



New insights into the relationship between seismic intensity measures and nonlinear structural response

Yeudy F. Vargas-Alzate¹ · Jorge E. Hurtado² · Luis G. Pujades¹

Received: 2 January 2021 / Accepted: 19 November 2021 / Published online: 30 December 2021
© The Author(s) 2021

Abstract

This paper focuses on the probabilistic analysis of Intensity Measures (*IMs*) and Engineering Demand Parameters (*EDPs*) in the context of earthquake-induced ground motions. Several statistical properties, which are desirable in *IMs* when they are used to predict *EDPs*, have been analysed. Specifically, efficiency, sufficiency and *steadfastness* have been quantified for a set of *IMs* with respect to two *EDPs*: the maximum inter-storey drift ratio, *MIDR*, and the maximum floor acceleration, *MFA*. Steadfastness is a new statistical property proposed in this article, which is related to the ability of *IMs* to forecast *EDPs* for large building suites. In other words, this means that efficiency does not significantly vary when different types of buildings are simultaneously considered in the statistical analyses. This property allows reducing the number of calculations when performing seismic risk estimations at urban level since, for instance, a large variety of fragility curves, representing specific building typologies, can be grouped together within a more generic one. The main sources of uncertainty involved in the calculation of the seismic risk have been considered in the analysis. To do so, the nonlinear dynamic responses of probabilistic multi-degree-of-freedom building models, subjected to a large data set of ground motion records, have been calculated. These models have been generated to simulate the dynamic behavior of reinforced concrete buildings whose number of stories vary from 3 to 13. 18 spectrum-, energy- and direct-accelerogram-based *IMs* have been considered herein. Then, from clouds of *IM-EDP* points, efficiency, sufficiency and steadfastness have been quantified. For *MIDR*, results show that *IMs* based on spectral velocity are more efficient and steadfast than the ones based on spectral acceleration; spectral velocity averaged in a range of periods, $A_v S_v$, has shown to be the most efficient *IM* with an adequate level of steadfastness. For *MFA*, spectral acceleration-based-*IMs* are more efficient than velocity-based ones. A comparison is also presented on the use of linear vs quadratic regression models, and their implications on the derivation of fragility functions. Concerning sufficiency, most of the 18 *IMs* analysed do not have this property. Nonetheless, multi-regression models have been employed to address this lack of sufficiency allowing to obtain a so-called ‘ideal’ *IM*.

Keywords Intensity measure · Engineering demand parameter · Efficiency · Sufficiency · Steadfastness

✉ Yeudy F. Vargas-Alzate
yeudy.felipe.vargas@upc.edu

Extended author information available on the last page of the article

1 Introduction

The main purposes of seismic risk assessment is to know the probability of occurrence of seismic actions—*hazard*—and that of the expected damages—*risk*. Concerning *design and retrofiting*, the aim is to act on the project and construction of structures to decrease the expected damage, resulting in a reduction of the catastrophic effects of earthquakes on society. Proper assessment and design methodologies require knowledge of how seismic actions, in terms of intensity measures (*IMs*), affect structural response. Variables representing this response are commonly known as engineering demand parameters (*EDPs*).

Several *IMs* have been proposed to better understand the seismic performance of buildings. In this respect, a relevant milestone has been the response spectra theory; see for instance Housner (1941, 1970), Housner and Jennings (1982) and Newmark and Hall (1982). Based on this, *IMs* considering the damped response of single-degree-of-freedom systems (*SDoF*) were introduced. This knowledge, combined with the increasing capacity of computers and collection of strong motion databases, have allowed to perform probabilistic calculations which had previously been unthinkable. In this way, statistical analyses of several spectrum-based *IMs* and *EDPs* (obtained through advanced nonlinear structural models) have been recently performed (Padgett et al. 2008; Jalayer et al. 2012; Ebrahimian et al. 2015, 2021; Bojórquez et al. 2017a,b; Pinzón et al. 2020; Du and Padgett 2021; O'Reilly 2021).

Efficiency, sufficiency and *steadfastness* are analysed and discussed in this article. These statistical properties are desirable for *IMs* when they are employed to predict the seismic response of civil structures. In short, efficiency is related to the predicting power; sufficiency means that the *IM* renders the conditional response distribution independent of other earthquake characteristics; *steadfastness*, introduced herein, is related to stability of efficiency for predicting the behavior of building populations with increasing diversity.

From the statistical point of view, the most desirable feature that an *IM* should exhibit is efficiency (Shome and Cornell 1999; Giovenale et al. 2004; Luco and Cornell 2007). An *IM* is considered efficient if it reduces the dispersion of the parameter that represents the structural response (*EDP*). Efficiency can be quantified through the dispersion of *IM-EDP* points; the lower this scattering the higher the efficiency. Several *IM* variables, that have been commonly used to quantify efficiency, are explored herein (Padgett et al. 2008; Ebrahimian et al. 2015; Ebrahimian and Jalayer 2021). In addition, new *IMs* based on average spectral values are also introduced.

As commented above, a sufficient *IM* makes the statistical properties of a cloud of *IM-EDP* pairs independent of other earthquake parameters (mainly those related to the earth crust rupture at the source of the event). This property can be quantified through the bivariate analysis of the residuals of a cloud of *IM-EDP* points and a specific seismological parameter (Luco and Cornell 2007). If these variables are not correlated, it means that the *IM* is sufficient with respect to the seismological parameter. Note that, lack of sufficiency can be seen as an opportunity to increase the efficiency of the *IM* by applying multi-regression analysis, as will be shown later (Elefante 2010); in this case, the seismological parameters become information variables, *IVs*. In general, for the purpose of this research, an *IV* can be any property that varies from simulation to simulation.

The last desirable *IM* statistical property analysed in this article, which has not heretofore been systematically investigated, at least up to the authors' knowledge, is steadfastness with respect to variations of structural properties. This means that the efficiency of steadfast *IMs*, after grouping results of building models with different physical characteristics, is stable. Note that a steadfast *IM* could be used to reduce the number of structural calculations when estimating seismic risk at urban level.

When assessing seismic risk, the most desirable feature for *EDPs* is that they quantify, as best as possible, the performance level of structures affected by earthquake-induced ground motions. One of the most used *EDP* for estimating this performance level is the maximum inter-story drift ratio, *MIDR* (Mayes 1995). Many other variables like the maximum displacement at the roof, base shear, maximum floor acceleration (*MFA*), or damage indices, can be considered as *EDPs*.

It is worth noting that both *IMs* and *EDPs* are random variables showing high dispersion. In the case of *IMs*, this is due to the complexity of the strong motion influencing a point at the ground level where the acceleration is recorded. Regarding *EDPs*, randomness comes not only from the seismic input itself, but also from the mechanical properties of the materials, the geometric features of the structure, directionality effects, amongst other aspects (Vargas et al. 2018). Therefore, efficiency, sufficiency and steadfastness should be measured by considering several sources of uncertainty. For instance, if a specific structure is analysed, uncertainties related to the mechanical properties of the materials and seismic action are the ones carrying more variability into the response (Vamvatsikos and Fragiadakis 2010; Jalayer et al. 2015; Vargas-Alzate et al. 2019a). In the case of urban environments, information related to the geometrical distribution of buildings belonging to the emplacement are also required (Silva et al. 2015; Vargas-Alzate et al. 2020). It is worth mentioning that the quantification of statistical properties in simulations that consider random sources of high variability can lead to identifying new properties of the *IMs*, as in the case of steadfastness.

There are many statistical approaches to extract information from *IM-EDP* pairs on the expected damage to buildings (Vamvatsikos and Cornell 2002; Jalayer 2003; Jalayer and Cornell 2009; Vargas et al. 2013; Jalayer et al. 2015; Vargas-Alzate et al. 2019a). The main differences between them lie in the sampling strategy used to obtain these pairs, the selection procedure of the strong motion records, and the *IMs* and *EDPs* selected. In general, results suggest that the selection of the most suitable *IM* to analyse the probabilistic distribution of an *EDP* highly depends on the structural type and seismogenic environment under consideration. In this article, the cloud analysis approach has been considered to analyze the relationship between *IMs* and *EDPs* (Jalayer et al. 2015).

Within the framework of the KaIROS project (<https://cordis.europa.eu/project/id/799553>), several numerical tools to deal with the aforementioned uncertainties have been developed. These tools are used herein to analyse the statistical relationships between a group of *IMs* and the response of buildings in terms of *MIDR* and *MAF*. Specifically, a set of a wide variety of probabilistic numerical models representing the behavior of Reinforced Concrete (RC) frame buildings are used as a testbed (Vargas-Alzate et al. 2019b). In order to characterize seismic hazard, a group of ground motion records, selected and scaled to get response values within and beyond the elastic limit, has been selected and scaled. Once building models and seismic actions have been characterized, *EDPs* are calculated using Nonlinear Dynamic Analysis, *NLDA*, as numerical tool. This latter has been considered the most reliable for numerically simulating the seismic response of buildings.

Then, from the ground motion set, a group of basic *IMs* is obtained and correlated with both *EDPs* (*MIDR* and *MAF*). As pointed out above, several statistical properties have been analysed: (1) the competence to predict *EDPs* (efficiency); (2) the sufficiency of *IMs* with respect to seismological parameters; and (3) the ability to group buildings with different features without increasing the variability of the bivariate distribution (steadfastness). It is shown that there is a set of *IMs* exhibiting efficiency, sufficiency and steadfastness. For insufficient *IMs*, it has been developed vector-valued *IMs* with increased efficiency (Elefante et al. 2010).

Note that the quantification of the expected damage of buildings can be improved if *IM-EDP* pairs exhibiting steadfastness and high-level of correlation are considered. This is relevant for assessment and design. For instance, fragility functions obtained from these highly correlated pairs would provide more accurate estimations on the expected damage state of structures. Moreover, safety factors used in design could be better quantified. Thus, this article focuses primarily on identifying ideal *IMs*, which are the ones that exhibit the highest efficiency and, for practicality, sufficiency and steadfastness.

2 Probabilistic structural models (exposure)

The success of developing suitable building models depends on the quantity and quality of information regarding the mechanical and geometrical properties of the structural elements. When the analysis is oriented to urban environments, it is also important to know about the geographical distribution of as many physical properties as possible. Of great importance is to know, for instance, the number of structures belonging to each structural type, distribution of the number of stories and spans, azimuthal position of buildings (Vargas-Alzate et al. 2021), amongst many others. In addition, the variability level of the mechanical and geometrical properties of the elements, as well as that of acting loads, should be available. A more detailed discussion of several of these random variables is provided below.

2.1 Gravity loads and mechanical properties

As commented above, to properly analyse the probabilistic seismic behavior of a group of buildings, it is necessary to consider the aleatory uncertainties regarding input variables. Concerning gravity loads and mechanical features of the materials, the following properties have been considered as random variables: live loads, *LL*, permanent loads, *DL*, compressive strength of concrete, *f_c*, tensile strength of steel, *f_y*, elastic modulus of concrete, *E_c*, and elastic modulus of steel, *E_s*. Continuous Gaussian distributions have been assumed for these variables; mean values, μ , standard deviations, σ , and coefficient of variation, c.o.v. of each variable are shown in Table 1. It is worth mentioning that properties assigned to *LL* correspond to a fraction (33%) of the design value (Eurocode 8).

2.2 Geometrical properties

Concerning geometrical properties, the following random variables have been considered: number of stories, N_{st} , number of spans, N_{sp} , story height, H_{st} , and span length, S_l . N_{st} and N_{sp} follow a uniform discrete distribution in the interval (3, 13) and (3, 6), respectively; H_{st} and S_l are distributed uniformly in the interval (2.8, 3.2) m and (4, 6) m, respectively.

Table 1 Mean values, μ , standard deviations, σ , and coefficient of variation, c.o.v., of the random variables related to the structural models

	RC frames		
	μ (kPa)	σ (kPa)	c.o.v
<i>LL</i>	1	0.1	0.1
<i>DL</i>	4.5	0.18	0.04
<i>f_c</i>	2.5E4	2.5E3	0.1
<i>f_y</i>	4.6E6	2.3E5	0.05
<i>E_c</i>	2.3E7	1.84E6	0.08
<i>E_s</i>	2.18E8	1.09E7	0.05

In order to assign the cross-section dimensions of the first story columns, the following equation has been used:

$$W_c \text{ or } D_c = c_1 * \ln(N_{st}) + c_2 * S_l + c_3 * \Phi_{1,0} + c_4 \tag{1}$$

where c_i are coefficients that may be adjusted depending on the data distribution of the analysed area. For this study, $c_1 = 0.4$, $c_2 = 0.05$, $c_3 = 0.01$ and $c_4 = -0.35$; $\Phi_{1,0}$ represents the standard normal distribution. Note that columns are not necessarily square, that is, one random sample is generated for the width, W_c , and one for the depth, D_c , of the element (Eq. 1). For upper stories, each side of the columns decreases systematically by 5 cm every three stories. Values generated are rounded to the nearest multiple of 5 cm to be consistent with real dimensions of RC elements.

The width of beams, W_b , depends on the number of stories and span length of the building model, and it has been calculated by using the following equation:

$$W_b = b_1 * N_{st} + b_2 * S_l + b_3 \tag{2}$$

where b_i are coefficients that depend on the characteristics of the study zone. For the hypothetical case study, $b_1 = 0.01$, $b_2 = 0.02$ and $b_3 = 0.19$. The depth of the beams, D_b , is obtained using the following equation:

$$D_b = g_1 * N_{st} + g_2 * S_l + g_3 \tag{3}$$

where $g_1 = 0.01$, $g_2 = 0.05$ and $g_3 = 0.17$. Notice that no random term is considered for beams.

As pointed out above, within the framework of the KaiROS project, a software package has been developed for probabilistic seismic analyses by using as structural solver the Ruaumoko program (Carr 2000). This package allows to generate probabilistic *MDOF* systems representing the behavior of RC frame buildings (see also Vargas-Alzate et al. 2019b); this module has been adapted to simulate the seismic response of other structural types located in other seismic environments (Pinzón et al. 2020). Using this algorithm, 1160 structural models have been generated (this number is related to the availability of records within the ground motion database, as will be shown later).

The modified Takeda hysteresis law, Otani (1974), has been used to represent the in-cycle behavior of the structural elements (beams and columns). This law allows considering strength degradation due to large deformations and excessive number of inelastic cycles. To do so, a strength degradation function, which depends on the ductility reached by the structural elements, as well as a degrading coefficient associated to the number of inelastic cycles has been considered in the hysteretic model (see Carr 2000). For

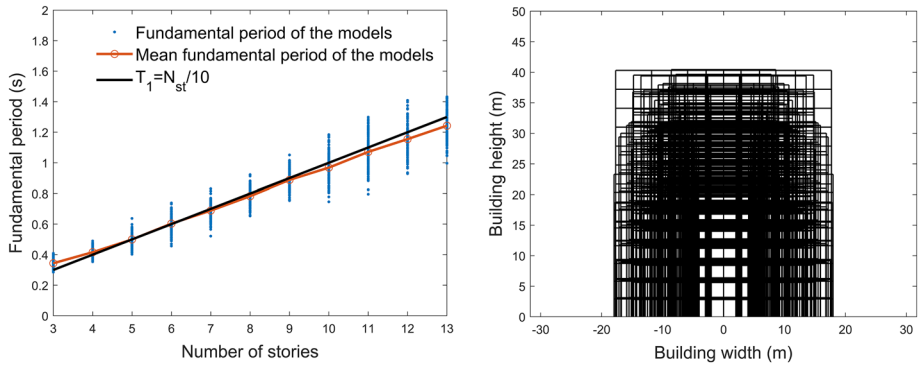


Fig. 1 Left: Variation of the fundamental period as a function of the number of storeys. Right: one-hundred building models generated with KaIROS software

stiffness degradation, the degrading factor has been fixed to 0.4. The yielding surfaces are defined by the bending moment–axial load interaction diagram for columns and bending moment–curvature for beams. P-Delta effects has been considered in the numerical model, as the structures analysed may experience deformations beyond the elastic range, sometimes exceeding their capacity to withstand gravity loads.

