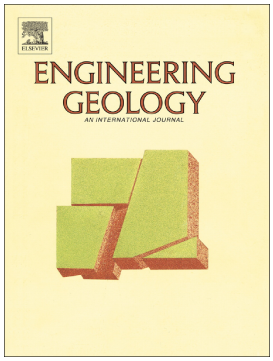

On the applicability of available regression models for estimating Newmark displacements for low to moderate magnitude earthquakes. The case of the Betic Cordillera (S Spain)

José Delgado, Julio Rosa, José A. Peláez, Martín J. Rodríguez-Peces, Jesús Garrido, Meaza Tsigé

PII: S0013-7952(20)30267-2

DOI: https://doi.org/10.1016/j.enggeo.2020.105710

Reference: ENGEO 105710

To appear in: *Engineering Geology*

Received date: 12 February 2020

Revised date: 29 May 2020

Accepted date: 2 June 2020

Please cite this article as: J. Delgado, J. Rosa, J.A. Peláez, et al., On the applicability of available regression models for estimating Newmark displacements for low to moderate magnitude earthquakes. The case of the Betic Cordillera (S Spain), *Engineering Geology* (2020), https://doi.org/10.1016/j.enggeo.2020.105710

On the applicability of available regression models for estimating Newmark displacements for low to moderate magnitude earthquakes. The case of the Betic Cordillera (S Spain)

José Delgado[a,*], Julio Rosa[b], José A. Peláez[c], Martín J. Rodríguez-Peces[d], Jesús Garrido[e], Meaza Tsigé[d]

(a) Dpto. Ciencias de la Tierra y Medio Ambiente, Universidad de Alicante, Ap. Correos 99, 03080 Alicante, Spain. Jose.delgado@ua.es.

(b) Dpto. Física, Ingeniería de Sistemas y Teoría de la Señal, Universidad de Alicante, Ap. Correos 99, 03080 Alicante, Spain. Julio.rosaherranz@ua.es.

(c) Dpto. Física, Campus Las Lagunillas, Universidad de Jaén, 23071 Jaén, Spain. japelaez@ujaen.es.

(d) Dpto. Geodinámica, Estratigrafía y Paleontología, Universidad Complutense de Madrid, C/José Antonio Novais 12, 28040 Madrid, Spain. martinjr@ucm.es, meaza@ucm.es.

(e) Dpto. Ingeniería Civil, Escuela Técnica Superior de Caminos, Canales y Puertos, Campus Fuentenueva, Universidad de Granada, 28071 Granada, Spain, jega@ugr.es.

(*) Corresponding author.

Abstract: Newmark displacement estimation is generally computed using empirical models. These models are estimated from large datasets that mainly comprise moderate-to-high magnitude events (Mw > 6.0). In this work, we study the performance of several of these models to study moderate-to-low magnitude scenarios. For this purpose, data from the Betic Cordillera, S Spain, with magnitudes ranging from Mw 3.5 to 6.3, were used to compare with model predictions. The results show that errors in the estimates depend on the magnitude of events or on the yielding acceleration considered to estimate the displacement. The

availability of an appropriate range of data (magnitude and yielding acceleration), when defining the regression model, may overcome the differences due to specific characteristics of the seismotectonic context of the area where data derives from. The results also show that performance of models including several ground motion predictors is better than those based on a single parameter, regardless of the combination these predictors. Furthermore, regression models with polynomial forms present a better performance than other functions based on the logarithm of these predictors. Finally, new specific models for the Betic Cordillera are proposed, especially suitable for low magnitude events (< 5.0) and low critical accelerations (< 0.1 g), based on simplified polynomial forms of models.

Keywords: Earthquake-induced landslide, Newmark displacement, regression models, low magnitude, Betic cordillera.

1. Introduction

Earthquake frequently induce landslides. The displacement of unstable masses contributes to increasing the damage caused by ground shaking, being the cause of significant damage to lifelines and urban areas (Bird and Bommer, 2004). The fast development our societies experienced over the last decades has dramatically increased the number of elements exposed to risk.

Likely, developing maps to predict the location of areas prone to experience such a phenomenon is the best way to cope with this problem, and then defining an appropriate land use planning. Among the available methodologies for the development of seismic-induced landslide hazard maps, the rigid-block methodology proposed by Newmark (1965), adapted to GIS environments (Jibson et al., 1998, 2000), has become the most widely used. Based on this

approach, potentially unstable materials may slide along a failure surface when the acceleration of ground motion overcomes a threshold value known as critical acceleration of slope (or yield acceleration, k_y), and the block continues sliding up until the velocity of the ground motion becomes zero. Critical acceleration measures the slope material resistance to slide. The final computed displacement, also known as Newmark displacement (Dn), is commonly used as an index for depicting areas prone to experience earthquake-induced instability (Jibson and Michael, 2009; Rodríguez-Peces et al., 2011b, 2014). In general, it is accepted that instabilities occur more frequently when Dn > 1 cm (Bray, 2007).

Newmark displacements have been computed by applying two approaches (Jibson et al., 2000; Jibson, 2011). In the first approach, a set of accelerograms are chosen, and the displacements are calculated by double integration of acceleration each time it exceeds the critical acceleration under consideration. Based on this approach, assumptions on seismic characteristics of the scenario (the magnitude of the event, the source mechanism, range of distances, etc.) have to be done, making more difficult the development of hazard maps for areas where multiple scenarios are expected. The second approach uses regression models which are obtained from the regression of computed displacements for a wide variety of accelerograms, usually extensive databases of accelerograms, against a single ground motion predictor (*scalar models, sensu* Saygili and Rathje, 2008) or a number of them (*vector models, sensu* Saygili and Rathje, 2008). This last approach is the method most frequently applied in practice. Used predictors are usually the peak ground acceleration (*PGA*), the peak ground velocity (*PGV*) and/or the Arias Intensity (*Ia*).

Several authors have proposed different regression models since Newmark (1965) proposed his pioneering method. These models can be divided into two different groups based on how displacements are calculated differentiating between rigid-block and flexible-block methods. In the rigid-block methods, no deformation occurs within the potentially unstable

mass, and its dynamic response is then neglected; likewise, as stated by Newmark (1965). Flexible-block methods consider the dynamic response of non-rigid masses when computing displacements. Flexible-block methods are useful to study deformable earth/waste potential sliding masses (Rathje and Bray, 1999, 2000; Bray and Travasarou, 2007; Bray *et al.*, 2018; Jafarian *et al.*, 2019). On the other hand, methods based on a rigid-block approach provide the best results to study shallow stiff slope failures such as rockfalls (Jibson, 2007, 2011; Saygili and Rathje, 2008; Bray *et al.*, 2018; Yigit, 2020). This is precisely the most frequent type of failure induced by earthquakes (Keefer, 1984). In other cases, nested failure surfaces are considered for computing the rigid-block displacement (Leshchinsky, 2018).

While focusing on the rigid-block regression analysis, some models were obtained from accelerograms recorded worldwide (Jibson, 2007; Saygili and Rathje, 2008; Rathje and Saygili, 2009; Hsieh and Lee, 2011; Table 1) and others from regional or country-based data: China (Jia-Liang *et al.*, 2018), Greece (Chousianitis *et al.*, 2014), Iran (Rajabi *et al.*, 2011), Italy (Romeo, 2000), Turkey (Bozbey and Gundogdu, 2011), among others. Evidently, regional or country models are the most suitable ones for the areas as determined by these authors because used data intrinsically include source and shaking characteristics (source models, anelastic attenuation, etc.) in such areas (Chousianitis *et al.*, 2014).

In a recent study, Du *et al.* (2018) considered up to 14 regression models, including both worldwide and regional-based models, to evaluate uncertainties in predictions of Dn related to the regression models. These authors considered three different scenarios, consisting of strike-slip source events with moment magnitudes (Mw) of 7.5, 6.5 and 5.5, respectively. Then, they calculated Dn with each of the 14 regression models. Because each model was established from different strong ground motion databases and its functional form was also different, it was not surprising that Dn fluctuated depending on the model considered. Such variability showed that the standard deviation of Ln(Dn), with Dn in cm, predictions was higher for low

magnitude events than for moderate to high magnitude scenarios: 0.4-1.0 for Mw 7.5-6.5 *vs.* 1.2-1.7 for Mw 5.5. Such result was interpreted because most regression models were obtained with data from moderate-to-large magnitude (Mw > 6) earthquakes. Table 1 compiles the characteristics of most regression models used by Du *et al.* (2018) and confirms that Hsieh and Lee (2011) did not use data from events with Mw < 5.5 when establishing their widely-used models. Such data were almost absent in the dataset used by Jibson (2007), and they were scarce in those used by Saygili and Rathje (2008).

Recently, there is a renewed interest in the contribution of low-to-moderate earthquakes (Mw 4.0-5.5) in the seismic hazard computation (Nievas *et al.*, 2020). Although seismic hazard is primarily controlled by events with higher magnitudes due to their severity and widespread damage, low-to-moderate events generate smaller damages but much more frequently. Hence, the population has to cope with these damages every year or every few years.

The Betic Cordillera (S Spain) is an area characterized by low-to-moderate magnitude seismicity. Although high magnitude (Mw > 6.5) events have occurred in the past, their return period is very high (centuries). On the contrary, events with Mw in the range 4.0-5.5 are relatively frequent (IGN, 2019). In the last decades, several of these events have occurred (Fig. 1) and triggered multiple shallow failures (disrupted landslides *sensu* Keefer, 1984) which affected road networks (Delgado *et al.*, 2011, 2013, 2015; Alfaro *et al.*, 2012). These infrastructures have proved to be especially vulnerable to seismically-induced landslides (Delgado et al., 2017). Similar results have been observed elsewhere (Martino et al., 2019).

The Spanish government is confronting hazards related to seismic-induced landslides affecting road infrastructure in the Betic Cordillera, through a project (EPILATES) which aims at developing hazard maps along roads specifically for shallow failures based on the rigid-block method. This project has to evaluate both the most severe and most frequent seismic scenarios (Mw < 5.5). The estimation of the Newmark displacement is a critical factor in

developing this project. Unfortunately, there is no explicitly established regression model for estimating Dn for this area. As a consequence, models obtained overseas have been used in the past (Rodríguez-Peces *et al.*, 2011b, 2014). Nevertheless, the high standard deviations in displacement predictions observed by Du *et al.* (2018) for low-to-moderate magnitude events (Mw < 5.5) pose the issue of the utility of current relations in these seismic scenarios. When preparing seismically-induced landslide hazard maps, it is imperative both to identify the threat in each point of the territory, and make a correct assignment to areas with/without risk (i.e., identifying areas where Dn > 1 cm from those where Dn is lower than this threshold value). However, given the high variability in the obtained results with presently available relationships for low-to-moderate magnitude scenarios, uncertainties about the correct assignment of the territory to its category (stable/unstable) are very high in such scenarios.

