Foulser-Piggott, R., & 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. DOI: 10.1785/0120140316

THE SEISMOLOGICAL SOCIETY OF AMERICA
400 Evelyn Ave., Suite 201
Albany, CA 94706-1375
(510) 525-5474; FAX (510) 525-7204
www.seismosoc.org

Bulletin of the Seismological Society of America, Vol. 105, No. 4, pp. 1903–1918, August 2015, doi: 10.1785/0120140316

Ground-Motion Prediction Models for Arias Intensity and Cumulative Absolute Velocity for Japanese Earthquakes Considering Single-Station Sigma and Within-Event Spatial Correlation

by Roxane Foulser-Piggott and Katsuichiro Goda

Abstract Arias intensity (I_A) and cumulative absolute velocity (CAV) are ground-motion measures that have been found to be well suited to application in a number of problems in earthquake engineering. Both measures reflect multiple characteristics of the ground motion (e.g., amplitude and duration), despite being scalar measures. In this study, new ground-motion prediction models for the average horizontal component of I_A and CAV are developed, using an extended database of strong-motion records from Japan, including the 2011 Tohoku event. The models are valid for magnitude greater than 5.0, rupture distance less than 300 km, and focal depth less than 150 km. The models are novel because they take account of ground-motion data from the 2011 Tohoku earthquake while incorporating other important features such as event type and regional anelastic attenuation. The residuals from the ground-motion modeling are analyzed in detail to gain further insights into the uncertainties related to the developed median prediction equations for I_A and CAV. The site-to-site standard deviations are computed and spatial correlation analysis is carried out for I_A and CAV, considering both within-event residuals and within-event single-site residuals for individual events as well as for the combined dataset.

Introduction

Arias intensity (I_A) is defined as the integral of the square of the acceleration time series (Arias, 1970), and is given by

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

in which g is the gravitational acceleration, $a(t)$ is the acceleration time series at time t, and t_{max} is the total duration of the time series. Cumulative absolute velocity (CAV) is defined as the integral of the absolute value of the acceleration time series (Electrical Power Research Institute, 1988; Campbell and Bozorgnia, 2010), and is given by

$$\mathrm{CAV} = \int_0^{t_{max}} |a(t)| dt, \tag{2}$$

in which $|\cdot|$ is the absolute value operator. Both of these measures are based on the idea that the amount of damage experienced by a structure is proportional to the energy dissipated by the structure per unit weight during the duration of the earthquake-induced ground motion. These duration-based ground-motion measures can capture the cumulative damage potential due to ground shaking and liquefaction (Kramer and Mitchell, 2006; Campbell and Bozorgnia, 2012b), particularly for structures that are susceptible to long-duration ground motion. I_A and CAV have also been found to correlate with seismic intensity measures such as modified Mercalli intensity (Campbell and Bozorgnia, 2012a) and spectral accelerations (Bradley, 2012). Modern predictive equations for I_A and CAV adopt complex functional forms for median and logarithmic standard deviation based on the state-of-the-art seismological findings (Campbell and Bozorgnia, 2010; Foulser-Piggott and Stafford, 2012). Moreover, they can be used to estimate I_A and CAV values simultaneously at multiple sites for a given earthquake scenario (Foulser-Piggott and Stafford, 2012; Du and Wang, 2013a). The latter extension is important to assess earthquake hazard and risk of spatially distributed infrastructure systems (Du and Wang, 2013b).

Japan is situated in a highly active seismic region influenced by complex interaction of tectonic plates. It hosts numerous major earthquakes in different seismotectonic regimes (e.g., shallow crustal, megathrust interface, and deep inslab events). An extensive analysis of observed ground motions in Japan by Oth *et al.* (2011) indicates that source and attenuation characteristics are spatially heterogeneous, affected by material properties and physical conditions of the crust (e.g., temperature and shear deformation). Regional variation of stress drop may be used as a proxy for a region-specific source parameter in ground-motion modeling (Oth, 2013).

The 11 March 2011 **M** 9 Great East Japan (Tohoku) earthquake was one of the largest recorded events in the world since 1900, and was extremely well recorded at national/regional strong-motion observation network stations, K-NET, KiK-net, and SK-net. The new data from the 2011 Tohoku event provide opportunities to investigate source-path-site characteristics for **M** 9 class subduction earthquakes and how these compare with those for smaller earthquakes. For instance, an investigation by Goda *et al.* (2013) indicated that the magnitude scaling effects of ground-motion parameters are different for peak-based measures (e.g., peak ground acceleration [PGA] and spectral acceleration) and for duration-based measures (e.g., I_A and CAV). The amplitudes for the former tend to saturate as the earthquake magnitude increases (between **M** 8 and 9), whereas amplitudes for the latter continue to increase due to long-duration features of the ground motions. Currently, there is a critical gap in conducting probabilistic seismic-hazard assessment in Japan using duration-based intensity measures. There are no modern ground-motion prediction equations (GMPEs) for I_A and CAV that are applicable to Japanese earthquakes. The situations differ from other standard intensity measures, such as PGA and spectral acceleration, for which models based on rich strong-motion datasets in Japan (and worldwide) are available (e.g., Kanno *et al.*, 2006; Zhao *et al.*, 2006), and a new generation of the models that incorporates key findings from the 2011 Tohoku earthquake has been developed (Morikawa and Fujiwara, 2013). Therefore, new development of predictive equations for I_A and CAV, which are based on an extended ground-motion database (including the 2011 Tohoku records), is warranted.

Statistical analysis of regression residuals offers further insights into the uncertainties associated with GMPEs and their dependency on explanatory variables (e.g., magnitude and distance). In the last decade, detailed analyses of between-event and within-event residuals have been carried out for various seismic regions (e.g., Chen and Tsai, 2002; Atkinson, 2006; Morikawa *et al.*, 2008; Anderson and Uchiyama, 2011; Lin *et al.*, 2011; Rodriguez-Marek *et al.*, 2011, 2013). The objectives of these studies are to distinguish sources of misfits in terms of source, path, and site characteristics by removing the ergodic assumption in developing empirical GMPEs. In the context of probabilistic seismic-hazard analysis, ergodic GMPEs imply that uncertainties associated with ground-motion predictions at a specific site can be obtained by averaging within-event residuals at different locations (with similar site conditions). However, by knowing systematic and repeatable features of ground-motion data (e.g., site-specific surface soil response and wave propagation path), uncertainties of ergodic GMPEs (e.g., standard deviation of within-event residuals) can be reduced significantly. The previous results indicate that accounting for the systematic site effects leads to reduced standard deviations of total residuals by about 10% to 15%, while further reduction of the standard deviation (up to about 35% to 50%) is possible by accounting for the path effects (Lin *et al.*, 2011). Another recent sophistication of empirical ground-motion modeling is related to spatial correlation of ground-motion intensity measures at multiple locations (e.g., Goda and Atkinson, 2010; Goda, 2011; Esposito and Iervolino, 2012; Sokolov and Wenzel, 2013). Spatial correlation models characterize spatial dependency of within-event residuals and previous investigations have been conducted by considering conventional ergodic GMPEs. However, the assessment of spatial correlation under the nonergodic assumption has not been explored extensively. This is important when the next generation of GMPEs based on the nonergodic assumption is applied in future seismic hazard and risk studies of spatially distributed structures.

This study develops new ground-motion models for I_A and CAV. For this purpose, an extended ground-motion database is compiled using high-quality digital recordings from the K-NET, KiK-net, and SK-net in Japan. The dataset is the updated version of Goda and Atkinson (2010), including recordings up to the end of 2012. From a ground-motion modeling standpoint, the database contains unique and valuable information in two respects. First, because the networks provide comprehensive spatial coverage of Japan, events are exceptionally well recorded. Second, the high rate of seismic activity in Japan and the surrounding area results in a dataset populated with a large number of well-recorded events in the past 16 years, with a range of magnitude–distance characteristics. The dataset is well suited for developing spatial correlation models of I_A and CAV, because it includes many well-recorded seismic events having more than 100 recordings. This facilitates the spatial correlation analysis for individual events, rather than combining heterogeneous events, as was done in previous studies (excluding Goda, 2011). We adopt an extended version of the functional form based on Foulser-Piggott and Stafford (2012), which takes into account a wider magnitude range (up to **M** 9) and different earthquake types (i.e., crustal versus interface versus inslab events). A nonlinear random-effect regression analysis is carried out to determine final sets of model coefficients for both median and logarithmic standard deviations of regression residuals. Spatial correlation models for I_A and CAV are evaluated empirically by constructing semivariograms of residuals and the spatial correlation of different residual components is evaluated. Specifically, within-event residuals are further decomposed into site-to-site residuals and within-event single-site residuals, and their correlation characteristics are compared with the conventional within-event spatial correlation. This work is novel and important for three main reasons. First, the developed models take into account ground-motion data from the 2011 Tohoku earthquake (having a very long duration) by modifying the magnitude scaling term. The models also incorporate other important features (e.g., depth and event type). Second, the site-to-site standard deviations are computed and the relative contributions from the between-event, site-to-site, and single-site components of the standard deviation are investigated. Third, spatial correlation modeling is conducted based on the extensive ground-motion dataset that has dense station-to-station spacing, and its features are evaluated in terms of event characteristics of individual earthquakes.

Ground-Motion Data for Japanese Earthquakes

A new ground-motion database for Japanese earthquakes is compiled for the purpose of ground-motion studies. It combines recordings from three national/regional ground-motion networks in Japan, K-NET, KiK-net, and SK-net, up to the end of 2012 (see Data and Resources). Records from different networks are first integrated by matching event information (occurrence time, location, earthquake size, etc.). Metadata, such as moment magnitude, fault mechanism (normal/reverse/strike slip), earthquake type (crustal/inslab/interface), and finite-fault plane information, are assigned to seismic events with the Japan Metrological Agency (JMA) magnitude $M_{\mathrm{JMA}} \geq 5.0$ individually. Whenever individual events are associated with those listed in the Harvard Centroid Moment Tensor catalog and F-net mechanism catalog, the moment magnitude **M** is used as a representative measure (the moment magnitudes based on the Harvard catalog are preferred). Otherwise, M_{JMA} is adopted. Using available finite-fault plane models, rupture distance is calculated and used as representative distance measure; otherwise, the hypocentral distance is used. Site information for the K-NET and KiK-net is obtained from the National Research Institute for Earth Science and Disaster Prevention websites (see Data and Resources). For the SK-net sites, an approach adopted by Goda and Atkinson (2010) is implemented, which combines multiple estimates of average shear-wave velocity in the uppermost 30 m (V_{S30}) from borehole logging, microtremor measurements, geomorphological information, and slope information. Priority is given in the following order: (1) adopt V_{S30} estimates from borehole logging when the logging depth exceeds 20 m depth; (2) use the geometric mean of V_{S30} estimates from borehole logging and those from microtremor measurements when the logging depth is between 10 and 20 m; (3) adopt V_{S30} estimates from microtremor measurements when the logging depth is shallower than 10 m or no borehole logging information is available; and (4) use the geometric mean of V_{S30} estimates based on the geomorphological information and the slope information when no physical measurements of soil profiles are available. By applying broad record selection criteria, the database contains 555,750 records from 6261 earthquakes. Subsequently, individual components in the record set are processed uniformly (i.e., tapering, zero padding, and band-pass filtering). The filter is a Butterworth fourth-order filter and the corner frequencies of the band-pass filter are 0.05–20 Hz. Various elastic ground-motion measures, including I_{A} and CAV, are computed using the processed record components.

