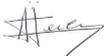# On the Selection of Natural and Synthetic Time Histories for Engineering Practice

Work Package 6 "Ground-motion for engineering"



| AUTHORS | | REVIEW | | APPROVAL | |
|---|---|---|---|---|---|
| Name | Date | Name | Date | Name | Date |
| Maria Infantino | 2021/09/30 | H. Kawase  A.Pecker | 2021/10/6 | E.VIALLET  **Public-access** ☒  **SIGMA-2 restricted** ☐ | 2023-02-15 |

## Document history

| DATE | VERSION | COMMENTS |
| :---: | :---: | :--- |
| *2021/04/09* | *1* | |
| *2021/09/30* | *2* | Corrections suggested by the reviewers have been taken into account. The main revision concerns a complementary analyses regarding the ductility demands obtained with synthetics and records inputs motions. |

## Executive summary

Nonlinear seismic structural analysis requires ground motion time histories as input. The latter can be records of past earthquakes, eventually scaled, or synthetics obtained through numerical simulations. The main objection against the use of artificial waveforms concerns the limit of reproducing the complexity of the natural ones. The purpose of this work is to analyze whether both sets of natural and artificial waveforms are comparable in terms of ground motion intensity parameters and nonlinear structural response. To this end, sets of records were selected from strong motion databases and compared with synthetics generated through Code Aster. Amongst the approaches proposed in literature to select ground motions, the methodology of Tarbali and Bradley (2015), extension of Bradley (2012), was adopted because able to take into account the variability of various ground motion intensities measures and their correlation. This approach was used for the selection of natural records and was extended for generating synthetics. As first application, 20 time histories were selected/generated compatible to the following target intensity measures differently weighted: Spectral Accelerations (weight, $w$ = 0.7), Significant strong motion Duration ($w$ = 0.1), Cumulative Absolute Velocity ($w$ = 0.1), and Arias Intensity ($w$ = 0.1). A number of SDoF systems were considered as combination of the following parameters: oscillation period, strength reduction factor, nonlinear behavior and hardening ratio. The results of this study show that, by using proper and consistent methods for both records selection and synthetics generation, the artificial time histories match well both the median and the variability of the peak seismic response and the mean of ductility demand produced by recorded motions.

# Table of Contents

# Introduction

Nonlinear seismic structural analysis requires reliable ground motion signals as input. Generally, the input ground motions are selected from recorded strong motion databases of past earthquakes representative to the target scenarios dominating the seismic hazard at the structure's site. Several selection approaches were proposed in literature (e.g. McGuire 1995; Shome 1998; Bommer and Acevedo 2004; Iervolino and Cornell 2005; Baker and Cornell 2006; Kottke and Rathje 2008; Baker 2010; Kohrangi et al. 2017; Jayaram et al. 2011; Wang 2011; Bradley 2010; Bradley 2012; Tarbali and Bradley 2015) a review of which, updated to 2010, can be found in Katsanos et al. (2010). Moreover, tools were developed for the automatic selection, and scaling, of spectrum-compatible ground motions for dynamic analysis of structures (e.g. Sgobba et al. 2019).

However, despite the increasing availability of high-quality databases of natural strong motion records, real accelerograms for some specific scenarios (such as large magnitude earthquakes recorded at close source-to-site distances) and/or site conditions (such as rock sites) are often scarce or non-existent. To overcome the paucity of recordings in some specific conditions, alternatives adopted in engineering practice include the use of modified real waveforms matching an elastic target spectrum (e.g. Atkinson and Goda 2010; Iervolino et al. 2010) or of synthetic time histories obtained by using stochastic (e.g. Vanmarcke et al. 1997; Rezaeian and Der Kiureghian 2011) or physics-based simulation approaches (e.g. Graves et al. 2011, Mazzieri 2013).

As regards the use of synthetics, the general concern amongst some engineers is that artificial waveforms may not produce the same seismic demand and, therefore, the same induced damages to nonlinear structures as real recordings (Naeim and Graves 2006) both in average, i.e. there is a systematic bias, and in record-to-record variability.

A number of works were conducted to assess whether nonlinear responses of structures subjected to synthetic and real input motions are comparable. Some studies aimed to validate simulations of past earthquakes in terms of nonlinear response of single (SDoF) (e.g. Bazzurro et al. 2004 and Galasso et al. 2012) and multi (MDoF) (e.g. Galasso et al. 2013) degree of freedom systems. Bijelic et al. (2018) and Teng and Baker (2019) focused on the nonlinear dynamic response of high-rise buildings. Iervolino et al. (2010) compared the response of nonlinear SDoF oscillators subjected to sets of records and synthetics selected in order to match the same design spectrum.

Similarly to the latter, the main goal of this work is to analyze whether sets of selected natural records and synthetic ground motions are comparable both in terms of ground motion intensity measures (IM) and of response of nonlinear SDoF systems. The synthetics herein considered were generated through the open source finite element code *Code Aster* (www.code-aster.org) by modeling ground motion as a stochastic process characterized by a power spectral density compatible to a target response spectrum. In this framework, the choice of the ground motion selection technique is fundamental to obtain ensembles of records and synthetics producing equivalent seismic demands and structural responses. The majority of the selection methods proposed in literature (Katsanos et al. 2010) are based on matching the (pseudo) acceleration response spectrum of the selected ground motions to a target spectrum. The latter can be a scenario Seismic Hazard Assessment (SHA), derived from a disaggregation of Probabilistic SHA (PSHA) analysis, or can be obtained from the design seismic code. Moreover, to identify input motions representative of the dominant scenarios ruptures and site conditions, conventional methods consider further implicit parameters (e.g. magnitude, source-to-site distance, site time-averaged shear-wave velocity in the top 30 m $V_{S30}$). However, constraining the selection only to a target mean spectrum, without taking into account the uncertainty, very likely will lead to sets of records and synthetics having the same mean but different variability. Indeed, if the uncertainty is not properly accounted for in the selection process, inevitably the selected sample of natural time histories will have a higher variability than the synthetics. The latter, in this work generated by following an approach 'spectrum-compatible', in fact will consist of a suite of signals having response spectra similar to each other because each matching the same target spectrum. While, on the other hand, being

the set of natural time histories selected from real records of different past events, inevitably will results in a more heterogeneous sample. This is evident observing Figure 1 showing samples of records (left) and synthetics (right) selected by matching a given target spectrum by Iervolino et al. (2010).
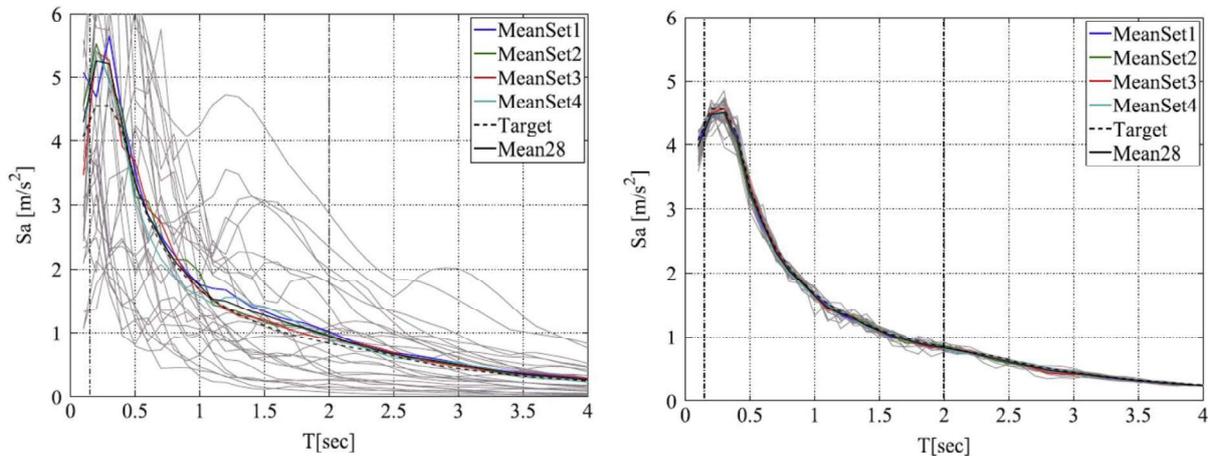


**Figure 1: Selected records (left) and synthetics (right) matching a target spectrum in Iervolino et al. (2010).**

Hence, since different distributions of the seismic demand will lead to different damage distributions, it is of paramount importance to properly account also for the variability in the ground motion IMs. This is addressed by some methodologies (e.g. Kottke and Rathje 2008; Jayaram et al. 2011; Wang 2011, Bradley 2010, Bradley 2012; Tarbali and Bradley 2015).

A second important consideration concerns the IMs used for the selection. Indeed, the severity of ground motion is not uniquely dependent on the spectral accelerations (SA) but also on other intensity parameters, such as duration, frequency content, cumulative effects, and their associated uncertainties. Bradley (2010) showed that considering only SA ordinates, as it is common in many conventional ground motions selection procedures, will typically result in ground motion sets misrepresentative of the cumulative and duration features of the appropriate ground motion target.

In this regard, the generalized conditional IM (GCIM) approach proposed by Bradley (2010) allows selecting ground motions considering the distribution of various IMs. Application of the GCIM method, based on the results of PSHA, was demonstrated by Bradley (2012) and, its extension to scenario SHA, by Tarbali and Bradley (2015).

In this work, the selection is based on a scenario SHA by following the procedure described in Tarbali and Bradley (2015) for real records and furthermore adapted to obtain compatible suites of synthetics. The seismic response of a large number of SDoFs with different backbones, hysteretic relationships, hardening parameters and strength reduction factors was investigated. Hypothesis tests on median and variability were carried out to assess if the SDoF response subjected to synthetics is systematically biased in comparison to recorded ground motion.

The work is organized as follows. First, in Section 1, the ground motion recorded and synthetics datasets are described. Then, Section 2 provides an explanation of the ground motion selection technique adopted in this work. Furthermore, the features of the SDoF oscillators considered are illustrated in Section 3. Therefore, the results are compared and discussed both in terms of ground motion IMs and structural response (Section 4). Finally, the statistical significance of the differences between SDoF demands for synthetics and natural input motions is quantified through hypothesis tests (Section 5).

# 1. Description of synthetic and real ground motion datasets

In this section the ground motion inputs used for the nonlinear analysis are presented: the recorded strong motion datasets and the simulation technique adopted to generate synthetic waveforms.

## 1.1 Recorded input motions

Two high-quality strong motion datasets were considered in this study: the ESM (Luzi et al. 2016) and the NGA-West2 (Ancheta et al. 2014). Both of them were used for several research studies and applications, such as calibration of ground motion prediction equations (GMPEs) or ground motion selection for dynamic structural analyses. Both datasets are completed with flatfile tables containing verified metadata and intensity measures of the waveforms included in each database. A systematic comparison between ESM and NGA-West2 flatfiles in terms of structure, data statistics and qualification of metadata can be found in Lanzano et al. (2020).

While for a detailed description of the two datasets the reader is referred to the relevant publications aforementioned, in the following some details of both datasets are provided.

The ESM, Engineering Strong-Motion, database contains waveforms relative to events with magnitude ≥ 4.0, mainly recorded in the European-Mediterranean regions and Middle-East. The dedicated website (https://esm-db.eu) provides a set of facilities to search, select, download and analyse ground motion data and associated metadata including an exhaustive flatfile (Lanzano et al. 2019) with a format particularly useful for the purposes of this work.

The ESM dataset includes 23,014 recordings from 2179 earthquakes and 2080 stations from Europe and Middle-East the distribution of which can be appreciated in Figure 2.
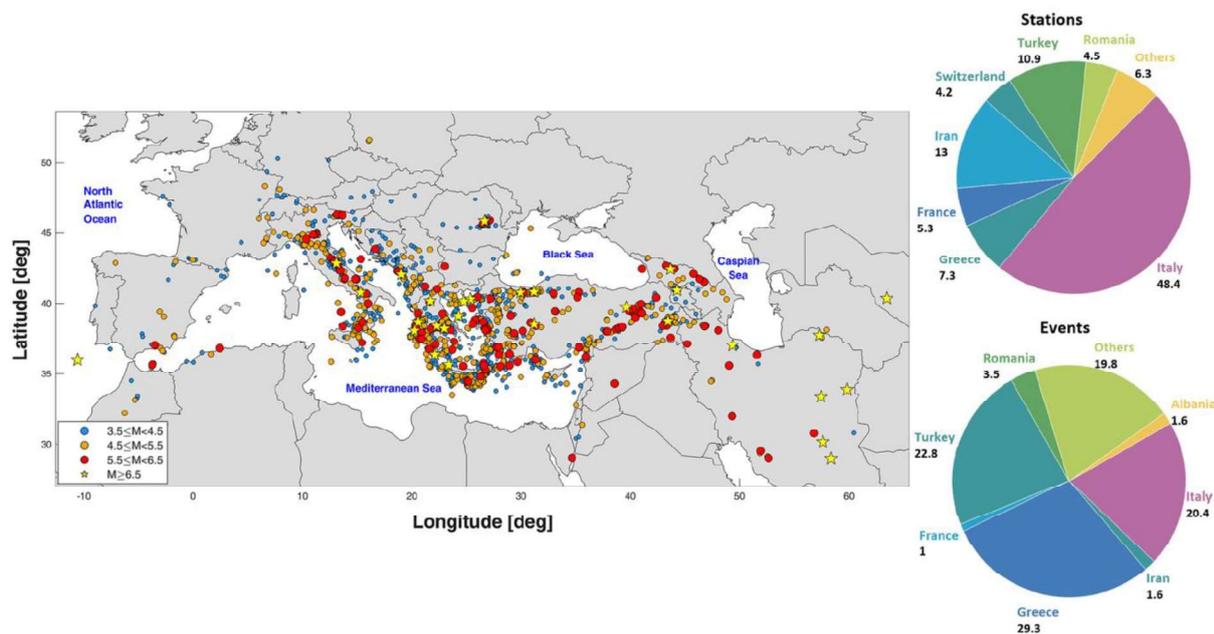


**Figure 2: Left: epicenter distribution of earthquakes in ESM database according to magnitude intervals. Right: pie-chart of the percentages (%) of stations (top) and events (bottom) within ESM database. From Lanzano et al. (2019).**

The NGA-West2 database (https://ngawest2.berkeley.edu/) includes 21,335 (mostly) three-component records from 599 events, from magnitude 3.0 to 7.9, closest distance from 0.05 to 1,533 km, and $V_{S30}$ from 94 m/s to 2100 m/s. The distribution of NGA-West2 data is shown in Figure 3.

The NGA-West2 database includes uniformly processed time series and response spectral ordinates for 111 periods ranging from 0.01 s to 20 s at 11 damping ratios.
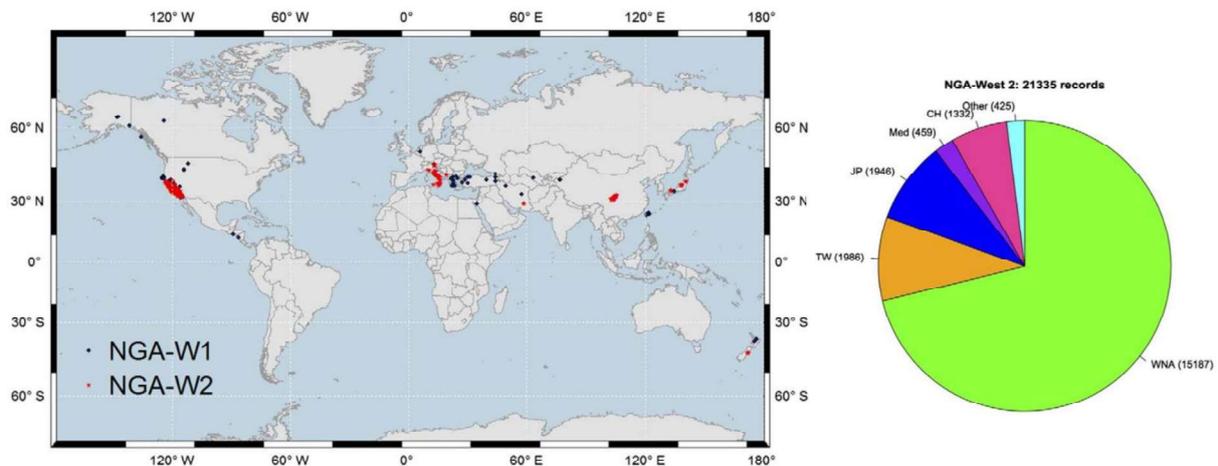


**Figure 3: Left: epicenter distribution of the 599 earthquakes in NGA-West2 database. Right: pie-chart of the records within NGA-West2 database. CH: China, JP: Japan, Med: Mediterranean, TW: Taiwan, WNA: Western North America (i.e., mostly California). From Ancheta at al. (2014).**

As general overview of comparison, Figure 4 shows the distribution of strong motion recording stations (2080 stations in ESM and 4149 in NGA-West2) as function of preferred $V_{S30}$ (left) and the magnitude-distance distribution (right) of the two datasets.
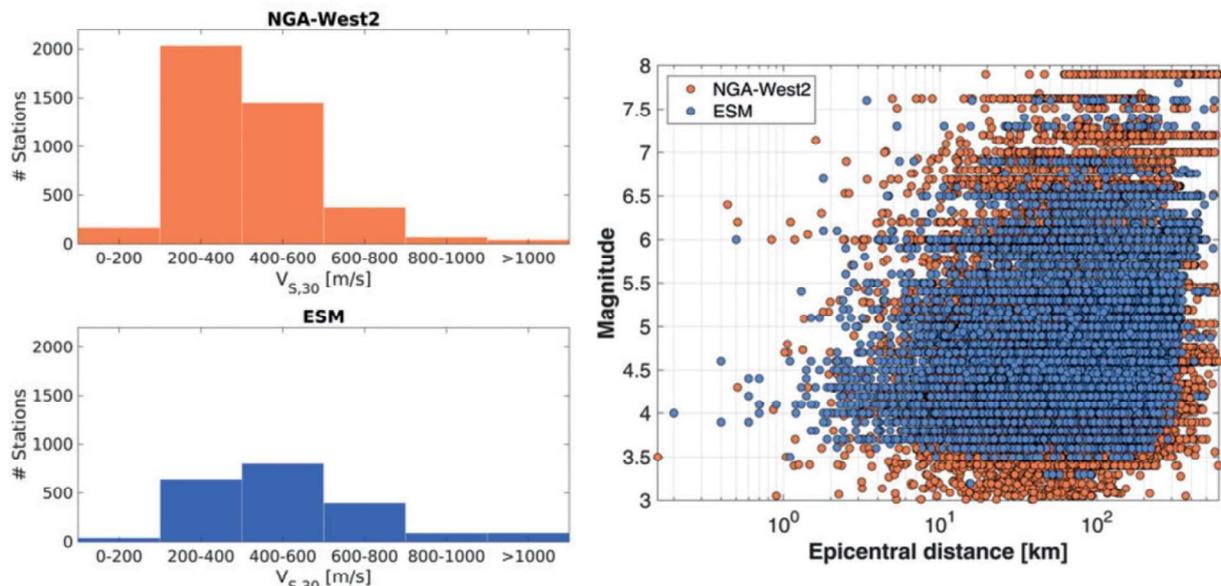


**Figure 4: Left: Distribution of strong motion recording stations as a function of $V_{S30}$ in NGA-West2 (top) and ESM (bottom). Right: Magnitude-distance distribution of ESM and NGA-West2 datasets. From Lanzano at al. (2020).**

## 1.2 Synthetic input motions

The synthetic accelerograms were generated through the open source finite element code *Code Aster* (www.code-aster.org).

