently been used for the derivation of predictive models for other ground-motion measures.

It is clear from the previous section that the most significant short-coming of existing models is their inability to adequately capture the modification of Arias intensity values that occurs as waves pass through near-surface deposits. The incorporation of the effects of near-surface geology in previous models for the prediction of Arias Intensity all make use of dummy variables for site classes. However, the model of Travasarou *et al*. (2003) attempts to incorporate some nonlinearity into these site class terms through the use of magnitude dependence. Similarly, Stafford *et al*. (2009) incorporate this effect more directly by developing a model that has site class terms that are dependent upon the strength of the predicted Arias intensity for rock site conditions. This latter approach is based upon the work of Abrahamson and Silva (1997) who were the first to incorporate nonlinear site response characteristics into empirical ground-motion models.

In the present study, recourse is again taken to the developments in empirical ground-motion modeling that have arisen through consideration of other intensity measures. The NGA project that has recently been completed has resulted in a suite of predictive models for spectral ordinates that all feature functional terms to account for nonlinear site response. Given that values of Arias intensity are heavily governed by the amplitudes of an accelerogram and that strong correlations have been found to exist between Arias intensity and peak ground motions (Baker, 2007; Stafford *et al*. 2009b) it is reasonable to initially adopt functional forms that are similar to those used by the NGA model developers.

The first step in modeling nonlinear site response is to obtain the functional form for soil amplification effects as a function of the median Arias Intensity on some reference site condition, I_a^{ref}. This reference site condition is typically chosen to correspond to rock sites upon which nonlinear effects should be minimal. During the process of predicting the Arias intensity, this reference ground motion is first predicted and then treated as an independent variable for the purposes of obtaining the final value of Arias intensity associated with the soil surface motion. Abrahamson and Silva (1997) first implemented a model that had a functional form of the type shown in Equation (1).

$$\ln\left(\text{Amp}\right) = \ln\left(\frac{I_a^{\text{soil}}}{I_a^{\text{rock}}}\right) = a + b\ln\left(I_a^{\text{rock}} + c\right) \tag{1}$$

The coefficients *a, b* and *c* can be interpreted in the following way: *a* represents a linear soil amplification factor that applies when the input rock motion is weak (actually the linear response is equivalent to the expression $a + b\ln(c)$ when the rock motion is very small); the coefficient *c* represents the reference ground-motion level, or 'corner' at which the transition from linear to nonlinear soil behavior occurs; and, finally, the coefficient *b* is the gradient of the amplification factor against reference ground-motion above the 'corner' in log-log space. Coefficient *b* therefore represents nonlinear soil behaviour with site amplification decreasing with increasing amplitude of the reference Arias intensity (*e.g.*, Walling *et al.,* 2008).

Chiou and Youngs (2008) have implemented a modified form of the expression shown in Equation (1) and the investigations made as part of the current study indicate that this form performs well when used to model the nonlinear site response of Arias intensity values. Before this functional form was blindly adopted, an external check was performed to see if the general scaling of Arias intensity followed the form of peak ground-motion parameters, including spectral ordinates. Recently, Papaspiliou (unpublished PhD thesis; Imperial College London) has conducted extensive parametric site response analyses in order to assess the nonlinear scaling of spectral amplitudes. One step of this process is to pass accelerograms through some reference soil profiles in order to obtain surface motions given some input motion at an assumed bedrock. A subset of the motions that have been used for these parametric analyses were taken and Arias intensity values were computed for both the bedrock and surface motions. From these calculations it was determined that the general scaling of Arias intensity resulting from nonlinear site response follows essentially the same form as models previous proposed for peak ground-motions (see Fig. 5). This exercise gave us confidence that the functional form of adopted by Chiou and Youngs (2008) could, at least, be tried and also provided very good staring estimates for the ensuing nonlinear random effects regression analyses.

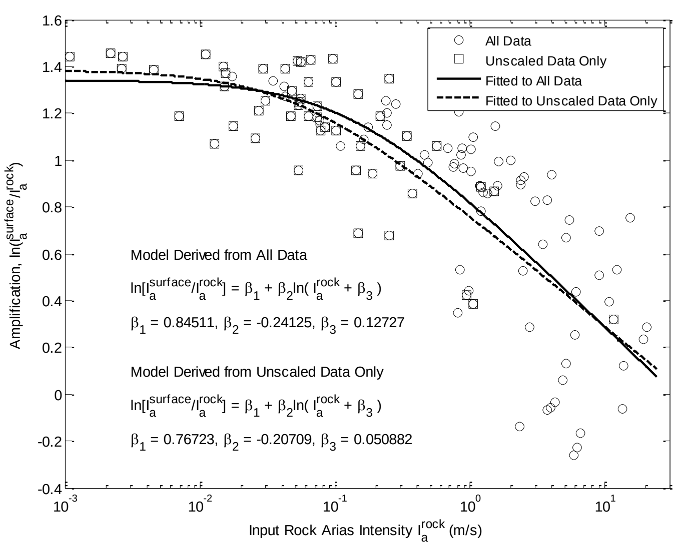

Fig. 5. Generic nonlinear scaling of Arias intensity for a soil site with an average shear-wave velocity of 280 m/s.

