Available online at www.eccomasproceedia.org
Eccomas Proceedia COMPDYN (2023) 1138-1156

ECCOMAS
Proceedia

COMPDYN 2023
9th ECCOMAS Thematic Conference on
Computational Methods in Structural Dynamics and Earthquake Engineering
M. Papadrakakis, M. Fragiadakis (eds.)
Athens, Greece, 12-14 June 2023

WHICH SIGNIFICANT DURATION OF GROUND MOTIONS TO USE? INVESTIGATIONS ON ARTIFICIAL NEURAL NETWORKS PREDICTING BUILDING SEISMIC COLLAPSE

Konstantinos Theodoros Tsalouchidis[1] **and Christoph Adam**[2]

[1]Universität Innsbruck, Unit of Applied Mechanics
Innsbruck, Austria
e-mail: konstantinos.tsalouchidis@uibk.ac.at

[2] Universität Innsbruck, Unit of Applied Mechanics
Innsbruck, Austria
e-mail: christoph.adam@uibk.ac.at

Abstract. *The effect of seismic duration on the collapse of building structures is recognized to be non-negligible. In this paper, a total of 24 spectral-based significant duration (SDs) metrics are proposed, which are an extension of the commonly employed significant durations (Ds). The 24 SDs metrics are inspired from well-known intensity measures (IM) derived from response spectra, such as first-mode spectral accelerations and average of spectral accelerations in a period range. The most intuitive SDs IM are first examined for their ability to capture the time-instance of collapse, as obtained from nonlinear response history analysis (NLRHA) of a 12-story case-study building. Following this study, the 24 SDs IM are utilized as input features to artificial neural networks (NN), in addition to spectral shape IM, with the goal of predicting seismic collapse fragilities of the case-study building as obtained by NLRHA. By developing a total of 26 NN to examine the influence of SDs, it is shown that taking into account multiple SDs IM is beneficial in collapse fragility prediction, as opposed to considering only Ds IM or no duration-based IM.*

Keywords: Seismic duration, Significant duration, Neural networks, Nonlinear response history analysis, Steel frames, P-Delta effect.

ISSN:2623-3347
Peer-review under responsibility of the organizing committee of COMPDYN 2023.
doi: 10.7712/120123.10464.20261

1 INTRODUCTION

The seismic response of a structure to ground motion (GM) excitation is affected by its amplitude, frequency content, and duration [10]. Amplitude and frequency content are typically accounted for by response spectra, either when utilizing rigorous nonlinear response history analyses (NLRHA) or simplified approaches such as pushover analyses or response spectrum analyses. Duration effects are commonly considered indirectly, for example, by considering GM with specific magnitude or other seismological characteristics that have been shown to be correlated with earthquake duration. Alternatively, there is also dedicated research to explicitly account for the effect of duration [10, 11], suggesting that the structural response can be much better identified and explained when the influence of duration is considered in addition to response spectra.

In the following, a brief overview of previous studies on the effect of GM duration on structural response is given, indication the motivation for the work presented here. This is followed by a detailed description of the current investigation, its aim and objectives.

1.1 Background and motivation

Early work on the effect of GM duration on seismic demand was inconclusive as to whether the effect of duration was statistically significant [8, 15]. Intuition suggests that the inelastic response of a structure depends on both the amplitude of the shaking and its duration, with the former responsible for cracking or yielding damage states and the latter affecting the number of cycles and the accumulated inelastic deformation [8]. In the study by [12], sensitivity analyses for structural model parameters indicated that structures with high deformation capacities and rapid rates of cyclic deterioration are most sensitive to duration, supporting the above-mentioned intuition.

An extensive list of studies (e.g., [33, 4]) has examined the effects of GM duration on structural response. A thorough literature review can be found in [10], which leads to the broad understanding that (i) the nature of GM studied; (ii) the metrics used to quantify duration; (iii) the characteristics of the structural models employed; and (iv) the examined structural demand parameters, affect the conclusions about the impact of GM duration on structural response. In that work [10], it is generally concluded that the hysteretic response of structural components is generally different under long-duration GM compared to short-duration GM. Moreover, it is argued that peak responses are not significantly affected, however an effect of duration on collapse capacity has been found, quantified by collapse fragility curves.

It is noted that over the last 50 years, a large number of studies have been conducted to examine the effects of GM duration, using a total of more than 30 duration-based metrics, as shown and explained in [9]. Duration-based metrics are commonly classified as "bracketed", "uniform", "significant", or "effective", as well as "structural" or "frequency-based" (e.g., [9, 26] for definitions). The wealth of duration-based metrics arises from the need to identify the "strong" part of the GM responsible for the structural responses (or damage) of interest. Therefore, the various duration metrics are the result of different definitions of the strong part of the GM.

Among them, the significant duration [27] has received the most attention and has been shown to be the most suitable metric to characterize ground motion duration for structural analysis since it is highly correlated with the structural response [9, 11, 12, 7, 26]. The calculation of the significant duration (Ds_{i-j}) is defined as the time interval in which $i\%$ and $j\%$ of the integral I in Eq. 1 is accumulated,

$$I(t_{max}) = \int_0^{t_{max}} a(t)^2 dt \tag{1}$$

where $a(t)$ is the acceleration of the GM at time t and t_{max} is the total time of the GM excitation.[1] It should be noted that the most commonly employed metrics for significant duration use the 5% threshold as the lower limit and 75% or 95% thresholds as the upper limit corresponding to the time instances t_5, t_{75} and t_{95} respectively, such that $\frac{I(t_5)}{I(t_{max})} = 5\%$, $\frac{I(t_{75})}{I(t_{max})} = 75\%$ and $\frac{I(t_{95})}{I(t_{max})} = 95\%$. This results in the significant durations of $Ds_{5-75} = t_{75} - t_5$ and $Ds_{5-95} = t_{95} - t_5$, which have numerous applications in geotechnical and structural engineering (e.g., [12, 11, 7, 18, 26, 3]). The choice between Ds_{5-75} and Ds_{5-95} is not apparent and is assumed to be structure dependent [12]. It is also noted that the integral I in Eq. 1 is analogous to the Arias intensity $I_A = \frac{\pi}{2g} I$ [2], which is well correlated with structural damage caused during earthquakes.

Similar to the calculation of significant durations, a proposal to include frequency-dependent duration metrics was explored in [16, 6] by employing single-degree-of-freedom (SDOF) oscillators and utilizing the SDOF acceleration signal instead of the GM as $a(t)$ in Eq. 1. Doing this for each period and determining the significant duration for each SDOF gives the "duration spectra". It can be assumed that the significant duration calculated from the SDOF pseudo-acceleration signal versus the one from the GM is the duration-equivalent to the peak ground acceleration (PGA) amplitude versus the spectral accelerations (Sa). In this respect, it is reasonable to study their effect considering the well-studied importance of the first-mode spectral acceleration ($Sa(T_1)$) or the average of spectral accelerations (Sa_{avg}) as intensity measures (IM) compared to PGA.

1.2 Current investigation, aim and objectives

The overwhelming list of available duration-based metrics, as well as the numerous studies on the impact of GM duration on structures, are indicative of its importance. Among the proposed duration-based IM, significant duration has attracted a lot of research attention. However, despite the extensive research, there are no conclusive observations and justification for the lower and upper bound used for significant duration, and no thorough investigation on the benefits of including frequency-dependent significant duration metrics.

In this study an attempt is made to investigate the range of bounds $i\%$ and $j\%$ of the significant duration Ds_{i-j}, such that the most meaningful and informative metrics are obtained. Moreover, multiple frequency-dependent (or spectral) significant duration metrics $SDs_{i-j}^{T=T_w}$ are investigated (for the same bounds $i\%$ and $j\%$), calculated from the pseudo-acceleration signal of an SDOF oscillator with period T_w and damping $\zeta = 5\%$. Additionally, more promising significant duration metrics are considered that extend the concept of $SDs_{i-j}^{T=T_w}$ as the spectral significant duration at period T_w to $SDs_{i-j}^{T_{low}-T_{high}}$ as the average of the significant durations in the period range $[T_{low}\ T_{high}]$.

The objective is to utilize these metrics of duration as input features to artificial neural networks (NN) to predict the collapse intensity of buildings. As a case study, a 12-story steel-frame building structure is investigated by NLRHA using a vast number of GM. The structure was designed according to modern seismic codes and modeled considering all important phenomena

[1] the time instance t_{max} can be considered any time such that the integral $I(t_{max})$ does not increase for any time $t > t_{max}$

occurring during GM excitation, such as P-Delta effect as well as cyclic and in-cycle material degradation, as explained in Section 2.