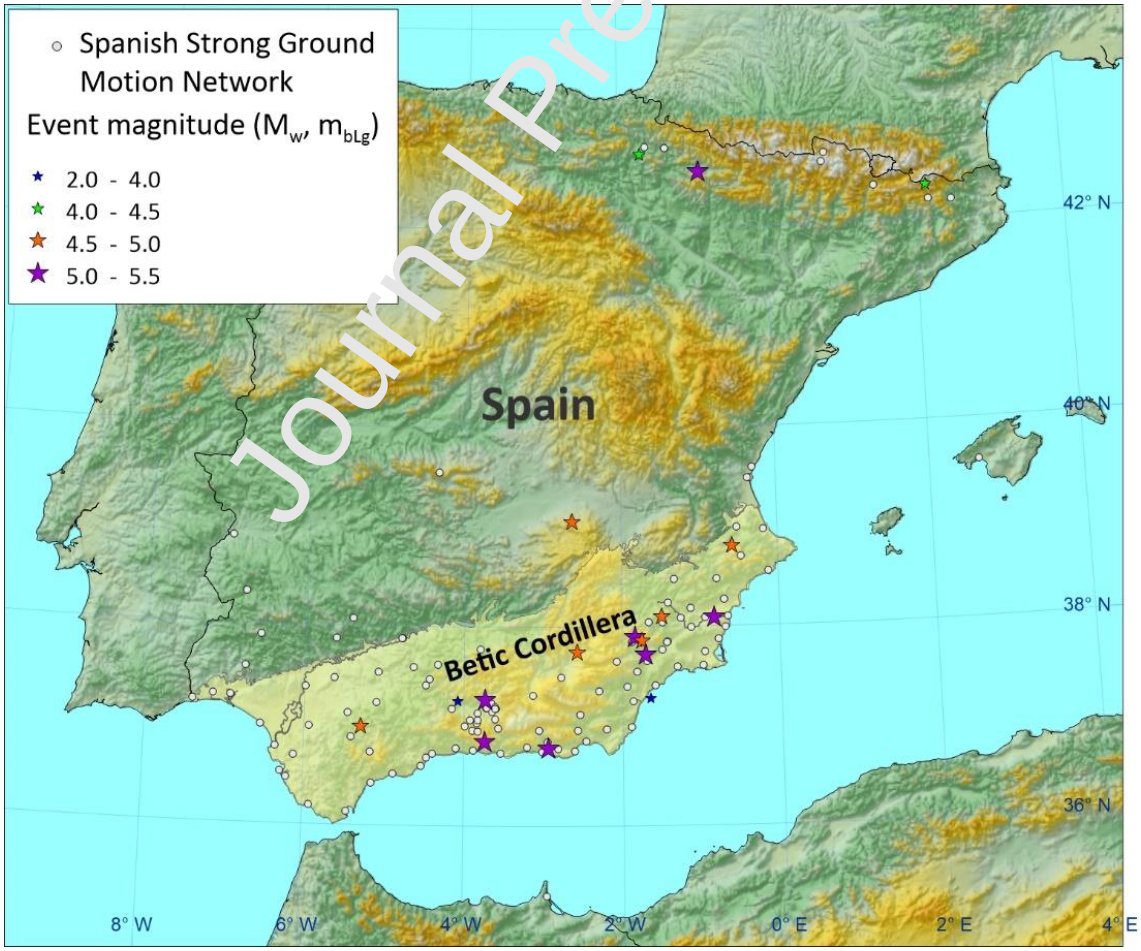

Figure 1. Map showing the location of stations of the Spanish Strong Ground Motion Network operated by the Spanish IGN. The map also shows the epicentral location of earthquakes known to have induced landslides for the period 1919 to 2019.

In this work, we analyzed a set of 13 regression models to evaluate their usefulness for low magnitude scenarios in the Betic Cordillera. For this purpose, recorded accelerograms in this area are used to compute Newmark displacements. Obtained data are subsequently compared with predictions made by the selected models, and the residuals are analyzed to evaluate the performance of each model considered. Finally, we propose regression models for the Betic Cordillera based on the best performing regression forms in the comparative analysis and Newmark displacements computed from accelerograms recorded at different sites of this region.

2. Data

The Spanish Strong Ground Motion Network which has been in operation since 1993 is operated by the Instituto Geográfico Nacional (IGN). At present, it consists of 128 three components stations, located mainly in and around the Betic Cordillera (Fig. 1). In the period 1993-2019, this network has registered up to 895 earthquakes, resulting in a database of 1879, 3-component, strong ground motion records (5637 single component records). The magnitude of events ranges from 1.2 to 6.3.

We have used the IGN database of records (up to October 2019), and among them, we have selected those meeting the following requirements: Mw ≥ 3.5 and *PGA* ≥ 0.02g (low magnitudes, usually in the m_{bLg} scale, were transformed to Mw scale following the relations as proposed in the latest update of the Spanish seismic hazard map; CNIG, 2013). Additionally,

only free-field or ground floor records were considered. Events of magnitude < 3.5 sometimes produced records with *PGA*> 0.02 g, but they were discarded because their duration and frequency content hardly induce landslides. The minimum *PGA* considered is in agreement with minimum acceleration reported as the threshold value to induce shallow disrupted landslides (Jibson and Harp, 2012; Delgado *et al.*, 2015). Taking into account these requirements, the resulting dataset included 87 single-component, horizontal records (Table 2), 10.4% on type A, 63.2% on type B and 26.4 % on type C site conditions (following the categories defined in the EC8 building code), according to data supplied by IGN.

These records were processed as follows:

- Baseline correction and bandpass filtering (0.1-25 Hz). Done through SeismoSignal software (©SeismoSoft, 2016).
- For each record, displacements were computed for both positive and negative polarities considering the largest value for the analysis (Saygili and Rathje, 2008).
- Displacements computed from orthogonal components recorded at the same station were considered as separate data for the analysis (Saygili and Rathje, 2008).
- Rigid block displacements were computed for k_y values of 0.02, 0.03, 0.04, 0.05, 0.06, 0.08, 0.1, 0.125, 0.15, 0.2, 0.25 and 0.3 g.

Specific software was coded for automatically computing displacements. Critical acceleration values were chosen taking into consideration that values greater than 0.1 g are not frequently recorded during earthquakes with Mw < 5.0.

After processing, the resulting database is characterized by *PGA* values lying in a range between 0.02 - 0.39 g, *PGV* values in a range between 0.35 - 33.15 cm/s, and *Ia* horizontal values in a range between 0.07 - 52.80 cm/s. A total of 242 Dn values were computed from this database of accelerograms. Displacements vary from almost zero (10^{-6}) up to 26 cm (Fig. 2).

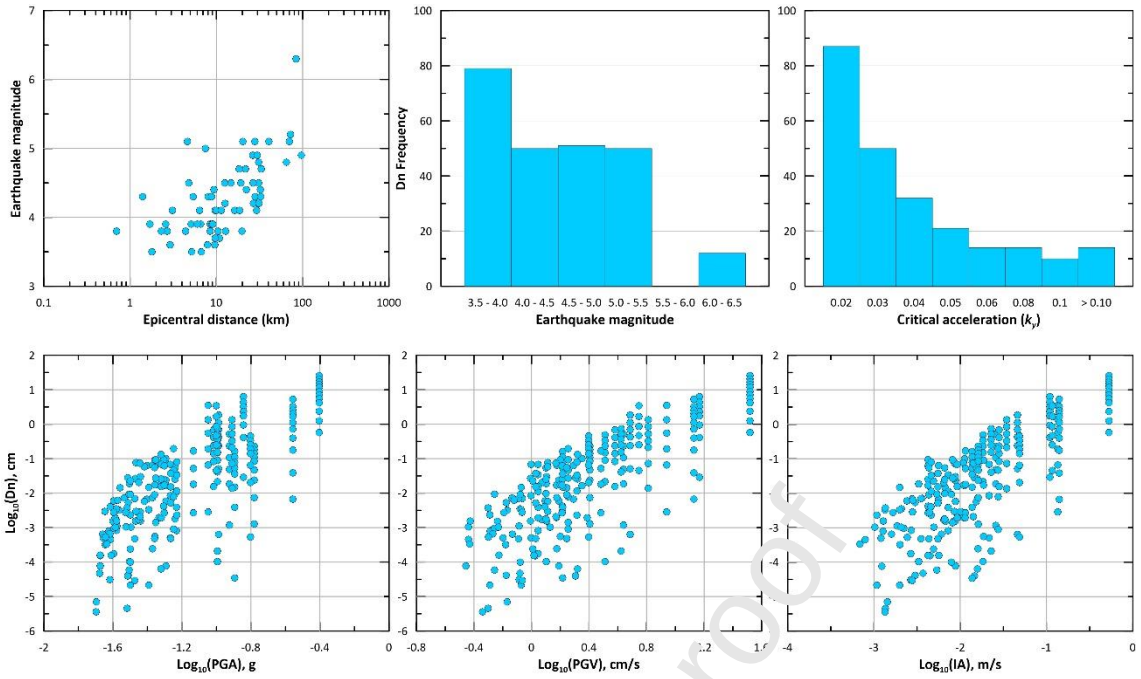

Figure 2. Distribution of seismic data used in our analysis as a function of magnitude and epicentral distance, and ground motion parameters

3. Analysis

3.1. Regression models

We have selected 12 regression models in our research (Table 1). These models present a wide variety of functional forms to compute Newmark displacements, including three different parameters for measuring the severity of ground motion when estimating Dn: *PGA*, *PGV* and *Ia*. For some models, earthquake magnitude together with the *PGA* add up as ground motion predictors.

Many of these models (9) were obtained from worldwide data and, consequently, include data from different geodynamic contexts. These models were obtained from extensive databases; however, few data from earthquakes in the range of magnitudes of interest for this study (Mw < 5.5) were used. Additionally, most of the critical accelerations considered in

establishing these models were high (0.2 g and higher), greater than accelerations usually recorded during low-to-moderate magnitude scenarios.

The remaining models were established from regional or country-based data (China, Greece, Iran and Italy). Including such models help us to monitor their usefulness, given the similarities in the geodynamic context (Italy) or in the range of magnitudes from the data used in establishing those models (Greece and Iran). The China model (Jia-Liang *et al.*, 2018) was used as an example of a model obtained with data from a single event of very different magnitude (Mw 7.9) in a geodynamic context that is somewhat different to that existing in the Betic Cordillera.

