Published on Earthquake Engineering and Structural Dynamics,doi: 10.1002/eqe.3067

Regional Ground-Motion Prediction Equations for amplitude-, frequency response- and duration-based parameters for Greece

Konstantinos Chousianitis[1], Vincenzo Del Gaudio[2], Pierpaolo Pierri[2], G.-Akis Tselentis[1]

[1] Institute of Geodynamics, National Observatory of Athens, Lofos Nymfon, 11810 Athens, Greece
[2] Dipartimento di Scienze della Terra e Geoambientali, Università degli Studi di Bari, via E. Orabona 4, 70125, Bari, Italy

Keywords: Ground-motion prediction equations; nonlinear regression; residual distribution; distance-dependent saturation effects; anelastic attenuation

Abstract

Although all of the main properties of a ground-motion cannot be captured through a single parameter, a number of different engineering parameters has been proposed that are able to reflect either one or more ground-motion characteristics concurrently. For many of these parameters, especially regarding Greece, there are relatively few or no published equations available for their prediction. In this context, we present a set of new regionally-calibrated equations for the prediction of the geometric mean of the horizontal components of ten amplitude-, frequency response- and duration-based parameters for shallow crustal earthquakes. These equations supersede previous empirical relationships for Greece since their applicability range for magnitude and epicentral distance has been extended down to M_w 4 and up to 200 km respectively, the incorporation of a term accounting for anelastic attenuation has been investigated, while their development was based on a ground-motion dataset spanning from 1973 to 2014. For all ground-motion parameters we provide alternative optimal equations relative to the availability of information on the different explanatory variables. In all velocity-based and contrary to the acceleration-based parameters, the anelastic attenuation coefficient was found statistically insignificant when it was combined with the geometric decay and the coefficient accounting for saturation with distance. In the regressions where the geometric decay coefficient simultaneously incorporated the contribution of anelastic attenuation, its increase was found to be much less considerable in the velocity-based than in the acceleration-based parameters, implying a stronger effect of anelastic attenuation on the parameters that are defined via the acceleration time history.

1. INTRODUCTION

A ground-motion prediction equation (GMPE) is a mathematical model that relates a dependent parameter to independent variables which characterize the earthquake source, the propagation path of the seismic waves, and the local site conditions at a particular site of interest. Since these relations provide estimates of the expected ground-motion triggered due to a specific earthquake scenario, they are especially useful in seismic hazard assessment and earthquake-resistant design. They are required to develop ground-motion hazard curves and are essential in either deterministic or probabilistic approaches, while are widely used for evaluation of the potential seismic performance of engineering structures. The lack of representative regional GMPEs is unquestionably a major source of uncertainty in seismic hazard assessment, although their applicability during the last decades has been extended to cover a variety of planning and environmental managing problems. In this context GMPEs have been incorporated into methodologies for the evaluation of the recurrence time of earthquake-induced landslide triggering (Del Gaudio et al., 2003; Du and Wang, 2014; Chousianitis et al., 2016), the evaluation of hazard from induced seismic events due to oil and gas activities (Bourne et al., 2015), as well as in the development of multi-hazard models (De Risi and Goda, 2016; Bathrellos et al., 2017).

As far as Greece is concerned, and despite the importance of proper regionally calibrated GMPE, during the last 15 years the only regional models for the prediction of ground-motion parameters for shallow crustal earthquakes were developed in the studies of Skarlatoudis et al. (2003) and Danciu and Tselentis (2007). The GMPEs of the former study were defined based on 619 records from 225 earthquakes in Greece, while the latter study used a ground motion dataset which contained 355 records from 151 earthquakes. Skarlatoudis et al. (2003) analyzed only peak ground motion parameters, while Danciu and Tselentis (2007) considered a number of engineering parameters along with spectral acceleration. The derived equations of both papers are valid for moment magnitudes between $4.5 \leq M_w \leq 7.0$ and for epicentral distance range of 0-160 km and 0-136 km respectively. Also, both studies used the same functional form with the same dummy variables and their strong-motion database, which consisted mainly of analogue records, covered the period from 1972 to 1999. The only worth mentioning difference between these studies is that in Skarlatoudis et al.

(2003) both horizontal components were used, while Danciu and Tselentis (2007) used the arithmetic average between the two horizontal components.

However, it is evident that the currently available GMPEs for Greece are based on a somewhat outdated strong-motion dataset. This is corroborated by the fact that since 2000 the analogue instruments were gradually replaced by digital, thus ensuring the elimination of digitization procedure errors, while after 2009 a serious instrumentation upgrade with high resolution instruments along with a targeted densification of the strong-motion network took place. Therefore, the development of new GMPEs for Greece based on an expanded ground-motion dataset spanning further than 1999 is warranted. Such an increase of the amount of strong-motion data should significantly improve the reliability and accuracy of GMPEs derived from them. Moreover, ground-motion parameters other than those commonly used to quantify the level of seismic shaking, have been demonstrated to be suitable in a variety of engineering applications. For instance, some of these parameters such as CAV are able to be used in various problems of geotechnical earthquake engineering. These include the rapid damage assessment after a major earthquake through near-real-time tools such as the instrumental intensity ShakeMaps, the evaluation of liquefaction potential and the assessment of hazard due to earthquake-induced landslides and slope failures. In addition, these parameters can be also useful in structural engineering, where scaling of earthquake ground motions enable nonlinear response-history analysis of buildings either for design or performance assessment purposes. Until now, GMPEs for some of these parameters have been developed only in the study of Danciu and Tselentis (2007). Accordingly, this singularity creates a shortage in conducting seismic-hazard assessment using, for instance, some frequency response- or duration-based parameters, emphasizing the need for the development of new GMPEs or the revision of existing ones.

Considering the aforementioned limitations that accompany the currently available GMPEs for Greece and justify the need for their update, we present a set of new equations for the prediction of ground-motion parameters for shallow crustal earthquakes. The proposed models have been calibrated for the estimation of the geometric mean of the horizontal components of peak horizontal acceleration, peak horizontal velocity, effective design acceleration, Housner spectrum intensity, acceleration spectrum intensity, velocity spectrum intensity, mean period, characteristic intensity, specific energy density and cumulative absolute velocity. For the development of the new GMPEs, we partitioned the entire dataset of available strong-motion records into a training and a validation subset, an approach which was not adopted in the Skarlatoudis et al. (2003) and Danciu and Tselentis (2007). The training dataset was used for the calculation of the regression coefficients, while the validation dataset was used for evaluating the effectiveness of the derived models in predicting the corresponding ground-motion parameters. This strategy is important in the assessment of any regression-generated model for a series of reasons. For example, upon building a predictive equation it is expected that the incorporation of more parameters may insubstantially improve the fit to the observed data, even only because it better fits a random "noise" due to observation errors, while upon comparing an equation obtained from a particular dataset with other different models, any conclusion should be made on the basis of a different dataset from that used for the regression of the inspected model. Thus, an objective evaluation on the predictive capability or appropriateness of a specific model can be obtained only by applying that equation to a dataset different from the one used to determine its coefficients.

The equations of the present study supersede previous GMPEs derived for Greece since they present a number of novelties with respect to the previous models. Firstly, their development is based on a updated amount of higher quality strong-motion data than was available in the previous studies. In this context we included records from earthquakes after 2000 that were not available in the studies of Skarlatoudis et al. (2003) and Danciu and Tselentis (2007), as well as we excluded all bad quality data with low signal-to-noise ratio existent in the database prior to 2000 which was utilized in the two aforementioned studies. This facilitates the robust estimation of the regression coefficients and the inclusion of data from small earthquakes at very short distances and from large events at regional distances. Moreover, in this study the derived equations predict the geometric mean of the horizontal components which is commonly employed in modern GMPE studies and was not the case in either of the previous papers. Also, contrary to the so far available GMPEs for Greece, we investigated the incorporation of a term accounting for anelastic attenuation among the different possible functional forms that we considered. Additionally, for all ground-motion parameters we provide alternative optimal equations relative to the availability of information on the different explanatory variables. Finally, we extended the range of the GMPEs applicability relative to moment magnitude and epicentral distance down to M_w=4.0 and up to 200 km, respectively.

2. GROUND-MOTION PARAMETERS EXAMINED

One of the most difficult but at the same time highly important tasks of earthquake engineering is the estimation of ground-motion parameters that facilitate the realistic estimation of the structural response and the quantification of the damage potential of earthquake strong ground motions. In this context amplitude, frequency content, duration as well as the number of picks above certain amplitudes in the recorded time history are the main properties of a ground motion

which are essential for describing its important characteristics. Amplitude has a critical role in the characterization of the strong ground motion, but it is the frequency content of that motion which affects the response of a structure and illustrates how the amount of energy is incorporated among the different frequencies. Strong ground motion duration depends on the time required for rupture to propagate across the fault and has a strong influence on earthquake damage, thus being important for seismic risk assessment. A number of engineering parameters has been proposed to represent either just one of the aforementioned characteristics of strong ground motions or more of them. However, because of the complex nature of the earthquake mechanism, the implementation of engineering needs such as damage potential assessment, probabilistic seismic hazard assessment and design criteria compilation through the identification of a single engineering parameter is not feasible. From this point of view, the parameters that need to be evaluated for specific engineering purposes depend on their intended use and their adequacy can be assessed by reviewing their definitions.

2.1 Amplitude-based parameters

The simplest as well as widely used parameters are the peak horizontal acceleration (PHA), which is adopted in many structural design codes worldwide, along with peak horizontal velocity (PHV). Both are the peak values of acceleration and velocity time histories of a strong ground motion and therefore are direct amplitude measures. Despite their similar definition these parameters are suitable for different spectral regions, in a way that PHA can be considered a high-frequency parameter, while the PHV a mid- to low-frequency parameter. This is caused by the integration process which has the ability to enhance low-frequency against the high-frequency components. Due to their simplicity and wide applicability in earthquake engineering, a large number of GMPEs have been published for their prediction (Douglas and Edwards 2016). The most recent predictive models for PHA where data from Greek earthquakes have been considered in the regression analyses are those of Ambraseys et al. (2005), Akkar and Bommer (2007b), Akkar and Bommer (2010), Akkar et al. (2014) and Kotha et al. (2016). However, only Skarlatoudis et al. (2003) and Danciu and Tselentis (2007) used solely Greek data to derive regionally calibrated GMPE for PHA, while as stated before, the time span of their database was until 1999. As far as PHV is concerned, Akkar and Bommer (2007a), Akkar and Bommer (2010), Akkar et al. (2014) and Kotha et al. (2016) proposed equations valid for Europe and Middle-East regions, and Skarlatoudis et al. (2003) and Danciu and Tselentis (2007) established regional models for the prediction of PHV within Greece.

Another amplitude-based parameter that is similar to the PHA is the Effective Design Acceleration (EDA) which corresponds to the peak value that remains after applying a low-pass filter with a cut-off frequency of 9 Hz to the acceleration time history (Electrical Power Research Institute [EPRI], 1988). Its concept emanated from the fact that high acceleration pulses located at high frequencies bring about little response in the majority of engineering structures. Accordingly, the Effective Design Acceleration can be considered as the acceleration which is efficient in triggering structural damage. Currently, this parameter has never been incorporated in a GMPE.

2.2 Frequency response-based parameters

The Acceleration Spectrum Intensity (ASI) was proposed by Von Thun et al. (1988) in an effort to characterize strong ground motion for analysis of reinforced concrete dams. It is a frequency response-based parameter which is defined through the following equation:

$$ASI = \int_{0.1}^{0.5} S_a(\xi = 0.05, T)dT \quad (1)$$

where S_a is the spectral acceleration for 5% damping and between periods of 0.1 sec and 0.5 sec, which comprise the typical range of the fundamental periods of concrete dams. Bradley (2010) concluded that ASI is a parameter with a better predictability than other common ground-motion parameters such as PHV and spectral accelerations. This was illustrated by comparing the total-event and intra-event lognormal standard deviations of these parameters and showing that ASI has the smaller variability.

In a similar way as for the ASI, Von Thun et al. (1988) defined the Velocity Spectrum Intensity (VSI):

$$VSI = \int_{0.1}^{2.5} S_v(\xi = 0.05, T)dT \quad (2)$$

where S_v is the spectral velocity for 5% damping and between periods of 0.1 sec and 0.5 sec. For none of these two parameters a GMPE exists, apart from an indirect method presented in Bradley (2010), predicting ASI through equations that provide spectral acceleration estimates.

The above three parameters make use of spectra to measure the frequency content of accelerograms. Mean Period (T_m), comprise an individual parameter that can be used as an alternative to complete spectra for the representation of the frequency content of a strong motion record. According to Rathje et al. (1998) it is the best simplified frequency content characterization parameter, and given that the dynamic response of structural systems presents strong dependence on the frequency content of the strong motion, mean period is of particular interest for seismic design purposes. Its definition is based on the Fourier amplitude spectrum through the following relation:

$$T_m = \frac{\sum C_i^2 / f_i}{\sum C_i^2} \quad (3)$$

where C_i are the Fourier amplitudes, and f_i represent the discrete Fourier transform frequencies over the range between 0.25 and 20 Hz. Mean period is a parameter for which very few GMPE have been developed. The most widely used relation is that of Rathje et al. (2004), which however was developed on the base of records including only a small number of events with $M_w < 5.5$, and consequently it should be used with caution below this limit. Recently, Yaghmaei-Sabegh (2015) developed a model for T_m prediction using data from earthquakes that occurred in Iran. As regards Greece, no predictive model for T_m is available and this is the first time that this parameter is incorporated in a GMPE.

2.3 Duration-based parameters

Characteristic Intensity, I_C, is a combined parameter which is defined as:

$$I_C = (a_{RMS})^{3/2} \cdot \sqrt{t_{tot}} \quad (4)$$

where a_{RMS} is the root mean square (rms) acceleration and t_{tot} is the duration of the strong motion. Due to its definition, it is a parameter that incorporates the effect of amplitude and frequency content of a strong motion record along with its duration. It has also been noted that this parameter represents well the destructive potential of ground motions since it exhibits good correlation with the mean values of the damage index (Park et al., 1985). Despite the usefulness of this parameter, only Danciu and Tselentis (2007) have developed a GMPE for it.

The Specific Energy Density, SED, is obtained by integrating velocity square over the duration of the strong motion of an earthquake:

$$SED = \int_0^{t_{tot}} v(t)^2 \, dt \quad (5)$$

where $v(t)$ is the ground velocity time history and t_{tot} is the duration of the strong motion. This parameter represents the total kinetic energy during an earthquake and captures its variation. Currently there is an absence of GMPE for SED and accordingly we will present the first model for its prediction derived from ground-motion records.

Cumulative Absolute Velocity, CAV, is defined as the area under the curve of the absolute value of the ground acceleration as function of time, i.e.:

$$CAV = \int_0^{t_{tot}} |\alpha(t)| dt \quad (6)$$

where $\alpha(t)$ is the ground acceleration and t_{tot} is the duration of the strong motion. It has been used to quantify the structural damage potential of the earthquake ground motion and since its value increases with time, it has the ability to include the cumulative effects of the ground motion duration. Despite this advantage of CAV over many amplitude-based ground motion parameters, only a few number of CAV prediction models are available in the literature. Among them, Campbell and Bozorgnia (2010) and Du and Wang (2013) used the Next Generation Attenuation strong motion database and proposed models on the basis of eight and four input parameters respectively. Sandıkkaya and Akkar (2017) presented a CAV predictive model using a database compiled from the broader European region, while Foulser-Piggott and Goda (2015) used Japanese data to develop linear and nonlinear site-response models accounting for linear magnitude scaling, fault mechanism, event type and region-specific anelastic attenuation. As far as Greece is concerned, the only available GMPE for CAV is the one developed by Danciu and Tselentis (2007) with the abovementioned dataset of Greek earthquakes that occurred up to 1999.