In Section 3, clues are given regarding the meaningful bounds ($i\%$ and $j\%$) for Ds_{i-j}, $SDs_{i-j}^{T=T_w}$ and $SDs_{i-j}^{T_{low}-T_{high}}$ for various choices of T_w and $[T_{low}\ T_{high}]$, by assessing the time instances at which the maximum interstory drift ratio (MIDR) occurred and translating these into the corresponding bounds. Finally, multiple NN are trained with the goal of predicting the collapse intensity of the case-study building subjected to GM excitation. The various NN differ in the input features used to determine the collapse intensity, with various duration-based metrics (in addition to spectral shape) employed to evaluate their influence. Comparisons are made through the predictive ability of the NN in terms of collapse fragilities, using the collapse fragilities obtained from the NLRHA as a benchmark. The predicted versus actual collapse fragilities (i.e., of NN versus NLRHA) demonstrate which of the significant duration-based metrics examined in this work are most informative in predicting seismic collapse of buildings by NN.

2 CASE STUDY BUILDING AND STRUCTURAL RESPONSE DATABASE

The 12-story steel moment frame structure used as a case study in this work is represented by the typical plan configuration in Fig. 1 and the mechanical model in Fig. 2. The building is designed in accordance with ASCE/SEI 7-05 (2006), AISC 341-05 (2005), and AISC358-05(2005) provisions and is analytically described in [17] (with the archetype ID number "5RSA"). The plan view in Fig. 1 is identical for all stories, with the blue area representing the tributary area of gravity loads applied directly to the frame and the green area representing the area of gravity loads affecting only the lateral behavior of the frame.

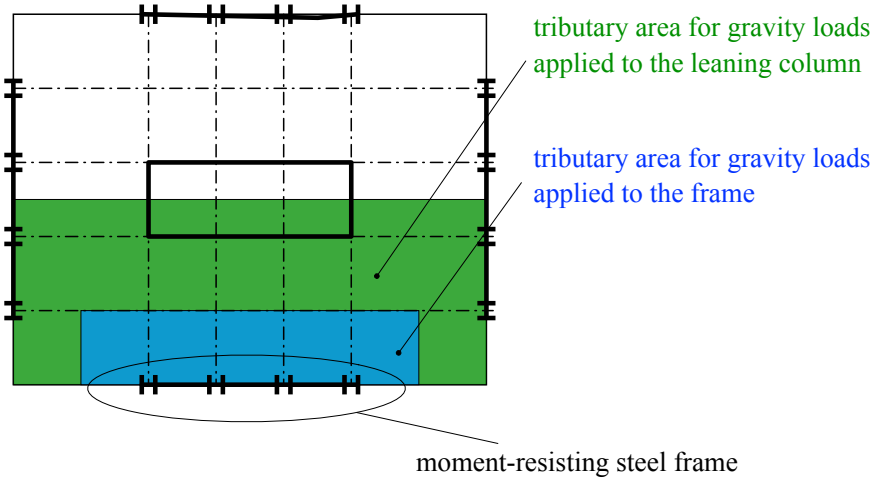

Figure 1: Characteristic plan configuration of the case study building (modified from [13])

To account for the influence of indirect loads on the modal properties and the P-Delta effect, the corresponding masses are applied to an axially rigid leaning column with zero flexural stiffness, arranged in parallel to the frame (Fig. 2). The planar structural model is implemented in OpenSees [22], where the masses are lumped and the nonlinear behavior of the members is modeled following the concentrated plasticity approach with two nonlinear springs at the ends of each element (Fig. 2). The moment-rotation properties of the plastic hinges follow the modified Ibarra-Medina-Krawinkler model [14, 20, 21] with bilinear hysteretic response.

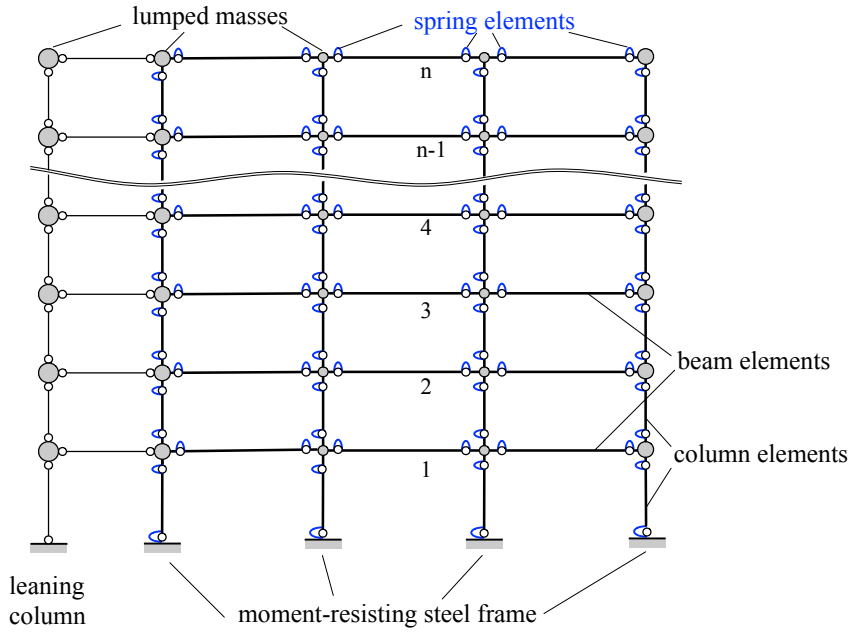

Figure 2: Concentrated plasticity model of the case study building (modified from [13])

The case-study building model is able to simulate all the important phenomena that occur during a GM excitation, such as the cyclic and in-cycle material degradation through the elaborate material model, and the P-delta effect through the implemented mechanical model. Moreover, it is particularly robust and reliable for NLRHA from linear response to collapse limit state- Hence, it is considered an excellent choice for investigating the research questions in a realistic building (and modeling) formulation.

The three fundamental periods of the structure are $T_1 = 3.22s$, $T_2 = 1.13s$ and $T_3 = 0.66s$, shown in Table 1. Additionally, the archetype ID in the report [17] and the three fundamental periods of vibration and the number of stories are also given. In parenthesis next to the fundamental period is the corresponding value calculated in [17], which shows good agreement (given the differences in the software employed, i.e., OpenSees as opposed to a "modified version of the DRAIN-2DX analysis program" used in [17]. This is a rather flexible structure, with both higher modes and P-delta effects affecting the response.

Table 1: Structural model and the first three periods of vibration (modified from [28])

Archetype ID	Stories	T_1 (s)	T_2 (s)	T_3 (s)
5RSA	12	3.22 - (3.12)	1.13	0.66

By extracting as many GM records as possible from the PEER NGA-West2 database [1], a total of 17,150 GM were collected for this study, which cover a wide range of earthquake scenarios and GM IM, with the vast majority being low-amplitude GM. In this collection of GM, with the exception of a small part,[2] each time history of accelerations is followed by information on the seismological characteristics of magnitude (M_w), source-to-site distance (R_{JB}), and

[2]For about one thousand GM (i.e., approximately 5%) the information other than the time histories of accelerations is missing

average shear-wave velocity in the top 30 m ($V_{s,30}$). Incremental dynamic analysis (IDA) [31] was performed for each GM resulting in 428,750 NLRHA. For each NLRHA, the scale factor was stored, an indication of whether a collapse occurred, and the displacements, drifts and accelerations of each story / floor were monitored as the engineering demand parameters (EDP) of interest. It should be noted that in this study, collapse was identified when an MIDR value greater than 0.1 was reached, which is a commonly considered a threshold [31, 28]. The EDP values to be stored were selected at the time instances of excitation when a peak EDP of each story / floor occurred, as well as residual values. Moreover, the value of these time instances were also stored, which is particularly useful for the investigations in this work and in particular in Section 3.

3 THE VARIOUS SIGNIFICANT DURATION METRICS

As mentioned in Section 1.1, the search for the "strong" part of the GM is the starting point for any duration-based metric, assuming that this part is responsible for the structural responses (or damage) of interest. To illustrate the commonly used bounds of significant duration metrics (i.e., 5%, 75% and 95% bounds), Fig. 3 shows the CHICHI_CHK-E GM as an example. The dashed vertical lines correspond to the time instances t_5, t_{75} and t_{95}, which define the significant duration metrics of $Ds_{5-75} \approx 17$ s and $Ds_{5-95} \approx 27$ s, respectively, while the entire GM record is up to approximately 60 s (and shown here up to 70 s). The question at hand is whether these time instances encompass the part of the GM that is essentially responsible for the structural responses of interest such as MIDR, so that these time duration metrics can successfully characterize this GM.