For developing GMPEs for I_{A} and CAV, the following record selection criteria are considered: (1) earthquake magnitude is greater than 5.0; (2) focal depth is less than 150 km; (3) average shear-wave velocity in the uppermost 30 m is between 150 and 1500 m/s; (4) magnitude–distance cutoff filter by Kanno *et al.* (2006), which was used by Morikawa and Fujiwara (2013), is applied (threshold for predicted median PGA value is set to 5 cm/s^2); and (5) the minimum number of eligible records that satisfy the five criteria above is more than 10. The criteria adopted herein are similar to those applied by Morikawa and Fujiwara (2013), but the PGA threshold is modified to consider the lower magnitude cutoff of 5.0, rather than 5.5 as considered by Morikawa and Fujiwara (2013). The application of the above five criteria has resulted in 68,567 records from 661 events, consisting of 112 crustal events, 331 interface events, and 218 inslab events, each of which has two orthogonal horizontal components that are combined using the geometric mean. Out of the 661 events, 586 are associated with the Harvard catalog moment magnitudes, 46 with F-net based moment magnitudes, and 29 with F-net based JMA magnitudes. In the selected dataset, there are 203 events defined as well recorded, that is, those that have more than 100 recordings; the total number of records for well-recorded events is 48,756. This subdataset is ideal for developing spatial correlation models for individual earthquakes (Goda, 2011) and for investigating the dependency of spatial correlation models on earthquake characteristics (e.g., magnitude and earthquake type).

Figure 1 shows the locations of the earthquakes, the magnitude–distance plot, and the histogram of V_{S30} of the selected records. The location map of earthquakes shown in Figure 1a indicates that the majority of earthquakes occur in the northeastern part of Japan; a boundary to separate such a region is shown in the figure. Figure 1b shows that the dataset lacks records with short distances (less than 5 km). Therefore, the GMPEs for I_{A} and CAV that are developed in this study should not be extrapolated to very short distances. The majority of the records are observed at sites on relatively soft soils (Fig. 1c). Preliminary investigations of the ground-motion data with respect to seismological parameters (results are omitted for brevity) indicate that I_{A} and CAV tend to increase with magnitude (as expected). For interface events, there is a clear increasing trend from the small-to-large magnitude range (although the variability of the data is significant due to distance and site variations) and no sign of saturated magnitude scaling in the large magnitude range is detected. The increasing trends at large magnitudes are due to long-duration motions for megathrust subduction earthquakes (e.g., 2003 Tokachi-oki and 2011 Tohoku earthquakes). Another important observation is that scatter of the data points for CAV is much less than that for I_{A}; this is related to how these two measures are defined/computed from accelerograms (equations 1 and 2).

Prediction Models

To conduct earthquake hazard and risk analyses in Japan in terms of I_{A} and CAV, stable empirical ground-motion models are required. Foulser-Piggott and Stafford (2012) developed a robust model (Foulser-Piggott and Stafford [FPS] model) for the prediction of I_{A} applicable to worldwide crustal earthquakes. This model was developed using a subset of the Pacific Earthquake Engineering Research–Next Generation Attenuation (PEER NGA) database, which includes recordings from shallow crustal earthquakes occurring worldwide.

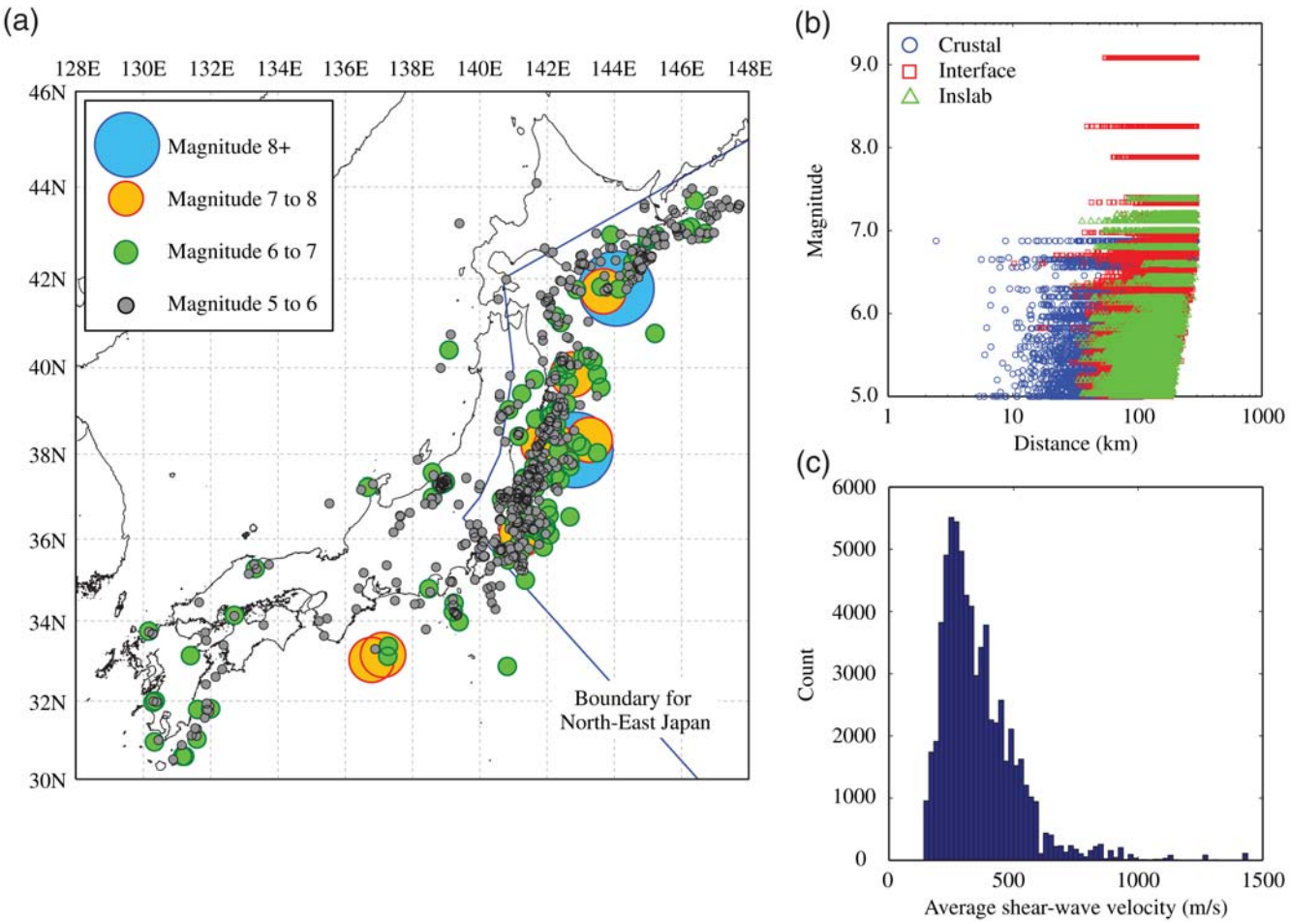

Figure 1. Ground-motion data characteristics: (a) spatial distribution of events, (b) magnitude–distance plot, and (c) average shear-wave velocity in the uppermost 30 m. The color version of this figure is available only in the electronic edition.

When applied to the Japanese dataset, the functional form of the FPS model does not converge. When the coefficients of the nonlinear site effects term are fixed, the model converges but has a relatively large logarithmic standard deviation (1.50). The unsatisfactory performance of the FPS model for the Japanese dataset can be attributed to the following facts. First, the model is based only on a dataset of crustal earthquakes and second, it does not include several terms that are relevant to Japanese earthquakes, such as regional anelastic attenuation terms, style of faulting, and type of earthquake. A counterpart of the I_A model for the prediction of CAV was developed by Campbell and Bozorgnia (2010) using the PEER NGA database. Because of the reasons above and because predictive models for I_A and CAV that are applicable up to **M** 9 events in Japan do not exist, new empirical models for the prediction of I_A and CAV in Japan are developed in this study. The new models developed in this study are based on functional forms that are basic variations of the theoretically motivated form presented in Foulser-Piggott and Stafford (2012); therefore, all terms have a physical meaning. In the development of new models for I_A and CAV, we use theoretically constrained functional forms but also adopt expressions that are commonly used for (1) the derivation of both worldwide predictive models for ground-motion measures (e.g., Campbell and Bozorgnia, 2010; Foulser-Piggott and Stafford, 2012), (2) earthquake event-type-specific GMPEs (e.g., Atkinson and Boore, 2003; Abrahamson *et al.*, 2015), and (3) regional GMPEs (e.g., Kanno *et al.*, 2006; Zhao *et al.*, 2006; Morikawa and Fujiwara, 2013). Multiple regression analyses are conducted with different functional forms and the results are compared using test statistics including log likelihood, the Akaike information criterion (AIC; Akaike, 1987), the Bayesian information criterion (BIC; Schwarz, 1978), model standard deviations, and significance of model coefficients (p-value). These statistical metrics are used to support the arguments for the inclusion or exclusion of variables. The efficacy of different functional forms is investigated using plots of residuals and identifying any trends.

Two functional forms are presented here: a linear site-response model and a nonlinear site-response model (see equations 4–6). Both functional forms give an expression for I_A and CAV in terms of the predictor variables: moment magnitude **M**, rupture distance R_{rup}, focal depth H, two dummy variables to indicate the type of event F_{inslab} and

$F_{\text{interface}}$, two dummy variables to indicate the type of faulting mechanism for crustal earthquakes F_{rv} and F_{nm}, and the shear-wave velocity V_{S30}. The equation also includes two dummy variables that take into account differences in attenuation for the forearc and backarc regions in northern Japan: F_{rfa} and F_{rba}. The four major parts of the functional form that are the focus of the model development are the magnitude scaling, depth effects, anelastic attenuation, and site-response terms.