3. METHODOLOGY

Functional forms commonly used for GMPEs include terms representing ground motion dependence on source properties (magnitude, focal mechanism) and wave propagation (attenuation, site effect). However, the estimate of equation parameters that optimize its predictive capacities, can be biased by correlations, within the regression analysis dataset, between the values of explanatory variables. Indeed, not all the possible combinations of these values can be evenly sampled by a database of ground motion recordings: small magnitude events can be recorded only within short

distances from the source, whereas, for strong events, recordings acquired at very short distances are often missing or very few in comparison to long distance observations. Since larger magnitudes and longer distances produce opposite effects on shaking, the prevalence of recordings of small events at short distances and strong events at long distances tends to produce, in equation predictions, a lower rate of ground motion decay with distance in comparison to reality (Fukushima and Tanaka, 1990). The databases used nowadays to develop GMPEs are much more complete than they used to be; in particular small events are much better recorded. Recordings of strong events at short distances are still missing though.

One way to counter this bias consists of decoupling the determination of coefficients of terms accounting for ground motion "intra-event" variability (among different sites, for the same event) from those representing "inter-event" variability (among different events, as effect of source differences) through a two-stage procedure (cf. Joyner and Boore, 1991, 1993). The starting point of this procedure is a general equation predicting a ground motion parameter Y, according to the form

$$\log Y = a + b \cdot M + c \cdot \log\sqrt{R^2 + h^2} + d \cdot \sqrt{R^2 + h^2} + \sum_{i=1,N} e_i \cdot s_i + \sum_{j=1,M} f_j \cdot m_j + \varepsilon_r + \varepsilon_e \quad (7)$$

where M is the event moment magnitude, R the epicentral distance in km, s_i and m_j are N and M dummy variables for site and focal mechanism types, respectively, a, b, c, d, e_i, f_j, h^2 are the regression parameters, ε_r and ε_e are the estimate errors related to unmodeled intra- and inter-event variability, respectively. During the development of GMPEs there are two main options regarding the adopted distance metrics that can be categorized as either point source or finite fault measures each one of which has advantages and disadvantages. In the present study we used epicentral distances due to the fact that the lack of detailed information about the exact rupture plane for the majority of the moderate and small events did not allow us to introduce a finite fault distance metric in our models. Although we acknowledge that the use of an extended fault geometry is crucial to correctly account the path effect in large magnitude events, we note that the vast majority of the earthquakes in our dataset is below $M_w = 6.4$. For such events, the discrepancy that is caused by the point-source approximation of an extended source model is relatively small due to the fact that the rupture dimensions are usually smaller than the distance to the strong motion stations. Hypocentral distance was not used so as to avoid bias due to poorly resolved focal depths particularly in the early part of the earthquake catalog. In addition, epicentral distance can be considered adequate for hazard assessment, if one calculates hazard through the definition on seismicity rates of different seismogenic zones in terms of the expected number of earthquakes of different magnitude having epicenter within each zone.

At the first stage, the regression is carried out introducing one dummy variable n_k for each of the L events recorded by the regression dataset, according to the equation

$$\log Y = c \cdot \log\sqrt{R^2 + h^2} + d \cdot \sqrt{R^2 + h^2} + \sum_{i=1,N} e_i \cdot s_i + \sum_{k=1,L} g_k \cdot n_k + \varepsilon_r \quad (8)$$

where g_k are coefficients to be determined together with those of the terms accounting for geometric spreading (c), anelastic attenuation (d) and site effect (e_i) and with the parameter representing the effect of ground motion saturation with distance (h). Since the last parameter makes the regression nonlinear, ordinary least-square regressions are iteratively carried out on a linear approximation of the equation (8), according to the matrix form

$$\mathbf{Y_1} = \mathbf{X_1}\mathbf{B_1} + \boldsymbol{\varepsilon_1} \quad (9)$$

where $\mathbf{Y_1}$ is the vector of the log$Y$ values for each observation, $\mathbf{X_1}$ is the matrix of the partial derivatives of (8) with respect to the equation parameters listed in the $\mathbf{B_1}$ vector, and $\boldsymbol{\varepsilon_1}$ is the vector of residuals. We adopted the criterion used by Joyner and Boore (1993) to stop iterations when the variation of h from the previous iteration is very small, as a good indicator of the closeness to a minimum of residual rms.

The final estimated values $\hat{g}_k$ resulting from the nonlinear regression are then employed in the second stage regression as response variable of a set of L equations

$$\hat{g}_k = a + b \cdot M + \sum_{j=1,M} f_j \cdot m_j + \varepsilon_g + \varepsilon_e \quad (10)$$

aimed at estimating the source-dependent parameters a, b, and f_i. Here ε_g is the error $\left(\hat{g}_k - g_k\right)$ affecting the estimate of g_k. In matrix form, these equations can be written as

$$\mathbf{Y_2} = \mathbf{X_2}\mathbf{B_2} + \boldsymbol{\varepsilon_2} \quad (11)$$

where $\mathbf{Y_2}$ is the vector of the values $\hat{g}_k$, $\mathbf{X_2}$ is the matrix reporting the values of the explanatory variables of (11), $\mathbf{B_2}$ is the vector of the parameters to be determined and $\boldsymbol{\varepsilon_2}$ is the vector of residuals given by the sum of errors ε_g and ε_e. Since these two kinds of errors can be considered independent and uncorrelated (cf. Joyner and Boore, 1993), the covariance of equation (11) residuals can be assumed equal to the sum of those of ε_g and ε_e, i.e.

$$\mathbf{V_2} = \mathrm{var}(\varepsilon_g) + \sigma_e^{\,2}\mathbf{I} \quad (12),$$

where var(ε_g) is derived from the first step regression, **I** is the identity matrix and errors ε_e deriving from inter-event variability are assumed to be characterized by a zero mean and a common variance σ_e^2. In general, the last assumptions cannot be applied also at the covariance of ε_g, in that ε_g values depend on the strength of constraints on g_k determination, which are related to the number of recordings available for each event. Thus, an ordinary least-square regression cannot be applied to equation (11) and needs to be replaced by a weighted least-squares approach, adopting the inverse of the square root of the covariance matrix (12) as weighting matrix.

Although the variance σ_e^2 is not known a priori, it can be determined by trials calculating iteratively the vector $\mathbf{B_2}$ from

$$\hat{\mathbf{B}}_2 = \left(\mathbf{X_2}^T \mathbf{V_2}^{-1} \mathbf{X_2}\right)^{-1} \mathbf{X_2}^T \mathbf{V_2}^{-1} \mathbf{Y_2} \quad (13)$$

until, following a bisection approach, a solution is found for $\mathbf{B_2}$, that makes the standard deviation of residuals of the weighted equations equal to 1 within a fixed approximation (cf. Hwang et al., 2004). The total variance σ^2 can then be estimated by summing those (σ_r^2 and σ_e^2) associated to errors ε_r and ε_e, respectively.

With regard to the development of the GMPE for T_m, we adopted the same functional form that was used by Rathje et al. (2004), excluding, as in the case of Yaghmaei-Sabegh (2015), the forward-directivity term due to lack of such information in our dataset, i.e.

$$\ln T_m = a + b \cdot (M - 6) + c \cdot R + \sum_{i=1,N} e_i \cdot s_i + \sum_{j=1,M} f_j \cdot m_j + \varepsilon_r + \varepsilon_e \quad (14)$$

The complete functional forms (7) and (14) may not be applicable in cases of missing information on source focal mechanism or on site category. Furthermore, the available regression dataset might be unable to provide good constraints for the calculation of some of the equation parameters. Therefore, following the scheme adopted in a previous study (Chousianitis et al., 2014), regressions were also performed using simpler functional forms that exclude some of the terms present in the complete equations. To evaluate if the inclusion of more terms actually improves equation predicting capability, the accelerometer database was subdivided into two parts, used as training and validation dataset, respectively. Regressions were carried out on the training dataset, which includes only events that were recorded by more than one station. The validation dataset was used to compare the effectiveness of different equations, and mainly consists of singly recorded events that were not incorporated into the training dataset. Yet, in order to adjust the range of recording distances up to the same limits as the training dataset, we supplemented the validation dataset with very few recordings that have been transferred from events belonging to the training dataset. Here we note that the main drawback of the two-stage approach is the impossibility to separate inter-event and intra-event component of residual in case of events recorded by a single station. That is why such kind of recording was excluded from the training dataset and therefore from regressions. However, such data can be used for the comparative validation stage. It is also crucial to acknowledge that although the dataset contains a number of events with only two records, which provide weaker constraints to the separation between inter-event and intra-event variability, these are not a dominant component of the total dataset and represent a minority when compared to the total number of recordings.

After each regression, Student's test t values were calculated to assess the significance of the obtained parameters. The goodness of fit provided by each tested equation was evaluated through the efficiency coefficient (Nash and Sutcliffe, 1970)

$$E = 1 - \frac{\sum \left(y_i - y^*_i\right)^2}{\sum \left(y_i - \bar{y}_i\right)^2} \quad (15),$$

which quantifies how much the equation estimations y^*_i outperform the simple mean value $\bar{y}_i$ as predictor of the experimental observations y_i. Furthermore, GMPEs are probabilistic models based on the assumption that ground-motion is log-normally distributed. The hypothesis that the actual value of a ground-motion parameter has a log-normal distribution around the median provided by the GMPE is an ideal model used to estimate the exceedance probability in hazard assessment, thus the agreement of observations with such a model should be verified to evaluate the reliability of hazard estimates. In this context, Scherbaum et al. (2009) introduced the LLH index which is based on the similarity of residual distribution to log-normality in terms of frequency for intervals of residual values and is often used to assess performance of different GMPEs. Finally, to compare the effectiveness of different functional forms in predicting ground motion observations distinct from those employed in regressions, root mean square of estimate errors and efficiency coefficient were estimated applying the equations to the validation dataset.

4. DATA AND RESULTS

The Greek accelerometric database that we used includes recordings acquired from 1973 to 2014. Since GMPEs are mainly used for the prediction of seismic damage scenarios, the calibration and validation of the models within the present study was carried out using only data relative to events of moment magnitude not less than 4.0. The dataset

incorporates the HEAD v1.0 strong-motion database for the period 1973–1999, where however we have excluded all bad quality data with low signal-to-noise ratio. After 2000 we have used selective data from the permanent network of the National Observatory of Athens (NOA) along with data from the EUROSEISTEST database (Pitilakis et al., 2013). The records which were not available in an already processed and filtered form, i.e. those of the permanent network of NOA, were corrected by applying a bi-directional second-order Butterworth filter with high-pass and low-pass cut-off frequencies equal to 0.2-0.3 Hz and 25-30 Hz, respectively. The final dataset, whose earthquakes and corresponding number of records are given in Table S1 of the supporting information, was further divided into a training and a validation dataset, as stated previously. The training dataset consists of 652 accelerograms, relative to 72 events, acquired by 124 stations, whereas the validation dataset includes 254 accelerograms relative to 123 events, acquired by 45 stations. The maximum moment magnitude for both datasets is 6.8. Figure 1 shows the geographic location of events and stations for the two datasets, whereas Figure 2 shows the corresponding source magnitude-distance distribution among the recordings of the two datasets. Although the training dataset includes a few recordings acquired very close to the seismic source, their number is too small to well constrain GMPE parameters for short distance predictions. Thus, the effectiveness of different GMPE functional forms was tested on a validation dataset including recordings acquired at epicentral distances not less than 5 km.

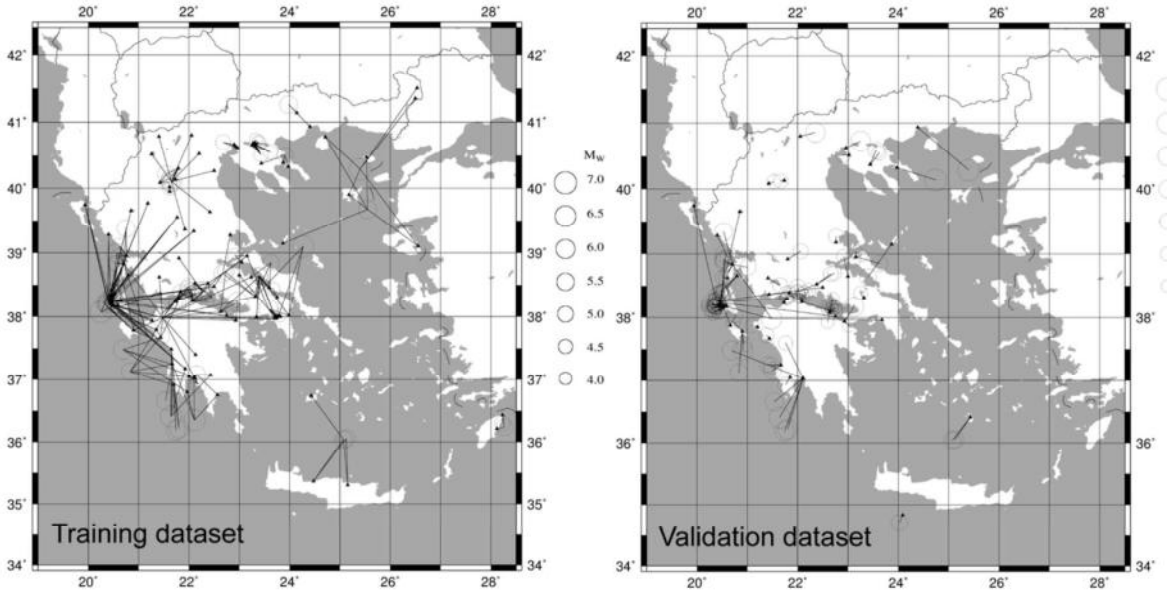


Figure 1. Geographic location of seismic stations (black triangles), event sources (circles sized according to magnitude) and source-station paths (black lines) relative to the seismic recordings included in the training dataset used for the calibration of GMPEs and the validation dataset used to compare the effectiveness of different GMPEs.

Regressions were carried out on the training dataset to derive equations predicting the ground-motion parameters mentioned earlier. As response variable, we used the geometric mean of the parameter values observed on the two horizontal components. With reference to the general form of the prediction equation (7) and (14), the categorization of event focal mechanisms was simplified unifying strike-slip and thrust fault types, following the outcome of previous studies (cf. Danciu and Tselentis, 2007), thus a single dummy variable m was included, setting its value to 0 for normal faults and to 1 otherwise. With regard to soil conditions, accelerometer sites in Greece are classified according to the criteria defined by the United States National Earthquake Hazards Reduction Program (NEHRP; see Building Seismic Safety Council [BSSC], 2003). All the stations were assigned to one of three classes, i.e. B (rock), C (stiff soil) or D (soft soil). Thus, two dummy binary variables s_i at most were needed, setting both their values to 0 for B type sites and setting, alternatively, one of the variable to 1 and the other to 0, for the C and D types.