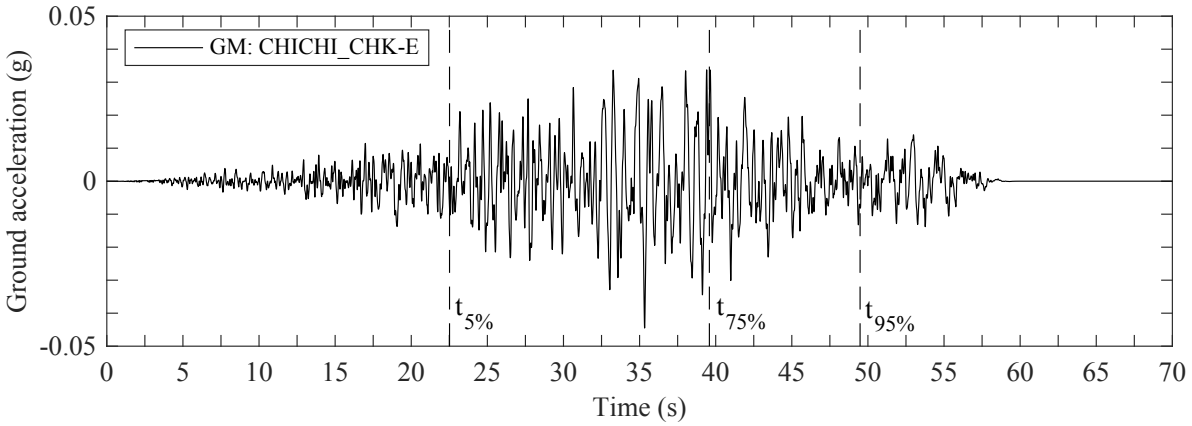

Figure 3: The time-history of accelerations of the GM record CHICHI_CHK-E and the time instances when 5%, 75% and 95% of the integral $I(t_{max})$ is reached

Before addressing this question, it should be noted that amplitude scaling of a GM as the one shown in Fig. 3 does not change any of the duration or time instances shown, but only the corresponding amplitudes. However, amplitude-scaled time series of accelerations originating from the same GM are likely to induce structural responses such as MIDR at different time instances (also shown later in Fig. 5). In this way, the significant durations remain the same for any amplitude-scaled version of the GM, whereas the time-occurrence of structural responses of interest is likely to be amplitude-dependent. Amplitude scaling is widely utilized in the literature [28], for example, in IDA where amplitude scaling is performed until the collapse

limit state is reached. Therefore, the above question also needs to examine the ability of Ds_{5-75} or Ds_{5-95} (or any other significant duration metric) to describe the part of the GM that is significant for the various (relevant) amplitude-scaling versions of the GM.

Following the example of Fig. 3, the IDA for the case-study building subjected to the CHI-CHI_CHK-E GM until the first collapse is shown in Fig. 4, where the circle markers correspond to non-collapse and the cross marker is the collapse NLRHA at the lowest intensity. For each of the NLRHA shown in Fig. 4, the time instances at which each story exhibits its peak interstory drift ratio (IDR) throughout the NLRHA are determined, as well as the maximum IDR among all stories (i.e., the MIDR). In Fig. 5, the time instances of the first and last occurred IDR among all stories as well as the time instance of the MIDR for each NLRHA in Fig. 4 are shown. The selection of these responses aims to show the occurrence of both local and global extrema of the NLRHA (i.e., the IDR and MIDR), shown in Fig. 5 with circle and asterisk markers respectively. It is clear that the occurrence of IDR and MIDR depends on the intensity, since the circle and asterisk markers do not correspond to the same time instances as intensity increases. The time of collapse corresponding to the asterisk marker at an intensity of 100% shows that collapse occurs at 57 s. This time instance is practically at the end of the GM excitation, significantly further from the t_{95} bound (Fig. 3) and corresponding to a time-bound of $t_{99.9}$. Examples such as this, where the commonly employed significant durations Ds_{5-75} and Ds_{5-95} do not account for the time instance of collapse, cast doubt on their success as predictors of structural response.

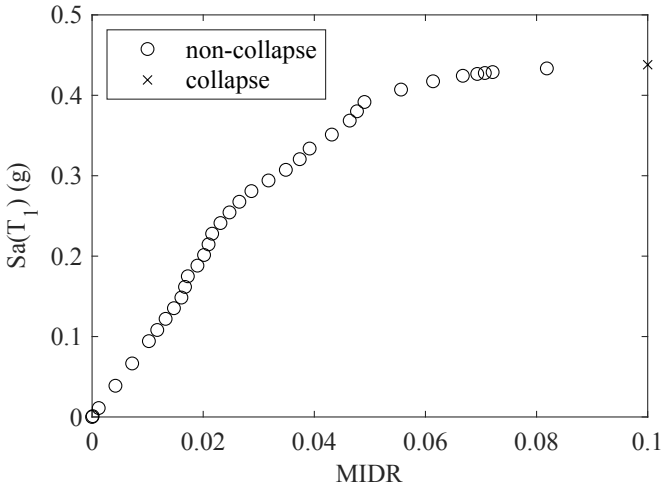

Figure 4: IDA results for the 12-story case study building subjected to the CHICHI_CHK-E GM as excitation

In an effort to investigate further the time instance of collapse and whether it can be better identified from the elastic response of an SDOF oscillator, the concept of spectral significant duration is now examined. In Fig. 6, the pseudo-acceleration elastic response of an SDOF oscillator is shown with a period equal to the fundamental period of the structure $T = T_1 = 3.22$ s and damping $\zeta = 5\%$. The vertical dashed lines correspond to the t_5, t_{75} and t_{95} time bounds and thus define the spectral significant durations of $SDs_{5-75}^{T=T_1} \approx 25$ s and $SDs_{5-95}^{T=T_1} \approx 33$ s. It is noticeable that the time instance of 57 s, when the collapse occurs, is included in the spectral significant duration $SDs_{5-95}^{T=T_1}$. This is because the response of this SDOF oscillator is not negligible at the end of the GM excitation and lasts approximately 70 s, as shown in Fig. 6, which seems to be more explanatory in terms of the expected structural response compared

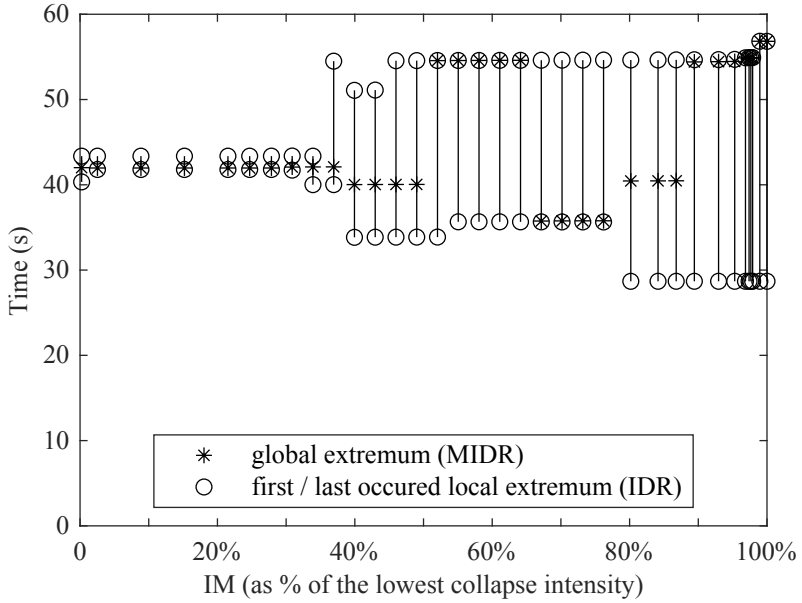

Figure 5: The time instances of first and last occurred local (IDR) and global (MIDR) extrema of drift responses of the the 12-story case study building subjected to the CHICHI_CHK-E GM as excitation

to the GM excitation. In this case, by applying Eq. 1 to calculate the integral I, it becomes clear that t_{max} should be larger than in the case of the GM excitation (Fig. 3) to ensure that the response has decayed.