The regression analysis consisted of initially trying various functional forms that were basic variations of the theoretically governed forms presented in Stafford *et al*. (2009) and Travasarou *et al*. (2003). Given that analytical, theoretical, considerations suggest that the logarithm of Arias intensity should scale linearly with magnitude this functional form was adopted. However, it is also clear that some nonlinearity in magnitude scaling can be observed. This nonlinearity can be captured through the use of both nonlinear site response terms as well as through the use of magnitude-dependent geometric spreading terms. As there is no pervasive physical reason why one should employ nonlinear magnitude scaling we opt for the use of a basic functional form that employs magnitude-dependent geometric spreading and then also accommodate nonlinearity in the site response.

All functional forms were evaluated using standard statistical metrics that are automatically provided through the use of the nlme package (Pinheiro *et al*., 2008) of the free software R (http://www.r-project.org/). In all cases the presented coefficients were found to be statistically significant with the exception of two of the terms within the nonlinear site response component of the model. However, as these parameters have a clear physical interpretation, and their *p*-values were still less than 0.1, *i.e.*, these coefficients were significant at the 90% confidence level, they were retained in the model.

The final functional form is presented in Equations (2) to (4). Equation (2) represents the generic expression that is used to obtain the logarithm of the Arias intensity for any given scenario.

$$\ln(I_a) = \ln\left(I_a^{\mathrm{ref}}\right) + f_{\mathrm{site}}\left(V_{s30}, I_a^{\mathrm{ref}}\right) \tag{2}$$

$$\begin{aligned}\ln\left(I_a^{\mathrm{ref}}\right) &= c_1 + c_2\left(M_{\mathrm{w}} - 6\right) \\ &\quad + \left(c_3 + c_4 M_{\mathrm{w}}\right)\ln\left(\sqrt{R_{\mathrm{RUP}}^2 + c_5^2}\right) + c_6 F_{\mathrm{RV}}\end{aligned} \tag{3}$$

$$\begin{aligned}f_{\mathrm{site}}\left(V_{s30}, I_a^{\mathrm{ref}}\right) &= v_1 \ln\left(\frac{V_{s30}}{V_{\mathrm{ref}}}\right) \\ &\quad + v_2\left[e^{v_3\left(V_{s30} - V_1\right)} - e^{v_3\left(V_{\mathrm{ref}} - V_1\right)}\right]\ln\left[\frac{I_a^{\mathrm{ref}} + v_4}{v_4}\right]\end{aligned} \tag{4}$$

In Equation (4), the two parameters V_{ref} and V_1 are equal to 1100 and 280 m/s respectively. All other parameters are found through the regression analysis and are presented in Table 2. In Table 2, the actual parameter values are provided along with the standard errors in the estimates which are used to construct confidence intervals on the parameter values. As previously mentioned, two parameters v_3 and v_4 are not statistically significant at the 95% confidence level. However, as can be seen in Table 2, these coefficients are significant at the 90% level and are retained in the model.

Table 2. Coefficients for the new model for Arias intensity outlined in Equations (2) to (4). The variance components σ_E, σ_A, and σ_T represent the inter-event, intra-event and total standard deviations respectively.

Coefficient	Value	Standard Error	Significant @ 95%	Significant @ 90%
c_1	3.5987	0.2979	Y	Y
c_2	1.3015	0.2293	Y	Y
c_3	-3.3901	0.3559	Y	Y
c_4	0.1852	0.0565	Y	Y
c_5	5.3239	0.9945	Y	Y
c_6	0.3688	0.1711	Y	Y
v_1	-1.1331	0.056	Y	Y
v_2	-1.033	0.408	Y	Y
v_3	-0.001	0.0006	N	Y
v_4	0.1425	0.0886	N	Y
σ_E	0.7042	0.0755	Y	Y
σ_A	0.8983	0.0133	Y	Y
σ_T	1.1414	0.0391	Y	Y

Note that the model that has been presented herein in homoskedastic. Further work is required to assess whether a more robust model may be obtained by incorporating heteroskedasticity into the variance components. In theory, there should at least be some nonlinear soil dependence in the aleatory variability, as demonstrated by Chiou and Youngs (2008).

Figure 6 shows the residual plots that have been obtained from the nonlinear random effects regression analysis. Visual inspection of these residuals suggests that the functional form is performing well and that there are no significant trends with respect to the predictor variables. These residual plots, and in particular the lower panel of Fig. 6, can be contrasted against the residuals shown earlier in Fig. 4. It is clear that the use of the continuous predictor variable of shear-wave velocity has enabled the site response to be adequately captured.

It is not possible, simply from looking at Fig. 6, to identify obvious signs that the use of a heteroskedastic variance structure would improve the quality of the fit that has been obtained. However, formal statistical analyses remain to be conducted in this regard and it may well be that such tests indicate that the use of a heteroskedastic variance structure would lead to an improved model. For the present study, and keeping in mind that the primary objective of deriving this new model for Arias intensity is actually to enable a spatial correlation model for this intensity measure to be developed, the model performance appears perfectly adequate.

Figure 7 shows a comparison between the generic scaling of the new model with respect to both magnitude and distance and compares this to the scaling of the TBA model. It is very clear from this figure that very significant differences exist between these two models.