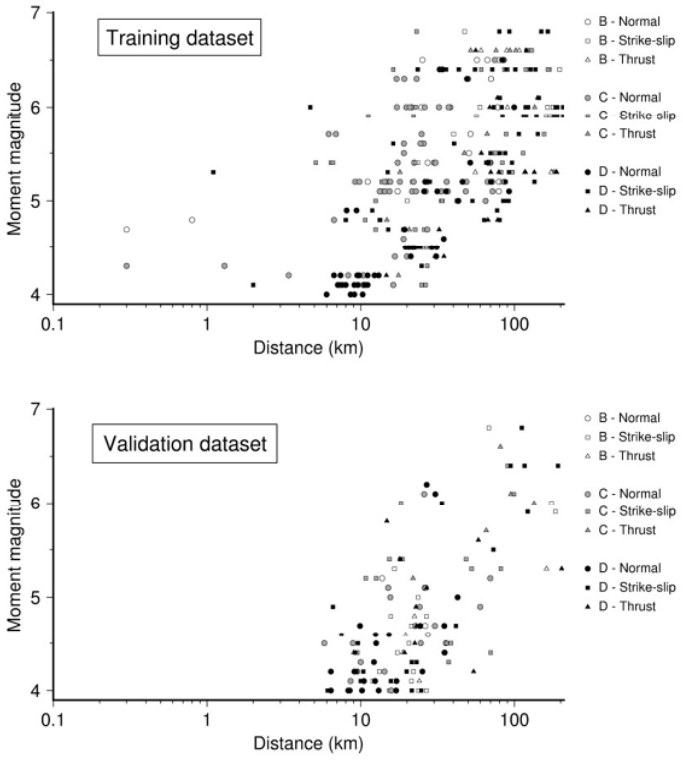


Figure 2. Magnitude of recorded events as function of distance of the recording station for the training and validation datasets. Focal mechanism of the recorded events and site class of the recording stations are represented through different symbols according to the legend.

All the functional forms tested in regression analyses for all the ground-motion parameters examined include the first three terms of equations (7) and (14). Alternative forms using all the possible combinations were tested choosing whether or not to include the coefficient h and the terms representing anelastic attenuation, focal mechanism and site effect. Ultimately, for each ground-motion parameter we carried out regressions for 24 alternative GMPE forms. For the different functional forms tested, Table S2, Table S3 and Table S4 present the regression results for the amplitude-based, frequency response-based and duration-based parameters examined in the present study. Accordingly, Table S5, Table S6 and Table S7 report the statistical parameters providing comparative indications about their quality. In these tables, the results relative to each functional form is labelled through an identification code having the form $c(1)d(0|1)e\#f(0|1)h(0|1)$. In this code, each of the optional parameters d, f and h present in the general form (7) and (14), is followed by 1 or 0, according that this parameter is calculated through regression analysis or not, whereas e is followed by the number # of dummy variables used for site classification. If just one site variable is used, soil type classes (C and D) are joint together into a single class, for which the dummy variable is set to 1.

In Table S2-S4 where the results of the regression analyses are summarized, we present only the functional forms where the t-statistics for all incorporated parameters are greater than 2 or less than -2. The opposite indicates the presence of coefficients which may well be removed from the model without hurting the fit substantially, thus implying the low reliability of the predictive power of that particular coefficient. In these tables the different functional forms are ordered by increasing root mean square of the errors affecting the logarithmic estimates of the ground-motion parameters, when the different GMPEs are applied to the validation dataset. This order corresponds to a decreasing effectiveness shown by the GMPE in predicting ground motion for events different from those whose recordings were used in regression analyses. Accordingly, in Table S5-S7 the t-values for each regression parameter, the efficiency coefficients E and LLH index are reported. For each ground-motion parameter, we end up with the most effective functional forms to be used depending on the availability of information on explanatory variables. The proposed GMPEs were chosen with primary criterion the effectiveness in predicting the validation dataset based on the *rmsl* values, followed by the log-likelihood index according to Scherbaum et al. (2009) and the efficiency coefficient calculated from the validation dataset. In the cases where the *LLH* and the E_{val} values between two models with the same available information are approximately the same, priority was given at the model that separates the propagation effects and has the largest number of independent variables. The latter choice, although can be considered counter-intuitive, is based on the fact that models that disregard

for instance the style-of-faulting predictor parameter or do not account for site effects, can be sometimes inadequate for hazard assessment studies that are site-specific and involve distinction between different soils and/or rock conditions, or require the inclusion of a particular rupture type.

5. DISCUSSION

5.1 Amplitude-based parameters

The aforementioned tables that summarize the regression analyses results and report various parameters obtained from either the training or the validation dataset, allow a quantitative evaluation of the quality of the results and the goodness of fit as well as a comparison of the predictive capabilities of the different formulae. In this context and with regard to PHA, one can notice that, although the smallest value of regression standard deviation (σ = 0.285) was obtained with a functional form using the highest number of explanatory variables (*c1d1e1f1h1*, see the row 7 in Table S2 for PHA), other simpler functional forms proved more effective in predicting PHA values of the validation dataset. In particular, the minimum *rmsl* (0.279) and the maximum efficiency coefficient (E = 0.676) was obtained with a GMPE that models the attenuation by including both an anelastic term and the saturation coefficient h, but not including terms accounting for focal mechanism and site effect (code *c1d1e0f0h1*). More in general, the inclusion of a saturation coefficient is the most frequent feature among the GMPEs with the best predictive performance. On the other hand, the best predictions are provided by equations that neglect the focal mechanism term, while when the site effect is accounted for, the inclusion of one variable jointing together into a single class the C and D classes proves to be adequate. It is also interesting to observe that the equations converge in defining a rate of attenuation with distance, deviating significantly from that expected as effect of body wave geometric spreading alone (i.e. with c = -1), but this deviation is obtained with different combinations of c, d and h values. Indeed, when the anelastic attenuation term is included, regression analyses always give c estimates close to -1 (provided that h is not omitted in the functional form), whereas, when it is excluded, the medium anelasticity effect appears incorporated in an increase of the rate of geometrical spreading, reflected by c values stably close to approximately -1.9. The value of the coefficient b controlling the PHA dependence on magnitude is very stable around a value of about 0.45 for all the regression results, and appears very well constrained, as well as high confidence levels characterize the estimates of c and d when determined from regression. In conclusion, for PHA estimates, expressed in cm/s^2, we propose the use of the following functional forms:

a) if focal mechanism and site typology are known, the model with code *c1d1e1f1h1:*

$$\log PHA = 0.787 + 0.478M - 1.092\log\sqrt{R^2 + 10.688^2} - 0.0044\sqrt{R^2 + 10.688^2} + 0.096s + 0.146m \quad (16)$$

b) if focal mechanism is unknown, the model with code *c1d1e1f0h1:*

$$\log PHA = 0.829 + 0.474M - 1.062\log\sqrt{R^2 + 10.772^2} - 0.004\sqrt{R^2 + 10.772^2} + 0.082s \quad (17)$$

c) if site category is unknown, the model with code *c1d1e0f1h1:*

$$\log PHA = 0.881 + 0.479M - 1.107\log\sqrt{R^2 + 10.802^2} - 0.0043\sqrt{R^2 + 10.802^2} + 0.142m \quad (18)$$

d) if both mechanism and site category are unknown, the model with code *c1d1e0f0h1:*

$$\log PHA = 0.907 + 0.474M - 1.074\log\sqrt{R^2 + 10.763^2} - 0.004\sqrt{R^2 + 10.763^2} \quad (19)$$

With regard to the PHV regressions, again the smallest standard deviation (σ = 0.306) was found for the GMPE using the highest number of independent parameters, i.e. with code *c1d0e2f1h1* (see the row 2 in Table S2 for PHV). However, in this case the PHV values of the validation dataset were better predicted by only one simpler functional form excluding the focal mechanism term. In particular, the best equation (*rmsl* = 0.283, E = 0.676) makes use of 2 dummy variables distinguishing site classes B, C and D (code *c1d0e2f0h1*). As for PHA, there is evidence that PHV decrease with distance is influenced by anelastic attenuation. Indeed, if the functional form does not include a specific term for anelastic attenuation, its influence appears as an increase of the rate of the geometric spreading effect (with c tending to values around -1.5 if the h coefficient is present in the functional form). It was also observed that the regression in every functional form for which the term for anelastic attenuation coexisted with the parameter representing the saturation with distance resulted in statistically insignificant d coefficient. Thus in the case of PHV it is inferred that when the distance-dependent saturation effects are taken into account through the "fictitious" depth coefficient, the anelastic attenuation is incorporated into the term of geometric spreading. Among the various functional forms of Table S2 for PHV in cm/s, we propose the use of the following four, depending on the available information:

a) if focal mechanism and site typology are known, the equation with code *c1d0e2f1h1:*

$$\log PHV = -1.082 + 0.692M - 1.614\log\sqrt{R^2 + 12.641^2} + 0.137s_1 + 0.31s_2 + 0.068m \quad (20)$$

b) if focal mechanism is unknown, the equation with code *c1d0e2f0h1:*

$$\log PHV = -1.095 + 0.691M - 1.577\log\sqrt{R^2 + 12.546^2} + 0.128s_1 + 0.306s_2 \quad (21)$$

c) if site category is unknown, the equation with code *c1d0e0f1h1:*

$$\log PHV = -0.869 + 0.661M - 1.542\log\sqrt{R^2 + 11.677^2} + 0.067m \quad (22)$$

d) if both mechanism and site category are unknown, the equation with code *c1d0e0f0h1:*

$$\log PHV = -0.887 + 0.662M - 1.504\log\sqrt{R^2 + 11.506^2} \quad (23)$$

Concerning EDA, the third amplitude-based parameter examined, the regression results resemble those of PHA. In this context, the same considerations reported for PHA regressions can be repeated with regard to the presence of the coefficient *h* in the functional forms providing the best predictive performance. Also, with the inclusion of the anelastic attenuation term together with the saturation factor *h,* the regression analyses result in *c* estimates close to -1, whereas the exclusion of the term for the anelastic attenuation causes an increase of the geometric decay coefficient at values close to -1.95. It is also noteworthy that similarly to PHA and in contrast to PHV, the anelastic attenuation coefficient was found to be statistically significant in the cases where the effect of ground-motion saturation with distance was jointly incorporated in the functional form. In these functional forms the regressions simultaneously estimated negative geometric and anelastic decay coefficients. With regard to the site effect term, it was observed a systematic statistic insignificance in one of the two dummy variables in almost every model of such type, implying that the EDA can be sufficiently modeled by jointing together into a single class both types of soil (C and D). In conclusion, for EDA estimates expressed in cm/s^2, we propose the use of the following equations:

a) if focal mechanism and site typology are known, the equation with code *c1d1e1f1h1:*

$$\log EDA = 0.683 + 0.477M - 1.079\log\sqrt{R^2 + 9.185^2} - 0.0047\sqrt{R^2 + 9.185^2} + 0.109s + 0.14m \quad (24)$$

b) if focal mechanism is unknown, the equation with code *c1d1e1f0h1:*

$$\log EDA = 0.724 + 0.475M - 1.05\log\sqrt{R^2 + 9.379^2} - 0.0043\sqrt{R^2 + 9.379^2} + 0.096s \quad (25)$$

c) if site category is unknown, the equation with code *c1d1e0f1h1:*

$$\log EDA = 0.795 + 0.479M - 1.097\log\sqrt{R^2 + 9.412^2} - 0.0047\sqrt{R^2 + 9.412^2} + 0.133m \quad (26)$$

d) if both mechanism and site category are unknown, the equation with code *c1d1e0f0h1:*

$$\log EDA = 0.82 + 0.476M - 1.065\log\sqrt{R^2 + 9.402^2} - 0.0044\sqrt{R^2 + 9.402^2} \quad (27)$$

5.2 Frequency response-based parameters

With regard to the ASI regressions, once more the smallest standard deviations (σ = 0.288 and 0.290) were found for the GMPEs using the highest number of independent parameters, i.e. with codes *c1d1e2f1h0* and *c1d0e2f1h1* at rows 5 and 6 in Table S3 for ASI. Yet, the same functional forms omitting the style-of-faulting coefficient are able to slightly better predict the ASI values. Like PHV, it is evident from Table S3 that no model exists that jointly incorporates the coefficients for geometric spreading, anelastic attenuation and distance-dependent saturation. However, in this case and contrary to PHV, the parameter for anelastic attenuation was found statistically significant in all the functional forms that included it. However, for every model where the *c*, *d*, and *h* coefficients were simultaneously present, the constant *a* was poorly constrained due to the large uncertainties affecting its estimates. It is also evident in Table S3 for ASI that the constant *a* varies considerably according to the functional form, reflecting the fact that it assumes a dominant role only for extrapolations of ASI predictions far beyond the limits of the explanatory variable values sampled by the training dataset. Moreover, the increase of the geometric decay coefficient at values close to -2.00 when the anelastic attenuation term is excluded from the model and providing that the "fictitious" depth is also incorporated, is comparable to the other two acceleration-based parameters (PHA, EDA). Since the site category is an important parameter that needs to be taken into account, for the ASI estimates expressed in cm/s, we propose the two following models:

a) if focal mechanism and site typology are known, the model with code *c1d0e2f1h1:*

$$\log ASI = 1.303 + 0.601M - 2.073\log\sqrt{R^2 + 17.802^2} + 0.151s_1 + 0.27s_2 + 0.09m \quad (28)$$

b) if focal mechanism is unknown, the model with code *c1d0e2f0h1:*

$$\log ASI = 1.266 + 0.6M - 2.013\log\sqrt{R^2 + 17.723^2} + 0.139s_1 + 0.265s_2 \quad (29)$$

The regressions for VSI gave similar results to those of ASI as regards the functional forms that have the best predictive capacity. In this context, the incorporation of either one or both dummy variables accounting for the local site

conditions appears important in almost all functional forms, while the faulting mechanism is not present at the models with the lowest root mean square of errors estimated for the validation dataset. The inability to simultaneously determine negative *c* and *d* coefficients is also experienced in the VSI regressions. However, in contrast to the ASI findings, it is not the intercept *a* that appears poorly constrained, but it is the estimated values of the anelastic decay coefficient that approach zero and are statistically insignificant. So similarly to the regressions of PHV, it can be assumed that the anelastic attenuation is incorporated into the term of geometric spreading when the distance-dependent saturation effects are also considered in the model. In that case the increase of the *c* coefficient due the fact that it accounts for both effects of wave propagation is at values around -1.45. This is almost equal to the increase found for PHV, the other velocity-based parameter. This also highlights the imbalanced increase of the geometric spreading coefficient when the velocity-based are compared to the acceleration-based parameters. Indeed, in the case of the acceleration-based parameters, the increase that is caused by the incorporation of the anelastic attenuation effect into the geometric decay coefficient is much larger and on the order of -2.00. This implies that the effect of the anelastic attenuation is stronger in the acceleration-based parameters and this is reflected by the fact that contrary to the velocity-based parameters, the anelastic decay coefficient was found to be statistically significant in all three parameters. In conclusion, and considering the importance of the information about the site category, for the VSI estimates expressed in cm, we propose the two following models:

a) if focal mechanism and site typology are known, the model with code *c1d0e2f1h1:*

$$\log VSI = -0.958 + 0.727M - 1.519\log\sqrt{R^2 + 10.124^2} + 0.156s_1 + 0.339s_2 + 0.062m \quad (30)$$

b) if focal mechanism is unknown, the model with code *c1d0e2f0h1:*

$$\log VSI = -0.969 + 0.726M - 1.487\log\sqrt{R^2 + 10.094^2} + 0.148s_1 + 0.336s_2 \quad (31)$$