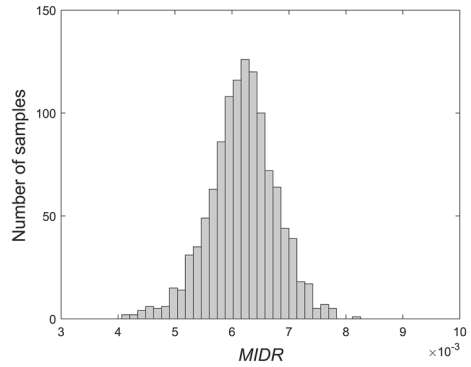
Coefficients considered in (1 to (3), as well as the mechanical property values (Table 1), have been selected to approximate the expected behavior of RC structures located in earthquake-prone regions. In order to verify this, three main aspects from the generated models have been analysed: (1) the fundamental period (T_1); (2) the horizontal stiffness; and (3) the nonlinear behavior given the design spectrum.

Regarding T_1 , the ones obtained from the generated models have been compared with those computed by means of the classical rough approximation $T_1 = N_{st}/10$. Figure 1(Left) shows the evolution of T_1 as a function of the number of stories for the 1160 building models; it can be seen that T_1 values agree with those expected for the building typology analysed. Figure 1(Right) shows a sketch of one hundred building models generated according to the conditions explained above.

In seismic design, when defining the horizontal stiffness of a building, it is common practice to limit the expected *MIDR* given a design spectrum (generally for a 475-years return period) provided by the seismic regulation of the area; this *MIDR* is generally limited to a maximum value of 0.01 (Eurocode 8). The most used methodology in practice to estimate this expected *MIDR* is the response spectrum method (Taghavi and Miranda 2010; Gupta and Hall 2017). Accordingly, the *MIDR* distribution of the generated models has been estimated by considering as seismic hazard the design spectrum type 1 (Eurocode 8), for a soil type A, and a basic acceleration, a_b , equal to 0.4 g. Figure 2 shows the distribution of the *MIDRs* after applying the response spectrum method; it can be seen that none of the generated models exceeds the threshold of 0.01 given the design level considered.

Regarding the nonlinear behavior, the capacity curves of the generated models have been calculated by using the adaptive pushover analysis (Satyarno 2000). This structural analysis type has been employed herein since it allows an automatic identification of the last point in the capacity curve, δ_u . The longitudinal steel percentages have been assigned to the structural elements so that the *life safety* damage state, assumed herein to be δ_u , is not exceeded by the expected demand in terms of the performance point, δ_m . It is worth

Fig. 2 MIDR distribution using the response spectrum method



mentioning that in the assignation of the reinforcement, minimum steel percentages equal to 1% for columns and 0.5% for beams have been adopted.

Then, the calculated capacity curves are transformed into capacity spectra (see Fig. 3a) and the respective δ_m s have been obtained by considering the design spectrum described above; the procedure A of the ATC-40 (ATC-40 1996, see chapter 8) has been employed to obtain δ_m . Thus, it has been quantified if the expected demand exceeds the last capacity displacement of the structure (in terms of δ_u). To do so, the ratio between δ_m and δ_u has been calculated for each of the capacity spectra; it has been found that only twice this ratio is slightly higher than one (which indicates that δ_m exceeds to δ_u , see Fig. 3b). In other words, the vast majority of the structures generated do not exceed the limit state related to *life safety* for the design spectrum with a 475 years return period, which is an important requirement established in current design guidelines.

Based on the above verifications, it can be said that the structural models generated have been developed to meet the main design requirements of RC buildings located in earthquake-prone regions.

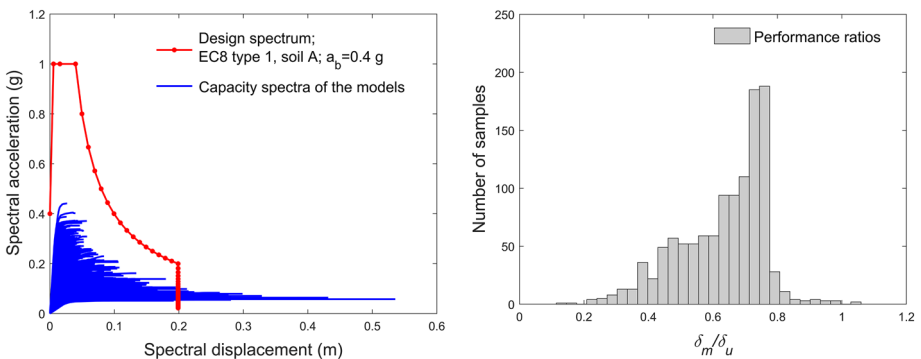


Fig. 3 Left: capacity spectra vs design spectrum type 1, soils A (Reference). Right: performance ratios

3 Seismic demand (hazard)

Seismic hazard is an important random variable that should be considered to properly estimate expected damage and risk. Actually, the main source of uncertainty when assessing the seismic response of civil structures is the variability of ground motions. The randomness in both seismic actions and, consequently, in the structural response, increases alongside with the severity of the earthquake.

Nowadays, there are excellent accelerometric databases at global, regional and country scale that have allowed to quantify uncertainties in seismic hazard. For instance, the *Istituto Nazionale di Geofisica e Vulcanologia*, INGV, provides the Engineering Strong Motion Database (ESM) and includes a link to the most important accelerometric databases in the world (Luzi et al. 2016a, b).

However, the availability of strong ground motion records covering high-intensity intervals at specific response periods is a common shortcoming found in current databases. Note that an improper sampling of ground motions can trigger excessive scaling of lower intensity records to fit them into high-intensity intervals. Besides, scaling the same record to different intensity levels would introduce false correlation between *IMs* and *EDPs*. Therefore, a proper selection of strong motions is paramount to the success of probabilistic structural analyses. The procedure employed to face these issues is detailed below.

There are several methodologies to properly select ground motion records from a database that, for instance, are consistent with a specific site-dependent spectral shape (Haselton et al. 2012). For the purpose of this study, the most important requirement is to have enough ground motion records which are scaled (if necessary) to analyse the generated structural models at different performance levels. Nonetheless, it is also important to avoid excessive scaling of the records so that bias introduced will be negligible. According to this, the main steps followed to pick and set up a proper suite of accelerograms are described in the following:

Step 1 Choose the strong ground motion database. The ESM database (Luzi et al. 2016a, b) compiled by the INGV has been used to select the ground motion records; this database contains records occurred in Europe

Step 2 Identify the *IM* for scaling the ground motion record set

Step 3 Define suitable intensity bands for the selected *IM*

Step 4 Compute the selected *IM* for each ground motion record

Step 5 Sort the records in descending order as a function of the calculated *IM* values

Step 6 If needed, the ground motion record with the highest *IM* value is scaled so that it falls into the highest intensity band; if this record naturally fulfills the interval condition, it is not scaled; this step is repeated with the subsequent records until having enough ground motion inputs belonging to this interval

Step 7 Step 6 is repeated for each interval considered; the scale factors estimated in STEP 6 (if any) are calculated so that *IM* values tend to be uniformly distributed within the intensity band

It seems reasonable to select (*Step 2*) an *IM* highly correlated with the structural response so that, by increasing or decreasing this variable, so does the structural response. Therefore, it is more appropriate to consider *IMs* calculated from the spectral response of *SDoF* systems, since they have been shown to be highly correlated with the *MIDR* (Miranda 2000). In this respect, several researchers have proven that it is more efficient

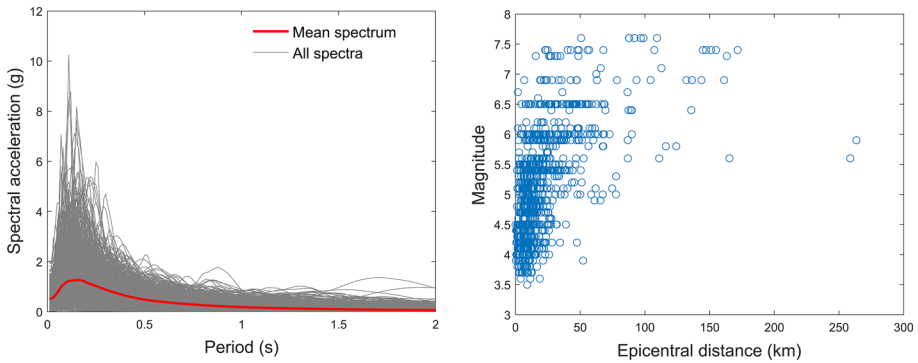


Fig. 4 Left: Response spectra of the 1160 selected and scaled accelerograms. Right: Distribution of magnitudes and epicentral distances

to use an *IM* based on the geometric mean of spectral acceleration values around the fundamental period, Sa_{avg} , than only using the spectral acceleration value associated to this period, $Sa(T_1)$, (Bianchini et al. 2009; Bojórquez and Iervolino 2011; Kazantzi and Vamvatsikos 2015; Eads et al. 2015; Adam et al. 2017). Analogous to this concept, the arithmetic mean of spectral acceleration values around the fundamental period, $AvSa$, has been considered herein as *IM* to select and scale ground motion records. It has been preferred to use the arithmetic mean than the geometric one since this tends to provide a more efficient and steadfast *IM*, as will be shown later.

Because of the probabilistic approach adopted in this study, there is not a single structural model but a group of them which, in addition, have highly variable fundamental periods (see Fig. 1). Therefore, the period range for averaging the spectral ordinates of the *IM* is established from the dynamic properties of the entire population of buildings (Vargas-Alzate et al., 2019b). In this way, ground motion records, whose mean spectral acceleration in the interval (0.2–1.6) s, lies in a band limited by two intensity levels are selected and scaled. This interval corresponds to 0.7 times the minimum and 1.12 times the maximum period observed in Fig. 1a; after several runs, they are the ones minimizing the dispersion in most of the *IM-EDP* clouds analysed. Note that these limits tend to be broader when analysing structures sharing similar number of stories.

The intensity levels (*Step 3*) defining the upper and lower limits of each band range from 0.035 to 0.7 g at intervals of 0.035 g (i.e., 20 bands are defined). The highest limit has been selected so that approximately 10% of the models exceed their capacity to withstand gravity loads.

By means of the selection and scaling process described above, the horizontal component related to the east–west direction of 1160 earthquake records (58 records per band) of the ESM database have been obtained (the entire database contains 1163 ground motion records). Figure 4 (Left) shows the response spectra of the 1160 selected records. The distribution of magnitudes, M , and epicentral distances, ED , of these records is shown in Fig. 4 (Right).

4 Intensity measures

The increased computing power and mathematical background along with proper numerical models allow to extract valuable information on variables related to strong seismic actions (*IMs*) as well as on the response of civil structures (*EDPs*). Ideally, an *IM* should contain enough information about the ground motion so that the structural response can be predicted with confidence (Pejovic and Jankovic 2015). Note that an *IM* may depend on the properties of the ground motion, or both the ground motion and the structural characteristics.

4.1 Spectral-based intensity measures

The dynamic response of structures subjected to ground motions induced by earthquakes has been correlated to the peak response of equivalent *SDoF* systems. The study of this simplified model gave rise not only to the response spectra but also to spectral *IMs*. It has been recognized that efficient *IMs* would be defined by response spectral ordinates. This is why response spectral ordinates are extensively used to quantify seismic hazard at a site. Such ordinates are calculated from the maximum time-history response after solving the dynamic equilibrium equation for *SDoF* systems given a set of specific period values:

$$m\ddot{u}(t) + c\dot{u}(t) + ku(t) = -m\ddot{u}_g(t) \quad (4)$$

where $\ddot{u}(t)$, $\dot{u}(t)$ and $u(t)$ are the acceleration, velocity and displacement time-history responses of the *SDoF*, respectively; $\ddot{u}_g(t)$ is the acceleration ground motion; m , c , and k represent mass, damping and stiffness of the system, respectively. *IMs* from 1 to 6 described in Table 2 belong to this category.

4.2 Energy-based intensity measures

For both design and assessing the performance of civil structures, the peak response of physical magnitudes like displacement, velocity or acceleration has been widely used. However, several researchers have found that the expected damage of structures is strongly tied to the amount of energy introduced to the system (see for instance Benavent-Climent et al. 2004; Yazgan 2012; Cheng et al. 2015; Güllü et al. 2019). In this respect, the equivalent velocity spectrum is a representation of the amount of energy introduced to a set of *SDoF* systems. This energy can be calculated by rewriting Eq. 4 in terms of energy. That is, each term of this equation is multiplied by the differential increment of displacement ($\dot{u}dt$) and then integrating in the time interval $(0, t)$, as follows (Akiyama 1985; Bertero and Uang 1992):

$$m \int_0^t \ddot{u} \dot{u} dt + c \int_0^t \dot{u}^2 dt + k \int_0^t u \dot{u} dt = -m \int_0^t \ddot{u}_g \dot{u} dt \quad (5)$$

where $E_k = m \int_0^t \ddot{u} \dot{u} dt$ is the kinetic energy; $E_\xi = c \int_0^t \dot{u}^2 dt$ is the energy dissipated by the inherent damping; $E_a = k \int_0^t u \dot{u} dt$ is the energy absorbed by the system; and $E_l = -m \int_0^t \ddot{u}_g \dot{u} dt$ is the energy introduced into the system by the ground motion. This

Table 2 Intensity measures description

Intensity measure	Id	Formula
Spectral acceleration at T_1	1	$Sa(T_1) = \max\left(\left \ddot{u}_g(t) + \ddot{u}(t, \xi, T_1)\right \right)$
Spectral velocity at T_1	2	$Sv(T_1) = \max(\dot{u}(t, \xi, T_1))$
Spectral displacement at T_1	3	$Sd(T_1) = \max(u(t, \xi, T_1))$
Average spectral acceleration ¹	4	$AvSa = \frac{\sum_{i=1}^{n_T} Sa(T_i)}{n_T}$
Average spectral velocity	5	$AvSv = \frac{\sum_{i=1}^{n_T} Sv(T_i)}{n_T}$
Average spectral displacement	6	$AvSd = \frac{\sum_{i=1}^{n_T} Sd(T_i)}{n_T}$
Equivalent velocity at T_1	7	$VE(T_1) = \sqrt{2E_f(T_1)}$
Average equivalent velocity	8	$AvVE = \frac{\sum_{i=1}^{n_T} VE(T_i)}{n_T}$
Peak ground acceleration	9	$PGA = \max\left(\ddot{u}_g(t) \right)$
Peak ground velocity	10	$PGV = \max\left(\dot{u}_g(t) \right)$
Peak ground displacement	11	$PGD = \max\left(u_g(t) \right)$
Arias intensity (Arias, 1970) ²	12	$I_A = \frac{\pi}{2g} \int_{t_{5\%}}^{t_{95\%}} \ddot{u}_g(t)^2 dt$
Root mean of the acceleration (Housner, 1975) ³	13	$acc_{RMS} = \sqrt{\frac{1}{\Delta} \int_{t_{5\%}}^{t_{95\%}} \ddot{u}_g(t)^2 dt}$
Root mean of the velocity	14	$vel_{RMS} = \sqrt{\frac{1}{\Delta} \int_{t_{5\%}}^{t_{95\%}} \dot{u}_g(t)^2 dt}$
Specific Energy Density (Sarma 1971; Sarma and Yang, 1987)	15	$SED = \int_{t_{5\%}}^{t_{95\%}} \dot{u}_g(t)^2 dt$
Cumulative Absolute Velocity (Reed, and Kassawara, 1990)	16	$CAV = \int_{t_{5\%}}^{t_{95\%}} \dot{u}_g(t) dt$
Characteristic intensity (Park et al., 1987)	17	$I_C = acc_{RMS}^{1.5} \sqrt{\Delta}$
Fajfar intensity (Fajfar et al. 1990; Pinzón et al. 2020)	18	$I_F = PGV * \Delta^{0.25}$

¹ n_T represents the number of periods to calculate the average

² $t_{5\%}$ and $t_{95\%}$ are the times related to the 5% and 95% of the Husid diagram (Trifunac and Novicova 1994).