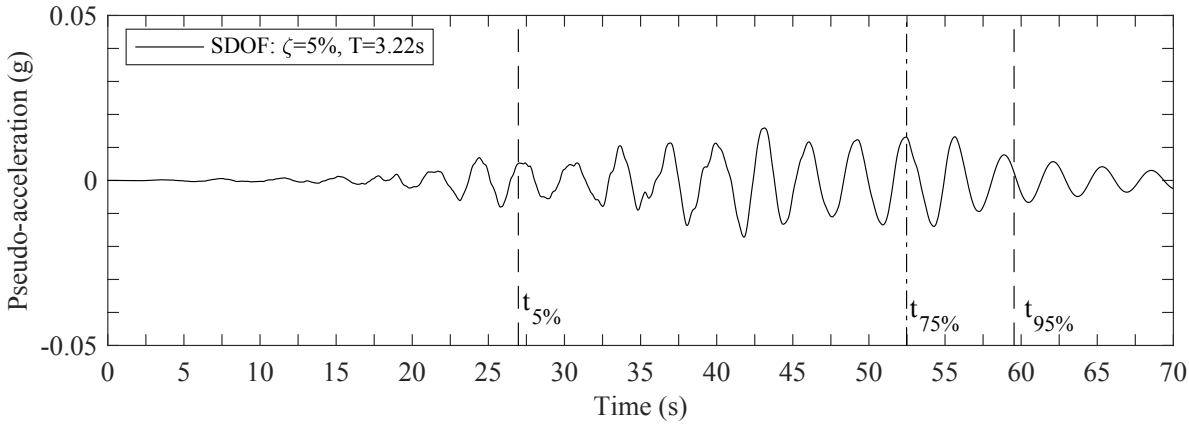

Figure 6: The pseudo-acceleration response of the elastic SDOF oscillator with $T = T_1 = 3.22$ s and $\zeta = 5\%$ subjected to the CHICHI_CHK-E GM and the time instances, when 5%, 75% and 95% of the integral $I(t_{max})$ is reached

To further demonstrate the sensitivity to the appropriate selection of t_{max}, Fig. 7 shows the pseudo-acceleration response of the SDOF oscillator with $T = 2T_1 = 6.44$ s and damping $\zeta = 5\%$. The pseudo-acceleration SDOF response depicted in black is computed up to $t_{max} = 60$ s (i.e., the end of the GM excitation) and yields the time bounds of t_5 and t_{95} plotted in dashed-black vertical lines. The pseudo-acceleration SDOF response shown in red is computed up to $t_{max} = \infty$ (i.e., a very large value) with the corresponding time bounds of t_5 and t_{95} plotted in

dashed-red vertical lines. The sensitivity of t_{95} to the selection of t_{max} is larger compared to t_5 and the resulting significant duration is $SDs_{5-95}^{T=2T_1} \approx 43$ s. It should also be noted that the time instance of collapse at 57 s is included both in $SDs_{5-95}^{T=2T_1}$ and $SDs_{5-75}^{T=2T_1}$ (which is not shown in Fig. 7 but is visible later in Fig. 9).

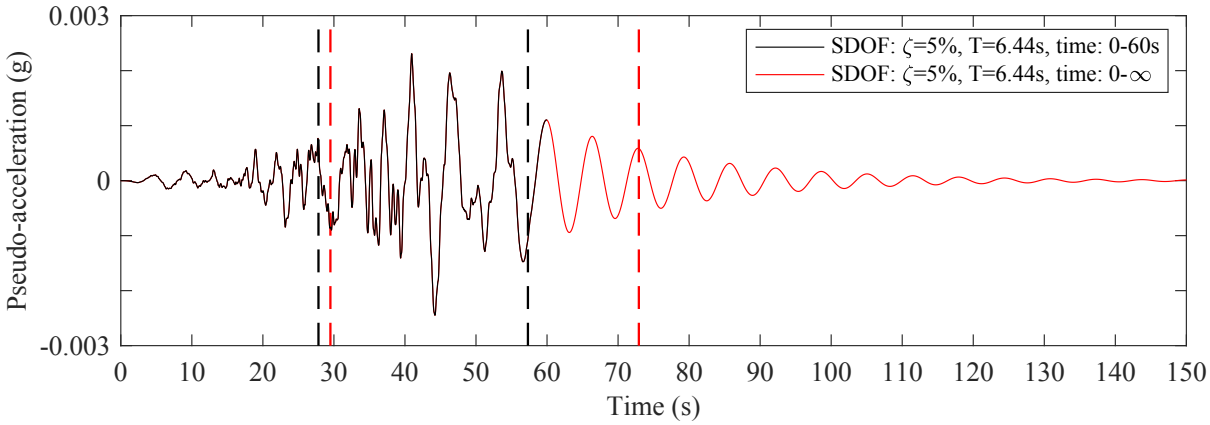

Figure 7: The pseudo-acceleration response of the elastic SDOF oscillator with $T = 2T_1 = 6.44s$ and $\zeta = 5\%$ subjected to the CHICHI_CHK-E GM and the time instances when 5% and 95% of the integral $I(t_{max})$ is reached for: 1) a selection of $t_{max} = 60s$ (in black) and 2) a selection of $t_{max} = \infty$ (in red)

From the above it is clear that the spectral significant durations can be substantially different depending on the oscillation period, and also deviate substantially compared to the more commonly employed significant durations (either considering an upper limit of 75%, 95% or any other bound). For instance, it was shown for CHICHI_CHK-E that $Ds_{5-95} \approx 27$ s, $SDs_{5-95}^{T=T_1} \approx 33$ s and $SDs_{5-95}^{T=2T_1} \approx 43$ s. Before approaching the question of which of the two metrics is more appropriate, the duration spectra for the CHICHI_CHK-E GM are shown in Fig. 8 for SDs_{5-95} (in black) and SDs_{5-75} (in red) to observe these differences over a wide range of periods. Note that for $T = 0$ s the significant duration spectra correspond to the significant duration of the GM (i.e., $SDs_{i-j}^{T=0} = Ds_{i-j}$), since the pseudo-acceleration of the corresponding SDOF oscillator is identical to the GM excitation.

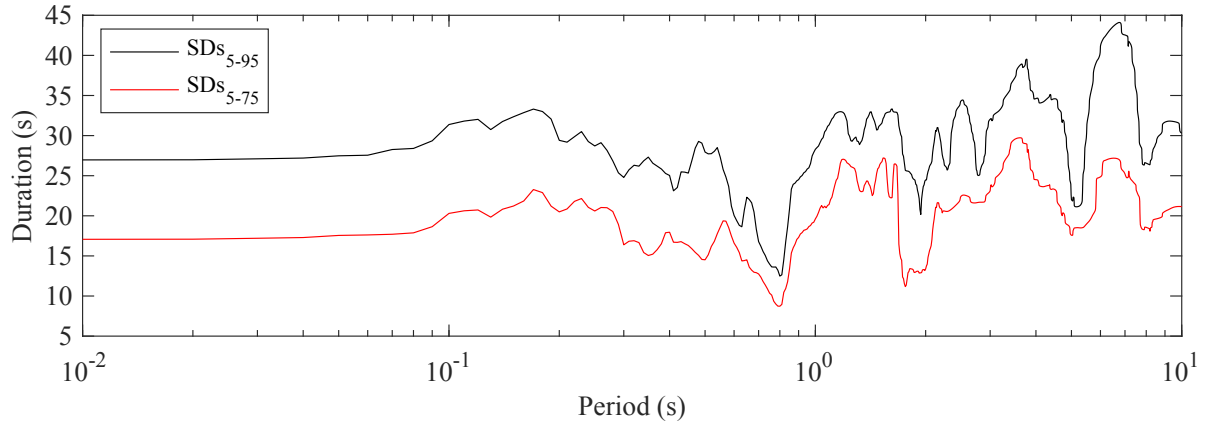

Figure 8: The significant duration spectra SDs_{5-95} (in black) and SDs_{5-75} (in red) for the CHICHI_CHK-E GM

When considering the significant duration spectra in a manner similar to the typical response spectra, various metrics in addition to those shown above (i.e., $SDs_{i-j}^{T=T_1}$ and $SDs_{i-j}^{T=2T_1}$) are of particular interest and are assessed here. Specifically, the spectral significant durations are examined at periods T_3, T_2, T_1, $2T_1$ and $2.5T_1$, as well as the average of significant durations $SDs_{i-j}^{T_{low}-T_{high}}$ for values of $\{T_{low}, T_{high}\} = \{0.2T_1, 2.5T_1\}$ and $\{T_{low}, T_{high}\} = \{T_1, 2.5T_1\}$. The motivation for this choice is primarily the already known significance of the corresponding values of the typical (amplitude) spectra of GM such as the average of spectral accelerations in the range $[0.2T_1\ 2.5T_1]$. Additionally, a period range of $[T_1\ 2.5T_1]$ was chosen because the effect of duration is expected to affect the nonlinear responses of the oscillation, where the effect of period elongation is prominent. As for the percentage bounds i and j, in addition to the typical choices of 5%, 75% and 95%, an additional lower bound of 1% and an additional upper bound of 99% are also investigated to assess their effect. Fig. 9 therefore shows the time instances of t_1, t_5, t_{75}, t_{95} and t_{99} calculated from the signal of the CHICHI_CHK-E GM, as well as the SDOF oscillators with periods T_3, T_2, T_1, $2T_1$ and $2.5T_1$ and the average values of the SDOF oscillators in the period ranges $[0.2T_1\ 2.5T_1]$ and $[T_1\ 2.5T_1]$ when excited with the CHICHI_CHK-E GM. When considering the time instance at which collapse occurs at 57 s, it can be seen from Fig. 9 that it is reflected in various percentages for the eight different metrics examined here (e.g., with a value of $t_{99.9}$ corresponding to the GM and a value of t_{75} corresponding to the SDOF oscillator with $T = 2T_1$).