Analytical and theoretical considerations suggest that the logarithm of I_{A} should either scale approximately linearly or mildly nonlinearly with magnitude (Foulser-Piggott and Stafford, 2012). This nonlinearity can be incorporated into a model in three ways: first through the use of direct nonlinear magnitude scaling terms, second using magnitude-dependent geometric spreading terms, or third using nonlinear site-response terms. Two options for the magnitude scaling of the model are investigated in this study: linear and quadratic. Regression analyses are conducted using linear magnitude scaling of the form $c_1(\mathbf{M}-5)$ (see equation 6) and quadratic magnitude scaling of the form $c_1(9.5-\mathbf{M})^2$. Using the linear over the quadratic magnitude term has minor effects on predicted ground-motion levels for small-to-moderate earthquakes, but has a significant impact for large earthquakes. Using quadratic scaling, the 2011 Tohoku motions (which are important for seismic hazard and risk assessments in Japan) are significantly underpredicted; the between-event residuals are +2.1 and +3.6 for the linear and nonlinear site-response models, respectively. Using linear magnitude scaling over quadratic scaling, the between-event residuals for the 2011 Tohoku records are reduced to 1.4 and 1.0 for the linear and nonlinear site-response models, respectively. A comparison of test statistics from both models reveals that the linear magnitude scaling produces better results in terms of AIC and BIC. These statistical metrics support the theoretical arguments for the adoption of linear scaling over quadratic scaling.

The depth scaling of the regression model is investigated to improve the fit of the model at small distances and large magnitudes. The model developed in this study has two depth terms. The first is for the near-source saturation effect (i.e., c_4 in equation 6). The term $c_5 \max(H-30, 0)$ captures the focal depth dependence for deep events, noting that the depth of 30 km is often used as the typical depth of the Moho boundary (e.g., Kanno *et al.*, 2006). This improves model fit to deep events, which in turn improves the performance of the model which is clear when examining a plot of the between-event residuals against magnitude. The inclusion of both of these depth terms reduces the model standard deviation by 2.5%.

Ghofrani and Atkinson (2011) and Oth (2013) discuss features of region-specific anelastic attenuation for Japanese earthquakes that are investigated for potential incorporation into the model. Ghofrani and Atkinson (2011) develop models for both Japanese crustal and inslab events to describe the ground motions in forearc and backarc regions and also show that distinguishing between these two regions reduces aleatory variability in ground motions. The results show that attenuation is greater in the backarc than in the forearc direction. Based on the work of Oth (2013), which discusses differences in attenuation in different regions, the dataset is divided into records from the regions of northeast Japan forearc or backarc and southwest Japan forearc or backarc. A number of options were considered for incorporating regional anelastic attenuation in the model, and regression analyses were conducted using anelastic attenuation terms of the form $(c_6 F_{\text{rfa}} + c_7 F_{\text{rba}})R_{\text{rup}}$ (see equation 6). Two different region combinations were also considered: first, F_{rfa} and F_{rba} take the values of 1 and 0 for forearc recordings in the northeast region of Japan, 0 and 1 for backarc recordings in the northeast region of Japan, and 0 and 0 otherwise. Second, F_{rfa} and F_{rba} take the values of 1 and 0 for forearc recordings in the northeast or southwest region of Japan, 0 and 1 for backarc recordings in the northeast or southwest region of Japan, and 0 and 0 otherwise. The first regionalization was selected because the values obtained from the regression analysis for the coefficients are physically reasonable (small, negative values), and both c_6 and c_7 are statistically significant. Therefore, the inclusion of these terms is warranted because they are empirically and analytically justified.

Studies on site amplification in Japan based on the KiK-net stations draw different conclusions regarding the linearity or nonlinearity of site response. Ghofrani *et al.* (2013) performed an extensive analysis of site amplification using surface and borehole ground-motion data from the KiK-net stations during the Tohoku and other Japanese events. They reported that site amplification effects in the high-frequency range are very large in Japan, with the exception of some localized nonlinearity and noted that the degree of nonlinearity increases with the intensity of shaking. Kaklamanos *et al.* (2013) assessed the accuracy and precision of site-response models and their results show that linear site-response analyses become inaccurate (due to nonlinearity in site response) for spectral acceleration prediction at vibration periods less than 0.5 s. In this study, two functional forms are trialed and presented, one with linear and the other with nonlinear site-response terms. The linear site-response functional form has a single coefficient v_1 to be estimated (equation 4) and the nonlinear functional form has four coefficients v_1, v_2, v_3, and v_4 (equation 5). For the nonlinear site-response case, a regression analysis is conducted using the nonlinear site-response functional form of the FPS model. This is of the form shown in equation (3), in which the amplitude of the surface ground-motion intensity $\hat{I}_{\text{soil}}$ is a function of the predicted ground-motion intensity on the reference site condition $\hat{I}_{\text{rock}}$, which is chosen to correspond to rock sites for which nonlinear effects should be minimal:

$$\ln(\text{Amp}) = \ln\left(\frac{\hat{I}_{\text{soil}}}{\hat{I}_{\text{rock}}}\right) = a + b(\ln \hat{I}_{\text{rock}} + c). \qquad (3)$$

The nonlinear site-response model shown in equation (3) was first implemented by Abrahamson and Silva (1997) and also

considered by Chiou and Youngs (2008). Foulser-Piggott and Stafford (2012) indicated that this form performs well when used to model the nonlinear site-response of I_A values. The regression analyses with both linear and nonlinear site-response terms resulted in statistical significance of all parameters and no visible trends in the plot of residuals against V_{S30}. The standard deviations as well as other statistical metrics (e.g., AIC and BIC) indicate that the nonlinear site-response model formulation performs in a similar way to the linear model. However, from a physical point of view, there are a number of important points. First, it is not expected that there would be a dramatic reduction in the standard deviation or other model performance measures because little data in the dataset are in the nonlinear site-response range. Second, there are mixed views on the occurrence and extent of nonlinear site-response in Japan. Third, the majority of modern GMPEs developed implement a nonlinear site-response model. Both models are presented in this study because for large earthquakes the linear site-response model will underpredict ground motions. The nonlinear site-response model overpredicts in the low- and mid-magnitude range; however, if low- or medium-magnitude earthquakes occur in an area where there is specific evidence of nonlinear site amplification, it would be more suitable to use the nonlinear site-response model. There were some issues with convergence for the coefficient v_4 for the nonlinear site-response model for CAV (the fitting was successful for I_A). The physical meaning of v_4, as discussed by Chiou and Youngs (2008), is the corner value, that is, the value of the ground motion where the transition from linear site amplification to nonlinear site amplification occurs. Using this physical meaning of v_4, a suitable value of v_4 for CAV was inferred from the value of v_4 obtained for I_A (within the normal model-fitting procedure) by checking the correspondence of the v_4 value for I_A data in terms of CAV data. To confirm that the obtained v_4 value for CAV is reasonable, we also carried out a large number of regression analyses by trying different values of v_4 to ensure that the fitting process leads to convergence and the inferred value of v_4 falls within such a suitable range.

In summary, the main differences between the functional form of the FPS model and the final functional forms shown in equations (4), (5), and (6) are as follows:

1. a linear rather than quadratic magnitude scaling;
2. an additional depth scaling term;
3. the addition of dummy variables to indicate the region, forearc or backarc: F_{rfa} and F_{rba}, these take the values of 1 and 0 for forearc recordings in the northeast region of Japan, 0 and 1 for backarc recordings in the northeast region of Japan, and 0 and 0 otherwise;
4. the addition of two dummy variables to indicate the event type under consideration: F_{inslab} and $F_{interface}$, these take the value of 1 and 0 for an inslab event, 0 and 1 for an interface event, and 0 and 0 for an crustal event; the inclusion of these terms is based on the finding by Zhao *et al.* (2006) that although PGA for different earthquake event types (crustal, interface, inslab) scales similarly with distance, the PGAs from inslab events are larger than those from crustal or interface events by a factor of 1.66;
5. additional terms to describe the effects due to style of faulting for crustal events: F_{rv} and F_{nm} for reverse and normal mechanisms, respectively.

At each stage of the development of the final functional forms, the trial models are evaluated using standard statistical metrics that are automatically provided through the use of the *nlme* package (Pinheiro *et al.*, 2008) of the software program *R*. Equations (4) and (5) are the functional forms for the linear (LIN) and nonlinear (NL) site-response models, respectively ($I = I_A$ or CAV):

$$\ln \hat{I}_{LIN} = \ln \hat{I}_{ref} + e_1 + v_1 \ln\left(\frac{V_{S30}}{1100}\right) \tag{4}$$

and

$$\begin{aligned}\ln \hat{I}_{NL} = {} & \ln \hat{I}_{ref} + e_1 + v_1 \ln\left(\frac{V_{S30}}{1100}\right) \\ & + v_2(\exp\{v_3[\min(V_{S30}, 1100) - 280]\} \\ & - \exp[v_3(1100 - 280)]) \\ & \times \ln\left(\frac{\exp[\ln \hat{I}_{ref} + e_1] + v_4}{v_4}\right),\end{aligned} \tag{5}$$

in which the reference ground-motion intensity $\hat{I}_{ref}$ is given by

$$\begin{aligned}\ln \hat{I}_{ref} = {} & c_0 + c_1(\mathbf{M} - 5) + (c_2 + c_3\mathbf{M}) \ln \sqrt{R_{rup}^2 + c_4^2} \\ & + c_5 \max(H - 30, 0) + (c_6 F_{rfa} + c_7 F_{rba}) R_{rup} \\ & + c_8 F_{inslab} + c_9 F_{interface} + c_{10} F_{rv} + c_{11} F_{nm}.\end{aligned} \tag{6}$$

In equations (4), (5), and (6), e_1 is the random-effect term for individual events, the coefficients v_1–v_4 are for the site-response models, and the coefficients c_0–c_{11} are for the reference ground motion. Regression coefficients for I_A and CAV (both linear and nonlinear site-response models) are summarized in Table 1; all these coefficients are found to be statistically significant at the 95% confidence level (i.e., for all coefficients, p-value ≈ 0). The model may be implemented for Japanese earthquakes with **M** greater than 5.0, R_{rup} less than 300 km, and H less than 150 km. These ranges of applicability are based upon the range of the data used in the model derivation (see Fig. 1) but the model will be less well constrained at the extreme values in this range. In particular, the maximum value of **M** that can be used for crustal and inslab events should be limited to 7.0 and 7.5, respectively.