The procedure implemented in Code Aster for generating artificial signals, a detailed description of which can be found in Zentner (2014, 2016), allows to simulate spectrum-compatible (Vanmarcke and

Gasparini 1976, Kaul 1978, Preumont 1980, Der Kiureghian 1981, Mertens and Preumont 1993, Cacciola 2010) accelerograms. More specifically, ground motion is modeled as a stochastic process characterized by a power spectral density (PSD) compatible to a target response spectrum. The resulting PSD allows simulating amplitude modulated ground motion whose response spectrum matches the target. The amplitude modulation is introduced by a deterministic function that is applied to the time histories. To account for the evolution of the frequency content, the model is parametrized and a time dependent central frequency is introduced. Once the PSD is constructed, ground motion time histories can be simulated using the classical spectral representation theorem.

A presentation of the methodology is provided in the following while a detailed description is given in Zentner (2014) and Zentner (2016).

Artificial ground motion time-histories are considered as realizations of a zero-mean Gaussian process $Y(t)$ defined by its evolutionary PSD (Priestley 1981):

$$S_Y(\omega, t) = |A(\omega, t)|^2 S_X(\omega) \quad (1)$$

Where $A(\omega, t)$ is a frequency dependent modulating function and $S_X(\omega)$ is the PSD of the Gaussian stationary process $X(t)$. In Code Aster the modulating function does not depend on the frequency, implying that amplitude varies only with respect to time:

$$S_Y(\omega, t) = q(t)^2 S_X(\omega) \quad (2)$$

In Code Aster the Jennings & Housner and the Gamma modulating functions are available. In this work the latter is considered, defined as (Saragoni and Hart 1973, Rezaeian and Der Kiureghian 2010):

$$q(t) = \alpha_1 t^{(\alpha_2 - 1)} exp(-\alpha_3 t) \quad (3)$$

Where $\alpha_1$ is a normalizing constant while $\alpha_2$ and $\alpha_3$ describe the shape and strong motion duration $D_{595}$, defined as the interval between 5% and 95% of Arias intensity, of the signal.

In order to comply with a target response spectrum, a "response-spectrum compatible" PSD $S_X(\omega)$ has to be identified. Vanmarcke and Gasparini (1976), amongst the first to study the problem, derived the fundamental relationship between the response spectrum $S_a(\omega_k, \xi)$ and the spectrum-compatible PSD $S_X(\omega)$ through the so-called "first passage problem":

$$S_X(\omega_k) = \frac{1}{\omega_k \left(\frac{\pi}{2\xi} - 2\right)} \left[\frac{S_a{}^2(\omega_k, \xi)}{\eta_{D_{595}}^2} - 2 \int_0^{\omega_k} S_X(\omega) d\omega\right] \quad (4)$$

Where $\xi$ is the damping ratio and $\eta_{D_{595}}$ is the peak factor, function of the strong motion duration $D_{595}$. Equation (4) allows to evaluate PSD for discrete positive frequencies $\omega_k$.

The spectrum-compatible and non-stationary ground motion model based on evolutionary PSD (Priestley, 1965) implemented in Code Aster uses a general formulation where ground motion is expressed as filtered white noise, close to the classical Kanai-Tajimi PSD model (Kanai, 1957; Tajimi, 1960). The evolutionary PSD is expressed as:

$$S_R(\omega, t) = \frac{R_1^2 + R_2^2 \omega^2}{(\omega^2 + \omega_0(t)^2)^2 + 2\xi^2 \omega^2 \omega_0(t)^2} \quad (5)$$

In which the evolution of the frequency content is introduced by considering a time dependent central frequency. Following Rezaeian and Der Kiureghian (2010) a linear relation is preferred:

$$\omega_0(t) = \omega_0 + \omega_p(t - T_{mid}) \quad (6)$$

The slope $\omega_p$ is assumed < 0 since central frequency is decreasing with time. $T_{mid}$ is the instant when half of the strong motion phase is reached. The parameters $R_1$, $R_2$ and $\omega_0$ are identified by minimizing the distance (at least squares) to the spectrum compatible PSD $S_X(\omega)$ obtained by means of Equation (4).

Finally, the equal energy criterion (Preumont 1985) guarantees the consistency between the spectrum-compatible PSD (Equation (4)) and the evolutionary PSD (Equation (5)):

$$\int_0^T q(t)^2 S_R(\omega, t) S_0(\omega) dt = \int_0^T q(t)^2 S_X(\omega) dt \quad (7)$$

where $S_0(\omega)$ is a corrective term.

# 2. Description of ground motion selection procedure for a scenario seismic hazard analysis

The selection procedure followed in this work is the one proposed and explained in Tarbali and Bradley (2015).

In order to take into account the ground motion variability for a given scenario rupture, it is necessary to select ground motions with an explicit representation of this variability. A computationally efficient way to achieve this goal is performing the selection by considering the distribution of various IMs (i.e. from the multivariate distribution) defined by the user (see Section 4). More specifically, the idea is to generate random realizations of the considered IMs from the multivariate distribution and then select ground motions that most closely match the random realizations.

The main steps of the process are described in the following. The original procedure (Sections 2.1-2.3) was further slightly adapted to generate suitable sets of synthetics (Section 2.4).

## 2.1 Constructing the multivariate distribution of the considered IMs

Let us define with **IM** = [*IM₁,…, IMᵢ,…, IMₙ*] the vector of chosen IMs according to which performing the selection. **IM** has length $N_{IM}$, where $N_{IM}$ is the number of IMs considered. Based on previous studies (e.g. Jayaram and Baker 2008), the lognormal multivariate distribution is used herein to describe the joint distribution of **IM**. It follows, that the marginal distribution of each intensity measure *IMᵢ* for a considered scenario earthquake rupture *ERup* can be written as:

$$f_{\Im_i \vee ERup} \, ln\big(\mu_{lnIM_i \vee ERup}, \sigma^2_{lnIM_i \vee ERup}\big) \quad (8)$$

Where $f_{\Im_i \vee ERup}$ is the probability density function of *IMᵢ* given the rupture *ERup*, while $\mu_{lnIM_i \vee ERup}$ and $\sigma^2_{lnIM_i \vee ERup}$ are the mean and variance of *ln(IMᵢ)* and they can be estimated by means of GMPEs. Once that the parameters $\mu_{lnIM_i \vee ERup}, \sigma^2_{lnIM_i \vee ERup}$ are computed, the marginal distribution of $\Im_i \vee ERup$ is defined.

In order to construct the multivariate distribution of the considered IMs, empirical correlation equations are then used to define the correlation matrix $\rho_{ln\Im \vee ERup}(i,j) = \rho_{i,j}$, where $\rho_{i,j}$ is the correlation coefficient between *ln(IMᵢ)* and *ln(IMⱼ)* for the considered rupture *ERup*.

## 2.2 Generating realizations of the considered IMs

Once that the multivariate distribution is defined, next step consists in generating random realizations of the considered IMs from the multivariate distribution. To this end, first a vector of uncorrelated random

numbers with standard normal distribution is generated ($u^{nsim}$) which is then converted to a vector of correlated random numbers ($v^{nsim}$) by means of correlation matrix $\rho_{ln\Im VERup}$:

$$v^{nsim} = L \cdot u^{nsim} \quad (9)$$

Where $L$ is from the Cholesky decomposition of the correlation matrix (i.e. $\rho_{ln\Im VERup} = \phantom{L}^T$) and $v^{nsim}$ is the resulting vector of correlated random numbers with a standard normal distribution. It follows that random realizations of the considered IMs can be computed as:

$$lnIM_i^{nsim} = \mu_{lnIM_i VERup} + \sigma_{lnIM_i VERup} \cdot v_i^{nsim} \quad (10)$$

Where $v_i^{nsim}$ is the $i^{th}$ element of $v^{nsim}$. Therefore, $\Im_i^{nsim}$ represents the $nsim^{th}$ realization of the $i^{th}$ intensity measure. $\Im_i^{nsim}$ is the $i^{th}$ element of the $\Im^{nsim}$ vector.

In this way, indicating with $N_{gm}$ the number of ground motion realizations desired, a total of $N_{gm}$ vectors $\Im^{nsim}$, with $1 \leq nsim \leq N_{gm}$ and of length = $N_{IM}$, are generated.

## 2.3 Selecting recorded ground motions from strong motion databases

Therefore for each vector $\Im^{nsim}$, a specific record can be selected from a strong motion database (e.g. ESM or NGA databases described in Section 2.1) based on the minimum mismatch:

$$r^{m,nsim} = \sum_{i=1}^{N_\Im} w_i \left[ \frac{lnIM_i^{nsim} - lnIM_i^m}{\sigma_{lnIM_i VERup}} \right]^2 \quad (11)$$

Where $\Im_i^{nsim}$ is the $nsim^{th}$ realization of the $i^{th}$ intensity measure computed from Equation (10), $\Im_i^m$ is the $i^{th}$ IM value of the $m^{th}$ record in the database and $w_i$ is the importance weight associated to the $i^{th}$ IM and defined by the user (see discussion at Section 4).

It is worth noting that, since this selection procedure is based on the generation of random realizations, performing the selection successive times may lead to different suites of selected records, especially if the number of records selected is small. By repeating the selection for several times, the "best" replicate can be chosen by comparing the distribution of the IMs from the selected sets with the target distribution from Equation (8). In particular, in this work the Kolomogorov-Smirnov (KS) test statistic was used. For each $\Im_i$, the KS statistic $D_{IM_i}$ measures the maximum difference between the empirical distribution, obtained from the set of selected motions, and the target distribution from Equation (8). Computing $D_{IM_i}$ for each $\Im_i$ and each replicate, the 'best' set will be the one characterized by the lower 'global' residual $Rg$:

$$Rg = \sum_{i=1}^{N_\Im} w_i \left( D_{IM_i} \right)^2 \quad (12)$$

## 2.4 Generating and selecting synthetic ground motions with Code Aster

As described in Section 1.2, the generation of synthetics with Code Aster (spec_unique option, see manual Zentner 2017) requires the definition of a target response spectrum $SA$ and of a target strong motion duration $D_{595}$. In this work, such targets are obtained by means of Equation (10) (defining $\Im = SA, D_{595}$). For each pair of target realizations ($SA^{nsim}, D_{595}^{nsim}$) a number $N^*$ of synthetics is first generated and then a specific synthetic ground motion is selected based on the minimum mismatch considering all the others IMs (besides $SA$ and $D_{595}$) considered in the selection process:

$$r^{s,nsim} = \sum_{i=1}^{N_{\Im_i}} w_i \left[ \frac{lnIM_i^{nsim} - lnIM_i^s}{\sigma_{lnIM_i VERup}} \right]^2 \quad (13)$$

Where $\mathfrak{I}_i^s$ is the $i^{th}$ IM value of the $s^{th}$ synthetic motion amongst the N* generated, while $w_i$, $\mathfrak{I}_i^{nsim}$ and $\sigma_{lnIM_i\backslash R}$ are explained in the previous sections.

# 3. Description of the SDoF Systems

The sets of ground motions, recorded (Section 1.1) and simulated (Section 1.2), selected according to the procedure described in Section 2, were used as seismic inputs to nonlinear analyses.

Overall, 238 SDoF systems were analyzed as combinations of the following parameters:

- SDoF fundamental period *T*. Fourteen (14) values of *T* were considered, i.e.: T=[0.1:0.05:0.3,0.4,0.5,0.75,1:1:5,10] s
- Nonlinear model. Four different models were considered: (i) elastoplastic with kinematic hardening; (ii) elastoplastic with isotropic hardening; (iii) Takeda model; (iv) nonlinear elastic. As an example, Figure 5 shows the four nonlinear responses of SDOF systems subjected to the same input motion.
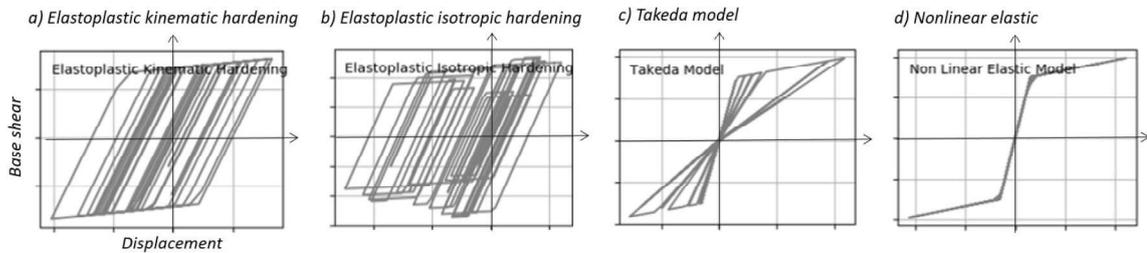


**Figure 5: SDOF nonlinear responses according to the four nonlinear models investigated in this study considering the same input motion.**

- Hardening parameter $\theta$ defining the post elastic slope. Two values of $\theta$ were considered: (i) $\theta=2\%$, (ii) $\theta=5\%$.
- Strength reduction factor *R*. Defined as the maximum elastic force over the yielding force, *R* defines the degree of inelasticity. Three values for *R* were considered: (i) *R* = 1 linear elastic (for completeness), (ii) *R* = 2 mildly inelastic; (iii) *R* = 8 severely inelastic.

Hence, 14 linear and 224 nonlinear SDoFs were considered. Indeed, it is worth noting that for linear elastic SDoFs (R=1), varying the hardening ratio $\theta$ or the nonlinear model has no effect.

As general consideration, it should be taken into account that the peak value of the elastic base shear force, or equivalently the peak elastic deformation, experienced by the structure is a ground motion specific quantity. Therefore it is possible to achieve the same value of *R* either (a) for each record in a dataset (referred to as *constant-R* approach) or (b) in an average sense for all the records in the same dataset (*constant-strength* approach). In the former *constant-R* approach (a), the same target *R* value is obtained by varying the yield strength *Fy*, or equivalently the yield displacement *dy*, from record to record. More precisely, for each record *Fy* (or *dy*) is set equal to the peak elastic force (or displacement) divided by the desired value of *R*. On the other hand, in a *constant-strength* approach (b), *Fy* (or *dy*) is kept constant for all the records and set equal to the median peak elastic force (or displacement) across all the records in the dataset divided by *R*.

In this study, the *constant-R* approach is considered to guarantee the expected levels of nonlinearity for all the input motions. Although the results may change if a *constant-R* or *constant-strength* approach is considered, Bazzurro et al. (2004) showed that the results of the comparison between recorded and

simulated ground motion using the *constant-strength* and the *constant-R* approaches for several SDoF systems do not differ significantly.

# 4. Application for an earthquake rupture scenario

In this section the comparison between sets of selected natural and artificial time histories is accomplished for a given earthquake scenario both in terms of ground motion intensity measures (Section 4.1) and SDoFs' response (Section 4.2).

The analysis were carried out for a $Mw6.0$ scenario and source-to-site distance of 40 km. The $V_{S30}$ was assumed equal to 800 m/s.

A key point concerns the choice of the IMs according to which perform ground motion selection (Section 2). Previous studies (Bradley 2010; Bradley 2012) recommended including also non-*SA* intensity measures in the selection process since a selection based only on *SA* ordinates will result in suites of ground motions which may provide a misrepresentation of the cumulative and duration features. Based on various sensitivity analysis, Tarbali and Bradley (2015) suggested for a generic case study to perform record selection based on: response spectra (*SA*), the 5-95% significant duration ($D_{595}$), to take into account shaking duration, Arias intensity (*AI*) and cumulative absolute velocity (*CAV*) to consider cumulative aspects correlated with the high (*AI*) and moderate-to-low (*CAV*) frequency content of ground motion respectively. Moreover, the authors recommended assigning a higher importance weight to *SA* and equal importance weights to non-*SA* IMs. Hence, following the authors' indications the IMs considered in this study include: *SA* for 11 vibration periods (T = 0.01,0.05,0.1,0.25,0.4,0.5,1,2,4,5 and 10s), $D_{595}$, *CAV* and *AI*; associated with the following importance weights: $w_{SA}$ = 0.7 (evenly distributed to 11 SA ordinates), $w_{D595}$ = 0.1, $w_{AI}$ = 0.1, and $w_{CAV}$ = 0.1.

The marginal distributions of these IMs were obtained using the following GMPEs: Campbell and Bozorgnia (2014) (referred to as CB14) for *SA*, Afshari and Stewart (2016) (AS16) for $D_{595}$ and Campbell and Bozorgnia (2019) (CB19) for *AI* and *CAV*. The empirical correlation equations: Campbell and Bozorgnia (2019) (CB19), Baker and Bradley (2017) (BB17), Bradley (2011) (B11), Bradley (2015) (B15), were used to define the correlation between the IMs as indicated in Table 1. Such GMPEs were selected because the most updated when the research was conducted.