³ Δ represents the significant duration (i.e. $t_{95\%} - t_{5\%}$)

latter is commonly expressed in terms of equivalent velocity, VE , and is normalized with respect to the mass of the structure as follows:

$$VE(T_1) = \sqrt{2E_f(T_1)} \tag{6}$$

Calculating VE for several elastic oscillators will produce the equivalent velocity spectra. IMs 7 and 8 described in Table 2 are obtained from the energy introduced into the system.

4.3 IMs based on direct computations of the ground motion record

IMs presented above take into account the spectral response of $SDoF$ systems. In other words, they consider not only the characteristics of ground motions but also some related to

the structural properties. However, several *IMs* can be obtained from direct computations of the ground motion record. The appeal of this type of *IMs* is that they do not depend on the dynamic properties of the structures. In this way, a generic fragility function could be developed to estimate the seismic performance of buildings with very different structural properties. However, this independency of the structural properties may cause a decrease of the efficiency for predicting *EDPs*. *IMs* from 9 to 18 described in Table 2 belong to this category.

5 Statistical analysis of *IM-EDP* pairs

Once the seismic hazard—intensity measures—and exposure—building models—have been defined and characterized, 1160 *NLDAs* have been performed. The RuauMoko software has been used to perform the structural analyses (Carr 2000).

It has been found that there is a number of simulations in which the capacity of the structure to withstand gravity loads is exceeded. In other words, the outcome of the analysis indicates structural collapse. These results are considered to be governed by chaos. In a few words, chaos is a phenomenon in which even deterministic models may lead to unpredictable outcomes due to the sensitivity of the governing dynamic equations to initial conditions. It has been observed that *MIDR* diverges after a collapse mechanism is activated. This fact agrees with the behavior of chaotic systems, which are predictable for a short time and then become excessively random. After analysing the response of building models exhibiting collapse, it has been noticed that this short time is in the order of a few hundredths of a second. A deeper view on chaotic systems and their properties can be found in Strogatz (2018). Thus, results related to collapse have been excluded from the statistical analysis. These outcomes have been identified as simulations providing damage index of Park and Ang higher than 1 (Park and Ang 1985). In recent and current research, this criterion has been shown to be very efficient in detecting structural collapses (Vargas et al. 2019b).

Note that, if collapses were identified as those samples exceeding a *MIDR* limit (or *MFA*), it is recommended to split data between censored and uncensored observations (see for instance Jalayer and Cornell 2009; Martins and Silva 2020). For the sake of simplicity, samples exhibiting collapse, according to the damage index limit described above, have been excluded from the statistical analysis. A potential problem of this simplification is that a significant number of the removed samples have very similar *IM* values, which may affect the statistical analysis (Aschheim et al. 2019). To check if this problem occurs, the coefficient of variations of each of the *IMs* related to the excluded samples have been calculated; the minimum observed value has been 0.23, which is significant with respect to the one of the entire population (0.56), meaning there is no concentration around specific *IM* values. Anyhow, note that the number of excluded simulations could be reduced if the superior limit of the *IM* for scaling the ground motion records is diminished. In addition, simulations in which $PGA > 1.5$ g, $PGV > 3$ m/s, $PGD > 1$ m or $Sa(T_1) > 4$ g have been also excluded from the statistical analyses. Thus, 119 ($\approx 10\%$) out of the 1160 generated samples have been excluded from the entire dataset. Consequently, 1041 cases are used to perform the statistical analysis of the *IMs*. Of this number, 141 are results in the elastic domain and 900 are results in the inelastic domain.

In the following, the resulting clouds of *IM-EDP* points are used to analyse several statistical properties related to the data variability.

5.1 Efficiency

Multi-regression models allow to define the optimal combination of several *IV*s to explain an *EDP*. In this study, *IV-EDP* relationships have been characterized by means of both multi-linear and -nonlinear regression models in the log–log space. As with linear least squares, nonlinear regression is based on determining values that minimize the sum of the squares of the residuals. In this sense, the following general linear least-square model allows several types of multi-regression analysis with respect to *y*:

$$y = \alpha_0 + \sum_{i=1}^{N_{IV}} \sum_{j=1}^n \alpha_{j+(i-1)*n} z_j + \varepsilon \tag{7}$$

where N_{IV} represents the number of *IV* considered in the regression model; n represents the polynomial degree of the model; $\alpha_0, \alpha_1, \dots, \alpha_{n*N_{IV}}$ are the coefficients providing the best fit between model and data; z_j are basic functions; ε represents the residuals. It can easily be seen how polynomial regression falls within this model. That is, $z_1 = x, z_2 = x^2, \dots, z_n = x^n$ (Chapra 2017). Substituting in Eq. 7 $y = \ln EDP$ and $z_j = (\ln IV_i)^j$, the general linear least-square model using polynomial functions can be used to extract statistical information from *IV-EDP* pairs according to the following equation (Eq. 8):

$$\ln EDP = \alpha_0 + \sum_{i=1}^{N_{IV}} \sum_{j=1}^n \alpha_{j+(i-1)*n} (\ln IV_i)^j + \varepsilon \tag{8}$$

This equation allows multi-linear ($n=1$) and -quadratic ($n=2$) regression models in the log–log space. In this way, several sources of information can be considered simultaneously to better predict a specific *EDP*. Further information regarding the development and implementation of this type of polynomial models in the log–log space can be found in Chapra (2017).

Thus, for $N_{IV} = 1$ Eq. 8 adopts the following form:

$$\ln EDP = \alpha_0 + \sum_{j=1}^n \alpha_j (\ln IM)^j + \varepsilon \tag{9}$$

IV_1 becomes *IM* in this equation.

In terms of the residuals, ε , there are several variables that can be used to quantify the variability of *IM-EDP* clouds. For instance, Ebrahimian et al. (2015) considered the logarithmic standard deviation of the regression, $S_{y/x}$, and the correlation coefficient, R^2 , to quantify the efficiency of *IMs* (for a perfect fit, $S_{y/x} = 0$ and $R^2 = 1$, signifying that the adopted function explains 100% of the data variability). These two variables (R^2 and $S_{y/x}$) are used herein to provide an estimation of the variability when analysing relationships involving *IM-EDP* pairs. The higher R^2 or the lower $S_{y/x}$, the lower the variability when predicting some *EDP* given an *IM*.

A more well-known and structure-specific measure of efficiency is the logarithmic standard deviation of the fragility curve, β , called proficiency (i.e., logarithmic standard deviation of the regression divided by the slope of regression; see Padgett et al. (2008) and Ebrahimian and Jalayer (2021)). This measure has been also calculated to analyse the efficiency of the *IMs*. Summarizing, the *IM* providing the highest R^2 and the lowest $S_{y/x}$ and β is the most efficient one, in the sense that it allows reducing the variability in seismic risk

Table 3 Summary of the statistical analysis considering as *EDP* to *MIDR*

<i>IM</i>	R_L^2	R_{NL}^2	S_{y/x_L}	$S_{y/x_{NL}}$	β	pV_M	pV_{ED}	η_{3-13}
<i>Sa</i> (T_1)	0.521	0.550	0.686	0.665	1.248	0.00E+00	1.46E-03	0.471
<i>Sv</i> (T_1)	0.832	0.834	0.406	0.404	0.411	2.60E-10	1.01E-03	0.895
<i>Sd</i> (T_1)	0.810	0.814	0.431	0.427	0.552	3.80E-03	1.97E-01	0.987
<i>AvSa</i>	0.648	0.652	0.588	0.585	0.631	0.00E+00	2.58E-06	0.527
<i>AvSv</i>	0.886	0.890	0.334	0.329	0.295	6.83E-10	1.47E-04	0.954
<i>AvSd</i>	0.852	0.854	0.380	0.379	0.451	6.53E-01	7.09E-01	0.964
<i>VE</i> (T_1)	0.779	0.791	0.465	0.453	0.541	1.62E-01	4.10E-01	0.809
<i>AvVE</i>	0.855	0.855	0.377	0.377	0.365	1.90E-01	1.48E-02	0.873
<i>PGA</i>	0.459	0.489	0.728	0.708	0.654	0.00E+00	0.00E+00	1.015
<i>PGV</i>	0.850	0.850	0.384	0.384	0.341	1.76E-10	4.03E-06	0.984
<i>PGD</i>	0.721	0.775	0.523	0.470	0.833	6.31E-06	1.02E-02	0.986
I_A	0.700	0.701	0.543	0.542	0.863	1.92E-10	8.46E-02	0.997
acc_{RMS}	0.444	0.482	0.738	0.713	0.741	0.00E+00	0.00E+00	1.016
vel_{RMS}	0.840	0.840	0.397	0.396	0.377	9.52E-07	1.65E-08	0.975
<i>SED</i>	0.825	0.837	0.414	0.401	0.911	9.02E-15	5.66E-11	0.966
<i>CAV</i>	0.655	0.675	0.582	0.565	0.593	9.73E-02	2.06E-10	0.992
I_c	0.570	0.581	0.649	0.641	0.850	0.00E+00	1.06E-12	1.003
I_F	0.851	0.852	0.383	0.382	0.364	1.05E-01	5.88E-03	0.979

Results corresponding to the most efficient *IM*, *AvSv*, have been highlighted; *p* values higher than 0.05, meaning that the *IM* is sufficient with respect to *M* or *ED*, have also been highlighted

estimations. In the following, variables described in section 4 are used as basic *IMs* whilst *MIDR* as *EDP*.

Results from two types of regression models have been compared (linear and non-linear, i.e. $n=1$ and $n=2$ in Eq. 9, respectively). The reason for considering a quadratic regression model is that several *IM-EDP* pairs seem to exhibit a non-linear dependence. In this regard, several authors have called into question the use of linear regression when characterizing these pairs (e.g. Baker 2007). It has been stated that a linear relationship may not be reasonable for the entire *IM* range of interest. To address this problem, the regression should be limited to the *IM* range where the linear form is reasonable, or piecewise linear relationships may be fitted (e.g. Mackie and Stojadinovic 2003). Herein, it has been opted to fit a quadratic regression model. It is worth mentioning that, when an *IM-EDP* set indicates that a linear fitting is more adequate to characterize this relationship in the entire interval, the quadratic model automatically makes the coefficient accompanying the nonlinear term close to zero.

Figure 5 shows the relationships between *IM-EDP* pairs in the log–log space for all the basic *IMs* described in Table 2. R^2 values for the linear (R_L^2) and nonlinear regression models (R_{NL}^2) are presented in this figure. Table 3 summarizes, and also complements, results shown in Fig. 5; subscripts ‘*L*’ and ‘*NL*’ in R^2 and $S_{y/x}$ variables stand for linear and non-linear regression model.

It can be observed that *IMs* related to the velocity are the ones most efficient to predict the *MIDR* of the analysed systems. It is also important to note that most of these variables are well fitted by both linear and nonlinear regression models. In other words, it has not

been observed a significant variation of the dispersion when using any of these models. In any case, note that the polynomial arrangement presented in Eq. 9 allows fitting linear functions if that is what the data suggests; for these cases, α_2 tends to zero. Notwithstanding, although similar levels of variability are observed for most of the *IMs* used (i.e. between linear and nonlinear regression models), it will be shown that including the non-linear term in the regression model strongly influences the derivation of analytical fragility functions. It can also be seen in Table 3 that *AvSv* is the *IM* most efficient to explain the *MIDR*.

5.2 Sufficiency

A sufficient *IM* makes the statistical properties of a cloud of *IM-EDP* pairs independent of other parameters related to the rupture (e.g., magnitude, *M*, epicentral distance, *ED*, etc.). This property can be defined through the correlation between the residuals of a cloud of *IM-EDP* pairs and the seismological parameter of interest, *SP*; the higher this correlation the lower the sufficiency of the *IM*.

The *p* value between the aforementioned residuals and *SPs* has been used to verify the sufficiency of the *IMs* presented in Fig. 5 (Luco and Cornell 2007). According to this, the *p* value between the residuals of each of the *IM-EDP* clouds have been calculated with respect to the magnitude (pv_M) and epicentral distance (pv_{ED}); Table 3 also shows these results. It can be seen that most of the *IMs* present *p* values lower than 0.05, meaning that they are not sufficient with respect to the *SP*. These *p* values have been calculated

Table 4 Summary of the statistical analysis considering as *EDP* to *MIDR*

<i>IM*</i>	R_L^2	R_{NL}^2	Sy/x_L	Sy/x_{NL}	β	pv_M	pv_{ED}	η_{3-13}
[<i>Sa</i> (<i>T</i> ₁), <i>M</i>]	0.570	0.604	0.650	0.625	1.446	9.44E-01	2.73E-02	0.425
[<i>Sv</i> (<i>T</i> ₁), <i>M</i>]	0.840	0.845	0.397	0.391	0.427	8.21E-01	6.97E-01	0.891
[<i>Sd</i> (<i>T</i> ₁), <i>M</i>]	0.812	0.820	0.430	0.421	0.569	7.62E-01	7.48E-01	0.981
[<i>AvSa</i> , <i>M</i>]	0.693	0.701	0.549	0.543	0.683	9.88E-01	1.54E-01	0.505
[<i>AvSv</i> , <i>M</i>]	0.891	0.894	0.327	0.323	0.303	7.52E-01	7.96E-01	0.957
[<i>AvSd</i> , <i>M</i>]	0.853	0.856	0.381	0.377	0.447	7.56E-01	9.14E-01	0.955
[<i>VE</i> (<i>T</i> ₁), <i>M</i>]	0.780	0.794	0.465	0.451	0.553	9.71E-01	8.75E-02	0.794
[<i>AvVE</i> , <i>M</i>]	0.855	0.856	0.377	0.377	0.359	9.84E-01	1.00E-01	0.868
[<i>PGA</i> , <i>M</i>]	0.655	0.669	0.582	0.571	0.594	9.11E-01	2.81E-01	1.012
[<i>PGV</i> , <i>M</i>]	0.858	0.859	0.374	0.373	0.353	8.27E-01	5.58E-01	0.968
[<i>PGD</i> , <i>M</i>]	0.730	0.775	0.515	0.470	0.732	5.47E-01	8.05E-01	0.990
[<i>I</i> _A , <i>M</i>]	0.716	0.721	0.528	0.524	0.931	9.79E-01	3.62E-02	1.007
[<i>acc</i> _{<i>RMS</i>} , <i>M</i>]	0.643	0.661	0.593	0.578	0.678	8.80E-01	9.89E-01	1.014
[<i>vel</i> _{<i>RMS</i>} , <i>M</i>]	0.845	0.847	0.390	0.389	0.391	9.22E-01	1.39E-02	0.964
[<i>SED</i> , <i>M</i>]	0.843	0.848	0.392	0.387	0.748	6.36E-01	5.09E-02	0.955
[<i>CAV</i> , <i>M</i>]	0.656	0.676	0.581	0.565	0.565	8.03E-01	2.28E-08	1.004
[<i>I</i> _c , <i>M</i>]	0.666	0.678	0.573	0.563	0.888	8.81E-01	2.03E-01	1.012
[<i>I</i> _F , <i>M</i>]	0.851	0.854	0.383	0.379	0.358	6.69E-01	4.34E-02	0.970

Results corresponding to the most efficient *IM**, [*AvSv*, *M*], have been highlighted; *p* values higher than 0.05, meaning that the *IM** is sufficient with respect to *M* or *ED*, have also been highlighted