The developed median ground-motion models for I_A and CAV are presented in Figure 2 by considering several interface earthquake scenarios. In the figure, two values of **M** (7.0 and 9.0) and V_{S30} (300 and 1100 m/s) are considered. For **M** 9.0 scenarios, the shortest distance is set to 40 km (instead of 20 km for **M** 7.0 scenarios), because megathrust

Table 1
Regression Coefficients for Arias Intensity (I_A) and Cumulative Absolute Velocity (CAV)

	Linear Site-Response Model (Equations 4 and 6)		Nonlinear Site-Response Model (Equations 5 and 6)	
Coefficients	I_A (m/s)	CAV (m/s)	I_A (m/s)	CAV (m/s)
c_0	3.056224	2.643261	2.16574	2.47814
c_1	2.639315	1.60688	3.508756	1.799346
c_2	2.352244	0.754765	1.294525	0.539751
c_3	0.080591	0.072283	0.256147	0.109694
c_4	12.682338	12.626135	7.244428	11.472109
c_5	0.009653	0.003811	0.009592	0.003831
c_6	0.001436	0.00059	0.001819	0.000685
c_7	0.006374	0.002767	0.006795	0.002882
c_8	1.869827	0.877694	1.886186	0.882441
c_9	1.639023	0.822831	1.650818	0.826529
c_{10}	0.573052	0.286527	0.570372	0.285578
c_{11}	1.856785	0.918286	1.854696	0.916566
v_1	1.030608	0.65776	1.060057	0.686706
v_2	-	-	0.629392	0.229981
v_3	-	-	0.006856	0.015479
v_4	-	-	0.346117	15.85
τ	0.9015	0.4114	0.9082	0.4149
ϕ	1.035	0.4900	1.0328	0.4893

subduction earthquakes in Japan are typically offshore events and it is rare to have distances less than 40 km. Figure 2 shows that for $V_{S30} = 1100$ m/s (rock sites), the linear model predictions are less than the nonlinear model predictions, whereas for $V_{S30} = 300$ m/s (soil sites) the effects of nonlinear site-response become significant and the nonlinear model predictions become less than the linear model predictions. The latter is particularly significant for I_A and for **M** 9.0 (Fig. 2a). The differences of the linear and nonlinear models for CAV are small.

Figure 3 compares developed median ground-motion models for I_A with existing models by Foulser-Piggott and Stafford (2012) and by Campbell and Bozorgnia (2012b) for crustal events. Two cases are selected: **M** 6.0 and **M** 7.0 (H is set to 10 km and V_{S30} is set to 300 m/s). Moreover, the median curves are compared with the I_A data (crustal earthquake records only) for similar conditions. Specifically, the considered magnitude and shear-wave velocity ranges are 5.75–6.25 for **M** 6.0; 6.75–7.25 for **M** 7.0; and 200–500 m/s for V_{S30}. The developed median equations pass through the

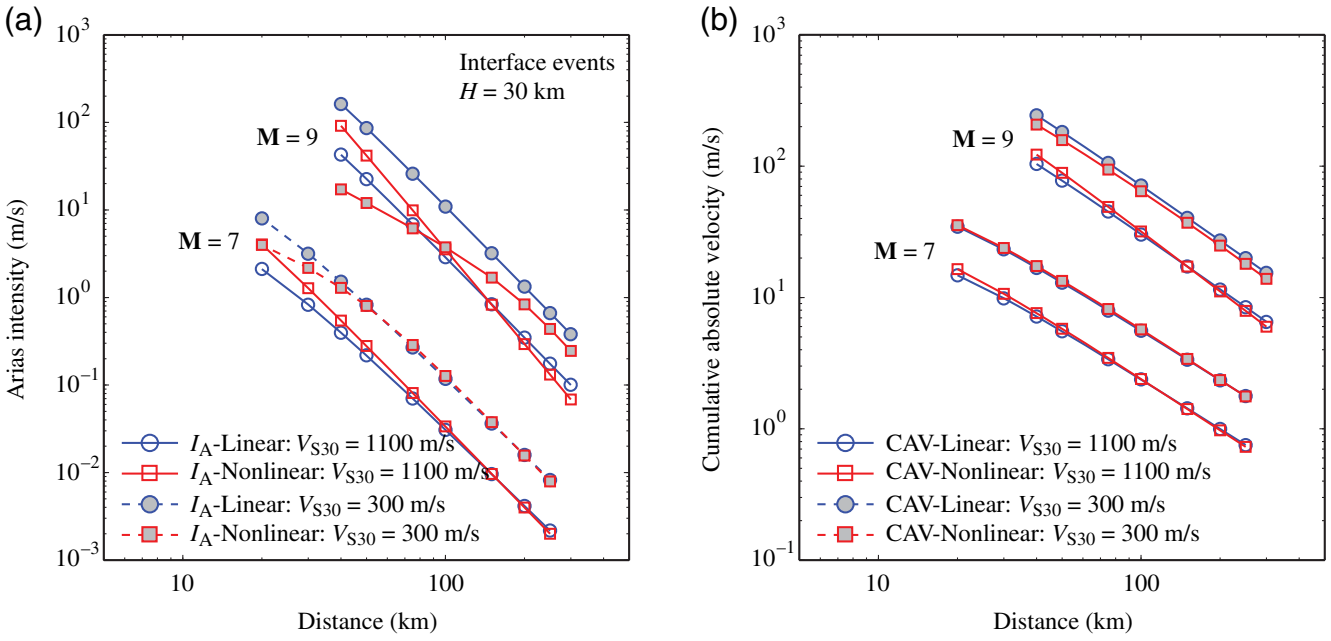

Figure 2. Comparison of median prediction models for linear and nonlinear site-response terms (**M** = 7 and 9, and V_{S30} = 300 and 1100 m/s): (a) Arias intensity (I_A) and (b) cumulative absolute velocity (CAV). The color version of this figure is available only in the electronic edition.

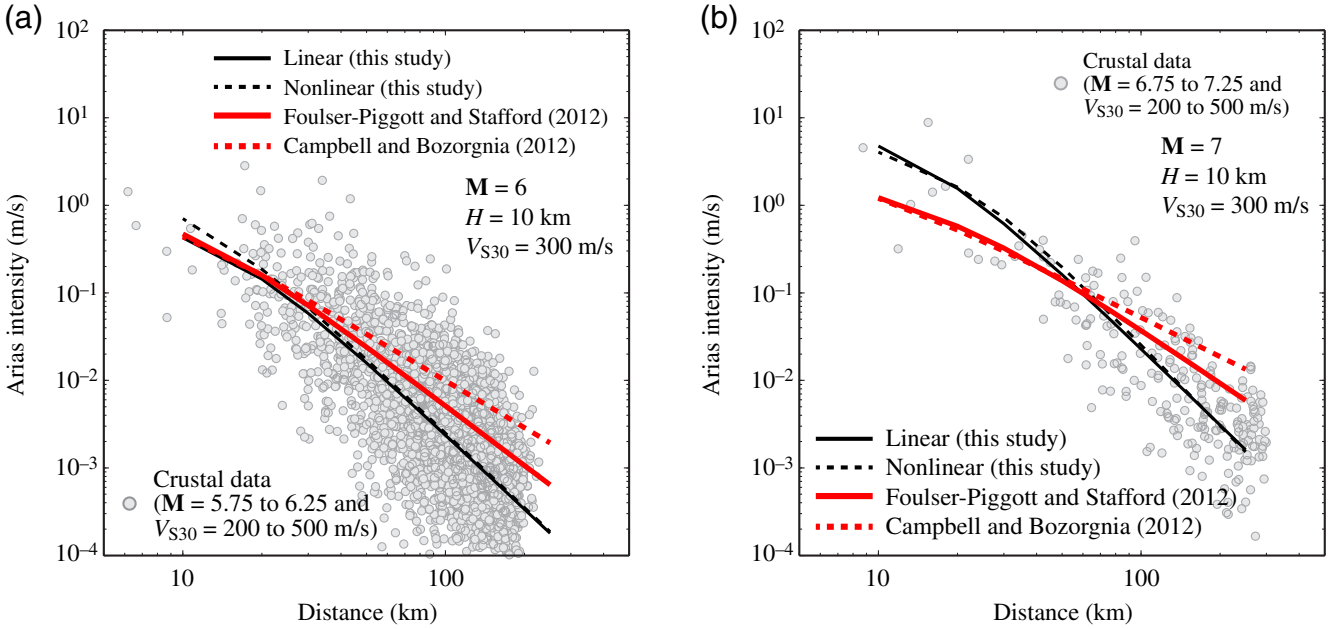

Figure 3. Comparison of developed median ground-motion models for I_A with existing models by Foulser-Piggott and Stafford (2012) and by Campbell and Bozorgnia (2012b) for crustal data: (a) $\mathbf{M} = 6$, $H = 10$ km, and $V_{S30} = 300$ m/s and (b) $\mathbf{M} = 7$, $H = 10$ km, and $V_{S30} = 300$ m/s. The color version of this figure is available only in the electronic edition.

center of the I_A data for both cases (as expected), noting that the scatter of the data is large. The comparison indicates that the attenuation rates for the worldwide crustal equations are more gradual than those for the Japanese equations. For **M** 6.0, differences in the predicted I_A values from different models become significant in the large distance range, whereas for **M** 7.0, the Japanese models and worldwide crustal models intersect at the intermediate distance range (around 60–70 km) and large differences in the predicted I_A values become noticeable in the short as well as large distance ranges. These differences are likely to be caused by the differences of the underlying data. Generally, the I_A dataset for the Japanese crustal earthquakes has more data points with smaller I_A values, in comparison with worldwide crustal earthquakes (obtained from the PEER NGA database).

Figure 4 shows the residual plots that have been obtained from the nonlinear random-effect regression analysis for I_A and CAV. 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 **M**, R_{rup}, and V_{S30}. As part of the model development, the applicability of the heteroskedastic variance structure to the above regression results was investigated. As discussed by Foulser-Piggott and Stafford (2012) and demonstrated by Chiou and Youngs (2008), the within-event residuals should display heteroskedasticity with respect to the reference rock motion. In addition, Foulser-Piggott and Stafford (2012) implemented a magnitude-dependent variance structure which predicted increasing variance with decreasing magnitude. Versions of such a heteroskedastic variance structure were trialed for the models presented in this study. However, modeling results were not physically defensible. In particular, the magnitude-dependent variance structure showed an increase with increasing magnitude. This was investigated and found likely to be an artifact of the data. In addition, site-dependent variance structure models resulted in extremely small coefficients that were not significant for I_A and no model convergence for CAV. Statistically, further tests for heteroskedasticity, such as Breusch–Pagan tests, confirmed the previous findings and the visual inspections did not identify any trends of variance with respect to predictor variables.

Statistical Analysis of Regression Residuals

Residuals from the ground-motion modeling are analyzed in detail to gain further insights into the uncertainties related to the developed median prediction equations for I_A and CAV. This section describes both single-station sigma and spatial correlation analyses.