Two of these models (J07_1 and SR08_1) use the *PGA* as a predictor of ground motion severity. Their functional forms are very different from each other. Other two models use the magnitude in addition to *PGA* (J07_2 and RS09). Once again, their respective functional forms are very different. Up to five models consider the *Ia* to quantify ground motion severity. Models J07_3 and Ra11 share the functional form for the equation while the others (HL11, CH14 and JL18) use different combinations of *Ia* and k_y to define the model. Other three models combine *PGA* and *Ia* to characterize the ground shaking level (R00, J07_4 and SR08_2). Models R00 and J07_4 have the same functional form, different from the SR08_2 model. Finally, there is only a single model that uses *PGA* and *PGV* to predict Dn (SR08_3).

In our analysis, we analyze the residuals obtained by using the regression models to predict Dn in the Betic Cordillera and compare them with observations computed as described in the previous section. *PGA*, *PGV* and *Ia* values compiled for the 87 records considered in this study were used to compute Dn according to each model. Residuals are defined as the difference between "$\log_{10}(Dn_{observed}) - \log_{10}(Dn_{predicted})$".

3.2. Residuals versus ground motion predictor

Figures 3 and 4 present plots with the residuals obtained from each model; Table 3 shows some statistics of them. Two parameters have been considered to control the performance of each model: the root mean square of errors (RMSE) and the efficiency coefficient (E). Chousianitis *et al.* (2014) already used the aforementioned coefficient to evaluate the quality of their models . Numerically, it is equal to:

$$E = 1 - \frac{\sum(y_i - y_i^*)^2}{\sum(y_i - \bar{y}_i)^2}$$

In this formula, y_i is the experimental observations (Dn computed from accelerograms), $\bar{y}_i$ the mean value of observations, and y_i^* the corresponding value predicted by the model. This coefficient quantifies the better functioning of the model regarding the average as the predictor of the variable. It is a negative value if its performance is worse than average, and a positive value if better than average (Chousianitis *et al.*, 2014).

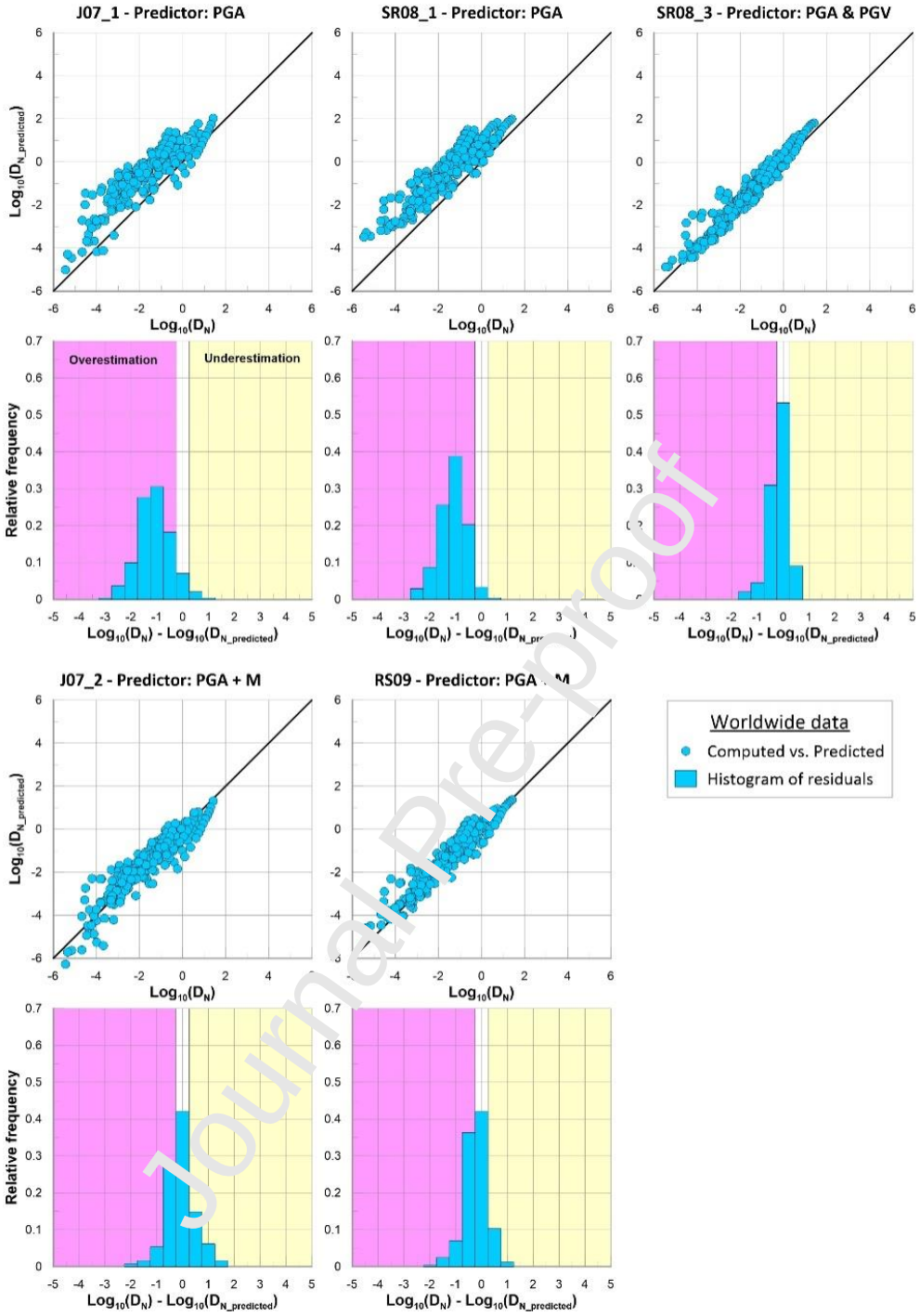

Figure 3. Residuals (and their relative frequency) obtained with regression models that use *PGA*, *PGA* and *PGV*, or *PGA* and magnitude as ground motion predictors. See Table 1 for more data about these models.

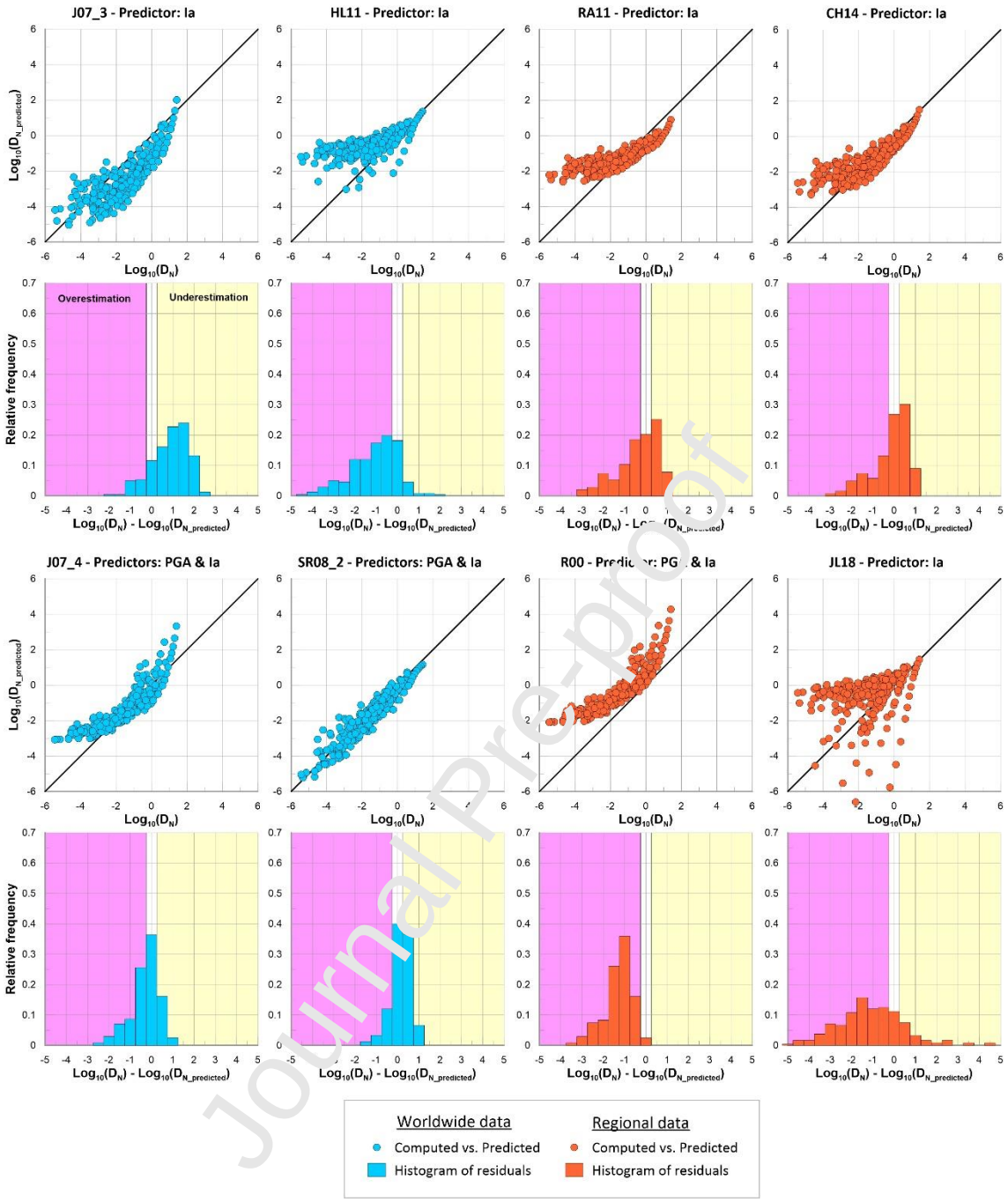

Figure 4. Residuals (and their relative frequency) obtained with regression models that use *Ia* or *PGA* and *Ia* as ground motion predictors. See Table 1 for more data about these models.

The scalar models, which are those based on the PGA as an estimator of ground motion (J07_1 and SR08_1), show very similar results. They greatly overestimate Dn (Fig. 3), and the corresponding statistics show a low performance for these models: RMSE is high and E

is low (Table 3). Because equations of both models are quite different from each other (Table 1), such low performance cannot be attributed to deficiencies in the form of the model to reflect the complexity of ground motion. When magnitude is added to these models (J07_2 and RS09 models), the results show a noticeable increase in their performance. Residuals consequently decrease less than a half, and efficiency increase above 0.8 (previous values are around 0.2).

The results obtained with models based on *Ia* show very high variability, from moderate to deficient performance (Fig. 4 and Table 3). Models J07_3 and RA11 share the functional form of the model (they include terms with $\log_{10}(Ia)$ and $\log_{10}(k_y)$), but the latter model gives slightly better statistics (lower RMSE and higher E, Table 3). Distribution of residuals (Fig. 4) shows that J07_3 underestimates Dn, while RA11 shows a bipolar behavior: an overestimation for low Dn values and an underestimation for high Dn values.