Table 5 Summary of the statistical analysis considering as *EDP* to *MIDR*

<i>IM*</i>	R_L^2	R_{NL}^2	Sy/x_L	Sy/x_{NL}	β	pV_M	pV_{ED}	η_{3-13}
[<i>Sa</i> (T_1), <i>ED</i>]	0.524	0.556	0.684	0.662	1.261	5.70E–14	2.40E–01	0.470
[<i>Sv</i> (T_1), <i>ED</i>]	0.834	0.836	0.404	0.402	0.411	6.94E–06	4.78E–01	0.897
[<i>Sd</i> (T_1), <i>ED</i>]	0.811	0.815	0.432	0.427	0.554	1.66E–02	5.78E–01	0.989
[<i>AvSa</i> , <i>ED</i>]	0.655	0.660	0.583	0.579	0.633	1.46E–15	1.81E–01	0.523
[<i>AvSv</i> , <i>ED</i>]	0.888	0.891	0.332	0.327	0.295	6.22E–05	4.65E–01	0.956
[<i>AvSd</i> , <i>ED</i>]	0.853	0.854	0.381	0.380	0.450	9.89E–01	7.73E–01	0.968
[<i>VE</i> (T_1), <i>ED</i>]	0.780	0.791	0.465	0.454	0.537	2.35E–02	6.10E–01	0.810
[<i>AvVE</i> , <i>ED</i>]	0.857	0.857	0.375	0.375	0.360	7.22E–01	8.66E–01	0.871
[<i>PGA</i> , <i>ED</i>]	0.514	0.536	0.691	0.676	0.601	0.00E+00	9.72E–02	1.014
[<i>PGV</i> , <i>ED</i>]	0.853	0.853	0.380	0.380	0.340	3.51E–04	5.02E–01	0.976
[<i>PGD</i> , <i>ED</i>]	0.725	0.775	0.520	0.470	0.812	3.48E–03	8.19E–01	0.983
[I_A , <i>ED</i>]	0.700	0.702	0.543	0.542	0.867	5.81E–09	3.47E–01	1.000
[<i>acc</i> _{<i>RMS</i>} , <i>ED</i>]	0.511	0.538	0.693	0.675	0.663	0.00E+00	5.52E–02	1.012
[<i>vel</i> _{<i>RMS</i>} , <i>ED</i>]	0.844	0.845	0.391	0.391	0.375	6.47E–02	1.80E–01	0.969
[<i>SED</i> , <i>ED</i>]	0.839	0.846	0.398	0.390	0.840	2.74E–05	9.64E–01	0.961
[<i>CAV</i> , <i>ED</i>]	0.681	0.692	0.560	0.550	0.525	1.27E–01	9.09E–01	0.989
[I_c , <i>ED</i>]	0.591	0.601	0.634	0.627	0.835	0.00E+00	1.25E–01	1.007
[I_F , <i>ED</i>]	0.852	0.853	0.381	0.380	0.358	9.57E–01	9.01E–01	0.973

Results corresponding to the most efficient *IM**, [*AvSv*, *ED*], have been highlighted; *p* values higher than 0.05, meaning that the *IM** is sufficient with respect to *M* or *ED*, have also been highlighted

with respect to the residuals provided by the linear regression model. However, similar conclusions can be drawn when performing this operation using the quadratic regression approach.

Regarding the group of *IMs* exhibiting sufficiency, only *AvSd* and *VE*(T_1) meet this condition simultaneously with respect to *M* and *ED*; the most efficient of them is *AvSd*.

Lack of sufficiency indicates that some *IMs* can gain efficiency when combined in a multi-regression model (see Eq. 8) with the respective seismological parameter (Elefante et al. 2010). This is analogous to the concept of vector-valued *IMs* (Baker 2007). Several researchers have applied this sophistication to develop enhanced *IMs* (*IM**), which consist in a vector of *IVs* providing more efficient estimates of *EDPs* (Modica and Stafford 2014; Yakhchalian et al. 2015, 2021; Bojórquez et al. 2015; Ge and Zhou 2018; Zengin and Abrahamson 2020). The amount of efficiency that can be gained by an *IM* depends on the correlation value between the residuals of the *IM-EDP* cloud and *SP*. The higher this correlation the higher the efficiency that can be gained.

Thus, for $n=1$ and $N_{IV}=2$, Eq. 8 can be used to include *SPs* in the linear regression analysis as follows:

$$\ln(EDP) = \alpha_0 + \alpha_1 * \ln(IM) + \alpha_2 * \ln(SP) + \epsilon \tag{10}$$

For $n=2$ and $N_{IV}=2$, Eq. 8 adopts the following form:

$$\ln(EDP) = \alpha_0 + \alpha_1 * \ln(IM) + \alpha_2 * (\ln(IM))^2 + \alpha_3 * \ln(SP) + \alpha_4 * (\ln(SP))^2 + \epsilon \tag{11}$$

Table 6 Summary of the statistical analysis considering as *EDP* to *MIDR*

<i>IM</i> *	R_L^2	R_{NL}^2	Sy/x_L	Sy/x_{NL}	β	pV_M	pV_{ED}	η_{3-13}
[<i>Sa</i> (T_1), <i>M</i> , <i>ED</i>]	0.579	0.613	0.644	0.618	1.493	9.83E-01	5.62E-01	0.415
[<i>Sv</i> (T_1), <i>M</i> , <i>ED</i>]	0.840	0.846	0.397	0.391	0.430	8.11E-01	6.87E-01	0.893
[<i>Sd</i> (T_1), <i>M</i> , <i>ED</i>]	0.813	0.821	0.430	0.421	0.573	7.54E-01	6.87E-01	0.979
[<i>AvSa</i> , <i>M</i> , <i>ED</i>]	0.697	0.706	0.546	0.539	0.696	9.58E-01	4.64E-01	0.497
[<i>AvSv</i> , <i>M</i> , <i>ED</i>]	0.891	0.894	0.327	0.323	0.304	7.50E-01	6.41E-01	0.958
[<i>AvSd</i> , <i>M</i> , <i>ED</i>]	0.853	0.856	0.381	0.377	0.448	7.54E-01	7.55E-01	0.952
[<i>VE</i> (T_1), <i>M</i> , <i>ED</i>]	0.783	0.797	0.463	0.448	0.560	1.00E+00	7.31E-01	0.795
[<i>AvVE</i> , <i>M</i> , <i>ED</i>]	0.857	0.857	0.376	0.375	0.362	9.90E-01	8.85E-01	0.867
[<i>PGA</i> , <i>M</i> , <i>ED</i>]	0.658	0.675	0.580	0.567	0.607	8.88E-01	4.82E-01	1.019
[<i>PGV</i> , <i>M</i> , <i>ED</i>]	0.858	0.859	0.374	0.373	0.353	8.24E-01	6.91E-01	0.971
[<i>PGD</i> , <i>M</i> , <i>ED</i>]	0.730	0.775	0.515	0.471	0.735	5.50E-01	6.78E-01	0.994
[I_A , <i>M</i> , <i>ED</i>]	0.722	0.729	0.523	0.517	0.946	9.84E-01	6.12E-01	1.012
[<i>acc</i> _{RMS} , <i>M</i> , <i>ED</i>]	0.644	0.664	0.592	0.576	0.690	8.69E-01	3.24E-01	1.021
[<i>vel</i> _{RMS} , <i>M</i> , <i>ED</i>]	0.846	0.847	0.390	0.389	0.385	9.03E-01	2.40E-01	0.969
[<i>SED</i> , <i>M</i> , <i>ED</i>]	0.846	0.850	0.390	0.385	0.754	6.68E-01	7.45E-01	0.959
[<i>CAV</i> , <i>M</i> , <i>ED</i>]	0.684	0.702	0.558	0.542	0.553	8.79E-01	8.33E-01	1.001
[I_c , <i>M</i> , <i>ED</i>]	0.671	0.686	0.570	0.557	0.907	8.56E-01	4.48E-01	1.017
[I_F , <i>M</i> , <i>ED</i>]	0.853	0.855	0.381	0.378	0.361	6.96E-01	8.78E-01	0.971

Results corresponding to the most efficient *IM**, [*AvSv*, *M*, *ED*], have been highlighted; *p* values higher than 0.05, meaning that the *IM** is sufficient with respect to *M* or *ED*, have also been highlighted

According to these arrangements, R^2 , $S_{y/x}$ and β are newly calculated by substituting *SP* for *M* and *ED* in Eqs. 10 and 11; Tables 4 and 5 show these results, respectively. Note that the proficiency terms (β) coming from multiple regression models have been calculated by considering that the slope of the linear regression analysis is α_1 in Eqs. 10 and 11.

It is important to note that several of the *IM**s considering to *M* in the arrangement presented in Eqs. 10 or 11 become sufficient with respect to *ED*; this is not that clear when considering *ED* in these equations (i.e. the resultant vector-valued *IM* is not sufficient with respect to the magnitude, see Table 5).

According to Tables 4 and 5, some *IM**s could gain even more efficiency if an enhanced arrangement combining both *M* and *ED* together with the basic *IM* is used (i.e., $N_{IV}=3$; $n = 1, 2$ in Eq. 8). For $n = 1$, the following linear regression model has been employed:

$$\ln(EDP) = \alpha_0 + \alpha_1 * \ln(IM) + \alpha_2 * \ln(M) + \alpha_3 * \ln(ED) + \epsilon \tag{12}$$

for the quadratic regression case (i.e. $n = 2$):

$$\begin{aligned} \ln(EDP) = & \alpha_0 + \alpha_1 * \ln(IM) + \alpha_2 * (\ln(IM))^2 + \alpha_3 * \ln(M) \\ & + \alpha_4 * (\ln(M))^2 + \alpha_5 * \ln(ED) + \alpha_6 * (\ln(ED))^2 + \epsilon \end{aligned} \tag{13}$$

Table 6 shows the statistical properties of the *IM**s by using the regression models presented in Eqs. 12 and 13. *p* values have been calculated again by considering the linear regression model. β has been calculated by using α_1 as the slope of the linear regression

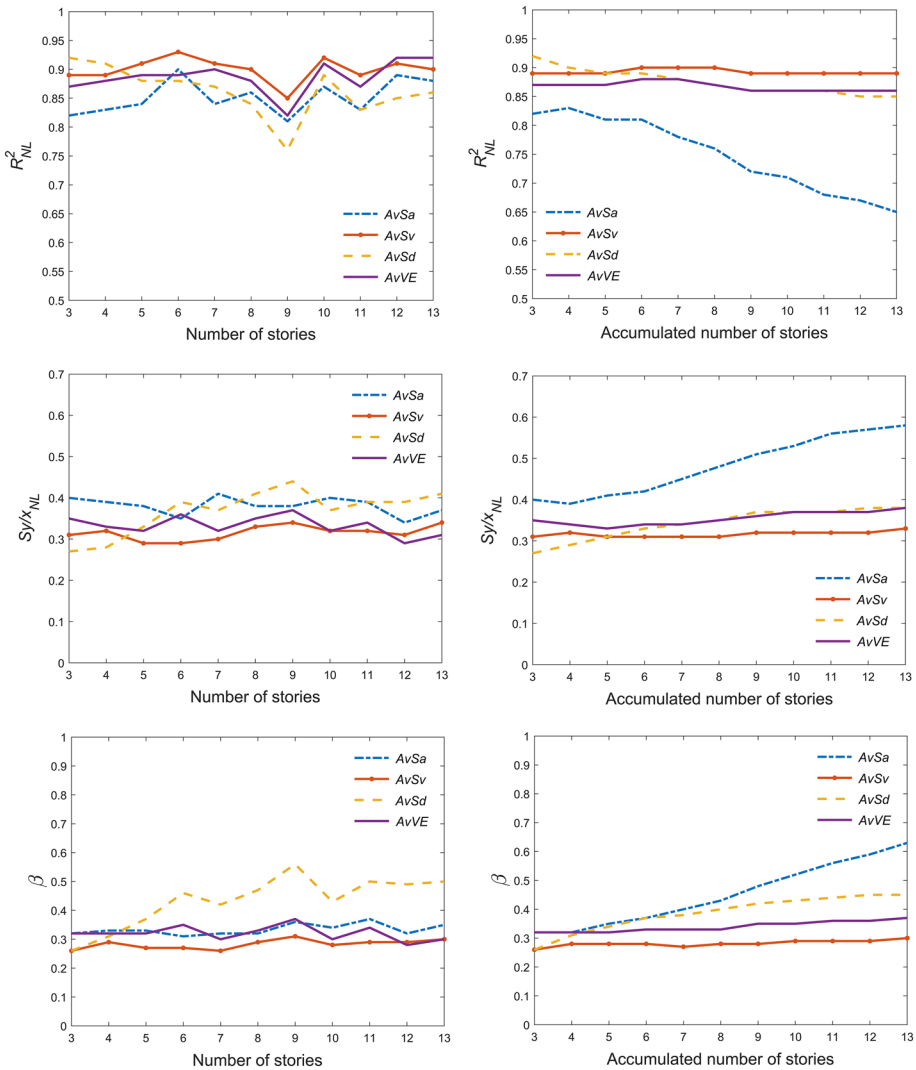


Fig. 6 Steadfastness analysis. Left: R^2 , $S_{y/x}$ and β values for buildings grouped according to the number of storeys. Right: R^2 , $S_{y/x}$ and β values for buildings grouped according to the accumulated number of stories. The EDP analysed has been MIDR

arrangement. It is worth noting that, for all the multi-regression models used, AvSv has been present in the vector that forms the most efficient IM^* .

5.3 Steadfastness

In Table 3, it can be seen that, except for AvSa, IMs defined in terms of average spectral values are best-correlated with MIDR. In general, IMs related to the spectral acceleration

do not correlate well with *MIDR*. To further analyse this somehow unexpected result, data- efficiency measures-have been recalculated by creating sub-categories according to the number of stories (see Fig. 6, left). Hence, R^2 , $S_{y/x}$ and β values have been obtained by using the quadratic regression model.

A significant increase of R^2 , as well as a reduction of $S_{y/x}$ and β values, can be seen for *AvSa*. In other words, the efficiency of *AvSa* has increased considerably, yet *AvSv* is best correlated with *MIDR* for most of the disaggregated structures. For low-rise structures—3- and 4-storey buildings—the most efficient *IM* is *AvSd*; for high-rise buildings—12- and 13-stories—*AvVE* is the most efficient one.

The fact that, when simultaneously analysing *IM-EDP* pairs coming from building models with different number of stories, an *IM* loses efficiency is a potential problem, since a higher number of structural calculations is required for obtaining reliable seismic risk estimations at urban level. In this article, when an *IM* or an *IM** does not lose efficiency when aggregating results of buildings with different properties, it is catalogued as steadfast. Another way to see steadfastness is to think about sufficiency, that is, a steadfast *IM* is sufficient with respect to the parameter analysed, in this case, number of stories.

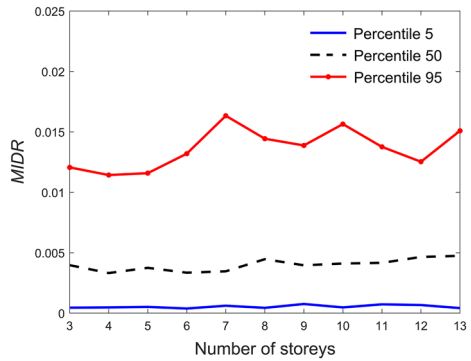
To analyse the steadfastness of *IMs* based on average spectral values, buildings with different number of stories have been progressively aggregated. Figure 6 (Right) shows the evolution of R^2 , $S_{y/x}$ and β values when aggregating buildings whose number of stories increases. In other words, an abscissa value corresponds to those buildings whose number of stories is equal to or less than it. An abscissa of 9, for instance, represents a subgroup of buildings having 9 stories or less. This way, R^2 , $S_{y/x}$ or β corresponding to an abscissa value of 3 are the same in Fig. 6, left and right. Also, when the abscissa is 13 in Fig. 6 (Right), parameters related to the efficiency are the same as those shown in Table 3 for the corresponding *IM*.