**Table 1: Empirical correlation equations between the considered IMs.**

| IM | SA | $D_{595}$ | CAV | AI |
| --- | --- | --- | --- | --- |
| SA | BB17 | BB17 | CB19 | CB19 |
| $D_{595}$ | BB17 | 1 | B11 | B15 |
| CAV | CB19 | B11 | 1 | BB17 |
| AI | CB19 | B15 | BB17 | 1 |

For each set (records and synthetics) a number of $N_{gm}$ = 20 motions was selected by matching the 20 ground motion target realizations chosen over 50 replicates (see Section 2.3, in particular Equation (12)). When the generation of synthetics is performed, for each ground motion target realization a specific synthetic motion is obtained as the best fit over $N^*$=10 artificial signals (Section 2.4).

The procedure, starting from the generation of random realizations described in Section 2.2 up to the generation of synthetics detailed in Section 2.4, was carried out two times: one time considering ESM database and another time considering NGA-West2 database.

The main assumptions for each step of the selection procedure are summarized in Table 2.

**Table 2: Assumptions chosen in this study for selecting records and synthetics compatible to a given earthquake scenario.**

| | |
| --- | --- |
| Earthquake Scenario and Site Conditions | $Mw$ = 6.0 |
| | Source-to-site distance = 40 km. |
| | $V_{S30}$ = 800 m/s |
| Construction of target realizations (Section 2.2) | Number of realizations generated: $N_{gm}$ = 20 |
| | Target IMs: $SA^{nsim}, D_{595}^{nsim}, AI^{nsim}, CAV^{nsim}$ with $1 \leq nsim \leq N_{gm}$ |
| | Number of replicates: $N_{rep}$ = 50 |
| | GMPEs* used: CB14, AS16, CB19, CB19, BB17, B11, B15 |
| Selection of natural records (Section 2.3) | Importance IMs' weights: $w_{SA}$ = 0.7, $w_{D595}$ = 0.1, $w_{AI}$ = 0.1, and $w_{CAV}$ = 0.1. |
| | Ground motion databases: ESM, NGA-West2 |
| Generation of synthetics (Section 2.4) | Code: *Code_Aster* |
| | $N^*$=10 synthetics generated for each pair of target realizations ($SA^{nsim}, D_{595}^{nsim}$) |
| Selection of synthetics (Section 2.4) | Importance IMs' weights: $w_{SA}$ = 0.7, $w_{D595}$ = 0.1, $w_{AI}$ = 0.1, and $w_{CAV}$ = 0.1. |

*the acronyms of the GMPEs are explained in the text.*

### 4.1 Comparison between recorded and synthetic ground motion intensity measures

In this section the IMs of recorded and synthetics ensembles, selected according to the assumptions discussed previously, are compared. As an example, Figures 6 and 7 show the *SA*, *AI*, *CAV* and *D₅₉₅* associated with the samples of records selected from the ESM database (Figure 6) and of synthetics (Figure 7) obtained by using the same target realizations (Section 2.2). Moreover, the GMPEs predictions (median +/- 1 and 2 sigma) are added.

In both cases, the sets appear representative samples of the IMs marginal distributions. The distributions of magnitude, source-to-site distances and site conditions of the samples of natural selected records are reported in Appendix 1. It is worth noting that further analyses, not reported here for sake of conciseness, were carried out by considering a prescreening according to magnitude, source to site distance and $V_{S30}$. Such analyses revealed that the quality of the selected records did not improve by performing a preselection while, on the contrary, the results worsened especially when a prescreening according to $V_{S30}$ was performed because of the paucity of available records. Since magnitude, source-to-site distance and $V_{S30}$ are used in GMPEs computation, they are considered in the ground motion selection process. Nevertheless, the selected ground motions might come from very different site contexts than those of the considered site (see Appendix 1).

A useful way to investigate whether recorded and synthetics sets are consistent each other, consists in comparing median and standard deviation of the relevant IMs. This check is performed in the following for *SA* (Figure 8), *D₅₉₅* (Figure 9), *AI* (Figure 10) and *CAV* (Figure 11). More specifically, in each figure top panels show the comparison between records selected from ESM database and synthetics, while bottom panels between records selected from NGA-West2 database and synthetics. On the left, they

are shown the IMs associated to each recorded (black) and synthetic (green) signal, while on the right the median and sigma computed from the two ensembles. Moreover, the estimations (median +/- 1, 2 sigma) from the considered GMPEs are added.

Overall, there is a good agreement between the two sets for the various IMs, especially in terms of response spectra and ground motion duration. The medians of CAV and AI computed from the synthetics are slightly higher than those from records.
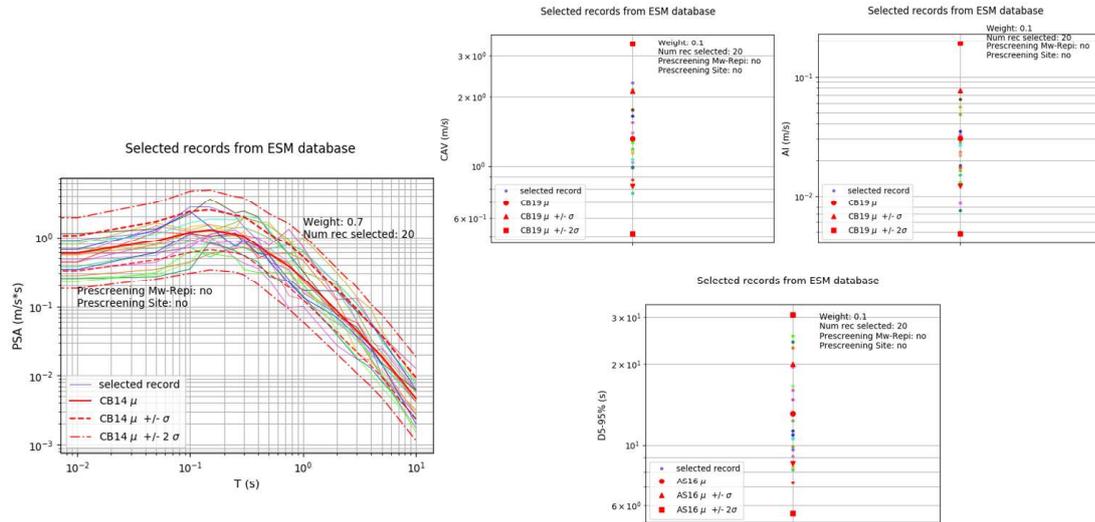


**Figure 6: IMs computed from the sample of records selected from ESM database (shown with different colors). Left: SA. Right: CAV (top left) and AI (top right) and D$_{595}$ (bottom). In red, the GMPEs predictions (median +/- 1 and 2 sigma) are added.**
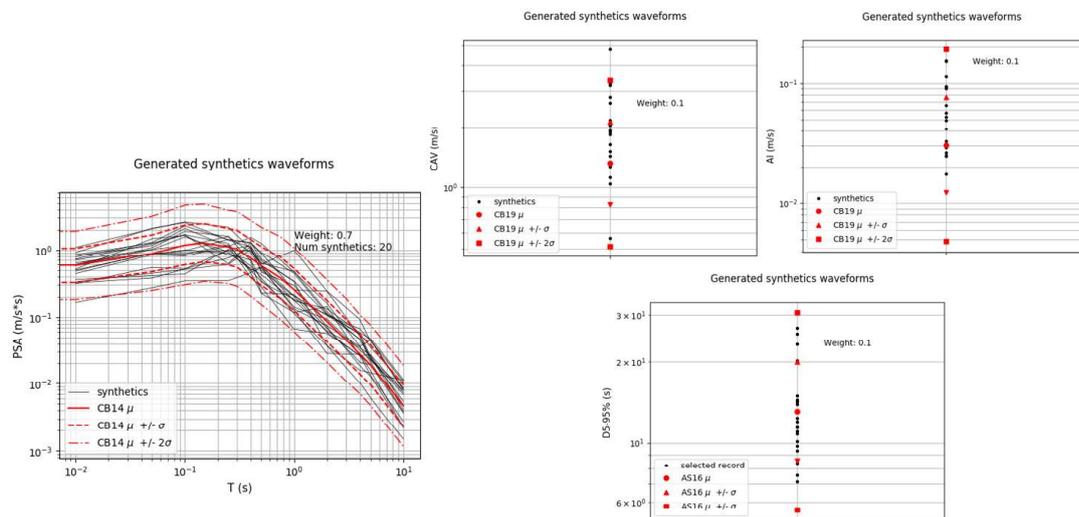


**Figure 7: IMs computed from the sample of synthetics (black). Left: SA. Right: CAV (top left) and AI (top right) and D$_{595}$ (bottom). In red, the GMPEs predictions (median +/- 1 and 2 sigma) are added.**

Then the correlations between the different IMs, computed from the sets of records and synthetics, were estimated and compared with the correlation values provided in literature. The correlation between two intensity measures (*x,y*) was computed by means of the Pearson correlation coefficient estimated as:

$$\rho_{x,y} = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2} \sqrt{\sum_{i=1}^{n}(y_i - \overline{y})^2}} \qquad (13)$$

Where $x_i, y_i$ are the pair of IMs (e.g. CAV and AI); $\overline{x}, \overline{y}$ represent the sample means of the considered IMs estimated from the sample, and $n$ is the sample size (in this case $n = 20$).

The Pearson correlation coefficient ($\rho_{x,y}$) ranges from -1 (*x* and *y* are perfectly negatively correlated) to +1 (*x* and *y* are perfectly positively correlated), while a correlation $\rho_{x,y}$ equal to 0 implies that *x* and *y* are uncorrelated. Figures 12 to 16 show the correlation coefficients between the IMs considered computed from the synthetics (red circles) and records (green triangles). Moreover, the correlation values provided in literature, i.e. relationships in Table 1, are added (grey squares). In general, in terms of correlation we can observe a better agreement between NGA-West2 records and synthetics (see for instance Figure 13), probably because the GMPEs employed in this work are calibrated on the NGA-West2 database.
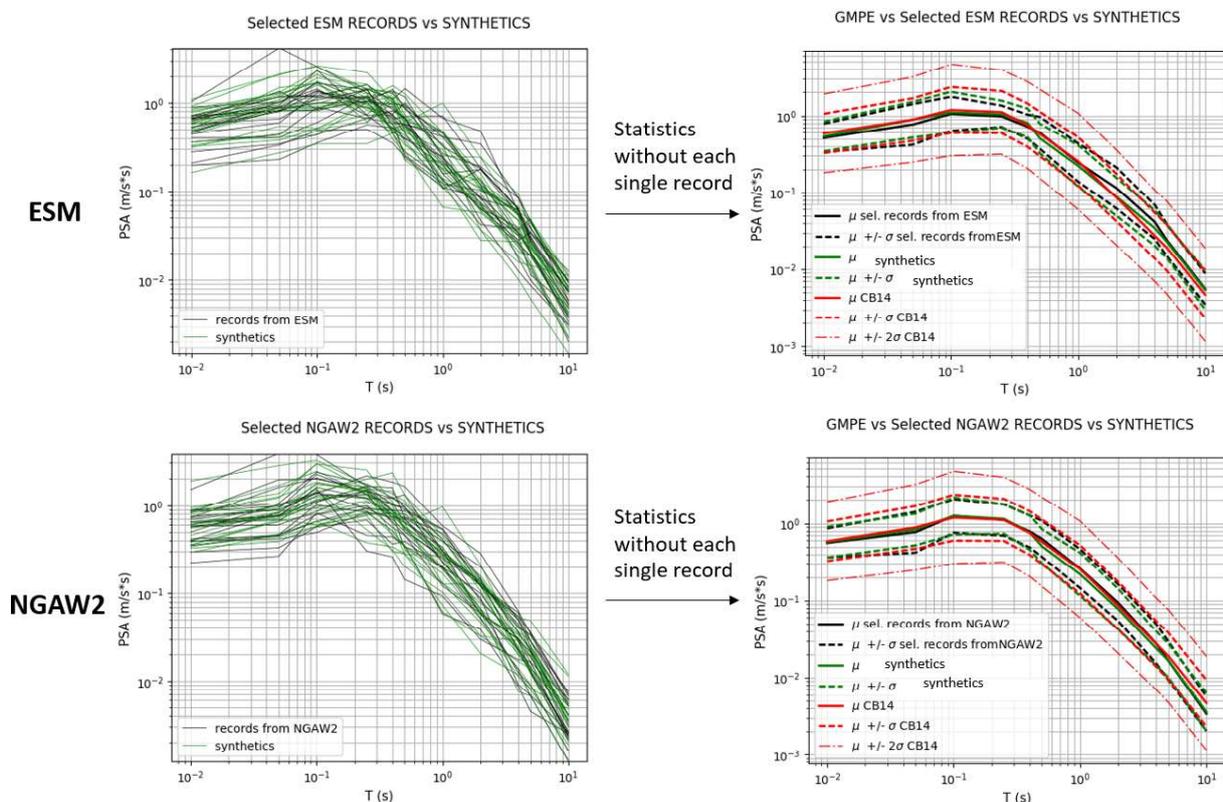


**Figure 8. Comparison between response spectra computed from the set of synthetics (green) and records (black) from ESM (top panels) and NGAW2 (bottom panels). On the left, the response spectra for each waveform is shown while on the right only the median +/- sigma for each set. On the right panels the GMPE prediction is added in red.**
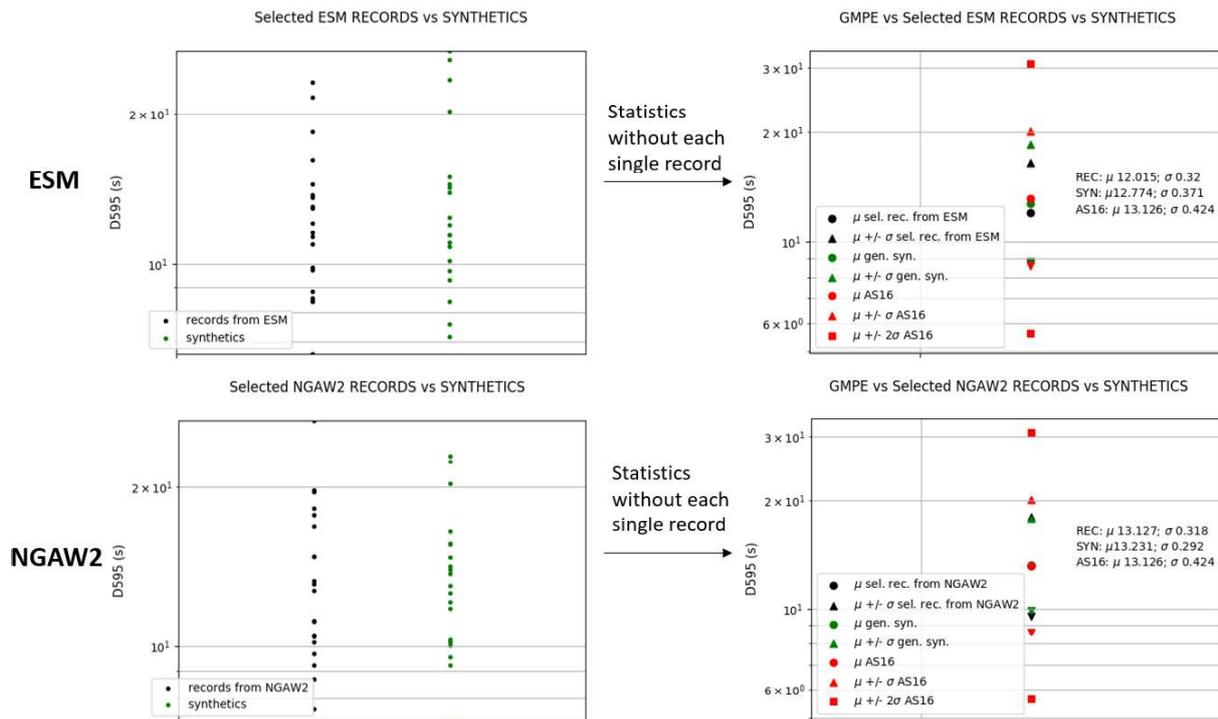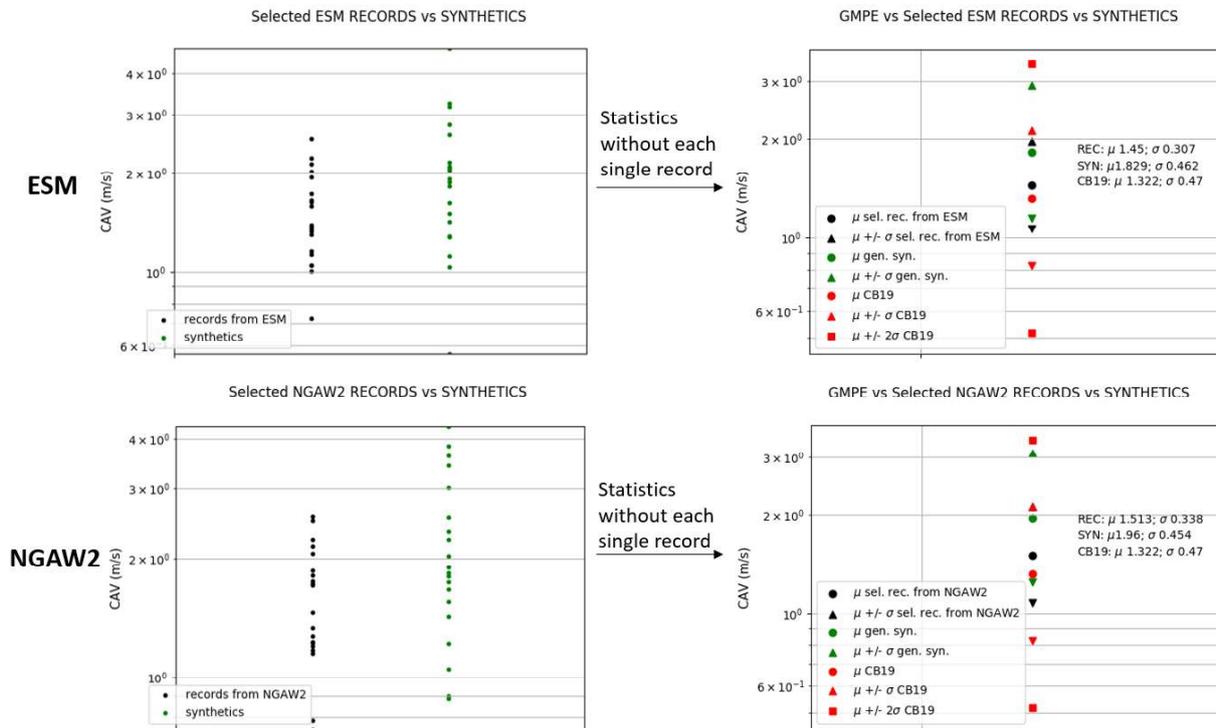
**Figure 9. Comparison between significant duration computed from the set of synthetics (green) and records (black) from ESM (top panels) and NGAW2 (bottom panels). On the left, the $D_{595}$ for each waveform is shown while on the right only the median +/- sigma for each set On the right panels the GMPE prediction is added in red.**
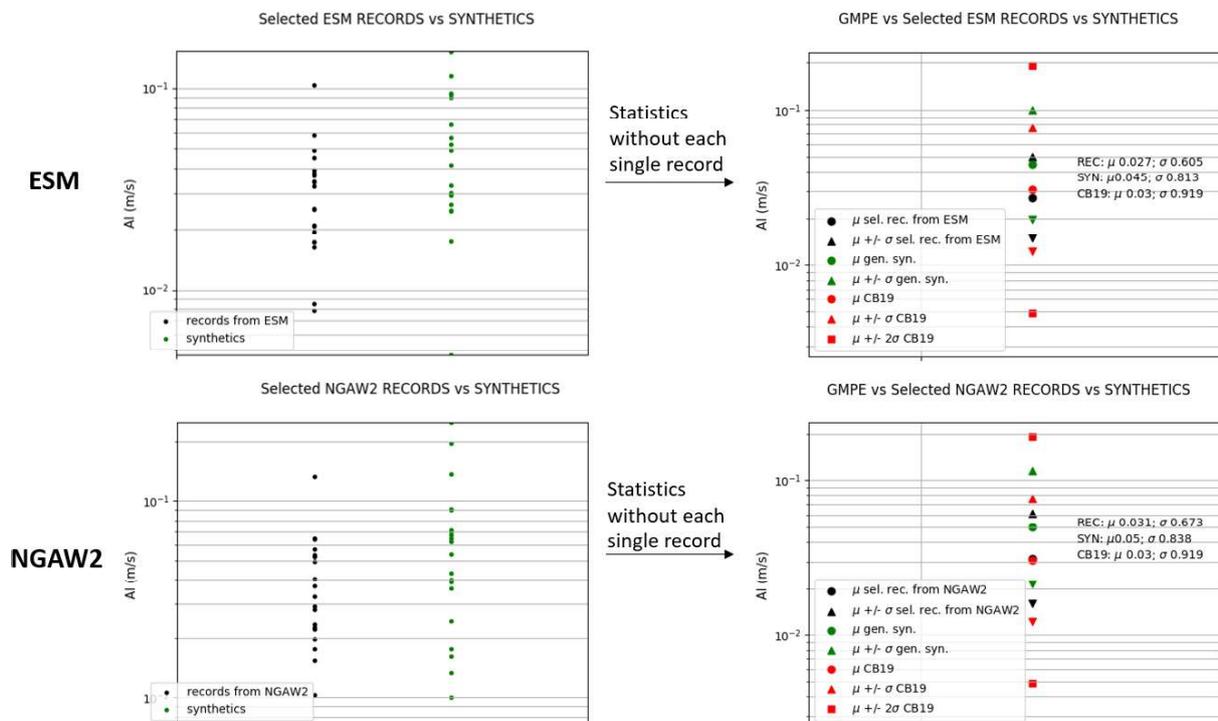