Models HL11 and JL18 share a functional form of the model. Additionally, they differ from the previous ones including a term with k_y (instead of its logarithm, as in J07_3 and RA11) and another term with k_y multiplying the $\log_{10}(Ia)$. The statistics of these models are the worst obtained in the analysis (Table 3). They tend to overestimate Dn (Fig. 4), although the behavior of the JL18 model is a bit more erratic, showing a high dispersion of residuals, so we do not consider this model in our research given its deficient performance.

The best performance for the *Ia*-based models is found in the CH14 model. It reduces RMSE to at least 50% concerning other *Ia*-based models, and increases the efficiency coefficient a minimum of 50%, although a trend to overestimate displacements when Dn is very low is recognizable (Fig. 4).

Regarding vector models, *PGA-Ia* models show high variability (Table 3). The R00 model shows poor performance, overestimating Dn in all ranges of values, and such overestimation is significant (one $\log_{10}$ unit, similar to that observed for J07_1 and SR08_1

models). The other two models of this type show a better performance, especially the SR08_2 one.

Finally, there is only one model that uses both *PGA* and *PGV* as ground motion predictors. The statistics are the strongest of all (Table 3) with small residuals RMSE values and a very high efficient value. The residual distribution reveals a normal distribution around zero with very narrow tails (Fig. 3).

3.3. Residuals versus magnitude of events

The next step is to study the performance of each model depending on the magnitude of the event. We believe this is of particular significance, given that most common seismic scenarios correspond to low magnitude events (Mw ≤ 5.0). For this purpose, computed displacements were grouped by magnitude, and RMSE was computed separately for each range of magnitude. Groups were of 0.5 units of magnitude (starting at 3.5).

Figure 5 shows the results of this analysis and it clearly reflects that RMSE is not evenly distributed with magnitude for all models. Starting with the scalar models based on worldwide data (*PGA*: J07_1 & SR08-1; *Ia*: J07_3 & HL11), they show the same relative behavior: RMSE is maximum for data with very low magnitudes; furthermore, it tends to be gradually reduced as magnitude becomes higher. RMSE is greater than 1 when M ≤ 5.0 (5.5 for some models) and goes below such threshold only when M > 6.0. On the contrary, regional scalar models (including data on low magnitude events in their determination) do not show such magnitude to be dependent on RMSE behavior. It is therefore essential to provide data of all magnitudes of interest when developing a regression model.

Results obtained with scalar models based in *PGA,* including the magnitude (models J07_2 and RS09; Fig. 5) seem to be very useful. They were acquired from the same dataset as

the scalar models, that is, without data from earthquakes with M < 5.0, though they do not show such clear magnitude-dependent behavior.

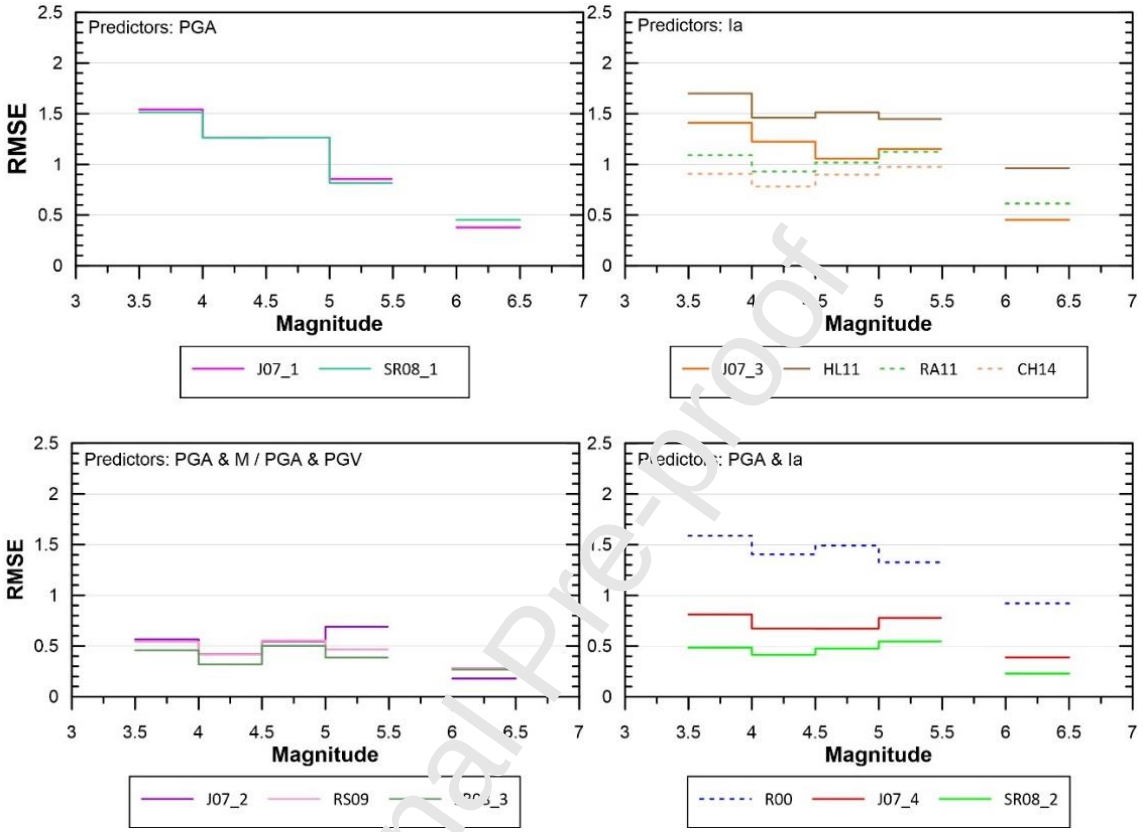

Figure 5. Variation of RMSE as a function of magnitude. Continuous lines depict models based on worldwide data. Dashed lines depict models based on regional data. See Table 1 for specific characteristics of each model.

Vector models (*PGA-Ia* or *PGA-PGV*) show a better performance than scalar models, and no evident relation exists among magnitude and RMSE. The findings demonstrate the poor efficiency of the R00 model (from Italy) to match the Betic Cordillera data once again.

3.4. Residuals versus critical acceleration

The final step in our analysis consisted of verifying the response of the models for low values of the critical acceleration. When they were established, although some models included low values of critical acceleration (Table 1), most data corresponded to critical accelerations that are out of the range of recorded accelerations for moderate to low magnitudes (k_y> 0.1 g). As in the previous section, the residuals have been grouped as a function of the critical acceleration considered when computing displacements. Because there are few displacements for critical accelerations greater than 0.1 g (only 14 values), they were grouped into a single category. The results are shown in Figure 6.

Scalar models show different behavior depending on the parameter considered quantifying ground motion (*PGA* vs. *Ia*). Those based on *PGA* show that for very low critical acceleration values *RMSE* is higher, and it stabilizes when k_y > 0.05 g. By comparison, *Ia*-based models display more complex behavior so that no specific pattern may be identified to this respect. Nevertheless, regional-based data models typically perform better than worldwide models.

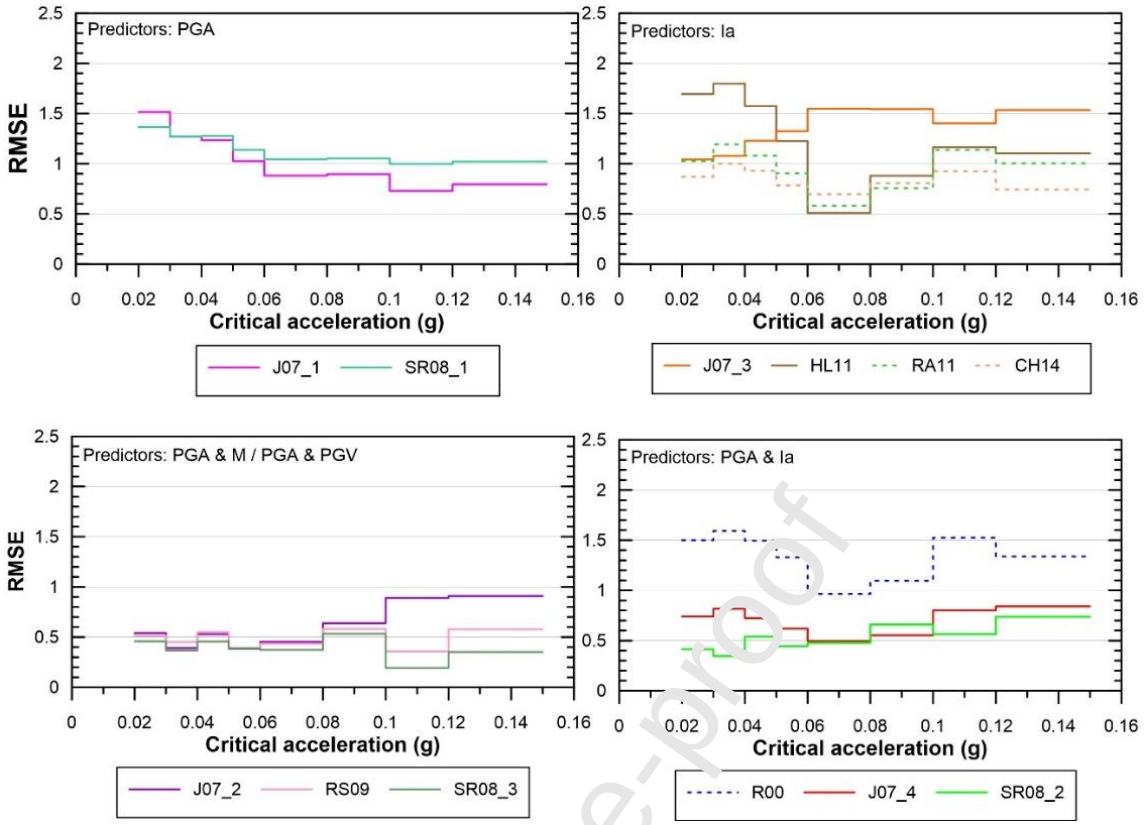

Figure 6. Variation of RMSE as a function of critical acceleration. Continuous lines depict models based on worldwide data. Dashed lines depict models based on regional data. See Table 1 for specific characteristics of each model.

Vector models once again perform better than scalar models. It is true not only for RMSE's absolute value but also for its variability as a k_y function (RMSE displays lower variance than scalar models).

4. Discussion

The evaluation of Newmark displacements is a basic requirement for the development of hazard maps of seismically-induced landslides. There are currently several models available for forecasting displacements. These are focused on a wide variety of predictors of ground motion, datasets for model development, and model functional forms.

Scalar models based on a single ground motion predictor were commonly used, as these predictors are typically the result of seismic hazard studies (expected *PGA* or *Ia* in a certain return period) and can be easily implemented (through regression models) to create seismically-induced landslide hazard maps. However, the analysis presented in the previous sections shows that present scalar models have several limitations.