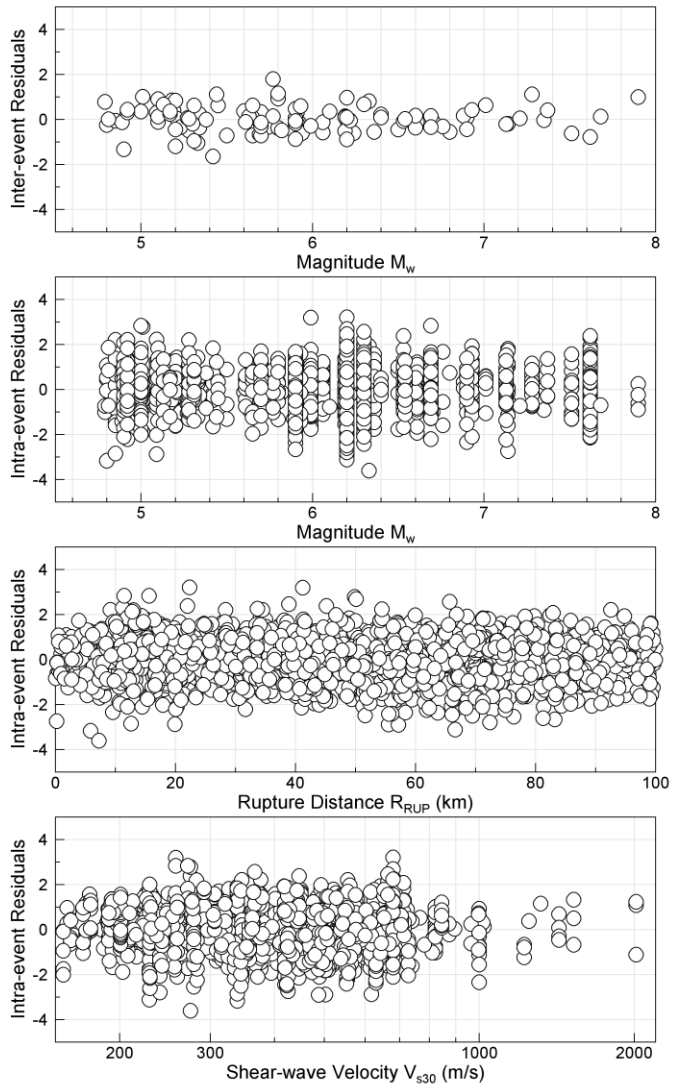

Fig. 6. Residual plots. The top panel shows inter-event residuals while the three remaining panels are for intra-event.

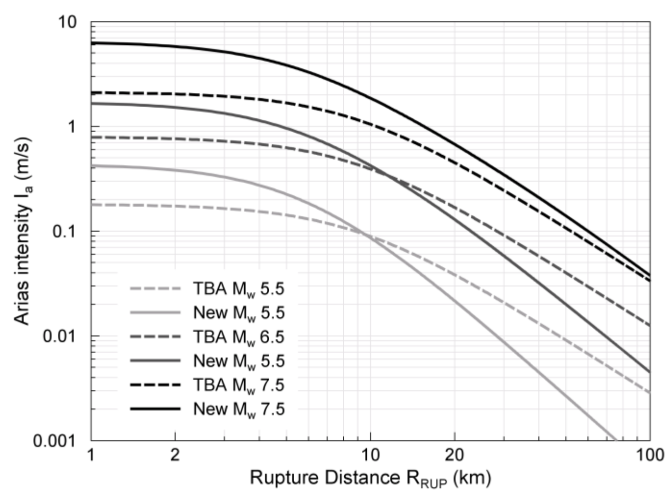

Fig. 7. Comparison of the median predictions of the new model with those of Travasarou et al. (2003). The predictions for the new model are provided for V_{s30}=760 m/s.

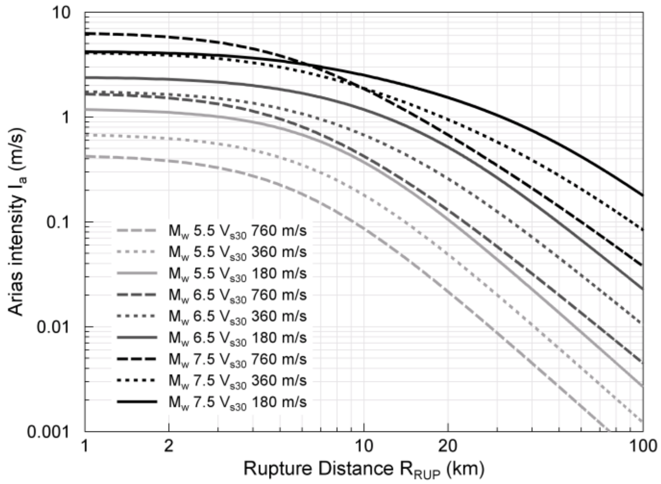

Fig. 8. Influence of site conditions on the median prediction of Arias intensity. Note the evidence of nonlinear site response that is apparent for the largest considered earthquakes.

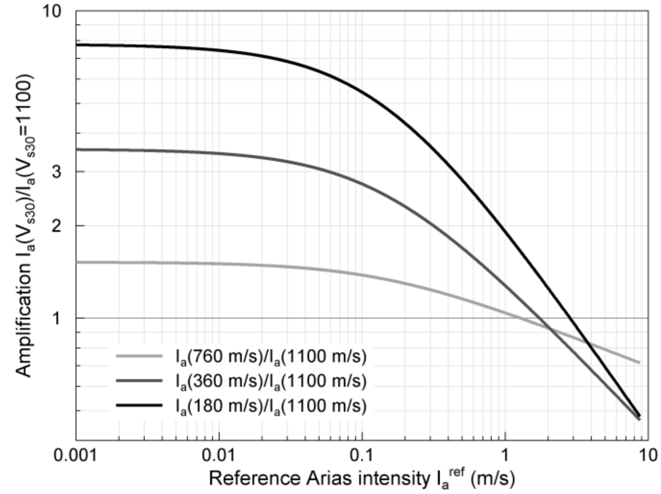

Fig. 9. Demonstration of the nonlinear scaling of amplification with reference Arias intensity. The three shear-wave velocities that are presented correspond to the NEHRP boundaries.