**Figure 10. Comparison between Arias Intensity computed from the set of synthetics (green) and records (black) from ESM (top panels) and NGAW2 (bottom panels). On the left, AI for each**

**waveform is shown while on the right only the median +/- sigma for each set. On the right panels the GMPE prediction is added in red.**



**Figure 11. Comparison between cumulative absolute velocity computed from the set of synthetics (green) and record (black) from ESM (top panels) and NGAW2 (bottom panels). On the left, CAV for each waveform is shown while on the right only the median +/- sigma for each set. On the right panels the GMPE prediction is added in red.**
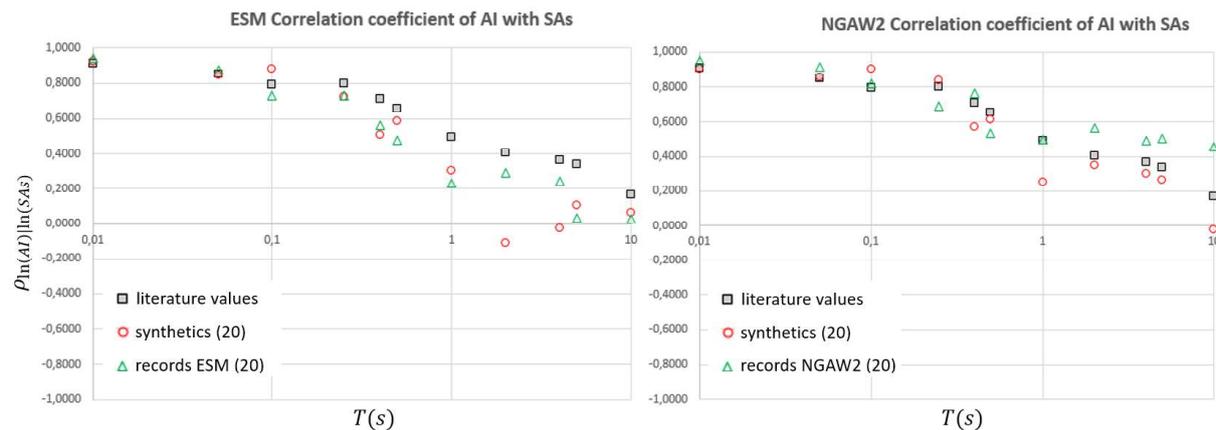


**Figure 12. Correlation coefficients $\rho$ between AI and SAs computed from records (green triangles), synthetics (red circles) and provided in literature (squares). Left: comparison ESM records vs synthetics. Right: comparison NGA-West2 records vs synthetics.**
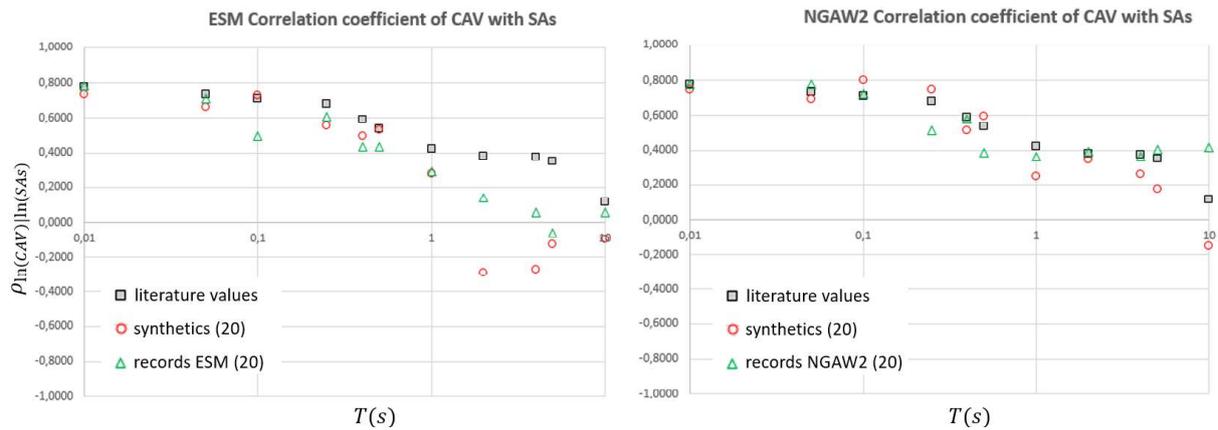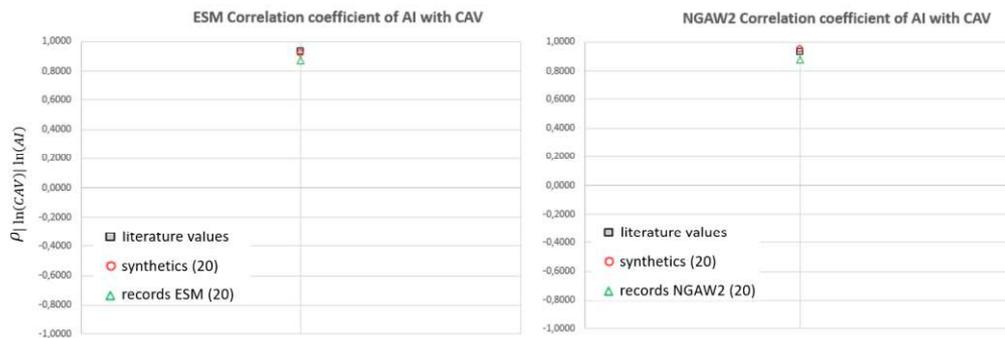
**Figure 13. Correlation coefficients $\rho$ between CAV and SAs computed from records (green triangles), synthetics (red circles) and provided in literature (squares). Left: comparison ESM records vs synthetics. Right: comparison NGA-West2 records vs synthetics.**
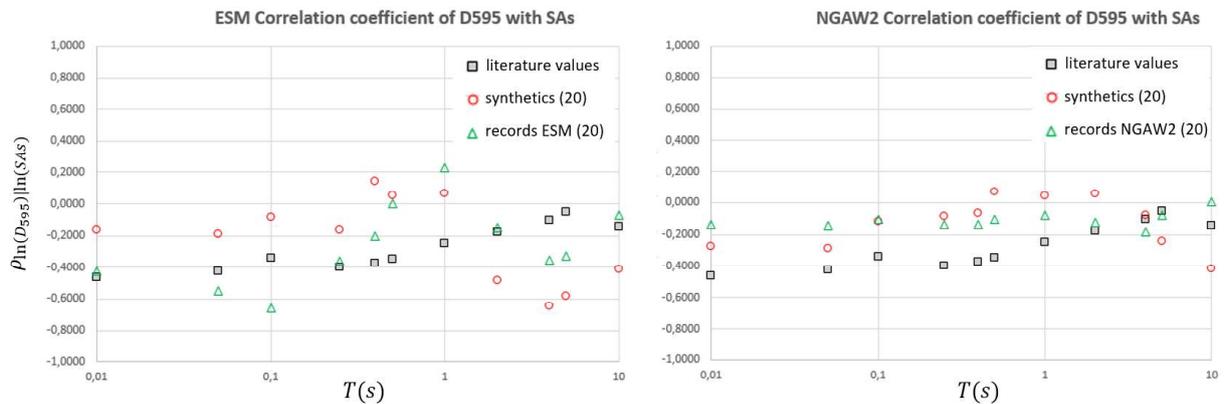


**Figure 14. Correlation coefficients $\rho$ between $D_{595}$ and SAs computed from records (green triangles), synthetics (red circles) and provided in literature (squares). Left: comparison ESM records vs synthetics. Right: comparison NGA-West2 records vs synthetics.**
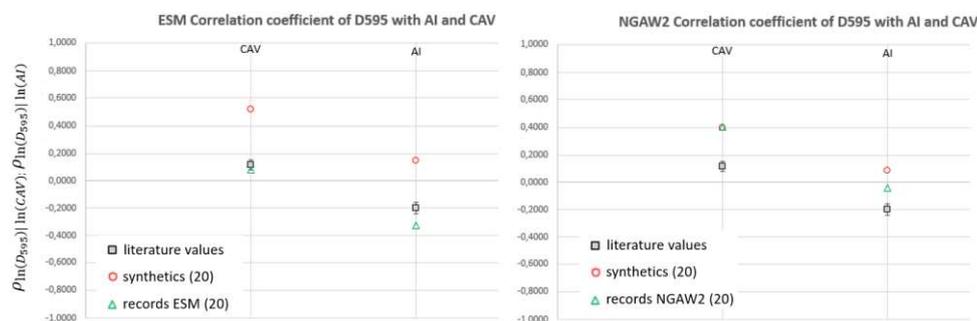


**Figure 15. Correlation coefficients $\rho$ between $D_{595}$ and CAV and AI computed from records (green triangles), synthetics (red circles) and provided in literature (squares). Left: comparison ESM records vs synthetics. Right: comparison NGA-West2 records vs synthetics.**

**Figure 16. Correlation coefficients $\rho$ between CAV and AI computed from records (green triangles), synthetics (red circles) and provided in literature (squares). Left: comparison ESM records vs synthetics. Right: comparison NGA-West2 records vs synthetics.**

### 4.2 Comparison in terms of SDOF response using recorded and synthetic ground motion

After verifying that ground motion intensity measures computed from the sets of records and synthetics are in agreement (in median, variance and correlation), the aim is now to check if they provide comparable results also in terms of structural response. Therefore, for each SDoF system described in Section 3, identified by a $T$, $\theta$, $R$ and constitutive model, the structural response in terms of peak displacement and the ductility demand was computed for each input motion belonging to each set. More specifically, the following computations were performed:

- peak displacement and ductility demand of each SDoF system subjected to each waveform (example Figure 17);
- median ($\mu$) and sigma (of logarithms $\sigma_{ln}$) of the two engineering demand parameters (i.e. peak displacement and ductility demand) obtained with both sets of records and synthetics for each SDoF system (example Figure 17);
- computation of the ratios: $\mu_{SYN}/\mu_{REC}$ and $\sigma_{ln,SYN}/\sigma_{ln,REC}$ to check if the structural responses to synthetic motions are in agreements, overestimate or underestimate the ones obtained with records.
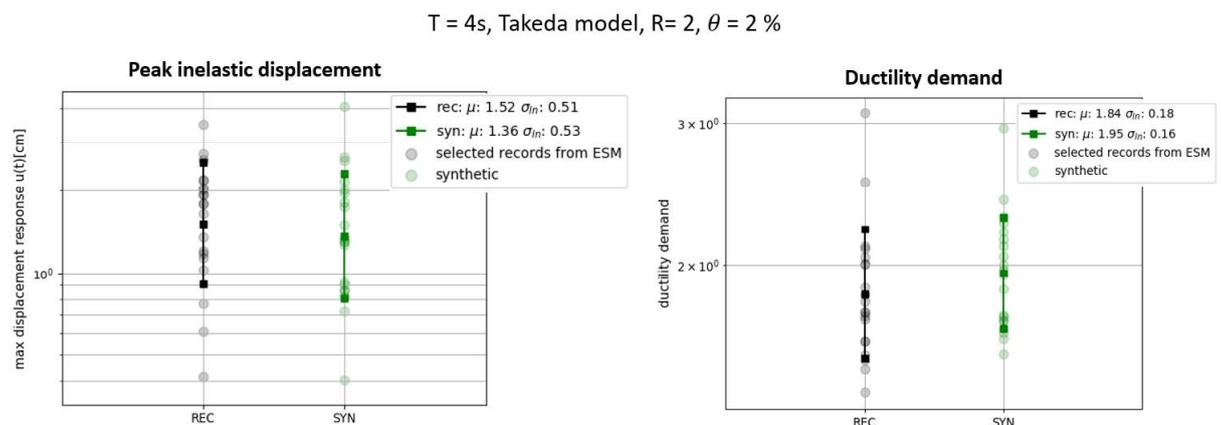


**Figure 17. Example of structural response (peak inelastic displacement on the left and ductility demand on the right) obtained using the set of synthetics (green) and records (black) from ESM database for a SDoF system with $T$ = 4s, $R$ = 2, Takeda model and $\theta$=2%.**

As an example, Figure 18 (top panels) shows the displacement spectra responses to records (black) and synthetics (green) in function of the period $T$, for SDoFs with an elastoplastic isotropic hardening

behavior and $\theta$ = 2%, for the three values of $R$ = 1, 2, 8. On the bottom, the ratios $\mu_{SYN}/\mu_{REC}$ and $\sigma_{ln,SYN}/\sigma_{ln,REC}$ are plotted as function of the period.