Generally, the observed ground motions can be decomposed as follows:

$$\ln I_{es} = \ln \hat{I}_{es} + \Delta_{es} = \ln \hat{I}_{es} + \delta B_e + \delta W_{es} \qquad (7)$$

(Al Atik *et al.*, 2010), in which $\ln I_{es}$ and $\ln \hat{I}_{es}$ are the logarithm of observed and predicted ground-motion intensity measures, respectively, for earthquake e and site s, and Δ_{es} are the total residuals. δB_e and δW_{es} are the between-event and within-event residuals, respectively. The between-event residuals are specific to individual events and represent the variability between earthquakes of the same magnitude; the within-event residuals represent the variability from the

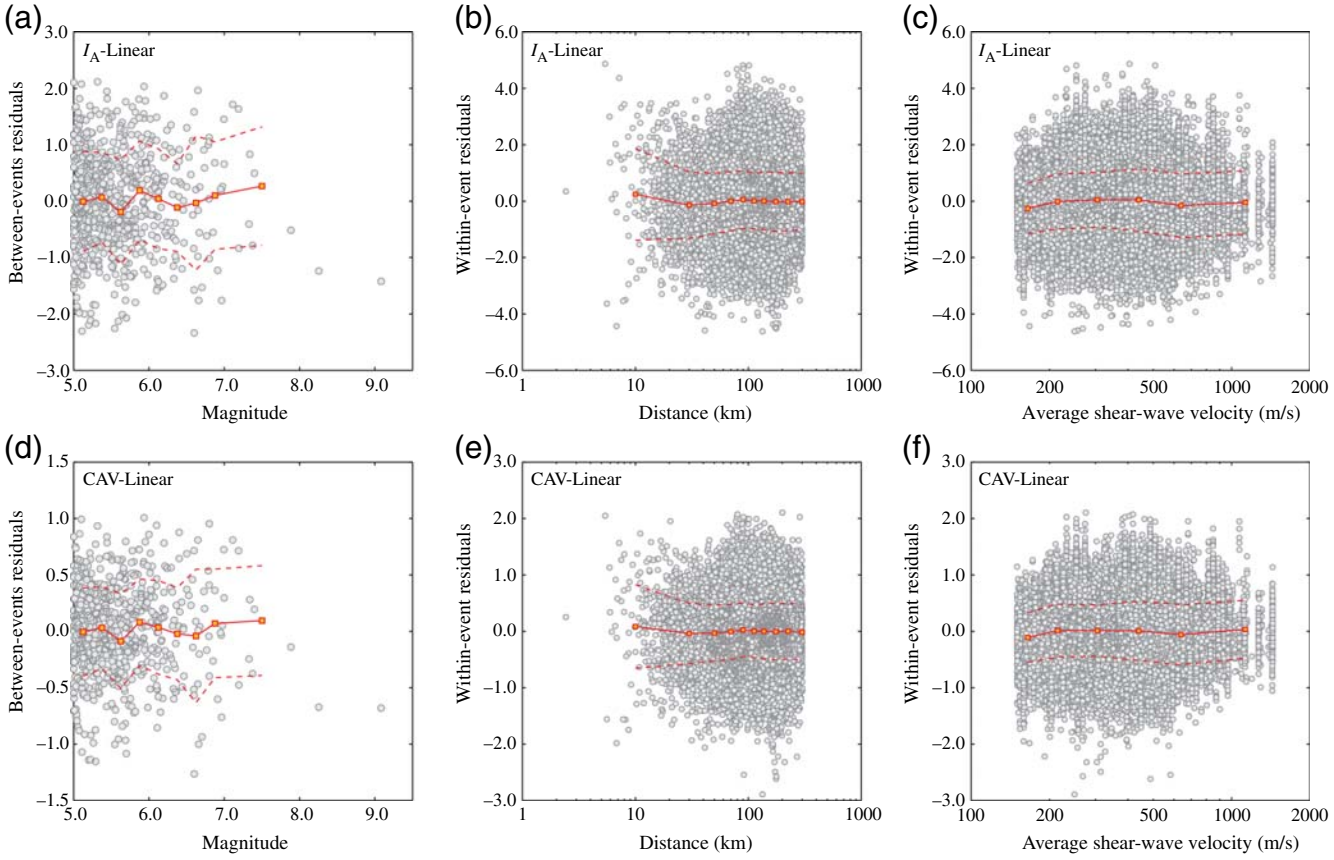

Figure 4. Residual characteristics: (a–c) interevent residuals versus magnitude, within-event residuals versus distance, and within-event residuals versus average shear-wave velocity for I_A and (d–f) for CAV. The color version of this figure is available only in the electronic edition.

median predicted value for a particular recording station in a given event. The regression residuals are approximated by zero-mean normal random variable and the degree of uncertainty is measured by the standard deviation. The standard deviations of Δ, δB, and δW are denoted by σ, τ, and ϕ, respectively. Assuming that τ and ϕ are uncorrelated, $\sigma^2 = \tau^2 + \phi^2$.

Single-Site Residuals

The within-event residuals can be further decomposed into systematic and nonsystematic components (e.g., site and path effects). Various approaches have been adopted in previous studies (Chen and Tsai, 2002; Atkinson, 2006; Morikawa *et al.*, 2008; Anderson and Uchiyama, 2011; Lin *et al.*, 2011; Rodriguez-Marek *et al.*, 2011, 2013). One useful way to account for repeatable site effects (with respect to the regression models) is to separate the systematic site effects from the within-event residuals (Anderson and Uchiyama, 2011; Rodriguez-Marek *et al.*, 2011, 2013):

$$\Delta_{es} = \delta B_e + \delta W_{es} = \delta B_e + \delta S2S_s + \delta WS_{es}, \quad (8)$$

in which $\delta S2S_s$ are the site-to-site residuals and can be obtained by averaging within-event residuals for different events at a particular recording site, and δWS_{es} are the within-event single-site residuals. Typically, the minimum number of 5–10 recordings per station is used in computing the site-to-site residuals. The site-to-site residuals account for site-specific effects which are not yet removed through the near-surface site-response variable V_{S30}. On the other hand, the within-event single-site residuals include all other unexplained misfits and are predominantly influenced by the path effects. The standard deviations of $\delta S2S_s$ and δWS_{es} are denoted by ϕ_{S2S} and ϕ_{SS}, respectively. In terms of variance of the residuals, the total variance consists of three subcomponents: $\sigma^2 = \tau^2 + \phi_{S2S^2} + \phi_{SS^2}$. It is important that the procedure adopted in this study assumes that the decomposition of residuals into different components has no influence on the regression coefficients of the median models. In reality, consideration of different residual compositions affects both regression coefficients (fixed effects) and residuals (random effects). A more rigorous approach has been proposed by Stafford (2014).

Table 2 summarizes values of τ, ϕ, ϕ_{S2S}, and ϕ_{SS} for I_A and CAV (based on equations 4, 5, and 6). The site-to-site standard deviations are computed for the stations having no less than five recordings (2199 stations). For τ, ϕ, and ϕ_{SS}, their values for crustal, interface, and inslab events

Table 2
Residual Components from the Single-Station Sigma Analysis for I_A and Cumulative Absolute Velocity

	Linear Site-Response Model		Nonlinear Site-Response Model	
Standard Deviation	I_A (m/s)	CAV (m/s)	I_A (m/s)	CAV (m/s)
Between events: τ	0.888	0.405	0.894	0.408
Between events: $\tau_{crustal}$	0.971	0.436	0.968	0.435
Between events: $\tau_{interface}$	0.858	0.396	0.866	0.400
Between events τ_{inslab}	0.892	0.403	0.902	0.408
Within event: ϕ	1.030	0.488	1.028	0.487
Within event: $\phi_{crustal}$	1.074	0.522	1.068	0.521
Within event: $\phi_{interface}$	0.984	0.468	0.981	0.467
Within event: ϕ_{inslab}	1.054	0.487	1.055	0.487
Site-to-site: ϕ_{S2S}	0.740	0.369	0.739	0.368
Within-event single site: ϕ_{SS}	0.717	0.319	0.714	0.319
Within-event single site: $\phi_{SScrustal}$	0.836	0.377	0.829	0.376
Within-event single site: $\phi_{SSinterface}$	0.655	0.299	0.651	0.298
Within-event single site: $\phi_{SSinslab}$	0.699	0.298	0.700	0.298

are also included, whereas for ϕ_{S2S}, such values are not applicable (i.e., for given prediction models, ϕ_{S2S} is computed using all relevant data from three earthquake types). The results suggest that when within-event residuals are further decomposed into nonrandom site-specific components and remaining random components, relative contributions from the between-event, site-to-site, and within-event single-site components are relatively similar to one another. The comparison of the standard deviations (i.e., τ, ϕ, and ϕ_{SS}) for different earthquake types indicates that standard deviations for crustal earthquakes are larger than those for interface and inslab events. The calculated standard deviations can be used for adjusting aleatory variability associated with the prediction models for different earthquake types, when they are applied in seismic-hazard analysis.

Spatial Correlation of Within-Event Residuals

It is important to recognize that the within-event residuals are spatially correlated (Goda and Atkinson, 2010; Goda, 2011; Esposito and Iervolino, 2012; Sokolov and Wenzel, 2013). The spatial correlation is due to the fact that ground motions at nearby sites are similar because they are affected by common factors such as propagation path, regional geology, and local soil condition, and regression models cannot account for the details of such common features completely. Hence, some systematic trends that are spatially dependent are present in the within-event residuals. The empirical spatial correlation model captures such trends as a function of separation distance between two sites Δ_s. The basic steps of evaluating spatial correlation of within-event residuals involve: (1) constructing data pairs and calculating within-event residual gap $\delta W_d = \delta W_i - \delta W_j$ (in which δW_i and δW_j are the within-event residuals at sites i and j, respectively); (2) assessing the sample semivariogram $[\phi_d(\Delta_S)]^2/2$ (Goovaerts, 1997), in which $[\phi_d(\Delta_s)]^2$ is the variance of $\delta W_d(\Delta_s)$ that falls within a separation distance bin represented by Δ_s; and (3) evaluating the within-event spatial correlation coefficient $\rho_{\delta W}(\Delta_s)$ as

$$\rho_{\delta W}(\Delta_s) = 1 - \frac{1}{2}\left(\frac{\varphi_d(\Delta_s)}{\varphi}\right)^2. \quad (9)$$

There are several ways to evaluate the spatial correlation model numerically. One approach is to calculate the standard deviation of within-event residuals for each event and use it for normalizing the semivariogram (referred to as normalization method 1). Alternatively, the standard deviation can be evaluated from within-event gap data $\delta W_d(\Delta_s)$ with long separation distances (i.e., the plateau level of the semivariogram, noting that $\phi \approx \phi_d(\Delta_s)/\sqrt{2}$ with a sufficiently large value of Δ_s). This is referred to as normalization method 2. For well-recorded events, estimates of the spatial correlation can be obtained individually, and are useful for characterizing interevent variability of the spatial correlation models (Goda, 2011). However, this may not be feasible when the available ground-motion dataset does not include well-recorded events (e.g., more than 100 records per event). In such cases, another common approach is to pool within-event residuals from all earthquakes in the dataset for regression analysis and to develop a combined spatial correlation model (Jayaram and Baker, 2009; Esposito and Iervolino, 2012; Sokolov and Wenzel, 2013). For this method, prior to combining the data from multiple earthquakes, δW_d data should be normalized by ϕ for individual events, because values of ϕ vary significantly depending on events (Goda, 2011). The advantage of the latter approach is that in contrast with results based on individual events alone, the estimates of spatial correlation at shorter separation distances can be obtained (because the number of data pairs per bin at short separation distance is increased and thus the estimates of $[\phi_d(\Delta_s)]^2$ can be computed). Because this study employs an extensive ground-motion dataset containing many well-recorded earthquakes (203 events with more than