The obvious differences that are observed in Fig. 8 act as very strong evidence in support of the adaptation of the new model. In particular, Jayaram and Baker (2009) have recently pointed out that biased estimates of spatial correlation can be obtained if the spatial correlation among soil deposits is not adequately accounted for. Spatial correlation among soil deposits can exist in reality for a number of geological reasons. However, such correlations can also arise artificially through the use of an inappropriate predictive model that does not capture the scaling of Arias intensity with respect to soil conditions. In this latter case one may observe clusters of motions higher or lower than expected as a result of a general biased in the model itself. The newly developed model is therefore far more likely to lead to a robust model for spatial correlation than the existing model of Travasarou *et al*. (2003).

Figure 8 demonstrates the influence that the shear-wave velocity has upon the median predictions of the new model. It is clear from inspection of this figure that site effects are very significant for Arias intensity as both linear and nonlinear scaling is very evident for the cases shown here. In particular, it can be noted that the general decay of amplitudes with respect to distance for the case of the small M_w5.5 earthquakes is constant. This implies that the site response is linear. In contrast, the rate of decay with respect to distance for the larger M_w7.5 events varies significantly. At short source-site distances in this latter case the amplitudes of the Arias intensity predictions tend to saturate and some de-amplification can also be observed.

Although the impact of nonlinear site response can be observed from Fig. 8, it is more clearly seen in Fig. 9 in which the amplification of Arias intensity is plotted with respect to the Arias intensity predicted for the reference site conditions of V_{s30} = 1100 m/s. In this figure the transition from linear to nonlinear site response is clearly seen. Furthermore, for all considered site classes de-amplification occurs once the reference Arias intensity exceeds 1 m/s.

SPATIAL CORRELATION MODEL

Now that the new model for Arias intensity has been presented we may turn our attention to the derivation of a model for the spatial correlation of Arias intensity values. In order to derive such a model we more or less follow the procedure outlined comprehensively in Jayaram and Baker (2009). In cases where insufficient detail is presented herein the article of Jayaram and Baker (2009) should be consulted. The key steps in the process, the derivation of key equations and the assumptions made for this study are outlined in the following sections. The focus of the following section is to develop and present the model for spatial correlation as well as to note points of departure from the process adopted by Jayaram and Baker (2009).

The Arias intensity at an individual site may be written as the sum of a median model prediction and two error components that together represent the total difference between the observed motion and the model prediction. This representation is shown mathematically in Equation (5).

$$\ln I_{a,ij} = \ln \hat{I}_{a,ij} + \eta_i + \varepsilon_{ij} \tag{5}$$

where the observed Arias intensity is represented by $I_{a,ij}$, the median Arias intensity is represented by $\hat{I}_{a,ij}$, η_i and ε_{ij} are the inter- and intra-event residuals and the subscripts i and j are indicies defining the event and recording respectively. For a single event the inter-event residual is common to all sites. Therefore, in order to assess the degree to which spatially separated sites are similarly above or below the median predicted value the only component of Equation (5) that can be interrogated is the intra-event residual ε_{ij}.

As shown by Jayaram and Baker (2009), a model for spatial correlation may be developed through construction of empirical semivariograms for the intra-event residuals of well-recorded earthquakes. Cressie (1985) has presented two different formulations that may be used to define the empirical semivariogram for a set of data; the classical and robust estimators. The classical estimator was used by Jayaram and Baker (2009) and takes the following form, in which z_{u_i} would represent an intra-event residual at position u_i, h is the separation distance between sites and N_h is the number of pairs of sites having this separation distance. The semivariogram itself is given by $\hat{\gamma}(h)$.

$$\hat{\gamma}(h) = \frac{1}{2N_h}\sum_{i=1}^{N_h}\left(z_{u_i+h} - z_{u_i}\right)^2 \tag{6}$$

Similarly, the robust estimator is defined as follows:

$$\hat{\gamma}(h) = \frac{\left(\frac{1}{N_h}\sum_{i=1}^{N_h}\left|z_{u_i+h} - z_{u_i}\right|^{1/2}\right)^4}{0.914 + 0.988/N_h} \tag{7}$$

Both estimators lead to similar empirical semivariograms, but the robust estimator is less sensitive to outliers. An example of the results obtained following application of both estimators to normalized intra-event residuals from the Chi-Chi mainshock are shown in Fig. 10. In this study we prefer to adopt the robust estimator over the classical estimator on the basis that it should perform better for events that are not particularly well-recorded and for which outliers are likely to be observed.