To provide a complete overview: Figures 19 to 26 show the ratios $\mu_{SYN}/\mu_{REC}$ and $\sigma_{ln,SYN}/\sigma_{ln,REC}$ of the SDoF peak displacements and Figures 27 to 34 show ratios between medians $\mu$ (top) and sigmas $\sigma_{ln}$ (bottom) of SDoFs ductility demands obtained with synthetics and records. It is worth noting that, in general, the ratios oscillate over the value of 1, implying that the estimates obtained with the synthetics are in overall agreement with the ones of records both in median and variability.
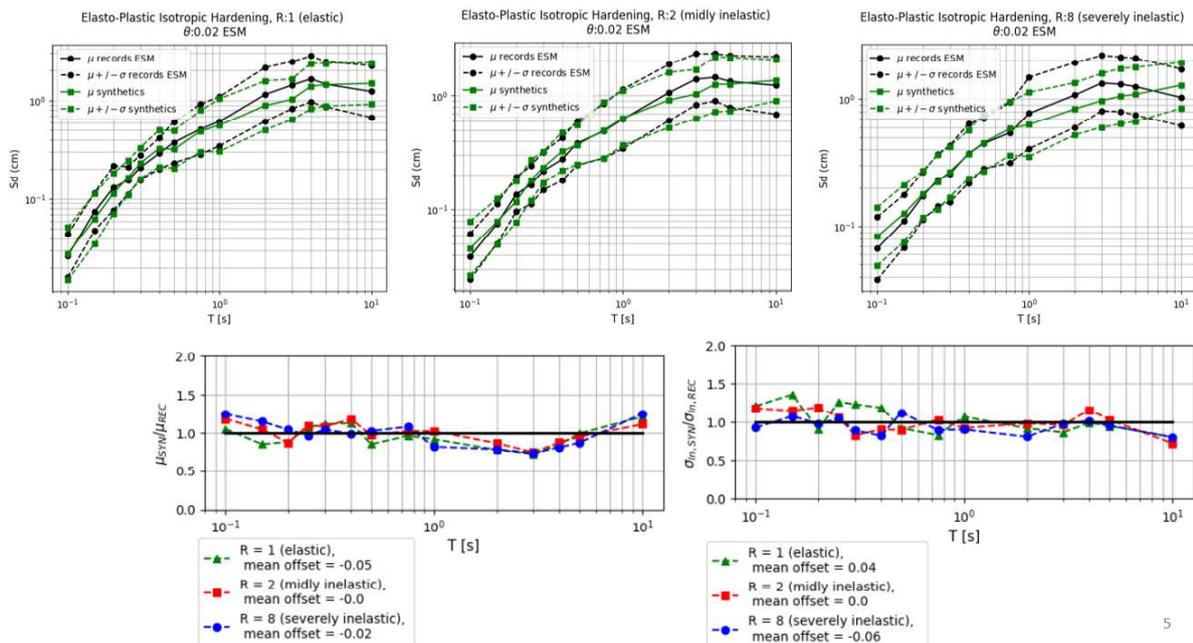


**Figure 18. Top: displacement spectra assuming an elastoplastic model with isotropic hardening and $\theta$=2%, for $R$ = 1 (left), 2 (center) and 8 (right) by using synthetics (green) and records (black) selected from ESM database. Bottom: ratios between medians $\mu$ (left) and sigmas $\sigma_{ln}$ (right) of peak displacements obtained with synthetics and records in function of $T$ and $R$ ($R$ = 1 in green, $R$ = 2 in red and $R$ = 8 in blue).**
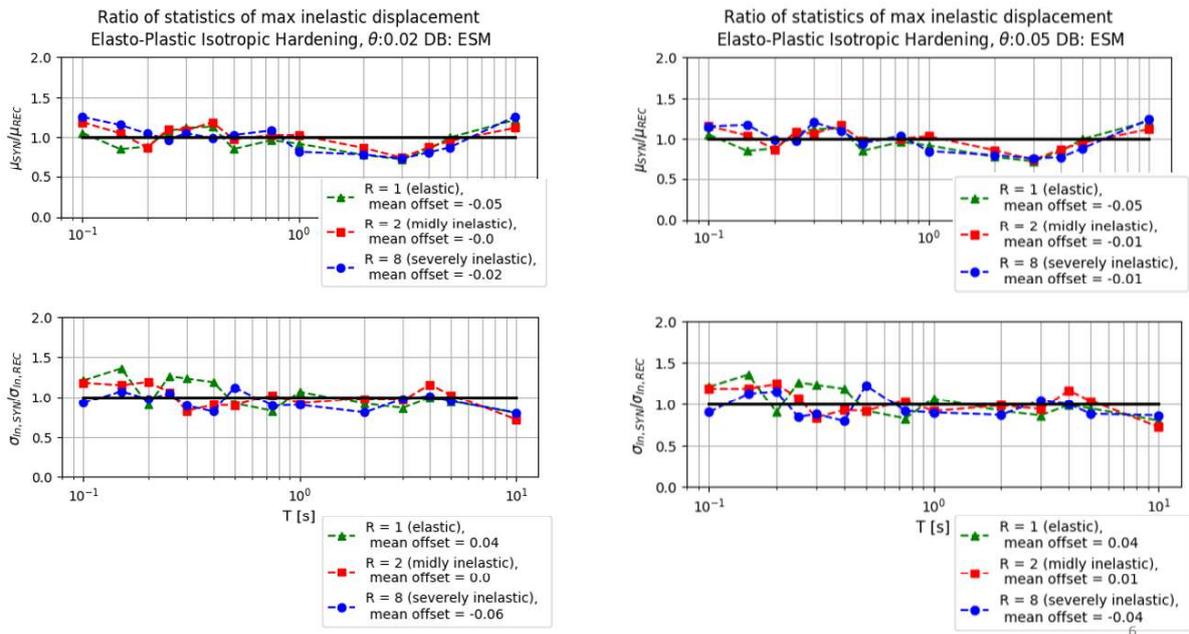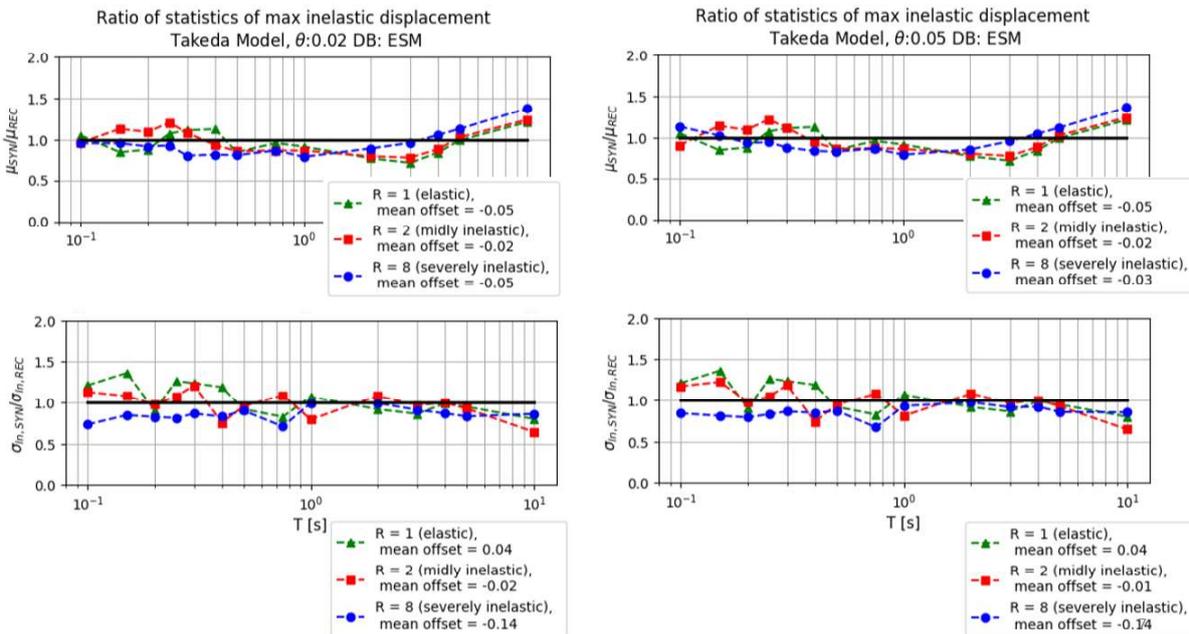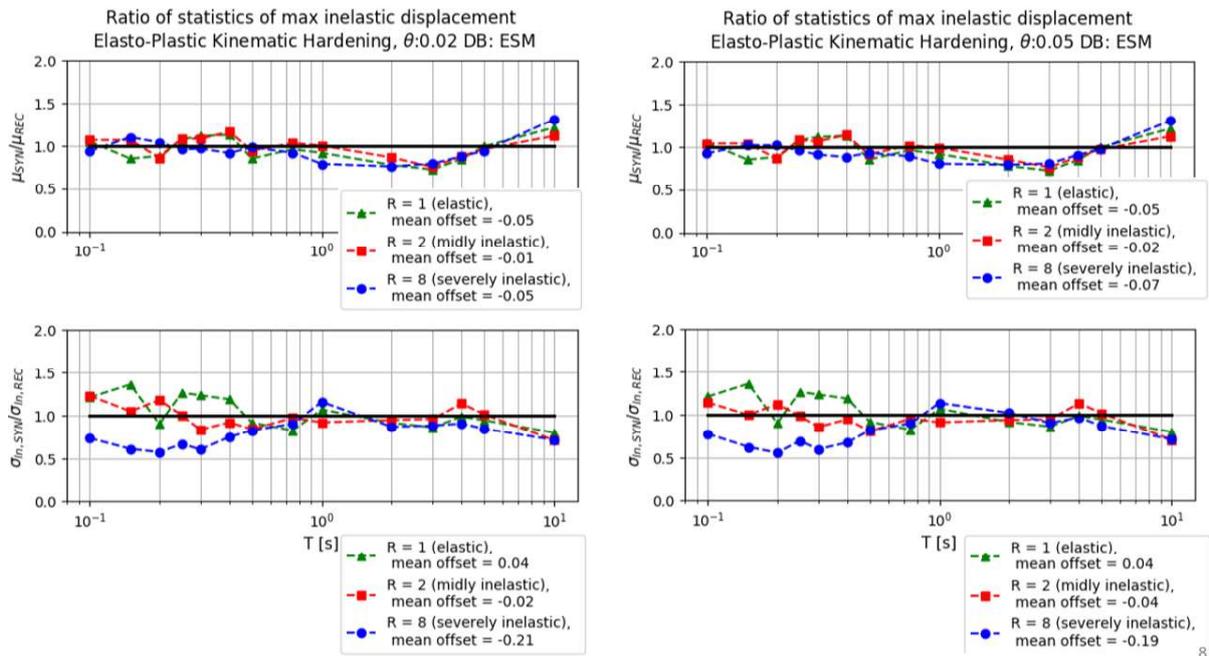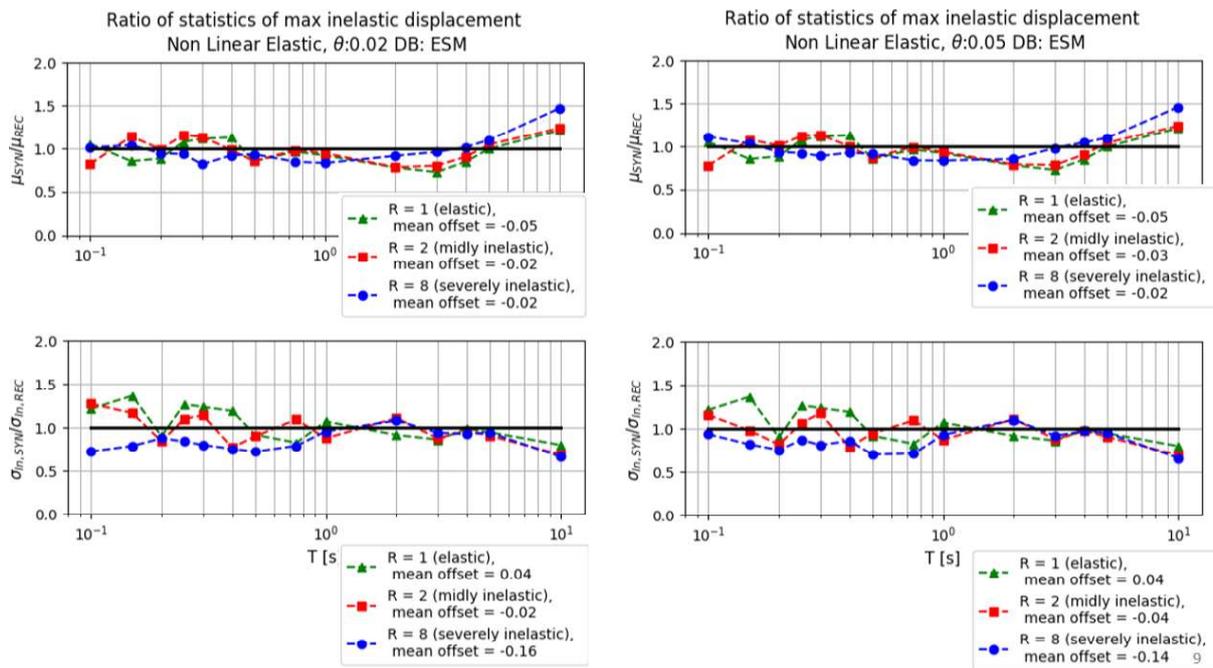
**Figure 19. Ratios between medians _µ_ (top) and sigmas _σIn_ (bottom) of SDoFs peak displacements obtained with synthetics and records from ESM database in function of _T, R_ (_R_ = 1 in green, _R_ = 2 in red and _R_ = 8 in blue) and _θ_ (_θ_=2% on the left, _θ_=5% on the right). An elasto-plastic isotropic hardening behavior is assumed.**



**Figure 20. Ratios between medians _µ_ (top) and sigmas _σIn_ (bottom) of SDoFs peak displacements obtained with synthetics and records from ESM database in function of _T, R_ (_R_ = 1 in green, _R_ = 2 in red and _R_ = 8 in blue) and _θ_ (_θ_=2% on the left, _θ_=5% on the right). A Takeda model is assumed.**
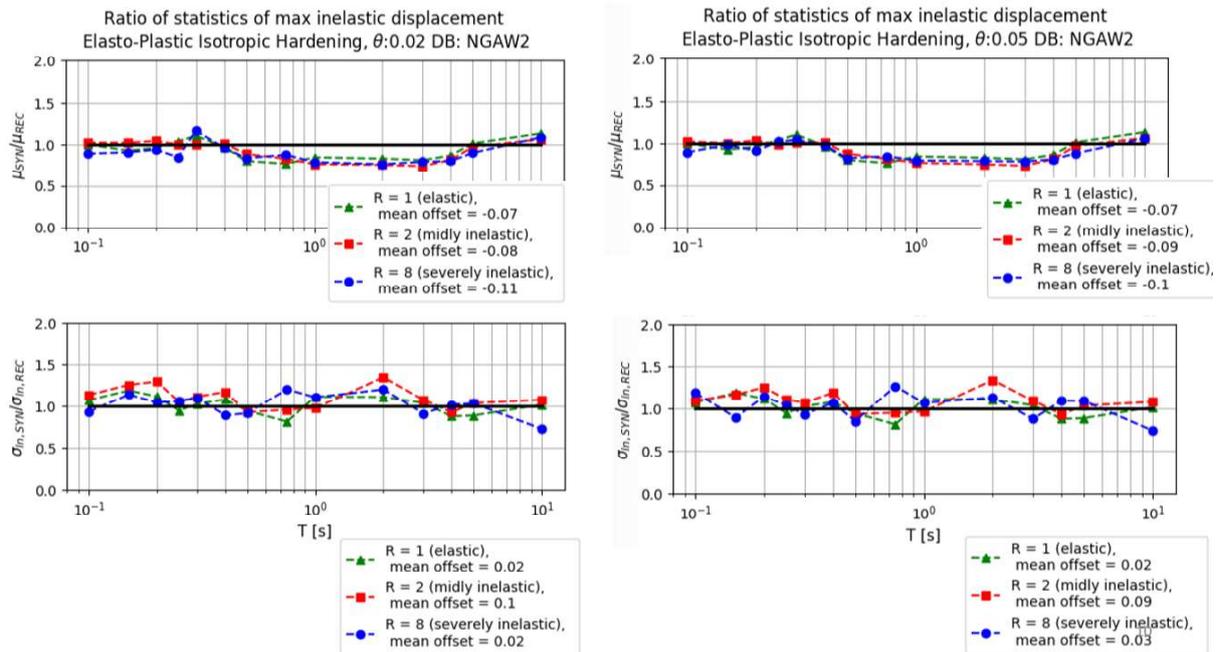
**Figure 21. Ratios between medians $\mu$ (top) and sigmas $\sigma_{ln}$ (bottom) of SDoFs peak displacements obtained with synthetics and records from ESM database in function of $T$, $R$ ($R$ = 1 in green, $R$ = 2 in red and $R$ = 8 in blue) and $\theta$ ($\theta$=2% on the left, $\theta$=5% on the right). An elasto-plastic kinematic hardening model is assumed.**



**Figure 22. Ratios between medians $\mu$ (top) and sigmas $\sigma_{ln}$ (bottom) of SDoFs peak displacements obtained with synthetics and records from ESM database in function of $T$, $R$ ($R$ = 1 in green, $R$ = 2 in red and $R$ = 8 in blue) and $\theta$ ($\theta$=2% on the left, $\theta$=5% on the right). A nonlinear elastic model is assumed.**
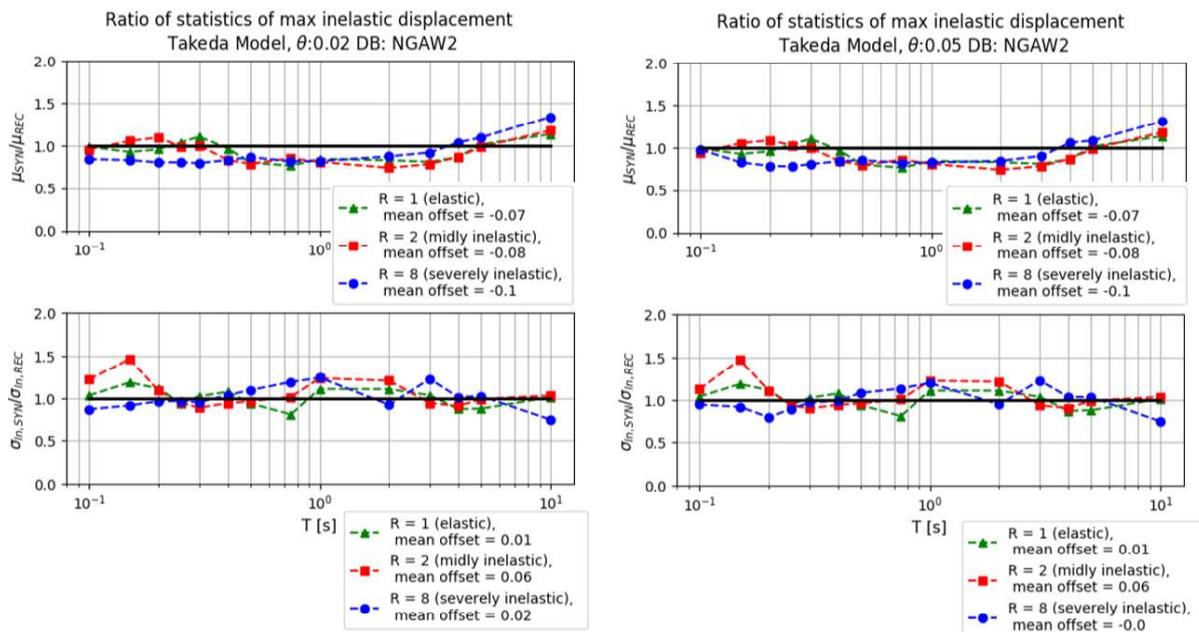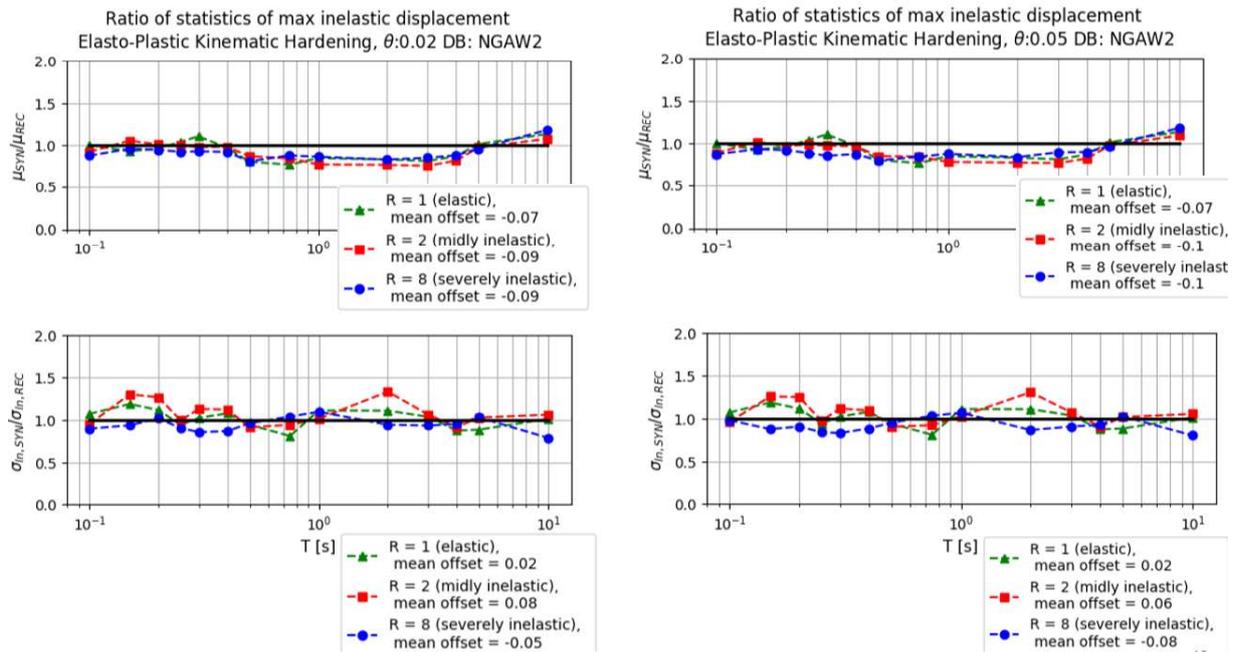
**Figure 23. Ratios between medians *μ* (top) and sigmas *σ*$_{In}$ (bottom) of SDoFs peak displacements obtained with synthetics and records from NGA-West2 database in function of *T*, *R* (*R* = 1 in green, *R* = 2 in red and *R* = 8 in blue) and *θ* (*θ*=2% on the left, *θ*=5% on the right). An elasto-plastic isotropic hardening model is assumed.**