In Fig. 6 (Right), it can also be seen that *AvSa* shows certain steadfastness when buildings holding 3 to 6-stories are grouped. This result is consistent with the approach followed by several methods for seismic risk analysis which groups, in different structural types, those buildings sharing similar number of stories. For instance, in previous seismic risk studies (see, for instance, Lantada et al. 2009), RC Low-, Mid- and High-rise building types, having 1 to 3, 4 to 6 and more than 6 stories, respectively, have been used to perform seismic risk estimations. This sort of grouping allows using the same fragility function for every building belonging to the same typology when estimating, for instance, the exceedance probability of a certain damage level.

AvSv, *AvSd* and *AvVE* are variables less prone to lose efficiency when grouping structures having different number of stories. Therefore, these *IMs* are candidates for considerably reducing the number of structural calculations when estimating seismic risk, since the same fragility function may represent an important variety of buildings (even the entire set analysed in this research, if desired) no matter the number of stories. For instance, classical Low- Mid- and High-rise typologies could be grouped within a comprehensive structural typology based on *AvSv*, which is not only the most efficient but also one of the *IMs* showing stability; it means that this *IM* is steadfast.

A simplified way to measure steadfastness is to quantify how much the efficiency decreases when grouping a set of buildings of different properties. To do so, the proficiency measure, β , has been selected in this article to estimate steadfastness. Thus, for a set of buildings, whose number of stories vary from I to S , the ratio between the average of the proficiency values for buildings grouped according to the number of stories, $\bar{\beta}_{NST}$, and the proficiency estimated from the entire set of buildings, β_{I-S} , is proposed as a measure of steadfastness (see Eq. 13).

Fig. 7 Percentiles 5, 50 and 95 for the *MIDR* distribution as a function of the number of stories



$$\eta_{I-S} = \frac{\bar{\beta}_{N_{ST}}}{\beta_{I-S}} \tag{14}$$

where $\beta_{N_{ST}}$ is a vector of *I-S+I* elements containing β values calculated by grouping the structures according to the number of stories (see Fig. 6, left). Thus, η_{3-13} provides an estimation of the amount of efficiency losses by an *IM*, when grouping buildings from *I*=3 to *S*=13 stories, with respect to subgroups of structures sharing the same number of stories. In this way, if η_{I-S} is close to one, it means that the *IM* is steadfast for grouping buildings from *I* to *S* number of stories.

Tables 3, 4, 5 and 6 (last columns) show the steadfastness estimation for all the basic *IMs* described in Sect. 4, and their vector-valued versions. It can be seen that *Sa*(*T*₁) and *AvSa* are the *IMs* showing the lowest steadfastness of the analysed group.

It is worth mentioning that the scaling criteria used in this article (see Sect. 3) allows reaching similar demand levels for building models grouped according to number of stories. Meaning, if percentiles 5, 50 and 95 of the *MIDR* distribution are calculated by grouping data based on the number of stories, similar values are observed (see Fig. 7).

5.4 *Sa*_{avg} versus *AvSa*

Several researchers have proven that it is more efficient to use an *IM* based on the geometric mean of spectral acceleration values around the fundamental period, *Sa*_{avg}, than *Sa*(*T*₁), (Bianchini et al. 2009; Bojórquez and Iervolino 2011; Kazantzi and Vamvatsikos 2015; Eads et al. 2015; Adam et al. 2017). Note that *Sa*_{avg} is analogous to *AvSa* since both *IMs* use averaged spectral acceleration ordinates over a period range around the fundamental period of the structure. The main difference between them is that *Sa*_{avg} uses the geometric mean whilst *AvSa* the arithmetic one. In this section, both *IMs* are compared in terms of efficiency and steadfastness. Figure 8 (Left) shows *R*², *S*_{*y/x*} and β values by grouping buildings according to the number of stories for *AvSa* and *Sa*_{avg}.

In terms of *R*² and *S*_{*y/x*}, it can be seen that for low-rise buildings (from 3 to 5 stories), the *MIDR* tends to exhibit less variability when predicted by *Sa*_{avg}; for higher buildings, *AvSa* is the *IM* providing less variability. In terms of β , only for 3-storey buildings *Sa*_{avg} produces less dispersion than *AvSa*. Regarding steadfastness, Fig. 8 (Right) shows the evolution of the efficiency-related parameters when grouping buildings of different number of stories; all variables indicate that *Sa*_{avg} loses more efficiency than *AvSa*.

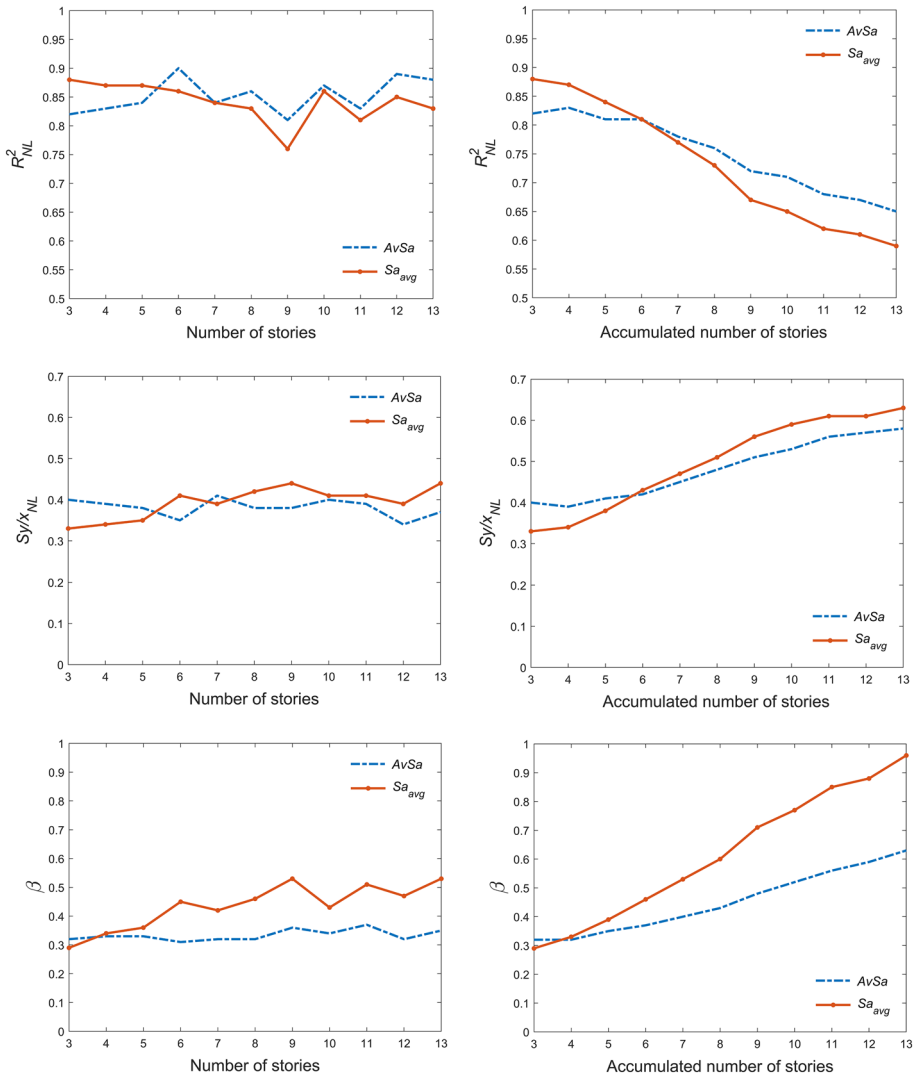


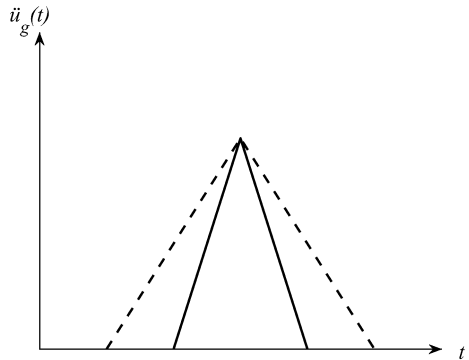
Fig. 8 Efficiency and steadfastness comparison between $AvSa$ and Sa_{avg}

The comparison between the use of arithmetic versus geometric mean has also been performed by considering $AvSd$, $AvSv$ and $AvVE$; similar conclusions have been found. For these reasons, throughout this article, it has been preferred to analyse the evolution of average-based IMs using the arithmetic mean rather than the geometric one.

5.5 ¿Why velocity-related IMs are more efficient to predict the MIDR?

Along this study, it has been shown that velocity- and energy-based IMs have the highest explanatory capacity in the above statistical relationships; that is, the maximum efficiency and steadfastness to predict $MIDR$. It seems reasonable to wonder why these IMs are better

Fig. 9 Two single peak triangular acceleration inputs; same PGA, different durations



proxies than the rest of them. To think about this question, let us consider two single peak triangular acceleration inputs, $\ddot{u}_g(t)$, having the same *PGA* but different durations (Fig. 9).

While the maximum input force is the same, i.e. *PGA* times the mass m , the total input force is:

$$\int_{t_1}^{t_2} m\ddot{u}_g dt = \int_0^{PGV} m d\dot{u}_g = mPGV \tag{15}$$

This integral equals the *PGV* that is obtained at the end of the interval, multiplied by the mass. Therefore, despite the maximum force for the two peaks is the same, the total input force is larger for the one having the largest *PGV*, which corresponds to the dashed line in Fig. 9. In other words, while *PGA* is associated to the maximum force, *PGV* represents the total input force; in this case, in a half cycle. Since the deformation of a structure is more related to the total input force, it is to be expected that *IMs* based on velocity are better predictors of the *MIDR* than those based on acceleration. Further insight into this fact and into their mutual relationship can be gained by considering the input energy, E_I , which is rewritten herein for the sake of convenience:

$$E_I = -m \int_0^t \ddot{u}_g \dot{u} dt \tag{16}$$

Integrating by parts Eq. 16 yields:

$$E_I = -m \left[\dot{u}_g \dot{u} - \int_0^t \dot{u}_g \ddot{u} dt \right] \tag{17}$$

which can be put in the form:

$$E_I = -m \left[\dot{u}_g \dot{u} - \int_0^t \dot{u}_g d\dot{u} \right] \tag{18}$$

since $d\dot{u} = \ddot{u} dt$.

In Eq. 18, the first term is an oscillatory function around zero energy. The backbone curve of the energy is constituted by the integral term, which gives the increasing component of the energy input. Recall that \dot{u}_g is related to the ground motion while \dot{u} is related to the structural response. Be it as it may, the meaning of the last presentation of the input energy with its two terms is clear: *the energy input can be expressed solely in terms of the*

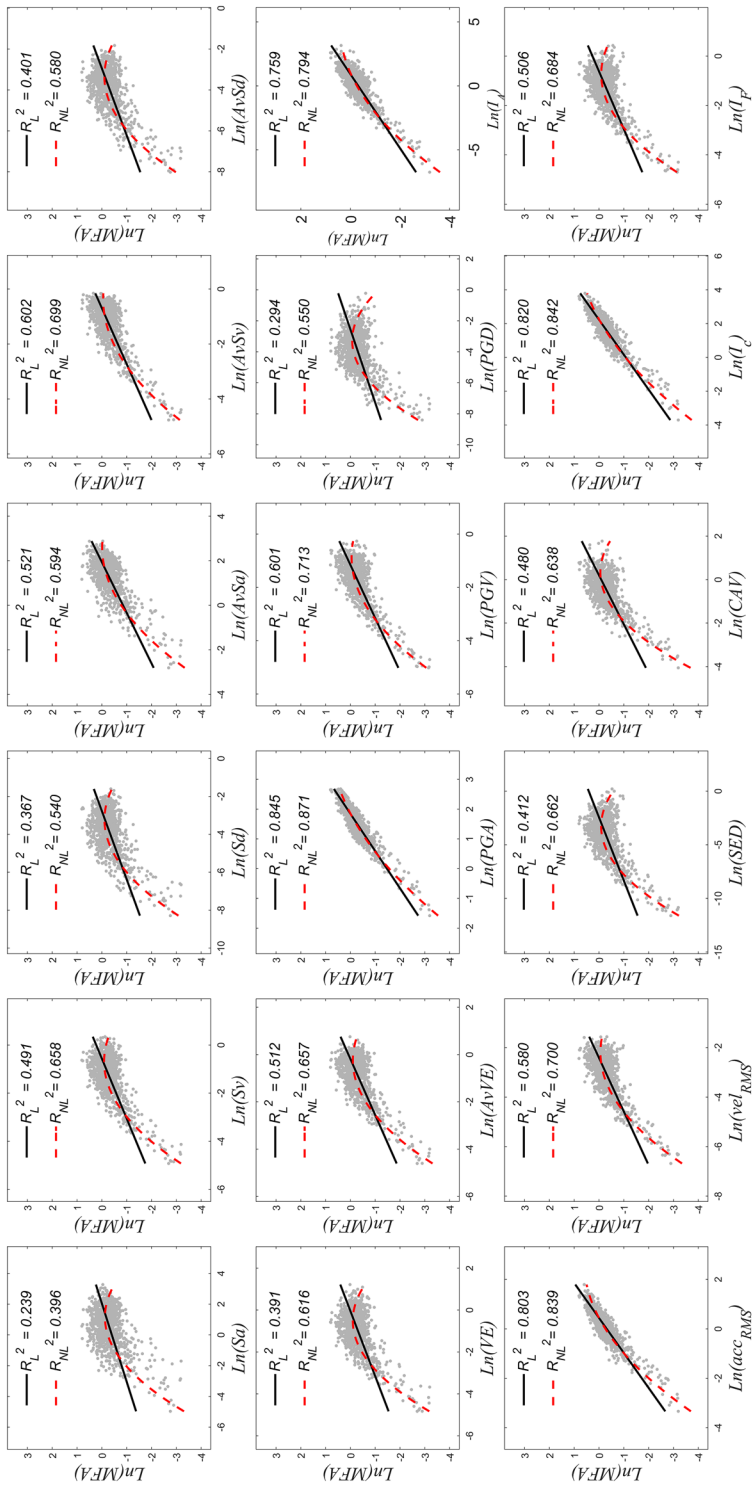


Fig. 10 R^2 calculated with both linear and nonlinear regression models; MFA has been considered as EDP

Table 7 Summary of the statistical analysis considering as *EDP* to *MFA*

<i>IM</i>	R_L^2	R_{NL}^2	Sy/x_L	Sy/x_{NL}	β	pV_M	pV_{ED}	η_{3-13}
<i>Sa</i> (T_1)	0.239	0.396	0.456	0.407	2.321	4.20E-01	2.99E-03	0.639
<i>Sv</i> (T_1)	0.491	0.658	0.373	0.306	0.929	2.66E-10	5.69E-07	0.965
<i>Sd</i> (T_1)	0.367	0.540	0.416	0.355	1.497	8.02E-07	3.48E-06	1.021
<i>AvSa</i>	0.521	0.594	0.362	0.333	0.822	1.97E-06	4.68E-05	0.662
<i>AvSv</i>	0.602	0.699	0.330	0.287	0.671	0.00E+00	8.04E-10	0.992
<i>AvSd</i>	0.401	0.580	0.405	0.339	1.325	2.81E-11	3.94E-08	1.005
<i>VE</i> (T_1)	0.391	0.616	0.408	0.324	1.268	2.19E-10	5.26E-09	0.924
<i>AvVE</i>	0.512	0.657	0.365	0.307	0.866	0.00E+00	9.84E-15	0.965
<i>PGA</i>	0.845	0.871	0.206	0.188	0.258	2.07E-08	3.48E-05	0.990
<i>PGV</i>	0.601	0.713	0.330	0.281	0.659	0.00E+00	2.53E-08	0.988
<i>PGD</i>	0.294	0.550	0.439	0.351	2.078	1.10E-11	1.51E-08	0.993
I_A	0.759	0.794	0.257	0.238	0.743	0.00E+00	0.00E+00	0.985
<i>acc</i> _{RMS}	0.803	0.839	0.232	0.210	0.328	5.39E-08	2.10E-09	0.989
<i>vel</i> _{RMS}	0.580	0.700	0.339	0.286	0.734	0.00E+00	2.00E-06	0.986
<i>SED</i>	0.412	0.662	0.401	0.304	2.363	0.00E+00	3.55E-16	0.991
<i>CAV</i>	0.480	0.638	0.377	0.315	0.849	0.00E+00	0.00E+00	0.994
I_c	0.820	0.842	0.222	0.208	0.458	4.07E-11	5.94E-04	0.986
I_F	0.506	0.684	0.368	0.294	0.859	0.00E+00	5.00E-15	0.991