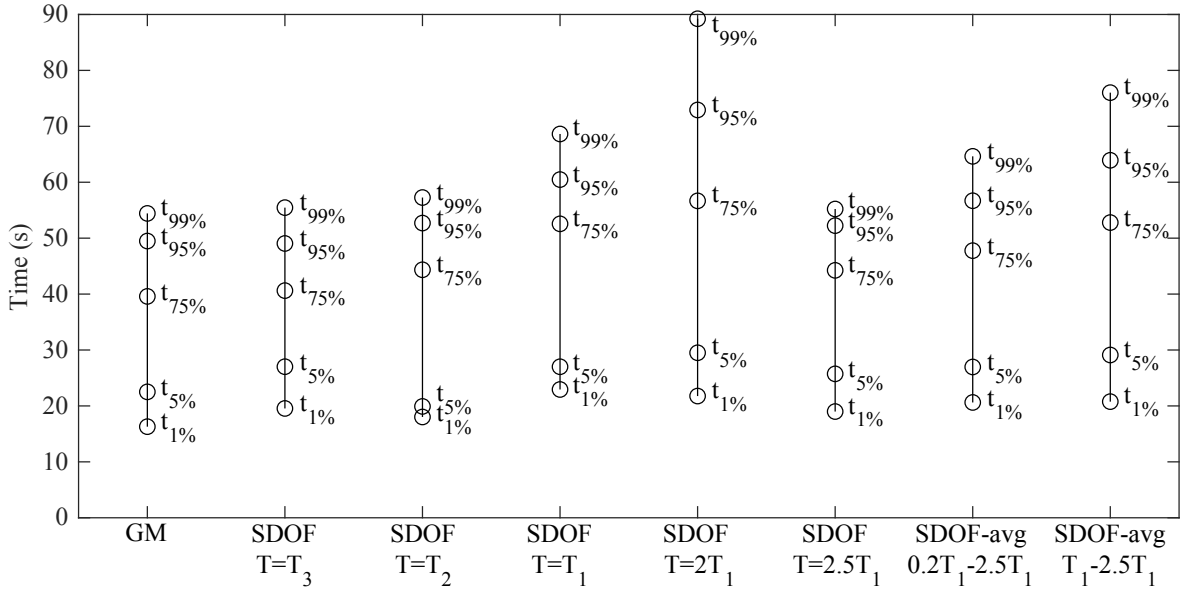

Figure 9: The time instances of t_1, t_5, t_{75}, t_{95} and t_{99} calculated from the signal of CHICHI_CHK-E GM GM, the SDOF oscillators with periods T_3, T_2, T_1, $2T_1$ and $2.5T_1$ and the average values of the SDOF oscillators in the period ranges $[0.2T_1\ 2.5T_1]$ and $[T_1\ 2.5T_1]$ subjected to the CHICHI_CHK-E GM

To explore possible trends related to these metrics, the above process is performed for a total of 730 GM (including CHICHI_CHK-E) obtained from an earthquake scenario with magnitude $M \in [6.9\ 8]$, Joyner-Boore distance $R_{JB} \leq 100$ km, and the average seismic shear-wave velocity from the surface to a depth of 30 m is $V_{s,30} \in [180\ 360]$ m/s, which is related to stiff soils and is classified as category D according to the National Earthquake Hazards Reduction Program (NEHRP). For each of the 730 GM, the time instances at which collapse occurred

($t_{collapse}$) are determined, and for comparison among the different GM, it is converted to the corresponding percentage for each of the eight metrics examined (i.e., $\frac{I(t_{collapse})}{I(t_{max})}$). The eight boxplots in Fig. 10 show the median values with a red line in each box and 25th / 75th percentiles as the box edges, as well as the minimum and maximum values with a black horizontal marker and the outliers with red cross markers. For the examined earthquake scenario yielding 730 GM, the $t_{collapse}$ values in terms of the GM signal are towards the end of it with a median value of t_{95}, while a more balanced distribution is obtained through the SDOF oscillators and the average values of SDOF oscillators. These results indicate that the significant duration metrics obtained from these SDOF oscillators may be particularly informative and superior to those obtained from the GM record.

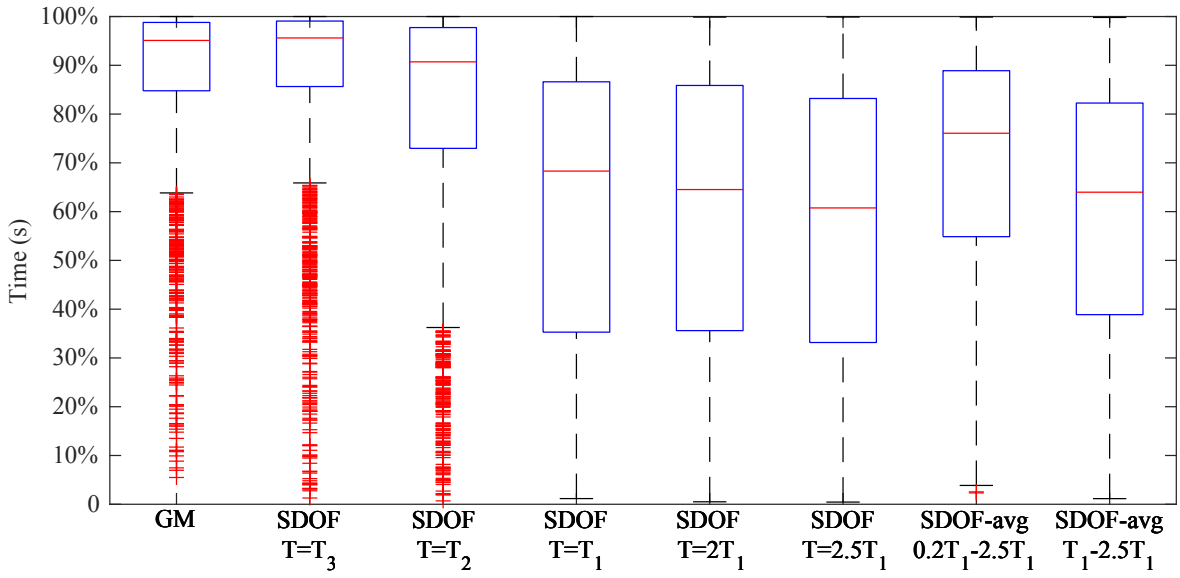

Figure 10: Summary statistics of the collapse time instances $t_{collapse}$ for 730 GM shown through box-plots and expressed as the percentage $\frac{I(t_{collapse})}{I(t_{max})}$

4 ARTIFICIAL NEURAL NETWORKS FOR THE PREDICTION OF SEISMIC COLLAPSE INTENSITY OF THE CASE-STUDY BUILDING

The duration metrics previously discussed are examined here in their usefulness as input features to NN that predict the seismic collapse intensity of buildings. More specifically, multiple NN are trained and tested for the case-study building using the developed database of structural responses. The objective is to assess the benefit of including each of the significant duration metrics as input features in addition to the spectral shape input features when the NN predict the GM intensity at which seismic collapse occurs. In the following, first a brief overview is provided on the scope of such NN as prediction models to better understand the motivation of their use. Then, the chosen formulation and architecture is discussed, which was chosen to be the same for all NN studied, with a total of 26 NN employed in this work to assess the influence of the significant duration metrics. Finally, the predictive ability of the NN is evaluated in order to draw meaningful conclusions.

4.1 Scope and input features of artificial neural networks

Multiple studies have been conducted that use NN or machine learning models to bypass the tedious NLRHA and thus perform seismic assessments rapidly (e.g., [19, 23, 24, 25, 5]). Of particular research interest among these studies is collapse estimation, since the collapse limit state (or collapse prevention) is of paramount importance in earthquake design and assessment. The prediction of structural collapse of a building subjected seismic excitation can be seen in different ways in terms of the scope and approach of interest. For example, classification schemes involving logistic regression and support vector machine binary classifiers have been developed to predict whether or not collapse occurs of the case-study building under investigation subjected to a given seismic excitation, see [5] . A different formulation of the same problem using regression involves predicting the intensity (i.e., the output is a real number in terms of an IM) at which collapse occurs given a particular seismic excitation, which was conducted in [30].

In this work, NN that predict the collapse intensity of the case-study building in terms of Sa_{avg} given a GM record (i.e., predict the intensity at which the given GM causes the case-study building to collapse) are of interest. In this way, by employing each of the prediction models developed in this work for several GM, the collapse intensities $Sa_{avg}@collapse$ can be determined in negligible time and thus the corresponding collapse fragilities. The output of the prediction models is thus a real-value output in terms of Sa_{avg}, while the input features (or predictors) on which the predictions are based are informative IM obtained from the GM (discussed in more detail below).