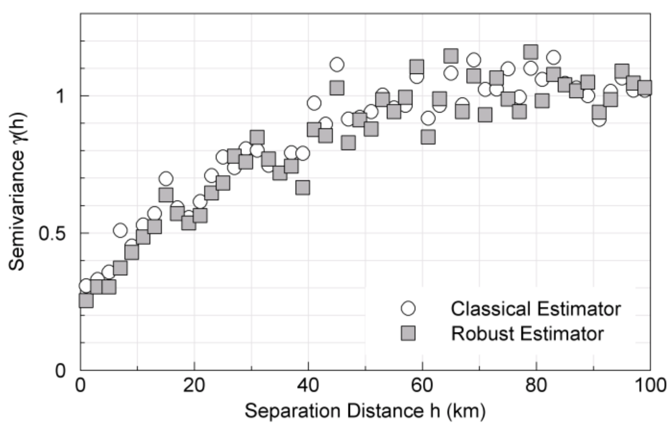

Fig. 10. Example empirical semivariograms computed using the classical and robust estimators of Cressie (1985). The example shown is for the Chi-Chi mainshock.

The application of the robust estimator results in a set of empirical values for the semivariogram at particular separation distances. The actual separation distances that are considered in practice are defined for discrete bins. In this study a bin size of 2 km was adopted after some cursory sensitivity checks.

In order to develop a predictive model for the spatial correlation of Arias Intensity at any two sites separated by h, a continuous function must first be fitted to the empirical semivariograms that are obtained for each considered event. Jayaram and Baker (2009) considered a series of common models for semivariograms that can generally be defined as a function of just two parameters: the sill of the semivariogram, a, and the range of the semivariogram, b. It is this possible to write a generic expression for common spatial correlation models as in Equation (8):

$$\gamma(h) = f(h; a, b) \tag{8}$$

As noted by Jayaram and Baker (2009), once a model for the semivariogram has been obtained, the model for the spatial correlation follows directly from the expression in Equation (9) in which $\rho(h)$ represents the correlation model that we eventually desire.

$$\gamma(h) = a\left[1 - \rho(h)\right] \tag{9}$$

It should be noted that the sill of a semivariogram is equivalent to the variance of the intra-event residuals when no spatial correlation is taken into account. Therefore, if one were to work directly with normalized intra-event residuals the expected variance of these residuals would be equal to unity and the expression in Equation (9) would simplify even further to become simply a function of a single parameter; the range of the semivariogram.

EXPERIMENTAL SEMIVARIOGRAMS FROM THE NORTHRIDGE AND CHI-CHI EARTHQUAKES

In this paper, the well recorded earthquakes of Northridge and Chi-Chi were used to investigate the feasibility of developing a model for the spatial correlation of Arias Intensity. For each earthquake, the coordinates of the recording stations were used to compute the separation distances of all sites and distance bins h_i were defined to have width δ_h = 2 km. Pairs of sites separated by distance h_i were then grouped into bins and the empirical semivariograms subsequently derived.

As previously mentioned, it is advantageous to work with the normalized intra-event residuals for the two earthquakes. Furthermore, it would be most desirable if the variance estimates for all earthquakes were identical (as is assumed in regression analyses). However, it is possible for the residuals from an individual earthquake to have a variance that differences from the intra-event variance of a predictive model and it is also possible for the predictions for a subset of observations from an individual event to also be biased. The use of inter-event residuals should act to ensure that the mean logarithmic motions for a particular event are equal to zero. However, in some cases one does not make use of all available records from a particular earthquake (it is possible for isolated sites to have large separation distances from all other sites and there is no point in including this observation among those used to develop the correlation model. For these reasons, the first step that was taken was to check that the mean and standard deviations of the residuals from each of the two considered events were very close to the values dictated from the regression analysis. While the mean residuals were found to be very close to zero (as expected) it was found that the standard deviations of the residuals from these two events were significantly lower than the intra-event standard deviation found from the regression analysis. The standard deviation of the residuals for the Chi-Chi and Northridge events were 0.768 and 0.732 respectively while the intra-event standard deviation of the model is 0.898. This discrepancy is most likely associated with the influence that some of the recordings from the Chi-Chi aftershocks are having in terms of inflating the variance and may also be due to the fact that these earthquakes are very-well studied and as such the metadata is very good for these events. Both of these aspects are the subject on ongoing investigation and it is likely that some modification to the variance structure of the proposed Arias intensity model will be required.

In order to overcome the problems associated with different standard deviations we normalized the intra-event residuals by their group-specific standard deviation. This normalization should act to ensure that the variance implied by the semivariogram tends to unity at large separation distances and also results in correlation values that tend to zero as the separation increases (as we would expect).

In addition to checking that the mean and variance of the residuals for the individual events were consistent with what we expected, it was also necessary to check that there were no trends in Chi-Chi and Northridge residuals when plotted against distance. When undertaking the regression analyses it is implicitly assumed that the distance scaling for all earthquakes is the same. It should not be surprising to learn that this is not generally the case. Earlier it was mentioned that it is important to appropriately model the nonlinear site response in order to ensure that artificial correlations are not implied through biased model predictions that systematically lead to groups of events with higher-than or lower-than average motions. For the same reason it is important to check that the distance scaling is appropriate for the individual events. This check was performed for both events and no statistically significant trends in the residuals with respect to distance were found.

After inspection of the empirical semivariograms, it appears appropriate to truncate the dataset at h = 100 km for the Chi-Chi event and to truncate at h = 50 km for the Northridge event. The reasons for this are threefold. Firstly, when fitting the model to the empirical data it is important to model the structure of the semivariogram well at small separation distances. Secondly, large separation distances (h >100km) are associated with low correlations which will have little effect on the joint distributions of Arias Intensity (Jayaram and Baker, 2009). Finally, it was found that spurious values were obtained when large separation distances were considered. Given that these values are of almost zero relevance to all practical cases it was decided to limit the range considered for the development of the models.