Results corresponding to the most efficient *IM*, *PGA*, with respect to *MFA* have been highlighted; *p* values higher than 0.05, meaning that the *IM* is sufficient with respect to *M* or *ED*, have also been highlighted

Table 8 Summary of the statistical analysis considering as *EDP* to *MFA*

<i>IM*</i>	R_L^2	R_{NL}^2	Sy/x_L	Sy/x_{NL}	β	pV_M	pV_{ED}	η_{3-13}
[<i>Sa</i> (T_1), <i>M</i>]	0.239	0.399	0.457	0.406	2.278	7.71E-01	7.79E-03	0.548
[<i>Sv</i> (T_1), <i>M</i>]	0.516	0.666	0.364	0.303	0.795	8.68E-01	1.90E-01	0.960
[<i>Sd</i> (T_1), <i>M</i>]	0.388	0.545	0.410	0.353	1.260	8.86E-01	9.24E-02	1.010
[<i>AvSa</i> , <i>M</i>]	0.532	0.602	0.358	0.330	0.754	8.12E-01	1.54E-01	0.599
[<i>AvSv</i> , <i>M</i>]	0.644	0.717	0.312	0.279	0.545	8.88E-01	4.93E-01	0.999
[<i>AvSd</i> , <i>M</i>]	0.439	0.588	0.392	0.336	1.035	8.78E-01	1.54E-01	0.988
[<i>VE</i> (T_1), <i>M</i>]	0.424	0.620	0.397	0.323	1.008	7.69E-01	2.98E-02	0.900
[<i>AvVE</i> , <i>M</i>]	0.583	0.673	0.338	0.299	0.620	7.57E-01	6.43E-02	0.959
[<i>PGA</i> , <i>M</i>]	0.850	0.874	0.203	0.186	0.258	9.81E-01	2.74E-01	0.988
[<i>PGV</i> , <i>M</i>]	0.640	0.730	0.314	0.272	0.543	6.52E-01	7.78E-01	0.995
[<i>PGD</i> , <i>M</i>]	0.351	0.559	0.422	0.348	1.352	5.53E-01	2.13E-01	1.010
[I_A , <i>M</i>]	0.836	0.852	0.212	0.201	0.510	7.12E-01	1.11E-01	0.992
[<i>acc</i> _{RMS} , <i>M</i>]	0.809	0.842	0.229	0.208	0.329	9.55E-01	2.34E-03	0.990
[<i>vel</i> _{RMS} , <i>M</i>]	0.624	0.719	0.321	0.278	0.589	7.14E-01	3.33E-01	0.995
[<i>SED</i> , <i>M</i>]	0.517	0.674	0.364	0.299	1.411	5.89E-01	3.14E-02	1.001
[<i>CAV</i> , <i>M</i>]	0.627	0.695	0.320	0.290	0.458	5.84E-01	1.59E-09	1.006
[I_c , <i>M</i>]	0.829	0.851	0.217	0.202	0.431	9.79E-01	7.38E-01	0.992
[I_F , <i>M</i>]	0.575	0.698	0.341	0.288	0.617	5.86E-01	3.74E-02	1.000

Results corresponding to the most efficient *IM**, [*PGA*, *M*], have been highlighted; *p* values higher than 0.05, meaning that the *IM** is sufficient with respect to *M* or *ED*, have also been highlighted

Table 9 Summary of the statistical analysis considering as *EDP* to *MFA*

IM^*	R_L^2	R_{NL}^2	Sy/x_L	Sy/x_{NL}	β	pV_M	pV_{ED}	η_{3-13}
[$Sa(T_1), ED$]	0.248	0.402	0.454	0.405	2.237	2.64E-01	8.31E-01	0.632
[$Sv(T_1), ED$]	0.505	0.666	0.368	0.303	0.896	5.68E-04	3.64E-01	0.968
[$Sd(T_1), ED$]	0.383	0.546	0.411	0.353	1.427	2.10E-02	4.60E-01	1.017
[$AvSa, ED$]	0.531	0.603	0.359	0.330	0.801	3.21E-02	6.51E-01	0.660
[$AvSv, ED$]	0.618	0.710	0.324	0.282	0.643	7.91E-09	2.36E-01	0.998
[$AvSd, ED$]	0.422	0.587	0.398	0.337	1.245	2.55E-04	3.64E-01	1.004
[$VE(T_1), ED$]	0.417	0.623	0.400	0.322	1.173	1.68E-03	4.10E-01	0.923
[$AvVE, ED$]	0.547	0.670	0.352	0.301	0.787	1.25E-09	2.11E-01	0.968
[PGA, ED]	0.849	0.873	0.203	0.187	0.253	2.52E-02	9.58E-01	0.992
[PGV, ED]	0.614	0.723	0.325	0.276	0.639	1.81E-09	2.36E-01	0.990
[PGD, ED]	0.323	0.554	0.431	0.350	1.869	5.31E-05	4.34E-01	0.993
[I_A, ED]	0.792	0.824	0.239	0.220	0.665	0.00E+00	1.33E-01	0.986
[acc_{RMS}, ED]	0.813	0.844	0.227	0.207	0.316	3.35E-01	6.10E-01	0.991
[vel_{RMS}, ED]	0.590	0.709	0.335	0.283	0.714	1.07E-10	4.41E-01	0.990
[SED, ED]	0.464	0.672	0.383	0.300	2.023	1.71E-10	2.69E-01	0.993
[CAV, ED]	0.582	0.687	0.339	0.293	0.637	1.77E-16	8.21E-02	0.993
[I_c, ED]	0.823	0.847	0.220	0.205	0.454	1.31E-05	6.02E-01	0.989
[I_F, ED]	0.541	0.697	0.355	0.288	0.780	9.42E-10	1.70E-01	0.993

Results corresponding to the most efficient IM^* , [PGA, ED], have been highlighted; p values higher than 0.05, meaning that the IM^* is sufficient with respect to M or ED , have also been highlighted

ground and structural velocities. This fact may help understand why velocity-based IMs are good descriptors of inelastic structural responses and damage represented in terms of the deformation field of the structure.

According to the previous rationalization, one would expect that $EDPs$ related to the maximum acceleration exhibited by the structure are better predicted by acceleration-based IMs than velocity-based ones. For instance, MFA , which is an EDP highly associated to damage in non-structural components (Elenas and Meskouris 2001), is better described by IMs based on acceleration than velocity (see Fig. 10). In this figure can be seen that the IM most able to explain MFA is the PGA . It can also be seen that the quadratic regression model fits much better the cloud of points than the linear one. Tables 7, 8, 9 and 10 present statistical results regarding properties like efficiency, sufficiency and steadfastness given that the EDP is MFA .

As a matter of example, the steadfastness evolution of the most efficient IMs to predict MFA are presented in Fig. 11. It is also shown the evolution of $AvSa$ which, again, exhibits moderate efficiency to predict the response of buildings sharing the same number of stories. However, this efficiency is lost when grouping a limited variety of buildings with different number of stories. In other words, this IM is not steadfast with respect to MFA .

In terms of sufficiency, when the EDP is MFA , $Sa(T_1)$ is the only IM exhibiting this statistical property with respect to M (none of the IMs has shown efficiency with respect to ED). For this reason, the statistical properties of $Sa(T_1)$ remain similar when integrated into more advanced vector-valued arrangements. Note that the lower the p value of the basic IM the higher the efficiency gained with the advanced IM^* s.

Table 10 Summary of the statistical analysis considering as *EDP* to *MFA*

<i>IM</i> *	R_L^2	R_{NL}^2	Sy/x_L	Sy/x_{NL}	β	PV_M	PV_{ED}	η_{3-13}
[<i>Sa</i> (T_1), <i>M</i> , <i>ED</i>]	0.251	0.408	0.453	0.404	2.395	8.02E-01	7.41E-01	0.535
[<i>Sv</i> (T_1), <i>M</i> , <i>ED</i>]	0.517	0.668	0.364	0.302	0.803	8.75E-01	5.01E-01	0.964
[<i>Sd</i> (T_1), <i>M</i> , <i>ED</i>]	0.390	0.547	0.409	0.353	1.288	8.99E-01	5.62E-01	1.015
[<i>AvSa</i> , <i>M</i> , <i>ED</i>]	0.534	0.605	0.358	0.330	0.764	8.27E-01	7.70E-01	0.591
[<i>AvSv</i> , <i>M</i> , <i>ED</i>]	0.644	0.718	0.312	0.279	0.544	8.86E-01	4.24E-01	1.004
[<i>AvSd</i> , <i>M</i> , <i>ED</i>]	0.440	0.590	0.392	0.336	1.050	8.87E-01	5.05E-01	0.990
[<i>VE</i> (T_1), <i>M</i> , <i>ED</i>]	0.428	0.624	0.396	0.322	1.030	7.89E-01	5.37E-01	0.902
[<i>AvVE</i> , <i>M</i> , <i>ED</i>]	0.584	0.676	0.338	0.299	0.626	7.70E-01	3.74E-01	0.961
[<i>PGA</i> , <i>M</i> , <i>ED</i>]	0.851	0.874	0.203	0.186	0.255	9.64E-01	8.12E-01	0.997
[<i>PGV</i> , <i>M</i> , <i>ED</i>]	0.640	0.730	0.314	0.273	0.540	6.45E-01	4.21E-01	1.002
[<i>PGD</i> , <i>M</i> , <i>ED</i>]	0.352	0.562	0.422	0.347	1.378	5.65E-01	6.03E-01	1.016
[I_A , <i>M</i> , <i>ED</i>]	0.836	0.853	0.212	0.201	0.512	7.23E-01	4.14E-01	0.997
[<i>acc</i> _{RMS} , <i>M</i> , <i>ED</i>]	0.813	0.844	0.227	0.207	0.318	9.84E-01	6.67E-01	0.999
[<i>vel</i> _{RMS} , <i>M</i> , <i>ED</i>]	0.626	0.719	0.321	0.278	0.578	6.96E-01	7.96E-01	1.005
[<i>SED</i> , <i>M</i> , <i>ED</i>]	0.520	0.678	0.363	0.298	1.433	6.09E-01	4.72E-01	1.007
[<i>CAV</i> , <i>M</i> , <i>ED</i>]	0.649	0.711	0.311	0.282	0.450	6.37E-01	1.56E-01	1.003
[I_c , <i>M</i> , <i>ED</i>]	0.829	0.851	0.217	0.202	0.430	9.85E-01	8.71E-01	0.997
[I_F , <i>M</i> , <i>ED</i>]	0.577	0.701	0.341	0.287	0.623	6.00E-01	2.93E-01	1.006

Results corresponding to the most efficient *IM**, [*PGA*, *M*, *ED*], have been highlighted; *p* values higher than 0.05, meaning that the *IM** is sufficient with respect to *M* or *ED*, have also been highlighted

6 Linear versus quadratic regression analysis to derive fragility curves

From sets of *IM-EDP* points, it is possible to derive fragility functions according to the cloud analysis approach (Jalayer et al., 2015). This methodology requires to calculate the best fit curve between a set of *IM-EDP* realizations in the log–log space. The resultant curve is used to estimate the median of a parametric statistical distribution, given an *IM* value. The variability of this parametric distribution is estimated as the logarithmic standard deviation of the regression ($S_{y/x}$). In this way, the probability of exceeding a certain damage threshold given an *IM* value can be calculated. These thresholds are particular realizations of the engineering demand parameter under consideration, *EDP_C*. Certainly, the variability of the fragility functions is directly related to the dispersion of the *IM-EDP* points. The higher this dispersion the higher the uncertainty when defining whether the structure presents one specific damage state or another. That is why it is important to identify efficient *IMs*.

In general, the aforementioned fitting is performed using linear regression analysis. However, as shown in Fig. 5, some *IM-EDP* clouds are slightly better represented by a quadratic regression analysis; in the case of *MFA* (see Fig. 10), this tendency is much clearer.

Figure 12 (Left) shows the linear and nonlinear fittings in the log–log space given the *IM-EDP* relationship between *AvSv* and *MIDR*. While no significant changes are observed with respect to $S_{y/x}$ (0.334 for the linear and 0.329 for the nonlinear fitting), probabilities provided by the fragility functions are affected because of the regression analysis type. In

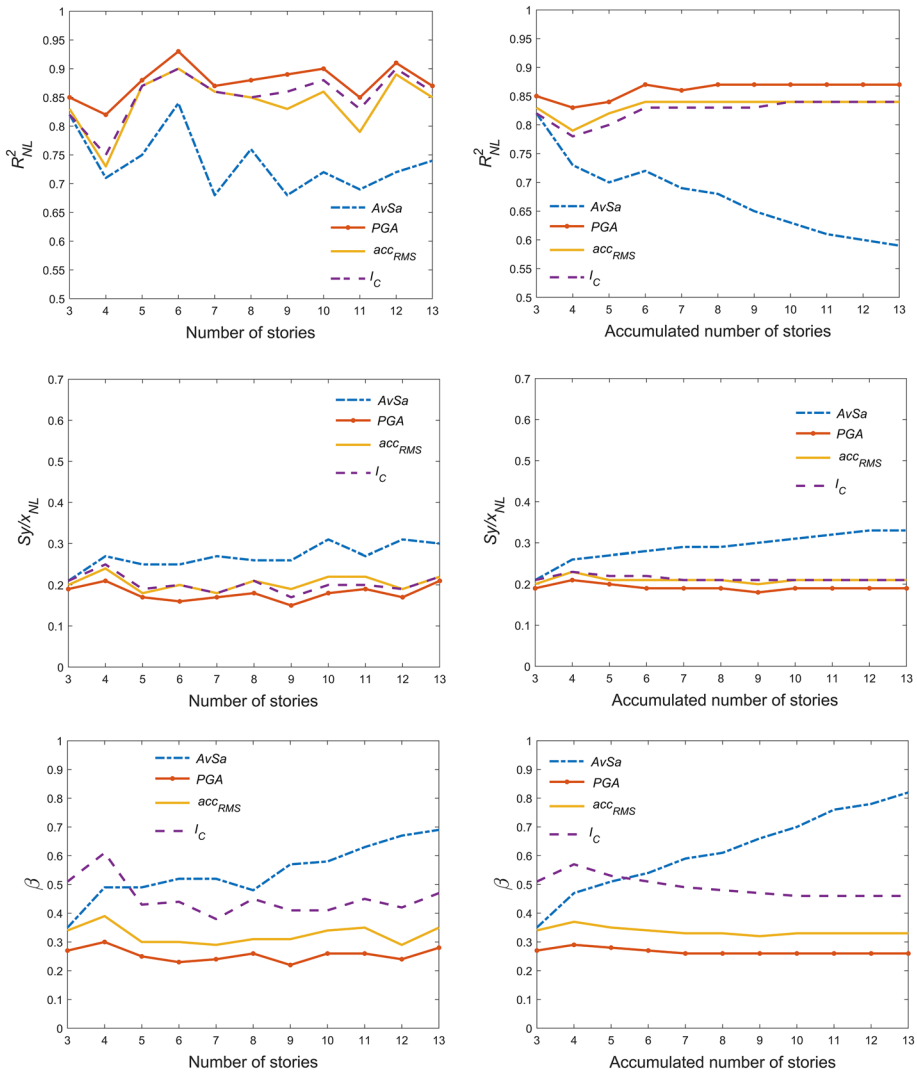


Fig. 11 Steadfastness analysis. Left: R^2 , $S_{y/x}$ and β values for buildings grouped according to the number of stories. Right: R^2 , $S_{y/x}$ and β values for buildings grouped according to the accumulated number of stories. The EDP analysed has been MFA

order to assess how much this probability of exceedance may vary, in Fig. 12 (Right) are compared fragility functions for AvSv using linear and nonlinear curves as best fit. The damage threshold to be exceeded is $MIDR_C=0.02$; this threshold can be related to a damage state identified as extensive. In this figure is also shown the absolute value of the difference between both curves (Blue dashed line). This function shows that the difference in probability may reach values in the order of 0.22.