Concerning T_m, it is clear from Table S3 that site category is an important parameter that needs to be taken into account. The values of the validation dataset were better predicted by the functional form that uses 2 dummy variables distinguishing site classes B, C and D, while additional information on faulting type slightly improves the predictability. It is demonstrated that in contrast to all other ground-motion parameters, normal faulting causes higher T_m values than strike-slip or thrust faulting. Concluding, for T_m we do not recommend the use of a functional form that excludes the site type information, thus we propose the use of the two following models which both use two dummy variables to represent local site conditions:

a) if focal mechanism and site typology are known, the model with code *c1d0e2f1h0:*

$$\ln T_m = -1.317 + 0.319(M - 6) + 0.0045R + 0.288s_1 + 0.446s_2 - 0.186m \quad (32)$$

b) if focal mechanism is unknown, the model with code *c1d0e2f0h0:*

$$\ln T_m = -1.408 + 0.316(M - 6) + 0.0042R + 0.315s_1 + 0.46s_2 \quad (33)$$

5.3 Duration-based parameters

Regarding CAV, the regressions gave results where the first four functional forms presented in Table S4 do not need information for the local site conditions, while the first one that best predicts the validation dataset do not need either the focal mechanism as explanatory variable. The statistic significance of the anelastic attenuation coefficient in all models is also confirmed for CAV, which is another acceleration-based parameter. However, for every functional form which simultaneously incorporates the *c*, *d*, and *h* coefficients, the standard error of the constant *a* was about the same as the value of the coefficient itself, leading to too small t-values to declare statistical significance. We also note that, contrary to the previously discussed acceleration-based parameters, when the coefficient for geometrical spreading also accounts for the effect of anelastic attenuation, it increases much less and is close to -1.20. This implies that the CAV decrease with distance is not so significantly influenced by anelastic attenuation as the rest of acceleration-based parameters. This is probably attributable to the definition of CAV which includes the cumulative effects of ground-motion duration through the consideration of the entire absolute accelerogram. So anelastic attenuation which is known to have a strong effect on ground-motion peak amplitudes, does not have the same influence on the duration-based CAV, as in the case of the aforementioned amplitude-based and frequency response-based parameters that are defined via the recorded acceleration time history. Also, the use of one dummy variable to represent local soil conditions appears adequate seeing that the more complex scheme with the two dummy variables that discriminate stiff from soft soils, give the largest *rmsl* values for the validation dataset. Concluding, for the CAV estimates expressed in cm/s, we propose the following models:

a) if focal mechanism and site typology are known, the model with code *c1d0e1f1h1*:

$$\log CAV = 0.472 + 0.598M - 1.228\log\sqrt{R^2 + 15.382^2} + 0.152s + 0.119m \quad (34)$$

b) if focal mechanism is unknown, the model with code *c1d0e1f0h1*:

$$\log CAV = 0.428 + 0.594M - 1.146\log\sqrt{R^2 + 14.882^2} + 0.144s \quad (35)$$

c) if site category is unknown, the model with code *c1d1e0f1h0*:

$$\log CAV = -0.472 + 0.566M - 0.313\log R - 0.0049R + 0.109m \quad (36)$$

d) if both mechanism and site category are unknown, the model with code *c1d1e0f0h0*:

$$\log CAV = -0.434 + 0.564M - 0.303\log R - 0.0046R \quad (37)$$

Another duration-based parameter for which we developed GMPEs is the I_C. Once more, it can be observed that the functional form with the largest number of independent variables does not have the best performance on predicting the validation dataset (see *c1d1e2f1h0* at row 7 for I_C). The I_C is based on the root mean square acceleration, thus it is also an acceleration-based parameter. Yet again the anelastic attenuation coefficient was found statistically significant in all models where it was simultaneously incorporated with the coefficient for geometrical spreading and the "fictitious" depth. These functional forms however failed to robustly estimate the intercept *a*, whose standard error was on the same order of magnitude with the coefficient itself, yielding very small t-values. It is also evident from Table S4 that the increase of the *c* coefficient in the models where the "fictitious" depth is incorporated and the coefficient for anelastic attenuation is absent, is larger than any ground-motion parameter examined and approximately at values of -2.6. Thus I_C is considerably influenced by anelastic attenuation due to the high increase of the *c* coefficient when it takes into account both path effects. The discrepancy in the increase of the *c* coefficient with CAV which is also a duration-based parameter is attributed to their different definitions, since the latter is calculated from the area enclosed by the absolute accelerogram, while I_C uses the integral sum of squared acceleration time-history. As a final point, for I_C predictions we propose the use of the following models:

a) if focal mechanism and site typology are known, the equation with code *c1d0e1f1h1*:

$$\log I_c = 1.259 + 0.834M - 2.647\log\sqrt{R^2 + 17.408^2} + 0.184s + 0.192m \quad (38)$$

b) if focal mechanism is unknown, the equation with code *c1d0e1f0h1*:

$$\log I_c = 1.177 + 0.831M - 2.509\log\sqrt{R^2 + 16.997^2} + 0.166s \quad (39)$$

c) if site category is unknown, the equation with code *c1d0e0f1h1*:

$$\log I_c = 1.413 + 0.838M - 2.659\log\sqrt{R^2 + 17.367^2} + 0.179m \quad (40)$$

d) if both mechanism and site category are unknown, the equation with code *c1d0e0f0h1*:

$$\log I_c = 1.322 + 0.833M - 2.53\log\sqrt{R^2 + 16.923^2} \quad (41)$$

The SED is the last duration-based parameter for which we developed GMPEs, but contrary to CAV and I_C it is defined through ground velocity time history and not acceleration. From Table S4 it is evident that the information about the local site conditions is essential for a good predictive performance, and in fact the more complex scheme that uses two dummy variables and categorizes rock sites, stiff and soft soils is preferable. On the contrary, the regressions of all functional forms gave statistically insignificant coefficients for the classification of different focal mechanism types, implying that different fault types does not affect SED. Statistically insignificant in all tested models was also the coefficient for anelastic attenuation, when it is combined with the geometric decay coefficient and the "fictitious" depth. This highlights the fact that in all velocity-based parameters, namely PHV, VSI and SED, the anelastic attenuation coefficient was found statistically insignificant, with very small t-values, when *h* is not omitted in the functional form. The increase of the term accounting for geometric spreading when the coefficient for anelastic attenuation is missing, whose coefficient is on the order of -2.4, is much more considerable compared to that of the rest of amplitude- and frequency response-based parameters that are defined via the velocity time history. This is attributed to the definition of SED, where the time integral of the velocity squared and not just velocity is used. In conclusion, and taking into account the insignificance of the focal mechanism, for SED estimates expressed in cm^2/s, we propose the use of the equation with code *c1d1e2f0h0* that uses two dummy variables to represent local site conditions, i.e.:

$$\log SED = -5.699 + 1.46M - 1.236\log R - 0.0071R + 0.361s_1 + 0.778s_2 \quad (42)$$

6. RESIDUAL DISTRIBUTION

Having obtained the most effective functional forms for the various ground-motion parameters examined, it is important to test their residual distribution. In this context we depict in Figure 3 plots of the ordinary residuals as well as the externally studentized residuals for each ground-motion parameter and for the model with the largest number of independent and dummy variables. The ordinary residuals are simply the difference between observed and predicted

values, while the studentized residuals are defined as the difference between observed $Y_{i(obs)}$ and predicted $Y_{i(pre)}$ values when the model is fit using all observations except the i-th and divided by the estimated standard error. The advantage of the studentized residuals over the ordinary residuals is that the quantification is performed in standard deviation units and they measure how far each value is from the fitted model when the entire dataset is used apart from the observation being considered each time; thus they can be straightforwardly detect outliers. This is important, since outliers can affect even simple analyses and they can severely bias a model if they are not detected and removed from further analyses. In this way, observations that have a studentized residual outside the ±2 range are considered statistically significant at the 95% level, while those that have studentized residuals larger than ±3 are considered outliers.

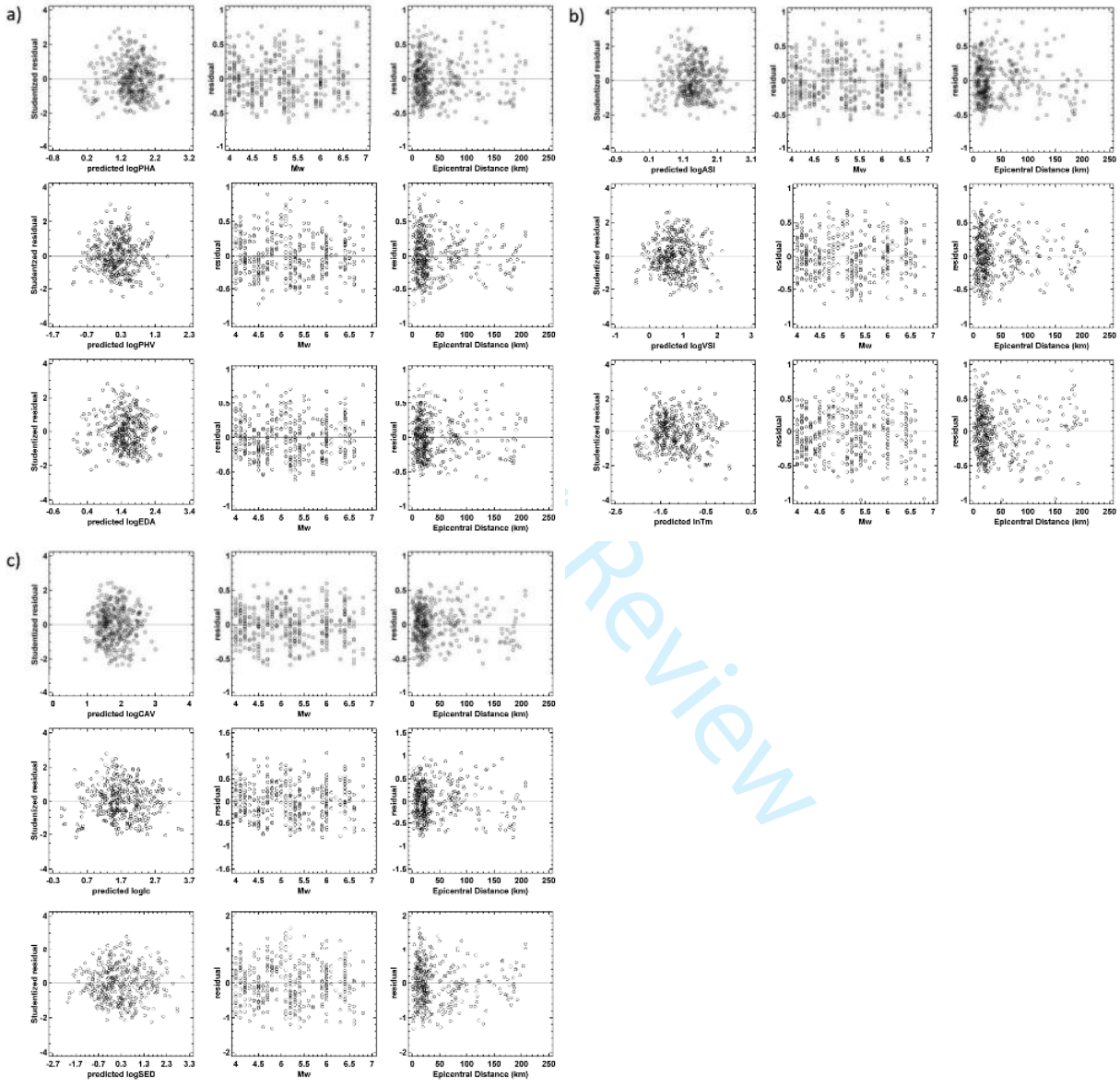


Figure 3. Residual plots of externally studentized residuals (left) and ordinary residuals (right) for a) amplitude-based, b) frequency response-based and c) duration-based parameters. They correspond to the most effective functional forms with the largest number of independent and dummy variables. The ordinary residuals are plotted versus magnitude and epicentral distance, while the studentized residuals versus the predicted ground-motion parameter values.

From Figure 3 it is evident that ordinary residual values do not show any notable trends in all cases and the corresponding means just slightly fluctuate about zero, indicating that the regression results are unbiased. The studentized residuals are also well behaved since no point falls outside the ±3 range confirming that outliers, which may have affected the fits, are absent in all ground-motion parameters. Additionally, about 95% of the studentized residuals are located within the ±2 range, thus they appear to be approximately normally distributed. Finally, a qualitative

assessment of normality of the residuals was performed by means of Probability Density Function (PDF) graphs and normal quantile-quantile (Q-Q) plots (Figure 4). With the PDF graphs we illustrate the shape of the residuals, while with the Q-Q plots we assessed their normality by comparing the actual versus the theoritical percentiles of the normal distribution. These plots should reveal a deployment of the residuals along a straight line without severe deviations. The latter is clearly confirmed for all the examined parameters, supporting the assumption that the error terms are normally distributed.

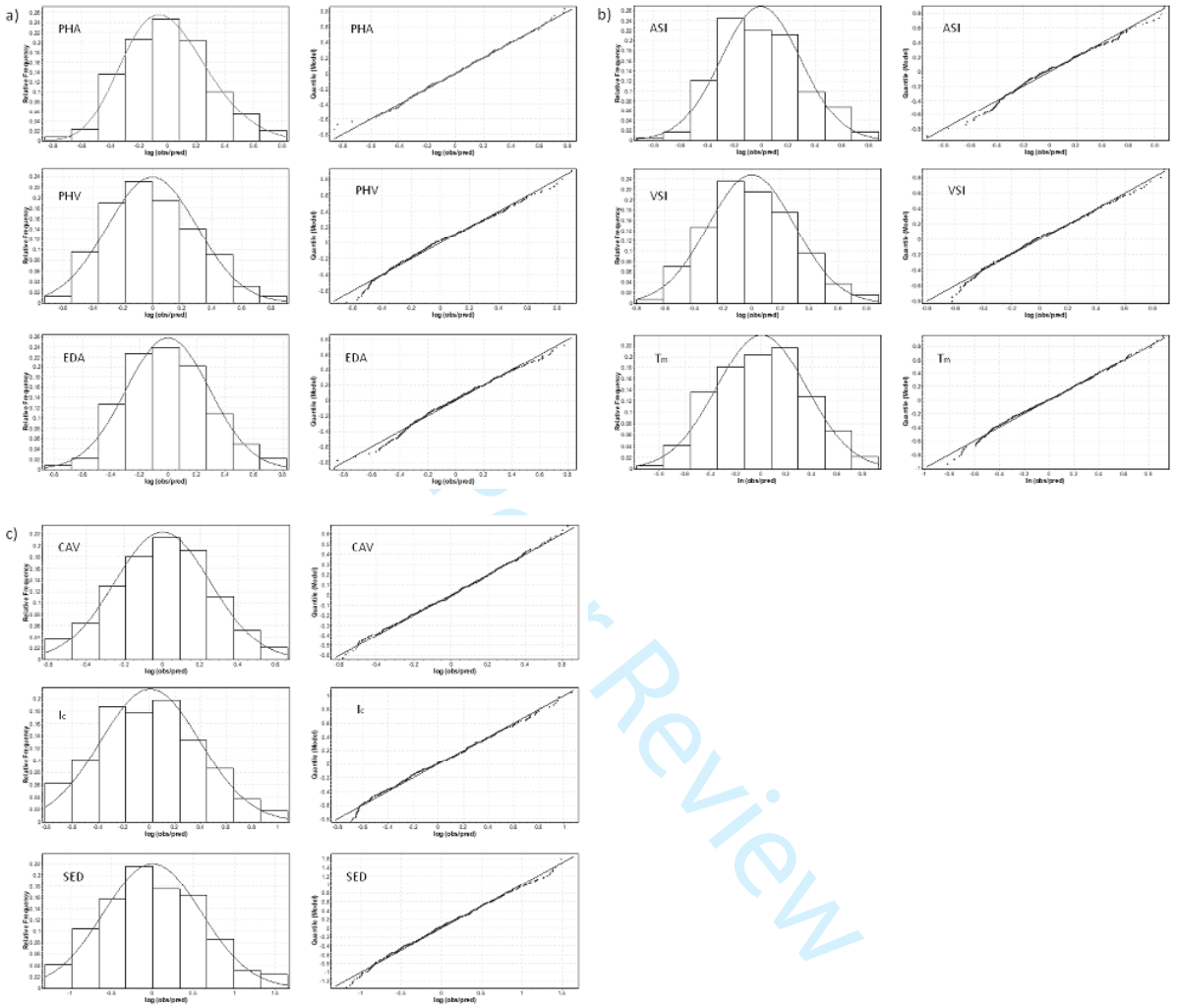
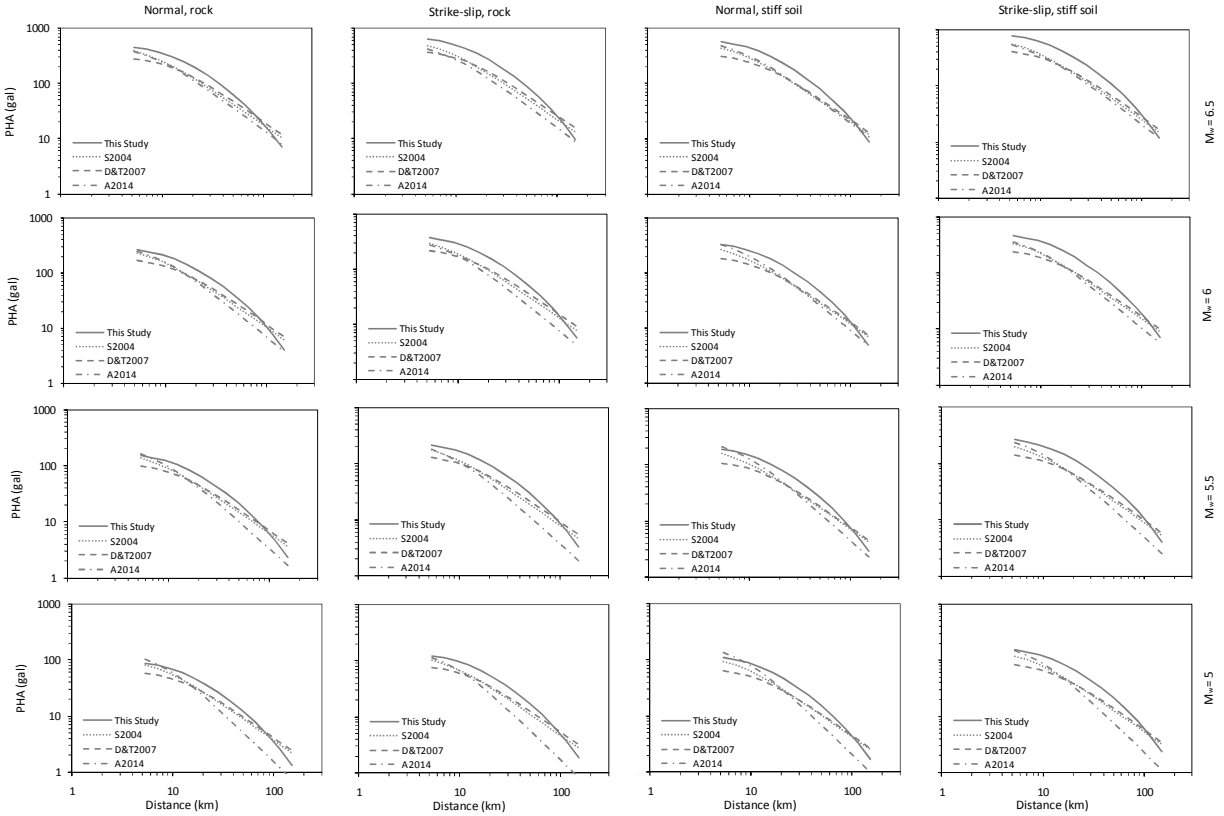