**Figure 24. Ratios between medians *μ* (top) and sigmas *σ*$_{In}$ (bottom) of SDoFs peak displacement obtained with synthetics and records from NGA-West2 database in function of *T*, *R* (*R* = 1 in green, *R* = 2 in red and *R* = 8 in blue) and *θ* (*θ*=2% on the left, *θ*=5% on the right). A Takeda model is assumed.**
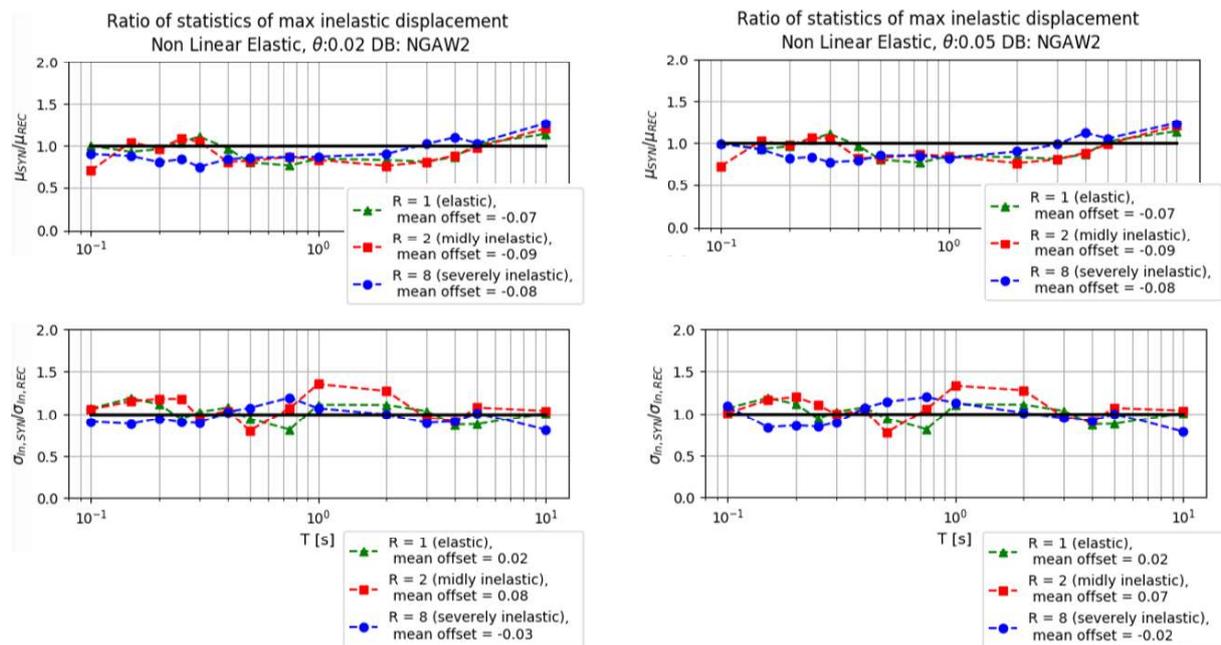
**Figure 25. Ratios between medians _μ_ (top) and sigmas _σ_In_ (bottom) of SDoFs peak displacements obtained with synthetics and records from NGA-West2 database in function of _T, R_ (_R_ = 1 in green, _R_ = 2 in red and _R_ = 8 in blue) and _θ_ (_θ_=2% on the left, _θ_=5% on the right). An elasto-plastic kinematic hardening model is assumed.**



**Figure 26. Ratios between medians _μ_ (top) and sigmas _σ_In_ (bottom) of SDoFs peak displacements obtained with synthetics and records from NGA-West2 database in function of _T, R_ (_R_ = 1 in green, _R_ = 2 in red and _R_ = 8 in blue) and _θ_ (_θ_=2% on the left, _θ_=5% on the right). A nonlinear elastic model is assumed.**
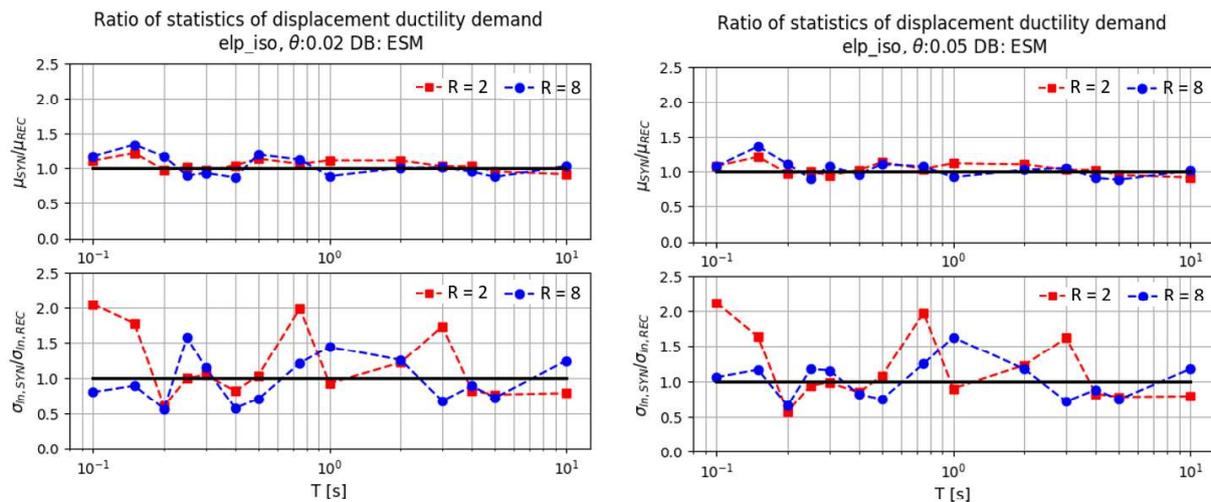
**Figure 27. Ratios between medians *μ* (top) and sigmas *σ$_{ln}$* (bottom) of SDoFs ductility demands obtained with synthetics and records from ESM database in function of *T, R* (*R* = 2 in red and *R* = 8 in blue) and *θ* (*θ*=2% on the left, *θ*=5% on the right). An elasto-plastic isotropic hardening behavior is assumed.**
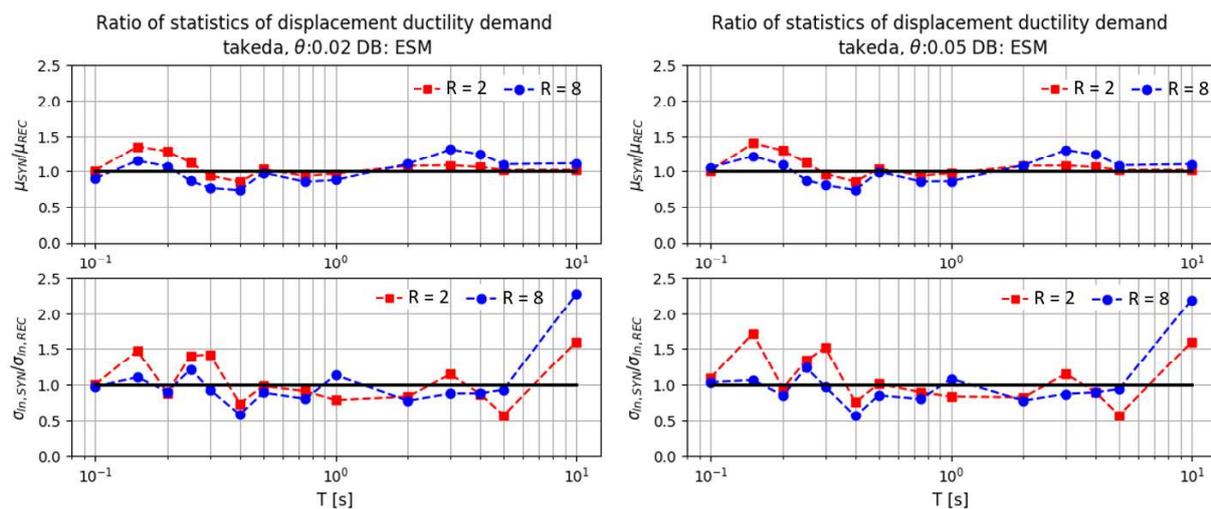


**Figure 28. Ratios between medians *μ* (top) and sigmas *σ$_{ln}$* (bottom) of SDoFs ductility demands obtained with synthetics and records from ESM database in function of *T, R* (*R* = 2 in red and *R* = 8 in blue) and *θ* (*θ*=2% on the left, *θ*=5% on the right). A Takeda model is assumed.**
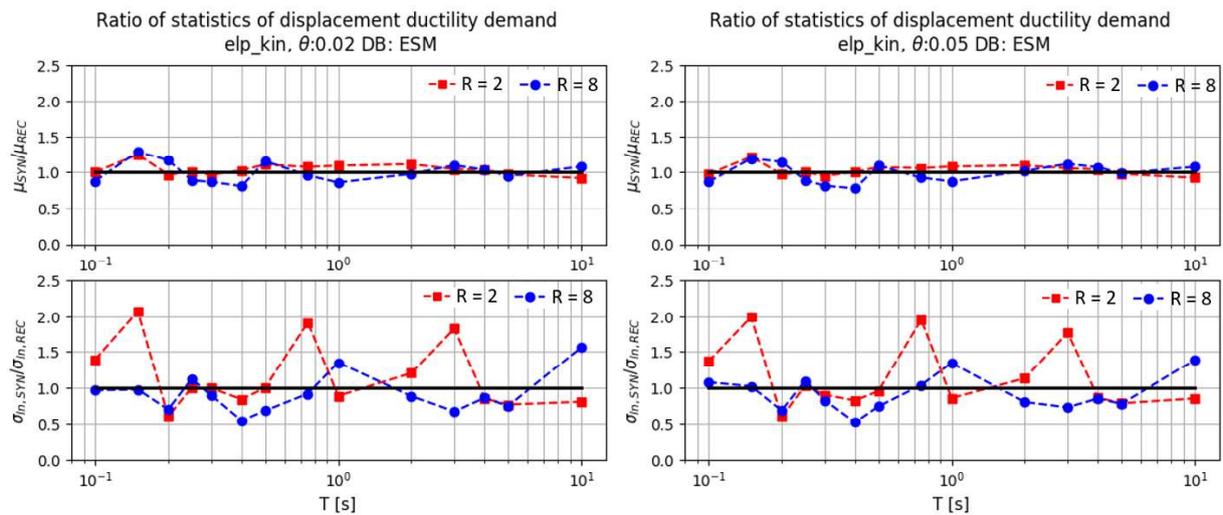
**Figure 29. Ratios between medians *μ* (top) and sigmas *σ$_{ln}$* (bottom) of SDoFs ductility demands obtained with synthetics and records from ESM database in function of *T, R* (*R* = 2 in red and *R* = 8 in blue) and *θ* (*θ*=2% on the left, *θ*=5% on the right). An elasto-plastic kinematic hardening model is assumed.**
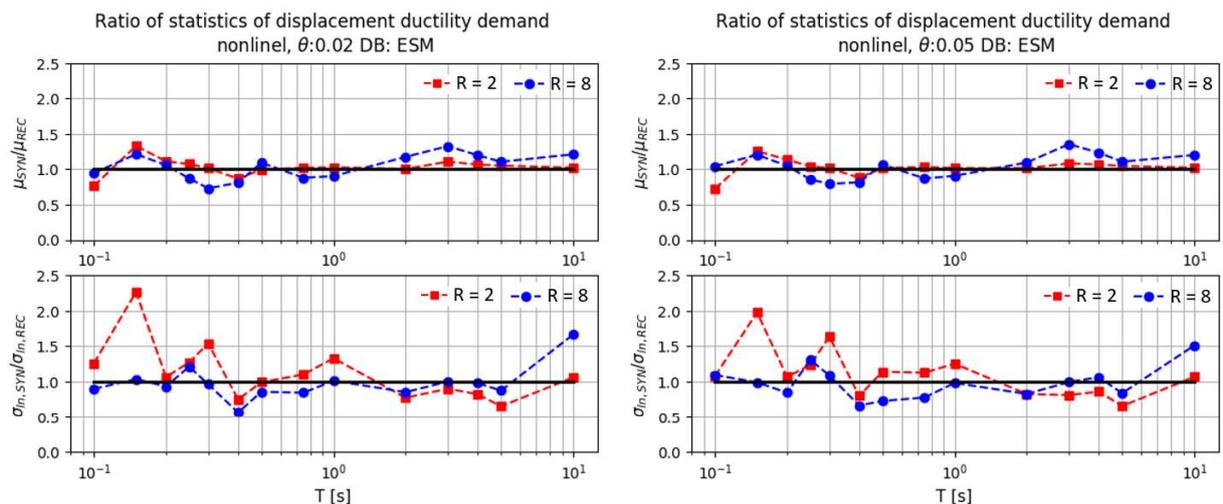


**Figure 30. Ratios between medians *μ* (top) and sigmas *σ$_{ln}$* (bottom) of SDoFs ductility demands obtained with synthetics and records from ESM database in function of *T, R* (*R* = 2 in red and *R* = 8 in blue) and *θ* (*θ*=2% on the left, *θ*=5% on the right). A nonlinear elastic model is assumed.**

**Figure 31. Ratios between medians *μ* (top) and sigmas *σ*$_{ln}$ (bottom) of SDoFs ductility demand obtained with synthetics and records from NGA-West2 database in function of *T, R* (*R* = 2 in red and *R* = 8 in blue) and *θ* (*θ*=2% on the left, *θ*=5% on the right). An elasto-plastic isotropic hardening model is assumed.**
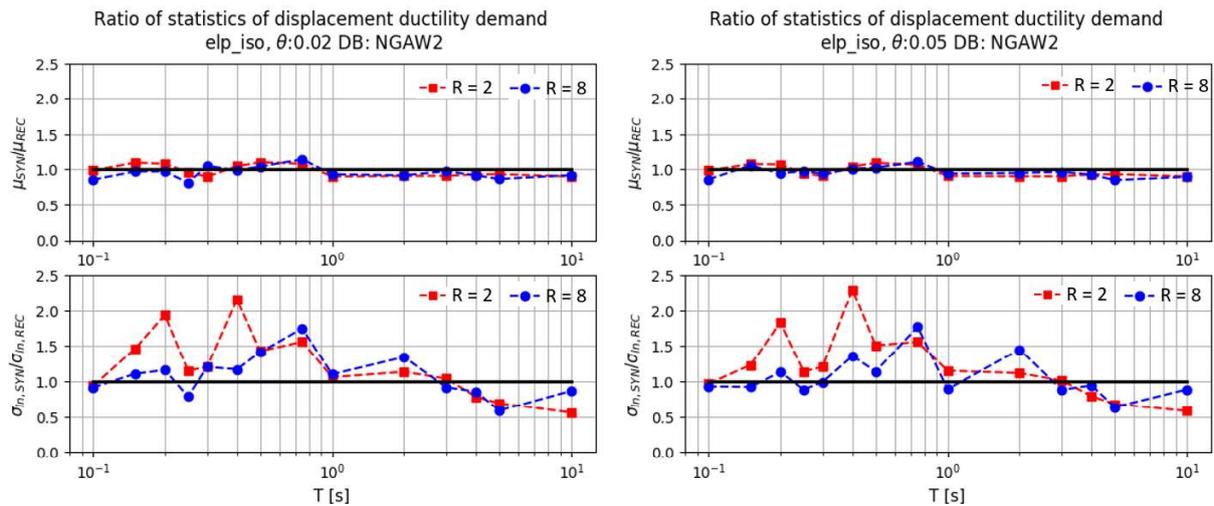