In the case of models based on the *PGA* as a predictor, the obtained results using data provided by the Betic Cordillera (Figs. 5 and 6) suggest that such models should be used only in the range of the magnitudes or critical accelerations from which they were established, because the resulting average residuals are greater than 1.0. From Figure 3, the displacements predicted systematically overestimate the real ones. It may contribute to an unnecessary overestimation of real hazard values, and maps elaborated based on these models can erroneously restrict the use of the land when applied to low-to-moderate magnitude scenarios.

Models that include the magnitude of events in addition to *PGA* show a better performance (Figs. 5 and 6). Residuals are significantly smaller, and the aforementioned tendency to overestimate displacements is now corrected (Fig. 4). Perhaps, this could be the best way to continue using the *PGA* as a predictor of ground motion severity.

Models based in the *Ia* as predictor also show clear limitations. This parameter has been usually considered a better estimator of ground shaking severity than *PGA* in seismically-induced landslide studies because the whole accelerogram and range of frequencies are considered for its computation, taking into account both the duration and frequency content of the record (Wilson and Keefer, 1985; Keefer and Wilson, 1989). In contrast, *PGA* only considers a single point in the accelerogram (that of maximum amplitude), in most cases related to high frequencies. Consequently, *PGA* may be controlled by high-frequency pulses of short duration and may not properly reflect the frequency content and the duration of the

ground shaking (Harp and Wilson, 1995). Du *et al.* (2018) pointed out another important limitation of models based on *PGA*: it may be indistinguishable from the *PGA* of a high magnitude event in the far-field from that of a low-to-moderate magnitude in the near field. However, the frequency content and duration differ significantly between both scenarios and the consequences on the stability of slopes could be quite different.

Even though the *Ia* is considered more robust to quantify the severity of ground shaking, displacements predicted by models based on this parameter show greater variability than *PGA* models. In some cases, they overestimate displacements; in others, the opposite occurs. It is remarkable that regional models (Ra11 and CH14, but not JL18) typically display lower residuals than models obtained using comprehensive data (J07_3, HL11), initially with utility restricted to the areas where they were developed. In addition to the observed differences given the origin of the used data in designing the models, there are also significant variations in the range of values used when establishing the models. In essence, regional models consider data from moderate-to-low magnitude events, and low values for the critical acceleration (although HL11 model also considered low values of critical acceleration). Model JL18 shows the worst results when predicting our data, with very high residuals that show very high variability. Interestingly, this model was derived from a particular scenario: data from a single high magnitude event (Mw 7.9) and very high critical accelerations (k_y> 0.2 g). These conditions are distinctly different from those observed in the study area, where no data are available for events with Mw > 6.3, and displacements for critical accelerations above 0.2 g are very scarce in the dataset. Once again, these findings seem to point out that working with models established from an appropriate range of data (of M, of k_y) is vital, perhaps more critical than utilizing models derived from only the same geodynamic background as the area being studied. These results (Figs. 5, 6 and 8) also seem to point out that functional forms of relationships based on AI are forecasting significantly worse than those based on other ground motion predictors. New functional forms ought to be developed for this purpose. Maybe, the

use of extensive datasets and data-driven machine learning techniques could help to improve these relationships.

Vector models are considered more robust to predict Newmark displacements (Saygili and Rathje, 2008; Rathje and Saygili, 2009). Having a second parameter to explain the severity of ground shaking allows the elimination of the observed indetermination, such as while dealing with *PGA* based models (Du *et al.*, 2018). The obtained results generally show an increase of performance with respect to scalar models (lower RMSE and higher efficiency coefficient, see Table 3). Even though there is a general improvement in predictions, this does not apply for all models in the same way. Differences could not be attributable to the predictors but to the functional form of models. Thus, distribution of real vs. predicted displacements with models R00 and J07_4 (*PGA-Ia*), show a distribution similar to a hyperbola, showing a trend to an important overestimation for small/large displacements (Fig. 4). In these two cases, the functional forms of the models were highly sensitive to small changes in the predictors, and a slight increase in them would lead to a notable change in the predicted displacements (large values). On the contrary, some kind of saturation of the model exists in the low range of predictor's values, and predicted displacements start to be very similar, departing from real values (and overestimating them). Differences between these R00 and J04_7 models come from the fact that R00 model predicts larger displacements (Fig. 4).

Model R00 was of especial interest for this research. It was established with data from Italy, an area which is commonly considered to have a similar geodynamic framework to that found in the Betic Cordillera. Several ground motion prediction equations developed based on Italian data were used for hazard studies in Spain (i.e., Peláez *et al.*, 2005, Rodríguez-Peces *et al.*, 2011a). Nevertheless, the overestimation of displacements, especially for the larger ones, prevents its use in the Betic cordillera because it can lead to a significant overestimation of hazard.

For the remaining vector models, the distribution of residuals shows that they can predict displacements very close to the obtained ones. Most of the residuals (>40%) are less than 0.25 $\log_{10}$ units. That implies that differences are under a factor of 2 (for example, differences were above 10 times the reported values for *PGA* models). Model SR08_2 considers just the same predictors as R00 and J07_4 models but it fits better here (Fig. 4), showing a quasi-lineal relation between observed and predicted displacements. Statistics of control of this model are also significantly better than J07_4 model. Moreover, Model SR08_3 uses *PGA* and *PGV* as predictors. Displacements predicted based on this model could be considered the best of all considered in our analysis (Table 3, Figs. 3 – 5). These two models share the same polynomial functional form to forecast displacements. The RS09 model (*PGA-M*) provides a similar functional form, which also shows a lineal relationship between real and predicted displacements.

5. Proposal of regression models for the Betic Cordillera

Regression models used in our research have shown wide variability in performance to predict Newmark displacements in the Betic Cordillera, using the available data from low-to-moderate magnitude earthquakes. It is remarkable to point out the low performance of scalar models among the presented results regardless of the predictor considered. Thus, RMSE is very high for low critical accelerations (k_y< 0.05 g) and low magnitude (Mw < 5.0) in the *PGA*-based models. Similarly, *Ia*-based models show a noticeable irregular but high RMSE.

Displacement data provided by earthquakes occurred the Betic Cordillera have been used to develop models valid for this area. In light of the presented results, it seems that polynomial models work better than other functional forms in this area. Then, only polynomial functional models were considered for the *PGA* model as well as a functional form similar to that of CH14

(Chousianitis *et al.*, 2014), which has been the *Ia*-model that better performed in our study. Additionally, polynomial models based in *PGA*-M, *PGA*-*Ia* and *PGA*-*PGV* also have been tested.

Data include the displacements computed as described in section 2. Since most data correspond to low values of critical acceleration (0.02 and 0.03 g; Fig. 2), we have used a weighted regression procedure to avoid an excessive influence of these data in the resulting models. For this purpose, data were binned into 8 categories as a function of critical accelerations (k_y = 0.02, 0.03, 0.04, 0.05, 0.06, 0.08, 0.10, and a final category that includes those data with K_y> 0.1 g). Each category was weighted by 1/8, and weight of individual datum within each category was set to '(1/8)/number of data in the bin'. The analysis was conducted using the Statgraphics software (©Statpoint Technologies, 2019).

For each considered regression model, results showed that P-values for the calculated coefficients of the term k_y/PGA in all polynomic models were always greater than 0.05. Therefore, that terms are not statistically significant at the 95.0% or higher confidence levels. Similar result is observed with the term $log_{10}(Ia)*log_{10}(k_y)$ in the model that uses *Ia* as predictor of ground motion. Consequently, we have removed them in the corresponding models. This may be due to the limited amount of available data for the analysis. Presumably, as more data becomes available, the coefficients for those terms may become statistically more significant and the models will retain these terms.

Once removed, data were fitted with the new functional forms. The corresponding models are presented in Table 4. Residuals obtained for each model are shown in Figures 7 and 8. The resulting models are usually characterized by high values of the correlation coefficient (ρ^2), above 0.85, and low standard deviations (between 0.3 and 0.5). It is remarkable the good quality of the model that uses *PGA* and *PGV*: ρ^2 = 0.94, standard deviation = 0.35 and residuals of almost 70% of data are lower than 0.25 $\log_{10}$ units (Fig. 7). Because these models were obtained from a dataset that includes the range of accelerations of interest for low-to-

moderate magnitude earthquakes, RMSE do not shows any dependence of k_y. This is especially remarkable in the case of the *PGA*-model, where currently available models show a strong relation of RMSE with k_y (Fig. 8).