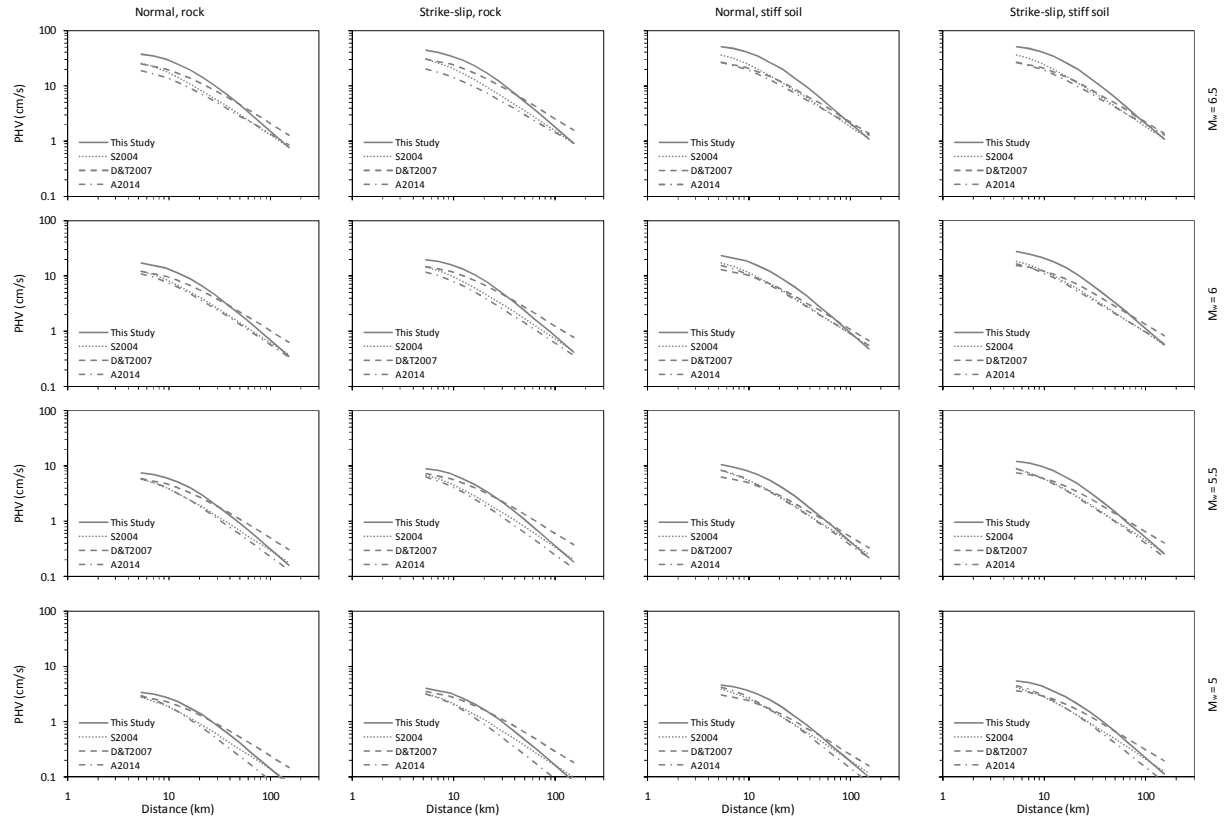
Figure 4. Probability Density Function (PDF) graphs and normal quantile-quantile (Q-Q) plots for a) amplitude-based, b) frequency response-based and c) duration-based parameters. They correspond to the most effective functional forms with the largest number of independent and dummy variables.

7. COMPARISON WITH OTHER GMPEs

For comparison purposes and for the parameters for which a number of GMPEs have already been published, we present in Figure 5 the models of the current study with the largest number of independent and dummy variables with those of different authors. In this context, trellis plots were prepared to illustrate PHA, PHV, CAV and T_m predictions for different combinations of moment magnitude, style-of-faulting and site categories. The Akkar et al. (2014) and the Sandıkkaya and Akkar (2017) models which are depicted in Figure 5 correspond to the formulations using epicentral distances, while regarding the models of Campbell and Bozorgnia (2010) and Du and Wang (2013) that employ Joyner-Boore distance and/or rupture distance one should consider that the use of these distance metrics imply the assumption of shorter distances, especially near the source zone. In the models where additional information was required, we incorporated a Vs_{30} of 500 m/s for stiff soils and 1000 m/s for rocks, a dip angle of the rupture plane equal to 30° and 90° for normal and strike-slip faulting respectively, and a depth to the top of the coseismic rupture plane equal to 0 km. Finally, for Vs_{30} = 500 m/s and Vs_{30} = 1000 m/s we incorporated a depth to the 2.5 km/s shear-wave velocity horizon

equal to 1.28 km and 0.55 km respectively, according to the equations of Abrahamson and Silva (2008) and Campbell and Bozorgnia (2007).





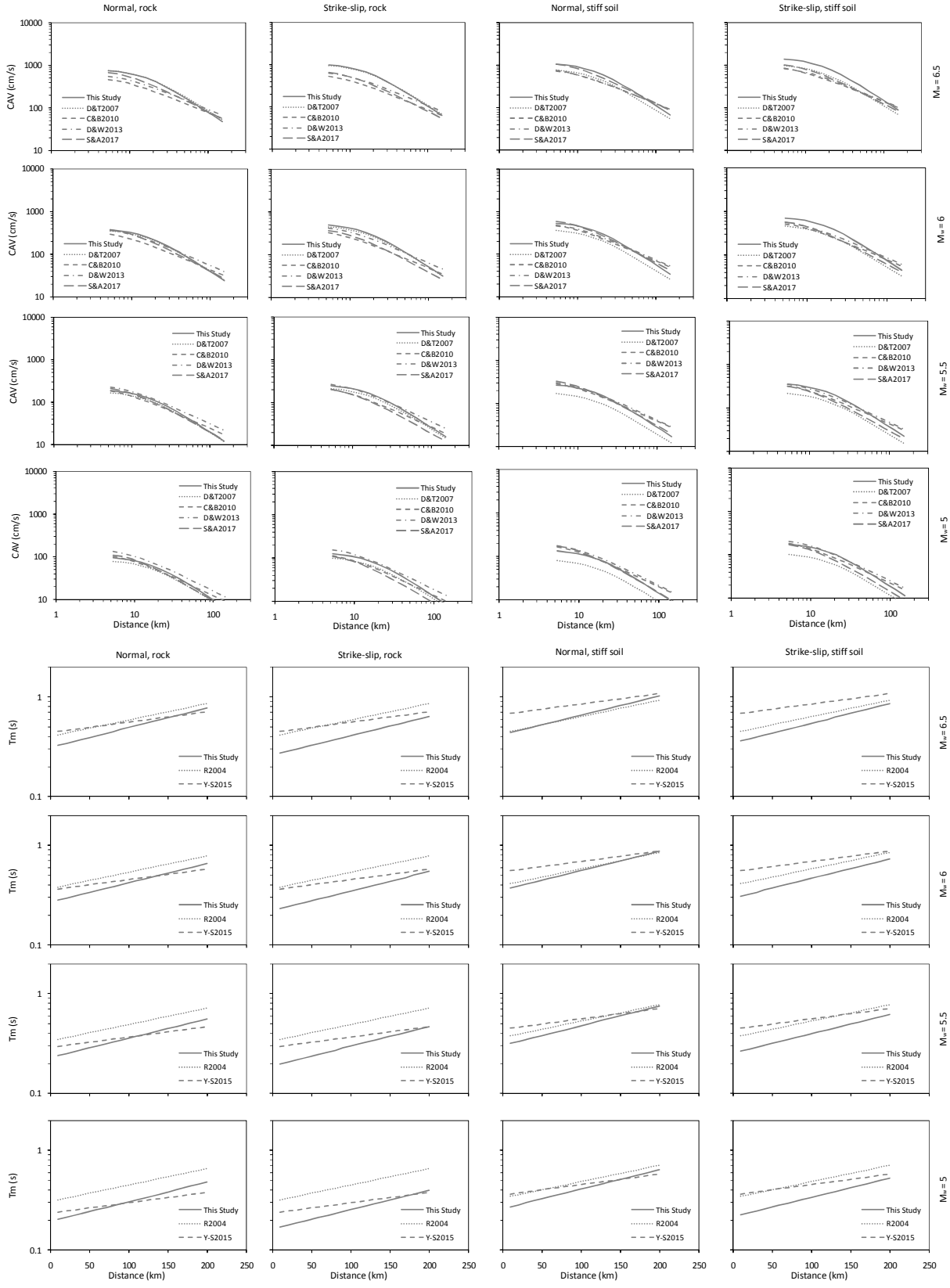


Figure 5. Trellis charts showing predicted PHA, PHV, CAV and T_m for different combinations of moment magnitude, style-of-faulting and site category. Legend abbreviations are as follows: S2004 = Skarlatoudis et al. (2003); D&T2007 =

Danciu and Tselentis (2007); A2014 = Akkar et al. (2014); C&B2010 = Campbell and Bozorgnia (2010); D&W2013 = Du and Wang (2013); S&A2017 = Sandıkkaya and Akkar (2017); R2004 = Rathje et al. (2004), Y-S2015 = Yaghmaei-Sabegh (2015).

It can be observed from Figure 5 that for the peak ground motion parameters, the curves differ in terms of shape as well as absolute values. As far as equations calibrated solely with Greek data is concerned, these deviations mainly reflect the differences of the adopted functional forms along with the number and quality of the employed dataset. Regarding the models of Akkar et al. (2014), the comparison reveals differences attributed to the diversities which characterize the Greek seismotectonic environment. It is true that the Akkar et al. (2014) models are based on a very large strong motion dataset, thus they are well calibrated. However, they were developed from data from the Middle East and from all seismically-active regions that surround the Mediterranean Sea. Thus the possible regional variations that exist among these regions have been smoothed out. In this context, the particularity of the seismotectonic environment of Greece in terms of attenuation has been pointed out in various studies since the late 80's, especially for the Aegean Sea and the upper 20 km of the lithosphere (Hashida et al., 1988; Ligdas et al., 1990; Stavrakakis et al., 1997). Thus it is expected the prevailing geologic and seismotectonic features of Greece to be incorporated in the strong-motion recordings of the earthquakes located within the Greek area. This generates differences in the properties of the acceleration time histories and has impact on the parameters related to seismic shaking. With this in mind, the latest pan-European models are expected to present differences compared to models calibrated on the basis of only Greek data. On the other hand, estimates for CAV are much more comparable at almost all distances. This is also the case for T_m, despite the diversities in stress regime and structural geology among the regions for which the other two curves were derived. The observed discrepancies on peak ground motion prediction can be attributed to differences in the fictitious depth coefficients, the regression approach, and the data distribution.

Special attention was given on peak ground motion, i.e. PHA and PHV, due to the fact that so far these parameters are traditionally used for seismic hazard analysis in Greece and therefore they are associated with increased interest to earthquake and engineering seismologists. For our detailed comparisons, we considered the latest published equations for PHA and PHV, i.e. by Danciu and Tselentis (2007) and by Akkar et al. (2014), which were obtained using a Greek accelerometric database and datasets acquired in the Europe – Middle East region, respectively. Danciu and Tselentis (2007) calibrated equations for PHA and PHV using a dataset including events of magnitude between 4.5 to 6.9, recorded in the Greek area from 1972 to 1999 by stations located at distances up to 136 km. They adopted the following functional form

$$\log Y = a + b \cdot M + c \cdot \log\sqrt{R^2 + h^2} + e_1 \cdot s_1 + f_1 \cdot m_1 \quad (43)$$

where Y is calculated as average between horizontal components, M is the moment magnitude, R the epicentral distance, s_1 a dummy variable accounting for site effect (set to 0, 1 and 2 for site classes B, C and D, respectively) and f_1 a dummy variable representing focal mechanism (set to 0 for normal faults and to 1 otherwise). The differences in Y definition (as arithmetic instead of geometric mean) and in magnitude/distance range of GMPE applicability, deriving from the datasets employed for the regression, complicates the comparison of the performances of these equations (hereafter D&T2007) with the present study best equations (hereafter PSBE). These problems required some care in arranging the comparative tests.