As mentioned earlier, a total of 26 NN are employed in this work to examine the influence of the significant duration metrics as input features. To ensure that the comparisons are accurate, each of the 26 NN has the same list of non-duration-based features described here. The non-duration-based features are: $\frac{Sa(T_3)}{Sa(T_1)}$, $\frac{Sa(T_2)}{Sa(T_1)}$, $\frac{Sa(2T_1)}{Sa(T_1)}$, $\frac{Sa(2.5T_1)}{Sa(T_1)}$ and $\frac{Sa_{avg}}{Sa(T_1)}$. It should be noted that each of these five features is informative about the spectral shape of the GM and not about its intensity (i.e., they remain the same for any amplitude scaling of the GM), similarly to the vector IM developed in [32]. The motivation behind normalizing the spectral features with $Sa(T_1)$ is to isolate the spectral shape information from the intensity (which is to be predicted from the NN for the collapse limit state). In addition to the five spectral-shape features, the MIDR is the final common input feature that takes values from 0 to 0.1 (i.e., until the collapse limit state). If the duration-based features of the k^{th} NN are $\{SDs\}_k$, then the entire list of input features is:

$$x_k = \left[\frac{Sa(T_3)}{Sa(T_1)},\ \frac{Sa(T_2)}{Sa(T_1)},\ \frac{Sa(2T_1)}{Sa(T_1)},\ \frac{Sa(2.5T_1)}{Sa(T_1)},\ \frac{Sa_{avg}}{Sa(T_1)},\ MIDR,\ \{SDs\}_k\right] \tag{2}$$

The next step is to identify the $\{SDs\}_k$ features for each NN. Following the investigations in Section 3, the significant durations Ds_{i-j}, the spectral significant durations SDs_{i-j}^{T} and average of significant durations $SDs_{i-j}^{T_{low}-T_{high}}$ are examined for four percentage bounds: $\{i,j\} = \{5,75\},\ \{5,95\},\ \{5,99\},\ \{1,99\}$. Moreover, the periods of vibration T_1, $2T_1$ and $2.5T_1$ and the ranges of periods $[0.2T_1\ 2.5T_1]$ and $[T_1\ 2.5T_1]$ (for the average of significant durations) are investigated. In this way, a total of 24 significant-duration-based predictors are developed, listed in Table 2. To examine their influence, for each one of the 24 significant-duration-based predictors, one NN is trained (i.e., having as $\{SDs\}_k$ one predictor from Table 2), while another NN is trained with all predictors (i.e., having as $\{SDs\}_k$ a vector with all 24 predictors from Table 2) and another is trained without a single predictor (i.e., having $\{SDs\}_k$ as empty). In this way, 24 NN have seven predictors, one NN has 30 predictors and one NN has six predictors. The reason for this choice is that there is one NN that predicts

$Sa_{avg}@collapse$ utilizing only the spectral shape information and no significant-duration-based input feature, one NN that uses all of predictors and one NN for each of the predictors. To refer to each of the developed NN with seven predictors, the notation to be used is the corresponding significant-duration-based predictor from Table 2, while the NN with 30 predictors is referred as "All" and the one with six predictors is referred as "None".

Table 2: List of the 24 significant-duration-based predictors

$\{i,j\}$	Ds_{i-j}	$SDs_{i-j}^{T_1}$	$SDs_{i-j}^{2T_1}$	$SDs_{i-j}^{2.5T_1}$	$SDs_{i-j}^{0.2T_1-2.5T_1}$	$SDs_{i-j}^{T_1-2.5T_1}$
$\{5,75\}$	Ds_{5-75}	$SDs_{5-75}^{T_1}$	$SDs_{5-75}^{2T_1}$	$SDs_{5-75}^{2.5T_1}$	$SDs_{5-75}^{0.2T_1-2.5T_1}$	$SDs_{5-75}^{T_1-2.5T_1}$
$\{5,95\}$	Ds_{5-95}	$SDs_{5-95}^{T_1}$	$SDs_{5-95}^{2T_1}$	$SDs_{5-95}^{2.5T_1}$	$SDs_{5-95}^{0.2T_1-2.5T_1}$	$SDs_{5-95}^{T_1-2.5T_1}$
$\{5,99\}$	Ds_{5-99}	$SDs_{5-99}^{T_1}$	$SDs_{5-99}^{2T_1}$	$SDs_{5-99}^{2.5T_1}$	$SDs_{5-99}^{0.2T_1-2.5T_1}$	$SDs_{5-99}^{T_1-2.5T_1}$
$\{1,99\}$	Ds_{1-99}	$SDs_{1-99}^{T_1}$	$SDs_{1-99}^{2T_1}$	$SDs_{1-99}^{2.5T_1}$	$SDs_{1-99}^{0.2T_1-2.5T_1}$	$SDs_{1-99}^{T_1-2.5T_1}$

4.2 Formulation of artificial neural network architecture

The most typical choice for such applications is the multilayer feedforward backpropagation NN, which is frequently employed in other studies (e.g., [23, 24, 25, 29]). In such an NN, a number of layers are sequentially arranged, each containing a number of artificial neurons. Starting from the input layer, which contains the input features (shown in Eq. 2), and forwardly performing mathematical operations in each of the following layers, called hidden layers, the prediction is obtained as the outcome of the final layer, called output layer. Therefore, the number of artificial neurons in the input layer is equal to the number of input features for each NN, and the number of artificial neurons in the output layer is one corresponding to the prediction $Sa_{avg}@collapse$.

The NN architecture refers to the selection of parameters (or hyperparameters) in terms of the number of layers and neurons in each layer, as well as other choices that determine the operation of the NN, which are briefly explained below. The mathematical operations to be performed in each artificial neuron are shown in Fig. 11 and expressed by the following equation.

$$y_n = \phi\left(\sum_{j=1}^{m} w_{nj}x_j + b_n\right) \tag{3}$$

where m is the total number of input signals (or scalars) x_j to the neuron, multiplied with a weight factor w_{nj} and their sum is added to the bias term b_n. The sum is the input to a function ϕ, commonly referred to as activation function (which is also a hyperparameter of the NN architecture), resulting in the output signal y_n. The choice of ϕ in this work is the hyperbolic tangent function, which is a common choice of a nonlinear activation function [24, 25].

As for the number of hidden layers and the number of neurons in each of them, as well as the connections between the layers, Fig. 12 shows graphically the choices in this work. The hidden layers are five in total (denoted HL in Fig. 12), each of them contains 35 neurons and the connections between them are conducted from each layer to the next. The input layer (denoted IL in Fig. 12) is connected to all subsequent layers including the output layer (denoted OL in Fig. 12). In this way, each neuron in the NN after the first hidden layer has as input signals the

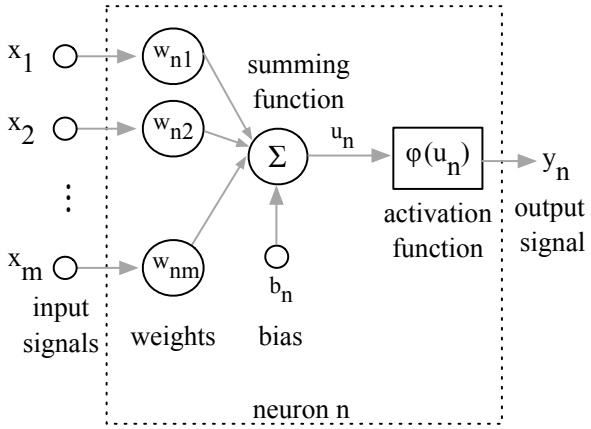

Figure 11: Schematic representation of the mathematical operations in an artificial neuron

35 output signals from the previous hidden layer and, in addition, the input features from the input layer. To illustrate the connections in Fig. 12, they are shown in detail with respect to the input layer and the first hidden layer, as well as the fifth hidden layer and the output layer, while a summary sketch is shown below to illustrate that the input layer is connected to all subsequent layers. It should be noted that the NN architecture shown in Fig. 12 is the same for all 26 NN developed in this work, with the only difference being the number of input features.