Jayaram and Baker (2009) outlined some of the most common models that have been fitted to spatial data in various fields. Of the models that they considered, they decided that the exponential model was the most appropriate for general application but found that common fitting procedures were not optimal for obtaining the model parameters. The reason for this is that while an optimal fit may be obtained in a statistical sense, such a fit will be governed by the performance of the fitted model over the full range of data that is considered. When the primary interest is upon the correlations at short separation distances it is desirable to ensure that a good fit is obtained in this region as a priority. To this end, Jayaram and Baker (2009) employed a manual fitting approach in which the model parameters were simply selected using visual judgement.

In the present study we prefer not to use the manual fitting procedure and instead employ a weighted least squares approach that has been proposed by Cressie (1985). This approach systematically gives higher weight to the observations at small separation distances and also takes into account the differing numbers of observations (pairs) that are contained in a bin at a given separation distance. The formal derivation is provided in Cressie (1985), but the method comes down to minimizing the loss function give in Equation (10).

$$\sum_{j=1}^{k} N_{h_j} \left[\frac{\hat{\gamma}(h_j)}{\gamma(h_j; \boldsymbol{\lambda})} - 1 \right]^2 \qquad (10)$$

In Equation (10), $\boldsymbol{\lambda}$ represents the vector of parameters that define the spatial correlation model and that are modified in order to minimise the loss function given in this expression. The term $\gamma(h; \boldsymbol{\lambda})$ represents the correlation model and in the present article two alternative models are considered. The first model is that preferred by Jayaram and Baker (2009) and is referred to as the exponential model. This model is given here in Equation (11).

$$\gamma(h) = a\left[1 - \exp\left(-3\frac{h}{b}\right)\right] \qquad (11)$$

In the case that normalized residuals are used we should normally expect the sill, a, to be very close to unity. Therefore, in practice the expression in Equation (11) really only has one free parameter. In order to consider another model that has more freedom to adapt itself to the empirical data a sigmoid function was employed. The sigmoid function is particularly flexible and can also be configured to have the desirable characteristics of tending to zero as the separation distance decreases and tending to unity as the separation distance increases. The sigmoid function for this application has just two parameters, just as the expression in Equation (11) does, but both parameters are able to make a significant contribution to the general form of the function. The expression for the sigmoid model is given below in Equation (12).

$$\gamma(h) = \frac{1}{1 + \exp\left[-\dfrac{\ln(h) + \phi_1}{\phi_2} \right]} \tag{12}$$

Examples of the models that have been obtained through the use of these two functional forms are given in Fig. 11. In this figure the data from the Chi-Chi and Northridge events are shown and the empirical semivariogram is determined using the robust estimator.

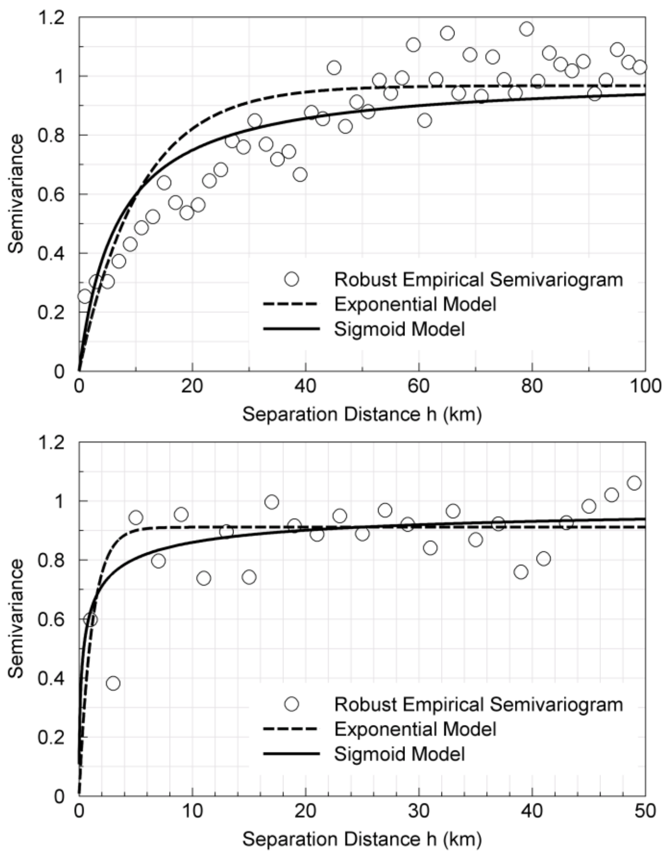

Fig. 11. Example fits of the spatial correlation models of Equations (11) and (12) to the Chi-Chi (upper panel) and Northridge (lower panel) earthquakes.

Very different model fits are obtained for both of these earthquake events. In both cases the fitted models do an adequate job of capturing the general features of the scaling of the correlation with separation distance. However, it is also clear that the exponential model significantly over-estimates the semi-variance of the Chi-Chi event over a considerable range of values that can be of potential importance to an earthquake loss analysis. The sigmoid model performs marginally better, but is still not able to capture the specific features of the empirical semivariograms.

In order to present a clearer picture of how these models are performing in the regions of greatest relevance to most engineering applications, *i.e.*, at very short separation distances, Fig. 11 is replotted using a logarithmic abscissa and is presented here as Fig. 12.