In particular, the comparison was carried out on a portion of the validation dataset consisting only of 63 recordings of events of $M \geq 4.5$ acquired at distances up to 136 km. Furthermore, prediction errors for PSBE and D&T2007 were calculated on the geometric and arithmetic mean of horizontal component peak values, respectively. Although the arithmetic mean is larger than the geometric mean, in the validation subset their differences are quite small (their logarithms differ by about 0.01 on average), thus the relative root mean squares of estimate errors (*rmsl*) can be considered comparable. Their values, calculated on the selected validation subset, were found significantly lower for PSBE (0.302 and 0.332 for PHA and PHV, respectively) in comparison to D&T2007 (0.377 and 0.385 for PHA and PHV, respectively). The left panels of Figure 6 show comparatively the trend of PHA and PHV values predicted by PSBE and D&T2007 as function of distance, for different magnitudes. One can notice that D&T2007 provides lower PHA estimates than PSBE at almost all distances (despite the arithmetic mean should be higher than the geometric one), but differences tend to decrease for lower magnitudes and longer distances. For PHV, estimate differences are smaller, but PSBE estimates are still significantly larger for PHV larger values (i.e. at higher magnitudes and shorter distances). Overall, these trends indicate that the rate of attenuation with distance resulting from D&T2007 appears lower than that derived from PSBE.

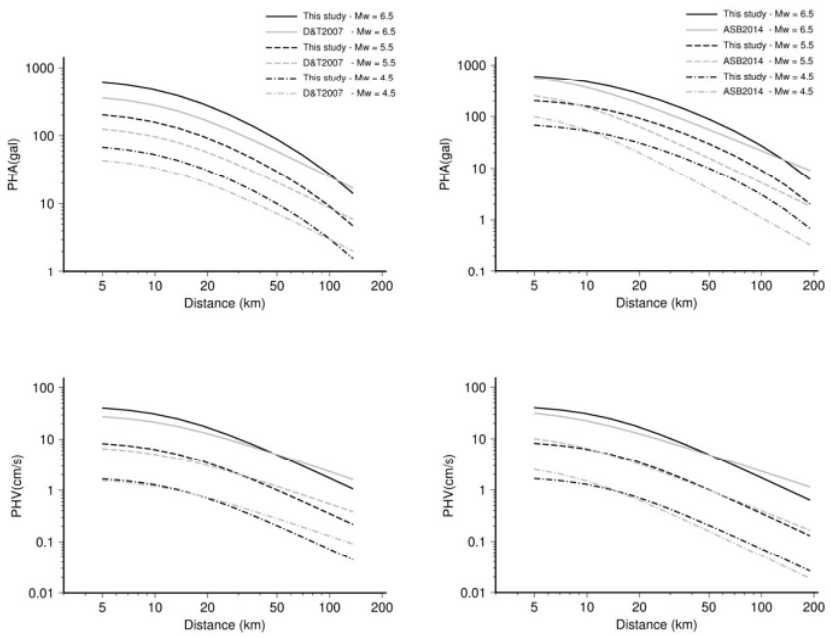


Figure 6. Values predicted for PHA (top) and PHV (bottom) by the best equations obtained in this study (black lines) and by Danciu and Tselentis (2007) (D&T2007; grey lines at left panels) as well as by Akkar et al. (2014) (ASB2014; grey lines at right panels) as function of distance, for different magnitudes.

The upper panels of Figure 7 show the logarithmic errors in predicting the peak values reported in the validation subset, as function of observed peak ground motions. It can be observed that error amount shows an inverse correlation with the real ground motion peak value, so that larger ground motion values tend to be underestimated and smaller values are overestimated. Within this context, the underestimation of larger peak ground motions appears more pronounced for D&T2007, whereas the overestimations of small peak ground motions are relatively smaller. The higher *rmsl* values resulting from D&T2007 predictions in comparison to PSBE likely reflect its worse performance at higher shaking levels, for which errors, being larger, tend to give a major contribution to *rmsl*.

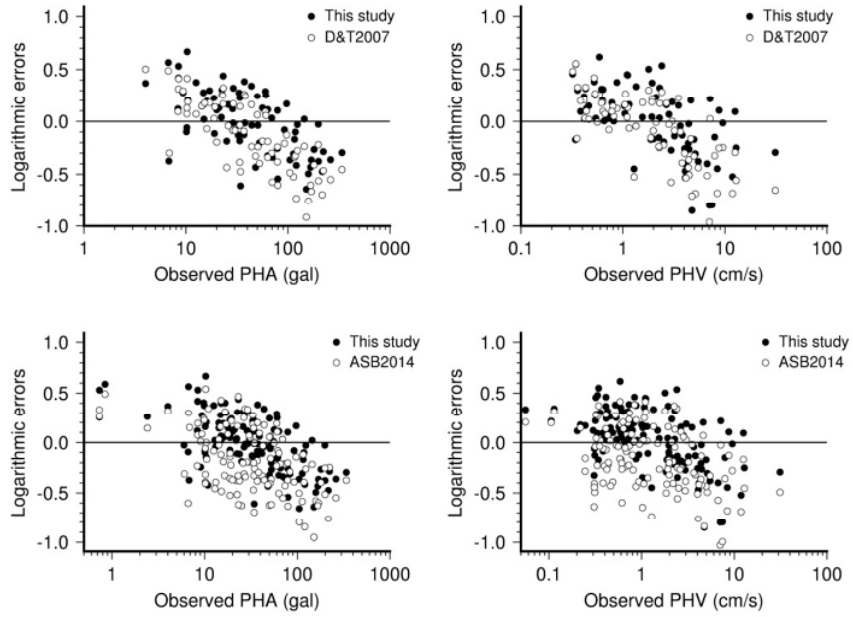


Figure 7. Logarithmic errors affecting the estimates of PHA (left) and PHV (right) of the validation subset, provided by the best equations obtained in this study (black dots) and by Danciu and Tselentis (2007) (D&T2007; open circles at top panels) as well as by Akkar et al. (2014) (ASB2014; open circles at bottom panels). Errors are plotted as function of the peak ground motion actually observed.

Akkar et al. (2014) calculated new GMPEs using an updated database of the Europe – Middle East region including events of magnitude from 4.0 to 7.6, recorded at distances up to 200 km. For regressions, functional forms consistent with the NGA models were used, which include a quadratic magnitude scaling, a magnitude-dependent attenuation (without, however, an anelastic term) and nonlinear site response modeling in terms of Vs_{30}, together with a couple of dummy variables representing the style of faulting. Equations were calculated also with formulations using epicentral distances, which can be directly compared with PSBE. Furthermore, the regression dataset was selected excluding events recorded by single stations, in order to analyze inter-event variability, thus the validation dataset of the present study can be integrally used for comparison with PSBE. Applying these equations (hereafter ASB2014) to the validation dataset, mean Vs_{30} values of different site classes were attributed to each recording station, i.e. 1130 m/s to class B, 560 m/s to class C and 270 m/s to class D. The resulting *rmsl* values for ASB2014 turned out again considerably higher (0.364 for PHA and 0.347 for PHV) in comparison to PSBE (0.278 and 0.295, respectively).

The right panels of Figure 6 display the trend of PHA and PHV decrease with distance predicted by ASB2014 for different magnitude events in comparison to PSBE. In comparison to D&T2007, curves of attenuation with distance show a better agreement with those obtained from PSBE, but a tendency to provide mostly lower estimates than PSBE is still present. This is also reflected by the distribution of estimate errors as function of PHA and PHV (Figure 6; bottom panels), which shows a prevalence of lower estimates, in comparison to PSBE, both for PHA and PHV at any level of ground motion peak values. This implies more pronounced underestimates of higher PHA and PHV values.

8. CONCLUSIONS

In this article we present newly developed equations for the prediction of ten different ground-motion parameters that are well-suited to a great variety of applications and problems in earthquake engineering. The examined measures reflect the characteristics of amplitude, frequency content and duration of the strong ground-motion and the consistency of their development in terms of the same extended database from earthquakes that occurred within the Greek region as well as the same regression approach, ensures that the set of GMPEs can be jointly used without compatibility issues. These models, which predict the geometric mean of the horizontal components of the examined ground-motion parameters, have been developed using an extended database of strong-motion records from Greece, spanning the period from 1973 up to 2015. The derived equations are applicable for shallow crustal earthquakes of magnitude from M_w=4.0 to 6.8 and for epicentral distances up to 200 km. For each ground-motion parameter, we adopted a general functional form including a logarithmic distance term to represent geometric attenuation, a linear term of the same distance standing for anelastic attenuation, a coefficient referred to as a "fictitious" depth measure which is well-known to provide a better fit to the data at short distances, as well as dummy variables to represent distinct categories for site classification and fault mechanism. The adopted functional form does not require a large number of complex input parameters, which in many cases may be either unknown or greatly uncertain, thus it compromises aleatory variability and model sophistication. Simpler equations were also tested, excluding one or both of the coefficients that account for distance-dependent saturation effects and anelastic attenuation, whereas site and mechanism effects were also taken into account or disregarded through the inclusion or not of the corresponding dummy variables. In this context, for all ground-motion parameters we provide alternative optimal equations relative to the available information on the different explanatory variables.

The evaluation of the regression results obtained with different formulae, as well as the comparison of their predictive capabilities, was achieved through a validation dataset including data different from those used in the regressions. These tests highlighted that CAV has the best predictability among the ground-motion parameters that we examined, while SED the lowest. The regression results of the various functional forms that we tested, highlighted some similar patterns either within one ground-motion parameter or between more parameters sharing common definition features. In this context, we investigated the separation of the effects of geometric spreading and anelastic attenuation and we found that in all velocity-based parameters, i.e. PHV, VSI and SED, and contrary to the acceleration-based parameters, the anelastic attenuation coefficient was found statistically insignificant, with very small t-values, when it was combined with the geometric decay coefficient and the "fictitious" depth *h* was not omitted in the functional form. In those cases, where the *c* coefficient accounts for both propagation effects, it should not be considered strictly as a geometric decay coefficient because it simultaneously incorporates the contribution of anelastic attenuation. With this in mind, these regressions highlighted that the increase of the term accounting for geometric decay is much less considerable in the velocity-based (PHV, VSI) than in the acceleration-based parameters (PHA, EDA, ASI, I_C), implying that in the latter case the effect of the anelastic attenuation is stronger. The only exceptions were found to be CAV and SED, where, in the former case, the influence of anelastic attenuation on CAV decrease with distance was not found so significant as on the rest of acceleration-based parameters, while, regarding SED, the increase of the *c* coefficient was found much larger compared to that of the rest of amplitude- and frequency response-based parameters that are defined via the velocity time history. On the contrary, when the anelastic attenuation term is separately included in the functional form,

i.e. for the acceleration-based parameters PHA and EDA, regression analyses always gave c estimates close to -1 (provided that *h* was not omitted). Finally, the regressions of all functional forms for SED, gave statistically insignificant coefficients for the classification of different focal mechanism types, implying that different fault types does not affect their values.

REFERENCES

Abrahamson NA, Silva WJ. Summary of the Abrahamson & Silva NGA ground-motion relations. *Earthq. Spectra* 2008;24(1):67–97.

Akkar S, Bommer JJ. Empirical Prediction Equations for Peak Ground Velocity Derived from Strong-Motion Records from Europe and the Middle East. *Bull. Seismol. Soc. Am.* 2007a;97:511–530.

Akkar S, Bommer JJ (). Prediction of elastic displacement response spectra in Europe and the Middle East. *Earthq. Eng. Struct. Dynam.* 2007b;36:1275–1301.

Akkar S, Bommer JJ. Empirical equations for the prediction of PGA, PGV and spectral accelerations in Europe, the Mediterranean region and the Middle East. *Seismol. Res. Lett.* 2010;81(2):195–206.

Akkar S, Sandıkkaya MA, Bommer JJ. Empirical ground-motion models for point- and extended-source crustal earthquake scenarios in Europe and the Middle East. *Bull. Earthq. Eng.* 2014;12:359–387;doi 10.1007/s10518-013-9461-4.

Ambraseys NN, Douglas J, Sarma SK, Smit PM. Equations for the estimation of strong ground motions from shallow crustal earthquakes using data from Europe and the Middle East: horizontal peak ground acceleration and spectral acceleration. *Bull. Earthq. Eng.* 2005;3(1):1–53.

Bathrellos GD, Skilodimou HD, Chousianitis K, Youssef AM, Pradhan B. Suitability estimation for urban development using multi-hazard assessment map. *Sci. Total Environ.* 2017;575:119–134.

Bourne SJ, Oates SJ, Bommer JJ, Dost B, van Elk J, Doornhof D. A Monte Carlo method for probabilistic hazard assessment of induced seismicity due to conventional natural gas production. *Bull. Seismol. Soc. Am.* 2015;105(3): 1721–1738.

Bradley BA. Site-Specific and Spatially Distributed Ground-Motion Prediction of Acceleration Spectrum Intensity. *Bull. Seismol. Soc. Am.* 2010;100(2):792–801;doi 10.1785/0120090157.

Building Seismic Safety Council (BSSC). National Earthquake Hazards Reduction Program Recommended Provisions and Commentary for Seismic Regulations for New Buildings and Other Structures (FEMA 450), Part 1, Federal Emergency Management Agency (FEMA), Washington, D.C., 2003.

Campbell KW, Bozorgnia Y. Campbell-Bozorgnia NGA ground motion relations for the geometric mean horizontal component of peak and spectral ground motion parameters. *PEER Report No. 2007/02, Pacific Earthquake Engineering Research Center, University of California, Berkeley*;2007.

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

Chousianitis K, Del Gaudio V, Kalogeras I, Ganas A. Predictive model of Arias intensity and Newmark displacement for regional scale evaluation of earthquake-induced landslide hazard in Greece. *Soil Dyn. Earthq. Eng.* 2014;65:11-29.

Chousianitis K, Del Gaudio V, Sabatakakis N, Kavoura K, Drakatos G, Bathrellos GD, Skilodimou HD. Assessment of earthquake-induced landslide hazard in Greece: from arias intensity to spatial distribution of slope resistance demand. *Bull. Seismol. Soc. Am.* 2016;106(1):174–188.

Danciu L, Tselentis G-A. Engineering Ground-Motion Parameters Attenuation Relationships for Greece. *Bull. Seismol. Soc. Am.* 2007;9(1B):162–183.

De Risi R, Goda K. Probabilistic Earthquake-Tsunami Multi-Hazard Analysis: Application to the Tohoku Region, Japan. *Front. Built Environ.* 2016;2;doi 10.3389/fbuil.2016.00025.

Del Gaudio V, Pierri P, Wasowski J. An approach to time-probabilistic evaluation of seismically induced landslide hazard. *Bull. Seismol. Soc. Am.* 2003;93:557-569.

Douglas J, Edwards B. Recent and future developments in earthquake ground motion estimation. *Earth Sci. Rev.* 2016;160:203–219.

Du W, Wang G. A simple ground-motion prediction model for cumulative absolute velocity and model validation. *Earthq. Eng. Struct. Dynam.* 2013;42(8):1189–1202.

Du W, Wang G. Fully probabilistic seismic displacement analysis of spatially distributed slopes using spatially correlated vector intensity measures. *Earthq. Eng. Struct. Dynam.* 2014;43:661–679;doi 10.1002/eqe.2365.

Electrical Power Research Institute (EPRI) (1988). EPRI Report NP-5930, A criterion for determining exceedance of the operating basis earthquake. Electrical Power Research Institute, Palo Alto, California; 1988.

Foulser-Piggott R, Goda K. Ground-Motion Prediction Models for Arias Intensity and Cumulative Absolute Velocity for Japanese Earthquakes Considering Single-Station Sigma and Within-Event Spatial Correlation. *Bull. Seismol. Soc. Am.* 2015;105(4):1903–1918.

Fukushima Y, Tanaka T. A new attenuation relation for peak horizontal acceleration of strong earthquake ground motion in Japan. *Bull. Seismol. Soc. Am.* 1990;80:757–783.

Housner GW. Spectrum intensities of strong-motion earthquakes. *Proceedings of the symposium on Earthquakes and Blast Effects on Structures*. Earthquake Engineering Research Institute 1952:20–36.