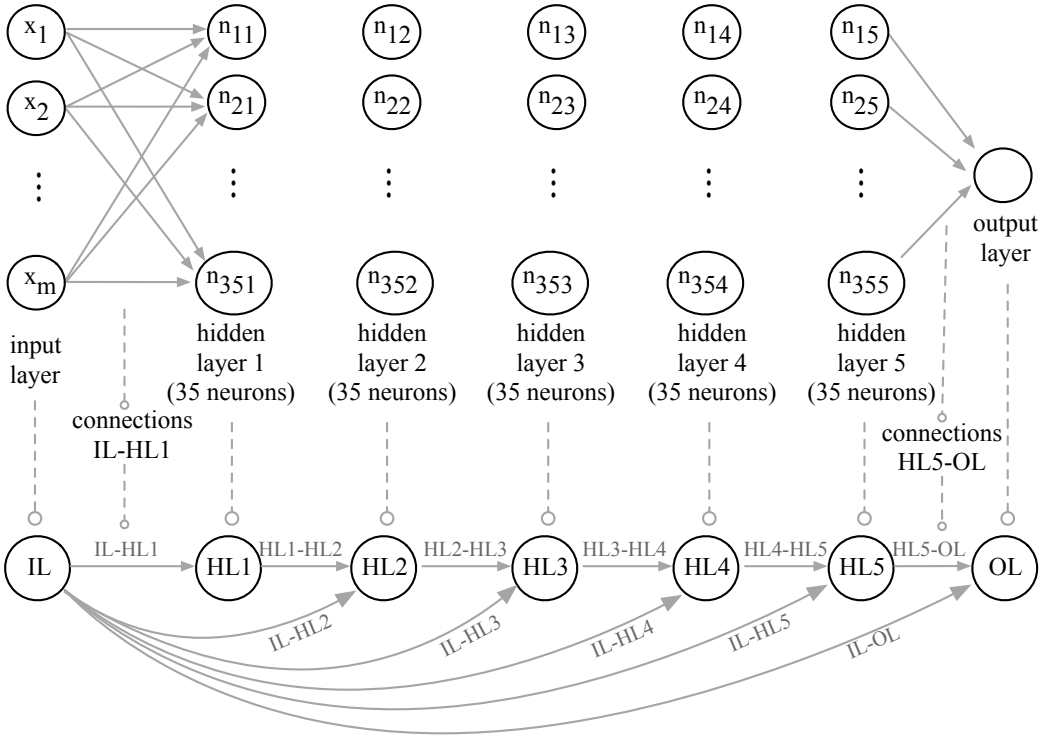

Figure 12: The architecture of each of the 26 developed NN

4.3 Evaluating the performance of artificial neural networks

The 26 NN are tested here in predicting seismic collapse, with the intention to investigate whether the inclusion of the SDs and Ds is beneficial. Seismic collapse is investigated in terms of collapse fragilities, and the comparisons are conducted between collapse fragilities obtained from the NLRHA and the NN referred as 'actual' and 'predicted' respectively. To ensure that the comparisons are robust, the database utilized to develop the NN is separated in training, validation and test GM sets following the 75%, 15% and 15% respectively, which is a common choice for such applications [24, 25]. More specifically, each GM record from the database including its amplitude scaled versions is included either in the training set, validation set or test set. The training and validation sets are used for the development of each of the NN [24], while the test set is only used to demonstrate their predictive ability. In this way, the predictions are reliable since they are demonstrated in respect to "unseen" data.

In order to identify the effect of the duration-based features, it is considered necessary to investigate GM sets with distinctive duration characteristics. For this reason, two subsets from the test set are obtained, with the one containing approximately the 10% longest GM and the other containing the 10% shortest GM (in terms of their duration), subsequently referred as 'long' and 'short' GM sets. It is mentioned that the 'long' GM set includes 278 GM with $Ds_{5-95} > 56$ s while the 'short' GM set includes 234 GM with $Ds_{5-95} < 9$ s. In this way, the effect of duration is amplified in the two sets.

The collapse fragilities of the 'long' and 'short' GM sets are shown in Fig. 13, both as the empirical cumulative distribution of the Sa_{avg} and the fitted lognormal distribution, for the actual (in red) and predicted (in black) using the "All" NN (i.e., containing all the features). It is noted that the collapse fragilities between the 'long' and 'short' GM sets differ substantially with a median of approximately 0.5 g and 0.8 g respectively, while the predictions fit the results of the NLRHA perfectly (also discussed in the following).

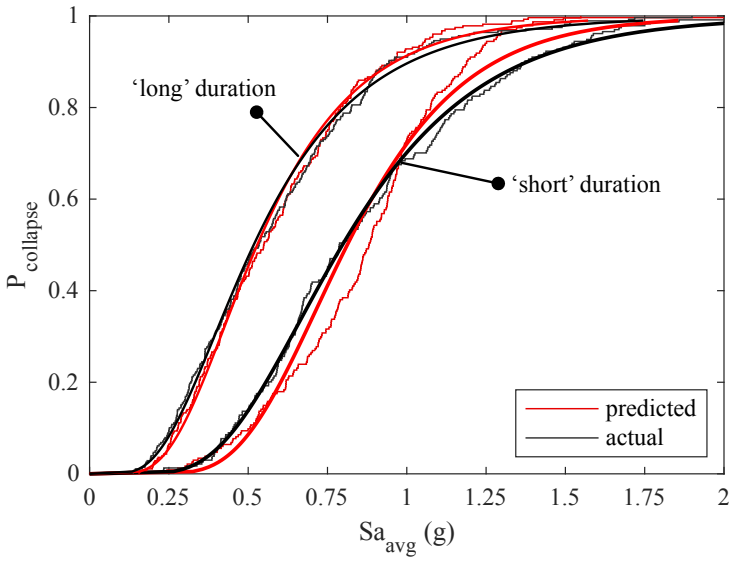

Figure 13: The collapse fragilities of the 'long' and 'short' GM sets, shown through their empirical and fitted lognormal distributions for the actual (black) and the predicted (red) using the "All" NN (i.e., containing all the features).

In order to compare efficiently the results of all NN with respect to the NLRHA, Fig. 14 and Fig. 15 summarize the 27 collapse fragilities (i.e., 26 from the NN and 1 from the NLRHA). The bar graphs show the median and standard deviations as computed from the fitted lognormal distribution for the 'short' and 'long' GM set respectively. The value on top of each bar corresponds to the ratio of the median or standard deviation of the bar over the corresponding value from the NLRHA. In both figures, the "All" NN appears to perform very well, arguably better (or equally well) than the others (in terms of both median and standard deviation predictions for the 'short' and 'long' GM sets). The "None" NN in Fig. 14 is under-predicting both the median and standard deviation (with 92% and 54% respectively) and the Ds collectively are performing slightly worse compared to the SDs. Even though the performance between the different NN is not dramatically different, it is argued from both Fig. 14 and Fig. 15, that there is a clear benefit from utilizing the SDs features.

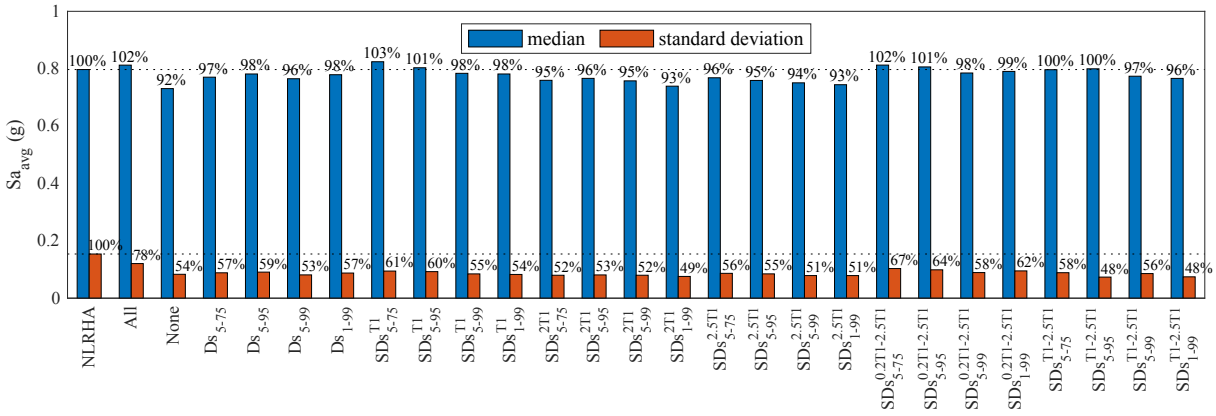

Figure 14: The performance of all 26 NN in terms of median and standard deviation, compared to the results from NLRHA for the 'short' GM set.

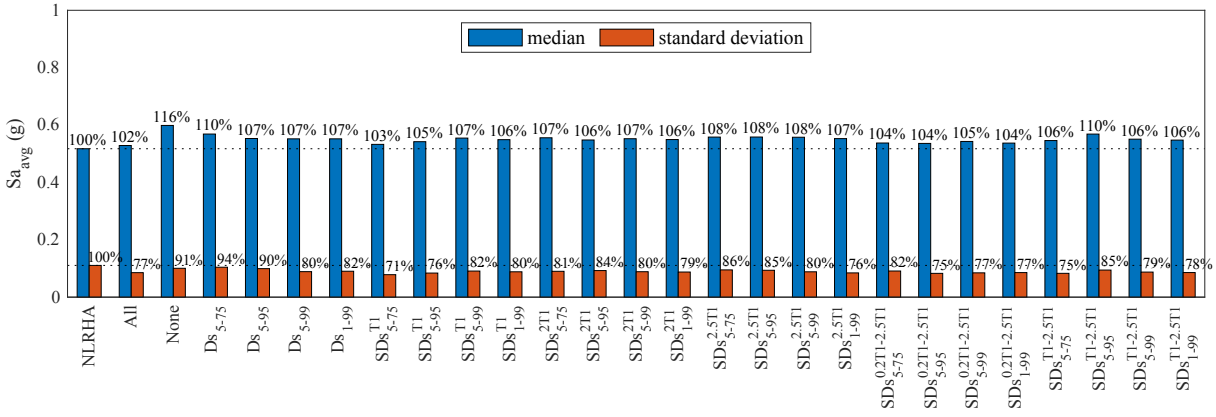

Figure 15: The performance of all 26 NN in terms of median and standard deviation, compared to the results from NLRHA for the 'long' GM set.