**Figure 32. Ratios between medians *μ* (top) and sigmas *σ*$_{ln}$ (bottom) of SDoFs ductility demand obtained with synthetics and records from NGA-West2 database in function of *T, R* (*R* = 2 in red and *R* = 8 in blue) and *θ* (*θ*=2% on the left, *θ*=5% on the right). A Takeda model is assumed.**
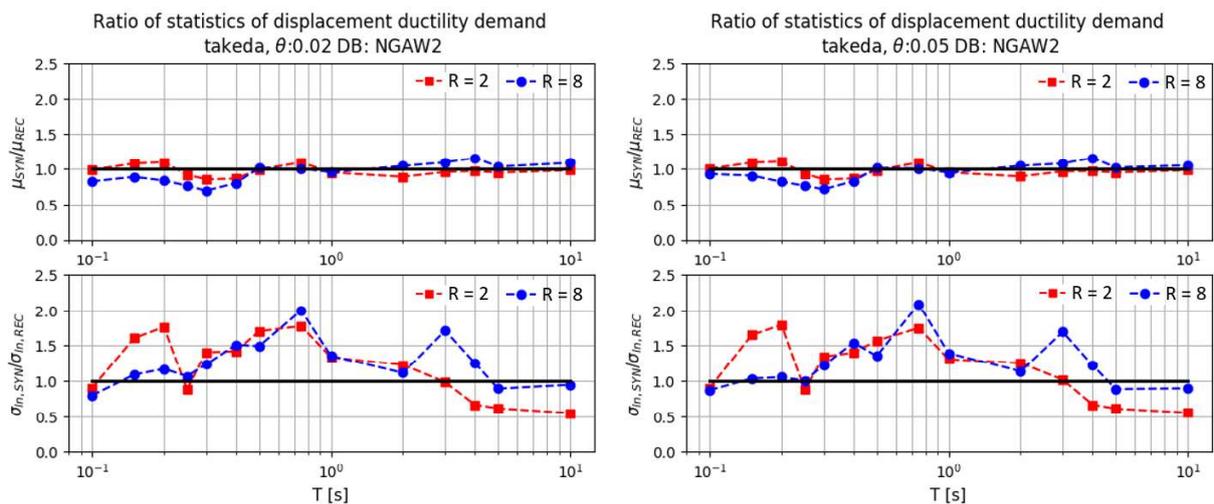
**Figure 33. Ratios between medians *μ* (top) and sigmas *σ*In (bottom) of SDoFs ductility demand obtained with synthetics and records from NGA-West2 database in function of *T*, *R* (*R* = 2 in red and *R* = 8 in blue) and *θ* (*θ*=2% on the left, *θ*=5% on the right). An elasto-plastic kinematic hardening model is assumed.**
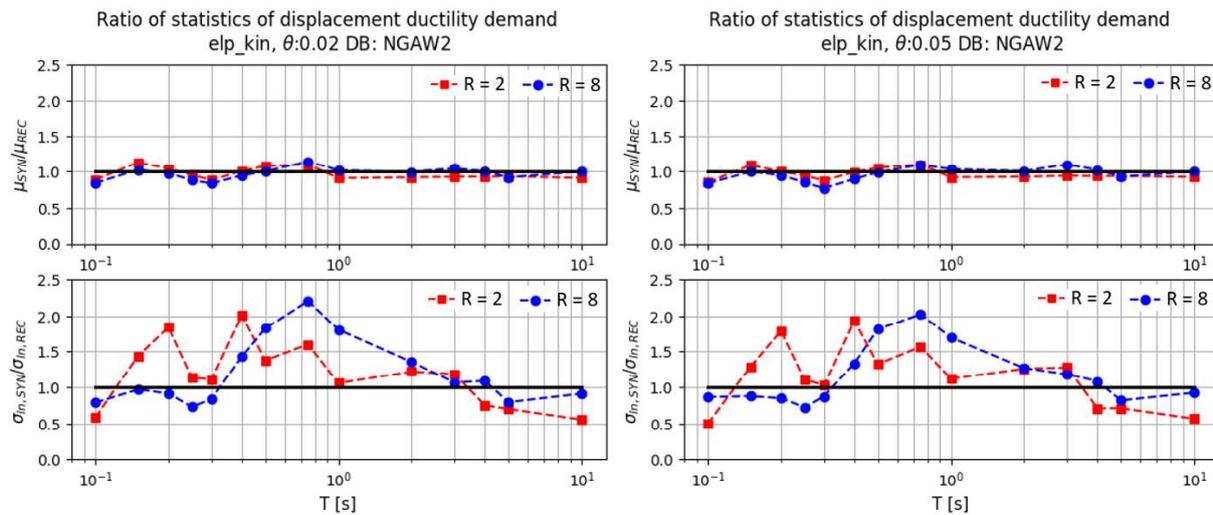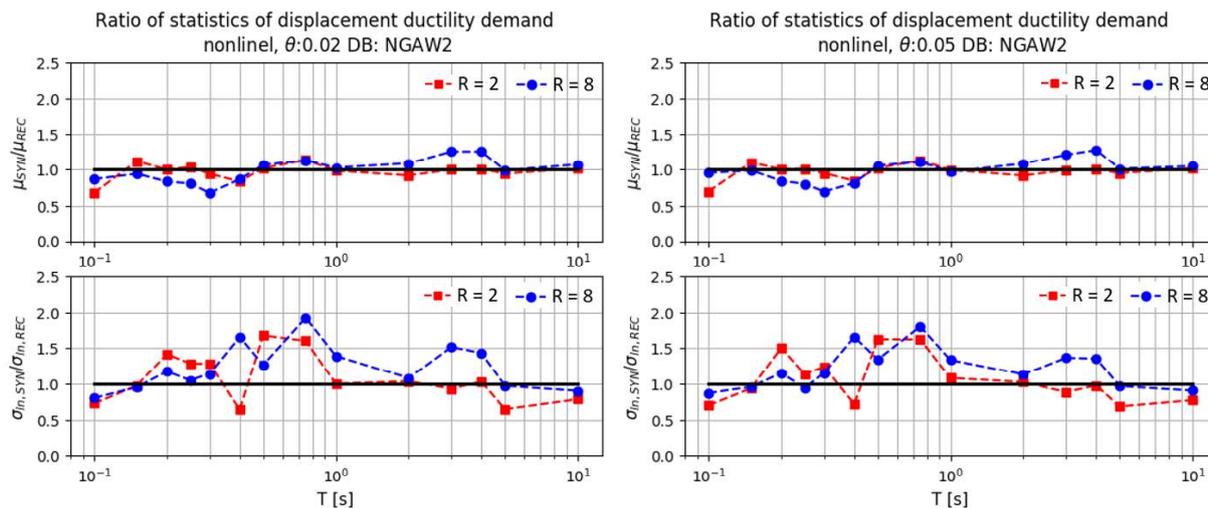


**Figure 34. Ratios between medians *μ* (top) and sigmas *σ*In (bottom) of SDoFs ductility demand obtained with synthetics and records from NGA-West2 database in function of *T*, *R* (*R* = 2 in red and *R* = 8 in blue) and *θ* (*θ*=2% on the left, *θ*=5% on the right). A nonlinear elastic model is assumed.**

# 5. Statistical significance between SDoF demands for synthetics and natural input motions.

Parametric hypothesis tests were performed to quantitatively assess the statistical significance of the differences observed in SDoFs' response (median and variance) to natural and synthetics input motions.

## 5.1 Hypothesis test for the median

The null hypothesis, $H_0$, is that the median of the engineering demand parameters considered (the mean of the natural logs) to synthetic input motions is equal to that from records. To this end, the two-tails Aspin–Welch test (Welch, 1938) was preferred over the standard Student's t-test, as the former does not require the assumption of equal, yet still unknown, variances of populations originating the samples. The employed test statistic is expressed as:

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\dfrac{s_1^2}{N_1} + \dfrac{s_2^2}{N_2}}} \qquad (14)$$

Where $\overline{X}_1$ and $\overline{X}_2$ are the sample means, $s_1$ and $s_2$ are the sample standard deviations, and $N_1$ and $N_2$ are the sample sizes (in this case $N_1 = N_2$). The test statistic, under $H_0$, has a Student's t-distribution with the number of degrees of freedom given by Satterthwaite's approximation (Satterthwaite, 1941).

The level of statistical significance is generally expressed through p-value. The latter is a number, between 0 and 1, calculated from the considered statistical test that describes the likelihood to find a particular set of observations if the null hypothesis was true. The smaller is the p-value, the stronger is the evidence that the null hypothesis $H_0$ should be rejected. Generally, a p-value less than 0.05 is considered statistically significant indicating strong evidence against the null hypothesis, as there is less than a 5% of probability that $H_0$ is correct. On the other hand, a p-value higher than 0.05 is commonly assumed not statistically significant and indicates strong evidence for the null hypothesis.

Figure 35 and 36 show the percentage of test rejection for the peak displacement (Figure 35) and ductility demand (Figure 36) obtained including all the SDoFs grouped for $T$ and $R$ values considering parametric tests between synthetics and ESM records (right) and synthetics and NGA-West2 records (left). The percentage of test rejections is low both for mildly (R=2) and severely (R=8) inelastic structures, especially when the peak displacements are considered.
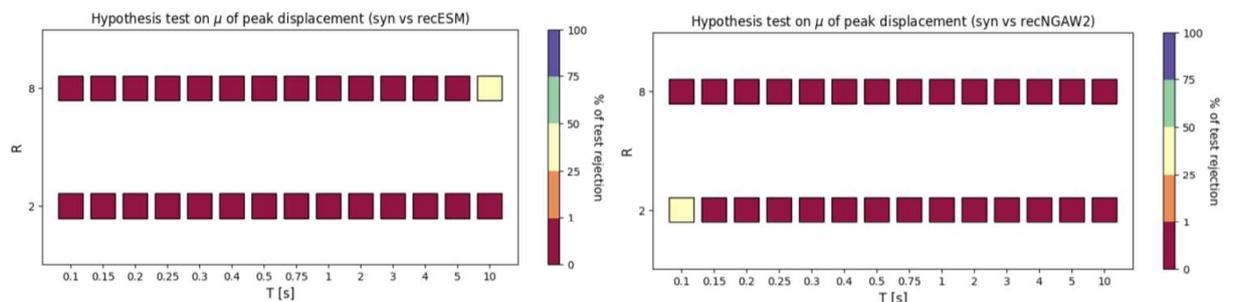


**Figure 35. Percentages of hypothesis test rejection ($\alpha$ = 0.05) of median of peak inelastic displacement comparing ESM (right) and NGAW2 (left) with the synthetics.**
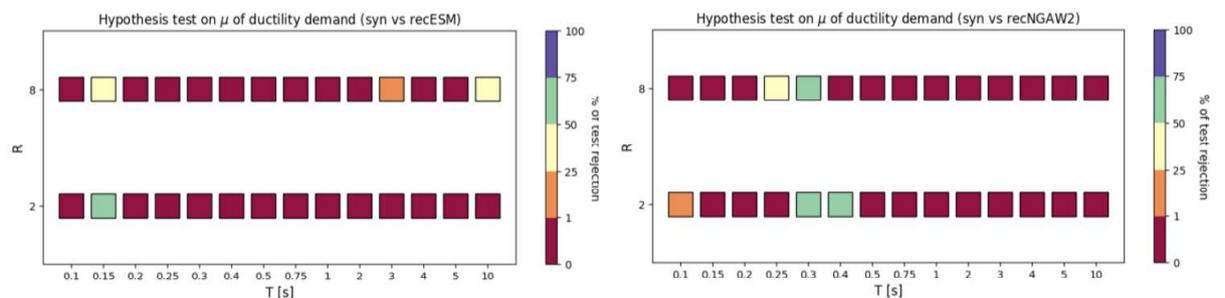


**Figure 36. Percentages of hypothesis test rejection ($\alpha$ = 0.05) for median of ductility demands comparing ESM (right) and NGAW2 (left) with the synthetics.**

## 5.2 Hypothesis test for the variance

In this case, the null hypothesis, $H_0$, is that the variance (in log term) of the engineering demand parameters from synthetics is equal to those from records. In this case the F-test for normally distributed data was performed. The employed statistics is:

$$F = \frac{s_1^2}{s_2^2} \qquad (15)$$

In which $s_1$ and $s_2$ are the sample standard deviations.

Figure 37 and 38 show the percentage of test rejection for the peak displacement (Figure 37) and ductility demand (Figure 38) obtained including all the SDoFs grouped for $T$ and $R$ values considering parametric tests between synthetics and ESM records (right) and synthetics and NGA-West2 records (left). The percentage of test rejections is low both for mildly (R=2) and severely (R=8) inelastic structures, especially when the peak displacements are considered.

The tests are accepted (i.e. there are not significant differences in variance) for all the SDoF systems when peak displacements from NGA-West2 database and synthetics are compared. On the other hand, the percentage of test rejection is not null for the tests considering ESM records versus synthetics and it increases with $R$ In general, in both cases (i.e. both comparing ESM and NGA-West2 records versus synthetics) the p-value distributions move towards lower values for increasing levels of nonlinearities. On the other hand, the percentage of tests rejection increases remarkably when the variance of the ductility demand is considered.

Moreover, Figures 41 and 42 provide the percentage of test rejections for each oscillation period $T$ and level of nonlinearity $R$.
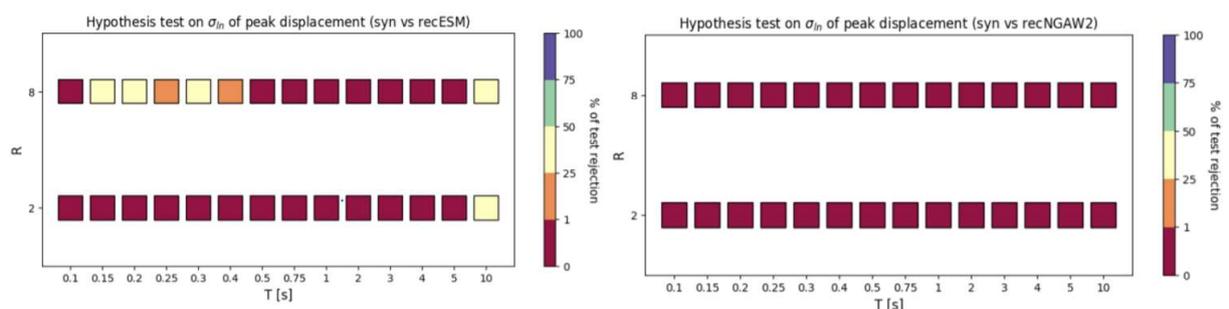


**Figure 37. Percentages of hypothesis test rejection ($\alpha$ = 0.05) of sigma of peak inelastic displacement comparing ESM (right) and NGAW2 (left) with the synthetics.**
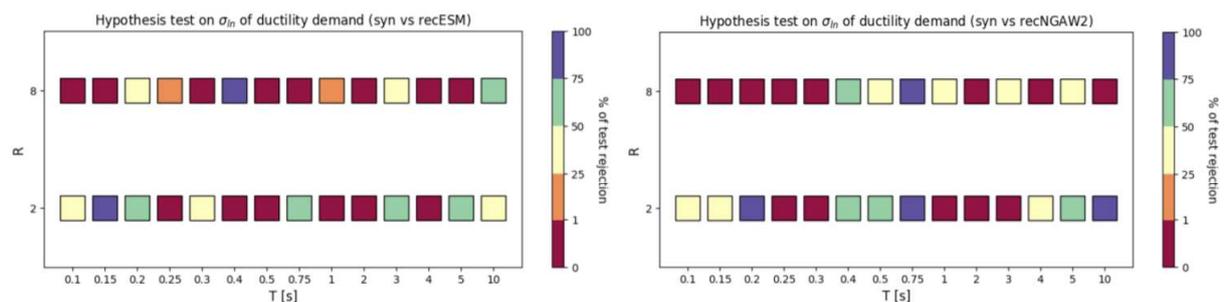
**Figure 38. Percentages of hypothesis test rejection ($\alpha$ = 0.05) of sigma of ductility demand comparing ESM (right) and NGAW2 (left) with the synthetics.**

# 6. Conclusions and further developments

This work aims to investigate whether natural and artificial ground motions generate nonlinear peak displacement responses statistically distinguishable. To this end, they were selected 20 real records from available strong motion databases and 20 synthetics generated with code-aster. The selection algorithm adopted is the one described in Tarbali and Bradley (2015), extension of Bradley (2012) for a scenario earthquake, based on the use of target random realizations from the conditional multivariate distribution of various ground motion intensity measures.

The spectral displacements were calculated for nonlinear SDoF oscillators having periods ranging from 0.1 to 10s, bilinear backbone curves, hardening ratio of 2 and 5% and strength reduction factor $R$ of 2 (mildly nonlinear) and of 8 (severely nonlinear).

In the context of the SDoF systems and the response parameter here considered, the results of this study show that, by using proper and consistent methods both for records selection and synthetics generation, the artificial time histories match well both the median and the variability of the peak seismic response produced by recorded motions. On the other hand, when the ductility demand is considered as engineering demand parameter, the comparison synthetics-records is satisfactory in median but not in variability.

Future studies will be devoted to compare other engineering demand parameters, such as the equivalent number of cycles Ne.

It is also worth noting that this first study takes into account very simple structures. Future studies may investigate more complex structures in order to better identify the limitation on the use of synthetics in realistic engineering applications.

Sensitivity analysis will be carried out in order to infer conclusions concerning the minimum number of time histories needed to achieve stable estimates of the structural response when natural and synthetics sets of signals are used as input motion for nonlinear dynamic analyses.

Finally, an extension of this work aiming to investigate also the performance of sets of synthetics properly selected from physics-based simulated ground motion databases is underway.

# 7. References

Ancheta T.D., R.B. Darragh, J.P. Stewart, E. Seyhan, W.J. Silva, B.S.-J. Chiou, K.E. Wooddell, R.W. Graves, A.R. Kottke, D.M. Boore, T. Kishida, J.L. Donahue (2014). NGA-West2 Database. *Earthquake Spectra*, **30**:989-1005.

Afshari K., and J. P. Stewart (2016). Physically parameterized prediction equations for significant duration in active crustal regions. *Earthquake Spectra,* **32**: 2057–2081.

Atkinson G. M., and K. Goda (2010). Inelastic seismic demand of real versus simulated ground motion records for the Cascadia subduction earthquakes. *Bulletin of Seismological Society of America,* **100**(1): 102–115.

Baker J.W., and Bradley (2017). Intensity Measure Correlations Observed in the NGA-West2 Database, and Dependence of Correlations on Rupture and Site Parameters. *Earthquake Spectral*, **33**(1): 145-156.

Baker J.W., and A.C. Cornell (2006). Spectral shape, epsilon and record selection*. Earthquake Engineering & Structural Dynamics*, **35**(9): 1077-1095.

Baker J.W. (2010). Conditional mean spectrum: Tool for ground-motion selection*. Journal of Structural Engineering*, **137**(3): 322-331.