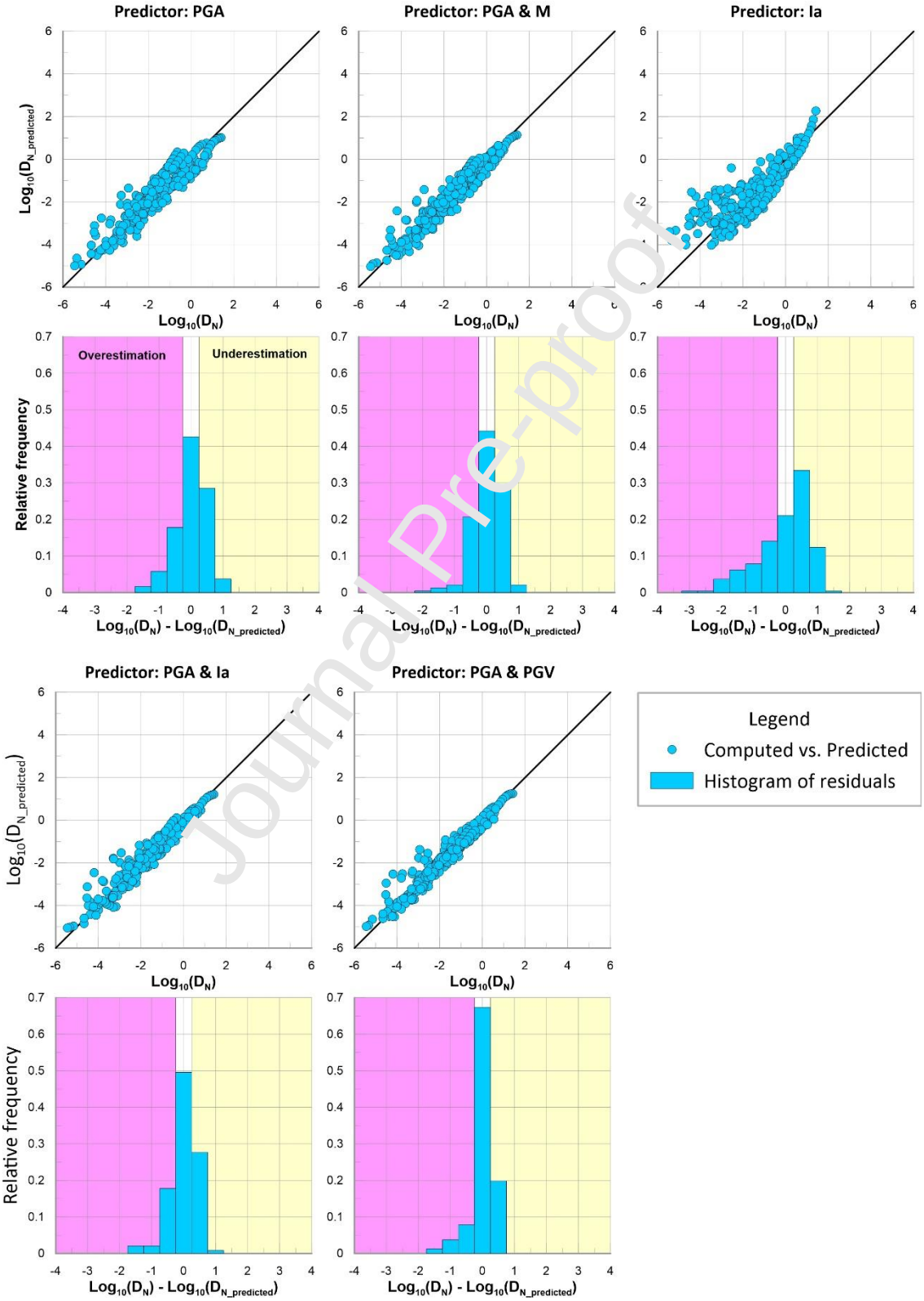

Figure 7. Residuals (and their relative frequency) obtained with regression models obtained for the Betic Cordillera. See Table 4 for more data about the models.

Yet again, the model that uses *Ia* as ground motion predictor shows the worst statistics (Table 4), being $\rho^2 = 0.67$ and the standard deviation is as high as 0.8. Figure 7 shows that only 20% of data show residuals below 0.25 $\log_{10}$ units. Distribution of data vs. predicted values of displacement shows the same hyperbolic shape (although less pronounced) already observed in other models based on this parameter. This may imply a possible deficiency of the functional form of these models.

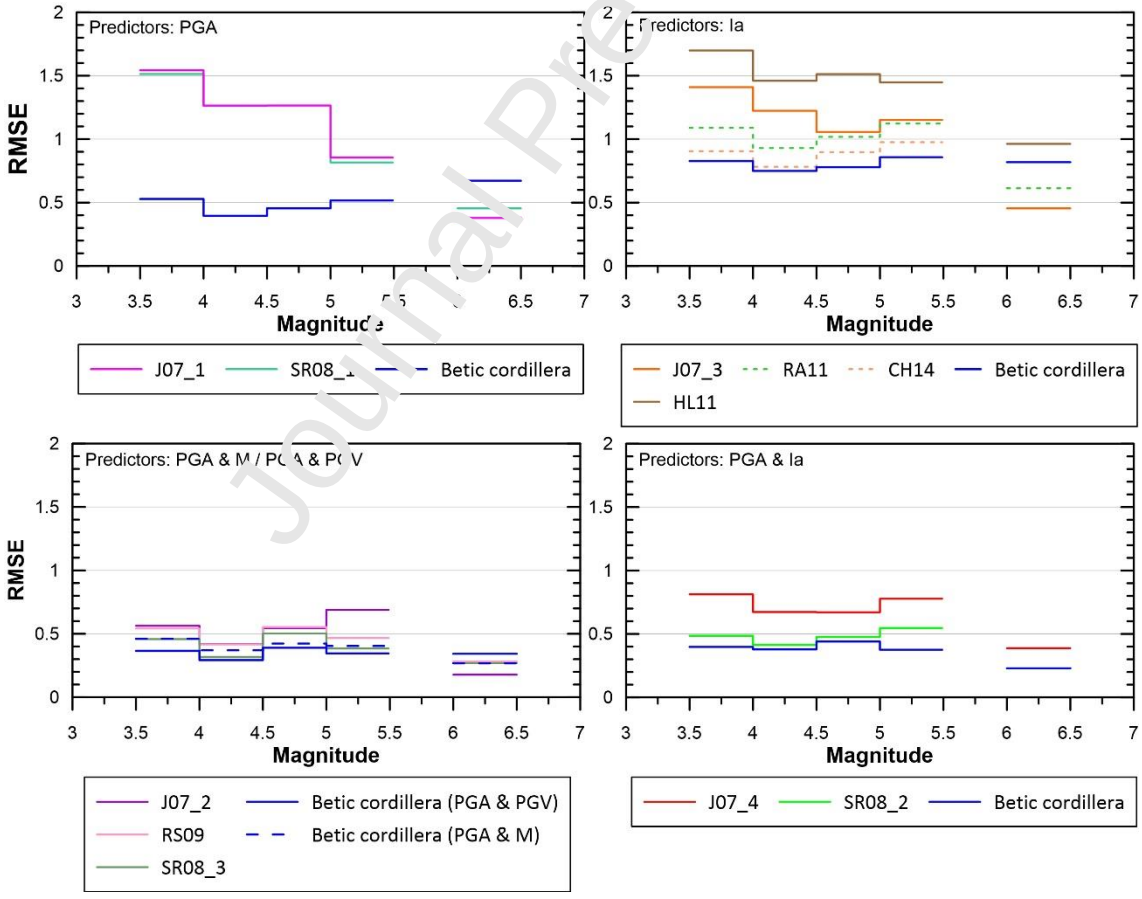

Figure 8. Variation of RMSE as a function of critical acceleration in the models obtained for the Betic Cordillera (blue lines). See Tables 1 and 4 for more details concerning regression models.

These results (Figs. 5, 6 and 8) seem to point out that functional forms of relations based on *Ia* are working worse than those based on other ground motion predictors. New functional forms should be considered for this purpose. Perhaps, the use of extensive datasets and data-driven machine learning techniques may help to improve these relations.

6. Conclusions

The use of regression models to estimate Newmark displacements is a standard procedure when estimating seismically-induced landslide hazard maps. At present, most available models have been obtained from data recorded during earthquakes of moderate-to-high magnitude (Mw > 6.5). However, many areas in the world are affected by moderate-to-low magnitude earthquakes. Thus, the reliability of existing models is not assured for these areas. In our study, an analysis of the behavior of some of these regression models was conducted by comparison with data from the Betic Cordillera (S Spain). The aim has been to evaluate their usefulness to predict reliable values of Newmark displacements for moderate-to-low magnitude scenarios.

The results have shown that the performance of scalar models is low. A general trend to overestimate displacements is detected in models that use the *PGA* as ground motion predictor, while the behavior is irregular in *Ia*-based models. In both cases, errors in displacement forecasting increase, considering lower magnitudes (Mw < 4.5-5.0) or lower critical acceleration scenarios (k_y< 0.05 g). Regional regression models, based on low magnitude and low critical accelerations, give more significant results than worldwide ones. It seems that it is imperative to have an appropriate range of data (both for Mw and k_y) to establish the regression models since they may overcome differences due to the geodynamic context of the area where data come from.

Vector models show a better performance regardless of the combination of ground motion predictors considered in the model. The inclusion of a second ground motion parameter allows a better characterization of ground motion severity and the extrapolation of existing models is possible (even though they did not include Mw or k_y found in moderate-to-low seismic scenarios).

The functional form of regression models displays some control over the quality of the results obtained in this study. Polynomial forms exhibit better behavior than other forms based on the logarithm of ground motion predictors. It becomes especially relevant when *Ia* is considered as a predictor from which no polynomial model exists, and the accuracy of predictions is lower than in *PGA* or *PGA-PGV* dependent models. New functional forms should be considered for this ground motion predictor. New studies are needed on this topic. Perhaps the usage of very large datasets and data-driven machine learning techniques may be of great benefit.

Finally, we propose new models for the Betic Cordillera. They have been developed using polynomial functional forms, and are capable of reducing the residuals in all situations with regard to currently available models. They are especially suitable for low magnitude events (Mw < 5.0) and low critical accelerations (k_y< 0.1 g).

7. Acknowledgements

This work was partially founded by the Spanish Ministry of Economy and EU FEDER funds (project EPILATES, CGL2015-65602-R) and by the Research Group VIGROB-184 (University of Alicante). The Instituto Geográfico Nacional (IGN) shared the strong ground motion data used in this study. Special thanks to Mr. J.M. Alcalde (IGN Madrid), who kindly explained all doubts about data provided. We also acknowledge suggestions made by J. J. Giner (Univ. Alicante) during the statistical analysis of data. Research was also partially funded by the Consejería de

Economía, Conocimiento, Empresa y Universidad (Junta de Andalucía), in the frame of the Programa Operativo FEDER Andalucía 2014-2020 – call made by the University of Jaén 2018.

8. References

Alfaro, P., Delgado, J., García-Tortosa, F.J., Lenti, L., López, A., López-Casado, C., Martino, S., 2012. Widespread landslides induced by the Mw 5.1 Lorca, SE Spain, earthquake of 11 May 2011. *Eng. Geol.* 137-138, 40-52.

Bird, J.F. and Bommer, J.J., 2004. Earthquake losses due to ground failure. *Eng. Geol.* 75, 147-279.

Bozbey, I., Gundogdu, O., 2011. A methodology to select seismic coefficients based on upper bound Newmark displacements using earthquake records from Turkey. *Soil Dyn. Earthq. Eng.* 31, 440-451.

Bray, J.D., Macedo, J., Travasarou, T., 2018. Simplified procedure for estimating seismic slope displacements for subduction zone earthquakes. *J. Geotech. Geoenviron. Eng.* 144, 04017124.

Bray, J. D., Travasarou, T., 2007. Simplified procedure for estimating earthquake-induced deviatoric slope displacements. *J. Geotech. Geoenviron. Eng.* 133, 381–392.

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

CNIG (2013). Update of seismic hazard maps of Spain. 2012 (in Spanish). Centro Nacional de Información Geográfica. Madrid, Spain.

Delgado, J., Peláez, J.A., Tomas, R., García-Tortosa, F.J., Alfaro, P., López-Casado, C.,2011. Seismically-induced landslides in the Betic Cordillera (S Spain). *Soil Dyn. Earthq. Eng.* 31, 1203–1211.

Delgado, J., Marques, F.S.F., Vaz, T., 2013. Movimientos de ladera inducidos por terremotos en la Península Ibérica. *Cuaternario y Geomorfología* 27, 5-32.

Delgado, J., García-Tortosa, F.J., Garrido, J., Loffredo, A., López-Casado, C., Martin-Rojas, I., Rodríguez-Peces, M.J., 2015. Seismically induced landslides by a low magnitude earthquake: the Ossa de Montiel event (Central Spain). *Eng. Geol.* 196, 280–285.