5 CONCLUSIONS

In the present work, the importance of various significant duration (Ds) metrics as predictors of structure collapse was investigated. First, the commonly employed Ds intensity measures (IM) were extended by using single-degree-of-freedom (SDOF) oscillators to obtain frequency-dependent or spectral-based significant durations (SDs). The SDs IM that were examined utilized the pseudo-acceleration response of elastic SDOF oscillators at specific periods of vibration related to the periods of vibration of the case-study building. Moreover, the average of SDs in period ranges around the fundamental period of vibration were examined, following the well-studied average of spectral accelerations around the first fundamental period of vibration. A total of 24 duration-based (SDs and Ds) IM were taken into account.

Initially, the duration-based IM were examined for their ability to capture the time-instances of collapse as revealed by extensive nonlinear response history analysis (NLRHA) of the case study building. The results from these investigations support the intuition that SDs IM can (in certain cases) explain structural responses (and the time-instance of collapse in particular) better than Ds.

Subsequently, the 24 SDs and Ds IM were used as input features to artificial neural networks (NN) in addition to spectral shape IM, with the goal of predicting seismic collapse fragilities of the case-study building as obtained by NLRHA. A total of 26 NN were developed to investigate the influence of the 24 IM. The results indicate that taking into account multiple SDs IM is beneficial in collapse fragility prediction, as opposed to considering only Ds IM or no duration-based IM.

References

[1] T. D. Ancheta, 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, et al. NGA-West2 database. *Earthquake Spectra*, 30(3):989–1005, 2014.

[2] A. Arias. A measure of earthquake intensity. *Seismic design for nuclear plants*, pages 438–483, 1970.

[3] M. Bahrampouri, A. Rodriguez-Marek, and R. A. Green. Ground motion prediction equations for significant duration using the kik-net database. *Earthquake Spectra*, 37(2):903–920, 2021.

[4] A. R. Barbosa, F. L. A. Ribeiro, and L. A. C. Neves. Influence of earthquake ground-motion duration on damage estimation: application to steel moment resisting frames. *Earthquake Engineering & Structural Dynamics*, 46(1):27–49, 2017.

[5] N. Bijelic, T. Lin, and G. Deierlein. Classification algorithms for collapse prediction of tall buildings and regional risk estimation utilizing SCEC CyberShake simulations. page 8.

[6] N. Bijelić, T. Lin, and G. G. Deierlein. Quantification of the influence of deep basin effects on structural collapse using SCEC cybershake earthquake ground motion simulations. *Earthquake Spectra*, 35:1845–1864, 2019.

[7] N. Bijelić, T. Lin, and G. G. Deierlein. Efficient intensity measures and machine learning algorithms for collapse prediction of tall buildings informed by scec cybershake ground motion simulations. *Earthquake Spectra*, 36(3):1188–1207, 2020.

[8] J. J. Bommer, G. Magenes, J. Hancock, and P. Penazzo. The influence of strong-motion duration on the seismic response of masonry structures. *Bulletin of Earthquake Engineering*, 2:1–26, 2004.

[9] J. J. Bommer and A. Martínez-Pereira. The effective duration of earthquake strong motion. *Journal of earthquake engineering*, 3(02):127–172, 1999.

[10] R. Chandramohan. *Duration of earthquake ground motion: Influence on structural collapse risk and integration in design and assessment practice.* PhD thesis, Stanford University, Stanford, CA, USA, 2016.

[11] R. Chandramohan, J. W. Baker, and G. G. Deierlein. Impact of hazard-consistent ground motion duration in structural collapse risk assessment. *Earthquake Engineering & Structural Dynamics*, 45(8):1357–1379, 2016.

[12] R. Chandramohan, J. W. Baker, and G. G. Deierlein. Quantifying the influence of ground motion duration on structural collapse capacity using spectrally equivalent records. *Earthquake Spectra*, 32(2):927–950, 2016.

[13] N. Gremer, C. Adam, R. A. Medina, and L. Moschen. Vertical peak floor accelerations of elastic moment-resisting steel frames. *Bulletin of Earthquake Engineering*, 17(6):3233–3254, 2019.

[14] L. F. Ibarra, R. A. Medina, and H. Krawinkler. Hysteretic models that incorporate strength and stiffness deterioration. *Earthquake Engineering & Structural Dynamics*, 34(12):1489–1511, 2005.

[15] I. Iervolino, G. Manfredi, and E. Cosenza. Ground motion duration effects on nonlinear seismic response. *Earthquake Engineering & Structural Dynamics*, 35(1):21–38, 2006.

[16] T. Ishii. A study on response duration time spectra of earthquake motions in tokyo. In *The 14th World Conference on Earthquake Engineering, Beijing, China, Paper ID*, pages 02–0020, 2008.

[17] C. Kircher, G. Deierlein, J. Hooper, H. Krawinkler, S. Mahin, B. Shing, and J. Wallace. Evaluation of the FEMA P-695 methodology for quantification of building seismic performance factors, 2010.

[18] S. Kitayama and M. C. Constantinou. Implications of strong earthquake ground motion duration on the response and testing of seismic isolation systems. *Earthquake Engineering & Structural Dynamics*, 50(2):290–308, 2021.

[19] N. D. Lagaros and M. Fragiadakis. Fragility assessment of steel frames using neural networks. *Earthquake Spectra*, 23(4):735–752, 2007.

[20] D. G. Lignos and H. Krawinkler. Deterioration modeling of steel components in support of collapse prediction of steel moment frames under earthquake loading. *Journal of Structural Engineering*, 137(11):1291–1302, 2011.

[21] D. G. Lignos and H. Krawinkler. Development and utilization of structural component databases for performance-based earthquake engineering. *Journal of Structural Engineering*, 139(8):1382–1394, 2013.

[22] F. McKenna. Opensees: a framework for earthquake engineering simulation. *Computing in Science & Engineering*, 13(4):58–66, 2011.

[23] C. C. Mitropoulou and M. Papadrakakis. Developing fragility curves based on neural network IDA predictions. *Engineering Structures*, 33(12):3409 – 3421, 2011.

[24] K. Morfidis and K. Kostinakis. Approaches to the rapid seismic damage prediction of r/c buildings using artificial neural networks. *Engineering Structures*, 165:120–141, 2018.

[25] K. Morfidis and K. Kostinakis. Comparative evaluation of mfp and rbf neural networks' ability for instant estimation of r/c buildings' seismic damage level. *Engineering Structures*, 197:109436, 2019.

[26] Y. Shi, J. Li, H. Qin, Z. Zhong, J. Wang, and F. Zhang. Correlation analysis of ground motion duration indexes and nonlinear seismic responses of a long-span continuous rigid-frame bridge with high-rise piers. *Journal of Earthquake Engineering*, 26(13):7011–7031, 2022.

[27] M. D. Trifunac and A. G. Brady. A study on the duration of strong earthquake ground motion. *Bulletin of the Seismological Society of America*, 65(3):581–626, 1975.

[28] K. T. Tsalouchidis and C. Adam. Amplitude scaling of ground motions as a potential source of bias: Large-scale investigations on structural drifts. *Earthquake Engineering & Structural Dynamics*, 51(12):2904–2924, 2022.

[29] K. T. Tsalouchidis and C. Adam. Building-specific artificial neural networks predicting structural drifts. *In: Proc. 3rd European Conference on Earthquake Engineering and Seismology (3ECEES), Bucharest, Romania, September 4-9 2022 (Arion, C., Scupin, A., Țigănesco, A., eds.) (ISBN 978-973-100-533-1), Paper-ID 2414, pp. 2956-2965*, 2022.

[30] K. T. Tsalouchidis, N. Bijelić, C. Adam, and L. Moschen. IDA-based seismic collapse patterns and their predictability by generalized linear models. *In: 17th World Conference on Earthquake Engineering (17WCEE), Sendai, Japan*, 2020.

[31] D. Vamvatsikos and C. A. Cornell. Incremental dynamic analysis. *Earthquake Engineering & Structural Dynamics*, 31(3):491–514, 2002.

[32] D. Vamvatsikos and C. A. Cornell. Developing efficient scalar and vector intensity measures for ida capacity estimation by incorporating elastic spectral shape information. *Earthquake Engineering & Structural Dynamics*, 34(13):1573–1600, 2005.

[33] S. Yaghmaei-Sabegh and S. Makaremi. Development of duration-dependent damage-based inelastic response spectra. *Earthquake Engineering & Structural Dynamics*, 46(5):771–789, 2017.