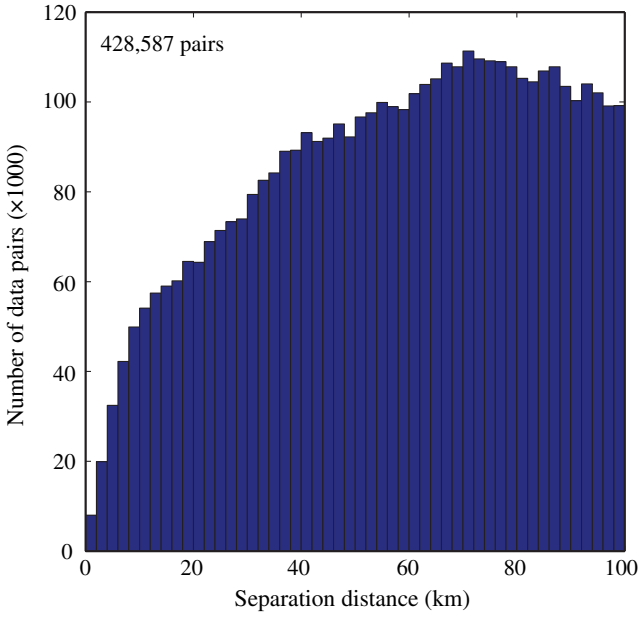

Figure 5. Separation distance for data pairs used for spatial correlation analysis. The color version of this figure is available only in the electronic edition.

100 recordings), both individual and combined approaches are pursued.

Conventionally, the above spatial correlation analysis procedure has been applied to the within-event residuals δW. The method can also be used for analyzing the spatial correlation of site-to-site residuals δS2S as well as within-event single-site residuals δWS (and relevant site-related variables such as V_{S30}; Jayaram and Baker, 2009; Sokolov and Wenzel, 2013). In particular, it is interesting to compare the spatial correlations of δW, $\delta S2S$, and δWS to examine the similarity and dissimilarity of different residual components which are related to different physical features.

Assessment of Spatial Correlation Models

The spatial correlation analysis of the within-event residuals (i.e., δW in equation 8) for I_A and CAV is carried out. The records that are used for the spatial correlation analysis are for well-recorded earthquakes having more than 100 recordings per event. Figure 5 shows the histogram of the separation distance for the data pairs used in the spatial correlation analysis. For the subset of residual data used for the spatial correlation analysis, trends of the explanatory regression variables are examined. No notable dependency of the residuals is observed for the subset (similar to Fig. 4). The following analysis focuses on the spatial correlation results based on the linear site-response model; the results based on the nonlinear site-response model are very similar and therefore, the observations and conclusions are applicable to both cases.

First, to show the variability of the within-event standard deviation ϕ for different events, estimates of ϕ based on normalization methods 1 and 2 are compared in Figure 6a. The two methods produce similar results. Importantly, it is clear that ϕ is highly variable for different earthquake events (each estimate of ϕ is calculated using more than 100 data points). It is noteworthy that there are no clear trends between ϕ and explanatory variables (e.g., H, **M**, and event type). This indicates that using a constant value of ϕ for the entire event set will result in significant bias of the spatial correlation model

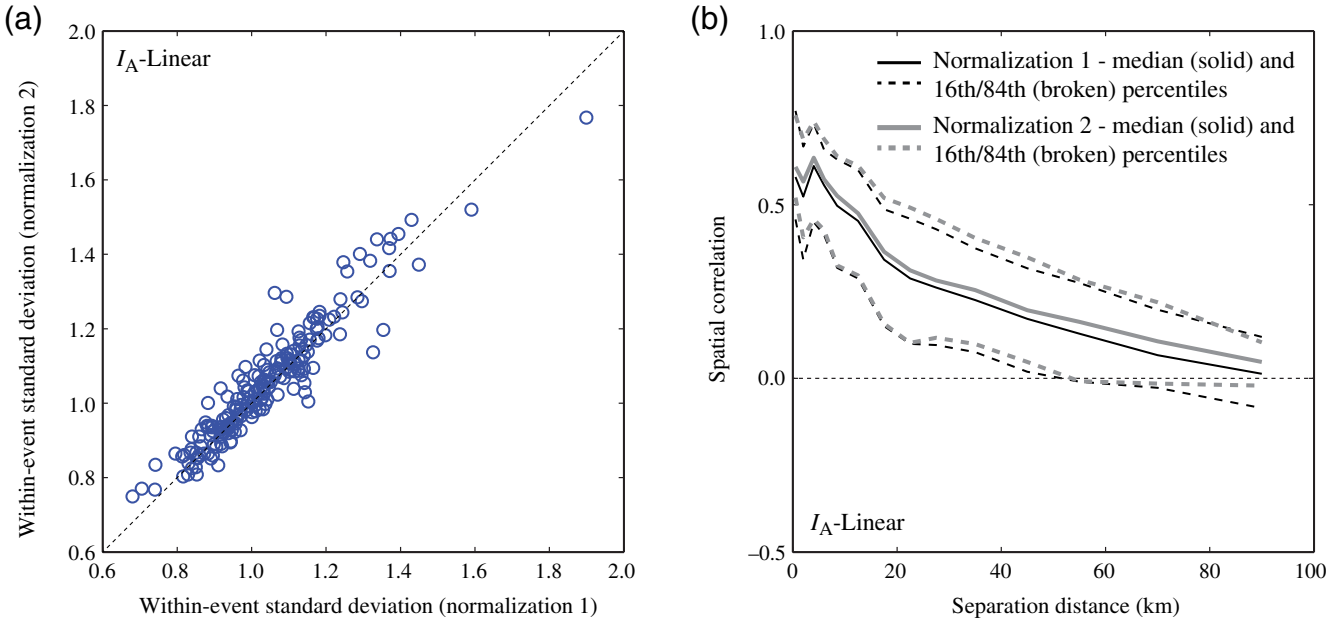

Figure 6. (a) Comparison of within-event standard deviations based on different normalization approaches and (b) comparison of within-event spatial correlations based on different normalization approaches (median and 16th or 84th percentile curves based on individual events). The color version of this figure is available only in the electronic edition.

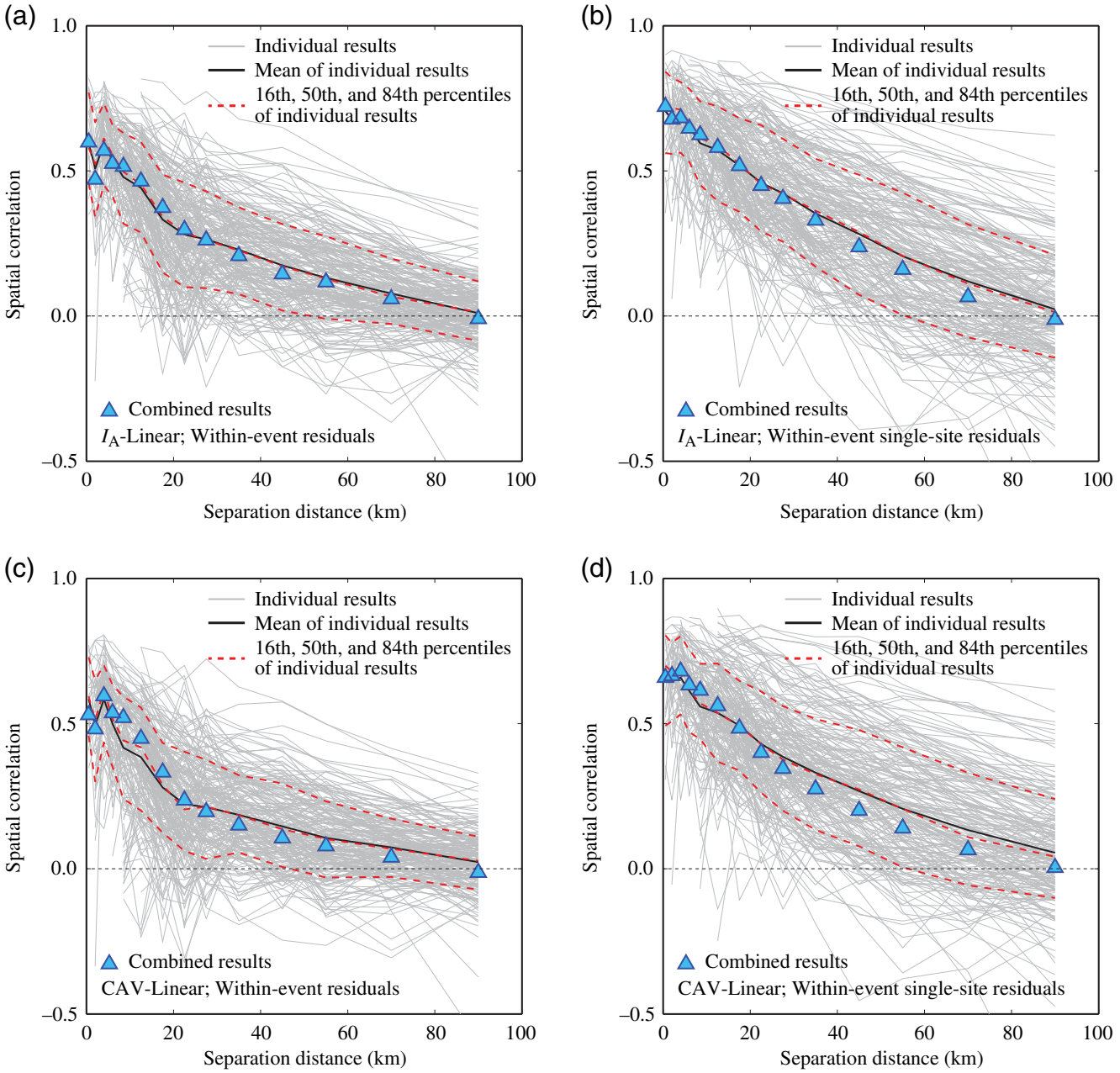

Figure 7. Spatial correlations of I_A based on (a) within-event residuals and (b) within-event single-station residuals and spatial correlations of CAV based on (c) within-event residuals and (d) within-event single-station residuals. The color version of this figure is available only in the electronic edition.

(equation 9). Figure 6b compares the obtained spatial correlation results for I_A based on the two normalization methods. In the figure, three percentile levels (i.e., median and 16th and 84th percentile curves based on 203 individual curves) are presented. The results indicate that for the well-recorded earthquakes in Japan, the impact of using different normalization methods is insignificant. Thus in the following analysis, normalization method 1 is adopted.