Note that using a quadratic function can bring out problems when estimating probabilities of exceedance before the minimum or after the maximum value of the parabola (in the

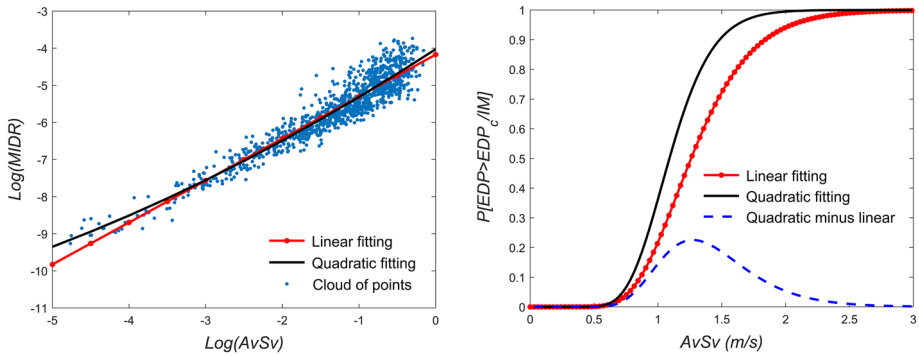


Fig. 12 Left: linear and nonlinear fitting considering $AvSv$ as IM and $MIDR$ as EDP . All the data have been included in the regression analyses. Right: Fragility curves considering both fittings; the blue dashed line indicates the absolute value of the difference in probabilities between both fragility functions

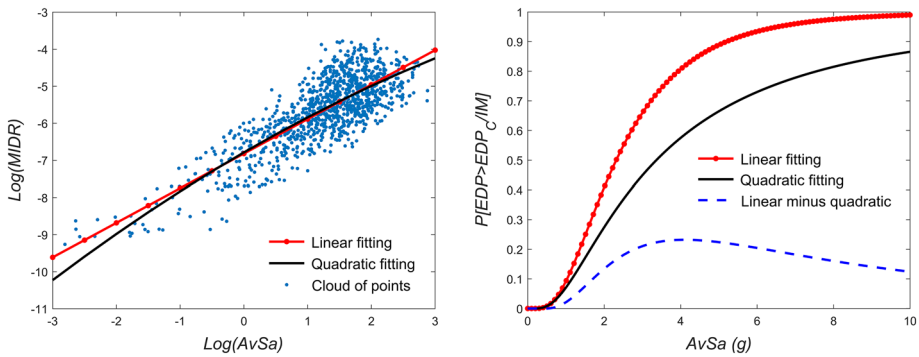


Fig. 13 Left: linear and nonlinear fittings considering $AvSa$ as IM and $MIDR$ as EDP . All the data have been considered in the regression analyses. Right: Fragility curves considering both fittings; the blue dashed line indicates the absolute value of the difference in probabilities between both fragility functions

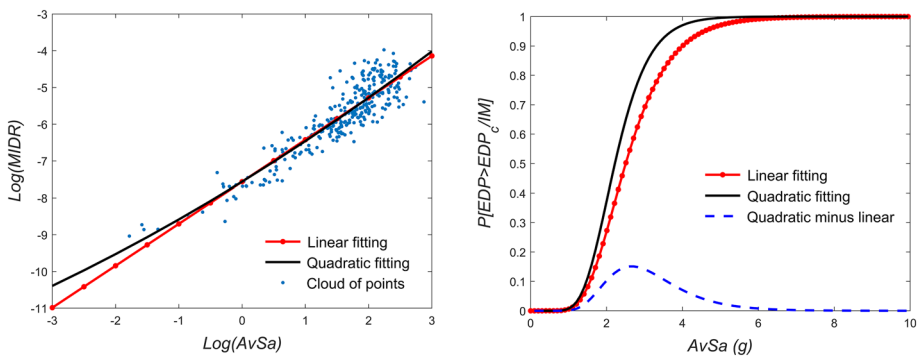


Fig. 14 Left: linear and nonlinear fittings considering $AvSa$ as IM and $MIDR$ as EDP . Only mid-rise structures have been considered in the regression analyses. Right: Fragility curves considering both fittings; the blue dashed line indicates the absolute value of the difference in probabilities between both fragility functions

case of parabolas that open downwards). In Fig. 12, for instance, the quadratic function providing the best fit is a vertical parabola opening upwards. This parabola has its minimum value in a point located in the log–log space at $(-12.0779, -12.1363)$ which, in the linear space, corresponds to $(5.6837e-06 \text{ m/s}, 5.3613e-06)$. In practical terms, the minimum value of the parabola fitted to the data (Fig. 12, left) will not produce significant problems, since there are no damage states associated with such small *MIDR* values.

If the *IM-EDP* relationship is analysed considering *AvSa* as *IM*, the resultant curve, given a nonlinear regression analysis, is a vertical parabola opening downwards (See Fig. 13, left). Its maximum value in the log–log space is located at $(9.9796, -1.8129)$ which, in the linear space, corresponds to a point located at $(2.1582e+04 \text{ g}, 0.1632)$. In practical terms, it is unlikely to have such extreme acceleration values, or this *MIDR* without reaching collapse. In Fig. 13 (Right) fragility functions for *AvSa* are derived using as best fit linear and nonlinear curves. The damage threshold to be exceeded is again $MIDR_C=0.02$. In this figure is also shown the absolute value of the difference between both curves (Blue dashed line). This function indicates that the difference in probability may reach values in the order of 0.23.

Fragility curves considering *AvSa* clearly reflect the high dispersion of this *IM-EDP* relationship when all the data are analysed simultaneously. This makes the identification of damage states more uncertain. However, as commented above, if subcategories are created by grouping buildings of similar stories, the dispersion between *AvSa-MIDR* pairs is reduced and, consequently, the quality of the fragility functions based on *AvSa* improves. For instance, Fig. 14 (Left) shows the linear and nonlinear regression fittings when considering data associated to RC mid-rise structures, i.e. building models ranging from 4 to 6 stories. A substantial reduction of the dispersion can be observed, which turns into an improvement of the quality of the fragility function (Fig. 14, right); for the nonlinear fitting, $S_{y/x}$ has decreased from 0.585, when all data is considered, to 0.415 for this subcategory.

In Fig. 14, it is also important to observe that the absolute value of the difference between the fragility functions (Blue dashed line) has reduced. This line shows that the difference in probability may reach values in the order of 0.15. Also, the resulting parabola now opens upwards.

In spite of the reduction of $S_{y/x}$ observed in Fig. 14 (Right), if fragility functions were derived for the same typology (mid-rise buildings) considering *AvSv* as *IM*, $S_{y/x}$ are 0.31 and 0.303 for the linear and nonlinear fittings, respectively. This indicates, again, the superior efficiency of *IMs* based on spectral velocity quantities to predict *EDPs* related to the deformation field of the structure.

7 Discussion

Variables related to the acceleration response of *SDoF* systems are commonly used in seismic design or risk assessment of civil structures. For instance, when applying the response spectrum method, which is one of the most invoked tools for seismic design, the acceleration design spectrum is commonly used to estimate seismic-induced forces. This approach is practical since most of the structural analysis techniques are applications based on the Newtonian-force concept. That is, the effects of equivalent static forces of various origin, static or even dynamic, are linearly combined to verify a number of stress conditions. In the case of seismic risk assessment, *IMs* based on the spectral acceleration have been commonly used to quantify the variability in the structural response. However, as demonstrated

in this article, the seismic response and damage of structural elements can be better related to the energy entering into them than to the forces exerted upon them (this is true when the damage is quantified through *EDPs* associated to the deformation field of the structure). It is important to note that this energy can be regarded as a function of the velocity of the system, as shown in Eq. 18. Therefore, results presented in this article indicate that earthquake design projects, as well as seismic risk assessments, would be more reliable if they were tied to *IMs* based on spectral displacement, velocity or equivalent velocity rather than acceleration. For instance, when analysing buildings according to the number of storeys, *AvSv* has shown to diminish $S_{y/x}$ for RC structures ranging from 5 to 11 storeys; *AvSd* not only minimizes $S_{y/x}$ for rigid structures (3- and 4-storey buildings) but it is also sufficient with respect to *M* and *ED*; the most efficient *IM* for high-rise (12- and 13-storey buildings) is *AvVE*. For design purposes, the outcomes of the present statistical analysis should allow a better quantification of the safety factors when acting loads are combined.

As discussed throughout this article, an ‘ideal’ *IM* should exhibit efficiency, sufficiency and steadfastness. The sufficiency property is the most difficult to fulfil for the 18 basic *IMs* presented herein. However, multiple regression models allow to overcome this issue. Thus, considering linear regression analysis for the sake of simplicity, the following *IM** meets all the statistical properties described in this article for an ‘ideal’ *IM*, when predicting *MIDR*:

$$AvSv * M^{0.48} \quad (19)$$

and for predicting *MFA*:

$$PGA * M^{0.29} \quad (20)$$

These *IM**s have been derived by considering coefficients given by Eq. 10. In both arrangements, the exponential of the independent term (e^{α_0}) has been set to 1; also, α_1 and α_2 have been normalized by α_1 .

In the context of earthquake engineering, probabilistic seismic hazard analysis (*PSHA*, Cornell 1968) has been and is the reference method to know about the seismic hazard on a site; see for instance Baker (2008). This method has been widely used for almost 50 years in many applications of the earthquake engineering (Mulargia et al. 2017). However, a number of researchers have questioned the validity of *PSHA* to properly quantify the expected seismic actions in urban environments, since many damaging earthquakes have occurred in regions rated as relatively low-risk areas (Stein et al. 2012; Mulargia et al. 2017). In this sense, using an *IM* highly correlated to the structural response, and revisiting some hypotheses of the *PSHA* with respect to the spatiotemporal distribution of earthquakes, could help improve the accuracy in seismic risk estimations. Note that some *IMs* based on average spectral values presented in this study can be easily incorporated within the conceptual framework of the *PSHA*. To do so, it would be necessary to develop ground motion predictive models to estimate the mean and dispersion values of these *IMs*. It is quite likely that, because of the steadfastness showed by them, the ground motion predictive models would be less dependent on the fundamental period of the system. Recently, Kohrangi et al. (2017) and Davalos and Miranda (2018) have presented ground motion predictive models for Sa_{avg} .

8 Conclusions

Advanced methodologies to quantify the probability of occurrence of a certain seismic damage state in buildings are based on the analysis of *IM-EDP* relationships. These relationships should be obtained using *NLDA*, since this is the most comprehensive numerical tool to estimate the dynamic behavior of a structure. It is also important to consider the main sources of uncertainty involved in seismic analysis. Therefore, a relevant objective for seismic engineering is to find which *IM* is the most capable of predicting a specific *EDP*. Identifying this *IM* will contribute to reducing uncertainties when quantifying seismic risk of the building stock in both rural and urban environments. This reduction is one of the key aspects to mitigating seismic risk worldwide, since stakeholders as well as policy- and decision- makers could prioritize actions to eradicate this risk based on more reliable estimations. The primary goal of this article has been to identify the most efficient *IMs* to predict *MIDR* and *MFA*. Other desirable properties expected in an ‘ideal’ *IM* like sufficiency and steadfastness have been also quantified. A set of 18 basic *IMs* was considered in the analysis; some of them are original proposals of this research.

The building stock in rural and urban environments is composed of very different structural typologies (Martins and Silva 2020). Materials, configuration, asymmetry level, and rise are physical properties generally considered to define these typologies. This is because when they vary, it is expected that the bivariate distribution of *IM-EDP* pairs also does. This fact demands the definition of a wide variety of structural types in order to include all the observed buildings in contemporary human settlements. However, this number of types can be reduced by employing *IMs* insensitive to variations of some of the aforementioned physical properties. This independence can be measured by means of a statistical property identified herein as *steadfastness*. An *IM* that exhibits steadfastness allows grouping several building types without significantly increasing the dispersion of the data. Hence, for instance, the same fragility function can be used to quantify damage to a wide variety of structures, which turns into a reduction of the number of calculations when estimating seismic risk. In other words, these *IMs* allow grouping a large variety of buildings within an enhanced structural type, since the dispersion of the *IM-EDP* pairs remains stable. For instance, similar R^2 , $S_{y/x}$ and β values have been observed whether results are grouped by the number of stories or not. This has been detected after quantifying the steadfastness of several *IMs* to predict *MIDR* and *MFA*, in the face of variations of the number of stories. Further research should be oriented to verify if these *IMs* also exhibit steadfastness for other structural types such as steel, masonry or wood buildings. The seismic response of other types of infrastructure can also be analysed with these enhanced *IMs* (Pinyol et al. 2021).

One of the most comprehensive approaches that currently exists for the derivation of fragility functions is the cloud analysis (Jalayer et al. 2015). This procedure is based on the statistical analysis of a cloud of *IM-EDP* points. A key matter of this procedure is the type of regression analysis employed to represent the mean curve of the cloud. This article has compared the use of linear and nonlinear regression models; differences in the order of 0.2 in probability have been observed. In the author’s view, fragility curves obtained when adjusting a nonlinear model will more adequately represent the structural performance when the *IM* varies. However, attention should be paid to the maximum or minimum values of the fitted parabola to avoid calculating unrealistic probabilities.

Regarding *IMs* obtained from direct computations of the ground motion record, it has also been shown that those related to the velocity are the most efficient to predict

the *MIDR*; in the case of the *MFA*, *IMs* related to acceleration play this role. Note that these *IMs* do not depend on the dynamic properties of the analysed structures. This fact implies that buildings holding very different structural properties can be analysed using the same fragility function. In other words, these *IMs* are steadfast according to the definition presented in this research. However, it has to be said that they generally produce more dispersion than those considering structural properties.

It has been observed that the effectiveness of insufficient *IMs* may increase when they are combined in a multi-regression model with seismological parameters like *M* and *ED*. Based on this concept, it would be of interest to analyse if new improved *IM**s can be developed using other *IVs* (for instance the ones presented in Table 1 or those described in Sect. 2.2). In this respect, polynomial regression models allow much better fits to some *EDPs* (e.g. *MFA*) than linear regression ones. This sophistication in the regression-model-type may help increase the accuracy in predictions of *EDPs*.

EDPs are important variables to consider when assessing seismic risk. They should describe as best as possible the expected damage of a structure. *MIDR* and *MFA* have been considered as *EDPs* in this research. They are related to the damage state of a single story (the one exhibiting the highest demand), but they may fail in describing the global damage state of the structure. Other *EDPs* like the damage index of Park and Ang (1985), which considers the response of all the structural elements should be explored. In doing so, the quantification of the damage state of a structure affected by an earthquake is expected to be more reliable.

Results presented in this article meet the principles of fair data (Wilkinson et al. 2016). Thus, potential users could find, access, interoperate and re-use data presented herein. A text file containing the main results as well a document describing the data arrangement can be downloaded from the following website: <http://kairoseq.upc.edu>. A MATLAB code has also been included, which can be used to reproduce some results presented in this article. This code could be easily operated or adapted, allowing potential users to develop further ad-hoc applications.