Delgado, J., Rodríguez-Peces, M., García-Tortosa, F.J., Garrido, J., Martín, I., Alfaro, P.,2017. Seismic-Induced Landslides: Lessons Learned from Recent Earthquakes in Spain. In: Mikos, M., Casagli, N., Yin, Y., Sassa, K. (eds.): *Advancing culture of living with landslides: Diversity of landslide forms*. 4th World Landslide Forum, Ljubljana (Slovenia). June 2017. Springer International Publishing AG, Cham (Switzerland). Volume 4, 111-117.

Du, W., Huang, D., Wang, G., 2018. Quantification of model uncertainty and variability in Newmark displacement analysis. *Soil Dyn. Earthq. Eng.* 109, 286-298.

Harp, E.L., Wilson, R.C., 1995. Shaking intensity thresholds for rock falls and slides: Evidence from 1987 Whittier Narrows and Superstition Hills earthquake strong-motion records. *Bull. Seism. Soc. Am.* 85, 1739-1757.

Hsieh, S.Y., Lee, C.T., 2011. Empirical estimation of the Newmark displacement from the Arias intensity and critical acceleration. *Eng. Geol.* 122, 34-42.

IGN, Instituto Geográfico Nacional, 2019. Estadísticas de terremotos en la Península Ibérica. http://www.ign.es/web/resources/docs/IGNCnig/SIS-Tablas_estadisticas_PIberica.pdf. Last accessed: 19 october 2019.

Jafarian, Y., Lashgari, A., Haddad, A., 2019. Predictive Model and Probabilistic Assessment of Sliding Displacement for Regional Scale Seismic Landslide Hazard Estimation in Iran. *Bull. Seism. Soc. Am.* 109, 1581-1593.

Jia-Liang, J., Yin, W., Dan, G., Ren-Mao, Y., Xiao-Yan, Y., 2018. New evaluation models of Newmark displacement for Southwest China. *Bull. Seism. Soc. Am.* 108, 2221-2236.

Jibson, R.W. (2007). Regression models for estimating coseismic landslide displacement. *Eng. Geol.* 91, 209-218.

Jibson, R.W. (2011). Methods for assessing the stability of slopes during earthquakes—A retrospective. *Eng. Geol.* 122, 43-50.

Jibson, R.W., Harp, E.L.,2012. Extraordinary Distance Limits of Landslides Triggered by the 2011 Mineral, Virginia, earthquake. *Bull. Seism. Soc. Am.* 102, 2368-2377.

Jibson, R.W., Harp, E.L., Michael, J.A. (1998). A method for producing digital probabilistic seismic landslide hazard maps: An example from the Los Angeles California area. US Geological Survey Open-File Rep. 98-113, 17 p.

Jibson, R.W., Harp, E.L., Michael, J.A. (2000). A method for producing digital probabilistic seismic landslide hazard maps. *Eng. Geol.* 58, 271-289.

Jibson, R. W., Michael, J. A., 2009. Maps showing seismic landslide hazards in Anchorage, Alaska. U.S. Geological Survey Scientific Investigations Map 3077, scale 1:25.000, 11-p.

Keefer, D.K., 1984. Landslides caused by earthquakes. *Geological Soc. America Bull.* 95, 406-421.

Keefer, D.K., Wilson, R.C., 1989. Predicting earthquake-induced landslides, with emphasis on arid and semi-arid environments. In: Sadler, P.M., Morton, D.M.(eds.): *Landslides in a semi-arid environment*. Inland Geological Society, Vol. 2, 118-149.

Leshchinsky, B.A. (2018). Nested Newmark model to calculate the post-earthquake profile of slopes. *Eng. Geol.* 233, 139-145.

Martino, S., Bozzano, F., Caporossi, P., D'Angio, D., Della Seta, M., Esposito, C., Fantini, A., Fiorucci, M., Giannini, L. M., Iannucci, R., Marmoni, G.M., Mazzanti, P., Missori, C., Moretto, S.,

Piacentini, D., Rivellino, S., Romeo, R.W., Sarandrea, P., Schiliro, L., Troiani, F., Varone, C., 2019. Impact of landslides on transportation routes during the 2016-2017 Central Italy seismic sequence. *Landslides* 16, 1221-1241.

Newmark, N.N., 1965. Effects of the earthquakes on dams and embankments. *Geotechnique* 15, 139-160.

Nievas, C., Bommer, J.J., Crowley, H., van Elk, J., 2020. Global occurrence and impact of small-to-medium magnitude earthquakes: a statistical analysis. *Bull. Earthq. Eng.* 18, 1-35.

Peláez, J.A., Delgado, J., López-Casado, C., 2005. A preliminary probabilistic seismic hazard assessment in terms of Arias intensity in southeastern Spain. *Eng. Geol.* 77, 139-151.

Rajabi, A.M., Mahdavifar, M.R., Khamehchiyan, M., Del Gaudio, V., 2011. A new empirical estimator of coseismic landslide displacement for Zagros Mountain región (Iran). *Nat. Hazards* 59, 1189-1203.

Rathje, E.M., Bray, J.D., 1999. An examination of simplified earthquake-induced displacement procedures for earth structures. *Can. Geotech. J.* 36, 72-87.

Rathje, E.M., Bray, J.D., 2000. Nonlinear coupled seismic sliding analysis of earth structures. *J. Geotech. Geoenviron. Eng.* 126, 1002-1014.

Rathje, E.M., Saygili, G., 2009. Probabilistic assessment of earthquake-induced sliding displacements of natural slopes. *Bull. New Zealand Soc. Earthq. Eng.* 42, 18-27.

Rodríguez-Peces, M.J. García-Mayordomo, J., Azañón, J.M., Insua-Arévalo, J.M., Jiménez Pintor, J., 2011a. Constraining pre-instrumental earthquake parameters from slope stability back-analysis: Palaeoseismic reconstruction of the Güevéjar landslide during the 1st November 1755 Lisbon and 25th December 1884 Arenas del Rey earthquakes. *Quaternary International* 242, 76-89.

Rodríguez-Peces, M.J., García-Mayordomo, J., Azañón, J.M., Jabaloy, A., 2014. GIS application for regional assessment of seismically induced slope failures in the Sierra Nevada Range, South Spain, along the Padul Fault. *Environmental Earth Sciences* 72, 2423-2435.

Rodríguez-Peces, M.J., Pérez-García, J.L., García-Mayordomo, J., Azañón, J.M., Insua-Arévalo, J.M., Delgado, J., 2011b. Applicability of Newmark method at regional, sub-regional and site scales: seismically induced Bullas and La Paca rock-slide cases (Murcia, SE Spain). *Nat. Hazards* 59, 1109-1124.

Romeo, R., 2000. Seismically induced landslide displacements: a predictive model. *Eng. Geol.* 58, 337-351.

Saygili, G., Rathje, E.M., 2008. Empirical predictive models for earthquake-Induced sliding displacements of slopes. *J. Geotech. Geoenviron. Eng.* 134, 790–803.

SeismoSoft, 2016. Earthquake Engineering Software Solutions: SeismoSignal. .

Statpoint Technologies (2019). Statgraphics 18 Centurion. http://www.statgraphics.com/.

Wilson, R.C., Keefer, D.K., 1985. Predicting areal limits of earthquake-induced landsliding. In: Ziony, J.I. (ed.): *Evaluating Earthquake Hazards in the Los Angeles Region: An Earth-Science Perspective*. U.S. Geological Survey Professional Paper 1360, Washington, 317-345.

Yigit, A. (2020). Prediction of amount of earthquake-induced slope displacement by using Newmark method. *Eng. Geol.* 264, 105385.

Author statement:

J. Delgado: Conceptualization of the problem; Development of software code for calculating Dn; Analysis; Writing manuscript.

J. Rosa: Data processing; Analysis; Manuscript review & editing.

J.A. Peláez: Data processing; Statistical analysis; manuscript review & editing.

M.J. Rodríguez-Peces: Statistical analysis; Analysis; manuscript review & editing.

J. Garrido: Analysis; manuscript review & editing.

M. Tsigé: Analysis; manuscript review & editing.

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

No conflict.

Figure caption

Figure 1. Map showing the location of stations of the Spanish Strong Ground Motion Network operated by the Spanish IGN. The map also shows the epicentral location of earthquakes known to have induced landslides for the period 1919 to 2019.

Figure 2. Distribution of seismic data used in our analysis as a function of magnitude and epicentral distance, and ground motion parameters.

Figure 3. Residuals (and their relative frequency) obtained with regression models that use *PGA*, *PGA* and *PGV*, or *PGA* and magnitude as ground motion predictors. See Table 1 for more data about these models.

Figure 4. Residuals (and their relative frequency) obtained with regression models that use *Ia* or *PGA* and *Ia* as ground motion predictors. See Table 1 for more data about these models.

Figure 5. Variation of RMSE as a function of magnitude. Continuous lines depict models based on worldwide data. Dashed lines depict models based on regional data. See Table 1 for specific characteristics of each model.

Figure 6. Variation of RMSE as a function of critical acceleration. Continuous lines depict models based on worldwide data. Dashed lines depict models based on regional data. See Table 1 for specific characteristics of each model.

Figure 7. Residuals (and their relative frequency) obtained with regression models obtained for the Betic Cordillera. See Table 4 for more data about the models.

Figure 8. Variation of RMSE as a function of critical acceleration in the models obtained for the Betic Cordillera (blue lines). See Tables 1 and 4 for more details concerning regression models.

Table caption

Table 1. List of regression models for estimating Newmark displacement (Dn) used in the analysis. *PGA*: Peak ground acceleration (in g's); *PGV*: Peak ground velocity (in cm/s); Ia: Arias intensity (in m/s); M: Moment magnitude; k_y: Critical acceleration of slope (in g's). NA: Not available.

Table 2. List of earthquakes and strong ground motion records used in the analysis. Data from IGN (available at https://www.ign.es/web/ign/portal/sis-catalogo-acelerogramas).

Table 3. Statistics of the residuals ($\log_{10}(Dn_{observed}) - \log_{10}(Dn_{predicted})$) resulting with different regression model (see Table 1). RMSE: Root Mean Square Error; E: Efficiency coefficient.

Table 4. Regression models developed for estimating Newmark displacements in the Betic Cordillera.

Table 1. List of regression models for estimating Newmark displacement (Dn) used in the analysis. *PGA*: Peak ground acceleration (in g's); *PGV*: Peak ground velocity (in cm/s); Ia: Arias intensity (in m/s); M: Moment magnitude; k_y: Critical acceleration of slope (in g's). NA: Not available.