Figure 7 shows the spatial correlation results for I_A and CAV by considering both within-event residuals (δW) and within-event single-site residuals (δWS). In each figure panel, results for individual events and statistics based on the individual event results (mean, median, and 16th and 84th percentile curves) are presented. In addition, results based on the combined dataset are included. The results shown in Figure 7 indicate that the estimates based on the individual events and those based on the combined dataset are consistent, particularly the average trends, and that spatial correlations for I_A and CAV are similar (although the degrees of site-to-site variability are different). Variability of the individual spatial correlation curves is relatively large; typically, the standard deviation of the spatial correlation coefficient

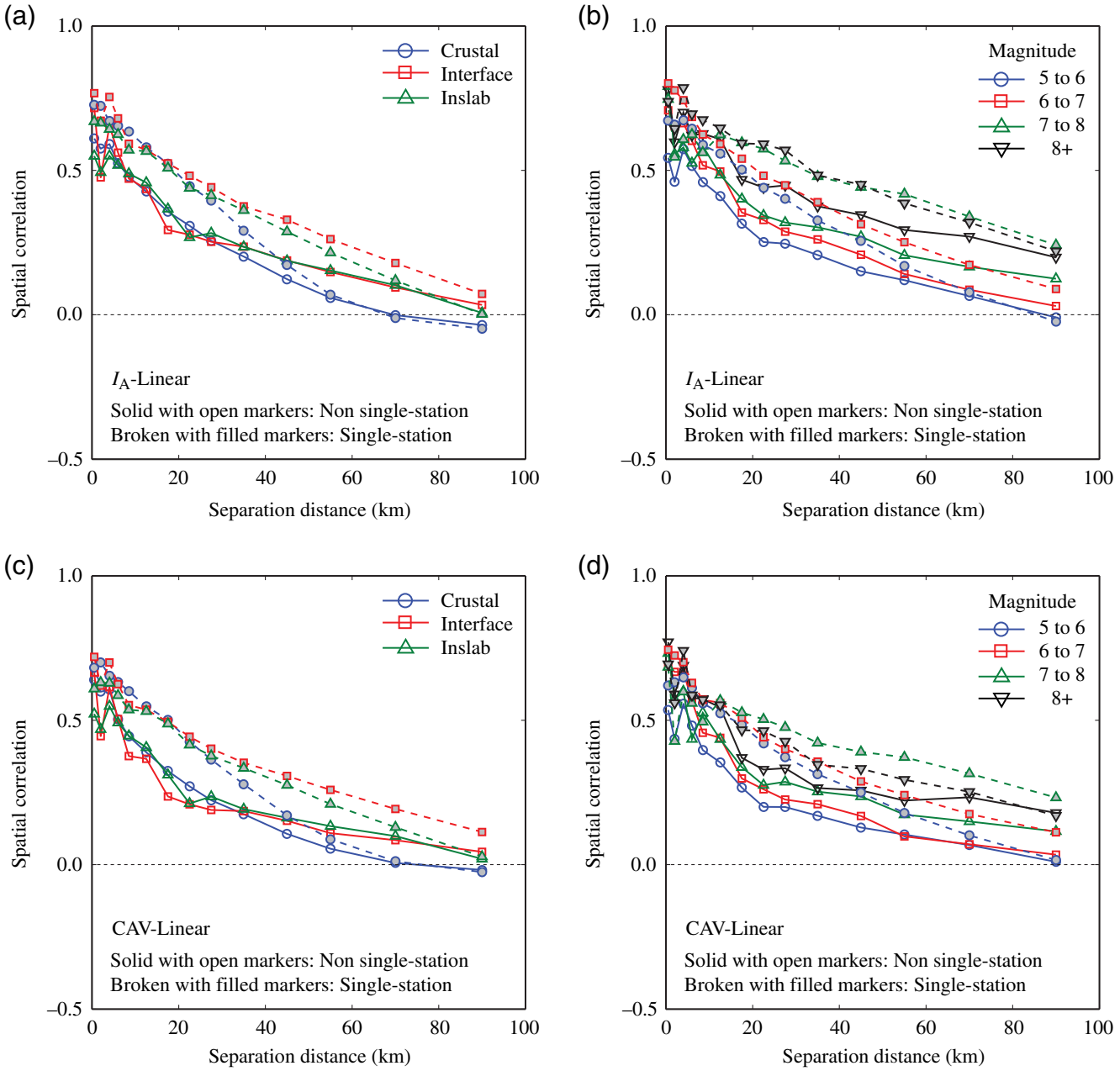

Figure 8. Dependency of spatial correlation of I_A on (a) earthquake type and (b) magnitude and dependency of spatial correlation of CAV on (c) earthquake type and (d) magnitude. The color version of this figure is available only in the electronic edition.

ranges from 0.1 to 0.2, depending on the separation distance range. Both cases with and without single-site adjustments exhibit decaying correlation with separation distance, indicating that spatial dependency of ground-motion parameters is attributed not only to site effects but also to path effects and other factors that are not associated with the site conditions. Furthermore, comparison of the spatial correlation curves for the within-event residuals and the within-event single-site residuals (e.g., Fig. 7a versus 7b and Fig. 7c versus 7d) shows that the correlations for the single-site cases are higher than those without single-site adjustments, and the variability (i.e., ranges between the 16th percentile and 84th percentile curves) is uniform in terms of separation distance (about 0.15).

To further investigate the dependency of spatial correlation of the within-event residuals and the within-event single-site residuals on event characteristics, results for individual events are grouped according to earthquake type (crustal/interface/inslab) and earthquake magnitude (5 to 6/6 to 7/7 to 8/8+). The results are shown in Figure 8; the solid lines with open markers are for the within-event residuals, whereas the broken lines with filled markers are for the within-event single-site residuals. The results shown in Figure 8a and 8c indicates that the spatial correlations for the three event types

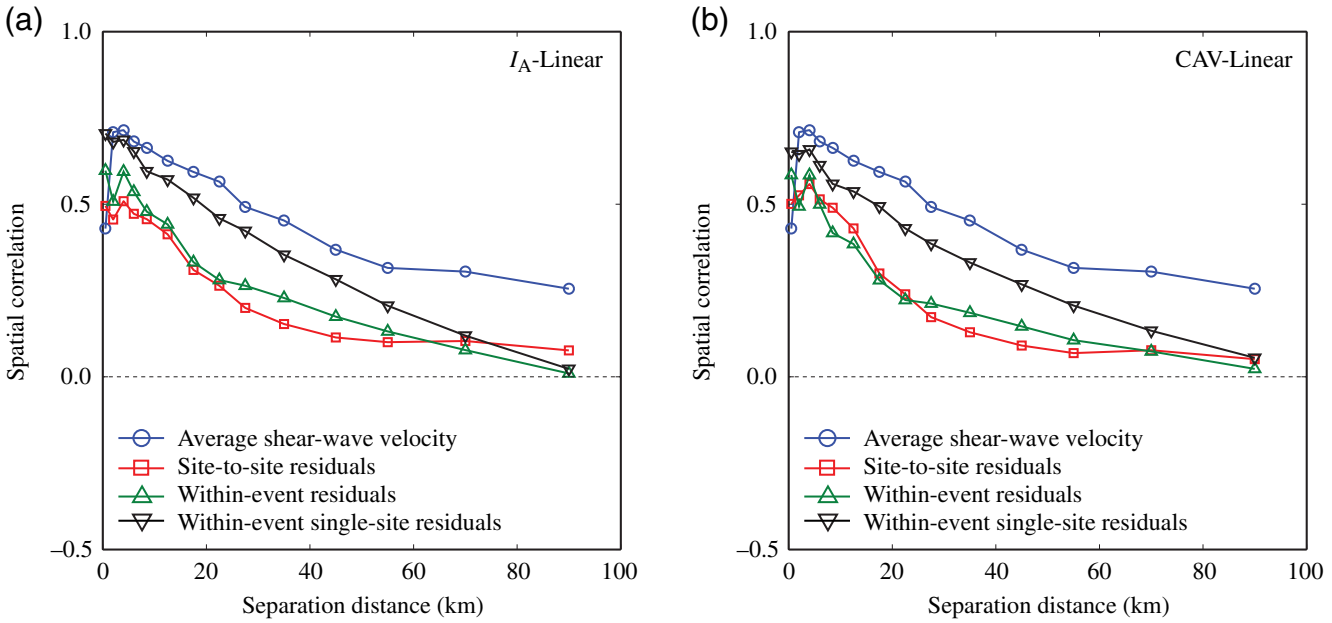

Figure 9. Comparison of spatial correlation based on average shear-wave velocity, site-to-site residuals, within-event residuals, and within-event single-station residuals for (a) I_A and (b) CAV. The color version of this figure is available only in the electronic edition.

are similar in the short separation distance range (up to 40 km); the differences are more noticeable when the single-site cases are considered. On the other hand, Figure 8b and 8d shows that the spatial correlations for larger events tend to be higher than those for smaller events (i.e., more nonrandom components), which is in agreement with the observations by Sokolov and Wenzel (2013).

Finally, Figure 9 compares the average spatial correlation curves for I_A and CAV based on the average shear-wave velocity (V_{S30}), site-to-site residuals ($\delta S2S$), within-event residuals (δW), and within-event single-site residuals (δWS). The results show that the correlation curves for V_{S30} are consistently higher than other curves and that the curves for V_{S30} and $\delta S2S$ do not return to zero correlation for large separation distances of 80–100 km, indicating that these two site parameters are affected by nonrandom components that are correlated at greater spatial scales (e.g., similarity of site conditions in Kanto plain; Sokolov and Wenzel, 2013). Overall, the results presented in Figures 8 and 9 suggest that different residual components, although exhibiting consistent decaying features with separation distance, have different degrees of spatial correlation, which are influenced by the underlying physical processes. When probabilistic seismic hazard and risk analysis is conducted for spatially distributed structures and infrastructure systems for specific scenarios, both reduction of uncertainties related to empirical GMPEs (by removing systematic and repeatable components) and changes to their spatial dependency need to be taken into account. The results presented in this study contribute to such improved assessment of ground-motion uncertainty and dependency related to duration-based ground-motion measures.

Conclusions

This article introduced new ground-motion models for I_A and CAV for Japan. The dataset that was used for developing the new GMPEs includes the 2011 Tohoku earthquake records, and thus the developed models can be used for future **M** 9 class megathrust subduction earthquakes in Japan. The models are valid for magnitude greater than 5.0, rupture distance less than 300 km, and focal depth less than 150 km. The functional form of the developed GMPEs captures key features of the available strong-motion data. It takes account of linear magnitude scaling, style-of-faulting mechanism, earthquake event type as well as regional anelastic attenuation. Both linear and nonlinear site-response models are considered. In particular, the linear magnitude scaling up to **M** 9, rather than saturation of magnitude scaling, is considered to be adequate as supported by various regression statistics. This is related to how I_A and CAV are defined, reflecting multiple characteristics of the ground motion (e.g., amplitude and duration). The final models were developed using nonlinear random-effect regression analysis and are robust; between-event and within-event residuals calculated from the final models do not show significant trends with respect to key explanatory variables.