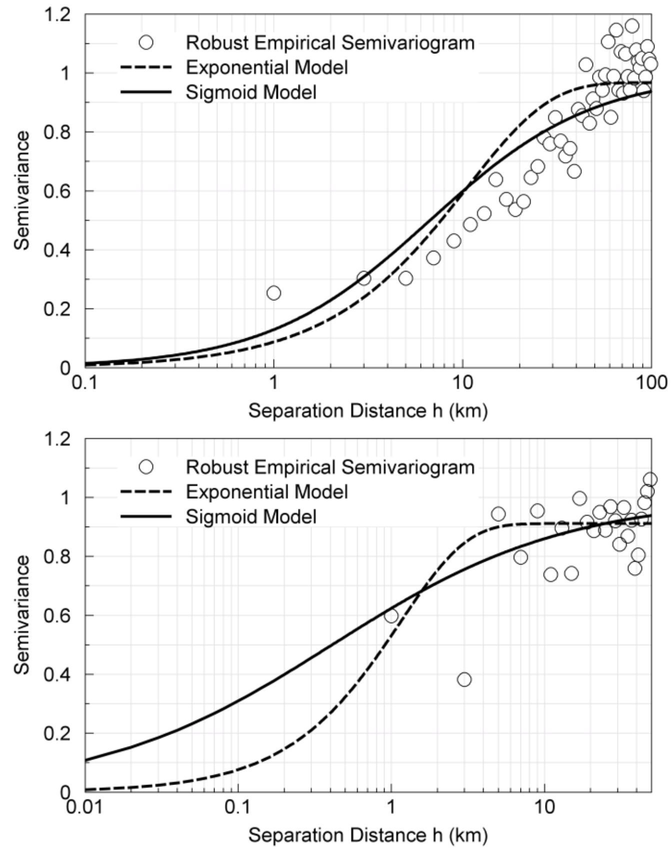

Fig. 12. The same fits shown in Fig. 11, but plotted using a logarithmic abscissa in order to accentuate the differences in the models for small separation distances.

The parameters of the models that have been obtained for these two events are presented in Table 3. As expected, the sill parameter is reasonably close to unity in both cases. However, the other parameters show considerable variability and it is clear that no generic trends can be inferred from just these two earthquake events.

Table 3. Model coefficients for the functional forms of Equations (11) and (12) and obtained from consideration of the Chi-Chi and Northridge earthquakes.

Model Parameter	Chi-Chi	Northridge
a	0.9675	0.9113
b	31.7127	3.4304
ϕ_1	-1.9046	0.8896
ϕ_2	0.9987	1.7603

DISCUSSION

This article has introduced a new empirical prediction equation for Arias intensity. The model has been shown to be robust and is also significantly different to the model of Travasarou *et al*. (2003) that is currently regarded as being the most robust model for Arias intensity that is generally applicable worldwide. The article has also investigated models for spatial correlation among Arias intensity values at multiple site locations. Thus far, just two earthquake events have been considered and the model parameters that have been obtained are markedly different for each of these events. It is clear from this work that further efforts must be directed at understanding the cause of these discrepancies so that we may know whether or not it will be feasible to develop a generic model for describing the spatial correlation among Arias intensity values. The new empirical prediction equation for Arias intensity may certainly be implemented for practical application in its current form. However, some minor modifications to this model are likely to arise from ongoing work that is investigating the nature of the variance structure that has been adopted so far. At this point in time, general recommendations for modeling spatial correlations cannot be made. It may well be that the most appropriate course of action that can be taken when the need arises to estimate the effects of ground motions on spatially-distributed systems is to implement spatial correlation models for spectral ordinates and to the obtain consistent models for Arias intensity through correlation relationships between these two intensity measures.

The ultimate aim of the work that is ongoing is to enable improved estimates of damage to spatially distributed systems to be made. In particular, initial efforts are being directed at incorporating spatial characteristics of ground motions and topography into models for landslide susceptibility and slope displacements. In any loss estimation analysis that includes the effects of ground failure, the probability of landslides occurring and their subsequent displacements are required. Modeling the spatial distribution of Arias intensity, obtained from equations such as those presented in this paper, is essential in order to accurately predict the probability of a landslide occurring as well as predicting the extent of the displacement related to this slide.

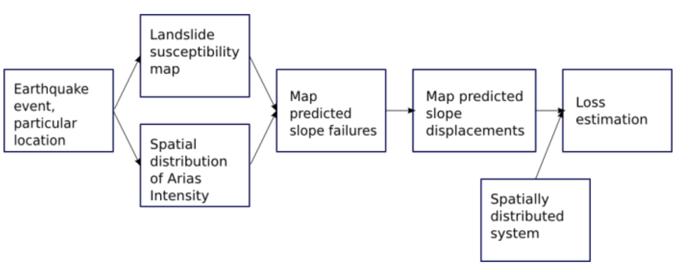

Fig. 13. Flowchart indicating the process via which spatially correlated ground motion fields may be incorporated within a loss estimation framework.

A schematic representation of the method via which these spatial features of ground motio may be implemented within a loss estimation framework that includes the effects of landslides is shown in the flowchart presented in Fig. 13.

For a given earthquake event, a landslide susceptibility map could be produced based on known local soil properties. The landslide susceptibility map, in conjunction with a map describing the spatial distribution of Arias intensity obtained using the new empirical relation presented in this paper and a model for the spatial correlation would be used to predict slope failures. Predicted slope displacements are then calculated and used to estimate landslide-induced losses to a spatially distributed system of interest, for example a lifeline network.