Hwang H, Lin CK, Yeh YT, Cheng SN, Chen KC. Attenuation relations of Arias Intensity based on the Chi-Chi Taiwan earthquake data. *Soil Dyn. Earthq. Eng.* 2004;24:509–517.

Joyner WB, Boore DM. Peak horizontal acceleration and velocity from strong motion records including records from the 1979 Imperial Valley, California, earthquake. *Bull. Seismol. Soc. Am.* 1981;81:2011–2038.

Joyner WB, Boore DM. Methods for regression analysis of strong-motion data. *Bull. Seismol. Soc. Am.* 1983;83:469–487.

Kotha SR, Bindi D , Cotton F. Partially non-ergodic region specific GMPE for Europe and Middle-East. *Bull. Earthq. Eng.* 2016;14(4):1245–1263.

Nash JE, Sutcliffe JV. River flow forecasting through conceptual models: Part I, a discussion of principles. *J. Hydrol.* 1970;10:282–290.

Park Y-J, Ang AH-S, Wen YK. Seismic Damage Analysis of Reinforced Concrete Buildings. *J. Struct. Eng. ASCE* 1985;111(4):740–757.

Pitilakis K, Roumelioti Z, Raptakis D, Manakou M, Liakakis K, Anastasiadis A, Pitilakis D. The EUROSEISTEST strong ground motion database and web portal. *Seismol. Res. Lett.* 2013;84(5):796–804.

Rathje EM, Abrahamson NA, Bray JD. Simplified frequency content estimates of earthquake ground motions. *J. Geotech. Geoenviron. Eng.* 1998;124:150–159.

Rathje EM, Faraj F, Russell S, Bray JD. Empirical relationships for frequency content parameters of earthquake ground motions. *Earthq. Spectra* 2004;20(1):119–144.

Sandıkkaya MA, Akkar S. Cumulative absolute velocity, Arias intensity and significant duration predictive models from a pan-European strong-motion dataset. *Bull. Earthq. Eng.* 2017;15(5):1881–1898.

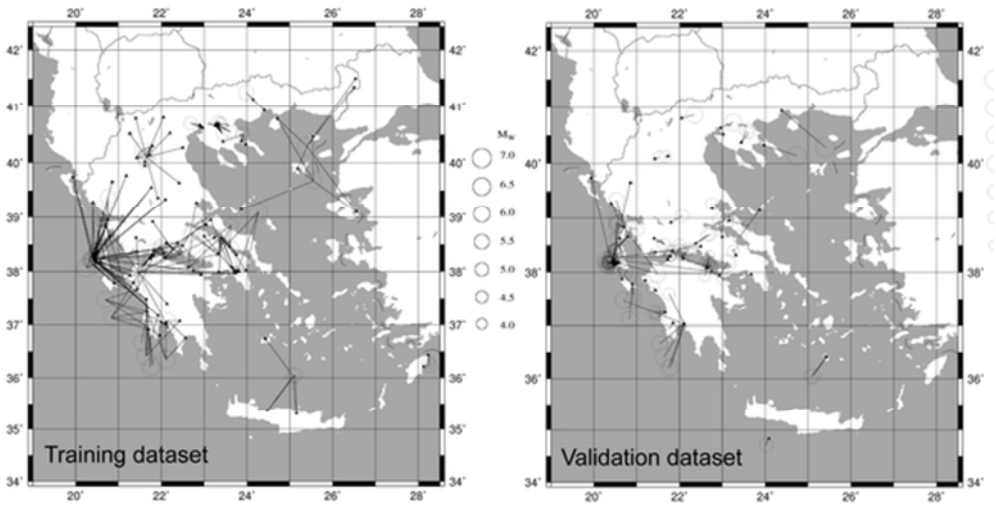
Scherbaum F, Delavaud E, Riggelsen C. Model selection in seismic hazard analysis: An information-theoretic perspective. *Bull. Seismol. Soc. Am.* 2009;99:3234–3247.

Skarlatoudis AA, Papazachos CB, Margaris BN, Theodoulidis N, Papaioannou CH, Kalogeras I, Scordilis EM, Karakostas VG. Empirical peak ground motion predictive relations for shallow earthquakes in Greece. *Bull. Seismol. Soc. Am.* 2003;93:2591–2603.

Theodulidis N, Kalogeras I, Papazachos C, Karastathis V, Margaris B, Papaioannou C, Skarlatoudis A. HEAD 1.0: a unified Hellenic accelerogram database. *Seismol. Res. Lett.* 2004;75:36–45.

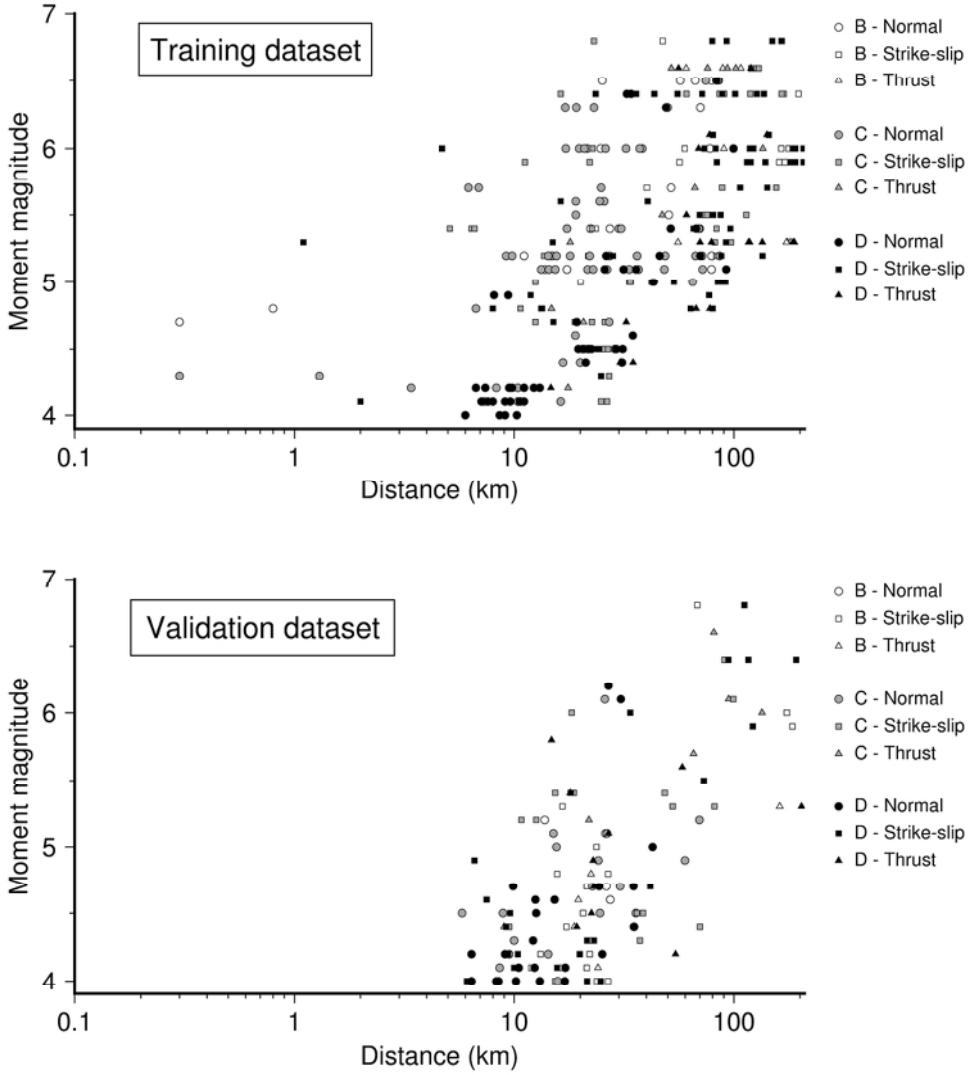
Von Thun, J., L. Roehm, G. Scott, and J. Wilson. Earthquake ground motions for design and analysis of dams. *Earthquake Engineering and Soil Dynamics II—Recent Advances in Ground-Motion Evaluation, Geotechnical Special Publication.* 1988;20:463–481.

Yaghmaei-Sabegh S. New models for frequency content prediction of earthquake records based on Iranian ground-motion data. *J. Seismol.* 2015;19:831–848.



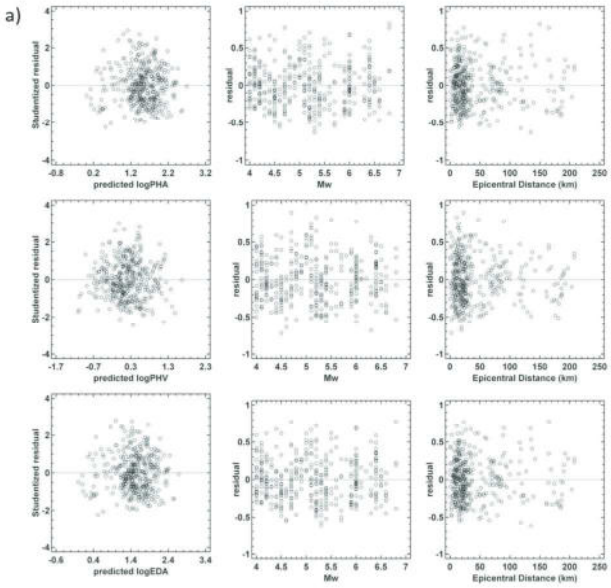
Geographic location of seismic stations (black triangles), event sources (circles sized according to magnitude) and source-station paths (black lines) relative to the seismic recordings included in the training dataset used for the calibration of GMPEs and the validation dataset used to compare the effectiveness of different GMPEs.

50x25mm (300 x 300 DPI)



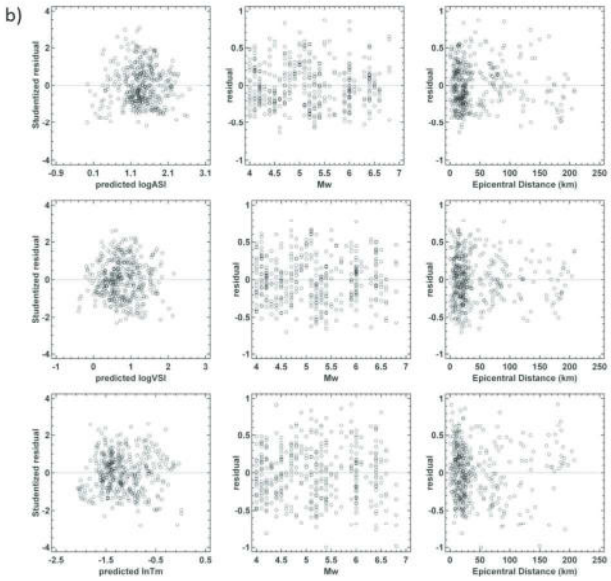
Magnitude of recorded events as function of distance of the recording station for the training and validation datasets. Focal mechanism of the recorded events and site class of the recording stations are represented through different symbols according to the legend.

90x99mm (300 x 300 DPI)



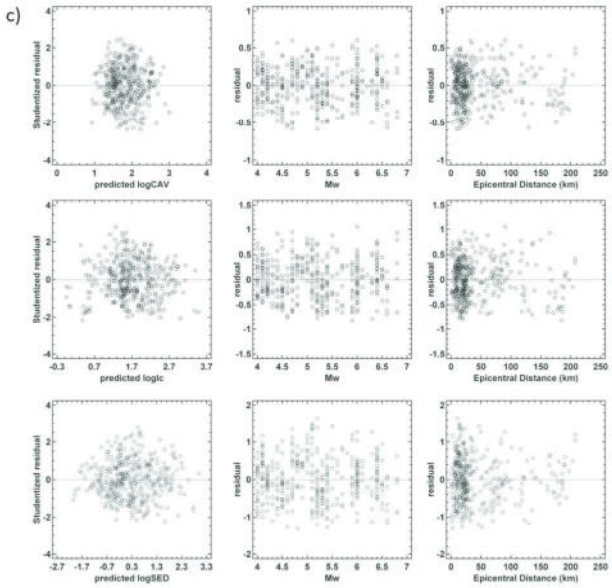
Residual plots of externally studentized residuals (left) and ordinary residuals (right) for a) amplitude-based, b) frequency response-based and c) duration-based parameters. They correspond to the most effective functional forms with the largest number of independent and dummy variables. The ordinary residuals are plotted versus magnitude and epicentral distance, while the studentized residuals versus the predicted ground-motion parameter values.

209x148mm (300 x 300 DPI)



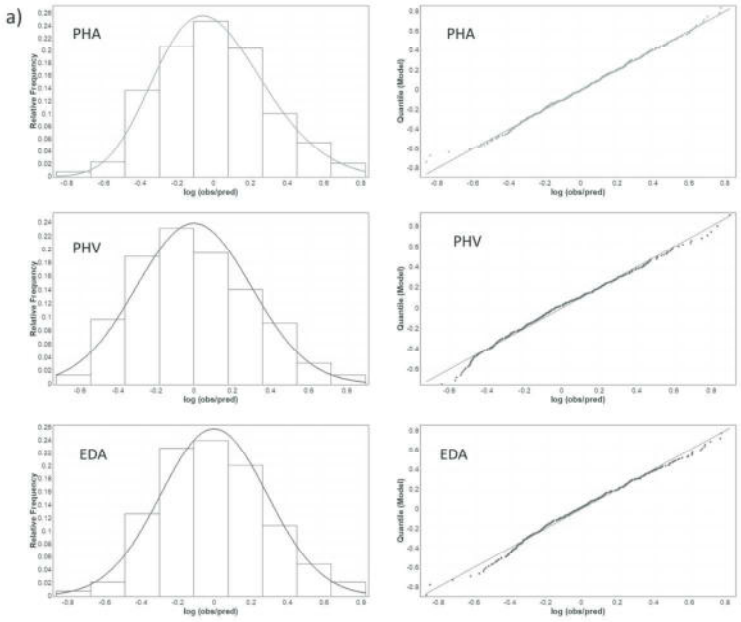
Residual plots of externally studentized residuals (left) and ordinary residuals (right) for a) amplitude-based, b) frequency response-based and c) duration-based parameters. They correspond to the most effective functional forms with the largest number of independent and dummy variables. The ordinary residuals are plotted versus magnitude and epicentral distance, while the studentized residuals versus the predicted ground-motion parameter values.

209x148mm (300 x 300 DPI)



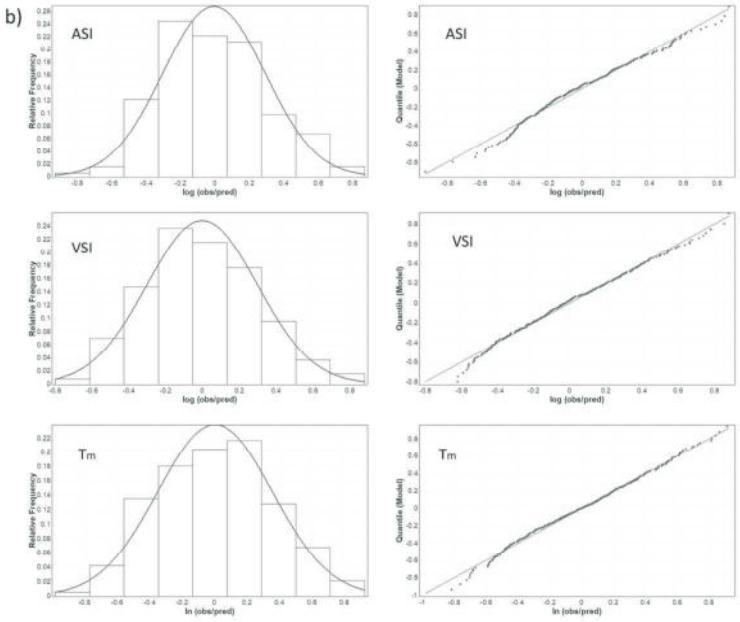
Residual plots of externally studentized residuals (left) and ordinary residuals (right) for a) amplitude-based, b) frequency response-based and c) duration-based parameters. They correspond to the most effective functional forms with the largest number of independent and dummy variables. The ordinary residuals are plotted versus magnitude and epicentral distance, while the studentized residuals versus the predicted ground-motion parameter values.