This study also analyzed the residuals from the ground-motion modeling to gain further insights into the uncertainties in the prediction equations for I_A and CAV. The first part of this analysis is the computation of the site-to-site standard deviations (i.e., single-station sigma). The results suggest that when within-event residuals are further decomposed into nonrandom site-specific components and remaining random components, relative contributions from the between-event,

site-to-site, and within-event single-site components are similar. The second part of the residual analysis is related to spatial correlation of ground-motion residuals. This was carried out for I_A and CAV, considering both within-event residuals (δW) and within-event single-site residuals (δWS) for individual events as well as for the combined dataset. The estimates based on the individual events and those based on the combined dataset were consistent. Variability of the individual spatial correlation curves was relatively large; the standard deviation of the spatial correlation coefficient ranges from 0.1 to 0.2, depending on the separation distance range. Comparison of the spatial correlation curves for the within-event residuals and the within-event single-site residuals showed that the correlations for the single-site cases are higher than those without station adjustments. This work is an important step toward the development of the next generation of GMPEs based on the nonergodic assumption and their application in future seismic hazard and risk studies of spatially distributed structures.

Data and Resources

Strong-motion data used in this study were obtained from K-NET and KiK-net at http://www.kyoshin.bosai.go.jp/ (last accessed August 2013) and SK-net at http://www.sknet.eri.u-tokyo.ac.jp/ (last accessed August 2013). Nonlinear regression analysis is carried out using the *nlme* package (Pinheiro *et al.*, 2008) of the software program *R* (http://www.r-project.org/; last accessed October 2014).

Acknowledgments

The authors appreciate the reviewers for their insightful comments and suggestions. This work is supported by the Alexander von Humboldt Fellowship for experienced researchers (K. G.).

References

Abrahamson, N. A., and W. J. Silva (1997). Empirical response spectral attenuation relations for shallow crustal earthquakes, *Seismol. Res. Lett.* **68,** 94–127.

Abrahamson, N. A., N. Gregor, and K. Addo (2015). BC Hydro ground motion prediction equations for subduction earthquakes, *Earthq. Spectra* doi: 10.1193/051712EQS188MR.

Akaike, H. (1987). Factor analysis and AIC, *Psychometrika* **52,** 317–332.

Al Atik, L., N. A. Abrahamson, J. J. Bommer, F. Scherbaum, F. Cotton, and N. Kuehn (2010). The variability of ground-motion prediction models and its components, *Seismol. Res. Lett.* **81,** 794–801.

Anderson, J. G., and Y. Uchiyama (2011). A methodology to improve ground-motion prediction equations by including path corrections, *Bull. Seismol. Soc. Am.* **101,** 1822–1846.

Arias, A. (1970). A measure of earthquake intensity, in *Seismic Design for Nuclear Power Plants*, R. J. Hansen (Editor), The MIT Press, Cambridge, Massachusetts, 438–483.

Atkinson, G. M. (2006). Single-station sigma, *Bull. Seismol. Soc. Am.* **96,** 446–455.

Atkinson, G. M., and D. M. Boore (2003). Empirical ground-motion relations for subduction zone earthquakes and their application to Cascadia and other regions, *Bull. Seismol. Soc. Am.* **93,** 1703–1729.

Bradley, B. A. (2012). Empirical correlations between cumulative absolute velocity and amplitude-based ground motion intensity measures, *Earthq. Spectra* **28,** 37–54.

Campbell, K. W., and Y. Bozorgnia (2010). A ground motion prediction equation for the horizontal component of cumulative absolute velocity (CAV) using the PEER-NGA database, *Earthq. Spectra* **26,** 635–650.

Campbell, K. W., and Y. Bozorgnia (2012a). Cumulative absolute velocity (CAV) and seismic intensity based on the PEER-NGA database, *Earthq. Spectra* **28,** 457–485.

Campbell, K. W., and Y. Bozorgnia (2012b). A comparison of ground motion prediction equations for Arias intensity and cumulative absolute velocity developed using a consistent database and functional form, *Earthq. Spectra* **28,** 931–941.

Chen, Y. H., and C. C. P. Tsai (2002). A new method for estimation of the attenuation relationship with variance components, *Bull. Seismol. Soc. Am.* **92,** 1984–1991.

Chiou, B. S. J., and R. R. Youngs (2008). An NGA model for the average horizontal component of peak ground motion and response spectra, *Earthq. Spectra* **24,** 173–215.

Du, W., and G. Wang (2013a). Intra-event spatial correlations for cumulative absolute velocity, Arias Intensity, and spectral accelerations based on regional site conditions, *Bull. Seismol. Soc. Am.* **103,** 1117–1129.

Du, W., and G. Wang (2013b). Fully probabilistic seismic displacement analysis of spatially distributed slopes using spatially correlated vector intensity measures, *Earthq. Eng. Struct. Dyn.* **43,** 661–679.

Electrical Power Research Institute (1988). A criterion for determining exceedance of the operating basis earthquake, *Report No EPRI NP-5930*, Palo Alto, California.

Esposito, S., and I. Iervolino (2012). Spatial correlation of spectral acceleration in European data, *Bull. Seismol. Soc. Am.* **102,** 2781–2788.

Foulser-Piggott, R., and P. J. Stafford (2012). A predictive model for Arias intensity at multiple sites and consideration of spatial correlations, *Earthq. Eng. Struct. Dyn.* **41,** 431–451.

Ghofrani, H., and G. M. Atkinson (2011). Forearc versus backarc attenuation of earthquake ground motion, *Bull. Seismol. Soc. Am.* **101,** 3032–3045.

Ghofrani, H., G. M. Atkinson, and K. Goda (2013). Implications of the 2011 M 9.0 Tohoku Japan earthquake for the treatment of site effects in large earthquakes, *Bull. Earthq. Eng.* **11,** 171–203.

Goda, K. (2011). Interevent variability of spatial correlation of peak ground motions and response spectra, *Bull. Seismol. Soc. Am.* **101,** 2522–2531.

Goda, K., and G. M. Atkinson (2010). Intraevent spatial correlation of ground-motion parameters using SK-net data, *Bull. Seismol. Soc. Am.* **100,** 3055–3067.

Goda, K., A. Pomonis, S. C. Chian, M. Offord, K. Saito, P. Sammonds, S. Fraser, A. Raby, and J. Macabuag (2013). Ground motion characteristics and shaking damage of the 11th March 2011 M_w 9.0 Great East Japan earthquake, *Bull. Earthq. Eng.* **11,** 141–170.

Goovaerts, P. (1997). *Geostatistics for Natural Resources Evaluation*, Oxford University Press, New York, New York.

Jayaram, N., and J. W. Baker (2009). Correlation model for spatially distributed ground-motion intensities, *Earthq. Eng. Struct. Dyn.* **38,** 1687–1708.

Kaklamanos, J., B. A. Bradley, E. M. Thompson, and L. G. Baise (2013). Critical parameters affecting bias and variability in site-response analyses using KiK-net downhole array data, *Bull. Seismol. Soc. Am.* **103,** 1733–1749.

Kanno, T., A. Narita, N. Morikawa, H. Fujiwara, and Y. Fukushima (2006). A new attenuation relation for strong ground motion in Japan based on recorded data, *Bull. Seismol. Soc. Am.* **96,** 879–897.

Kramer, S. L., and R. A. Mitchell (2006). Ground motion intensity measures for liquefaction hazard evaluation, *Earthq. Spectra* **22,** 413–438.

Lin, P. S., B. Chiou, N. A. Abrahamson, M. Walling, C. T. Lee, and C. T. Cheng (2011). Repeatable source, site, and path effects on the standard deviation for empirical ground-motion prediction models, *Bull. Seismol. Soc. Am.* **101,** 2281–2295.

Morikawa, N., and H. Fujiwara (2013). A new ground motion prediction equation for Japan applicable up to M9 mega-earthquake, *J. Disaster Res.* **8,** 878–888.

Morikawa, N., T. Kanno, A. Narita, H. Fujiwara, T. Okumura, Y. Fukushima, and A. Guerpinar (2008). Strong motion uncertainty determined from observed records by dense network in Japan, *J. Seismol.* **12,** 529–546.

Oth, A. (2013). On the characteristics of earthquake stress release variations in Japan, *Earth Planet. Sci. Lett.* **377/378,** 132–141.

Oth, A., D. Bindi, S. Parolai, and D. D. Giacomo (2011). Spectral analysis of K-NET and KiK-net data in Japan, part II: On attenuation characteristics, source spectra, and site response of borehole and surface stations, *Bull. Seismol. Soc. Am.* **101,** 667–787.

Pinheiro, J., D. Bates, S. DebRoy, D. Sarkar, and the R Core team (2008). nlme: Linear and nonlinear mixed effects models, *R package version 3.1-89*.

Rodriguez-Marek, A., F. Cotton, N. A. Abrahamson, S. Akkar, L. Al Atik, B. Edwards, G. A. Montalva, and H. M. Dawood (2013). A model for single-station standard deviation using data from various tectonic regions, *Bull. Seismol. Soc. Am.* **103,** 3149–3163.

Rodriguez-Marek, A., G. A. Montalva, F. Cotton, and F. Bonilla (2011). Analysis of single-station standard deviation using the KiK-net data, *Bull. Seismol. Soc. Am.* **101,** 1242–1258.

Schwarz, G. (1978). Estimating the dimension of a model, *Ann. Stat.* **6,** 461–464.

Sokolov, V., and F. Wenzel (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, *Bull. Earthq. Eng.* **11,** 1909–1926.

Stafford, P. J. (2014). Crossed and nested mixed-effects approaches for enhanced model development and removal of the ergodic assumption in empirical ground-motion models, *Bull. Seismol. Soc. Am.* **104,** 702–719.

Zhao, J. X., J. Zhang, A. Asano, Y. Ohno, T. Oouchi, T. Takahashi, H. Ogawa, K. Irikura, H. K. Thio, P. G. Somerville, Y. Fukushima, and Y. Fukushima (2006). Attenuation relations of strong ground motion in Japan using site classification based on predominant period, *Bull. Seismol. Soc. Am.* **96,** 898–913.

Centre for Risk Studies
Judge Business School
University of Cambridge
Trumpington Street
Cambridge CB2 1QA, United Kingdom
r.foulser-piggott@jbs.cam.ac.uk
(R.F.-P.)

Department of Civil Engineering
University of Bristol
Office 2.41
Queen's Building
University Walk
Clifton BS8 1TR Bristol, United Kingdom
(K.G.)

Manuscript received 22 October 2014;
Published Online 21 July 2015