Bazzurro, P., B. Sjoberg, and N. Luco (2004). Post-elastic response of structures to synthetic ground motions. *Report for Pacific Earthquake Engineering Research (PEER) Center Lifelines Program Project*, Berkeley, CA, 65–112.

Bijelic N., T. Lin, and G.G. Deierlein (2018). Validation of the SCEC Broadband Platform simulations for tall building risk assessments considering spectral shape and duration of the ground motion. *Earthquake Engineering & Structural Dynamics,* **47**: 2233–2251.

Bommer J.J. and A.B Acevedo (2004). The use of real earthquake accelerograms as input to dynamic analysis*. Journal of Earthquake Engineering*. **8**(spec01): 43-91.

Bradley B.A. (2010). A generalized conditional intensity measure approach and holistic ground-motion selection*. Earthquake Engineering & Structural Dynamics*, **39**(12): 1321-1342.

Bradley B.A. (2011). Correlation of Significant Duration with Amplitude and Cumulative Intensity Measures and Its Use in Ground Motion Selection. *Journal of Earthquake Engineering*, **15**(6): 809-832

Bradley B.A. (2012). Ground motion selection algorithm based on the generalized conditional intensity measure approach. *Soil Dynamics and Earthquake Engineering*, **48**:48–61.

Bradley B.A. (2015). Correlation of Arias intensity with amplitude, duration and cumulative intensity measures. *Soil Dynamics and Earthquake Engineering,* **78**: 89-98

Cacciola P. (2010). A stochastic approach for generating spectrum compatible fully non stationary earthquakes. *Computers and Structures*, **88**(15-16): 889–901.

Campbell K. W., and Y. Bozorgnia (2014). NGA-West2 ground motion model for the average horizontal components of PGA, PGV, and 5% damped linear acceleration response spectra. *Earthquake Spectra*, **30**: 1087–1115.

Campbell K. W., and Y. Bozorgnia (2019). Ground motion models for the horizontal components of Arias intensity (AI) and cumulative absolute velocity (CAV) using the NGA-West2 database. *Earthquake Spectra*, 35(3): 1289–1310.

Der Kiureghian A. (1981). A response spectrum method for random vibration analysis of MDF systems. *Earthquake Engineering and Structural Dynamics*; **9**(1981):419–35.

Galasso C., F. Zareian, I. Iervolino, and R. Graves (2012). Validation of ground-motion simulations for historical events using SDoF systems. *Bulletin of the Seismological Society of America,* **102**: 2727–2740.

Galasso C., P. Zhong, F. Zareian, I. Iervolino, and R.W. Graves (2013). Validation of groundmotion simulations for historical events using MDoF systems. *Earthquake Engineering & Structural Dynamics*, **42**: 1395–1412.

Graves R., T.H. Jordan, S. Callaghan, E. Deelman, E. Field, G. Juve, C. Kesselman, P. Maechling, G. Mehta, K. Milner, D. Okaya, P. Small, and K. Vahi (2011). CyberShake: A physics-based seismic hazard model for Southern California, *Pure and Applied Geophysics*, **168**: 367–381.

Iervolino I., C.A. Cornell (2005). Record selection for non linear seismic analysis of structures. *Earthquake Spectra*, **21**(3):685–713.

Iervolino I., F. De Luca, and E. Cosenza (2010). Spectral shape-based assessment of SDoF nonlinear response to real, adjusted and artificial accelerograms. *Engineering Structures*, **32**(9): 2776–2792.

Jayaram N., and J.W. Baker (2008). Statistical tests of the joint distribution of spectral acceleration values. *Bulletin of the Seismological Society of America*, **98**(5): 2231-2243.

Jayaram N., T. Lin, and J.W. Baker (2011). A computationally efficient ground-motion selection algorithm for matching a target response spectrum mean and variance. *Earthquake Spectra*, **27**(3): 797-815.

Kanai K. (1957). Semi-empirical formula for the seismic characteristics of the ground motion. *Bulletin of the Earthquake Research Institute*, **35**: 308–325.

Katsanos, E.I., A.G. Sextos, and G.D. Manolis (2010). Selection of earthquake ground motion records: A state-of-the-art review from a structural engineering perspective. *Soil Dynamics and Earthquake Engineering*. **30**(4): 157-169.

Kaul MJ. (1978). Stochastic characterization of earthquakes through their response spectrum. *Earthquake Engineering and Structural Dynamics*; 6: 497–509.

Kohrangi M., P. Bazzurro, D. Vamvatsikos, A. Spillatura (2017). Conditional spectrum based ground motion record selection using average spectral acceleration. *Earthquake Engineering & Structural Dynamics,* **46**(10): 1667–1685

Kottke A. and E.M. Rathje (2008). A semi-automated procedure for selecting and scaling recorded earthquake motions for dynamic analysis. *Earthquake Spectra*, **24**(4): 911-932.

Lanzano G., S. Sgobba, L. Luzi, R. Puglia, F. Pacor, C. Felicetta, M. D'Amico, F. Cotton, and D. Bindi (2019). The pan-European engineering strong motion (ESM) flatfile: compilation criteria and data statistics. *Bulletin of Earthquake Engineering,* **17**(2): 561-582.

Lanzano G., S. Sgobba, L. Luzi, F. Pacor, R. Puglia, C. Felicetta, and M. D'Amico (2020). The pan-European Engineering Strong Motion (ESM) flatfile: comparison with NGA-West2 database. *BGTA-Bollettino di Geofisica Teorica ed Applicata*.

Luzi L., R. Puglia, E. Russo, M. D'Amico, C. Felicetta, F. Pacor, G. Lanzano, U. Çeken, J. Clinton, G. Costa, L. Duni, E. Farzanegan, P. Gueguen, C. Ionescu, I. Kalogeras, H. Özener, D. Pesaresi, R. Sleeman, A. Strollo, and M. Zare (2016). The Engineering Strong-Motion Database: A Platform to Access Pan-European Accelerometric Data. *Seismological Research Letters*, **87**(4): 987–997.

Mazzieri I., M. Stupazzini, R. Guidotti, and C. Smerzini (2013). SPEED: SPectral elements in elastodynamics with discontinuous Galerkin: a nonconforming approach for 3D multi-scale problems. *International Journal for Numerical in Methods Engineering*, **95**: 991-1010.

McGuire R.K. (1995). Probabilistic seismic hazard analysis and design earthquakes: closing the loop. *Bulletin of the Seismological Society of America*, **85**(5): 1275-1284.

Mertens P.G., A. Preumont (1993). Improved generation of PSD functions, artificial accelerograms and spectra, fully compatible with a design response spectrum. Transactions of SMIRT 12. Elsevier Science Publishers B.V Munich.

Naeim F., and R. W. Graves (2006). The case for seismic superiority of well-engineered tall buildings. *Structural Design of Tall Special Buildings*, **14**(5):401–416.

Preumont A. (1980). A method for generating artificial earthquake accelerograms. *Nuclear Engineering and Design*, **59**(2): 357–68.

Preumont A. (1985). The generation of non-separable artificial earthquake accelerograms for the design of nuclear power plants. *Nuclear Engineering and Design*, **88**: 185

Priestley M.B. (1965). Evolutionary spectra and nonstationary processes. *Journal of the Royal Statistical Society*, **27**(2):204 – 237, 1965.

Priestley M.B. (1981). Spectral analysis and time series. Academic Press London UK.

Rezaeian S., A. Der Kiureghian (2010). Simulation of synthetic ground motions for specified earthquake and site characteristics. *Earthquake Engineering and Structural Dynamics*, **39**(10):1155–80.

Rezaeian S., and A. Der Kiureghian (2011). Simulation of orthogonal horizontal ground motion components for specified earthquake and site characteristics. *Earthquake Engineering and Structural Dynamics*, **41**(2): 335–353.

Saragoni G.R., G.C. Hart (1973). Simulation of artificial earthquakes. *Earthquake Engineering and Structural Dynamics*, **2**(3):249–67.

Satterthwaite F. E. (1941). Synthesis of variance. *Psychometrika*, **6**(5): 309–316.

Sgobba S., R. Puglia, F. Pacor, L. Luzi, E. Russo, C. Felicetta, C., G. Lanzano, M. D'Amico, R. Baraschino, G. Baltzopoulos, and I. Iervolino (2019). REXELweb: A tool for selection of ground-motion records from the Engineering Strong Motion database (ESM). *7th International Conference on Earthquake Geotechnical Engineering*, 7-20 June, Rome, Italy.

Shome N., C.A. Cornell, P. Bazzurro, and J.E. Carballo (1998). Earthquakes, records, and nonlinear responses. *Earthquake Spectra*, **14**(3): 469-500.

Tajimi H. (1960). A statistical method of determining the maximum response of a building structure during an earthquake. *2nd World Conference Earthquake Engineering*, Tokyo, 1960.

Tarbali K., and B. A. Bradley (2015). Ground-motion Selection for Scenario Ruptures Using the Generalized Conditional Intensity Measure (GCIM) method. *Earthquake Engineering & Structural Dynamics,* **44**(10): 1601–1621.

Teng G., and J. Baker (2019). Evaluation of SCEC CyberShake Ground Motions for Engineering Practice. *Earthquake Spectra*, **35**(3): 1311-1328.

Vanmarcke E.H., D.A. Gasparini (1976). Simulated earthquake motions compatible with prescribed response spectra. SIMQKE User's manual and documentation. MIT Report R76-4.

Vanmarcke E.H., G.A. Fenton, and E. Heredia-Zavoni (1997). SIMQKE-II, Conditioned Earthquake Ground Motion Simulator: User's Manual, version 2. *Princeton University, Princeton, New Jersey*, pp. 25

Wang G. (2011). A ground motion selection and modification method capturing response spectrum characteristics and variability of scenario earthquakes*. Soil Dynamics and Earthquake Engineering*, **31**(4): 611-625.

Welch B.L. (1938). The significance of the difference between two means when the population variances are unequal. *Biometrika*, **29**: 350–362.

Zentner I. (2014). A procedure for simulating synthetic accelerograms compatible with correlated and conditional probabilistic response spectra. *Soil Dynamics and Earthquake Engineering*, **63**(2014): 226-233.

Zentner I. (2016). Code_Aster Documentation R4.05.05 Generation of Seismic Signals. https://www.code-aster.org/V2/doc/default/en/man_r/r4/r4.05.05.pdf

Zentner I. (2017). Code_Aster Documentation U4.36.04 Operator GENE_ACCE_SEISME. https://www.code-aster.org/V2/doc/default/en/man_u/u4/u4.36.04.pdf

# APPENDIX 1

In this Appendix the distribution of Magnitude, soil conditions and source-to-site distances of the ground motions selected from ESM (Figures A.1 and A.2) and NGA-West2 (Figures A.3 and A.4) databases are shown. The same symbols and colors refer to the records from the same seismic events. The target values (i.e. used in the GMPEs computations) are indicated with an orange solid line.
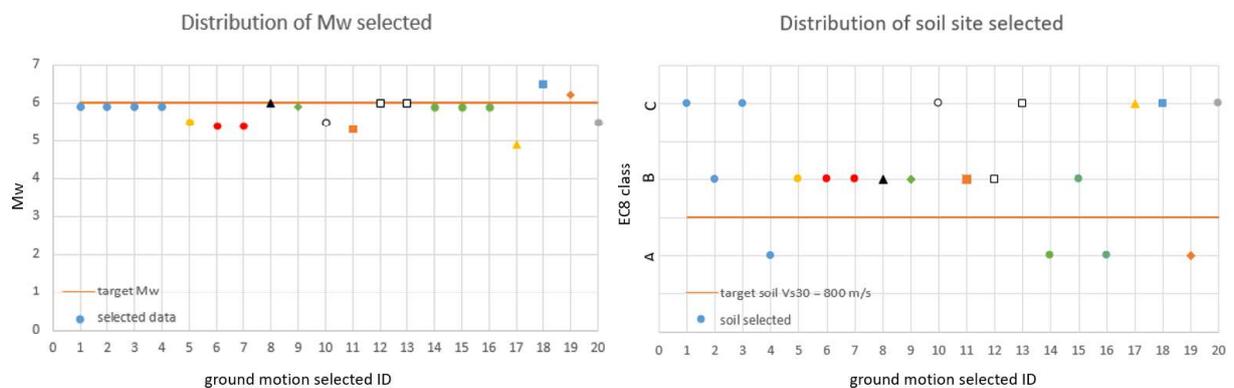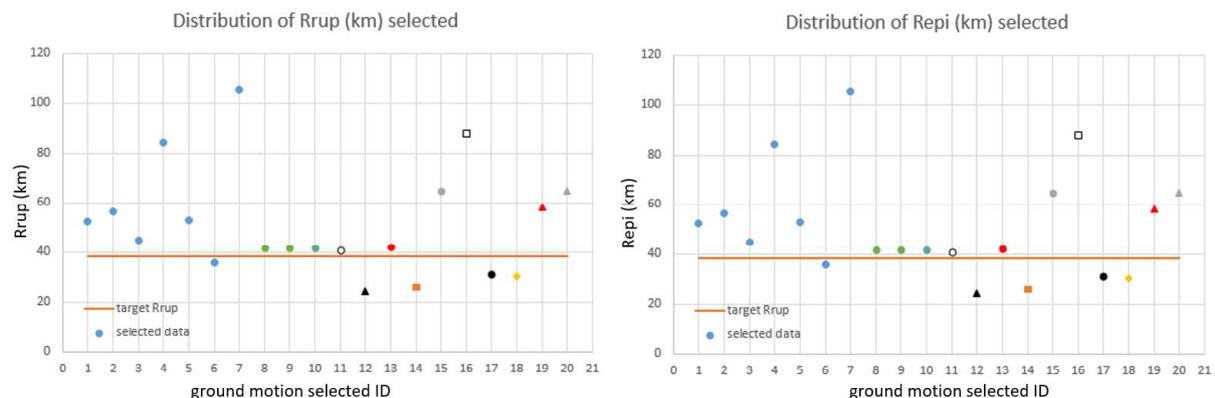


**A1. Magnitude (left) and soil site, in terms of EC8 classes, (right) distributions of the 20 ground motions selected from the ESM database. The same symbols and colors refer to the records from the same seismic events. The target values (used in the GMPEs computations) are indicated with an orange solid line. EC8 Class: A: $V_{S30}$ > 800m/s, B: $V_{S30}$ = 360-800m/s, C: $V_{S30}$ = 180-360 m/s, D: $V_{S30}$ < 180m/s.**
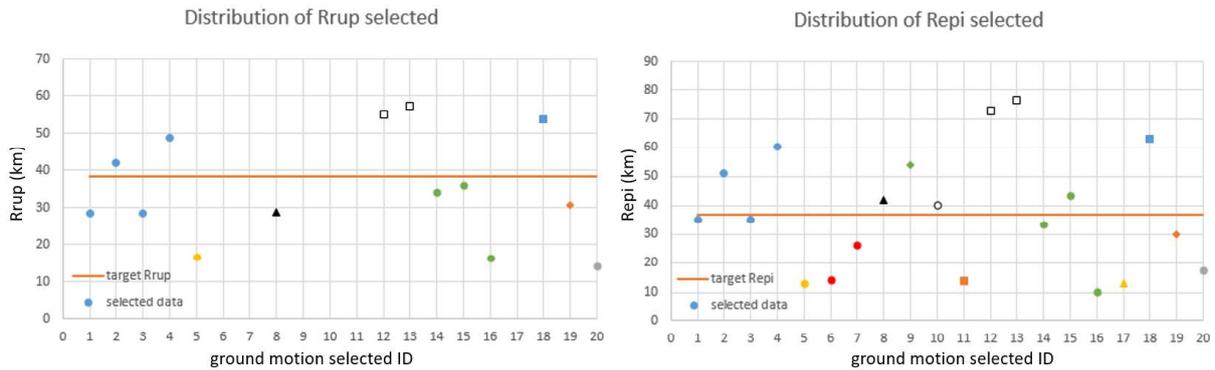
**A2. Source-to-site distances (rupture distance, Rrup, on the left and epicentral distance, Repi, on the right) distributions of the 20 ground motions selected from the ESM database. The same symbols and colors refer to the records from the same seismic events. The target values (used in the GMPEs computations) are indicated with an orange solid line.**
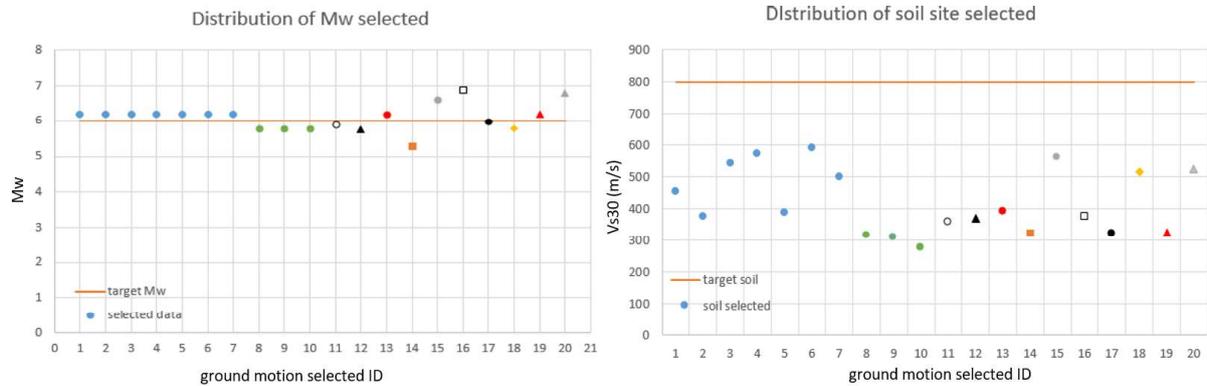


**A3. Magnitude (left) and soil site, in terms of $V_{S30}$, (right) distributions of the 20 ground motions selected from the NGA-West2 database. The same symbols and colors refer to the records from the same seismic events. The target values (used in the GMPEs computations) are indicated with an orange solid line. EC8 Class: A: $V_{S30}$ > 800m/s, B: $V_{S30}$ = 360-800m/s, C: $V_{S30}$ = 180-360 m/s, D: $V_{S30}$ < 180m/s.**



**A4. Source-to-site distances (rupture distance, Rrup, on the left and epicentral distance, Repi, on the right) distributions of the 20 ground motions selected from the NGA-West2 database. The**

same symbols and colors refer to the records from the same seismic events. The target values (used in the GMPEs computations) are indicated with an orange solid line.