This article has taken a significant step towards enabling the framework shown in Fig. 13 to be implemented in practice. However, it is clear that more work is required before generic recommendations can be made regarding the nature of a spatial correlation model for Arias intensity.

REFERENCES

Abrahamson, N. A. and W. J. Silva [1997]. “Empirical response spectral attenuation relations for shallow crustal earthquakes”, Seismological Research Letters, No. 68(1), 94-127.

Abrahamson, N. A. and W. J. Silva [2008]. "Summary of the Abrahamson & Silva NGA ground-motion relations", Earthquake Spectra, No. 24(1), 67–97.

Arias, A. [1970]. “A measure of earthquake intensity”, in Seismic design for nuclear power plants, (R.J. Hansen, ed.) MIT Press, Cambridge MA, pp. 438-483.

Baker, J. W. [2007]. “Correlation of ground motion intensity parameters used for predicting structural and geotechnical response”, Proceedings of the 10th International Conference on Applications of Statistics and Probability in Civil Engineering (ICASP10), Tokyo, Japan.

Baker, J. W. and N. Jayaram [2008]. "Correlation of spectral acceleration values from NGA ground motion models", Earthquake Spectra, No. 24(1), 299-317.

Bommer, J. J., P. J. Stafford and J. E. Alarcón [2009]. “Empirical equations for the prediction of the significant, bracketed and uniform duration of earthquake ground motion.” Bulletin of the Seismological Society of America, No. 99(6), *in press*.

Boore, D. M. and G. M. Atkinson [2008]. "Ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods between 0.01 s and 10.0 s", Earthquake Spectra, No. 24(1), 99-138.

Chiou, B., R. Darragh, N. Gregor and W. Silva [2008]. “NGA project strong-motion database.” Earthquake Spectra, No. 24(1), 23-44.

Chiou, B. S.-J. and R. R. Youngs [2008]. "An NGA model for the average horizontal component of peak ground motion and response spectra", Earthquake Spectra, No. 24(1), 173-215.

Cressie, N. [1985]. "Fitting variogram models by weighted least squares", Mathematical Geology, No. 17(5), 563-586.

Egan, J. A. and D. Rosidi [1991]. “Assessment of earthquake-induced liquefaction using ground-motion energy characteristics”. Proceedings of the Pacific Conference on Earthquake Engineering, Auckland, New Zealand.

Goda, K. and H. P. Hong [2008]. “Spatial correlation of peak ground motions and response spectra.” Bulletin of the Seismological Society of America, No. 98(1), 354-365.

Harp, E. L. and R. C. Wilson [1995]. "Shaking intensity thresholds for rock falls and slides: Evidence from 1987 Whittier Narrows and Superstition Hills earthquake strong-motion records", Bulletin of the Seismological Society of America, No. 85(6), 1739-1757.

Jayaram, N. and J. W. Baker [2008]. "Statistical tests of the joint distribution of spectral acceleration values", Bulletin of the Seismological Society of America, No. 98(5), 2231-2243.

Jayaram, N. and J. W. Baker [2009]. “Correlation model for spatially distributed ground-motion intensities”, Earthquake Engineering & Structural Dynamics, *in press*.

Jibson, R. W. [2007]. "Regression models for estimating coseismic landslide displacement", Engineering Geology, No. 91(2-4), 209-218.

Kayen, R. E. and J. K. Mitchell [1997]. “Assessment of liquefaction potential during earthquakes by Arias intensity”. Journal of Geotechnical and Geoenvironmental Engineering, No. 123(12), 1162-1174.

Keefer, D. K. [2002]. "Investigating landslides caused by earthquakes - A historical review", Surveys in Geophysics, No. 23(6), 473-510.

Kramer, S. L. [1996]. *Geotechnical Earthquake Engineering*. Prentice Hall, Upper Saddle River, NJ

Lee, R. and A. S. Kiremidjian [2007]. "Uncertainty and correlation for loss assessment of spatially distributed systems", Earthquake Spectra, No. 23(4), 753-70.

Pinheiro, J., D. Bates, S. DebRoy, D. Sarkar and the R Core team [2008]. “nlme: linear and nonlinear mixed effects models”, R package version 3.1-87.

Stafford, P. J. and J. J. Bommer [2009]. "Empirical equations for the prediction of the equivalent number of cycles of earthquake ground motion", Soil Dynamics and Earthquake Engineering, No. 29, 1425-1436.

Stafford, P. J., J. B. Berrill and J. R. Pettinga [2009a]. "New predictive equations for Arias intensity from crustal earthquakes in New Zealand", Journal of Seismology, No. 13, 31-52.

Stafford, P. J., S. Sgobba and G. C. Marano [2009b]. “An energy-based envelope function for the stochastic simulation of earthquake accelerograms”, Soil Dynamics and Earthquake Engineering, No. 29(7), 1123-1133.

Travasarou, T., J. D. Bray and N. A. Abrahamson [2003]. "Empirical attenuation relationship for Arias Intensity", Earthquake Engineering & Structural Dynamics, No. 32(7), 1133-1155.

Walling, M., W. J. Silva and N. A. Abrahamson [2008]. "Nonlinear site amplification factors for constraining the NGA models", Earthquake Spectra, No. 24(1), 243-255.

Wang, M. and T. Takeda [2005]. Macrospatial correlation model of seismic ground motions. Earthquake Spectra, No. 21(4), 1137-1156.