Predictor	Code (Reference)	$\log_{10}$ Dn (Dn in cm)	σ	ρ^2	Num. EQs	Magnitude range	Num. Records	N (*
PGA& Ia	R00 (Romeo, 2000)	$0.852 + 0.607\log_{10}\left(\frac{Ia}{100}\right) - 3.719\log_{10}\left(\frac{k_y}{PGA}\right)$	0.365	0.89	17	4.6 – 6.8	190	
PGA	J07_1 (Jibson, 2007)	$0.215 + 2.341\log_{10}\left(1 - \frac{k_y}{PGA}\right) - 1.438\log_{10}\left(\frac{k_y}{PGA}\right)$	0.510	0.84	30	5.3 – 7.6	2770	
PGA& M	J07_2 (Jibson, 2007)	$-0.710 + 2.335\log_{10}\left(1 - \frac{k_y}{PGA}\right) - 1.478\log_{10}\left(\frac{k_y}{PGA}\right) + 0.424M$	0.454	0.87				
Ia	J07_3 (Jibson, 2007)	$-3.230 + 2.401\log_{10}(Ia) - 3.481\log_{10}(k_y)$	0.656	0.71				
PGA& Ia	J07_4 (Jibson, 2007)	$-1.474 + 0.56\log_{10}(Ia) - 3.833\log_{10}\left(\frac{k_y}{PGA}\right)$	0.616	0.75				
PGA	SR08_1 (Saygili & Rathje, 2008)	$0.4343\left[5.52 - 4.43\left(\frac{k_y}{PGA}\right) - 20.39\left(\frac{k_y}{PGA}\right)^2 + 42.61\left(\frac{k_y}{PGA}\right)^3 - 28.74\left(\frac{k_y}{PGA}\right)^4 + 0.72\ln(PGA)\right]$	1.13	NA	54	5.0 – 7.9	2383	
PGA& Ia	SR08_2 (Saygili & Rathje, 2008)	$0.4343\left[2.39 - 5.24\left(\frac{k_y}{PGA}\right) - 18.78\left(\frac{k_y}{PGA}\right)^2 + 42.01\left(\frac{k_y}{PGA}\right)^3 - 29.15\left(\frac{k_y}{PGA}\right)^4 - 1.56\ln(PGA) + 1.38\ln(IA)\right]$	0.46 + 0.56(k_y/*PGA*)	NA				
PGA&*PGV*	SR08_3 (Saygili & Rathje, 2008)	$0.4343\left[-1.56 - 4.58\left(\frac{k_y}{PGA}\right) - 20.84\left(\frac{k_y}{PGA}\right)^2 + 44.75\left(\frac{k_y}{PGA}\right)^3 - 30.5\left(\frac{k_y}{PGA}\right)^4 - 0.64\ln(PGA) + 1.55\ln(PGV)\right]$	0.41 + 0.52(k_y/*PGA*)	NA				
PGA& M	RS09 (Rathje & Saygili, 2009)	$0.4343\left[4.89 - 4.85\left(\frac{k_y}{PGA}\right) - 19.64\left(\frac{k_y}{PGA}\right)^2 + 42.49\left(\frac{k_y}{PGA}\right)^3 - 29.06\left(\frac{k_y}{PGA}\right)^4 + 0.72\ln(PGA) + 0.89(M-6)\right]$	0.732 + 0.789(k_y/*PGA*) – 0.539(k_y/*PGA*)	NA	54	5.0 – 7.9	2383	
Ia	HL11 (Hsieh & Lee, 2011)	$1.84 + 0.847\log_{10}(Ia) - 10.62k_y + 6.587k_y\log_{10}(Ia)$	0.259	0.89	6	6.7 – 7.6	1343	
Ia	Ra11 (Rajabi *et al.*, 2011)	$-1.454 + 1.202\log_{10}(Ia) - 1.585\log_{10}(k_y)$	0.358	NA	80	3.6 - 7	108	
Ia	CH14 (Chousianitis *et al.*, 2014)	$-1.495 + 2.228\log_{10}(Ia) - 2.498\log_{10}(k_y) + 0.273\log_{10}(Ia)\log_{10}(k_y)$	0.231	0.95	98	3.2 – 6.7	205	
Ia	JL18 (Jia-Liang *et al.*, 2018)	$2.092 + 0.465\log_{10}(Ia) - 22.201k_y + 12.896k_y\log_{10}(Ia)$	0.148	0.92	1	7.9	33	

*Displacements were computed, for each accelerogram, by setting the ratio k_y/PGA instead of fixing the value of k_y.

**Figure 3 of Saygili and Rathje (2008) plots accelerograms versus distance and magnitude of events. Some accelerograms were used in the range Mw 5.0-5.5 but the overlap of symbols prevents estimating the number of them or the percentage with respect to total data used in the analysis.

Table 2. List of earthquakes and strong ground motion records used in the analysis. Data from IGN (available at https://www.ign.es/web/ign/portal/sis-catalogo-acelerogramas).

Date	Time	Lat.	Lon.	Depth (km)	Mw	Num. Records
23/12/1993	14:22:35	36.78	-2.9367	8	5.0	2
04/01/1994	8:03:15	36.5717	-2.815	2	4.9	3
19/04/1994	23:51:59	37.3083	-1.9467	5	3.6	1
17/03/1995	14:04:14	37.175	-3.7733	2	3.9	3
09/01/1996	7:36:59	37.055	-3.92	2	3.8	2
28/12/1996	7:30:37	37.1617	-3.7167	1	4.1	7
24/02/1997	7:09:51	37.02	-3.835	6	4.3	6
18/11/1998	23:18:11	36.9678	-3.7792	3	3.8	4
02/02/1999	13:45:17	38.0963	-1.5014	1	4.7	2
10/09/2003	20:22:47	37.1189	-3.7966	3	3.6	2
29/01/2005	7:41:32	37.8535	-1.7555	11	4.8	2
04/01/2007	23:32:32	37.2008	-3.7447	-	3.8	3
30/06/2007	3:53:45	37.0784	-5.3732	8	4.4	1
06/02/2008	17:53:00	36.8928	-2.1944	10	4.3	2
21/07/2008	2:30:03	39.013	-0.433	[illegible]	3.6	2
02/10/2008	4:02:53	37.0442	-5.4112	8	4.7	2
05/11/2009	5:39:55	37.0517	-3.8217	10	3.9	6
11/05/2011	15:05:13	37.7196	-1.7076	2	4.5	2
11/05/2011	16:47:26	37.7175	-1.7114	4	5.1	8
11/05/2011	20:37:45	37.7308	-1.701[illegible]	4	3.9	2
26/02/2012	15:31:35	37.0618	-3.8[illegible]63	2	3.5	4
05/02/2013	21:23:48	38.0399	-2.27[illegible]5	3	3.7	2
05/02/2013	21:24:12	38.0476	-3.[illegible]651	-	3.9	2
23/11/2015	8:02:11	36.8027	-[illegible]1812	7	3.8	2
25/01/2016	4:22:01	35.6004	-3.8056	12	6.3	2
31/01/2016	16:25:27	36.5694	-3.0897	-	4.5	2
22/02/2016	3:46:03	35.634[illegible]	-3.6097	12	5.1	2
12/03/2016	15:04:07	35.5[illegible]	-3.5971	22	4.8	2
15/03/2016	4:40:40	35.[illegible]26	-3.6483	11	5.2	1
03/05/2016	11:56:38	3[illegible].7181	-1.5876	-	3.7	1
02/03/2018	19:08:22	[illegible].931	-1.4223	12	3.9	2
13/12/2018	7:58:27	37.6749	-1.7424	1	3.5	1
25/10/2019	9:35:48	36.9835	-5.2931	-	4.5	2

Table 3. Statistics of the residuals ($\log_{10}(Dn_{observed}) - \log_{10}(Dn_{predicted})$) resulting with different regression model (see Table 1). RMSE: Root Mean Square Error; E: Efficiency coefficient.

Predictor	Regression model	RMSE	E
PGA	J07_1	1.266	0.203
	SR08_1	1.249	0.223
PGA& M	J07_2	0.550	0.850
	RS09	0.495	0.878
Ia	J07_3	1.215	0.265
	HL11	1.531	-0.166
	RA11	1.031	0.471
	CH14	0.882	0.613
	JL18	1.944	-0.880
PGA& Ia	R00	1.451	-0.047
	J07_4	0.733	0.733
	SR08_2	0.474	0.888
PGA&*PGV*	SR08_3	0.421	0.912

Table 4. Regression models developed for estimating Newmark displacements in the Betic Cordillera.

Predictor	Primary type of Equation	Equation for $\log_{10}$ Dn (Dn in cm)	Σ	ρ^2	E
PGA	SR08_1	$1.655\ (\pm0.166) - 13.755\ (\pm3.193)\left(\frac{k_y}{PGA}\right)^2 + 26.429\ (\pm7.012)\ \left(\frac{k_y}{PGA}\right)^3 - 16.897\ (\pm4.090)\left(\frac{k_y}{PGA}\right)^4 + 1.487\ (\pm0.127)\ \log_{10}(PGA)$	0.537	0.85	0.88
PGA & M	RS09	$-0.813\ (\pm0.261) - 12.428\ (\pm2.602)\left(\frac{k_y}{PGA}\right)^2 + 22.873\ (\pm5.717)\left(\frac{k_y}{PGA}\right)^3 - 14.487\ (\pm3.337)\left(\frac{k_y}{PGA}\right)^4 + 1.268\ (\pm0.105)\log_{10}(PGA) + 0.486\ (\pm0.044)\ M$	0.438	0.90	0.91
Ia	CH14	$-1.014\ (\pm0.240) + 2.185\ (\pm0.102)\log_{10}(Ia) - 2.291\ (\pm0.235)\log_{10}(k_y)$	0.806	0.67	0.68
PGA & Ia	SR08_2	$1.416\ (\pm0.127) - 11.110\ (\pm2.423)\left(\frac{k_y}{PGA}\right)^2 + 20.421\ (\pm5.321)\ \left(\frac{k_y}{PGA}\right)^3 - 13.303\ (\pm3.105)\ \left(\frac{k_y}{PGA}\right)^4 - 0.279\ (\pm0.163)\ \log_{10}(PGA) + 1.056\ (\pm0.079)\ \log_{10}(Ia)$	0.406	0.92	0.92
PGA & PGV	SR08_3	$-1.419(\pm0.203) - 10.713\ (\pm2.092)\left(\frac{k_y}{PGA}\right)^2 + 19.787\ (\pm4.565)\left(\frac{k_y}{PGA}\right)^3 - 13.065\ (\pm2.680)\left(\frac{k_y}{PGA}\right)^4 - 0.530\ (\pm0.140)\log_{10}(PGA) + 1.632\ (\pm0.091)\log_{10}(PGV)$	0.351	0.94	0.94

Highlights:

- Performance of existing models in low-to-moderate magnitude scenarios is evaluated.
- Scalar models show magnitude and critical acceleration dependent behavior.
- Best performance is found with polynomial vectorial models.
- Models for the Betic Cordillera are proposed.