Acknowledgements Special thanks are given to the three anonymous reviewers for this article; their thoughtful comments were very helpful in improving the quality of this research. Yeudy F. Vargas-Alzate has been granted an Individual Fellowship (IF) in the research grant program of the Marie Skłodowska-Curie Actions (MSCA), European Union/European (H2020-MSCA-IF-2017) No 799553. This author is deeply grateful to this institution. Support for this research from the Universidad Nacional de Colombia has been received by the second author. This support is gratefully acknowledged.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. H2020 Marie Skłodowska-Curie Actions, 799553, Yeudy Felipe Vargas-Alzate.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Adam C, Kampenhuber D, Ibarra LF (2017) Optimal intensity measure based on spectral acceleration for P-delta vulnerable deteriorating frame structures in the collapse limit state. *Bull Earthq Eng* 15:4349–4373
- Akiyama H (1985) Earthquake-resistant limit-state design for buildings. University of Tokyo Press, Tokyo
- Arias A (1970) A measure of earthquake intensity. MIT Press, Cambridge, pp 438–483
- Aschheim M, Hernandez-Montes E, Vamvatsikos D (2019) Design of reinforced concrete buildings for seismic performance: practical deterministic and probabilistic approaches protection of built environment against earthquakes. CRC Press, London
- ATC-40 (1996) Seismic evaluation and retrofit of concrete buildings, Applied Technology Council, Redwood City, California
- Baker JW (2007) Probabilistic structural response assessment using vector-valued intensity measures. *Earthq Eng Struct Dyn* 36(13):1861–1883
- Baker JW (2008) An introduction to probabilistic seismic hazard analysis (PSHA). Version 1.3. October 1st, 2008. <https://www.researchgate.net/publication/237770309>. Last Accessed 2 Oct 2020
- Benavent-Climent A, Akiyama H, López-Almansa F, Pujades LG (2004) Prediction of ultimate earthquake resistance of gravity-load designed RC buildings. *Eng Struct* 26(8):1103–1113
- Bertero VV, Uang CH (1992) Issues and future directions in the use of an energy approach for seismic-resistant design of structures. In: Fajfar P, Krawinkler H (eds) *Nonlinear seismic analysis and design of reinforced concrete buildings*. Elsevier, pp 3–22
- Bianchini M, Diotallevi P, Baker JW (2009) Prediction of inelastic structural response using an average of spectral accelerations. In: *Proceedings, 10th international conference on structural safety and reliability (ICOSSAR09)*, Osaka
- Bojórquez E, Iervolino I (2011) Spectral shape proxies and nonlinear structural response. *Soil Dyn Earthq Eng* 31(7):996–1008
- Bojórquez E, Astorga L, Reyes-Salazar A, Teran-Gilmore A, Velazquez J, Bojórquez J, Rivera L (2015) Prediction of hysteretic energy demands in steel frames using vector-valued IMs. *Steel Compos Struct* 19(3):697–711
- Bojórquez E, Chávez R, Reyes-Salazar A, Ruiz SE, Bojórquez J (2017a) A new ground motion intensity measure IB. *Soil Dyn Earthq Eng* 99:97–107
- Bojórquez E, Baca V, Bojórquez J, Reyes-Salazar A, Chávez R, Barraza M (2017b) A simplified procedure to estimate peak drift demands for mid-rise steel and R/C frames under narrow-band motions in terms of the spectral-shape-based intensity measure INp. *Eng Struct* 150:334–345
- Carr AJ (2000) *Ruaumoko-inelastic dynamic analysis program*. University of Canterbury, Department of Civil Engineering, Christchurch
- Chapra S (2017) *Applied numerical methods with MATLAB for engineers and scientists*, 3rd edn. McGraw-Hill, New York
- Cheng Y, Lucchini A, Mollaioli F (2015) Correlation of elastic input energy equivalent velocity spectral values. *Earthquakes and Structures* 8(5):957–976
- Cornell CA (1968) Engineering seismic risk analysis. *Bull Seismol Soc Am* 58:1583–1606
- Davalos H, Miranda E (2018) A ground motion prediction model for average spectral acceleration. *J Earthq Eng*. <https://doi.org/10.1080/13632469.2018.1518278>
- Du A, Padgett JE (2021) Entropy-based intensity measure selection for site-specific probabilistic seismic risk assessment. *Earthq Eng Struct Dyn* 50(2):560–579
- Eads L, Miranda E, Lignos DG (2015) Average spectral acceleration as an intensity measure for collapse risk assessment. *Earthquake Eng Struct Dynam* 44(12):2057–2073
- Ebrahimian H, Jalayer F (2021) Selection of seismic intensity measures for prescribed limit states using alternative nonlinear dynamic analysis methods. *Earthq Eng Struct Dyn* 50(5):1235–1250
- Ebrahimian H, Jalayer F, Lucchini A, Mollaioli F, Manfredi G (2015) Preliminary ranking of alternative scalar and vector intensity measures of ground shaking. *Bull Earthq Eng* 13(10):2805–2840
- Elefante L, Jalayer F, Iervolino I, Manfredi G (2010) Disaggregation-based response weighting scheme for seismic risk assessment of structures. *Soil Dyn Earthq Eng* 30(2):1513–1527
- Elenas A, Meskouris K (2001) Correlation study between seismic acceleration parameters and damage indices of structures. *Eng Struct* 23(6):698–704
- Eurocode 8: Design of structures for earthquake resistance-part 1: general rules, seismic actions and rules for buildings. Brussels: European Committee for Standardization
- Fajfar P, Vidic T, Fischinger M (1990) A measure of earthquake motion capacity to damage medium-period structures. *Soil Dyn Earthq Eng* 9(5):236–242

- Ge P, Zhou Y (2018) Investigation of efficiency of vector-valued intensity measures for displacement-sensitive tall buildings. *Soil Dyn Earthq Eng* 107:417–424
- Giovenale P, Cornell CA, Esteva L (2004) Comparing the adequacy of alternative ground motion intensity measures for the estimation of structural responses. *Earthq Eng Struct Dyn* 33(8):951–979
- Güllü A, Yüksel E, Yalçın C, Anıl-Dindar A, Özkaynak H, Büyükoztürk O (2019) An improved input energy spectrum verified by the shake table tests. *Earthq Eng Struct Dyn* 48(1):27–45
- Gupta AK, Hall WJ (2017) Response spectrum method: in seismic analysis and design of structures. Routledge
- Haselton CB, Whittaker AS, Hortacsu A, Bray J, Grant DN (2012) Selecting and scaling earthquake ground motions for performing response-history analyses. In: 15th World conference on earthquake engineering, Lisbon, Portugal, 24–28 Sep
- Housner GW (1941) An investigation on the effects of earthquakes on buildings. PhD dissertation. California Institute of technology, Pasadena, California. 1941. https://thesis.library.caltech.edu/1694/1/Housner_gw_1941.pdf. Last accessed Oct 1st 2020
- Housner GW (1970) Design spectrum. Chapter 5. In: Wiegel RL (ed) *Earthquake engineering*. Prentice Hall, Englewood Cliffs, pp 93–106
- Housner GW (1975) Measures of severity of earthquake ground shaking. In: *Proceedings of the US national conference earthquake engineering*, pp 25–33
- Housner GW, Jennings PC (1982) *Earthquake design criteria*. Earthquake Engineering Research Institute (EERI). ISBN 10: 0943198232 ISBN 13: 9780943198231
- Jalayer F (2003) *Direct probabilistic seismic analysis: implementing nonlinear dynamic assessments*. Doctoral thesis, Stanford University
- Jalayer F, Cornell CA (2009) Alternative non-linear demand estimation methods for probability-based seismic assessments. *Earthq Eng Struct Dyn* 38(8):951–972
- Jalayer F, Beck J, Zareian F (2012) Analyzing the sufficiency of alternative scalar and vector intensity measures of ground shaking based on information theory. *J Eng Mech* 138(3):307–316
- Jalayer F, De Risi R, Manfredi G (2015) Bayesian cloud analysis: efficient structural fragility assessment using linear regression. *Bull Earthq Eng* 13:1183–1203
- Kazantzi AK, Vamvatsikos D (2015) Intensity measure selection for vulnerability studies of building classes. *Earthq Eng Struct Dyn* 44(15):2677–2694
- Kohrangi M, Kotha SR, Bazzurro P (2017) Ground-motion models for average spectral acceleration in a period range: direct and indirect methods. *Bull Earthq Eng* 16:45–65
- Lantada N, Pujades LG, Barbat AH (2009) Vulnerability index and capacity spectrum based methods for urban seismic risk evaluation. A comparison. *Nat Hazards* 51:501–524
- Luco N, Cornell CA (2007) Structure-specific scalar intensity measures for near-source and ordinary earthquake ground motions. *Earthq Spectra* 23(2):357–392
- Luzi L, Puglia R, Russo E, ORFEUS WG5 (2016a) Engineering strong motion database, version 1.0. Istituto Nazionale di Geofisica e Vulcanologia, Observatories and Research Facilities for European Seismology. <https://doi.org/10.13127/ESM>
- Luzi L, Puglia R, Emiliano Russo E, D'Amico M et al (2016b) The engineering strong-motion database: a platform to access Pan-European accelerometric data. *Seismol Res Lett* 87(4):987–997. <https://doi.org/10.1785/0220150278>
- Mackie K, Stojadinovic B (2003) Seismic demands for performance-based design of bridges. PEER 2003–16, Pacific Earthquake Engineering Research Center, University of California at Berkeley, Berkeley
- Martins L, Silva V (2020) Development of a fragility and vulnerability model for global seismic risk analyses. *Bull Earthq Eng*. <https://doi.org/10.1007/s10518-020-00885-1>
- Mayes R (1995) Interstorey drift design and damage control issues. *Struct Des Tall Build* 4(1):15–25
- Miranda E (2000) Inelastic displacement ratios for displacement-based earthquake resistant design. In: 12th World conference on earthquake engineering, Auckland, New Zealand, 30 January–28 February
- Modica A, Stafford P (2014) Vector fragility surfaces for reinforced concrete frames in Europe. *Bull Earthq Eng* 12(4):1725–1753
- Mulargia F, Stark PB, Geller RJ (2017) Why is Probabilistic Seismic Hazard Analysis (PSHA) still used? *Phys Earth Planet Inter* 264:63–75
- Newmark NM, Hall WJ (1982) *Earthquake spectra and design*. Earthquake Engineering Research Institute. (EERI). ISBN: 0-943198-22-4
- O'Reilly GJ (2021) Limitations of $S_a(T_1)$ as an intensity measure when assessing non-ductile infilled RC frame structures. *Bull Earthq Eng*. <https://doi.org/10.1007/s10518-021-01071-7>
- Otani S (1974) Inelastic analysis of RC frame structures. *J Struct Div ASCE* 100:1433–1449
- Padgett JE, Nielson B, DesRoches R (2008) Selection of optimal intensity measures in probabilistic seismic demand models of highway bridge portfolios. *Earthq Eng Struct Dyn* 37(5):711–725

- Park YJ, Ang AHS (1985) Mechanistic seismic damage model for reinforced concrete. *J Struct Eng* 111(4):722–739
- Park YJ, Ang AHS, Wen YK (1987) Damage-limiting aseismic design of buildings. *Earthq Spectra* 3:1–26
- Pejovic J, Jankovic S (2015) Selection of ground motion intensity measure for reinforced concrete structure. *Procedia Eng* 117:588–595
- Pinyol N, Carluccio G, Alonso E (2021) A slow and complex landslide under static and seismic action. *Eng Geol* 106478
- Pinzón LA, Vargas-Alzate YF, Pujades LG, Diaz SA (2020) A drift-correlated ground motion intensity measure: application to steel frame buildings. *Soil Dyn Earthq Eng* 132(5):106096
- Reed JW, Kassawara RP (1990) A criterion for determining exceedance of the operating basis earthquake. *Nucl Eng Des* 123:387–396
- Sarma SK (1971) Energy flux of strong earthquakes. *Tectonophysics* 11:159–172
- Sarma SK, Yang KS (1987) An evaluation of strong motion records and a new parameter A95. *Earthq Eng Struct Dynam* 15:119–132
- Satyarno I (2000) Adaptive pushover analysis for the seismic assessment of older reinforced concrete buildings. PhD thesis, Department of Civil Engineering, University of Canterbury
- Shome N, Cornell CA (1999) Probabilistic seismic demand analysis of nonlinear structures, RMS Program, Stanford University, Report No. RMS35
- Silva V, Crowley H, Varum H, Pinho R, Sousa L (2015) Investigation of the characteristics of Portuguese regular moment-frame RC buildings and development of a vulnerability model. *Bull Earthq Eng* 13:1455–1490
- Stein S, Geller RJ, Liu M (2012) Why earthquake hazard maps often fail and what to do about it. *Tectonophysics* 562–563:1–25
- Strogatz SH (2018) *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering*, 2nd edn. CRC Press, Taylor & Francis Group. ISBN-10: 0813349109 ISBN-13: 978-0813349107
- Taghavi S, Miranda E (2010) Response spectrum method for estimation of peak floor acceleration demand. In: *Improving the seismic performance of existing buildings and other structures*, pp 627–638
- Trifunac MD, Novikova EI (1994) State of the art review on strong motion duration. In: *Proceedings of the Tenth European conference on earthquake engineering*, Vienna, Austria, vol 1, pp 131–140
- Vamvatsikos D, Cornell CA (2002) The incremental dynamic analysis. *Earthq Eng Struct Dyn* 31(3):491–514
- Vamvatsikos D, Fragiadakis M (2010) Incremental dynamic analysis for estimating seismic performance sensitivity and uncertainty. *Earthq Eng Struct Dyn* 39(2):141–163
- Vargas YF, Pujades LG, Barbat AH, Hurtado JE (2013) Capacity, fragility and damage in reinforced concrete buildings: a probabilistic approach. *Bull Earthq Eng* 11(6):2007–2032
- Vargas YF, Pujades LG, Barbat AH, Hurtado JE, Diaz-Alvarado SA, Hidalgo-Leiva DA (2018) Probabilistic seismic damage assessment of reinforced concrete buildings considering directionality effects. *Struct Infrastruct Eng* 14(6):817–829
- Vargas-Alzate YF, Pujades LG, Barbat AH, Hurtado JE (2019a) An efficient methodology to estimate probabilistic seismic damage curves. *J Struct Eng ASCE* 145:04019010
- Vargas-Alzate YF, Pujades LG, González-Drigo JR, Alva RE, Pinzón LA (2019b) On the equal displacement approximation for mid-rise reinforced concrete buildings. In: *Computational methods in structural dynamics and earthquake engineering*. Crete, Greece, 24–26 June 2019
- Vargas-Alzate YF, Lantada N, González-Drigo R, Pujades LG (2020) Seismic risk assessment using stochastic nonlinear models. *Sustainability* 12:1308
- Vargas-Alzate YF, Silva V, Vamvatsikos D, Pujades LG (2021) A simplified approach for including the incidence angle effect in seismic risk assessment. *Earth Eng Struct Dyn*. <https://doi.org/10.1002/eqe.3562>
- Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G (2016) The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* 3:160018. <https://doi.org/10.1038/sdata.2016.18>
- Yakhchalian M, Nicknam A, Amiri GG (2015) Optimal vector-valued intensity measure for seismic collapse assessment of structures. *Earthq Eng Eng Vib* 14(1):37–54
- Yakhchalian M, Yakhchalian M, Asgarkhani N (2021) An advanced intensity measure for residual drift assessment of steel BRB frames. *Bull Earthq Eng* 19(4):1931–1955
- Yazgan U (2012) Proposal of energy spectra for earthquake resistant design based on Turkish registers. Doctoral thesis, Universidad Politécnica de Cataluña
- Zengin E, Abrahamson NA (2020) A vector-valued intensity measure for near-fault ground motions. *Earthq Eng Struct Dyn* 49(7):716–734

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Yeudy F. Vargas-Alzate¹  · Jorge E. Hurtado² · Luis G. Pujades¹

¹ Departament d'Enginyeria Civil i Ambiental, Universitat Politècnica de Catalunya - BarcelonaTech (UPC), Barcelona, Spain

² Departamento de Ingeniería Civil, Universidad Nacional de Colombia, Manizales, Colombia