209x148mm (300 x 300 DPI)



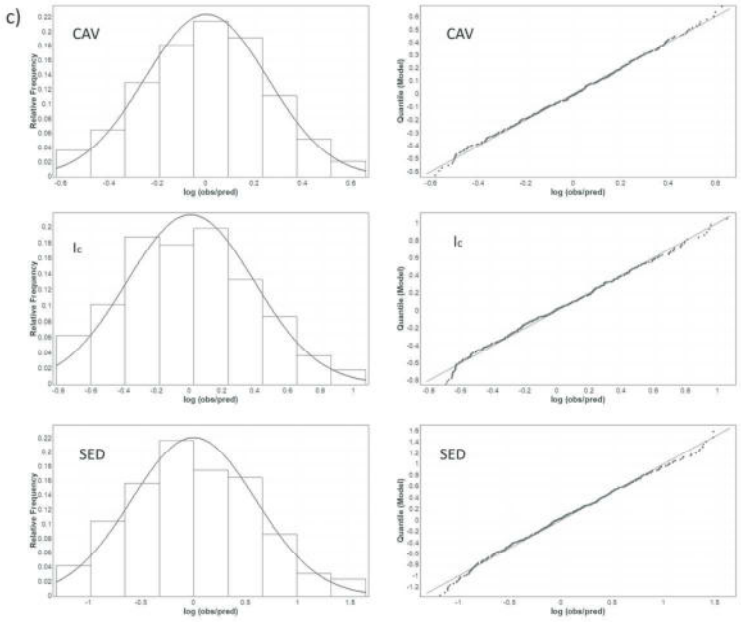
Probability Density Function (PDF) graphs and normal quantile-quantile (Q-Q) plots for a) amplitude-based, b) frequency response-based and c) duration-based parameters. They correspond to the most effective functional forms with the largest number of independent and dummy variables.

209x148mm (300 x 300 DPI)



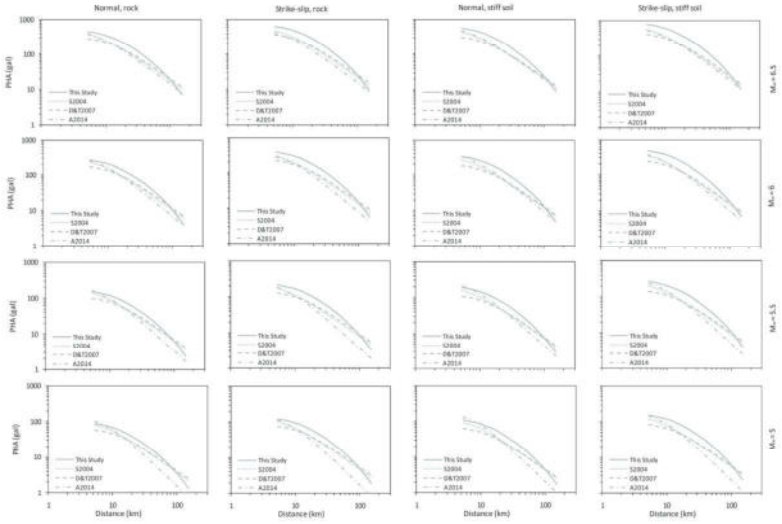
Probability Density Function (PDF) graphs and normal quantile-quantile (Q-Q) plots for a) amplitude-based, b) frequency response-based and c) duration-based parameters. They correspond to the most effective functional forms with the largest number of independent and dummy variables.

209x148mm (300 x 300 DPI)



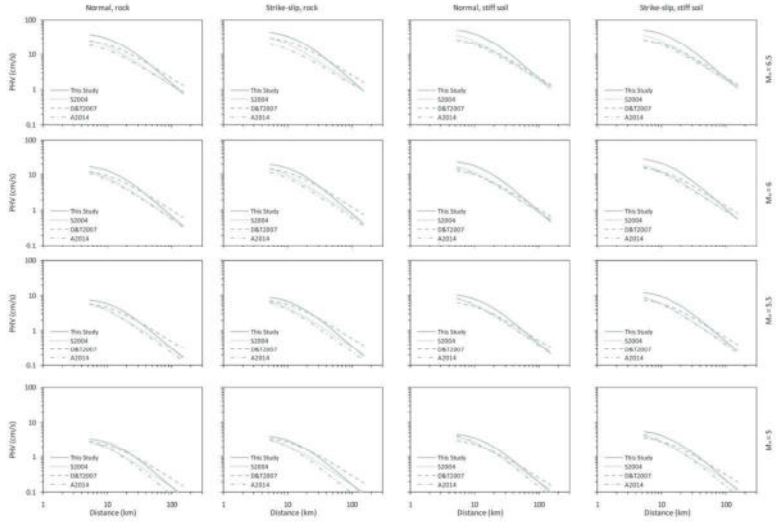
Probability Density Function (PDF) graphs and normal quantile-quantile (Q-Q) plots for a) amplitude-based, b) frequency response-based and c) duration-based parameters. They correspond to the most effective functional forms with the largest number of independent and dummy variables.

209x148mm (300 x 300 DPI)



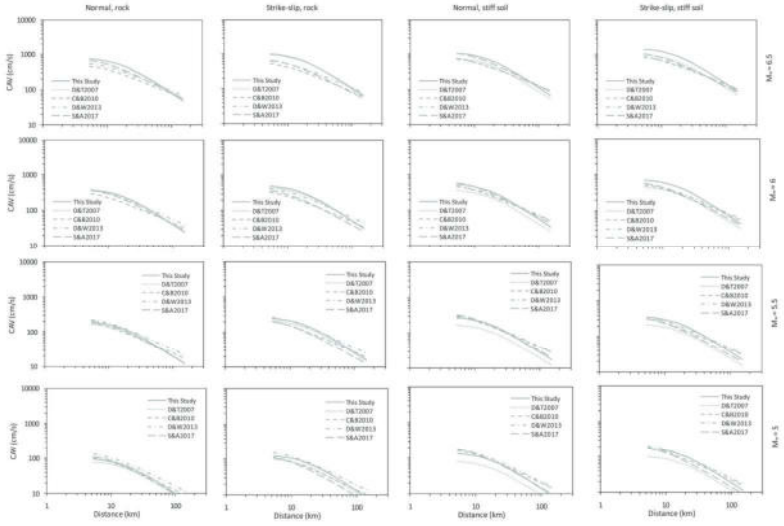
Trellis charts showing predicted PHA, PHV, CAV and Tm for different combinations of moment magnitude, style-of-faulting and site category. Legend abbreviations are as follows: S2004 = Skarlatoudis et al. (2003); D&T2007 = Danciu and Tselentis (2007); A2014 = Akkar et al. (2014); C&B2010 = Campbell and Bozorgnia (2010); D&W2013 = Du and Wang (2013); S&A2017 = Sandıkkaya and Akkar (2017); R2004 = Rathje et al. (2004), Y-S2015 = Yaghmaei-Sabegh (2015).

209x148mm (300 x 300 DPI)



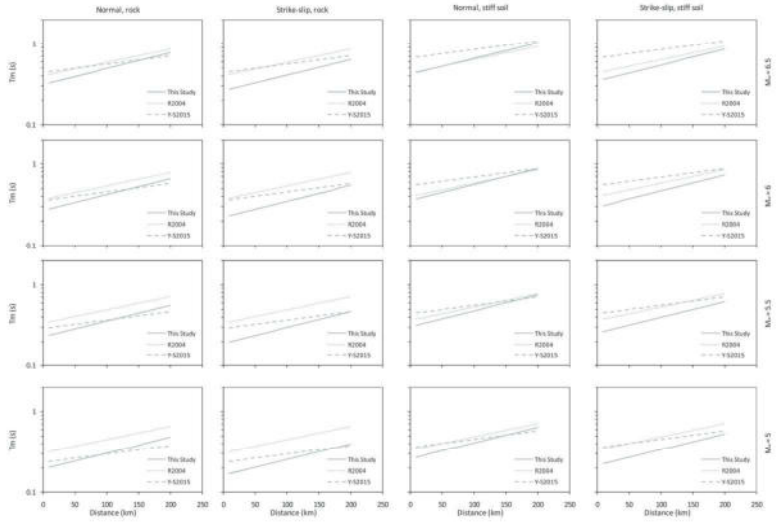
Trellis charts showing predicted PHA, PHV, CAV and Tm for different combinations of moment magnitude, style-of-faulting and site category. Legend abbreviations are as follows: S2004 = Skarlatoudis et al. (2003); D&T2007 = Danciu and Tselentis (2007); A2014 = Akkar et al. (2014); C&B2010 = Campbell and Bozorgnia (2010); D&W2013 = Du and Wang (2013); S&A2017 = Sandıkkaya and Akkar (2017); R2004 = Rathje et al. (2004), Y-S2015 = Yaghmaei-Sabegh (2015).

209x148mm (300 x 300 DPI)



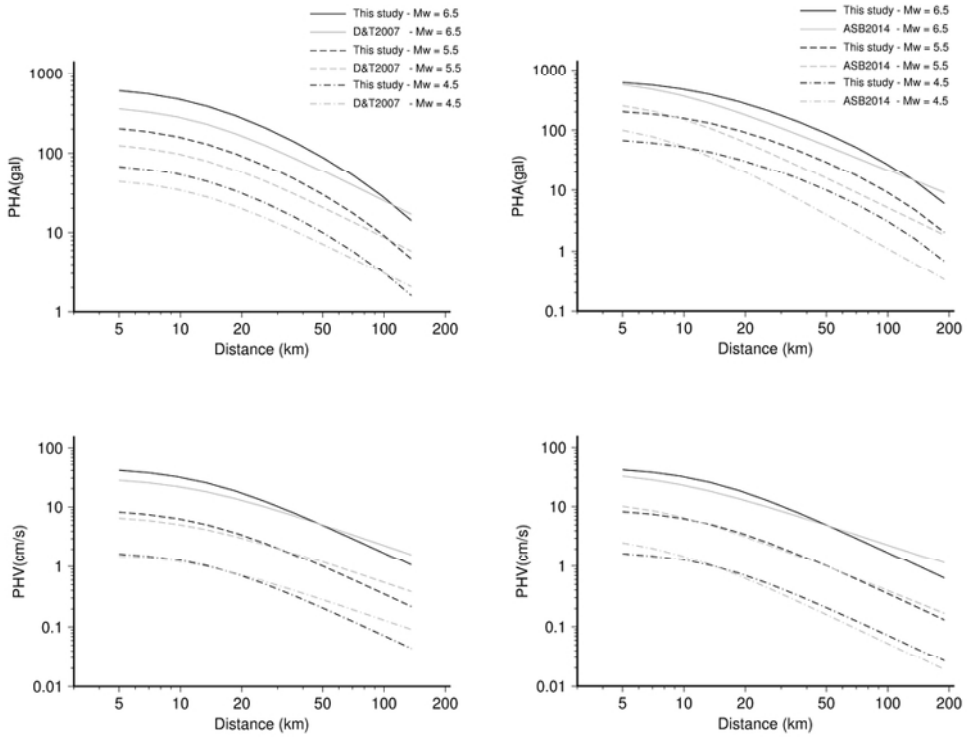
Trellis charts showing predicted PHA, PHV, CAV and Tm for different combinations of moment magnitude, style-of-faulting and site category. Legend abbreviations are as follows: S2004 = Skarlatoudis et al. (2003); D&T2007 = Danciu and Tselentis (2007); A2014 = Akkar et al. (2014); C&B2010 = Campbell and Bozorgnia (2010); D&W2013 = Du and Wang (2013); S&A2017 = Sandıkkaya and Akkar (2017); R2004 = Rathje et al. (2004), Y-S2015 = Yaghmaei-Sabegh (2015).

209x148mm (300 x 300 DPI)



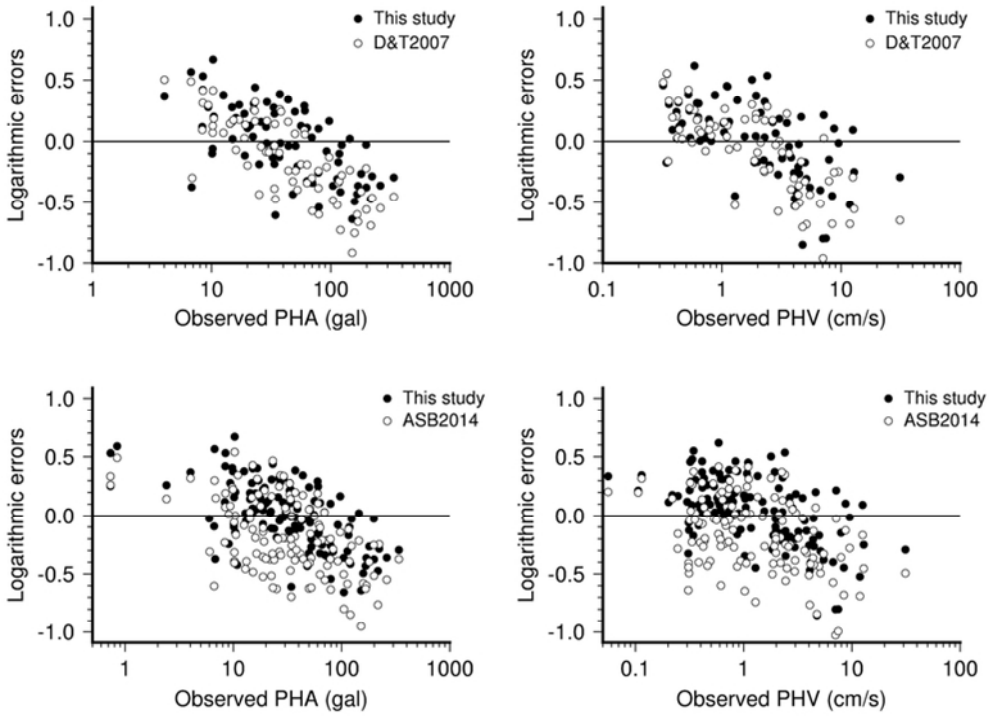
Trellis charts showing predicted PHA, PHV, CAV and Tm for different combinations of moment magnitude, style-of-faulting and site category. Legend abbreviations are as follows: S2004 = Skarlatoudis et al. (2003); D&T2007 = Danciu and Tselentis (2007); A2014 = Akkar et al. (2014); C&B2010 = Campbell and Bozorgnia (2010); D&W2013 = Du and Wang (2013); S&A2017 = Sandıkkaya and Akkar (2017); R2004 = Rathje et al. (2004), Y-S2015 = Yaghmaei-Sabegh (2015).

209x148mm (300 x 300 DPI)



Values predicted for PHA (top) and PHV (bottom) by the best equations obtained in this study (black lines) and by Danciu and Tselentis (2007) (D&T2007; grey lines at left panels) as well as by Akkar et al. (2014) (ASB2014; grey lines at right panels) as function of distance, for different magnitudes.

77x59mm (300 x 300 DPI)



Logarithmic errors affecting the estimates of PHA (left) and PHV (right) of the validation subset, provided by the best equations obtained in this study (black dots) and by Danciu and Tselentis (2007) (D&T2007; open circles at top panels) as well as by Akkar et al. (2014) (ASB2014; open circles at bottom panels). Errors are plotted as function of the peak ground motion actually observed.

66x48mm (300 x 300 DPI)